

# COMPARATIVE ANALYSIS OF PARAMETRIC LORENZ CURVES AND ALTERNATIVE CONVEX MODELS WITH ISOTONIC REGRESSION FOR ESTIMATING THE LORENZ CURVE

Z. BEHDANI  

Article type: Research Article

(Received: 02 May 2025, Received in revised form 30 November 2025)

(Accepted: 05 February 2026, Published Online: 05 February 2026)

**ABSTRACT.** This study develops a systematic comparative framework for estimating Lorenz curves and Gini coefficients, addressing key methodological gaps in measuring income inequality. We employ the Generalized Mean Squared Error (GMSE) to compare several parametric models (such as polynomial, beta, and established functional forms) with isotonic regression as a non-parametric alternative. Extensive Monte Carlo simulations using log-normal and Pareto distributions show that isotonic regression consistently achieves higher accuracy than parametric approaches. An application to Iranian household income data ( $n = 18,809$ ) further confirms these results. Based on data characteristics and research objectives, the findings offer practical guidance for selecting appropriate estimation methods.

*Keywords:* Lorenz curve; Gini coefficient; isotonic regression; parametric models; income inequality; Generalized Mean Squared Error.

*2020 MSC:* 62P20, 62G08, 91B82.

## 1. Introduction

Income distribution has changed over the years, from early economic ideas to modern studies of inequality. Classical economists such as David Rodrik [36] and Adam Smith [41] advanced the analysis of economic inequality, with Smith's concept of the "invisible hand" and Pareto's "80/20 rule" laying the groundwork for understanding income disparity ([28]). Amartya Sen [38] broadened the study of inequality by examining individuals' real freedoms and opportunities, focusing on the capability approach. Theories have been formulated to explain income inequality, encompassing neoclassical economic models, structuralist theories, and empirical studies of the relationship between industrial expansion and income disparity. However, recent research indicates that inequality may persist or intensify in advanced economies ([32]).

---

✉ behdani@bkatu.ac.ir, ORCID: 0000-0002-0029-4833

<https://doi.org/10.22103/jmmr.2026.25196.1795>

Publisher: Shahid Bahonar University of Kerman

How to cite: Z. Behdani, *Comparative analysis of parametric Lorenz curves and alternative convex models with isotonic regression for estimating the Lorenz curve*, J. Mahani Math. Res. 2026; 15(2): 229-249.



© the Author(s)

The Lorenz curve (LC) is a widely used tool to measure income inequality, illustrating the distribution of wealth among individuals. It is crucial for identifying and addressing economic inequality, as it adapts to changes in income distribution. Accurately estimating the LC can be challenging, especially with real-world data that exhibit diverse distributional structures. Researchers have developed parametric and non-parametric models to capture the subtleties of income distribution, allowing for more flexible and precise estimation of LCs. The Gini coefficient, derived from the LC, remains a central measure for summarizing inequality.

Regression techniques have long been applied to model relationships between variables ([2, 9, 16, 20]). Isotonic regression, a non-parametric method, fits a monotonic function to the data and has been widely applied in economics, medicine, and quality control for modeling monotonic relationships ([4, 26, 35]). Its sole requirement of monotonicity aligns naturally with the fundamental properties of the Lorenz curve.

Several studies have explored the effectiveness of various parametric models for fitting the Lorenz curve. Felman ([14, 15]) demonstrated the utility of polynomial regression in capturing the shape of empirical LCs. Sitthiyot and Holasut [43] critically evaluated the single-parameter functional form proposed by Paul and Shankar [31] using income data from 40 countries, showing that more flexible functional forms generally provide a better fit to real income distributions. Shen and Dai [40] further examined the estimation of the Gini index using Kakwani's three-parameter Lorenz curve [21], comparing regression-based approaches with traditional error minimization techniques across sixteen countries. Their results highlighted the superior performance of regression methods in estimating Gini coefficients and fitting income shares by deciles, particularly in economies with medium to high inequality.

Key methodological contributions include Paul and Shankar's [31] single-parameter functional form, which offered a parsimonious approach to represent inequality, and Giudici and Raffinetti's [19] work on Lorenz model selection, providing practical guidance for choosing among competing functional forms. Recent comparative studies, such as those by Sitthiyot and Holasut [42, 43], have systematically assessed the performance of various parametric forms under different distributional scenarios, while Fajar et al. [13] introduced new measures of Lorenz curve asymmetry along with associated hypothesis testing frameworks.

Historical reviews, including Shen [39], trace the evolution of Lorenz curve modeling and highlight the progressive refinement of estimation techniques over decades. Methodological innovations continue to emerge: Fajar and Iriawan [12] proposed adjusted signal-to-noise ratio measures for selecting Lorenz functions; Shen and Dai [40] developed regression-based methods for estimating the Gini index using decile data; and more recently, Dai and Shen [11] introduced novel approaches for Gini coefficient estimation from quantile data, improving accuracy under certain distributional conditions. In parallel, the

development of specialized statistical distributions, such as the Unit New Half Logistic Distribution by Karakaya and Sağlam [24], has expanded the tools available for modeling income inequality and fitting Lorenz curves.

Despite these substantial advances, a comprehensive comparative framework that evaluates both parametric and non-parametric approaches, including isotonic regression, under unified simulation conditions remains underdeveloped. Our study addresses this gap by providing an empirical comparison of multiple estimation techniques across various distributional scenarios and sample sizes. The primary contributions of this paper are threefold: (1) providing the first systematic evaluation of isotonic regression against multiple parametric models under controlled simulation conditions; (2) introducing a rigorous framework for assessing estimator performance across diverse scenarios; and (3) offering practical guidelines for researchers in selecting appropriate estimation methods based on empirical evidence rather than theoretical assumptions alone.

The remainder of the paper is organized as follows: Section 2 provides the theoretical background and literature review, covering essential concepts such as the Lorenz curve, the Gini coefficient, and established parametric models of the LC. Section 3 details the methodology, including parametric and non-parametric regression models for Lorenz curve estimation, the model evaluation framework using Generalized Mean Squared Error (GMSE), and the comparative framework philosophy. Section 4 presents a simulation study comparing model performance using data generated from log-normal and Pareto distributions. Section 5 applies the models to real Iranian household income data. Section 6 describes the software and implementation details. The final section summarizes the findings and conclusions of the study.

## 2. Theoretical Background and Literature Review

This section lays the theoretical foundation for the study by reviewing essential concepts related to the Lorenz curve, the Gini coefficient, and established parametric models used for their estimation.

**2.1. The Lorenz Curve.** The Lorenz curve is a fundamental graphical tool for representing the distribution of income or wealth within a population. It plots the cumulative proportion of income received against the cumulative proportion of households or individuals when the latter are ranked from poorest to richest ([27]). A curve that coincides with the 45-degree line (the line of perfect equality) indicates perfect equality, where each percentile of the population receives the same percentile of total income. As inequality increases, the Lorenz curve bows away from the line of equality.

Mathematically, for a cumulative distribution function  $F(x)$  with a finite mean  $\mu > 0$ , the Lorenz curve  $L(p)$  is defined as:

$$(1) \quad L(p) = \frac{1}{\mu} \int_0^p F^{-1}(t) dt, \quad 0 \leq p \leq 1,$$

where  $F^{-1}(t) = \inf\{x : F(x) \geq t\}$  is the quantile function [25]. From this definition, key theoretical properties of any valid Lorenz curve follow:

- **Boundary Conditions:**  $L(0) = 0$  and  $L(1) = 1$ .
- **Monotonicity:**  $L(p)$  is a non-decreasing function of  $p$ .
- **Convexity:**  $L(p)$  is a convex function, i.e.,  $L''(p) \geq 0$  for  $p \in (0, 1)$ .

The empirical Lorenz curve is constructed from sample data by ordering incomes  $x_1 \leq x_2 \leq \dots \leq x_n$ , calculating the cumulative proportions of the population  $p_i = i/n$ , and the cumulative share of income

$L(p_i) = \left(\sum_{j=1}^i x_j\right) / \left(\sum_{j=1}^n x_j\right)$ . The connection between the Lorenz curve and the underlying income distribution is profound. Parzen [29] showed that if  $F(x)$  has a continuous density  $f(x)$ , the quantile function's derivative is  $\frac{dF^{-1}(t)}{dt} = 1/f(F^{-1}(t))$ . Building on this, Arnold and Villasenor [3] established that if the second derivative of the Lorenz function  $L''(p)$  is continuous and positive on an interval, the density of the distribution can be recovered as:

$$(2) \quad f(x) = \frac{1}{\mu L''(F(x))}.$$

This highlights the Lorenz curve's utility not just as a descriptive tool, but also as an instrument for analyzing statistical distributions themselves. See some applications of the Lorenz curve in various sciences in [5–8]

**2.2. The Gini Coefficient.** The Gini coefficient is the most widely used summary measure of inequality derived from the Lorenz curve. Proposed by [18], it measures the extent to which the actual income distribution deviates from perfect equality. Geometrically, it is defined as the ratio of the area between the line of equality and the Lorenz curve (area  $C$ ) to the total area under the line of equality (area  $C + D = 1/2$ ). Thus,

$$(3) \quad G = \frac{C}{C + D} = \frac{C}{1/2} = 2C.$$

Since  $D = \int_0^1 L(p) dp$  is the area under the Lorenz curve and  $C = 1/2 - D$ , the Gini coefficient can be equivalently expressed as:

$$(4) \quad G = 1 - 2 \int_0^1 L(p) dp.$$

This integral formula is particularly useful for computation. The value of  $G$  ranges from 0 (complete equality) to 1 (complete inequality). The interested reader is referred to Kleiber and Kotz [25] for further reading on this topic.

**2.3. Parametric Families of Lorenz Curves.** Estimating the Lorenz curve from empirical data can be approached non-parametrically (e.g., connecting empirical points) or parametrically. Parametric estimation involves fitting a specific functional form that satisfies the properties of a Lorenz curve. This approach can provide smooth estimates and often requires fewer parameters

than a fully non-parametric fit, especially with limited data. Over the years, numerous parametric forms have been proposed.

One of the earliest and simplest parametric forms is based on the Pareto distribution ([28]):

$$(5) \quad L(p) = 1 - (1 - p)^{1-1/\alpha}, \quad \alpha > 1.$$

Kakwani and Podder ([22, 23]) proposed a flexible two-parameter form:

$$(6) \quad L(p) = p^\delta e^{-\eta(1-p)}, \quad \eta > 0, 1 < \delta < 2.$$

Rasche et al. ([33]) suggested a generalization of the Pareto Lorenz curve:

$$(7) \quad L(p) = [1 - (1 - p)^\eta]^{1/\delta}, \quad \eta, \delta > 0.$$

Chotikapanich ([10]) proposed a form using the exponential function:

$$(8) \quad L(p) = \frac{e^{\alpha p} - 1}{e^\alpha - 1}, \quad \alpha > 0.$$

Aggarwal ([1]) introduced a model based on a ratio of polynomials:

$$(9) \quad L(p) = \frac{(1 - \alpha)^2 p}{(1 + \alpha)^2 - 4\alpha p}, \quad 0 < \alpha < 1.$$

Sarabia et al. ([37]) developed a hierarchical family of Lorenz curves, generalizing simpler forms:

$$\begin{aligned} L_0(p) &= 1 - (1 - p)^\alpha, \\ L_1(p) &= p^\alpha [1 - (1 - p)^\beta], \\ L_2(p) &= [1 - (1 - p)^\alpha]^\gamma, \\ L_3(p) &= p^\eta [1 - (1 - p)^\alpha]^\gamma. \end{aligned}$$

More recent models include the polynomial regression model, which offers outstanding flexibility, and the beta distribution model, which is naturally suited for bounded data on the unit interval ([10]). The Rasch model has also been adapted for income distribution analysis ([34]). Each of these models has an associated formula for the Gini coefficient, allowing for direct estimation of inequality based on the model parameters.

For parametric models whose functional forms inherently satisfy boundary conditions, we ensured that the parameters were estimated within ranges preserving monotonicity and convexity. Consequently, these constrained parametric models also maintain the convex combination property.

**2.4. Overview of Regression Techniques.** Regression analysis is a cornerstone of statistical modeling, used to describe and quantify relationships between variables. Its origins trace back to the work of Galton ([17]) on heredity. The core idea is to model a dependent variable  $Y$  as a function of one or more

independent variables  $X$ , often with an error term  $\epsilon$  accounting for unexplained variation. The simplest form, linear regression, is expressed as:

$$(10) \quad Y = \beta_0 + \beta_1 X + \epsilon.$$

This framework was extended to multiple linear regression, incorporating several predictors:

$$(11) \quad Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon.$$

Multiple regression allows for analyzing the effect of each predictor while controlling for others, which is invaluable in fields like economics [2, 9].

### 3. Methodology: Regression Models for Lorenz Curve Estimation

This section details the regression methodologies employed for estimating the Lorenz curve, encompassing both parametric and non-parametric approaches. It also outlines the procedures implemented to ensure the estimated curves adhere to the essential theoretical properties of a valid Lorenz curve.

**3.1. Parametric Models.** Parametric approaches assume a specific functional form for the Lorenz curve, defined by a set of parameters. The model is fit by estimating these parameters from the data. In the following, some parametric models that are suitable for fitting the Lorenz curve are introduced.

**3.1.1. Polynomial Regression.** Polynomial regression models the relationship between  $p$  and  $L(p)$  as an  $k$ th-degree polynomial:

$$(12) \quad L(p) = \beta_0 + \beta_1 p + \beta_2 p^2 + \dots + \beta_k p^k + \epsilon.$$

This form provides considerable flexibility in capturing non-linear patterns, but the main challenge lies in selecting an appropriate polynomial degree to balance goodness-of-fit and model complexity while avoiding overfitting. In this study, we imposed structural constraints on the model coefficients. Specifically, the polynomial coefficients were restricted to be non-negative, with the intercept of zero and their sum bounded by one. These constraints ensure that the fitted Lorenz curve satisfies the monotonicity property (i.e., non-negativity of the first derivative  $(dL/dp)$  as well as the convexity requirement over the interval  $([0,1])$ ). Consequently, only models adhering to these fundamental properties of the Lorenz curve were retained, guaranteeing both validity and interpretability of the fitted curve.

**3.1.2. Beta Regression.** Beta regression is a useful technique for analyzing continuous data like ratios, rates, or percentages constrained to the interval  $(0,1)$ . Paolino ([30]) was the first to use the beta distribution to represent continuous response variables, including ratios. He characterized the beta distribution parameters as the distribution of the response variable and provided the maximum likelihood estimate of the model parameters. The beta distribution is versatile, as its probability density function can take many forms based on its

parameter values. This adaptability has led to its applicability across various fields. In 2004, Ferrari and Cribari-Neto [16] defined beta regression for ratio data as the reparametric beta distribution. Beta regression offers multiple benefits, including addressing the bounded characteristics of the response variable, modeling non-linear associations between responses and predictors, and yielding interpretable parameters expressed as odds ratios when using a logit link. The following is a beta regression model of the relationship between  $p$  and  $L(p)$ .

$$L(p; \alpha, \beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)},$$

where  $B(\alpha, \beta)$  denotes the beta function, functioning as a normalisation constant. Beta regression is employed in this paper not because we assume income shares follow a beta distribution, but because the beta model is naturally suited for modeling response variables bounded between 0 and 1. The Lorenz curve, representing cumulative proportions, inherently satisfies this boundedness condition.

The beta density function's flexibility allows it to capture various shapes of Lorenz curves through different parameter configurations.

**3.1.3. Other Parametric Forms.** We also evaluated several established parametric Lorenz curves from the literature, as summarized in Table 1. These include the models by Kakwani and Podder [22], Aggarwal [1], Chotikapanich [10], and the Pareto-based model. These were estimated using non-linear least squares or maximum likelihood estimation, with constraints applied to ensure theoretical validity.

**3.2. Non-Parametric Model: Isotonic Regression.** Isotonic regression is a non-parametric method that estimates a monotone (non-decreasing) function from data without assuming a parametric form. It minimizes the weighted least-squares criterion

$$\min_{\hat{L}_1, \dots, \hat{L}_n} \sum_{i=1}^n w_i (L_i - \hat{L}_i)^2,$$

subject to the monotonicity constraints

$$\hat{L}_i \leq \hat{L}_j \quad \text{whenever } p_i \leq p_j.$$

Here,  $(p_i, L_i)$  are observed data points, and  $w_i \geq 0$  are optional weights (often set to 1).

When  $p_i$  are totally ordered (e.g.,  $p_1 \leq p_2 \leq \dots \leq p_n$ ), the constraints reduce to adjacent pairs:

$$E = \{(i, i+1) : 1 \leq i < n\}.$$

This formulation can be solved efficiently using the *Pool Adjacent Violators Algorithm (PAVA)*, which has complexity  $O(n)$  for sorted data.

**3.3. Pool Adjacent Violators Algorithm (PAVA).** PAVA works iteratively: if two adjacent fitted values violate monotonicity ( $\hat{L}_{i-1} > \hat{L}_i$ ), they are pooled and replaced with their average. This process repeats until all constraints are satisfied. The result is a monotone fit that minimizes the sum of squared errors. For a formal treatment and proofs of optimality we follow standard references. [4, 35]

---

**Algorithm 1:** PAVA Algorithm for Unweighted Isotonic Regression

---

**Input:** Ordered points  $(p_1, L_1), \dots, (p_n, L_n)$  with  $p_1 < \dots < p_n$ .

**Output:** Monotone fitted values  $\hat{L}_1, \dots, \hat{L}_n$ .

**1. Initialization:** Set  $\hat{L}_i = L_i$  for all  $i$ .

**2. Iterative pooling:**

**for**  $i = 2$  **to**  $n$  **do**

**if**  $\hat{L}_{i-1} > \hat{L}_i$  **then**

        Pool observations  $i - 1$  and  $i$ :

        Compute their average  $m = \frac{\hat{L}_{i-1} + \hat{L}_i}{2}$ .

        Set  $\hat{L}_{i-1} = \hat{L}_i = m$ .

        Move backward while monotonicity is violated:

**while**  $i > 2$  **and**  $\hat{L}_{i-2} > \hat{L}_{i-1}$  **do**

                Average the three pooled values and update them.

$i \leftarrow i - 1$ ;

**3. Return** the monotone fitted values  $\hat{L}_1, \dots, \hat{L}_n$ .

---

The fitted function  $\hat{L}(p)$  is usually defined as a continuous, piecewise linear interpolation between the fitted points  $(p_i, \hat{L}_i)$ :

$$\hat{L}(p) = \begin{cases} \hat{L}_1 & p \leq p_1, \\ \hat{L}_i + \frac{p-p_i}{p_{i+1}-p_i}(\hat{L}_{i+1} - \hat{L}_i) & p_i \leq p \leq p_{i+1}, \\ \hat{L}_n & p \geq p_n. \end{cases}$$

The key advantage of isotonic regression is its flexibility; it adapts to the data's shape without strong prior assumptions about the functional form. A potential drawback is that the raw PAVA fit is piecewise constant. To ensure convexity, a post-processing smoothing step was applied.

Table 1 summarizes some of the most important parametric Lorenz curves and their corresponding Gini coefficients, which are central to the comparative analysis in this study.

Isotonic regression inherently guarantees monotonicity through the Pool Adjacent Violators Algorithm (PAVA). The resulting fit is piecewise linear and non-decreasing. While convexity is not guaranteed by PAVA alone, a postprocessing smoothing step can be applied to approximate it. Since the empirical data include the boundary points (0,0) and (1,1), the isotonic fit naturally

TABLE 1. Some of the most important parametric Lorenz curves and their Gini coefficients

Name	Model	Gini
Isotonic Regression	$\begin{cases} L_1 & \text{if } p \leq p_1 \\ L_i + \frac{p-p_i}{p_{i+1}-p_i}(L_{i+1}-L_i) & \text{if } p_i \leq p \leq p_{i+1} \\ L_n & \text{if } p \geq p_n \end{cases}$	$1 - 2 \sum_{i=1}^n (p_i - p_{i-1}) \left( \frac{L(p_i) + L(p_{i-1})}{2} \right)$
Kakwani and Podder [22]	$pe^{-\alpha(1-p)}, \alpha > 0$	$1 - \frac{2(\alpha-1)}{\alpha^2} - \frac{2e^{-\alpha}}{\alpha^2}$
Aggarwal [1]	$\frac{(1-\alpha)^2 p}{(1+\alpha)^2 - 4\alpha p}, 0 < \alpha < 1$	$\frac{(1+\alpha)^2}{2\alpha} \left[ \frac{(1-\alpha)^2}{4\alpha} \ln \left( \frac{1-\alpha}{1+\alpha} \right)^2 + 1 \right] - 1$
Chotikapanich [10]	$\frac{e^{\alpha p} - 1}{e^\alpha - 1}, \alpha > 0$	$\frac{(\alpha-2)e^\alpha + (\alpha+2)}{\alpha(e^\alpha - 1)}$
Pareto	$1 - (1-p)^{1-\frac{1}{\alpha}}, \alpha > 1$	$\frac{1}{2\alpha} - 1$
Polynomial Model	$\beta_1 p + \dots + \beta_k p^k$	$1 - 2 \left( \frac{1}{2} \beta_1 + \dots + \frac{1}{k+1} \beta_k \right)$

passes through these points. Consequently, isotonic regression satisfies the essential Lorenz curve properties and the convex combination property when combining two estimated curves.

**3.4. Model Evaluation Framework.** To impartially compare the performance of the different Lorenz curve estimators, a robust and integrated evaluation metric is essential. This study employs a Generalized Mean Squared Error (GMSE) metric, specifically designed to assess the accuracy of the estimated Lorenz curve across its entire domain, from the poorest to the richest segments of the population.

**3.4.1. Generalized Mean Squared Error (GMSE).** The standard Mean Squared Error (MSE) measures the average squared difference between estimated and true values at the observed data points. However, for a comprehensive evaluation of a curve estimator, it is crucial to consider its performance at many points along the theoretical domain  $[0, 1]$ . To this end, we define a dense set of evaluation points  $T = \{0, 0.01, 0.02, \dots, 1.00\}$ , which divides the interval  $[0, 1]$  into 100 equal subintervals, yielding  $|T| = 101$  evaluation points.

For each Monte Carlo replication  $i$  ( $i = 1, 2, \dots, N$ ), the Mean Squared Error is calculated over this entire set  $T$ :

$$(13) \quad MSE_i = \frac{1}{|T|} \sum_{t \in T} \left[ L(t) - \hat{L}_i(t) \right]^2,$$

where:

- $L(t)$  is the true value of the Lorenz curve at point  $t \in T$ , derived from the known theoretical distribution used to generate the data.
- $\hat{L}_i(t)$  is the estimated value of the Lorenz curve at point  $t$  obtained from the  $i$ -th simulation replication.
- $|T|$  is the number of evaluation points.

The Generalized Mean Squared Error (GMSE) is then defined as the average of these  $MSE_i$  values across all  $N$  replications:

$$(14) \quad GMSE = \frac{1}{N} \sum_{i=1}^N MSE_i.$$

This two-stage averaging process provides a single, integrated measure of estimator performance. The inner average ( $MSE_i$ ) ensures the evaluation is comprehensive across the curve's domain for a single dataset, while the outer average ( $GMSE$ ) ensures the result is statistically stable across different random samples from the same population. A lower GMSE value indicates an estimator that is, on average, more accurate across the entire problem spectrum.

**3.5. Comparative Framework Philosophy.** The primary contribution of this study lies in its comparative framework rather than in the novelty of any single method. While isotonic regression is well-established, and parametric models like Kakwani-Podder and Chotikapanich are standard in economics literature, their systematic comparison under a unified simulation setting represents a significant gap in current research.

Our study addresses this gap by evaluating all estimators under identical conditions—using the same data-generating processes (log-normal and Pareto distributions), sample sizes, and performance metrics. This allows for a direct and fair assessment of their relative strengths and weaknesses. This approach shifts the focus from methodological novelty to empirical evidence-based model selection, reflecting the philosophical understanding that different models capture different aspects of income inequality.

In the next two sections, we present two examples: one using simulated data from lognormal and Pareto distributions, and the other analyzing real income distribution data from Iran. These examples demonstrate the effectiveness of different regression models in fitting LC.

#### 4. Simulation Study

This section intends to compare parametric and non-parametric isotonic models in fitting the Lorenz curve with simulated data. In this study, data from two Pareto and lognormal distributions were simulated, and empirical Lorenz curves were calculated. For each sample, we fitted five classical models: Kakwani, Aggarwal, Chotikapanich, Pareto, and a constrained polynomial model with non-negative coefficient constraints, along with the nonparametric isotonic model. The empirical Gini coefficient and the Gini coefficient for each model were calculated from the predicted Lorenz curve. The results showed that in both distributions, the isotonic model was able to follow the empirical Lorenz curve with the highest accuracy, and the Gini coefficient predicted by it was closer to the empirical value, while other parametric models showed slight

deviations in some areas of the curve.

The first model we examine is the Pareto distribution, represented as  $F(x) = 1 - x^{-\alpha}$  for  $\alpha > 1$ , signifying that the mean is finite. The LC is expressed as  $L(p) = 1 - (1 - p)^{\frac{1}{\alpha}}$ , with the mean represented by  $\mu = \frac{1}{\alpha - 1}$  and the Gini coefficient determined by  $G = \frac{1}{2\alpha} - 1$ . Figure 1 (left side) illustrates the LC for the Pareto distribution with different parameters ( $\alpha = 2, 3, 5$ ). As you can see in this figure, as the parameter increases, the level of inequality, or in other words, the Gini coefficient, decreases.

In this analysis, we assume the parameter  $\alpha$  is equal to 2. Consequently, we find that  $G = 0.3333$  and  $\mu = 1$ . In addition, the estimated parameter values and estimated Gini coefficients for a sample of 100 from this distribution are recorded in Table 2. In the constrained polynomial model, all coefficients were positive, and their sum was less than or equal to one, which ensured that the predicted Lorenz curve remained increasing and convex. The empirical Lorenz curve and estimated models for this distribution are shown in Figure 2 (left). According to the graph and the information in the table, it is quite clear that there is a very clear difference between isotonic regression and other data fitting models.

We are examining the log-normal distribution as a secondary model here. Its cumulative distribution function is  $F(x) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$  where  $\Phi$  is the cumulative distribution function of the standard normal. It is defined, in terms of  $\mu$  and  $\sigma$ . They are the mean and standard deviation of the logarithm of the data, respectively. In this study, we assume the parameter values are  $\mu = 0$  and  $\sigma = 0.5$ . The following formula calculates the LC for the log-normal distribution.

$$L(p) = \Phi(\Phi^{-1}(p) - \sigma)$$

The Gini coefficient is expressed as  $G = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$ . For the log-normal distribution with the parameters we considered, its value is equal to  $G = 0.2763$ . The obtained empirical Gini coefficient is equal to  $G = 0.254$ . In Figure 1 (right side), the LC for the parameters 0.5, 1, and 1.5 is plotted. In the figure, it is evident that for the log-normal distribution, with the increase of this parameter, inequality and the Gini coefficient increase. For a sample size of one hundred, the parameter values and the Gini coefficients that correspond to the various models have been computed. This procedure is similar to what was done for the Pareto distribution. As can be seen from the information shown in Table 2, the use of isotonic regression results in the Gini coefficient that is most similar to the empirical Gini coefficient of the data.

Figure 2 shows the different models fitted to the Pareto data (left) and log-normal data (right). The results indicate that isotonic regression provides the best fit to the empirical data for both distributions, and the Gini coefficient

predicted by it is almost equal to the experimental Gini coefficient (0.25405 vs. 0.254 for lognormal and 0.3737 vs. 0.3737 for Pareto). Other parametric models, including Kakwani, Aggarwal, Chotikapanich, and Pareto, have predicted the Gini coefficient to be lower or higher than the empirical value at certain points. For example, in the lognormal distribution, the Gini coefficient of the Pareto model is less than the empirical value (0.24189 versus 0.254), and in the Pareto distribution, the Polynomial model predicts a Gini coefficient of 0.33883, which is less than the empirical value of 0.3737. We also provide fitted parameters for parametric models, suggesting that these parameters can analyze the intensity of inequality. The limited polynomial model has positive coefficients, and their sum is less than or equal to one, so the predicted Lorenz curve remains upward and valid. Overall, these results indicate that isotonic regression is the best model for fitting the empirical Lorenz curve, and its predicted Gini coefficient has the closest match to the actual data, while other parametric models may provide slightly higher or lower values than the empirical Gini coefficient.

To further examine the models' performance, curve fitting was also performed for 40 and 100 sample sizes in 100 independent iterations. We then fit the different models introduced in Table 1 for these two distributions. The GMSE value for each model is calculated and presented in Table 3. Box plots of the error values for the different methods are also plotted in Figures 3 and 4.

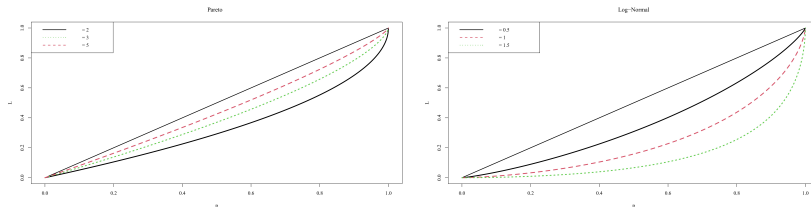


FIGURE 1. The Lorenz Curve for Pareto ( $\alpha = 2, 3, 5$ )(left) and lognormal ( $\sigma = 0.5, 1, 1.5$ )(right) distributions with different parameters.

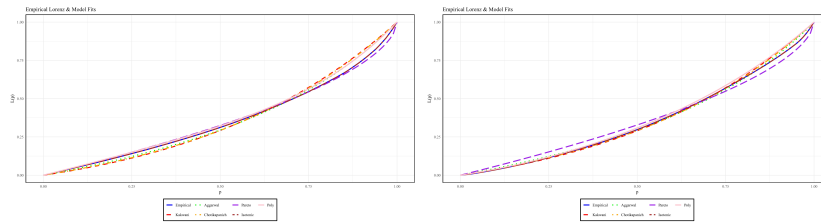


FIGURE 2. Experimental Lorenz curve and Lorenz curves fitted to simulated data from the Pareto distribution  $\alpha = 2$  (left) and log-normal distribution  $\mu = 0, \sigma = 0.5$  (right).

TABLE 2. Gini Coefficient and Estimated Parameters for the Lorenz Curve of Lognormal( $\mu = 0, \sigma = 0.5$ ) and Pareto ( $\alpha = 2$ ) Distributions (Sample Size: 100) with Different Model Fits

Model	Log-Normal( $\mu = 0, \sigma = 0.5$ )			Pareto( $\alpha = 2$ )	
Model	Gini	Estimated Parameters	Gini	Estimated Parameters	
Empirical Gini	0.254	-	0.3737	-	
Model Gini	0.2763264	-	0.3333	-	
Isotonic Regression	0.25405	-	0.3737	-	
Kakwani and Podder	0.25098	$\alpha = 0.937$	0.36813	$\alpha = 1.566$	
Aggarwal	0.25055	$\alpha = 0.189$	0.37807	$\alpha = 0.288$	
Chotikapanich	0.25092	$\alpha = 1.566$	0.36903	$\alpha = 2.422$	
Pareto	0.24189	$\alpha = 1.638$	0.37668	$\alpha = 2.209$	
Polynomial Model	0.253	$\beta_1 = 0.436, \beta_2 = 0.391, \beta_5 = .151$	0.3753	$\beta_1 = 0.491, \beta_2 = 0.086, \beta_5 = 0.229$	

Additionally, Table 3 summarizes the GMSE for each method, categorized by the different sample sizes. This table shows the values of the generalized mean squared error (GMSE) for different models fitted to simulated data from two distributions: lognormal and Pareto. Comparing these values for sample sizes of 40 and 100 shows that isotonic regression performs best in both distributions, with error values ranging from  $(10^{-32})$  to  $(10^{-31})$ , which is practically close to zero and indicates a very accurate fit to the empirical Lorenz curve. Other models have larger errors compared to the isotonic model. Specifically, for the lognormal distribution, the Aggarwal model with a GMSE of  $(1.1 \times 10^{-4})$  in the 40-sample and  $(8.01 \times 10^{-5})$  in the 100-sample after isotonic regression performs best, and the Kakwani model also provides results close to that. The Pareto and Polynomial models, on the other hand, displayed higher errors. For the Pareto distribution, it is also observed that the original Pareto model performs better than other parametric models, with errors of  $(1.96 \times 10^{-4})$  and  $(9.85 \times 10^{-5})$  for sample sizes of 40 and 100, respectively. Conversely, the Polynomial model exhibits the highest error and outperforms all other methods. Overall, this table confirms that isotonic regression is the most accurate method for fitting a Lorenz curve in both distributions and for different sample sizes, with other models following only with varying degrees of accuracy. An

important observation is that for data generated from a Pareto distribution, the Pareto parametric model consistently yields the best performance among parametric alternatives (see Table 3 and Figure 3). This result is theoretically expected: when the data-generating mechanism aligns with the assumed functional form of a parametric model, the model becomes both efficient and coherent. This finding reinforces the broader message of this paper that the suitability of a model is context-dependent and closely tied to the underlying structure of the data.

TABLE 3. Generalized Mean Squared Error (GMSE) values for different models fitted to simulated data from log-normal and Pareto distributions in 100 repetitions

Model	Log-Normal Distribution		Pareto Distribution	
	$n = 40$	$n = 100$	$n = 40$	$n = 100$
Isotonic Regression	$4.88 \times 10^{-32}$	$4.96 \times 10^{-31}$	$4.84 \times 10^{-32}$	$4.36 \times 10^{-31}$
Kakwani and Podder	$2.21 \times 10^{-4}$	$1.79 \times 10^{-4}$	$2.34 \times 10^{-3}$	$2.4 \times 10^{-3}$
Aggarwal	$1.1 \times 10^{-4}$	$8.01 \times 10^{-5}$	$1.2 \times 10^{-3}$	$1.32 \times 10^{-3}$
Chotikapanich	$1.82 \times 10^{-4}$	$1.42 \times 10^{-4}$	$2.12 \times 10^{-3}$	$2.17 \times 10^{-3}$
Pareto	$8.32 \times 10^{-4}$	$8.51 \times 10^{-4}$	$1.96 \times 10^{-4}$	$9.85 \times 10^{-5}$
Polynomial Model	$2.47 \times 10^{-4}$	$2.20 \times 10^{-4}$	$2.96 \times 10^{-3}$	$3.26 \times 10^{-3}$

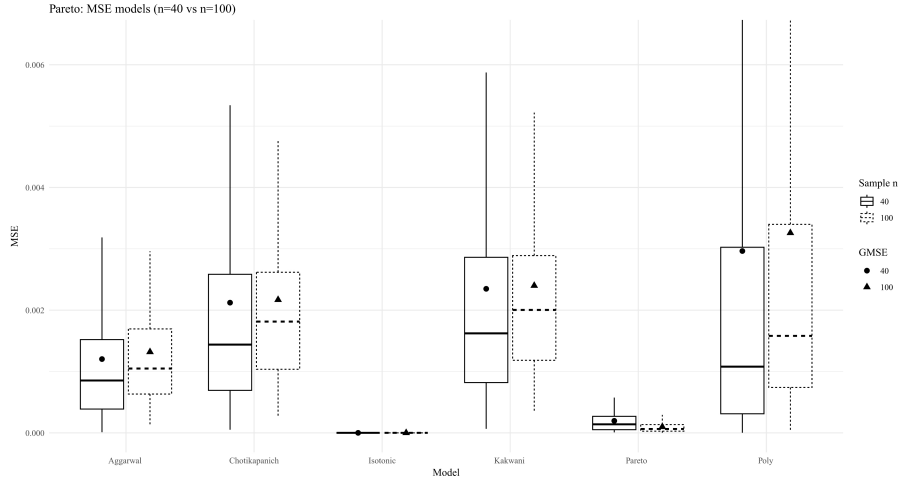


FIGURE 3. Boxplots of Mean Squared Errors for Various Lorenz Curve Models for Pareto Distribution.

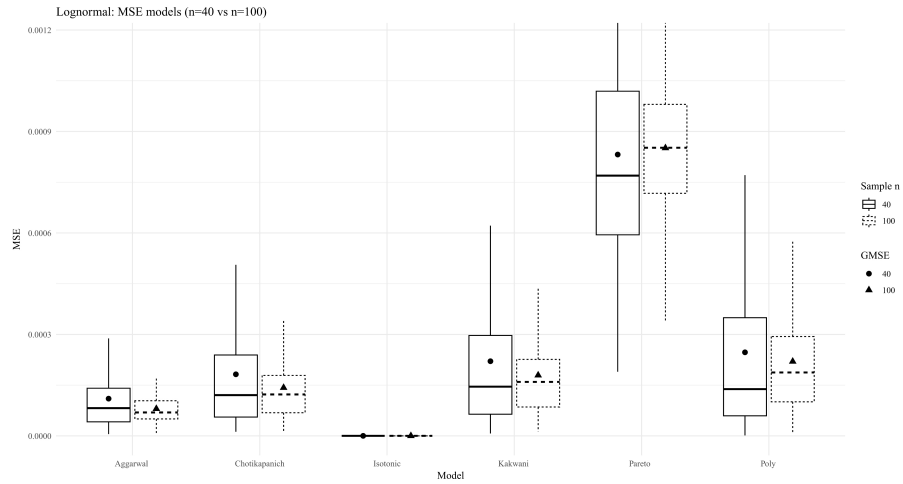


FIGURE 4. Boxplots of Mean Squared Errors for Various Lorenz Curve Models for Log-Normal Distribution.

## 5. Real Data Analysis: Income Distribution in Iran

In this section, we analyze income distribution in Iran using data from 18,809 households collected through the Household Expenditure and Income Survey conducted annually by the Statistical Center of Iran. The dataset was obtained via a stratified random sampling framework, covering both urban and rural households across all 31 provinces. Household incomes were sorted in ascending order, and cumulative income shares were calculated to construct the Lorenz curve, where the horizontal axis represents the cumulative population share and the vertical axis represents the cumulative income share.

To evaluate the accuracy of Lorenz curve estimation, we fitted eight different models to the data, including parametric regression models, isotonic regression, beta regression, and other advanced approaches. The shapes of these fitted models are shown alongside the empirical Lorenz curve in Figure 5, and the estimated parameters and Gini coefficients for each model are summarized in Table 4.

TABLE 4. Gini Coefficient and Estimated Parameters for the Lorenz Curve of Income Data with Different Model Fits

Model	Gini	Estimated Parameters
Empirical Gini	0.3556	-
Isotonic Regression	0.3556	-
Kakwani and Podder	0.349	$\alpha = 0.451$
Aggarwal	0.346	$\alpha = 0.264$
Chotikapanich	0.348	$\alpha = 0.263$
Pareto	0.332	$\alpha = 0.996$
Polynomial Model	0.3544	$\beta_1 = 0.2367, \beta_2 = 0.5234, \beta_3 = 0, \beta_4 = 0, \beta_5 = 0.178$
Beta	0.3553	$\alpha = 0.461, \beta = 0.694$
Rasch	0.3515	$\eta = 0.71569, \delta = 0.695$

The empirical Gini coefficient of 0.3556 serves as a benchmark for comparison. The isotonic regression reproduces the empirical Gini exactly, highlighting its ability to capture the observed income distribution without imposing parametric assumptions. Among the parametric models, the Beta and Polynomial models provide estimates very close to the empirical Gini (0.3553 and 0.3544, respectively), indicating strong flexibility and accurate fit. Models such as Kakwani and Podder, Aggarwal, and Chotikapanich slightly underestimate inequality, while the Pareto model shows a larger deviation (0.332), suggesting it underestimates income inequality relative to the observed data. The Rasch model produces an intermediate estimate of 0.3515.

The fitting accuracy of these models is further quantified in Table 5, which presents performance metrics including mean absolute error (MAE), root mean squared error (RMSE),  $R^2$ , and AIC. Notably, isotonic regression achieves nearly perfect fit metrics ( $R^2 = 1$ ,  $MAE \approx 0$ ), demonstrating its superiority in capturing the detailed empirical distribution. Rasch and other flexible models also perform very well, whereas simpler parametric models show slightly higher errors.

TABLE 5. Performance Metrics for Lorenz Curve Fitting Models (Iran Household Income)

Model	MAE	$R^2$	RMSE	AIC
Polynomial Regression	0.01988	0.99099	0.02526	-199.41
Beta Distribution	0.00464	0.99953	0.00579	-144131.02
Rasch Model	0.00052	0.99999	0.000647	-228756.41
Kakwani and Podder	0.014737	0.994533	0.019672	-96901.74
Aggarwal	0.011091	0.99759	0.0130477	-112756.90
Chotikapanich	0.013347	0.99554	0.017757	-100851.34
Pareto	0.032421	0.98175	0.035941	-73630.5844
Isotonic Regression	$3.22 \times 10^{-8}$	1.00000	0.0000016	0.0

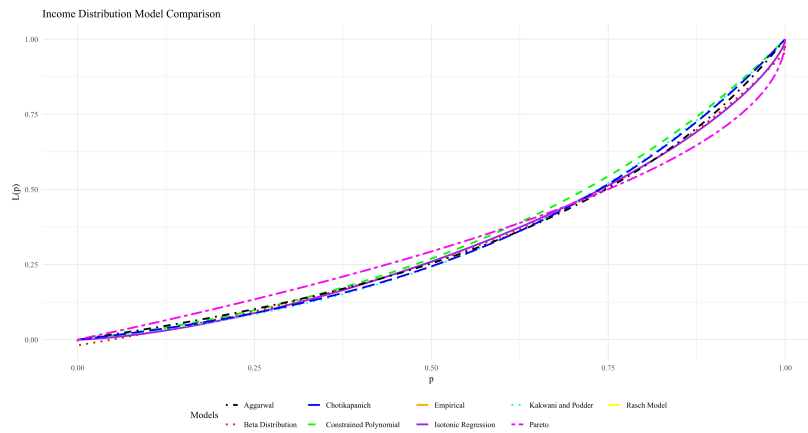


FIGURE 5. Fitted Lorenz Curves from Different Models for Iranian Income Data

The graphical comparison in Figure 5 illustrates how closely flexible parametric and non-parametric models approximate the empirical Lorenz curve, whereas simpler models such as Pareto deviate more noticeably. The nearly perfect fit of isotonic regression reflects the descriptive reality of the observed data. Given the large sample size, the detailed interpolation is appropriate for describing income inequality. To ensure that the polynomial model satisfies the defining boundary condition of Lorenz curves, we imposed the restriction that the sum of coefficients equals one, guaranteeing that  $L(1) = 1$ . The revised polynomial curve passes through  $(1,1)$ , as reflected in the Figure 5.

In the real-data analysis of Iranian household income, we evaluated the performance of multiple Lorenz curve estimation methods. The isotonic regression model provided a nearly perfect fit to the empirical data, closely matching the observed Lorenz curve and accurately reproducing the Gini coefficient. Parametric models such as Rasch, Beta, and Aggarwal also offered good fits, though their estimated Gini coefficients showed slight deviations compared to the empirical value. These results indicate that while non-parametric methods excel in descriptive accuracy, parametric models remain valuable when interpretability and theoretical grounding are important.

To further assess the robustness of the models, we performed sensitivity analyses considering the potential influence of extreme values. No outliers were removed, as the Lorenz curve and Gini coefficient are designed to account for the full income distribution, including high-income households that are essential for accurate inequality measurement. The models' performance was also examined under simulated conditions reflecting heavy-tailed (Pareto) and log-normal income distributions, yielding consistent results with the real data analysis. This demonstrates that the relative performance of isotonic

and parametric models is stable across both empirical and simulated scenarios, highlighting the reliability of our comparative framework.

## 6. Software and Implementation Details

In this study, all calculations and analyses were performed using the R software, which is considered one of the most powerful environments for statistical simulation and model fitting. Various packages were utilized to implement the methods and calculations, including the `ineq` package for calculating the Gini coefficient and empirical Lorenz curve, the `VGAM` package for generating simulated data from the Pareto distribution, the `dplyr` and `tidyr` packages for data preparation and processing, and the `ggplot2` package for plotting graphs and presenting results visually. The simulation study was conducted with specialized packages, including the `'isotone'` package for isotonic regression and the `'betareg'` package for beta regression. Furthermore, the constrained polynomial model was fitted using the `constrOptim` function to impose constraints such as non-negativity of coefficients and the sum of coefficients being less than or equal to one. These constraints ensure that the estimated Lorenz curve retains the key characteristics of a valid curve, namely being upward sloping and lying within the range of  $[0,1]$ . The results meet the necessary conditions for accuracy and validity, both theoretically and computationally.

## 7. Conclusion

This study offers an extensive comparison of parametric and non-parametric methods for estimating the Lorenz curve, with an emphasis on income distribution in Iran. Our analysis demonstrates that isotonic regression attains remarkably precise fits for both the Lorenz curve and the Gini coefficient, closely mirroring empirical data. Nonetheless, the selection of an estimation method ought to be contingent upon research objectives, data attributes, and pragmatic considerations. Isotonic regression is great for descriptive accuracy, but parametric models have parameters that are easy to understand, smooth functional forms, and theoretical consistency. Such approaches can be useful for policy analysis or smaller datasets. Large samples favor adaptable non-parametric methods, while parametric techniques may offer regularization advantages for constrained datasets or heavy-tailed distributions. The ideal selection is contingent upon the context, requiring a balance of accuracy, interpretability, and theoretical foundation. The results show that choosing the right model is important for getting useful information about income distribution. Isotonic regression is advised for comprehensive empirical description; however, parametric models are advantageous when interpretability, computational efficiency, or economic theory are emphasized. Subsequent research may broaden these comparisons to longitudinal data, supplementary parametric models, and datasets from various countries to improve generalizability.

While isotonic regression provides a perfectly adaptive descriptive fit for the observed data, it may not be optimal for prediction or inference on new samples, where parametric models can offer advantages due to their inherent smoothing. This distinction highlights that practical model choice depends on the specific analytic objective.

This study provides an evidence-based framework for choosing the right Lorenz curve estimators. This will help us measure income inequality more accurately and make better policy decisions.

### Conflict of interest

The authors declare that they have no conflict of interest.

### Data Availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study

### Funding

This study has not received any external funding.

### References

- [1] Aggarwal V.(1984). On optimum aggregation of income distribution data. *Sankhyā: The Indian Journal of Statistics, Series B*, 46: 343–355. <https://www.jstor.org/stable/25052360>
- [2] Arkes, J. (2023). *Regression analysis: a practical introduction*. Routledge. <https://doi.org/10.4324/9781003285007>
- [3] Arnold, B. C., & Villasenor, J.A. (1998). Lorenz Ordering of Order Statistics and Record Values. In: Balakrishnan, N, & Rao C.R. (Eds), *Handbook of Statistics*, Vol. 16. Elsevier, Amsterdam, 75-87. [https://doi.org/10.1016/S0169-7161\(98\)16006-6](https://doi.org/10.1016/S0169-7161(98)16006-6)
- [4] Barlow RE, Bartholomew DJ, Bremner J, & Brunk H.(1972). *Statistical inference under order restrictions: The theory and application of isotonic regression*. Wiley; New York. MR0326887. <https://doi.org/10.2307/2345150>
- [5] Behdani, Z., Mohtashami Borzadaran, G. R., & Sadeghpour Gildeh, B. (2019). Connection of Generalized Failure Rate and Generalized Reversed Failure Rate with Inequality Curves. *International Journal of Reliability, Quality and Safety Engineering*, 26(02), 195-223. <https://doi.org/10.1142/S0218539319500062>
- [6] Behdani, Z., Mohtashami Borzadaran, G. R., & Sadeghpour Gildeh, B. (2020). Concepts of Reliability Theory and Their Relationship with the Economic Inequality indicators. *Mathematical Researches*, 6(1), 25-38. <https://doi.org/10.52547/mmr.6.1.25>
- [7] Behdani, Z., Mohtashami Borzadaran, G. R., & Sadeghpour Gildeh, B. (2018). Relationship between the weighted distributions and some inequality measures. *Communications in Statistics-Theory and Methods*, 47(22), 5573-5589. <https://doi.org/10.1080/03610926.2017.1397171>
- [8] Behdani, Z., Mohtashami Borzadaran, G. R., & Sadeghpour Gildeh, B. (2020). Some properties of double truncated distributions and their application in view of income inequality. *Computational Statistics*, 35, 359-378. <https://doi.org/10.1007/s00180-019-00890-2>

- [9] Carter Jr, W. H., Wampler, G. L., & Stablein, D. M. (2024). Regression analysis of survival data in cancer chemotherapy. CRC Press.
- [10] Chotikapanich D. (1993). A comparison of alternative functional forms for the Lorenz curve. *Econ Lett.*; 41: 129–138. [https://doi.org/10.1016/0165-1765\(93\)90186-G](https://doi.org/10.1016/0165-1765(93)90186-G)
- [11] Dai, P., & Shen, S. (2025). Estimation of the Gini coefficient based on two quantiles. *PloS one*, 20(2), 1-13. <https://doi.org/10.1371/journal.pone.0318833>
- [12] Fajar, M., Setiawan, S., & Iriawan, N. (2023). The adjusted SNR and its application for selection Lorenz function of income inequality analysis. *Procedia Computer Science*, 227, 1-16. <https://doi.org/10.1016/j.procs.2023.10.497>
- [13] Fajar, M., Setiawan, S., & Iriawan, N. (2024). The New Measures of Lorenz Curve Asymmetry: Formulation and Hypothesis Testing. *Decision Making: Applications in Management and Engineering*, 7(1), 99-130. <https://doi.org/10.31181/dmame712024875>
- [14] Fellman, J. (2018). Regression analyses of income inequality indices. *Theoretical Economics Letters*.8:1793–1802. <https://doi.org/10.4236/tel.2018.810117>
- [15] Fellman, J. (2021). Empirical analyses of income: Finland (2009) and Australia (1967-1968). *Journal of Statistical and Econometric Methods*. 10(1), 35-53. <https://doi.org/10.47260/jsem/1013>
- [16] Ferrari, S.L.P. & Cribari-Neto, F. (2004), Beta Regression for Modelling Rates and Proportion, *Journal of Applied Statistics*, 7, 799-815. <https://doi.org/10.1080/0266476042000214501>
- [17] Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246-263. <https://doi.org/10.2307/2841583>
- [18] Gini, C. (1912). Variabilità e mutabilità. *Reale Accademia dei Lincei*.
- [19] Giudici, P., & Raffinetti, E. (2020). Lorenz model selection. *Journal of Classification*, 37(3), 754-768. <https://doi.org/10.1007/s00357-019-09358-w>
- [20] Huang, H. H., & He, Q. (2024). Nonlinear regression analysis. *arXiv preprint arXiv:2402.05342*. <https://doi.org/10.48550/arXiv.2402.05342>
- [21] Kakwani, N. (1980). Functional forms for estimating the Lorenz curve: A reply. *Econometrica*, 48(4), 1063-1064. <https://doi.org/10.2307/1912949>
- [22] Kakwani, N.C., and Podder, N. (1973). On the estimation of Lorenz curve from grouped observations. *International Economic Review*. 14. 278-292. <https://doi.org/10.2307/2525921>
- [23] Kakwani, N.C., & Podder, N. (1976). Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations. *Econometrica*. 44. 137-148. [https://doi.org/10.1007/978-0-387-72796-7\\_4](https://doi.org/10.1007/978-0-387-72796-7_4)
- [24] Karakaya, K., & Sağlam, Ş. (2025). Unit New Half Logistic Distribution: Theory, Estimation, and Applications with Novel Regression Analysis. *Statistics, Optimization & Information Computing*. 14, 526-540. <https://doi.org/10.19139/soic-2310-5070-2360>
- [25] Kleiber, C., & Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*, John Wiley.
- [26] De Leeuw, J., Hornik, K., & Mair, P. (2010). Isotone optimization in R: pool-adjacent-violators algorithm (PAVA) and active set methods. *Journal of statistical software*, 32, 1-24. <https://doi.org/10.18637/jss.v032.i05>
- [27] Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70), 209-219. <https://doi.org/10.1080/15225437.1905.10503443>
- [28] Pareto, V. (1895). La Legge della Domanda, *Giornale degli Economisti*, Gennaio. 59-68. <https://www.jstor.org/stable/23219874>
- [29] Parzen, E. (1980). Non-parametric statistical data modeling. *J. Amer. Statist. Assoc.*, 74, 105-131. <https://doi.org/10.1080/01621459.1979.10481621>

- [30] Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis*, 9(4), 325-346. <https://doi.org/10.1093/oxfordjournals.pan.a004873>
- [31] Paul, S., & Shankar, S. (2020). An alternative single parameter functional form for Lorenz curve. *Empirical Economics*, 59(3), 1393-1402. <https://doi.org/10.1007/s00181-019-01715-3>
- [32] Piketty, T. (2014). Capital in the twenty-first century. *Trans. Arthur Goldhammer/Belknap*. <https://doi.org/10.4159/9780674369542-intro>
- [33] Rasche, R.H., Gaffney, A.Y.C., and Koo, N. Obst. (1980). Functional forms for Estimating the Lorenz Curve. *Econometrica*. 48 (4), 1061-1062. <https://doi.org/10.2307/1912948>
- [34] Rasch, G. (1993). Probabilistic Models for Some Intelligence and Attainment Tests. ERIC.
- [35] Robertson T, Wright F. T., & Dykstra R. L.(1988). Order Restricted Statistical Inference. New York: John Wiley. ISBN 9780471917878. URL. <https://www.jstor.org/stable/2096708>
- [36] Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of economic growth*, 4, 385-412. <https://doi.org/10.1023/A:1009863208706>
- [37] Sarabia, J., Castillo, E., & Slottje, D. (1999). An ordered family of Lorenz curves. *Journal of Econometrics*, 91. 43-60. [https://doi.org/10.1016/S0304-4076\(98\)00048-7](https://doi.org/10.1016/S0304-4076(98)00048-7)
- [38] Sen, A. (2014). Development as freedom (1999). *The globalization and development reader: Perspectives on development and global change*, 525.
- [39] Shen, B. W. (2023). A Review of Lorenz's Models from 1960 to 2008. *International Journal of Bifurcation and Chaos*, 33(10), 2330024–2330220. <https://doi.org/10.1142/S0218127423300240>
- [40] Shen, X., & Dai, P. (2024). A regression method for estimating Gini index by decile. *Humanities and Social Sciences Communications*, 11(1), 1-8. <https://doi.org/10.1057/s41599-024-03701-2>
- [41] Smith, A. (1937). *The wealth of nations [1776]* (Vol. 11937). na.
- [42] Sitthiyot, T., & Holasut, K. (2021). A simple method for estimating the Lorenz curve. *Humanities and Social Sciences Communications*, 8(1), 1-9. <https://doi.org/10.1057/s41599-021-00948-x>
- [43] Sitthiyot, T., & Holasut, K. (2023). An investigation of the performance of parametric functional forms for the Lorenz curve. *Plos one*, 18(6), 1-21. <https://doi.org/10.1371/journal.pone.0287546>

ZAHRA BEHDANI

ORCID NUMBER: 0000-0002-0029-4833

DEPARTMENT OF MATHEMATICS AND STATISTICS

FACULTY OF ENERGY AND DATA SCIENCES

BEHBAHAN KHATAM ALANBIA UNIVERSITY OF TECHNOLOGY

BEHBAHAN, IRAN

*Email address:* [behdani@bkatu.ac.ir](mailto:behdani@bkatu.ac.ir)