Is the wealth of the Forbes 400 lists really Pareto distributed?

DOI:
10.1016/j.econlet.2016.12.017

Citation for published version (APA):
Chan, S., Chu, J., & Nadarajah, S. (2017). Is the wealth of the Forbes 400 lists really Pareto distributed? *Economics Letters*, *152*, 9-14. https://doi.org/10.1016/j.econlet.2016.12.017

Published in:
Economics Letters

Is the wealth of the Forbes 400 lists really Pareto distributed?

by

Stephen Chan, Jeffrey Chu and Saralees Nadarajah
School of Mathematics, University of Manchester, Manchester M13 9PL, UK

Abstract: A number of researchers have studied the wealth distribution of the Forbes 400 lists (for example, Klass et al. (2006)). They argue that the wealth is Pareto distributed. We ask the question: does the Pareto distribution really give a statistically adequate fit? We find other distributions giving statistically adequate fits.

Keywords: Anderson Darling test; Beta Pareto distribution; Estimation

1 Introduction

The wealth distribution of the Forbes 400 lists have been investigated by a number of researchers. Using the Forbes 400 lists for 1988-2003, Klass et al. (2006) showed that wealth is distributed according to a Pareto type I distribution with an average exponent of 1.49. A similar conclusion is given in Klass et al. (2007).

The Pareto type I distribution has its probability density function (pdf) and cumulative distribution function (cdf) specified by

$$f(x) = \frac{aK^a}{x^{a+1}} \tag{1}$$

and

$$F(x) = 1 - \frac{K^a}{x^a}, \tag{2}$$

respectively, for $x > K > 0$ and $a > 0$. The parameter a is referred to as the exponent. The parameter K represents the lower end point or the minimum wealth (for the application considered in this note).

The purpose of this note is two folded: i) does the Pareto type I distribution provide a statistically adequate fit to the wealth distribution of the Forbes 400 lists? ii) if not, are there distributions providing statistically adequate fits to the wealth distribution of the Forbes 400 lists? The adequacy is tested using the well known Anderson Darling test for goodness of fit (Anderson and Darling (1952) Anderson and Darling (1954)).

The contents of this note are organized as follows: Section 2 provides a description of the data used; Section 3 answers the two questions; some conclusions are noted in Section 4. All computations for this note were performed using the R software (R Development Core Team (2016)).

2 Data

The data are the wealth in billions of dollars of the 400 richest individuals in the United States for the years from 1988 to 2016. The data were extracted from the website www.forbes.com. The following summary statistics of the data are given in Table 1: minimum (Min), first quartile (Q1), median (Med), mean, third quartile (Q3), maximum (Max), variance (Var), coefficient of variation (CV), skewness (Skew) and kurtosis (Kurt).

Year	Min	Q1	Med	Mean	Q3	Max	Var	CV	Skew	Kurt
2016	1.600	2.400	3.300	5.946	5.000	81.700	82.690	1.529	4.682	29.089
2015	1.700	2.300	3.300	5.844	5.000	76.000	69.227	1.424	4.315	26.019
2014	1.550	2.200	3.000	5.779	4.900	81.000	72.665	1.475	4.472	28.826
2013	1.300	1.900	2.600	4.850	3.400	72.000	55.231	1.532	4.613	30.452
2012	1.050	1.500	2.100	3.813	3.400	59.000	29.427	1.423	4.960	37.471
2011	1.000	1.500	2.000	3.779	3.400	56.000	30.947	1.472	5.227	39.042
2010	1.000	1.400	1.900	3.366	3.225	53.000	23.957	1.454	5.691	47.081
2009	0.950	1.300	1.800	3.172	3.000	50.000	20.268	1.419	5.597	45.913
2008	1.300	1.675	2.300	3.932	3.700	57.000	27.619	1.336	5.476	44.975
2007	1.300	1.600	2.300	3.850	3.500	59.000	27.607	1.365	5.898	51.683
2006	1.000	1.300	1.800	3.136	2.800	53.000	20.747	1.452	6.297	58.398
2005	0.900	1.000	2.000	2.855	3.000	51.000	18.366	1.501	6.275	57.807
2004	0.750	1.000	1.500	2.505	2.200	48.000	17.018	1.647	6.586	60.312
2003	0.600	0.900	1.200	2.389	2.000	46.000	17.367	1.745	5.864	47.516
2002	0.600	0.865	1.200	2.366	2.000	54.000	18.342	1.810	6.886	66.695
2001	0.725	0.980	1.500	2.994	2.600	63.000	31.208	1.866	6.823	62.815
2000	0.700	1.000	1.500	2.808	2.700	85.000	31.008	1.983	9.633	127.275
1999	0.560	0.810	1.100	2.054	1.900	58.400	14.609	1.861	9.704	128.064
1998	0.540	0.724	1.000	1.754	1.800	39.800	7.478	1.559	8.591	104.396
1997	0.520	0.650	0.920	1.369	1.300	18.500	2.537	1.164	5.906	52.018
1996	0.435	0.534	0.750	1.154	1.100	14.800	1.894	1.192	5.197	40.373
1995	0.395	0.500	0.640	1.033	1.000	10.000	1.363	1.130	4.212	25.560
1994	0.360	0.470	0.600	0.959	1.000	9.000	1.086	1.087	4.030	23.118
1993	0.350	0.430	0.575	0.897	0.950	8.600	0.918	1.068	3.972	22.359
1992	0.350	0.430	0.575	0.897	0.950	8.600	0.918	1.068	3.972	22.359
1991	0.340	0.425	0.565	0.828	0.853	8.600	0.587	0.926	4.579	35.860
1990	0.340	0.425	0.565	0.828	0.853	8.600	0.587	0.926	4.579	35.860
1989	0.350	0.429	0.550	0.826	0.887	8.000	0.573	0.916	4.272	30.320
1988	0.290	0.350	0.453	0.686	0.750	6.800	0.456	0.984	5.021	38.686

Table 1: Some summary statistics of the data.

The rich are getting richer. So, it is not surprising that the minimum, first quartile, median, mean, third quartile and the maximum are increasing with the increasing year. The variance is also increasing with the increasing year. The coefficient of variation, skewness and kurtosis initially increase and then decrease. They appear to reach their largest values around the year 2000. The wealth distribution for each year is positively skewed and its peakedness is sharper than that of the normal distribution.

3 Results and discussion

We fitted the Pareto type I distribution to the data in Section 2. The method of maximum likelihood was used. The Akaike information criterion values, and p-values of the Anderson Darling test are given in Table 2. The AIC is due to Akaike (1974). The smaller the value of the AIC the better the fit. For more discussion on the AIC, see Burnham and Anderson (2004) and Fang (2011). All of the p-values for the Pareto type I distribution are so small (zero up to three decimal places), suggesting that its fit is hardly adequate for any of the years.

	Pareto type I distribution		Beta Pareto distribution	
Year	AIC	p-value	AIC	p-value
2016	1803.8	0.000	1768.6	0.605
2015	1745.1	0.000	1738.1	0.392
2014	1754.7	0.000	1749.1	0.163
2013	1599.3	0.000	1592.8	0.054
2012	1463.8	0.000	1444.5	0.351
2011	1482.7	0.000	1453.4	0.296
2010	1317.7	0.000	398.4	0.051
2009	1287.8	0.000	1272.0	0.390
2008	1399.5	0.000	992.9	0.060
2007	1372.1	0.000	798.5	0.058
2006	1223.6	0.000	269.1	0.055
2005	1103.6	0.000	1047.2	0.053
2004	1037.4	0.000	1033.6	0.095
2003	1049.0	0.000	1029.2	0.386
2002	1041.1	0.000	1021.4	0.594
2001	1197.4	0.000	1192.7	0.496
2000	1198.3	0.000	1185.0	0.475
1999	956.8	0.000	930.9	0.813
1998	822.7	0.000	803.6	0.615
1997	554.7	0.000	546.4	0.464
1996	386.7	0.000	382.4	0.557
1995	293.1	0.000	283.4	0.060
1994	281.8	0.000	264.6	0.570
1993	196.6	0.000	123.7	0.061
1992	196.6	0.000	123.7	0.065
1991	157.9	0.000	152.8	0.114
1990	157.9	0.000	152.8	0.114
1989	118.7	0.000	-363.7	0.066
1988	-32.4	0.000	-152.5	0.054

Table 2: Results of the fits of the Pareto type I and beta Pareto distributions.

The Pareto type I distribution has been generalized by many authors. One of the simplest generalizations is the beta Pareto distribution (Akinsete et al. (2008)) with the pdf and cdf specified by

$$f(x) = \frac{aK^{a\beta}}{B(\alpha,\beta)x^{a\beta+1}}\left(1 - \frac{K^a}{x^a}\right)^{\alpha-1} \tag{3}$$

and

$$F(x) = I_{1-\frac{K^a}{x^a}}(\alpha,\beta), \tag{4}$$

respectively, for $x > K > 0$, $a > 0$, $\alpha > 0$ and $\beta > 0$, where $B(a,b)$ and $I_x(a,b)$ are the beta function and the incomplete beta function ratio defined by

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt \tag{5}$$

and

$$I_x(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt, \tag{6}$$

respectively, for $0 < x < 1$, $\alpha > 0$ and $\beta > 0$. The parameters a and K can be referred to as the exponent and minimum wealth, respectively, as for the Pareto type I distribution. α and β are shape parameters. α controls the lower tail behavior. β controls the upper tail behavior.

The beta Pareto distribution can be interpreted as follows: suppose there are $(a+b-1)$ independent Pareto type I random variables representing say the wealth of $(a+b-1)$ individuals; the beta Pareto distribution is the distribution of the ath smallest of these variables; in other words, the beta Pareto distribution is the distribution of the wealth of the ath poorest of the individuals. For example, if $a = 10$ and $b = 1$ then the beta Pareto distribution is the distribution of the wealth of the richest of the 10 individuals. On the other hand, if $a = 1$ and $b = 10$ then the beta Pareto distribution is the distribution of the wealth of the poorest of the 10 individuals.

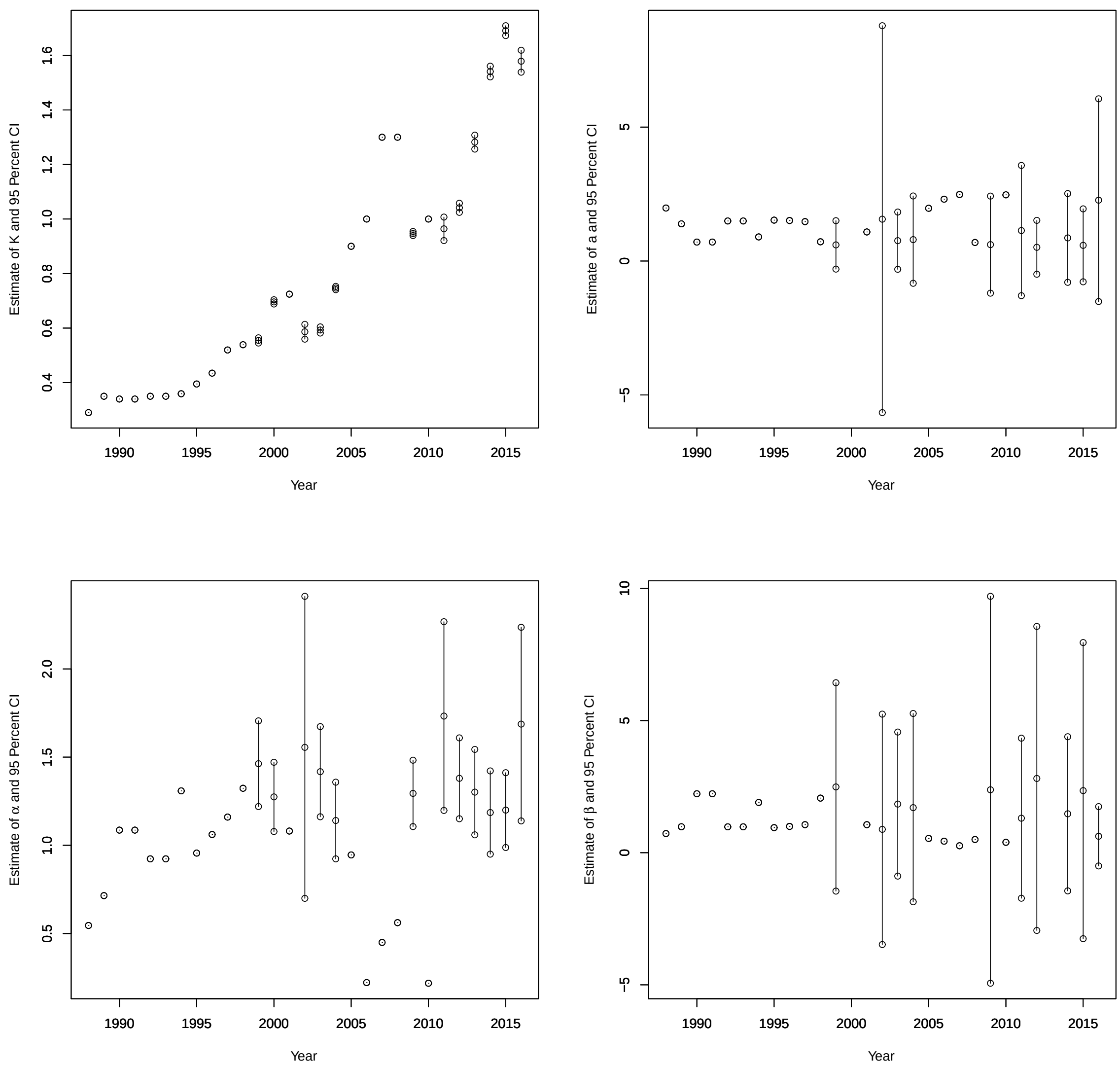

Figure 1: Parameter estimates and their 95 percent confidence intervals for the fitted beta Pareto distribution.

We fitted this distribution to the same data to see if it would give better fits. The AIC values and p-values of the Anderson Darling test are also given in Table 2. We see that the beta Pareto distribution has smaller AIC and much larger p-value than the Pareto type I distribution for each year. The beta Pareto distribution actually gives an adequate fit for each year.

The maximum likelihood estimates of the parameters of the beta Pareto distribution are plotted in Figure 1. Also shown in the figure are 95 percent confidence intervals of the parameter estimates. The estimate of K increases with time since it represents minimum wealth and the rich get richer with time. There is variability in the estimates of a, α and β over the period from 1999 to 2016. However, taking account of the confidence intervals, there is no evidence that these parameter estimates actually vary with time, especially from 1999 to 2016.

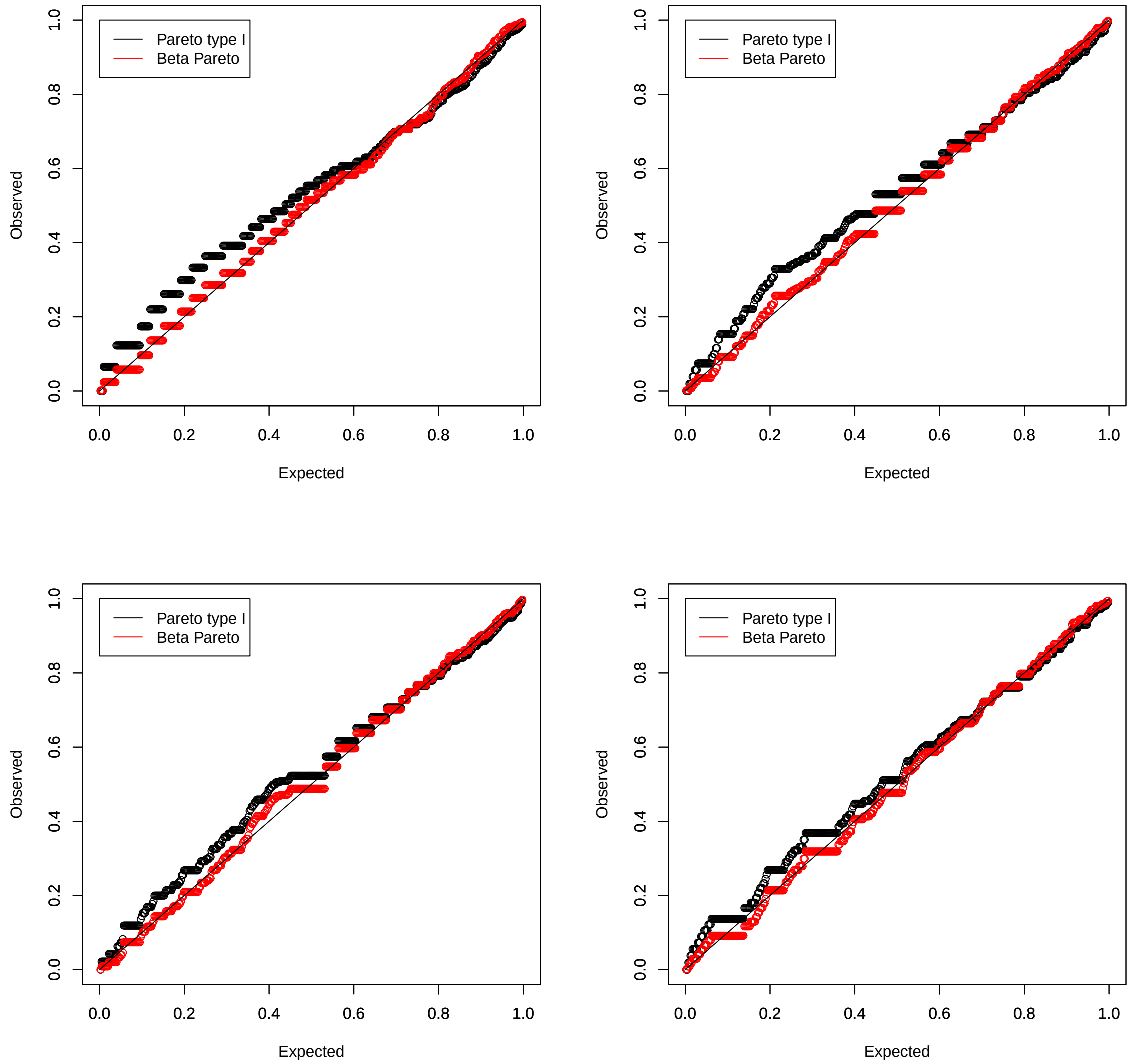

Figure 2: Probability plots for the fits of the Pareto type I and beta Pareto distributions for selected years: 2016 (top left), 1999 (top right), 1998 (bottom left) and 1994 (bottom right).

The adequacy of the beta Pareto distribution is further assessed by Figures 2, 3 and 4. Figure 2 gives the probability plots for selected years. The values plotted on the x axes are the expected probabilities, that is, $i/401$ for $i = 1, 2, \ldots, 400$. The values plotted on the y axes are the observed probabilities under the fitted Pareto type I distribution,

$$1 - \frac{\widehat{K}^{\widehat{a}}}{x_{(i)}^{\widehat{a}}}, \quad i = 1, 2, \ldots, 400,$$

and the observed probabilities under the fitted beta Pareto distribution,

$$I_{1 - \frac{\widehat{K}^{\widehat{a}}}{x_{(i)}^{\widehat{a}}}} \left(\widehat{\alpha}, \widehat{\beta} \right), \quad i = 1, 2, \ldots, 400,$$

where $\widehat{K}, \widehat{a}, \widehat{\alpha}, \widehat{\beta}$ are the maximum likelihood estimates, and $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(400)}$ are the wealth of the 400 billionaires arranged in increasing order.

The probability plots show that the Pareto type I distribution does not give satisfactory fits in both the lower and upper tails of the wealth distribution. The fits appear especially unsatisfactory in the lower tails of the wealth distribution.

The probability plots also show that the beta Pareto distribution does give satisfactory fits in both the lower and upper tails of the wealth distribution. The fits appear especially good in the upper tails of the wealth distribution.

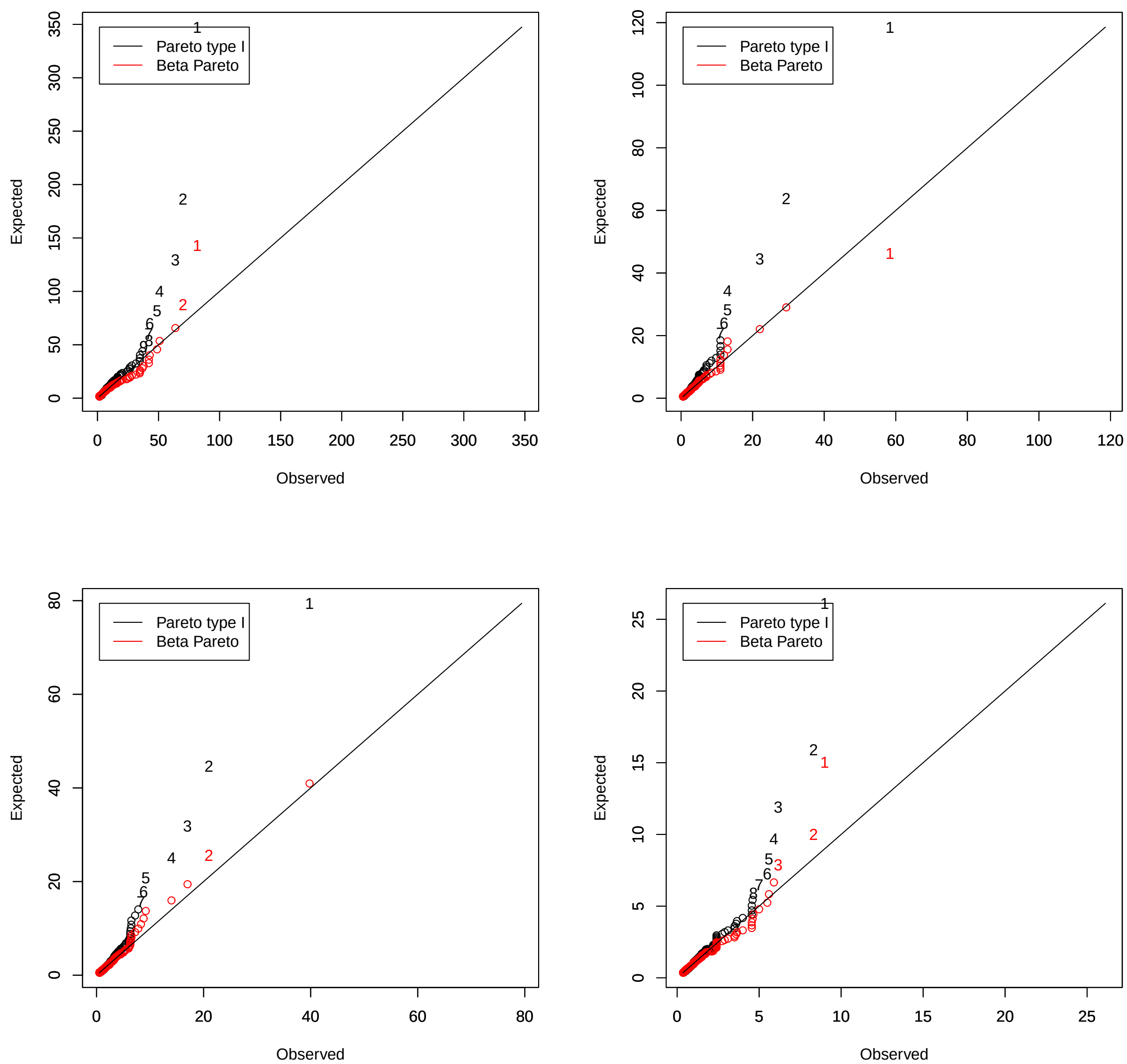

Figure 3: Quantitle plots for the fits of the Pareto type I and beta Pareto distributions for selected years: 2016 (top left), 1999 (top right), 1998 (bottom left) and 1994 (bottom right).

Figure 3 gives the quantile plots for selected years. The values plotted on the x axes are the observed quantiles, that is, $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(400)}$, the wealth of the 400 billionaires arranged in increasing order. The values plotted on the y axes are the expected quantiles under the fitted Pareto type I distribution,

$$\widehat{K}\left(1-\frac{i}{401}\right)^{-1/\widehat{a}}, \quad i=1,2,\ldots,400,$$

and the expected quantiles under the fitted beta Pareto distribution,

$$\widehat{K}\left[1-I^{-1}_{\frac{i}{401}}\left(\widehat{\alpha},\widehat{\beta}\right)\right]^{-1/\widehat{a}}, \quad i=1,2,\ldots,400,$$

where $\widehat{K},\widehat{a},\widehat{\alpha},\widehat{\beta}$ are the maximum likelihood estimates, and $I_p^{-1}(\alpha,\beta)$ denotes the inversion function of the incomplete beta function ratio.

The quantile plots show that the Pareto type I distribution does not give satisfactory fits in the upper tails of the wealth distribution. The top nine richest billionaires (numbered in black from 1 to 9) appear to differ significantly from the 45 degree line for the year 2016. They are Bill Gates, Warren Buffet, Jeff Bezos, Mark Zuckerberg, Larry Ellison, Michael Bloomberg, Charles Koch, David Koch and Larry Page. The top seven richest billionaires (numbered in black from 1 to 7) appear to differ significantly from the 45 degree line for the year 1999. They are Bill Gates, Paul Allen, Warren Buffett, Steven Ballmer, Michael Dell, Alice Walton and Helen Walton. The top seven richest billionaires (numbered in black from 1 to 7) appear to differ significantly from the 45 degree line for the year 1998. They are Bill Gates, Warren Buffett, Paul Allen, Michael Dell, DuPont family, Steven Ballmer and Alice Walton. The top eight richest billionaires (numbered in black from 1 to 8) appear to differ significantly from the 45 degree line for the year 1994. They are DuPont family, Bill Gates, Warren Buffett, Rockefeller family, John Kluge, Edward Johnson, Mellon family and Richard DeVos.

The quantile plots also show that the beta Pareto distribution does give satisfactory fits except for a few data points in the very extreme upper tails. These data points appear to correspond to: the top two billionaires (numbered in red as 1 and 2), Bill Gates and Warren Buffet, for the year 2016; the richest billionaire (numbered in red as 1), Bill Gates, for the year 1999; the second richest billionaire (numbered in red as 2), Paul Allen, for the year 1998; the top three billionaires (numbered in red as 1, 2 and 3), DuPont family, Bill Gates and Warren Buffett, for the year 1994.

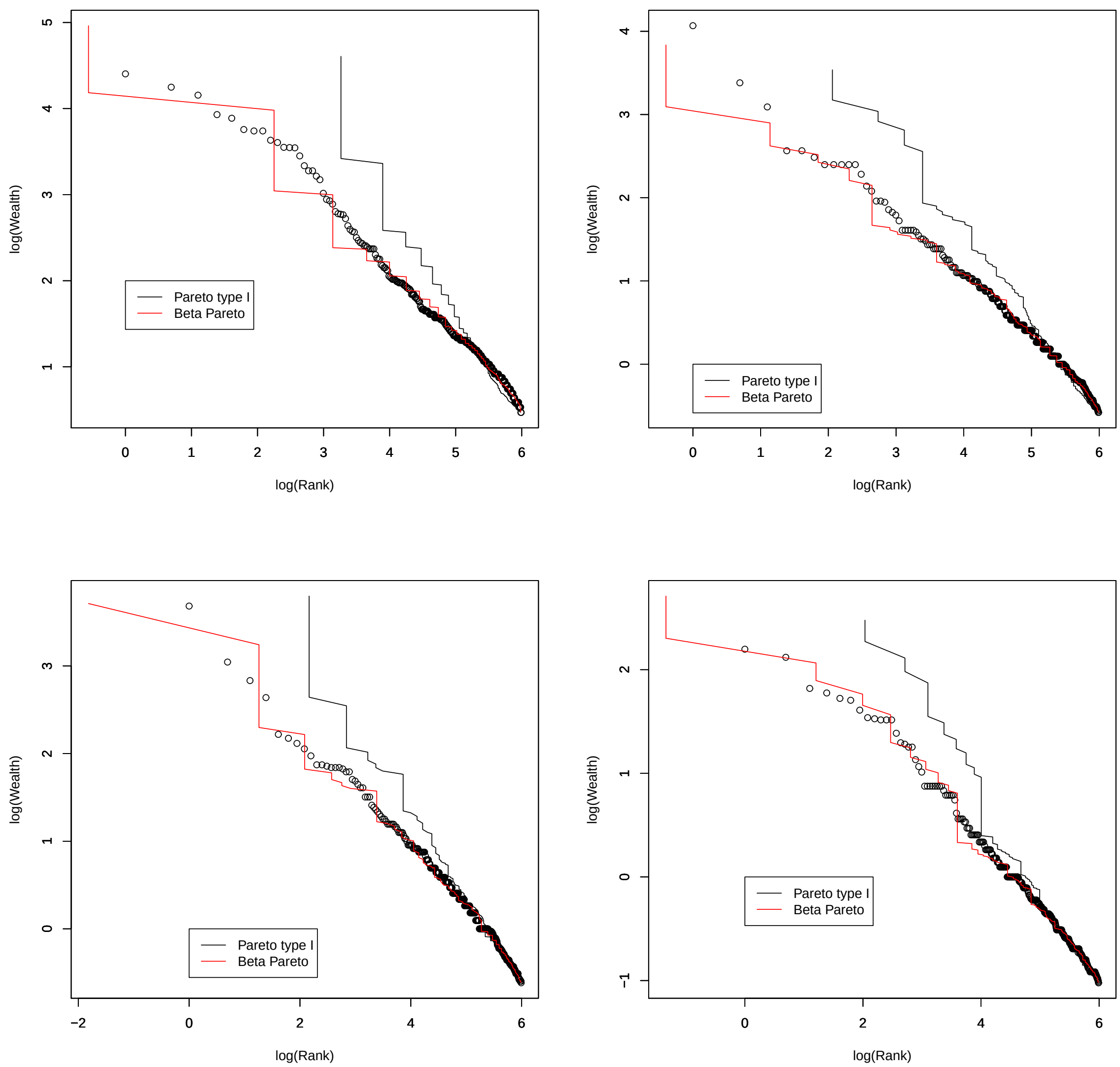

Figure 4: The rank-wealth relationships for the fits of the Pareto type I and beta Pareto distributions for selected years: 2016 (top left), 1999 (top right), 1998 (bottom left) and 1994 (bottom right).

Figure 4 shows the rank-wealth relationship for selected years (Levy (2009)). The x and y coordinates of the plotted points are the logarithms of the rank and the logarithms of the wealth of the 400 billionaires. The rank 1 is used to denote the richest billionaire and the rank 400 is used to denote the poorest billionaire. The x and y coordinates of the curve in black (for the fitted Pareto type I distribution) are

$$\left\{ \log \left[400 \left(1 - \frac{\widehat{K}^{\widehat{a}}}{x_{(i)}^{\widehat{a}}} \right) \right], \log \left[\widehat{K} \left(1 - \frac{401 - i}{401} \right)^{-1/\widehat{a}} \right] \right\}$$

for $i = 1, 2, \ldots, 400$. The x and y coordinates of the curve in red (for the fitted beta Pareto distribution) are

$$\left\{ \log \left[400 I_{1-\frac{\widehat{K}^{\widehat{a}}}{x_{(i)}^{\widehat{a}}}} \left(\widehat{\alpha}, \widehat{\beta} \right) \right], \log \left[\widehat{K} \left[1 - I^{-1}_{\frac{401-i}{401}} \left(\widehat{\alpha}, \widehat{\beta} \right) \right]^{-1/\widehat{a}} \right] \right\}$$

for $i = 1, 2, \ldots, 400$.

The rank-wealth relationships show that the beta Pareto distribution gives much better fits than the Pareto type I distribution for almost all ranks. The Pareto type I distribution performs equally well only when the rank is very close to 400.

The beta Pareto distribution gives much better fits because its two shape parameters add much flexibility. The two parameters allow the tails to behave freely. The averages of the parameter estimates for the beta Pareto distribution are: 1.244 for the exponent parameter; 1.076 for the shape parameter controlling the lower tail; 1.384 for the shape parameter controlling the upper tail.

4 Conclusions

We have studied the wealth distribution of the Forbes 400 lists for the years from 1988 to 2016. We have shown that the Pareto type I distribution does not give a statistically adequate fit to the wealth for any of the years. We fitted a generalization of the Pareto type I distribution referred to as the beta Pareto distribution. It was shown to provide statistically adequate fits to the wealth for all of the years. The adequacy was assessed in terms of Akaike information criterion values, Anderson Darling tests, probability plots, quantile plots and rank-wealth relationships.

The beta Pareto distribution was chosen because it is one of the simplest generalizations of the Pareto type I distribution. A future work is to find other generalizations or other distributions providing statistically adequate fits to the wealth for all of the years.

Acknowledgments

The authors would like to thank the Editor and the referee for careful reading and comments which greatly improved the paper.

References

H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

A. Akinsete, F. Famoye, and C. Lee. The beta pareto distribution. *Statistics*, 42:547–563, 2008.

T. Anderson and D. Darling. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23:193–212, 1952.

T. Anderson and D. Darling. A test of goodness-of-fit. *Journal of the American Statistical Association*, 49:765–769, 1954.

K. Burnham and D. Anderson. Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods and Research*, 33:261–304, 2004.

Y. Fang. Asymptotic equivalence between cross-validations and akaike information criteria in mixed-effects models. *Journal of Data Science*, 9:15–21, 2011.

O. Klass, O. Biham, M. Levy, O. Malcai, and S. Solomon. The forbes 400 and the pareto wealth distribution. *Economics Letters*, 90:290–295, 2006.

O. Klass, O. Biham, M. Levy, O. Malcai, and S. Solomon. The forbes 400 and the pareto wealth distribution. *European Physical Journal B*, 55:143–147, 2007.

M. Levy. Gibrat's law for (all) cities: Comment. *The American Economic Review*, 99:1672–1675, 2009.

R Development Core Team. *R: A Language and Environment for Statistical Computing*, 2016.