Int. J. Nonlinear Anal. Appl. In Press, 1–20
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2026.36731.5390

Some features of a new measure of income inequality based on the ratios between lower and upper medians

Shahryar Mirzaei[a,*], Iman Makhdoom[b]

[a]*Department of Statistics, Payame Noor University, Tehran, Iran.*

[b]*Department of Statistics, Payame Noor University, Tehran, Iran.*

(Communicated by Mouquan Shen)

Abstract

This paper introduces a median-based measure of income inequality, defined as the ratio between the lower and upper medians of the income distribution. The index is defined pointwise, and its corresponding inequality curve is presented. We examine the essential properties of the index, including normalization, scale invariance, location sensitivity, and the Pigou–Dalton transfer principle. The behavior of the inequality curve is investigated within common parametric income models. A suitable estimator is proposed, and its strong consistency and asymptotic normality are established. Additionally, a stochastic order based on the proposed curve is defined and studied. Simulation studies show that the proposed estimator performs competitively compared to the Gini coefficient in terms of bias, mean squared error, and coverage probabilities. Finally, the measure is applied to real household income data, illustrating its practical usefulness.

Keywords: Income inequality curve, Lower and upper medians, Quantile functions, Median inequality index.
2020 MSC: 62P20, 91B82, 62G05.

---

1 Introduction

Understanding income inequality heavily relies on the concept of quantiles. Quantiles, essentially "breakpoints" within a dataset, divide the population into groups based on where they fall within the income distribution. The p-quantile of a dataset is a value (x_p) that marks a point where a certain proportion (p) of the data falls below or at that value. Many measures of income inequality use these quantiles to compare the income levels of different groups within a population. By examining the ratios between higher and lower quantiles, we can see how income is distributed. A larger ratio means a wider gap between the rich and the poor, indicating higher inequality.

Additionally, several inequality indices and criteria are based on comparing the average or total income of lower-income groups to the total income of the entire population. This provides insight into how much wealth is concentrated in the hands of those with lower incomes.

In 1905, Lorenz [10] proposed a way to understand income inequality by dividing society into groups based on income. He defined the Lorenz function, which measures the proportion of total income earned by those below a

*Shahryar Mirzaei
Email addresses: sh_mirzaee@pnu.ac.ir (Shahryar Mirzaei), makhdoom@pnu.ac.ir (Iman Makhdoom)

Received: January 2025 *Accepted:* September 2026

certain income level (x_p) compared to the total income of everyone. Later, in 1932, Bonferroni [3] built upon this concept by introducing his index. Bonferroni's index specifically focused on the ratio of the average income of those below the income threshold (x_p) to the average income of the entire population. Zenga [18], in his 2007 research, categorized society into two groups based on a specific income level (x_p). He then introduced the Zenga function, which measures inequality by comparing the average income of those earning less than x_p to the average income of those earning more than x_p. More recently, Davydov and Greselin [6] proposed a symmetric modification of the Zenga index.

However, when dealing with income data that often displays extreme values and skewed distributions, relying solely on the mean can be misleading. To address this challenge, researchers suggest using the median instead, as it provides a more robust measure that is less influenced by outliers. The median, being a more stable indicator of central tendency in skewed distributions, offers a deeper understanding of income disparities.

Gastwirth [7] suggested using the median instead of the mean to adjust the Gini mean difference (GMD), a measure of income inequality. This adjustment, while potentially useful, may result in an index that does not fit the standard definition of a normalized index. The issue arises because the index compares the mean income of the poor to the median income of the entire population.

Despite the availability of various inequality measures, a fully median-based index that compares lower and upper medians has not been systematically studied. This paper aims to fill this gap.Building upon Zenga's work and the benefits of using the median, a new income inequality index can be defined by comparing the median income of the lower segment of the population to the median income of the higher segment. This approach offers a more nuanced and accurate representation of income inequality, particularly in situations where traditional mean-based measures might not provide a complete picture. This research introduces a novel measure of income inequality that utilizes the ratio of lower and upper medians.

The paper is organized as follows: Section 2 reviews classical inequality measures and introduces notation. Section 3 defines the new index and establishes its properties. Section 4 studies its behaviour in common income models. Section 5 proposes an estimator and establishes its asymptotic properties. Sections 6 and 7 present simulation results and real-data applications, respectively. Section 8 concludes.

2 Foundations and Review of Inequality Measures

Before introducing the proposed median-based index, we establish the necessary notation and review the main classical inequality measures used in the literature. This review reveals that all classical indices rely on the mean of the distribution, which makes them sensitive to extreme values. This observation motivates the central idea of the present work: replacing the mean with the median in the structure of the inequality index, thereby enhancing robustness to heavy tails and extreme observations.

2.1 Notation and Preliminaries

Let X be a non-negative random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with cumulative distribution function (cdf) $F_X(x) = P(X \leq x)$, continuous and strictly increasing on its support $S_X \subseteq [0,\ \infty)$, and quantile function defined as the generalized inverse

$$Q_X(p) = F_X^{-1}(p) = \inf\{x :\ F_X(x) \geq p\}, \qquad 0 \leq p \leq 1.$$

The probability density function (pdf) $f_X(x) = F'_X(x)$ is assumed to exist. The distribution is assumed to have a finite positive mean

$$\mu = \mathbb{E}[X] = \int_0^\infty x dF_X(x) = \int_0^1 Q_X(t)dt,$$

with $0 < \mu < \infty$. For $p \in (0,\ 1)$, the lower mean, $\mu_p^- = \frac{1}{p}\int_0^p Q_X(t)dt$ is the mean income of those below the quantile $Q_X(p)$, and the upper mean, $\mu_p^+ = \frac{1}{1-p}\int_p^1 Q_X(t)dt$, is the mean income of those above $Q_X(p)$. It is obvious that

$$\mu = p \times \mu_p^- + (1-p) \times \mu_p^+.$$

The median is defined as $m = Q(0.5)$, i.e. the unique value satisfying $F(m) = 0.5$. For $p \in (0,\ 1)$, the lower median, $Q_X(\frac{p}{2})$ is the median income of those below the quantile $Q_X(p)$, and the upper median, $Q_X(\frac{1}{2} + \frac{p}{2})$ is the median income of those above $Q_X(p)$.

These basic quantities form the foundation for all subsequent definitions and analyses.

2.2 Overview of Related Inequality Measures

To situate the proposed median-based index within the existing literature, this subsection briefly reviews several classical inequality measures that have been widely employed in income distribution studies. These measures are all defined in terms of the quantile function and the mean of the income distribution. For instance, the Lorenz curve is based on the ratio of the cumulative income of the lower group to the total income; the Bonferroni index relies on the ratio of the partial mean to the overall mean; the Zenga and Davydov–Greselin indices compare the mean incomes of lower and upper groups; the Gini Mean Difference serves as the building block for the classical Gini coefficient; and another index is defined as a combination of the GMD and the median, obtained by replacing the mean with the median in the structure of the Gini coefficient. A brief review of these measures not only situates our proposed index within the existing literature but also highlights its distinguishing feature: the use of medians rather than means in the comparison of lower and upper income groups, which enhances robustness to heavy tails and extreme observations.

- **The Lorenz curve:** The Lorenz curve, introduced by Lorenz (1905) [10], is a fundamental graphical tool for representing income or wealth inequality within a population. It plots the cumulative proportion of total income earned by the bottom $\%100 \times p$ of the population against the cumulative proportion p of the population, for $p \in [0, 1]$. Formally, for a non-negative random variable X with finite positive mean, the Lorenz curve is defined as

$$L_X(p) = \frac{1}{\mu} \int_0^p Q_X(t)dt, \qquad \forall p \in (0,\ 1) \tag{2.1}$$

  As shown in Figure 1, the curve is convex, lies below the diagonal line of perfect equality, and coincides with the diagonal only under complete equality. The further the curve lies below the diagonal, the greater the inequality in the distribution. The Lorenz curve not only provides a visual summary of inequality but also serves as the foundation for many scalar measures, most notably the Gini coefficient (G), which is defined as twice the area between the Lorenz curve and the line of perfect equality.

$$G = 2 \int_0^1 [p - L_X(p)]dp. \tag{2.2}$$

- **Bonferroni Index:** Let X be a non-negative random variable representing income with finite mean $\mu = \mathbb{E}(X)$. The Bonferroni inequality index [3] is defined based on the ratio of the mean income of low-income groups to the mean income of the entire population. The Bonferroni curve is defined as:

$$B_X(p) = 1 - \frac{\mu_p^-}{\mu}, \qquad \forall p \in (0\ , 1), \tag{2.3}$$

  where μ_p^- is the mean income of individuals at or below the p-th quantile: The Bonferroni index (B) is equal to the area between the Bonferroni curve and the line of perfect equality:

$$B = \int_0^1 B_X(p)dp.$$

  The Bonferroni index lies between 0 and 1; a value of 0 indicates perfect equality, while a value of 1 indicates complete inequality. Compared to the Gini coefficient, the Bonferroni index is more sensitive to changes in income among the poorest individuals, making it particularly suitable for poverty and vulnerability analysis.

- **The Zenga index:** The Zenga index, introduced by Zenga (2007) [18], is a recent inequality measure defined through its associated Zenga curve. For each proportion $p \in (0, 1)$, the Zenga curve $Z(p)$ compares the mean

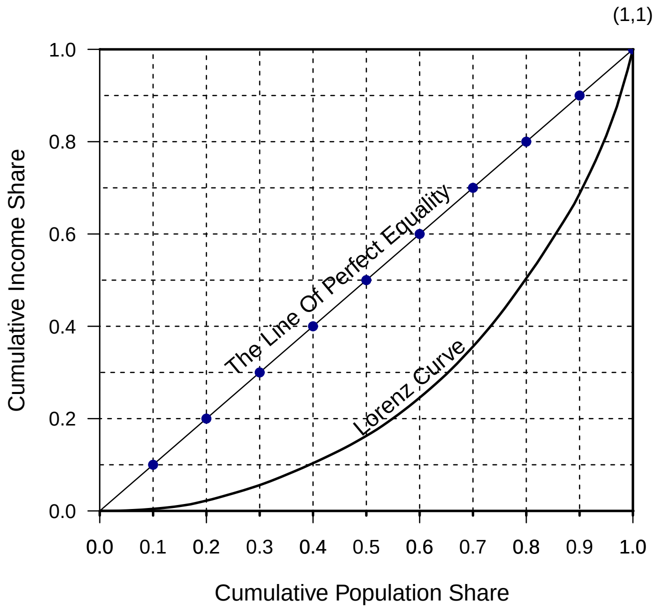

Figure 1: **An example of the Lorenz curve**

income of the poorest $\%100\times p$ of the population with the mean income of the richest $\%100\times(1-p)$. Specifically, it is defined as

$$Z_X(p) = 1 - \frac{\mu_p^-}{\mu_p^+}, \qquad \forall p \in (0\ ,1), \tag{2.4}$$

where $\mu_p^- = \frac{1}{p}\int\limits_0^p Q_X(t)dt$ is the mean income of those below the quantile $Q_X(p)$, and $\mu_p^+ = \frac{1}{1-p}\int\limits_p^1 Q_X(t)dt$ is the mean income of those above $Q_X(p)$. The global Zenga index is then obtained by integrating the curve over all $p \in (0,\ 1)$:

$$Z = 1 - \int\limits_0^1 \frac{\frac{1}{p}\int\limits_0^p Q_X(t)dt}{\frac{1}{1-p}\int\limits_p^1 Q_X(t)dt} dp. \tag{2.5}$$

- **The Davydov–Greselin index:** This index, introduced by Davydov and Greselin (2019) [6], offers a symmetric modification of the original Zenga index. While the Zenga curve compares the mean income of the lower group below the quantile $Q_X(p)$ with the mean income of the upper group above the same quantile $Q_X(p)$, the Davydov–Greselin index compares the mean income of the lower $\%100 \times p$ of the population (below $Q_X(p)$) with the mean income of the upper $\%100 \times p$ of the population (above $Q_X(1-p)$). This modification ensures that both groups involved in the comparison have the same size p, which simplifies interpretation and enhances symmetry. The index is defined as

$$D = 1 - \int\limits_0^1 \frac{\frac{1}{p}\int\limits_0^p Q_X(t)dt}{\frac{1}{p}\int\limits_{1-p}^1 Q_X(t)dt} dp. \tag{2.6}$$

  Like the Zenga index, D lies between 0 and 1, attains zero under perfect equality, and is scale invariant. However, by comparing equal-sized lower and upper groups, the Davydov–Greselin index provides a more balanced assessment of inequality, particularly in distributions with heavy tails, and offers a useful alternative to mean-based measures such as the Gini coefficient and the classical Zenga index.

- **GMD index:** For a random variable X with cdf F_X, the GMD index is given by

$$GMD(X) = \mathbb{E}|X_1 - X_2|, \tag{2.7}$$

where X_1 and X_2 are independent copies of the same distribution. In terms of quantiles, it is expressed as

$$GMD(X) = \int_0^1 \left(\int_0^1 |Q_X(u) - Q_X(v)| du \right) dv, \tag{2.8}$$

representing the average absolute difference over all pairs of quantiles. Unlike the variance, the GMD does not require finite second moments and is more robust to heavy-tailed distributions, making it particularly suitable for income and wealth data. The GMD serves as the building block for the classical Gini coefficient, given by

$$G = \frac{GMD}{2\mu}, \tag{2.9}$$

but retains the original scale, allowing its use as a direct measure of dispersion.

Gastwirth [7] introduced a median based version of the Gini coefficient,

$$G_m = \frac{GMD}{2m}, \tag{2.10}$$

where the mean in the denominator is replaced by the median(X) (m) in 2.9 . While this adjustment improves robustness to extreme values and increases sensitivity to upper tail inequality, it comes with drawbacks. First, G_m is not a purely median based index, since the numerator (the GMD) remains a mean based quantity; it is therefore a hybrid measure. Second, G_m may exceed 1 in highly skewed distributions, violating the standard normalization property that makes the Gini coefficient easily interpretable. In contrast, the index we propose in the next section is fully median based, relying solely on comparisons of lower and upper medians, and thus avoids these limitations.

3 The new inequality index

Here, the median income of those below the quantile $Q_X(p)$ needs to be compared with the median income of those above the quantile $Q_X(p)$.This point of view has led the author to a definition of the new index based on the ratio of the two quantiles:

$$I_X(p) = \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}, \qquad \forall p \in (0,1).$$

In addition, $I_X(p)$ is a ratio between the lower and the upper partial medians; the meaning of $I_X(p)$ is very simple. Furthermore, when all income values are equal, the value of the inequality index should be zero. Therefore, the point value of the new index should be defined as follows:

$$\mathcal{M}_X(p) = 1 - \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}, \qquad \forall p \in (0,1). \tag{3.1}$$

A model of this curve is shown in Figure 2. This point index takes the value of zero in the case of complete equality and the value of one in the case of complete inequality. By averaging $\mathcal{M}_X(p)$, the new inequality index $\mathcal{M}$ is obtained. The new index is linked to the new inequality curve, which is obtained by plotting the points $(p, \mathcal{M}(p))$. Moreover, in the continuous case, it is easy to say that:

$$\mathcal{M} = \int_0^1 \mathcal{M}_X(p) dp, \tag{3.2}$$

and we call it the median-inequality index.

If all members of society have the same income, the inequality curve coincides with the line of perfect equality. In this case, the curve connects the points $(0,1)$, $(0,0)$, and $(1,0)$ on the coordinate axis. Furthermore, it can also be said that the value of the new inequality index is equal to the area between the inequality curve and the line of perfect equality (as shown in Figure 2).

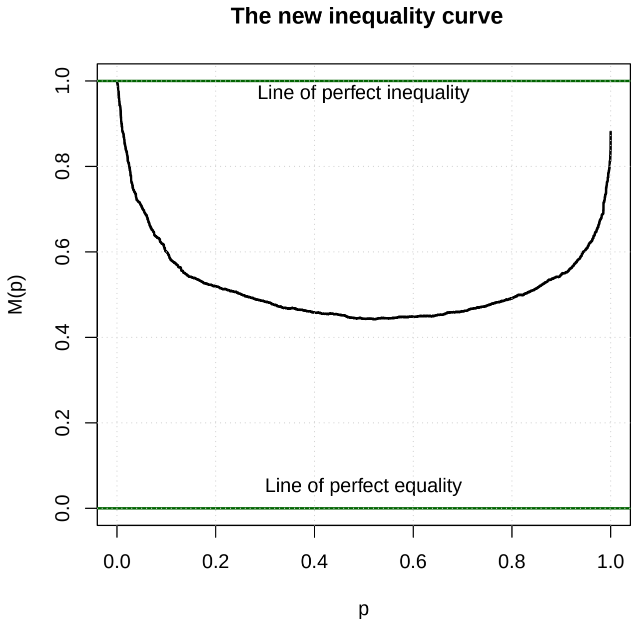

Figure 2: **An example of the new inequality curve**

3.1 Some main properties of the new index

When evaluating income inequality indices, it's crucial to assess their properties to determine their suitability for measuring income disparities. These properties, known as principles, act as guidelines for a good inequality index. The more principles an index fulfills, the more reliable it is considered. Here's a breakdown of key principles:

- **Normalization:** The index should fall between 0 and 1, allowing for easy comparison and interpretation.
- **Replication reliability:** If incomes are duplicated, the index should remain unchanged, reflecting that true inequality hasn't shifted.
- **Symmetry:** Swapping the income of two individuals shouldn't affect the index, indicating that it focuses purely on income differences, not individual identities.
- **Sensitivity to location reliability:** Adding or subtracting a fixed amount to everyone's income should, respectively, decrease or increase the index, demonstrating its responsiveness to income shifts.
- **Insensitivity to scale reliability:** If everyone's income is multiplied by the same factor, the index should remain constant, highlighting its focus on relative income differences.
- **Decomposability:** If society is divided into groups, the overall inequality should be a combination of the inequalities within each group, allowing for analysis at different levels.
- **Sensitivity to income transfer:** Transferring income from a richer person to a poorer person (without reversing their relative positions) should lower the index, reflecting a decrease in inequality.

By adhering to these principles, income inequality indices become more reliable and provide a clearer picture of income disparities in a society.

This section focuses on key principles that govern the effectiveness of income inequality indicators. We will explore specific properties such as normalization, scale reliability, sensitivity to location (or income shifts), and sensitivity to income transfers (or Pigou-Dalton property). These properties, which will be examined in the following propositions, are crucial for ensuring that inequality indices accurately reflect and measure income disparities within a society.

Proposition 3.1. (**Normalization**) Let X be a non-negative random variable with $0 < \mathbb{E}(X) < \infty$, and let $\mathcal{M}$ be its associated inequality index as defined in 3.2. Then,

$$0 \leq \mathcal{M} \leq 1.$$

Proof . From the monotonicity of the quantile function Q, we have

$$0 \leq Q_X(\frac{p}{2}) \leq Q_X(\frac{1}{2} + \frac{p}{2}), \quad \forall p \in (0,\ 1)$$

Since $Q_X(\frac{1+p}{2}) > 0$, it follows that

$$0 \leq \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})} \leq 1$$

Thus, the pointwise curve $\mathcal{M}(p) = 1 - \dfrac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}$ satisfies

$$0 \leq \mathcal{M}_X(p) \leq 1, \quad \forall p \in (0,\ 1)$$

Integrating with respect to p over $[0, 1]$ gives

$$0 \leq \int_0^1 \mathcal{M}_X(p)dp \leq 1.$$

By definition, $\mathcal{M} = \int_0^1 \mathcal{M}_X(p)dp$, hence $0 \leq \mathcal{M} \leq 1$. □

Proposition 3.2. (**Scale Invariance**) Let X be a non-negative random variable with $0 < \mathbb{E}(X) < \infty$, and let $\mathcal{M}_X(p)$ be its inequality curve as defined in 3.1. Then, for any $a > 0$,

$$\mathcal{M}_{aX}(p) = \mathcal{M}_X(p), \qquad \forall p \in (0,\ 1),$$

and consequently,

$$\mathcal{M}_{aX} = \mathcal{M}_X.$$

Proof . Let $Y = aX$ for any $a > 0$, then

$$Q_Y(p) = aQ_X(p), \qquad \forall p \in (0, 1), \tag{3.3}$$

where Q_X and Q_Y are the quantile function of X and Y respectively. Also, we have

$$\begin{aligned} \mathcal{M}_Y(p) &= 1 - \frac{Q_Y(\frac{p}{2})}{Q_Y(\frac{1}{2} + \frac{p}{2})} \\ &= 1 - \frac{aQ_X(\frac{p}{2})}{aQ_X(\frac{1}{2} + \frac{p}{2})} \\ &= \mathcal{M}_X(p). \end{aligned} \tag{3.4}$$

Integrating over p gives $\mathcal{M}_X = \mathcal{M}_Y$. As a result, insensitivity to scale reliability or the scale invariant of the $\mathcal{M}$ index is obtained. □

Proposition 3.3. (**Location Sensitivity**) Let X be a non-negative random variable with $0 < \mathbb{E}(X) < \infty$, and let $\mathcal{M}_X(p)$ be its inequality curve as defined in 3.1. Then, for any $b > 0$,

$$\mathcal{M}_{X+b}(p) < \mathcal{M}_X(p), \qquad \forall p \in (0,\ 1),$$

and consequently,

$$\mathcal{M}_{X+b} < \mathcal{M}_X.$$

Proof . Let $Y = X + b$ with $b > 0$. By the properties of quantile functions, we have

$$Q_Y(p) = Q_X(p) + b, \quad \forall p \in (0,\ 1).$$

Using the definition of the inequality curve,

$$\begin{aligned}\mathcal{M}_Y(p) &= 1 - \frac{Q_Y(\frac{p}{2})}{Q_Y(\frac{1}{2} + \frac{p}{2})} \\ &= 1 - \frac{Q_X(\frac{p}{2}) + b}{Q_X(\frac{1}{2} + \frac{p}{2}) + b}\end{aligned}$$

Since $Q_X(\frac{p}{2}) < Q_X(\frac{1}{2} + \frac{p}{2})$ and $b > 0$, it follows that

$$\frac{Q_X(\frac{p}{2}) + b}{Q_X(\frac{1}{2} + \frac{p}{2}) + b} > \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}.$$

Therefore,

$$\mathcal{M}_Y(p) < 1 - \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})} = \mathcal{M}_X(p).$$

Thus, $\mathcal{M}_Y(p) < \mathcal{M}_X(p)$ for all $p \in (0, 1)$. Integrating both sides with respect to p over $[0, 1]$ gives

$$\int_0^1 \mathcal{M}_Y(p)dp < \int_0^1 \mathcal{M}_X(p)dp$$

which is equivalent to $\mathcal{M}_Y < \mathcal{M}_X$. Hence the index is sensitive to location, decreasing when all incomes are increased by a fixed positive amount. □

Remark 3.4. A similar consideration holds for any $b < 0$, in this case $\mathcal{M}_Y > \mathcal{M}_X$.

Pigou-Dalton's property states that if some of the rich person's income is transferred to a poorer person, provided that the position of the poor and rich person is not changed, the inequality index will decrease.

Proposition 3.5. (**Pigou–Dalton Transfer Principle)** Let X be a non-negative random variable with $0 < \mathbb{E}(X) < \infty$, and let $\mathcal{M}_X(p)$ be its inequality curve as defined in 3.1. If Y is obtained from X by a progressive transfer of income from a richer to a poorer individual (without rank reversal), then

$$\mathcal{M}_Y < \mathcal{M}_X$$

Proof . Let X denote the pre transfer distribution and Y the post transfer distribution, where Y is obtained from X by a progressive transfer of amount $\delta > 0$ from a richer individual to a poorer individual, without rank reversal. We aim to show that

$$\mathcal{M}_Y(p) < \mathcal{M}_X(p), \quad \forall p \in (0,\ 1)$$

Assume, for contradiction, that for some $p \in (0, 1)$,

$$\mathcal{M}_Y(p) \geq \mathcal{M}_X(p),$$

This inequality is equivalent to

$$1 - \frac{Q_Y(\frac{p}{2})}{Q_Y(\frac{1}{2} + \frac{p}{2})} \geq 1 - \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}$$

Then, we obtain

$$Q_Y(\frac{p}{2})Q_X(\frac{1}{2} + \frac{p}{2}) \leq Q_X(\frac{p}{2})Q_Y(\frac{1}{2} + \frac{p}{2}) \tag{3.5}$$

But the transfer raises lower quantiles and lowers upper quantiles:

$$Q_X(\frac{p}{2}) \leq Q_Y(\frac{p}{2}), \quad Q_Y(\frac{1}{2} + \frac{p}{2}) \leq Q_X(\frac{1}{2} + \frac{p}{2}).$$

Multiplying these yields

$$Q_Y(\frac{p}{2})Q_X(\frac{1}{2} + \frac{p}{2}) \geq Q_X(\frac{p}{2})Q_Y(\frac{1}{2} + \frac{p}{2}) \tag{3.6}$$

Inequalities 3.5 and 3.6 force equality, which implies $\delta = 0$, contradiction. Thus $\mathcal{M}_Y(p) < \mathcal{M}_X(p)$ for all p. Integrating over p: $\mathcal{M}_Y < \mathcal{M}_X$. Hence the new index satisfies the Pigou–Dalton principle. □

3.2 Stochastic orders based on $\mathcal{M}(p)$ curve

In defining the new income inequality curves, it is important to define the stochastic orders based on these curves. Stochastic orders determine the role of distribution parameters in income inequality. The role of income distribution parameters can be direct or indirect.

The first stochastic order of the inequality curve was defined based on the Lorenz curve. After that, stochastic orders were defined based on Benferroni, Zenga, and Canberra inequality curves. The equivalence of the stochastic order based on the famous Lorenz curve with the newly defined inequality curves has been considered (see Sarbia [15], Kleiber and Kotz [9], Polisicchio and Porro [14], and Arnold [2]). Here, we first define the stochastic order based on the Lorenz curve.

Definition 3.6. Let X and Y be random variables both with finite and positive expected value. The Lorenz order ($\leq_L$) is defined by,

$$X \leq_L Y \Leftrightarrow L_X(p) \geq L_Y(p), \qquad \forall p \in (0,1). \tag{3.7}$$

If $X \leq_L Y$, then X expresses less inequality than Y in the Lorenz's sense. From a graphic perspective, the random variable X is smaller than Y in this order, if its Lorenz curve lies above the Lorenz curve of Y for all $p \in (0,1)$.

In analogy with the Lorenz order, the order curves for the new inequality curve are defined as follows.

Definition 3.7. Let X and Y be random variables both with finite and positive expected value. The new curve order $\leq_{\mathcal{M}}$ is defined by

$$X \leq_{\mathcal{M}} Y \Leftrightarrow \mathcal{M}_X(p) \leq \mathcal{M}_Y(p), \qquad \forall p \in (0,1). \tag{3.8}$$

If $X \leq_{\mathcal{M}} Y$, then X shows less inequality than Y in new curve sense.

4 The Behavior of $\mathcal{M}(p)$ curve in some income models

In this part, we study the new inequality curve in famous income models. Also, by changing the model's parameters, the behavior of this curve is compared with the behavior of the Lorenz inequality curve. The Lorenz inequality curve is always an ascending convex curve, and the bisector of the first quadrant of the Cartesian coordinate plane evokes the state of perfect equality of incomes in this curve. The Lorenz curve shows an easy and intuitive interpretation of inequality in society. To study and compare the Lorenz curve with other income inequality curves, see Polisicchio and Porro [14] and Arcagni and Porro [1].

In the realm of modeling income distribution, various probability density functions have been proposed in research literature. Notably, these distributions are typically designed for a positive support domain. Among the simpler and commonly used distributions are the exponential, power, Pareto, and uniform distributions, each characterized by one or two parameters. Consequently, their curves exhibit distinct behaviors. To improve the fit for the tails of the distribution, more complex three-parameter models have been introduced. In this analysis, we will examine the Singh-Maddala distribution, a representative three-parameter model used to explore income distribution patterns.

4.1 Exponential model

Let X be a random variable with exponential distribution $F_X(x) = 1 - e^{-\lambda x} I_{(x>0)}$, where $\lambda > 0$, is its parameter distribution, then the quantile function is

$$Q_X(p) = \frac{-\ln(1-p)}{\lambda}, \qquad \forall p \in (0,1),$$

and from that, the median inequality curve is obtained as

$$\mathcal{M}_X(p) = \frac{\ln(1-p) - \ln(2-p)}{\ln(1-p) - \ln(2)}, \qquad \forall p \in (0,1).$$

It is important to pay attention to this point that the scale parameter λ is not an inequality indicator, in fact, the traditional Lorenz curve $L(p)$ in the left panel of Figure 3 and the new curve $\mathcal{M}(p)$ in the right panel of Figure 3 don't depend on λ.

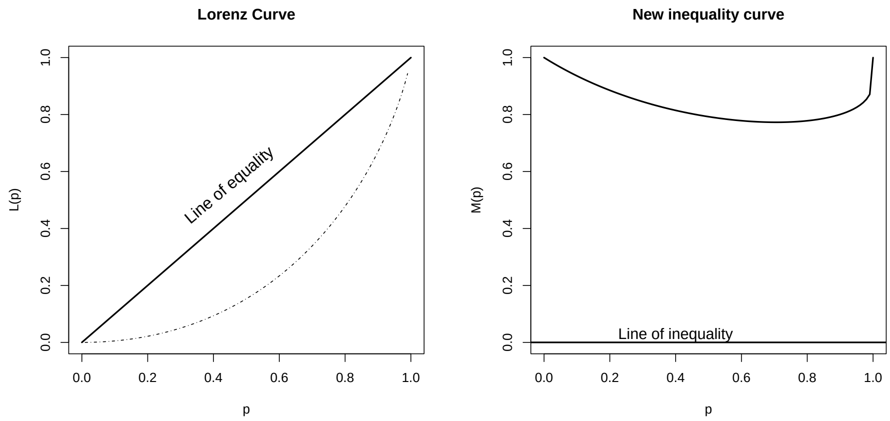

Figure 3: The Lorenz curves (left panel) and the new curves (right panel) of exponential distribution

4.2 Power model

For random variable with power distribution with ζ parameter, with

$$F_X(x) = x^{\zeta}, \qquad 0 < x < 1, \quad \zeta > 0,$$

the quantile function and the new inequality curve are:

$$Q_X(p) = p^{\frac{1}{\zeta}}, \qquad \forall p \in (0,1),$$

and

$$\mathcal{M}_X(p) = (p+1)^{\frac{-1}{\zeta}}\left[(p+1)^{\frac{1}{\zeta}} - (p)^{\frac{1}{\zeta}}\right], \qquad \forall p \in (0,1).$$

Figure 4 illustrates the impact of the distribution parameter ζ on the shape of the income inequality curves in the power model. The right panel displays several $\mathcal{M}(p)$ curves, each corresponding to a different value of ζ, highlighting how changes in ζ influence the curve's form. The left panel of the same figure shows the associated $L(p)$ curves for various values of the model parameter.

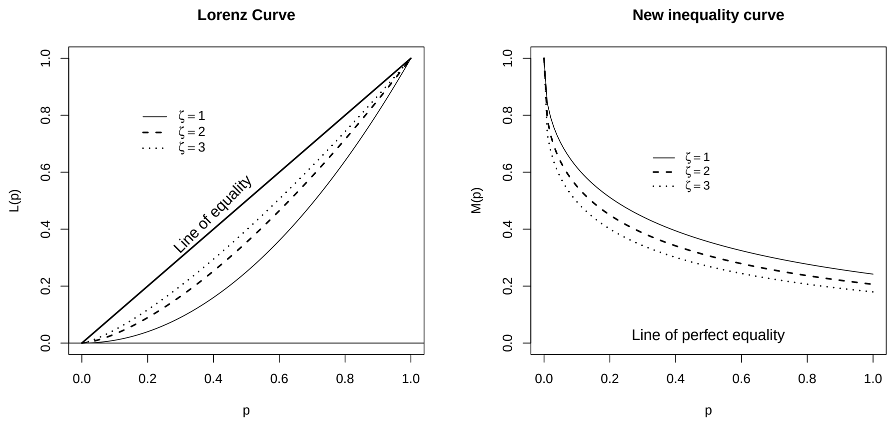

Figure 4: The Lorenz curves (left panel) and the new curves (right panel) of power distribution

We notice an intriguing link between the inequality curve and the distribution parameter ζ. When ζ increases for any fixed $p \in (0,1)$, the inequality curves get closer to the state of perfect equality, signifying a reduction in inequality. This relationship is illustrated through the $\mathcal{M}(p)$ and $L(p)$ curves. In simpler terms, a higher ζ value implies lower inequality, serving as an inverse indicator of inequality for these curves. Furthermore, by changing the parameter value, the inequality patterns represented by both curves $\mathcal{M}(p)$ and $L(p)$ are aligned and harmonious.

4.3 Pareto model

A random variable X has a Pareto distribution if its distribution function is

$$F_X(x) = 1 - (\frac{x}{\eta})^{-\nu} I_{(x>\eta)},$$

here $\eta > 0$ and $\nu > 1$. In this case,

$$Q_X(p) = \eta(1-p)^{\frac{-1}{\nu}}, \qquad \forall p \in (0,1),$$

and from that, the new curve can be obtained as

$$\mathcal{M}_X(p) = (2-p)^{\frac{-1}{\nu}}[(2-p)^{\frac{1}{\nu}} - (1-p)^{\frac{1}{\nu}}], \qquad \forall p \in (0,1). \tag{4.1}$$

The behavior of the inequality curve remains consistent regardless of changes to the scale parameter (η). This means that the measurement of inequality, whether represented by indices or curves, is not influenced by adjustments to the scale parameter.

Figures 5 showcases several examples of $L(p)$ (left panel) and $\mathcal{M}(p)$ (right panel) curves for the Pareto model. Each curve represents a distinct value of the distribution parameter (ν), demonstrating that variations in ν affect the curve's shape.

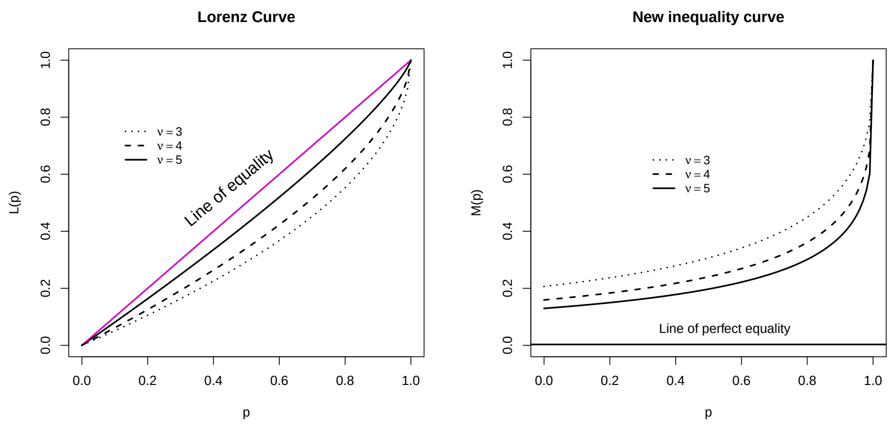

Figure 5: The Lorenz curves (left panel) and the new curves (right panel) of Pareto distribution with different values of ν

We observe an interesting relationship between the inequality curve and the distribution parameter (ν). For any fixed $p \in (0,1)$, as ν increases, the inequality curves are getting closer to the state of perfect equality, but this means the level of inequality decreases. In simpler terms, a higher ν value indicates less inequality, acting as an inverse indicator of inequality for these curves.

4.4 The uniform model

If X is a uniform random variable in the interval (α, β), then the random variable X has a classical uniform model with

$$F_X(x) = \frac{x-\alpha}{\beta-\alpha} I_{(\alpha<x<\beta)},$$

where $0 < \alpha < \beta$ are parameters distribution, In this case, the quantile function is as follows

$$Q_X(p) = \alpha + p(\beta - \alpha), \qquad \forall p \in (0,1).$$

Therefore the corresponding inequality curve can be shown as

$$\begin{aligned}\mathcal{M}_X(p) &= 1-\frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2}+\frac{p}{2})}\\ &= 1-\frac{\alpha+\frac{p}{2}(\beta-\alpha)}{\alpha+(\frac{1}{2}+\frac{p}{2})(\beta-\alpha)}\\ &= \frac{\beta-\alpha}{\beta(p+1)-\alpha(p-1)}\ .\end{aligned}$$

Here, the role of the two parameters α and β for the uniform model in relation to the new curve has been checked. Figure 6 presents two sets of $\mathcal{M}(p)$ curves. The left panel displays curves with varying values of α while keeping β constant at 5. Conversely, the right panel demonstrates the impact of changing β with a fixed α value of 2.

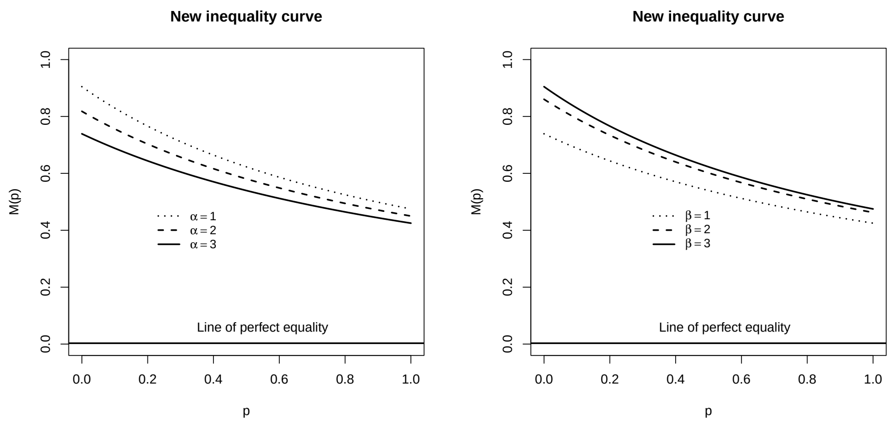

Figure 6: Comparing the new curves for uniform models with varying parameters α and β

It is obvious that, if the value of the distribution parameter β is fixed then α is an inverse inequality indicator, and if α is fixed, then β is a direct indicator.

4.5 The Singh-Maddala distribution

Let X be a non-negative random income variable with Singh-Maddala model and corresponding cdf,

$$F_X(x) = 1-\frac{1}{[1+(\frac{x}{\beta})^{\alpha}]^{\gamma}}, \qquad x>0,$$

where $\alpha,\beta,\gamma>0$. In this case, the quantile function is as follows

$$Q_X(p) = \beta[(1-p)^{\frac{-1}{\gamma}}-1]^{\frac{1}{\alpha}}, \qquad \forall p\in(0,1).$$

Therefore the corresponding inequality curve can be shown as:

$$
\begin{aligned}
\mathcal{M}_X(p) &= 1 - \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2}+\frac{p}{2})} \\
&= 1 - \frac{\beta[(1-\frac{p}{2})^{\frac{-1}{\gamma}} - 1]^{\frac{1}{\alpha}}}{\beta[(1-\frac{1+p}{2})^{\frac{-1}{\gamma}} - 1]^{\frac{1}{\alpha}}} \\
&= \frac{[(\frac{2}{1-p})^{\frac{1}{\gamma}} - 1]^{\frac{1}{\alpha}} - [(\frac{2}{2-p})^{\frac{1}{\gamma}} - 1]^{\frac{1}{\alpha}}}{[(\frac{2}{1-p})^{\frac{1}{\gamma}} - 1]^{\frac{1}{\alpha}}} .
\end{aligned}
\tag{4.2}
$$

Here, the role of the parameters α and γ for the Singh-Maddala model about the new curves has been checked. We note that since β is the scale parameter in the Singh-Maddala model, the inequality indices and curves are invariant to it, and so our results are not affected by the choice. Figure 7 presents two sets of $\mathcal{M}(p)$ curves, each exploring the impact of a different parameter on the curves. The left panel illustrates the effect of varying α while keeping γ fixed at 5. The right panel, on the other hand, showcases how the curves change as γ varies with a constant α value of 2.

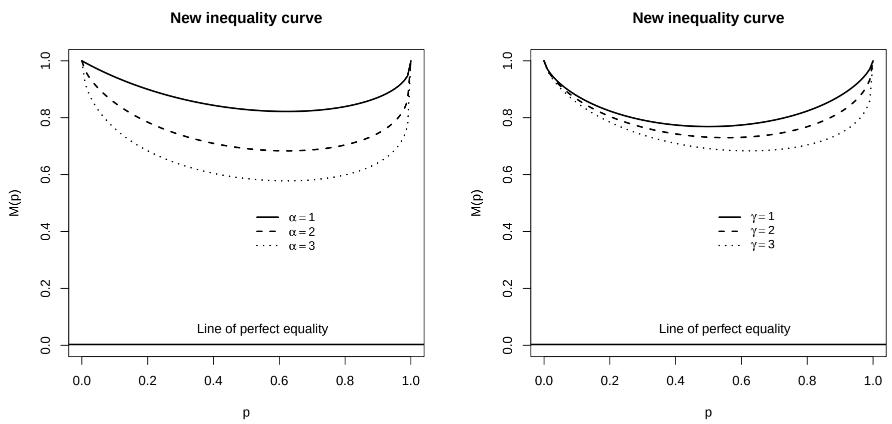

Figure 7: Comparing the new curves for Singh-Maddala models with varying parameters α and γ

The distribution parameters α and γ are inversely related in their ability to indicate inequality. When α is constant, γ acts as a measure of inequality, with higher values of γ implying less inequality. Similarly, when γ is fixed, α serves as an inverse indicator of inequality, meaning higher values of α suggest less inequality.

5 Sampling estimator of the new index

Consider a random sample $X_1, \ldots, X_n$ of size n drawn from a cdf F_X. The corresponding order statistics from this sample are denoted as $X_{1:n} \leq \ldots \leq X_{i:n} \leq \ldots \leq X_{n:n}$. We can define the new index as follows:

$$
\mathcal{M}_n = 1 - \frac{1}{\lfloor \frac{n}{2} \rfloor} \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \frac{X_{k:n}}{X_{\lceil \frac{n}{2} \rceil + k:n}}, \tag{5.1}
$$

where, for every real $X \geq 0$, $\lfloor X \rfloor$ is the largest integer that does not exceed X, and $\lceil X \rceil$ is the smallest integer that is not below X.

5.1 Consistency

Strong consistency ensures that an estimator converges almost surely to the true population parameter as the sample size increases. This property is essential for income inequality indices because it guarantees that the estimated value reliably reflects the true inequality level with probability one, irrespective of sampling variability or outliers. Furthermore, strong consistency lays the groundwork for all subsequent asymptotic procedures, including confidence intervals and hypothesis testing, thereby reinforcing the estimator's statistical credibility and practical applicability.

Theorem 5.1. Let X be a non negative random variable with continuous cdf F_X and $0 < \mathbb{E}[X] < \infty$. Let $X_1, \ldots, X_n$ be a random sample of size n drawn from F_X, then

$$\mathcal{M}_n \xrightarrow[n\to\infty]{\text{a.s.}} \mathcal{M}. \tag{5.2}$$

That is $\mathcal{M}_n$ is a strongly consistent estimator of the population index.

Proof . By the Glivenko–Cantelli theorem and standard results on empirical quantiles (see, e.g., David & Nagaraja [4]; Serfling [16]), we have, uniformly in $p \in (0, 1)$

$$Q_n(p) \xrightarrow[n\to\infty]{\text{a.s.}} Q_X(p)$$

Consequently

$$Q_n(\frac{p}{2}) \xrightarrow[n\to\infty]{\text{a.s.}} Q_X(\frac{p}{2}), \quad Q_n(\frac{1}{2} + \frac{p}{2}) \xrightarrow[n\to\infty]{\text{a.s.}} Q_X(\frac{1}{2} + \frac{p}{2}).$$

Since $0 \leq Q_X(\frac{p}{2}) \leq Q_X(\frac{1}{2} + \frac{p}{2})$ for all $p \in [0, 1]$, the continuous mapping theorem gives:

$$\frac{Q_n(\frac{p}{2})}{Q_n(\frac{1}{2} + \frac{p}{2})} \xrightarrow[n\to\infty]{\text{a.s.}} \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})}$$

Define

$$g_n(p) = 1 - \frac{Q_n(\frac{p}{2})}{Q_n(\frac{1}{2} + \frac{p}{2})}, \tag{5.3}$$

Then $0 \leq g_n(p) \leq 1$ for all p and n. The Lebesgue Dominated Convergence Theorem applies:

$$\lim_{n\to\infty} \mathcal{M}_n = \lim_{n\to\infty} \int_0^1 g_n(p)dp = \int_0^1 \lim_{n\to\infty} g_n(p)dp = \int_0^1 (1 - \frac{Q_X(\frac{p}{2})}{Q_X(\frac{1}{2} + \frac{p}{2})})dp = \mathcal{M}.$$

Therefore,

$$\mathcal{M}_n \xrightarrow[n\to\infty]{\text{a.s.}} \mathcal{M}.$$

which proves that $\mathcal{M}_n$ is a strongly consistent estimator of the new inequality index. □

5.2 Asymptotic Normality

The derivation of the asymptotic distribution of the estimator $\mathcal{M}_n$ is a fundamental step in statistical inference, as it provides the theoretical basis for quantifying estimation uncertainty and making probabilistic statements about the population parameter. The following theorem establishes the asymptotic normality of the estimator $\mathcal{M}_n$, which is essential for constructing confidence intervals and performing hypothesis tests in empirical analyses of income inequality.

Theorem 5.2. Let X be a non-negative random variable with continuous cdf F_X, probability density function (pdf) f_X and finite second moment, i.e. $0 < \mathbb{E}[X^2] < \infty$. Let $X_1, \ldots, X_n$ be a random sample of size n drawn from F_X and let $\mathcal{M}_n$ be the empirical estimator of the population index $\mathcal{M}$ as defined in 3.2. Then, as $n \to \infty$,

$$\sqrt{n}(\mathcal{M}_n - \mathcal{M}) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\ \sigma^2_{\mathcal{M}}), \tag{5.4}$$

with

$$\sigma^2_{\mathcal{M}} = \int_0^1 \int_0^1 [\frac{\partial}{\partial t}\mathcal{M}(Q_X(t))][\frac{\partial}{\partial s}\mathcal{M}(Q_X(s))] \frac{\min(t, s) - ts}{f_X(Q_X(t))f_X(Q_X(s))} dtds. \tag{5.5}$$

where $\xrightarrow[n\to\infty]{\mathcal{D}}$ denotes convergence in distribution.

Proof .

Let F be the cdf of X, and let F_n be the empirical distribution function based on the sample $X_1, \ldots, X_n$. The population index can be written as a statistical functional:

$$\mathcal{M} = \Phi(F) = \int_0^1 (1 - \frac{Q(\frac{p}{2})}{Q(\frac{1}{2} + \frac{p}{2})})dp,$$

where $Q(p)$ is the quantile function. Similarly, the empirical estimator is $\mathcal{M}_\mathrm{n} = \Phi(\mathrm{F_n})$.

The functional Φ is Hadamard differentiable at F with respect to the sup-norm, provided $f > 0$ at the relevant quantiles and $\mathbb{E}[X^2] < \infty$. Its influence function is obtained by differentiating Φ in the direction of a point mass δ_X:

$$\Psi(x) = \Phi'_F(\delta_x - F). \tag{5.6}$$

Explicitly,

$$\Psi(x) = \frac{1}{2}\int_0^1 \frac{1_{\{x<Q(\frac{p}{2})\}-\frac{p}{2}}}{f(Q(\frac{p}{2}))} \times (-\frac{1}{Q(\frac{1}{2}+\frac{p}{2})})dp + \frac{1}{2}\int_0^1 \frac{1_{\{x<Q(\frac{1}{2}+\frac{p}{2})\}-(\frac{1}{2}+\frac{p}{2})}}{f(Q(\frac{1}{2}+\frac{p}{2}))} \times \frac{Q(\frac{p}{2})}{Q(\frac{1}{2}+\frac{p}{2})}dp \tag{5.7}$$

By the functional delta method (van der Vaart [17]), since $\sqrt{n}(F_n - F)$ converges weakly to a Gaussian process $\mathbb{G}$ with covariance structure

$$Cov(\mathbb{G}_F(s), \mathbb{G}_F(t)) = F(\min(s,\ t)) - F(s)F(t), \tag{5.8}$$

we have

$$\sqrt{n}(\Phi(F_n) - \Phi(F)) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Psi(X_i) + o_p(1). \tag{5.9}$$

Since $\mathbb{E}(\Psi(X)) = 0$ (by the properties of influence functions) and $\mathbb{E}(\Psi^2(X)) < \infty$ (guaranteed by the finite second moment condition), the classical Lindeberg–Lévy Central Limit Theorem applies:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \Psi(X_i) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\ \mathbb{E}[\Psi(X)^2]). \tag{5.10}$$

Therefore,

$$\sqrt{n}(\mathcal{M}_n - \mathcal{M}) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\ \sigma^2_{\mathcal{M}}),$$

where

$$\sigma^2_{\mathcal{M}} = \mathbb{E}[\Psi(X)^2].$$

Using the covariance structure of the empirical quantile process, the variance can be expressed in terms of the quantile function and density:

$$\sigma^2_{\mathcal{M}} = \int_0^1\int_0^1 [\frac{\partial}{\partial t}\mathcal{M}(Q(t))][\frac{\partial}{\partial s}\mathcal{M}(Q(s))]\frac{\min(t,s) - ts}{f(Q(t))f(Q(s))}dtds. \tag{5.11}$$

Thus, we have established that

$$\sqrt{n}(\mathcal{M}_n - \mathcal{M}) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,\ \sigma^2_{\mathcal{M}}),$$

which proves the asymptotic normality of the estimator. □

This result enables the construction of approximate confidence intervals and hypothesis tests for the new inequality index.

Table 1: Relative frequency distribution table for American households in 2019

Total	[0,15)	[15,25)	[25,35)	[35,50)	[50,75)	[75,100)	[100,150)	[150,200)	[200,∞)
100	9.1	8	8.3	11.7	16.5	12.3	15.5	8.3	10.3

6 Simulation study

This data analysis uses information on the total money income of U.S. households from 2019. This information was gathered by the U.S. Census Bureau as part of their Current Population Survey (CPS). Table 1 shows these income figures, specifically highlighting the relative frequency of each group, which was calculated from a dataset of 128,451 individual incomes. This data is available on the Census Bureau's official website.

This analysis involves a sample of $128,451$ household income records, which have been organized into nine distinct groups. Because the income data is presented in this grouped format, a statistical method known as maximum likelihood estimation (specifically, by maximizing the multinomial log-likelihood function) was employed to derive the most accurate estimates (see McDonald, 1984 [11]).

The results of estimation and goodness of fit show that the Generalized Beta type 2 (GB2) distribution with pdf

$$f_X(x;a,b,p,q) = \frac{a x^{a\,p-1}}{b^{a\,p} B(p,q) \left[1 + (\frac{x}{b})^a\right]^{p+q}}, \tag{6.1}$$

with $x > 0,\ \ a, b > 0$ and where $B(p,q)$ is the Beta function given by

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^\infty \frac{t^{p-1}}{(1+t)^{p+q}} dt, \qquad p, q > 0.$$

is the most suitable model for representing the pattern of these household incomes. The GB2 model is particularly beneficial because it encompasses many other probability distributions as special instances. This characteristic makes it an excellent and comprehensive foundation for detailed discussions and analysis.

In this analysis, we are generating synthetic data using a Monte Carlo simulation based on a GB2 model. This model is configured with specific shape parameters: $a = 1.544, p = 0.851, q = 2.314$, and a scale parameter of 158.461. It's important to note that the b parameter, which also functions as a scale parameter in the GB2 distribution, does not influence the level of inequality observed in the data. These specific parameter values for the GB2 model were determined by applying maximum likelihood estimation to the actual total money income data for U.S. households in 2019, which was sourced from the U.S. Census Bureau's Current Population Survey (CPS).

We performed a simulation study to compare a newly developed inequality index with the widely recognized Gini coefficient (G), which is referenced in Davidson [5]. This evaluation utilized the statistical distribution (with previously mentioned parameters) that was found to represent the actual data best. To compile summary statistics, we generated 10,000 estimates for both indices, using varying sample sizes of 10, 20, 30, 50, 100, and 500.

6.1 Bias and MSE

We conducted an analysis to compare the Bias and Mean Squared Error (MSE) of the two income inequality estimates, using a distribution that was fitted to real-world data. The results of this comparison are displayed in Table 2 for easier interpretation.

The results show that the Gini estimator is affected by negative bias, that is, it underestimates the values of the related index. Also, the new estimator is affected by positive bias, that is, it overestimates the values of the related index. Furthermore, the estimator of the new index presents a smaller MSE than the estimator of the Gini index.

6.2 Normality

Here we compare the new index with the usual Gini inequality index . To do this, it simulates, 10000 simulated samples of size $n = 100$ from the fitted distribution to income data and evaluates both indices in each sample. The histograms in Fig 8 show the corresponding sampling distributions for this set of simulations.

The illustrations indicate that the sample distributions of the new index and the Gini coefficient are roughly normal. Additionally, the sampling distributions for the two estimators are alike.

Table 2: Comparison the Bias and MSE for Gini and the new measures

	G		$\mathcal{M}$	
n	Bias	MSE	Bias	MSE
10	-0.0406	0.0106	0.0385	0.0103
20	-0.0316	0.0078	0.0274	0.0065
30	-0.0232	0.0055	0.0198	0.0048
50	-0.0162	0.0037	0.0153	0.0033
100	-0.0099	0.0018	0.0087	0.0015
500	-0.0022	0.0009	0.0018	0.0008

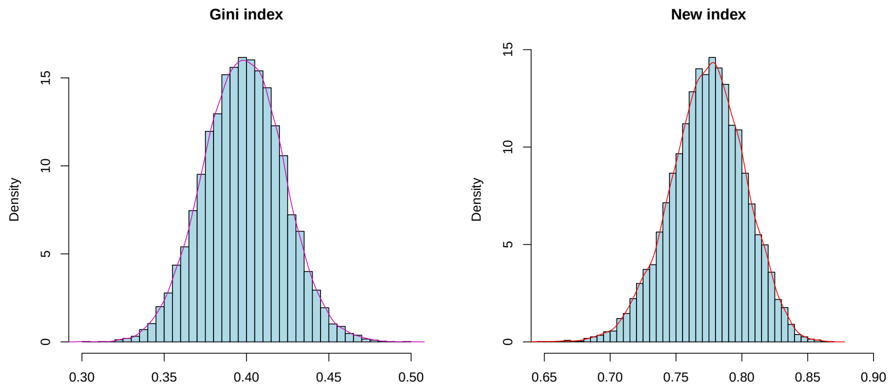

Figure 8: Histogram of the distributions of Gini and the new estimates computed on 10000 samples of size $n = 100$

We now conduct simulations to assess the performance of the new inequality index in comparison to the traditional Gini coefficient. Our analysis focuses on skewness and kurtosis derived from a distribution that has been fitted to real data using predetermined parameters. We generate 10,000 estimates for both indices across sample sizes of 10, 20, 30, 50, 100, and 500, and we present our results in Table 3. The findings presented in Table 3 indicate that the sample distributions of the Gini coefficient and the new index ($\mathcal{M}$) exhibit similar characteristics concerning skewness and kurtosis. Furthermore, as the sample size increases, the distributions of these estimators tend to converge towards a normal distribution.

Table 3: Comparison the skewness and kurtosis for Gini and the new measures

	Skewness		Kurtosiss	
n	G	$\mathcal{M}$	G	$\mathcal{M}$
10	0.307	0.257	0.356	0.320
20	0.216	0.189	0.229	0.207
30	0.174	0.168	0.172	0.169
50	0.133	0.124	0.139	0.132
100	0.089	0.077	0.085	0.081
500	0.058	0.053	0.058	0.053

6.3 Coverage probability

We compared the effectiveness of asymptotic confidence intervals for a new measure of income inequality (which competes with the $\mathcal{M}$-index) against those for the conventional Gini coefficient. We evaluated three types of confidence intervals—Normal, t-bootstrap, and BCa (as described by Mills and Zandvakili [12])—using a statistical distribution that was fitted to real income data.

In Table 4, we report the coverage probabilities (relative frequencies of confidence intervals that include the true value of the inequality measurements) for 10,000 confidence intervals, which include three types: normal, t-bootstrap, and BCa with a nominal confidence level of 0.95.

Table 4: Coverage probability of the Gini index and the new measure

	Normal		t-bootstrap		BCa	
n	G	$\mathcal{M}$	G	$\mathcal{M}$	G	$\mathcal{M}$
10	0.825	0.883	0.863	0.927	0.837	0.842
20	0.887	0.918	0.908	0.930	0.885	0.885
30	0.891	0.923	0.913	0.932	0.889	0.891
50	0.915	0.926	0.921	0.939	0.903	0.904
100	0.931	0.937	0.935	0.941	0.919	0.923
500	0.941	0.946	0.938	0.945	0.920	0.926

The results show that the coverage accuracy of the resampling techniques' confidence intervals is reasonably close to the nominal confidence level for a large sample size. As can be seen, there is no significant difference in CP for the two inequality indices. Nevertheless, we observe from Table 4 that the new index estimator plays a good role based on coverage probabilities of confidence intervals compared to the Gini estimator.

7 Application to real data

Here, to present the performance of this new inequality curve, some empirical median-inequality and Lorenz curves for the distribution of households by total money income are shown. The household income distributions considered regarding the years 2000, 2006, and 2013 are also available from the US Census Bureau. The income data used in this study comes from the US Census Bureau and adheres to their privacy regulations. This data is publicly available on the Census Bureau website. Figure 9 presents the Lorenz curves $L(p)$ and the proposed median-based curves $\mathcal{M}(p)$ for 2000, 2006, and 2013. The $\mathcal{M}(p)$ curves exhibit a more pronounced upward trend, reflecting an increase in household income inequality over the period, while the Lorenz curves show a less distinct response to these changes. This indicates that the median-based curve is more sensitive to changes in the income distribution and can serve as a useful tool for measuring income inequality.

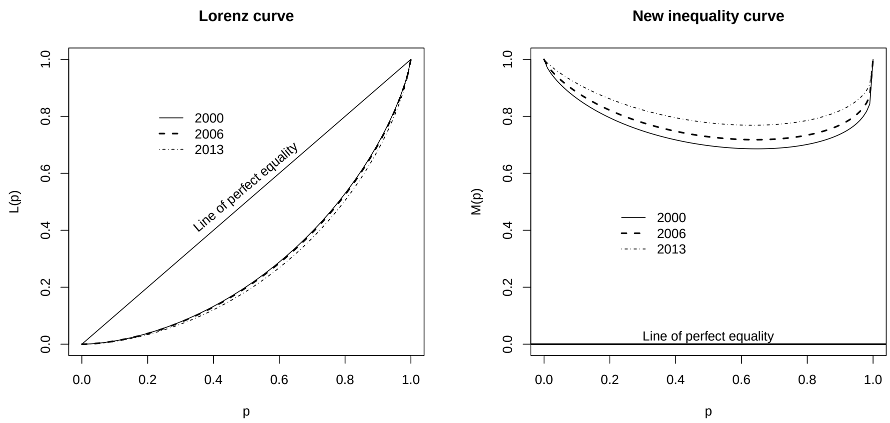

Figure 9: Lorenz curves (left panel) and median-inequality curves (right panel) of household income in different years.

The changes in household income inequality are influenced by the behavior of distribution parameters. In Table 5, we summarize the indicator parameters of Singh-Madala and the inequality indices of the household incomes in 2000, 2005, and 2013. For a discussion of the fit of Singh Madala's model to income data, see Mizaei et al [13] and McDonald [11].

The results of Table 5 show that the shape parameters of the distribution have an inverse property, and the inequality increases with their decrease. Also, the new inequality index ($\mathcal{M}$) has increased during the mentioned years and has been aligned with the famous indices of inequality, Gini and Zenga coefficients (to learn more about these indices, see Greselin et al, 2013 [8]).

Table 5: The role of distribution parameters in inequality indices

Year	α_i	γ_i	Gini	Zenga	$\mathcal{M}$
2000	1.426	4.338	0.432	0.784	0.752
2005	1.416	4.150	0.437	0.788	0.756
2013	1.355	3.701	0.459	0.807	0.774

8 Discussion

In this paper, we introduced a new inequality index and its associated curve, based on the ratio of lower and upper medians. We first verified that the proposed index satisfies the essential principles expected of any reliable inequality measure, including normalization, scale invariance, location sensitivity, and the Pigou–Dalton transfer principle, all of which were rigorously proven. We then proposed a suitable estimator for this index and established its key theoretical properties, including strong consistency and asymptotic normality, which justify its use for reliable statistical inference. We also defined a stochastic order based on the new curve, providing a framework for comparing inequality across different income distributions. The behaviour of the new curve was examined across several well known income models exponential, power, Pareto, uniform, and Singh–Maddala and its relationship with the Lorenz curve was explored. Through simulations and real data applications, we compared the performance of the new index with the Gini coefficient, finding that the proposed measure is competitive in terms of bias, mean squared error, and coverage probabilities. While these results are promising, several questions remain open. However, further research is needed to fully understand their capabilities. Specifically, we need to investigate:

- Statistical Inference: How well do these indices perform in statistical tests and how reliable are the conclusions drawn from them?
- Real-World Performance: How do these indices and their corresponding curves work in analyzing various real-world income distributions?
- Theoretical Behavior: How do these indices behave in theoretical models commonly used to represent income distributions?
- Decomposition Analysis: Can these indices be used to effectively break down inequality into different contributing factors?
- Recovering the Density: Is it possible to determine the distribution of income solely from the new inequality curve?
- Relationship to Lorenz: Is there a connection between the new inequality curve and the widely-used Lorenz curve?

Addressing these questions will provide a more complete understanding of the strengths and limitations of the proposed indices and their potential for improving our understanding of income inequality.

References

[1] A. Arcagni and F. Porro, *The graphical representation of inequality*, Revista Colombiana de Estadistica **37** (2014), no. 1, 419–436.

[2] B. C. Arnold, *On zenga and bonferroni curves*, Metron **73** (2015), no. 1, 25–30.

[3] C. E. Bonferroni, *Elementi di statistica generale*, Università commerciale L. Bocconi, 1941.

[4] H. A. David and H. N. Nagaraja, *Order statistics*, John Wiley and Sons, 2004.

[5] R. Davidson, *Reliable inference for the gini index*, Journal of Econometrics **150** (2009), 30–40.

[6] Y. Davydov and F. Greselin, *Inferential results for a new measure of inequality*, The Econometrics Journal **22** (2019), 153–172.

[7] J. L. Gastwirth, *Median-based measures of inequality: Reassessing the increase in income inequality in the u.s. and sweden*, Statistical Journal of the IAOS **30** (2014), 311–320.

[8] F. Greselin, L. Pasquazzi, and R. Zitikis, *Contrasting the gini and zenga indices of economic inequality*, Journal of Applied Statistics **40** (2013), 282–297.

[9] C. Kleiber and S. Kotz, *Statistical size distributions in economics and actuarial sciences*, John Wiley and Sons, Hoboken, New Jersey, 2003.

[10] M. O. Lorenz, *Methods of measuring the concentration of wealth*, Journal of the American Statistical Association **9** (1905), no. 70, 209–219.

[11] J. B. McDonald, *Some generalized functions for the size distribution of income*, Econometrica **52** (1984), 647–663.

[12] J. A. Mills and S. Zandvakili, *Statistical inference via bootstrapping for measures of inequality*, Journal of Applied Econometrics **12** (1997), 133–150.

[13] S. Mirzaei, G. R. Mohtashami Borzadaran, M. Amini, and H. Jabbari, *Two competitor inequality indices in burr xii distribution*, International Journal of Statistics and Economics **18** (2017).

[14] M. Polisicchio and F. Porro, *A comparison between lorenz $L(p)$ curve and zenga $I(p)$ curve*, Statistica Applicata **21** (2011), no. 3–4, 289–301.

[15] J. M. Sarabia, *Parametric lorenz curves: Models and applications*, Modeling Income Distributions and Lorenz Curves, Springer, New York, 2008, pp. 167–190.

[16] R. J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley and Sons, 2009.

[17] A. W. van der Vaart, *Asymptotic statistics*, Cambridge University Press, 2000.

[18] M. Zenga, *Inequality curve and inequality index based on the ratios between lower and upper arithmetic means*, Statistica Applicazioni **5** (2007), no. 1, 3–28.