

# A review on Quantile functions, Income distributions, and Income inequality measures

Ashlin Varkey<sup>1</sup>, Haritha N Haridas<sup>2</sup>

•

<sup>1,2</sup>Farook College (Autonomous), Department of Statistics, Kozhikode, 673632, Kerala, India.

[ashlinvarkey@gmail.com](mailto:ashlinvarkey@gmail.com) , [haritasuhas@gmail.com](mailto:haritasuhas@gmail.com)

## Abstract

*The modeling of income data has originated a century ago with the work of Vilfredo Pareto. Since then several authors have added immense literature about income distributions and income inequality measures. In the present paper, we have pointed out some recent works in income distributions and income inequality measures. Recently the potential of the quantile function has been discussed in parallel with the distribution function for modeling reliability and income data. So we have also included some quantile functions existing in the literature and their potentiality to model income data in the review. We derived the Lorenz curve, Gini index, Pietra index, Bonferroni index, Bonferroni curve, and Zenga curve for six distributions and checked the model adequacy of three distributions using real income data.*

**Keywords:** Income distribution, Income Inequality measures, Quantile based reliability models, Quantile function, Distribution function.

## 1. Introduction

We can consider the introduction of the Pareto model by Vilfredo Pareto as the beginning of the study of income distributions. A variety of models such as Lognormal, Dagum, Singh Maddala, etc., came into the literature for modeling income data. Several income inequality measures also exist in the literature. Kakwani [1] made a detailed study on income distributions, income inequalities, government policies affecting personal income distributions, and the measurement of poverty. Dagum [2] introduced the economic distance ratio, which is used to determine the degree of income inequality between two populations. For a detailed review of income distributions and income inequality measures since the beginning one can refer to [3]. This review tried to bring all the income distributions and income inequality measures that came in the literature under one umbrella.

Most of the income distributions adopted the distribution function approach. Only a very few have used the quantile function approach to model income data [3, 4]. So in the present review, we have considered quantile functions that can be used as income models.

The study of probability distribution in applied problems can be accomplished in two different ways; one by specifying the distribution function and the other through the quantile function. The quantile function is defined as,

$$Q(u) = F^{-1}(u) = \inf(x | F(x) \geq u), 0 \leq u \leq 1. \quad (1)$$

Since  $F(x) \geq u$  iff  $Q(u) \leq x$ , the knowledge of the form of  $Q(u)$  is equivalent to the knowledge of the functional form of  $F(x)$ . Gilchrist [5] provided a thorough explanation of quantile function and

model-building principles.

After this introduction in Section 2, we have reviewed some income distributions in which the distribution function is in the closed form. In Section 3 we have discussed some quantile functions which can be used as income models. We have reviewed some more quantile functions which have applications in reliability in Section 4. In Section 5 we have discussed income inequality measures existing in the literature and derived inequality measures for six distributions having closed-form quantile functions. We carried out a real data analysis in Section 6. Finally, the conclusions of the study are given in Section 7.

## 2. Distribution function based Income distributions

McDonald et al. [6] found that the Weibull, the Dagum, the Generalized beta of the second kind (GB2) gave the best fit among 5 parameter Generalized beta and its 10 special cases, that were fitted into 82 income data sets. These data sets comprise income data for 23 nations, covering both developed and emerging economies, and were taken from Luxembourg Income Study (LIS) database. He also observed that for almost every country, the Gini index increases monotonically over time.

Exponential Kumaraswamy - Dagum (EKD) distribution for analysis of income and lifetime data was introduced by [7]. The Cumulative Distribution Function (CDF) of EKD distribution is given by,

$$F_{EKD}(x; \alpha, \lambda, \delta, \phi, \theta) = \left\{ 1 - \left[ 1 - \left( 1 + \lambda x^{-\delta} \right)^{-\alpha} \right]^{\phi} \right\}^{\theta} \quad (2)$$

where  $\alpha, \lambda, \delta, \phi, \theta > 0$  and  $x > 0$ . Its basic statistical properties like moments, hazard functions, mean and median deviations, measures of income inequalities like Lorenz and Bonferroni curves, reliability measures,  $k^{\text{th}}$  order statistics, and Renyi entropy were derived. Maximum likelihood estimation was done to estimate the parameters of EKD and its application to real data was also presented.

An application of Gamma distribution to income distributions was explained by [8]. Its parameters were estimated based on the income quantile data, using non-linear optimization and the least square method. The future distribution of income and hunger issues were also discussed in this paper. They used the Gamma distribution to compare the estimates of people living in absolute poverty to the World Bank's report and estimated the number of people who are hungry in each country based on the results of a crop market model.

Fabio Clementi et al. [9] have surveyed k-generalized distribution for modeling income and wealth size distributions. They studied the k-generalized model [10, 11, 12] for the distribution of income, the k-generalized mixture model for the distribution of wealth, and the extended k-generalized distributions of the first and second kinds. They discussed its basic properties, interrelations with other distributions; income inequalities like Gini Index and Lorenz Curve. They concluded that a good fit for the distribution of income and wealth is given by k-generalized models.

A three-parameter Weibull-Pareto (WP) distribution with CDF,

$$F(x; b, \alpha, \theta) = 1 - e^{-\left[ \left( \frac{x}{\theta} \right)^{\alpha} - 1 \right]^b}; x > \theta > 0, \alpha > 0, b > 0 \quad (3)$$

was proposed by [13]. Various properties of WP distribution including moments, incomplete moments, mean deviations, mode, reliability measures, generating functions, quantile functions, Bonferroni and Lorenz curves were derived. They also obtained density of order statistics, Renyi, and q-entropy of the above distribution. The parameters of the above distribution were estimated using the method of maximum likelihood. On two real-life data sets, WP distribution provides a better fit than comparable lifetime models.

Calderin-Ojeda et al. [14] proposed two extensions of the Exponential distribution ie, the Exponential Arc Tan (EAT) model and the composite EAT-Lognormal model. These models can

describe income distributions even with zero income. The adequacy of these models was evaluated using Australian income data for the years 2001-2012 and found that the proposed models provide a better fit for the data than exponential and gamma distribution.

Six mixture distributions based on Weibull ie, Weibull-Paralogistic, Weibull-Weibull, Weibull-Fisk, Weibull-Gamma, Weibull-Burr, and Weibull-Dagum to model income data were proposed by [15]. Here the method of maximum likelihood is used to estimate the parameters of Weibull mixture models and are evaluated with respect to average income per tax unit data for ten countries. Also, the application of income inequality measures like Gini, Bonferroni, and generalized index and mathematical expression of poverty measures like headcount ratio and poverty gap ratio was discussed in this paper.

A comparative study was taken by [16] on the distribution of income of poor households in Malaysia. They compared the models Reverse Pareto [17, 18] shifted reverse exponential, shifted reverse stretched exponential [19], and shifted reverse lognormal. They derived expressions for the Lorenz curve and Gini index for the four models. Results from Kolmogorov Smirnov (KS) test and R<sup>2</sup> coefficient conclude that the Reverse Pareto distribution adequately describes the distribution of income in poor households. Also, the Lorenz curve and Gini index based on Reverse Pareto models showed that poor households have evenly distributed incomes.

The historical evidence, empirical properties, the relationship between distribution functions, and model selection of top income distributions were reviewed by [20]. He summarized that for modeling income, the Generalized Pareto and Generalized Beta family of 3-4 parameters, which includes Dagum, Singh-Maddala, and GB2 distributions are invariably successful.

### 3. Quantile function based Income distributions

Sodomova et al. [21] did the statistical analysis and modeling on a sample of the yearly net income of 1566 households in the Slovak Republic in the year 2002. Results showed that except for the intervals having the lowest and highest income, the Weibull distribution with parameters from maximum likelihood estimation provides the best fit. Whereas for the intervals two fitted distributions from the Weibull-Pareto class, which is obtained by the method of modeling with quantile distributions provided the best fit.

Hankin and Lee [22] introduced and studied distribution with quantile function,

$$Q(u) = \frac{Cu^{\lambda_1}}{(1-u)^{\lambda_2}} \tag{4}$$

where  $C, \lambda_1, \lambda_2 > 0$  and  $0 \leq u \leq 1$ . They studied the properties, shape, comparison with other distributions and calculated the moments of distribution in equation (4). The parameters of this distribution were estimated using maximum likelihood and the method of logged regression. In this paper, they compared the efficiencies of the above two estimation methods using simulation and found when the sample size is small and parameters are roughly equal we can use the regression method otherwise the maximum likelihood method. This distribution was also applied for modeling the toxic gas release. The quantile function in equation (4) is obtained as the product of quantile functions of Pareto and Power distribution, hence it has the potential for income modeling.

Haritha et al. [3] worked on the Modified Lambda family (MLF) with quantile function,

$$Q(u) = \lambda_1 + \frac{1}{\lambda_2} \left[ \frac{u^{\lambda_3} - 1}{\lambda_3} - \frac{(1-u)^{\lambda_4} - 1}{\lambda_4} \right] \tag{5}$$

where  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  are real. The distributional characteristics of MLF were studied in detail. Major income distributions were obtained either as special or limiting cases or by approximation from MLF. They also expressed commonly used income inequality measures in quantile terms and calculated those measures for MLF. The parameters of MLF were evaluated using a new estimation

procedure involving location, dispersion, skewness, and kurtosis in quantile measures. Through a simulation study, it was shown that the above method of estimation is better than the method of percentiles and moments. They characterized income distributions using truncated Gini index and Income Gap Ratio at various ranges of poverty and affluence limit.

Using Dagum's three-parameter type 1 model [23] examined the change in personal income in Spain between 1995 and 2005. The quantile function of this distribution is given as,

$$Q(u) = \lambda^{\frac{1}{\delta}} \left[ \left( \frac{1}{u} \right)^{\frac{1}{\beta}} - 1 \right]^{\frac{-1}{\delta}} \tag{6}$$

where  $0 \leq u \leq 1$  and  $\lambda, \beta, \delta > 0$ . The model in equation (6) fits the empirical income distribution of Spain quite well. The economic interpretation of the Dagum model parameters was also studied. The data from the European Community Household Panel (ECHP) and the European Union Statistics on Income and Living Conditions (EU-SILC) is used to study the effect of parameter changes in the growth of inequality in Spain as well as on different income percentiles.

The distributional and geometric properties of partial moments of first and second order in quantile terms were studied by [24]. The  $r^{\text{th}}$  order partial moment is given by,

$$P_r(u) = \int_u^1 (Q(p) - Q(u))^r dp \tag{7}$$

where  $Q(\cdot)$  is a quantile function. Stop-loss transform based on quantile function is mainly discussed in this paper. Relationships of income inequality measures like Lorenz, Gini, Bonferroni, and Leimkuhler curves with scaled stop-loss transform curves were also developed.

The capabilities of the Zenga curve as an inequality measure were explored by [25]. They made a detailed study on the properties of the Zenga curve and stochastic orders based on this curve were used to prove some results. They established a relationship between the Zenga curve and inequality measures like the Bonferroni curve and the Leimkuhler curve. Similarly, they established an association between the Zenga curve and reliability measures like mean residual quantile function and reversed mean residual quantile function. A study on quantile-based income distributions like the Govindarajulu distribution, quantile model with linear hazard quantile form, and Power  $\times$  Pareto distribution was made and derived measures of income inequality like the Lorenz curve, Bonferroni curve, etc. In this thesis, they also studied the relationship between L-moments and income inequality measures. Bivariate reliability concepts based on copula were also discussed.

Ekum et al. [26] proposed a six-parameter distribution named the exponentiated-exponential Dagum {Lomax} (EEDL) from T- Dagum{Y} family using T-R{Y} framework. Its quantile function is given by

$$Q_x(u) = \phi^{\left(\frac{1}{p}\right)} \left\{ \left[ 1 - \left( 1 - \frac{1}{\beta} \log \left( 1 - u^{\left(\frac{1}{\alpha}\right)} \right) \right)^{-\gamma} \right]^{\left(\frac{1}{q}\right)} - 1 \right\}^{\left(\frac{1}{p}\right)} \tag{8}$$

where  $\alpha, \gamma, q, p$  defines the shape and  $\phi, \beta$  defines the spread of the distribution and  $0 \leq u \leq 1$ . The basic distributional and reliability characteristics including stochastic ordering, asymptotes, analysis of stress-strength, and Shannon entropy were studied. The parameters of EEDL distribution were estimated using the method of maximum likelihood. On two real data sets, the EEDL distribution did well in comparison with the Exponential Kumaraswamy Dagum (EKD), the Exponentiated Generalized Exponential Dagum (EKD), and the Mc Dagum (McD) distributions.

#### 4. New quantile functions in reliability analysis

Nair et al. [27] discussed the method of constructing quantile functions for lifetime models by utilizing the relationship between the hazard quantile function and the Parzens score function [28]. Three models based on score function were illustrated in this paper and many known distributions exist as its special case. Various reliability properties of the Parzen score function were also studied. The reliability ideas in quantile terms were broadly explained by [29]. They also discussed distributions having closed-form quantile functions, ageing concepts, total time on test transforms, L-moments of residual life, hazard quantile function, stochastic orders, and modeling in a quantile framework.

Thomas et al. [30] introduced a software reliability model with a quantile function,

$$Q(u) = k\beta(u, a+1, b+1) \quad (9)$$

where  $k > 0$ ,  $a, b$  are real numbers,  $0 \leq u \leq 1$  and  $\beta(u, a+1, b+1)$  is an incomplete beta function. Various distributional and reliability characteristics of the above quantile function were studied. They approximated equation (9), to two well-known distributions like Inverse Gaussian and Weibull. The method of L-moments was used to estimate the parameters of the model and applied the model in (9) to a real data set.

The reliability properties of the quantile-based proportional hazard model (PHM) were studied by [31]. The ageing properties and characterizations for the PHM were derived and demonstrated with examples. Certain important stochastic orders in the context of PHM were discussed. They also proposed the quantile-based dynamic cumulative residual Kullback-Leibler divergence of PHM.

A new class of distribution as the product of quantile functions of Weibull and Pareto distributions was proposed by [32] and it is given as,

$$Q(u) = \sigma(1-u)^{-\alpha} (-\log(1-u))^\beta \quad (10)$$

where  $0 < u < 1$ ;  $\alpha, \beta, \sigma > 0$ . The distributional characteristics and reliability properties of distribution in equation (10) were studied in detail. The inference was done using L-moments and the model was applied to two real datasets.

Kumar and Paduthol [33] introduced a new class of distributions with quantile functions,

$$Q(u) = \sigma \left( \log \left( 1 + \frac{u}{\beta(1-u)} \right) + \frac{\gamma u}{\beta + (1-\beta)u} \right) \quad (11)$$

where  $\sigma > 0, \beta > 0$  and  $\gamma > -1$ , this class of distributions was obtained as an extension of distributions with linear mean residual quantile function. The distributional characteristics, reliability properties, and L-moments of equation (11) were calculated. The method of percentiles was employed to estimate the parameters of the model in (11). As an application, the above model was applied to real data reported in [34], consisting of the strength of glass fibers.

A class of distributions with quadratic hazard quantile function was developed by [35] and is given as,

$$Q(u) = \frac{1}{2(\beta-\alpha)} \log(1+u) - \frac{1}{2(\beta+\alpha)} \log(1-u) - \frac{\alpha}{(\beta^2-\alpha^2)} \log\left(\frac{\beta+\alpha u}{\beta}\right) \quad (12)$$

where,  $\beta > 0, \beta > |\alpha|, \beta \neq \alpha$ . They studied distributional properties, reliability characteristics, and characterization of the distribution. Estimation was done using the method of least squares and demonstrated the models' utility using real data.

Ghosal et al. [36] used subject-specific quantile functions to capture the distributional nature of wearable data. They used these quantile functions L-moment representations in Scalar-On-Function Regression (SOFR) model [37], Functional Generalized Additive Model (FGAM) [38], and

Joint and Individual Variation Explained(JIVE) method [39]. As an application, they illustrated the proposed method in the study of Alzheimer's Disease(AD).

Power exponential geometric distribution was introduced by [40] as the sum of the quantile function of the power and the exponential geometric distributions. The distributional and reliability properties were studied. A simulation study was done and applied the model to real data.

## 5. Income inequality measures

A book on modeling income distributions and the Lorenz curves was edited by [41]. A compilation of five major papers in this area makes the first part and four survey papers on Lorenz functions, as well as generalizations and extensions of a few income distributions, are included in Part two. Eight papers on recent research and advancement in this field are included in the last section.

The distinction between the Lorenz curve in economics and the Leimkuhler curve in information science is that the Lorenz curve is used to arrange sources in increasing productivity order whereas for Leimkuhler is arranged in decreasing order. A general definition for the Leimkuhler curve was introduced by [42] in terms of theoretical CDF and is given as,

$$K_X(u) = \frac{1}{\mu_X} \int_{1-u}^1 F_X^{-1}(y) dy \quad (13)$$

where  $0 \leq u \leq 1$ . The discrete, continuous, and mixed random variables are covered by the equation (13). In this paper, they derived the Leimkuhler curves expressions for five continuous, one mixed, and one discrete distribution.

The Bonferroni curve and Bonferroni index have applications not only in the field of economics to study poverty and income but also in the field of insurance, medicine, demography, and reliability. For thirty-five continuous distributions, [43] provided explicit expressions for Bonferroni Curve, Bonferroni Index, Lorenz curve, and Gini index.

Fellman [44] studied two optimal cases, where the transformed variable Lorenz dominates the initial variable and the initial variable Lorenz dominates the transformed one. The first case has more practical application than the second because it results in policies that decrease inequality. The properties and limits of the transformed Lorenz curve were also analyzed. In this work, the limits found are valid for a wide range of distributions and transformations but the inequalities that result from pursuing general conditions cannot be improved.

Chotikapanich et al. [45] discussed poverty measures like the head-count ratio, the Foster-Greer-Thorbecke (FGT) measure, the Atkinson index, the Watts index, the Sen index, and the Gini index and derived their expression for GB2 distribution. An analysis of poverty trends in South and South East Asian nations is done using beta 2 distribution, which is a special case of GB2.

Using the semiparametric method, [46] estimated the Lorenz curve and Gini index for the exponential distribution. The above estimation was done under type 1, type 2, and interval censoring. From Monte Carlo simulation studies they found, that as sample size increases, the mean square error (MSE) of the estimator decreases.

Fellman [47] comprehensively explained the Lorenz curve and gave a brief description of the Gini index and Pietra index. As an application, the changes in Lorenz curves, Gini, and Pietra indices with respect to model parameters of Pareto [48], the simplified Rao Tam [49], and the Chotikapanich [50] distributions were also given. When the above three models have the same Gini index, the Lorenz curves for the simplified Rao Tam and the Chotikapanich models are rather similar but Pareto's is different.

Behdani & Mohtashami Borzadaran [51] reformulated certain income inequality measures using quantiles. They used the relationship between the Lorenz curve and reliability concepts like mean residual quantile function and reversed mean residual quantile function to characterize probability distributions. They also studied ageing concepts using the Lorenz curve and quantile function.

Kattumannil et al. [52] proposed a non-parametric estimation of the Gini index for right-censored observations in the sample. The propound estimator has an asymptotic normal distribution and is consistent. Monte Carlo simulation was used to determine the potential of the above estimator. The simulation study showed that the confidence interval of the Gini index based on the proposed estimator has good coverage probability and can be implemented very easily.

A detailed literature review on the association between income inequality and economic growth was done by [53]. To comprehend how income inequality and growth are related, theoretical and empirical literature is studied and analyzed.

Table 1 lists major income inequality measures and curves of the six distributions covered in this review. The distributions having explicit quantile functions and those for which income inequality measures exist in closed form are taken for the calculation of inequality measures.

**Expressions used in Table 1**

- (i) LC, GI, PI, BC, BI, and ZC denote the Lorenz curve, Gini index, Pietra index, Bonferroni curve, Bonferroni index, and Zenga curve respectively.
- (ii)  $\beta_u(.,.)$  denotes the incomplete beta function.
- (iii)  $u_0$  in Pietra indexes can be obtained by solving  $u$  the equation  $\mu = Q(u)$ .
- (iv)  $\text{csc}(\cdot)$  denotes the *coscant*( $\cdot$ ) function in trigonometry.
- (v)  $Har.no(n)$  denotes the  $n^{th}$  harmonic number and is given as  $Har.no(n) = \sum_{k=1}^n k^{-1}$ .
- (vi)  $\Gamma(\cdot)$  denotes the gamma function.
- (vii)  $\Gamma(.,.)$  denotes incomplete gamma function.
- (viii)  $I_u(.,.)$  is the regularized incomplete beta function and is given as  $I_u(.,.) = \frac{\beta_u(.,.)}{\beta(.,.)}$ .
- (ix)  $\frac{1}{\beta(a+1, b+2)} \left\{ \frac{\pi \text{csc}(b\pi) \Gamma(a+2) [Har.no(a+1) - Har.no(a+b+2)]}{\Gamma(-b) \Gamma(a+b+3)} \right\}$
- (x)  $\frac{2ku_0(1-u_0)^{-\frac{1}{2}\alpha} (1-(1-u_0)^{2k})^{\frac{1}{2}\alpha} - \beta_{1-(1-u_0)^{2k}}(\frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha})}{\beta(\frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha})}$
- (xi)  $\frac{1}{\mu} \left\{ (\gamma-1)^{-1} (\gamma-2)^{-1} (1-u)^{-\gamma} \left[ (1-u)^\gamma + u(u+\gamma-u\gamma) - 1 \right] \eta + \lambda u \right.$   
 $\left. + 2^{-(\beta+1)} \eta \left[ \Gamma(\beta+1) - \Gamma(\beta+1, -2\ln(1-u)) \right] \right\}$
- (xii)  $\frac{-3^{-(\beta+1)} \eta \left[ 6^{\beta+1} (\gamma+1) + (2^{\beta+2} - 3^{\beta+1}) (\gamma-1) (\gamma-2) (\gamma-3) \Gamma(\beta+1) \right]}{(\gamma-3) \left\{ 2^{\beta+1} [\eta + (\gamma-1) (\gamma-2) \lambda] + \eta (\gamma-1) (\gamma-2) \Gamma(\beta+1) \right\}}$
- (xiii)  $\frac{\eta}{\mu} \left\{ u_0 (1-u_0) \left[ -\ln(1-u_0) \right]^\beta + (\gamma-1)^{-1} (\gamma-2)^{-1} (1-u_0)^{-\gamma} \left[ u_0^2 (\gamma-1)^2 - u_0 \gamma + 1 - (1-u_0)^\gamma \right] \right.$   
 $\left. - 2^{-(\beta+1)} \left[ \Gamma(\beta+1) - \Gamma(\beta+1, -2\ln(1-u_0)) \right] \right\}$
- (xiv)  $\frac{1}{\mu} \left\{ (\gamma-1)^{-1} (\gamma-2)^{-1} u^{-1} (1-u)^{-\gamma} \left[ (1-u)^\gamma + u(u+\gamma-u\gamma) - 1 \right] \eta + \lambda \right.$   
 $\left. + 2^{-(\beta+1)} \eta u^{-1} \left[ \Gamma(\beta+1) - \Gamma(\beta+1, -2\ln(1-u)) \right] \right\}$

$$(xv) \quad 1 - \int_0^1 \frac{1}{\mu} \{(\gamma - 1)^{-1}(\gamma - 2)^{-1}u^{-1}(1 - u)^{-\gamma} [(1 - u)^\gamma + u(u + \gamma - u\gamma) - 1]\} \eta + \lambda \\ + 2^{-(\beta+1)} \eta u^{-1} [\Gamma(\beta + 1) - \Gamma(\beta + 1, -2 \ln(1 - u))] du$$

(xvi) In expressions (xi), (xiii), (xiv), and (xv),  $\mu$  is the mean of the Weibull-Pareto distribution and is given as  $\mu = \lambda + \frac{\eta}{(\gamma - 1)(\gamma - 2)} + 2^{-(\beta+1)} \eta \Gamma(\beta + 1)$

**Table 1:** Income inequality measures and inequality curves

Quantile function		Income inequality measures/ inequality curves
		$\frac{[u\beta_u(a + 1, b + 1) - \beta_u(a + 2, b + 1)]}{\beta(a + 1, b + 2)}$
	LC	
	GI	$(a + 1)(a + b + 3)^{-1} \frac{\beta_{u_0}(a + 2, b + 1)}{\beta(a + 1, b + 2)}$
	PI	
Bijamma Thomas distribution	BC	$\frac{1}{\beta(a + 1, b + 2)} \left[ \beta_u(a + 1, b + 1) - \frac{1}{u} \beta_u(a + 2, b + 1) \right]$
$Q(u) = k\beta(u, a + 1, b + 1)$	BI	(ix)
	ZC	$\frac{\beta(a + 1, b + 2) - \beta_u(a + 1, b + 1) + \frac{1}{u} \beta_u(a + 2, b + 1)}{\beta(a + 1, b + 2) - u\beta_u(a + 1, b + 1) + \beta_u(a + 2, b + 1)}$
		$\frac{u(\ln u + \lambda x_0 - 1)}{\lambda x_0 - 1}$
	LC	
	GI	$\frac{1}{2(\lambda x_0 - 1)}$
	PI	$\frac{u_0}{\lambda x_0 - 1}$
Shifted reverse exponential distribution	BC	$\frac{\ln u + \lambda x_0 - 1}{\lambda x_0 - 1}$
$Q(u) = x_0 + \frac{1}{\lambda} \ln(u)$	BI	$\frac{1}{\lambda x_0 - 1}$
	ZC	$\frac{\ln u}{1 + u(\ln u - 1) - \lambda x_0(1 - u)}$
		$\frac{x_0 u - \lambda \Gamma(1 + \frac{1}{\rho}, -\ln u)}{x_0 - \lambda \Gamma(1 + \frac{1}{\rho})}$
	LC	
	GI	$(1 - 2^{-\frac{1}{\rho}}) \frac{\lambda \Gamma(1 + \frac{1}{\rho})}{\left[ x_0 - \lambda \Gamma(1 + \frac{1}{\rho}) \right]}$
Shifted reverse stretched exponential distribution	PI	$\frac{\lambda \Gamma(1 + \frac{1}{\rho}, -\ln u_0) - \lambda u_0 (-\ln u_0)^{\frac{1}{\rho}}}{x_0 - \lambda \Gamma(1 + \frac{1}{\rho})}$
$Q(u) = x_0 - \lambda(-\ln u)^{\frac{1}{\rho}}$	BC	$\frac{x_0 - \lambda u^{-1} \Gamma(1 + \frac{1}{\rho}, -\ln u)}{x_0 - \lambda \Gamma(1 + \frac{1}{\rho})}$



	BI	$\frac{\lambda\Gamma(1 + \frac{1}{\rho})}{x_0\rho - \lambda\Gamma(\frac{1}{\rho})}$
	ZC	$\frac{\lambda \left[ \Gamma(1 + \frac{1}{\rho}, -\ln u) - u\Gamma(1 + \frac{1}{\rho}) \right]}{u \left\{ x_0(1 - u) + \lambda \left[ \Gamma(1 + \frac{1}{\rho}, -\ln u) - \Gamma(1 + \frac{1}{\rho}) \right] \right\}}$
	LC	$I_{1-(1-u)^{2k}} \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right)$
	GI	$1 - \frac{2\beta \left( \frac{1}{\alpha} + 1, \frac{1}{k} - \frac{1}{2\alpha} \right)}{\beta \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right)}$
K-Generalized distribution	PI	$\frac{1}{u} I_{1-(1-u)^{2k}}^{(x)} \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right)$
$Q(u) = \beta \left[ \ln_K \left( \frac{1}{1-u} \right) \right]^{\frac{1}{\alpha}}$	BC	$1 - \int_0^1 \frac{1}{u} I_{1-(1-u)^{2k}} \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right) du$
	BI	$\frac{u\beta \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right) - \beta_{1-(1-u)^{2k}} \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right)}{u \left[ \beta \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right) - \beta_{1-(1-u)^{2k}} \left( \frac{1}{\alpha} + 1, \frac{1}{2k} - \frac{1}{2\alpha} \right) \right]}$
	ZC	$\frac{\beta + 2}{(\beta + 2)\theta + 2\sigma} \left[ \theta u + \sigma u^{\beta+1} \left( \frac{\beta + 2 - \beta u}{\beta + 2} \right) \right]$
		LC
Govindarajulu distribution	PI	$\frac{\sigma\beta u_0^{\beta+1}}{[(\beta + 2)\theta + 2\sigma]} [(\beta + 2) - (\beta + 1)u_0]$
$Q(u) = \theta + \sigma[(\beta + 1)u^\beta - \beta u^{\beta+1}]$	BC	$\frac{\beta + 2}{(\beta + 2)\theta + 2\sigma} \left[ \theta + \sigma u^\beta \left( \frac{\beta + 2 - \beta u}{\beta + 2} \right) \right]$
	BI	$\frac{\sigma\beta(2\beta + 3)}{(\beta + 1)(\beta + 2)[(\beta + 2)\theta + 2\sigma]}$
		$\frac{(\beta + 2)}{(\beta + 2)\theta + 2\sigma} - \left[ \theta + \sigma u^\beta \left( \frac{\beta + 2 - \beta u}{\beta + 2} \right) \right]$
	ZC	$\frac{(\beta + 2)}{(\beta + 2)\theta + 2\sigma} - u \left[ \theta + \sigma u^\beta \left( \frac{\beta + 2 - \beta u}{\beta + 2} \right) \right]$
Weibull-Pareto distribution	LC	(xi)
	GI	(xii)
$Q(u) = \lambda + \eta \left\{ (1-u)[- \ln(1-u)]^\beta + \frac{u}{(1-u)^\gamma} \right\}$	PI	(xiii)
	BC	(xiv)
	BI	(xv)
	ZC	(xvii)

$$(xvii) \quad \frac{\eta \left\{ \frac{2(1-u)^{-\gamma+1} [(1-u)^\gamma + \gamma u - u - 1]}{(\gamma-1)(\gamma-2)} + 2^{-\beta} [\beta\Gamma(\beta) - u\Gamma(\beta+1) - \Gamma(\beta+1, -2\ln(1-u))] \right\}}{u(1-u) \left[ \frac{(1-u)^{-\gamma} (\gamma u - u - 1)\eta - \lambda}{(\gamma-1)(\gamma-2)} - 2^{-(\beta+1)} \eta u \Gamma(\beta+1, -2\ln(1-u)) \right]}$$

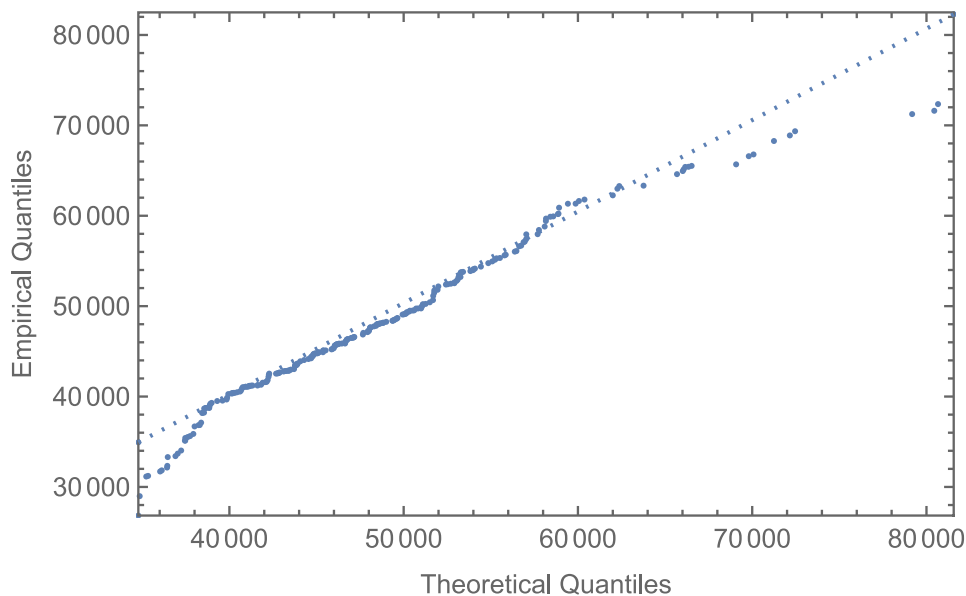
## 6. Data Analysis

This section analyses 2020's per capita personal income (in dollars) of 254 counties in Texas, US. This data is available from <https://www.bea.gov> and is used for studying the potential of Shifted Reverse Exponential, Govindarajulu, and Weibull Pareto distribution in income modeling. Here we use the method of percentiles for estimating the parameters of the above three distributions.

Here, the Chi-square ( $\chi^2$ ) test and Q-Q plot are used to determine model adequacy. Table 2 provides the parameter estimates,  $\chi^2$  test statistics, and p-values of Shifted Reverse Exponential, Govindarajulu, and Weibull Pareto distributions. Table 2 and the Q-Q plot given in Figure 1 make it evident that the Weibull Pareto distribution provides the best fit for the real data.

**Table 2:** Parameter Estimates,  $\chi^2$  Statistic, p-value

Distribution	Parameter Estimates	$\chi^2$ Statistic	p-value
Shifted Reverse Exponential	$x_0 = 5.472113 \times 10^4$ $\lambda = 8.922956 \times 10^{-5}$	68.39832	$8.41825 \times 10^{-8}$
Govindarajulu	$\theta = 3.85826 \times 10^4$ $\sigma = 2.605615 \times 10^4$ $\beta = 2.943906$	78.84319	$1.36464 \times 10^{-9}$
Weibull Pareto	$\lambda = 3.429882 \times 10^4$ $\eta = 1.177798 \times 10^4$ $\beta = 5.374932 \times 10^{-1}$ $\gamma = 4.088145 \times 10^{-1}$	16.73783	0.54119



**Figure 1:** Q-Q plot corresponding to Weibull Pareto distribution

## 7. Conclusion

In this work, we reviewed recent income distributions, income inequalities, and quantile functions that appeared in the literature. This study is carried out in five sections comprising income distributions, income models based on quantiles, new quantile functions in reliability analysis, income inequality measures, and data analysis. For the six distributions examined in this work, the Lorenz curve, Gini index, Pietra index, Bonferroni index, Bonferroni curve, and Zenga curve were determined. Three models were applied to the per capita personal income data of 254 counties in Texas State and found that the Weibull Pareto distribution provides the best fit. In future works, we can check whether more quantile functions used in reliability analysis have potential in income modeling.

### Acknowledgment

The first author is thankful to Kerala State Council for Science, Technology and Environment (KSCSTE) for the financial support.

### Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

### References

- [1] N. C. Kakwani, *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. 1980.
- [2] C. Dagum, "Inequality Measures between Income Distributions with Applications," *Econometrica*, vol. 48, no. 7, pp. 1791, 1980, doi: 10.2307/1911936.
- [3] H. Haritha N., K. R. M. Nair, and N. U. Nair, "Income Modeling Using Quantile functions," 2007.
- [4] A. Tarsitano, "Fitting the Generalized Lambda," *Compstat 2004 Symposium*, 2004, pp. 1861–1867.
- [5] W. G. Gilchrist, *Statistical Modeling with Quantile Functions*. Chapman and Hall/CRC, New York, 2000.
- [6] J. B. McDonald, R. Bandourian, and R. S. Turley, "A Comparison of Parametric Models of Income Distribution Across Countries and Over Time," *SSRN Electron. J.*, 2002.
- [7] S. Huang and B. O. Oluyede, "Exponentiated Kumaraswamy-Dagum distribution with applications to income and lifetime data," *J. Stat. Distrib. Appl.*, vol. 1, 2014, doi: 10.1186/2195-5832-1-8.
- [8] S. Mori, D. Nakata, and T. Kaneda, "An Application of Gamma Distribution to the Income Distribution and the Estimation of Potential Food Demand Functions," *Mod. Econ.*, vol. 06, no. 09, pp. 1001–1017, 2015, doi: 10.4236/me.2015.69095.
- [9] F. Clementi, M. Gallegati, G. Kaniadakis, and S. Landini, " $\kappa$ -generalized models of income and wealth distributions: A survey," *Eur. Phys. J. Spec. Top.*, vol. 225, no. 10, pp. 1959–1984, Oct. 2016, doi: 10.1140/epjst/e2016-60014-2.
- [10] F. Clementi, M. Gallegati, and G. Kaniadakis, " $\kappa$ -generalized statistics in personal income distribution," *Eur. Phys. J. B*, vol. 57, no. 2, pp. 187–193, 2007, doi: 10.1140/epjb/e2007-00120-9.
- [11] F. Clementi, T. Di Matteo, M. Gallegati, and G. Kaniadakis, "The  $\kappa$ -generalized distribution: A new descriptive model for the size distribution of incomes," *Phys. A Stat. Mech. its Appl.*, vol. 387, no. 13, pp. 3201–3208, 2008, doi: 10.1016/j.physa.2008.01.109.
- [12] F. Clementi, M. Gallegati, and G. Kaniadakis, "A  $\kappa$ -generalized statistical mechanics approach to income analysis," *J. Stat. Mech. Theory Exp.*, vol. 2009, no. 2, 2009, doi: 10.1088/1742-5468/2009/02/P02037.
- [13] M. H. Tahir, G. M. Cordeiro, A. Alzaatreh, M. Mansoor, and M. Zubair, "A New Weibull–Pareto Distribution: Properties and Applications," *Commun. Stat. Simul. Comput.*, vol. 45, no. 10, pp. 3548–3567, 2016, doi: 10.1080/03610918.2014.948190.

- [14] E. Calderín-Ojeda, F. Azpitarte, and E. Déniz, "Modelling income data using two extensions of the exponential distribution," *Phys. A Stat. Mech. its Appl.*, vol. 461, pp. 756–766, 2016, doi: 10.1016/j.physa.2016.06.047.
- [15] S. A. A. Bakar and D. Pathmanathan, "Income modeling with the Weibull mixtures," *Commun. Stat. - Theory Methods*, 2020, doi: 10.1080/03610926.2020.1800737.
- [16] M. A. M. Safari, N. Masseran, K. Ibrahim, and S. I. Hussain, "Modeling the income distribution of poor households in Malaysia," *AIP Conf. Proc.*, vol. 2266, no. October, 2020, doi: 10.1063/5.0018066.
- [17] M. A. M. Safari, N. Masseran, K. Ibrahim, and S. I. Hussain, "A robust and efficient estimator for the tail index of inverse Pareto distribution," *Phys. A Stat. Mech. its Appl.*, vol. 517, pp. 431–439, 2019, doi.org/10.1016/j.physa.2018.11.029.
- [18] C. Kleiber and S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons, Ltd, 2003.
- [19] M. Brzezinski, "Power laws in citation distributions: evidence from Scopus," *Scientometrics*, vol. 103, no. 1, pp. 213–228, 2015, doi: 10.1007/s11192-014-1524-z.
- [20] V. Hlasny, "Parametric representation of the top of income distributions: Options, historical evidence, and model selection," *J. Econ. Surv.*, vol. 35, no. 4, pp. 1217–1256, 2021, doi.org/10.1111/joes.12435.
- [21] E. Sodomova, B. Pacak, and L. Sipkova, "Models of households in the Slovak Republic .," *Statistics in management of social and economic development: 11th Ukrainian-Polish-Slovak scientific conference, 20–22 October 2004*, 2005, pp. 80–91.
- [22] R. K. S. Hankin and A. Lee, "A new family of non-negative distributions," *Aust. New Zeal. J. Stat.*, vol. 48, no. 1, pp. 67–78, 2006, doi: 10.1111/j.1467-842X.2006.00426.x.
- [23] C. G. Pérez and M. P. Alaiz, "Using the Dagum model to explain changes in personal income distribution," *Appl. Econ.*, vol. 43, no. 28, pp. 4377–4386, 2011, doi: 10.1080/00036846.2010.491459.
- [24] N. U. Nair, P. G. Sankaran, and S. M. Sunoj, "Quantile based stop-loss transform and its applications," *Stat. Methods Appl.*, vol.22, no.2, pp.167–182, 2013,doi:10.1007/s10260-012-0213-4.
- [25] N. Sreelakshmi and K. Nair, "A quantile based analysis of income data," 2014.
- [26] M. I. Ekum, M. O. Adamu, and E. E. Akarawak, "T-Dagum: A Way of Generalizing Dagum Distribution Using Lomax Quantile Function," *J. Probab. Stat.*, vol. 2020, 2020, doi: 10.1155/2020/1641207.
- [27] N. U. Nair, P. G. Sankaran, and B. V. Kumar, "Modelling lifetimes by quantile functions using Parzen's score function," *Statistics (Ber.)*, vol. 46, no. 6, pp. 799–811, 2012, doi: 10.1080/02331888.2011.555551.
- [28] E. Parzen, "Nonparametric Statistical Data Modeling," *J. Am. Stat. Assoc.*, vol. 74, no. 365, pp. 105–121, Mar. 1979, doi: 10.1080/01621459.1979.10481621.
- [29] N. Nair, P. G. Sankaran, and N. Balakrishnan, *Quantile-Based Reliability Analysis*. 2013.
- [30] B. Thomas, M. N. Nellikkattu, and S. Godan Paduthol, "A software reliability model using quantile function," *J. Probab. Stat.*, vol. 2014, no. March, 2014, doi: 10.1155/2014/951608.
- [31] N. U. Nair, P. G. Sankaran, and S. M. Sunoj, "Proportional hazards model with quantile functions," *Commun. Stat. - Theory Methods*, vol. 47, no. 19, pp. 4710–4723, 2018, doi: 10.1080/03610926.2018.1445858.
- [32] S. Paduthol and D. Kumar, "Power Pareto quantile function," *J. Appl. Probab. Stat.*, vol. 13, pp. 81–95, 2018.
- [33] D. Kumar Maladan and P. G. Sankaran, "A new family of quantile functions and its applications," *Commun. Stat. - Theory Methods*, vol. 50, no. 18, pp. 4216–4235, 2020, doi: 10.1080/03610926.2020.1713368.
- [34] R. L. Smith and J. C. Naylor, "A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-Parameter Weibull Distribution," *J. R. Stat. Soc. Ser. C (Applied Stat.)*, vol. 36, no. 3, pp. 358–369, 1987, doi.org/10.2307/2347795.

- [35] I. C. Aswin, P. G. Sankaran, and S. M. Sunoj, "A Class of Distributions with Quadratic Hazard Quantile Function," *J. Indian Soc. Probab. Stat.*, vol. 21, no. 2, pp. 409–426, 2020, doi: 10.1007/s41096-020-00088-6.
- [36] R. Ghosal *et al.*, "Distributional data analysis via quantile functions and its application to modeling digital biomarkers of gait in Alzheimer's Disease," *Biostatistics*, 2021, doi: 10.1093/biostatistics/kxab041.
- [37] J. S. Morris, "Functional Regression," *Annu. Rev. Stat. Its Appl.*, vol. 2, no. 1, pp. 321–359, Apr. 2015, doi: 10.1146/annurev-statistics-010814-020413.
- [38] M. W. McLean, G. Hooker, A.-M. Staicu, F. Scheipl, and D. Ruppert, "Functional Generalized Additive Models," *J. Comput. Graph. Stat.*, vol. 23, no. 1, pp. 249–269, Jan. 2014, doi: 10.1080/10618600.2012.729985.
- [39] E. F. Lock, K. A. Hoadley, J. S. Marron, and A. B. Nobel, "Joint and individual variation explained (JIVE) for integrated analysis of multiple data types," *Ann. Appl. Stat.*, vol. 7, no. 1, pp. 523–542, Mar. 2013, doi: 10.1214/12-AOAS597.
- [40] J. Joseph and A. A.P., "Power -Exponential Geometric Quantile Function," *Reliab. Theory Appl.*, vol. 16, no. 4, pp. 294–307, 2021.
- [41] D. Chotikapanich, Ed., *Modeling Income Distributions and Lorenz Curves*. Springer, 2008.
- [42] J. M. Sarabia, "A general definition of the Leimkuhler curve," *J. Informetr.*, vol. 2, no. 2, pp. 156–163, 2008, doi: 10.1016/j.joi.2008.01.002.
- [43] G. M. Giorgi and S. Nadarajah, "Bonferroni and Gini indices for various parametric families of distributions," *Metron*, vol. 68, no. 1, pp. 23–46, 2010, doi: 10.1007/BF03263522.
- [44] J. Fellman, "Properties of Lorenz Curves for Transformed Income Distributions," *Theor. Econ. Lett.*, vol. 02, no. 05, pp. 487–493, 2012, doi: 10.4236/tel.2012.25091.
- [45] D. Chotikapanich, W. Griffiths, W. Karunarathne, and D. S. Prasada Rao, "Calculating Poverty Measures from the Generalised Beta Income Distribution," *Econ. Rec.*, vol. 89, no. S1, pp. 48–66, 2013, doi.org/10.1111/1475-4932.12031.
- [46] P. P. C. Pillai, G. Rajesh, and E. I. Abdul-Sathar, "Semi-Parametric Estimation of Lorenz Curve and Gini-index in an Exponential Distribution," *Sri Lankan J. Appl. Stat.*, vol. 15, no. 1, 2014, doi: 10.4038/sljastats.v15i1.6749.
- [47] J. Fellman, "Income Inequality Measures," *Theor. Econ. Lett.*, vol. 08, no. 03, pp. 557–574, 2018, doi: 10.4236/tel.2018.83039.
- [48] R. H. Rasche, J. Gaffney, A. Y. C. Koo, and N. Obst, "Functional Forms for Estimating the Lorenz Curve," *Econometrica*, vol. 48, no. 4, pp. 1061–1062, 1980, [Online]. Available: <http://www.jstor.org/stable/1912948>.
- [49] U. L. Gouranga Rao and A. Yuk-Pang Tam, "An empirical study of selection and estimation of alternative models of the Lorenz curve," *J. Appl. Stat.*, vol. 14, no. 3, pp. 275–280, Jan. 1987, doi: 10.1080/02664768700000032.
- [50] D. Chotikapanich, "A comparison of alternative functional forms for the Lorenz curve," *Econ. Lett.*, vol. 41, no. 2, pp. 129–138, 1993, doi.org/10.1016/0165-1765(93)90186-G.
- [51] Z. Behdani and G. R. Mohtashami Borzadaran, "Measures of Income Inequality Based on Quantile Function," 2019.
- [52] S. K. Kattumannil, I. Dewan, and S. N., "Non-parametric estimation of Gini index with censored observations," *Stat. Probab. Lett.*, vol. 175, no. April, 2021, doi: 10.1016/j.spl.2021.109113.
- [53] K. Mdingi and S.-Y. Ho, "Literature review on income inequality and economic growth," *MethodsX*, vol. 8, p. 101402, 2021, doi.org/10.1016/j.mex.2021.101402.