International Statistical Review

Check for updates

International Statistical Review (2026) doi: 10.1111/insr.70052

A Comparative Note on the Reliability of Inequality Indicator Estimates

Silvia Scarpa[1], Stefan Sperlich[2] and Maria Rosaria Ferrante[3]

[1]*Department of Economics Marco Biagi, University of Modena, Modena, Italy*
[2]*Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland*
[3]*Department of Statistical Sciences, University of Bologna, Bologna, Italy*
Corresponding to: Silvia Scarpa, Department of Economics Marco Biagi, University of Modena, Modena, Italy. Email: silvia.scarpa@unimore.it

Summary

The importance of accurate measurement of economic or social inequality is generally accepted. Estimators of its most popular measure, the Gini coefficient, are commonly criticised for becoming increasingly imprecise for increasingly skewed distributions and small to moderate samples, i.e., two phenomena that we face more frequently today. More robust inequality measures, typically based on quantile ratios, are well studied in theory, but still attract little attention in practice. We compare through simulations bias, mean squared error and sensitivity to outliers of several inequality estimators, namely, the Gini and quantile-based indicators, including the recently proposed quantile ratio index (QRI). Our results, based on Italian SILC and synthetic data, demonstrate that the QRI estimator offers superior precision and robustness, making it a reliable tool for monitoring inequality.

Key words: Gini index; inequality indicators; quantile ratio index; quantile-based indicators.

1 Introduction

Monitoring economic or social inequality requires statistical estimators that are reliable given the available data in practice. The literature offers a wide array of indices that summarise inequality in a single value, each of them giving different weights to different segments of the distribution of interest. The choice of the indicator in practice may reflect the researcher's objectives and beliefs (Sen & Foster, 1997), or other preferences, such as those for certain theoretical properties (Jenkins & Van Kerm, 2009). The focus of this note is on the statistical properties of inequality indicators in finite samples; and for the sake of notation, let us concentrate on income. Many of the widely used measures are moment-based or rely heavily on measures of central tendency of the target variable's distribution, like the mean. Such dependency can easily make them sensitive to extreme values, and their estimates may suffer from large variance under complex survey designs or when the underlying distribution is heavy-tailed, see for instance Cowell & Flachaire (2007), which, in turn, has serious consequences for subsequent inference (Davidson & Flachaire, 2007). This issue is particularly severe for indices that assign considerable weight to top incomes (Atkinson, 2007), which have grown substantially in recent decades and represent a disproportionately large share of total income as highlighted by Atkinson *et al.* (2011).

Unlike moment-based measures, quantile-based inequality indicators are expected to be much more robust because they focus on order statistics. This has been studied and emphasised by various authors in different contexts, see for instance (Aaberge & Atkinson, 2013; Brazauskas *et al.*, 2024) or Jokiel-Rokita & Piątek (2025). However, most of them consider only particular portions of the distribution, say, arbitrary quantiles in the tails, which have to be selected a priori, neglecting all other information contained in the modal or middle range. Although heaviness and length of the upper and lower tails are key drivers of overall inequality, the economic well-being of middle-income individuals also plays a fundamental role in shaping the degree of heterogeneity as was emphasised by Cobham & Sumner (2013) among others. Consequently, a proper measure of inequality should integrate both the tails and the centre, without sacrificing robustness or precision (Gastwirth, 2017). This explains the popularity of the Gini index, together with its simplicity, its intuitive interpretation and the compliance with some of the standard axioms, see Giorgi & Gigliarano (2017) for a recent review. However, its estimation often suffers downward biases in small samples, especially under heavy-tailed distributions, as studied by (Deltas, 2003; De Nicolò *et al.*, 2024), further accentuated by under-reporting of top incomes in surveys (Jorda & Niño-Zarazúa, 2019). Because upper-tail growth increases both the aggregate income gaps in the numerator and the mean in the denominator, its overall increase is attenuated, as observed and discussed by various authors like (Davydov & Greselin, 2020; Gastwirth, 2014) and Cowell & Flachaire (2024). Finally, its unbounded influence function makes it sensitive to extreme values, producing large sampling variances (Cowell & Victoria-Feser, 1996; Giorgi & Gigliarano, 2017). This explains why it is often accompanied by a simple quantile-based indicator such as the quintile share ratio (QSR) provided by Eurostat.

This note provides a comparison of commonly used quantile-based inequality measures under complex survey data, namely, the QSR, inter-decile ratios and the Palma index, together with the recently introduced quantile ratio index (QRI) (Prendergast & Staudte, 2018), and the Gini coefficient as a benchmark. Although moment-based indicators are well studied, comparative evaluations of quantile-based indicators in finite populations are rare, particularly for small samples, which are increasingly relevant for monitoring small domains or clusters. To date, most studies have focused on their theoretical properties (Brazauskas *et al.*, 2024). Our computational note complements these by comparing the performance of their estimators presented in Section 2.1—in terms of bias, variability and sensitivity to skewed distributions and outliers—particularly in small samples under potentially complex designs. We generate data from the IT-SILC survey (see Section 2.2 for details) and evaluate the indicators' estimators through simulations, also examining influence functions as key determinants of robustness and precision (Section 3). All the indicators together with their influence functions are estimated via the R package `inequantiles` (Scarpa & Sperlich, 2026). Our main finding is that the QRI outperforms the other indicators in bias and mean squared error, capturing the entire distribution like the Gini, but with greater robustness. For the ease of presentation we consider as target variable positive income, although our statements hold equally true for other variables that range from zero to an arbitrarily large but bounded value like gross wealth and expenditures.

2 Considered Indicators and Data

2.1 The Inequality Indicators

A simple and intuitive way to analyse inequality is to compute (not necessarily symmetric) percentile ratios. A popular choice is to compare the 90th percentile to that at the 10th percentile, say P90/P10. Provided with quantile estimator $\widehat{Q}(p)$ for $p \in [0, 1]$ one would obtain the estimator

$$\widehat{P90/P10} = \frac{\widehat{Q}(p = 0.9)}{\widehat{Q}(p = 0.1)}.$$

Its main advantage is robustness to extreme observations, making it less affected by top-coding and under-reporting of top incomes in survey data, although it requires the a priori selection of the quantiles and does not take into account anything beyond those two quantiles; see Van Kerm (2007) and Jenkins & Van Kerm (2009) for further discussion.

It is more common to consider the ratio of percentile aggregates, the so-called inter-decile share ratios. A frequently used one is the Palma index (Palma, 2006; 2011) which divides the share earned by the richest 10% by the share of the poorest 40%. It is featured, for instance, in the UN Sustainable Development Goals Report. Consider a sample s containing observations y_j with sampling weights w_j for $j = 1, \ldots, n$. The sampling weights – usually the inverse of the first-order inclusion probabilities often calibrated to known population totals and adjusted for nonresponse – are used to produce design-consistent estimates (Pfeffermann, 1993). The Palma index can be estimated by

$$\widehat{Palma} = \frac{\sum_{j \in s} w_j y_j \mathbb{1}\left\{y_j \geq \widehat{Q}(0.9)\right\}}{\sum_{j \in s} w_j y_j \mathbb{1}\left\{y_j \leq \widehat{Q}(0.4)\right\}}.$$

Similarly, an indicator of inequality often used in rich countries (e.g., by Eurostat) is the QSR. It compares the total income earned by the richest 20% to that earned by the poorest 20% and can be estimated similarly to the above; see Langel & Tillé (2011) for more details by

$$\widehat{QSR} = \frac{\sum_{j \in s} w_j y_j \mathbb{1}\left\{y_j \geq \widehat{Q}(0.8)\right\}}{\sum_{j \in s} w_j y_j \mathbb{1}\left\{y_j \leq \widehat{Q}(0.2)\right\}}.$$

QSR and the Palma index focus on the tails of distributions, based on the idea that middle income is stable such that inequality depends mainly on the tails. They have the advantage of being easily formulated and quite intuitive. Yet, they neglect the majority of the population, are functions of arbitrarily chosen percentiles and are extremely sensitive to very small or large incomes (Van Kerm, 2007). One may also criticise their dependence on the chosen quantile estimators as there exist many of them. For the rest of this article, all results for the P90/P10, the QSR and the Palma index are based on the quantile estimator recommended by Eurostat (see Alfons & Templ (2012) for more details):

$$\widehat{Q}_{Euro}(p) = \begin{cases} \frac{1}{2}(y_{(k)} + y_{(k+1)}) \text{ if } \sum_{j=1}^{k} w_j = p \sum_{j=1}^{n} w_j, \\ y_{(k+1)} \text{ if } \sum_{j=1}^{k} w_j < p \sum_{j=1}^{n} w_j < \sum_{j=1}^{k+1} w_j, \end{cases}$$

where $y_{(k)}$ are the k-th order statistic of our observations. We checked by simulations that our final findings and conclusions do not change when other reasonable quantile estimators are used.

Today, the coefficient proposed by Gini (1914) is the most widely used indicator of inequality. Therefore, although not based on quantiles, this so-called Gini coefficient is considered a benchmark in this study. Many articles have studied its statistical properties; see for instance (Giorgi & Gigliarano, 2017; Langel & Tillé, 2013) for more recent reviews. Apart from its clear interpretation, its popularity is also based on the fact that it accounts for the whole distribution of the population, and it respects some nice axioms. In practice, this coefficient is typically estimated from survey data. We take a popular estimator under potentially complex sampling design, see again Alfons & Templ (2012):

$$\widehat{G} = \frac{2\sum_{j \in s} w_j y_j \sum_{j \in s} w_j - \sum_{j \in s} w_j^2 y_j}{\sum_{j \in s} w_j \sum_{j \in s} w_j y_j} - 1.$$

Various attempts have been made to address the criticism that the quantile-based indicators ignore large parts of the distribution on one side and the downsides of the Gini on the other side. The latter refer to its dependence on the mean causing serious problems in practice for increasingly long-tailed income distributions. Among many others, (Davydov & Greselin, 2020; Prendergast & Staudte, 2016; Zenga, 2007) propose and study a series of those 'attempts', either by using quantile-based Lorenz curves for Gini-like indicators or by extending the concept of integration over all quantile ratios.

A recently proposed alternative is the so-called QRI of Prendergast & Staudte (2018), although it is still hardly used in practice. This QRI measures the distance between the equi-distribution case and the integral over the ratio between symmetric quantiles around the median, say $QRI = \int_0^1 (1 - Q(p/2)/Q(1 - p/2))\, dp$, where $Q(p)$ denotes the quantile function for $p \in (0, 1)$. It can be represented as the area between the equi-distribution line and the corresponding inequality curve. Scarpa *et al.* (2025) introduced an estimator of the QRI for the context of finite populations and survey data under potentially complex design, together with an intensive study of its properties. As this indicator is not well known, we briefly give some details about the proposed estimator.

Let $W_j = \sum_{i \in s} w_i \mathbb{1}\{i \leq j\}$ denote the cumulative sum of the weights up to the j-th ordered observation. Consider an equi-distance grid on $(0, 1)$ of M points $p_m = \frac{m - 1/2}{M}$, $m = 1, \ldots, M$, on which the ratio of symmetric quantiles is evaluated to approximate the integral over the entire distribution. One can estimate the finite population quantiles at any point $0 < p_m < 1$ by

$$\widehat{Q}(p_m) = y_{(k-1)} + \left(y_{(k)} - y_{(k-1)}\right)\left(\frac{(W_n + w_n)p_m - W_{k-1}}{w_k}\right), \tag{1}$$

where $y_{(k)}$ is the k-th order statistic and $W_{k-1} \leq (W_n + w_n)p_m < W_k$ by definition of W_j. For all p_m, estimate the ratio of symmetric quantiles $R(p_m) = \frac{\widehat{Q}(p_m/2)}{\widehat{Q}(1 - p_m/2)}$, with $R(0) = 0$ and $R(1) = 1$. Then, an estimator of the QRI is given by

$$\widehat{QRI} = \frac{1}{M}\sum_{m=1}^{M}\left(1 - \frac{\widehat{Q}(p_m/2)}{\widehat{Q}(1 - p_m/2)}\right), \tag{2}$$

where one can verify that $0 \le \widehat{QRI} \le 1$; under perfect equality the ratio equals 1 for all m. Consequently, the larger the distance between the lower and upper quantiles, the greater the inequality and the closer $\hat{QRI}$ is to 1. Because income typically has a strongly positively skewed distribution, like many other economic variables, one expects the symmetric quantile ratios to be close to 0 when comparing income values at opposite tails of the distribution.

Recall that the average in Equation (2) is intended to approximate the integral in the QRI definition. Therefore, under the assumption of having reasonable estimates for all quantiles $\widehat{Q}(p_m/2)$ and $\widehat{Q}(1 - p_m/2)$, one expects larger M to yield better approximations. Although this suggests that there may be an optimal trade-off between the quality of quantile estimates and the integral approximation, extensive simulations revealed that the quality of the QRI estimate was remarkably robust for $M = 100$. The potential effects of poor quantile estimates for large M but small sample size appeared to cancel out in formula (2). At the same time, a finer grid beyond $M = 100$ hardly improved the integral approximation. Consequently, increasing M beyond 100 did not give better results, whereas smaller values of M offered no advantage and sometimes gave clearly worse QRI estimates. We admit that one could think about automatic choices of M that adapt to sample size and distribution. As this, however, would clearly go beyond the scope of this work without a significant gain in practice, we decided to continue this study with $M = 100$ grid points.

2.2 Statistics on Income and Living Conditions (SILC)

SILC is a household and individual data survey about income distribution, social exclusion and poverty. The Italian SILC survey 2017 (IT-STILC) used in this work employs a two-stage stratified sampling method for the selection of each sample. The primary sampling units (PSUs) are the municipalities that distinguish between self-representatives (SRs) and not-self-representatives (NSRs). The former are larger and considered as strata themselves. For NSRs, a two-stage stratified sampling design is performed: first, PSUs are selected with inclusion probabilities proportional to the dimension of the municipality, stratified according to geographical and demographic characteristics; afterward, a simple random sampling without replacement of households is used in the drawn municipalities. Note that here, the IT-SILC survey data are only used for simulation purposes. The target variable is the equivalised disposable income defined as the net household income adjusted for household size and composition according to the OECD-modified scale. Only positive values are considered, as all the studied indicators are designed to handle variables with strictly positive support. We believe that this restriction has negligible practical impact because less than 0.1% of all observations in this dataset take negative values.

3 Comparison Through Simulations

3.1 Accuracy and Precision of Inequality Indicators Estimators

From the IT-SILC data, $R = 1\,000$ samples are drawn using a two-stage stratified sampling design that aims to mimic the original sampling design, subject to the limitations of the available survey information. The 20 Italian regions are used as strata and treated as target domains. Samples are drawn under three scenarios, approximating 10%, 5% and 3% of the target dataset size, with PSUs and households selected at rates designed to achieve these proportions. The regions' average sample sizes vary for each of these sampling rates as follows: between 82 and 549, 37 and 256, 21 and 134, with average numbers of observations being 243, 123 and 74, respectively.

As quality measures for estimators $\widehat{\theta}_r$ ($r = 1, \ldots, R$) of an inequality parameter θ, consider

$$\mathrm{RB}(\widehat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{\widehat{\theta}_r}{\theta} - 1\right),$$

for the relative bias and, analogously, for the relative root mean squared error

$$\mathrm{RRMSE}(\widehat{\theta}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\frac{\widehat{\theta}_r}{\theta} - 1\right)^2}.$$

These measures are calculated for the 20 regions and reported in terms of summary statistics in Table 1 in percentages for each indicator and sampling rate. The RB values for QRI are close to 0 on average for samples of large to moderate size, but -3.66% for very small samples. Although the estimator for the Gini coefficient behaves similarly in underestimating the true index value, it clearly performs worse than the QRI estimator. The $\widehat{P90/P10}$, $\widehat{\mathrm{Palma}}$ and $\widehat{QSR}$ are quite imprecise for small sampling rates. Only the Palma estimates seem to outperform those for the Gini in terms of bias, but not those for the QRI.

Table 1. *Summary statistics of RB and RRMSE (in %) of all inequality indicators on 1000 samples with different sampling rates (10%, 5%, 3%).*

Rate	*Indicator*	*Measure*	*Min*	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*Max*
10%	$\widehat{P90/P10}$	RB	0.87	2.28	3.23	4.34	5.28	16.45
		RRMSE	10.84	14.53	21.86	23.16	30.71	54.17
	$\widehat{Palma}$	RB	−1.62	−0.29	0.31	0.27	0.85	2.77
		RRMSE	12.6	15.5	17.92	19.37	22.76	28.03
	$\widehat{QSR}$	RB	−15.26	−9.28	−6.57	−7.29	−4.26	−2.75
		RRMSE	12.02	14.16	19.17	19.57	24	30.19
	$\widehat{G}$	RB	−3.59	−2.52	−1.35	−1.71	−1.17	−0.52
		RRMSE	6.43	7.67	9.2	9.51	11.11	12.45
	$\widehat{QRI}$	RB	−1.19	−0.49	−0.22	−0.31	−0.05	0.24
		RRMSE	3.41	4.67	5.52	5.81	7.21	8.3
5%	$\widehat{P90/P10}$	RB	3.59	7.18	16.66	18.78	20.14	71.11
		RRMSE	20.82	33.89	63.03	98.32	87.34	591.33
	$\widehat{Palma}$	RB	−0.5	0.62	2.7	3.25	4.66	10.88
		RRMSE	20.93	25.92	36.75	35.31	42.2	56.6
	$\widehat{QSR}$	RB	−32.95	−21.96	−16.19	−17.77	−11.82	−6.54
		RRMSE	20.29	28.53	33.69	36.58	44.86	61.44
	$\widehat{G}$	RB	−8.36	−5.57	−4.59	−4.45	−2.67	−1.65
		RRMSE	9.9	13.05	15.26	15.14	16.93	19.91
	$\widehat{QRI}$	RB	−3.5	−1.4	−0.86	−1.02	−0.51	0.17
		RRMSE	6.31	8.04	9.08	9.65	11.65	13.4
3%	$\widehat{P90/P10}$	RB	−5.56	10.05	18.18	49.73	80.68	201.88
		RRMSE	41.61	63.97	84.74	537.09	511.07	3064.02
	$\widehat{Palma}$	RB	0.12	2.87	8.45	15.03	18.81	61.79
		RRMSE	34.82	42.76	55.63	85.25	80.26	352.28
	$\widehat{QSR}$	RB	−64.08	−42.68	−26.75	−19.82	−21.14	251.25
		RRMSE	35.55	40.15	57.62	280.62	77.36	4169.42
	$\widehat{G}$	RB	−21.3	−13.76	−8.25	−10.1	−6.07	−3.96
		RRMSE	14.04	16.99	20.86	21.33	23.97	31.04
	$\widehat{QRI}$	RB	−11.97	−4.92	−2.14	−3.66	−1.08	−0.23
		RRMSE	9.79	10.85	14.75	15.37	17.49	26.95

Next, we ran simulations with synthetic data, drawing samples from a fixed artificial population. This allowed us to study the estimators' performance under controlled distributional settings, including scenarios with heavier tails than those observed in the real data. As such a reference distribution, we chose the Log-Normal, not arbitrarily, as it is one of the most widely used parametric models for income data (Cowell, 2011). To guarantee that this distribution somewhat resembled our observed ones, we first fitted a Log-Normal density to our SILC data by maximum likelihood, obtaining $\mathcal{LN}(10.27,\ 0.6)$.[1] This was used to generate a synthetic finite population of $N = 100\,000$ individuals. For each sample size n from 5 to 100, we drew $R = 1\,000$ samples using Midzuno's proportional-to-size scheme in which the auxiliary weights were created as in Alfons *et al.* (2013), mimicking a complex unequal probability design by assigning larger inclusion probabilities to poorer people. The solid lines in Figure 1 show the RB and RRMSE of our estimates in percentages as functions of sample size n. Certainly, these figures are not fully comparable with the results shown in Table 1 because the latter is based on the real rather than the synthetic data. The much larger biases observed for the real data suggest the presence of outliers in the SILC samples, which our synthetic population does not fully reproduce. We therefore repeated the simulation exercise using a much more skewed population, i.e. $\mathcal{LN}(10.27,\ 4)$, in which the location parameter is preserved, whereas the scale parameter is deliberately set to a larger value to investigate the effect of heavy tails. The dashed lines in Figure 1 show the effects of this increased skewness on our estimators. Overall, $\widehat{G}$ and $\widehat{QRI}$ are again much less affected by the increased skewness than the other indicators, with $\widehat{QRI}$ showing the smallest RB and RRMSE throughout.

3.2 Robustness to Anomalous Values

Inequality indicators are estimated on data that often exhibit heavy-tailed distributions, including outliers. They may arise from errors during recording or may be true but extreme observations, as observed by (Cowell & Flachaire, 2007). In either case, their existence can cause serious biases and large sampling variances depending on their presence or absence in the sample (Van Kerm, 2007). Such absence is particularly frequent due to the systematic under-representation of top-income households caused by unit non-response in household surveys (Hlasny & Verme, 2018). For this reason, it is often recommended in the literature to adjust estimators using methods such as modelling the upper tail (Alfons *et al.*, 2013; Jenkins, 2017) or applying trimming and winsorising (Van Kerm, 2007). However, the choice of the procedures and their hyper-parameters is arbitrary, and they all lead to different estimates. It might therefore be preferable to use indicators and estimators that are more robust and allow us to disregard those arbitrary adjustments.

In this context, the boundedness of the influence function is of special interest, as it ensures a limited impact of anomalous values (Cowell & Victoria-Feser, 1996; Cowell & Flachaire, 2007). We therefore first compare the influence functions of the QRI, the Gini and the other quantile-based indicators treated in this study (P90/P10, QSR and Palma), before presenting further simulation outcomes for the estimators themselves. Deville (1999) extended the classical concept of influence functions from robust statistics theory, which was first studied by Hampel *et al.* (1986), to finite population survey sampling, linearising complex estimators under general designs. In this way, Scarpa *et al.* (2025) derived the influence function of the ratio between two quantiles:

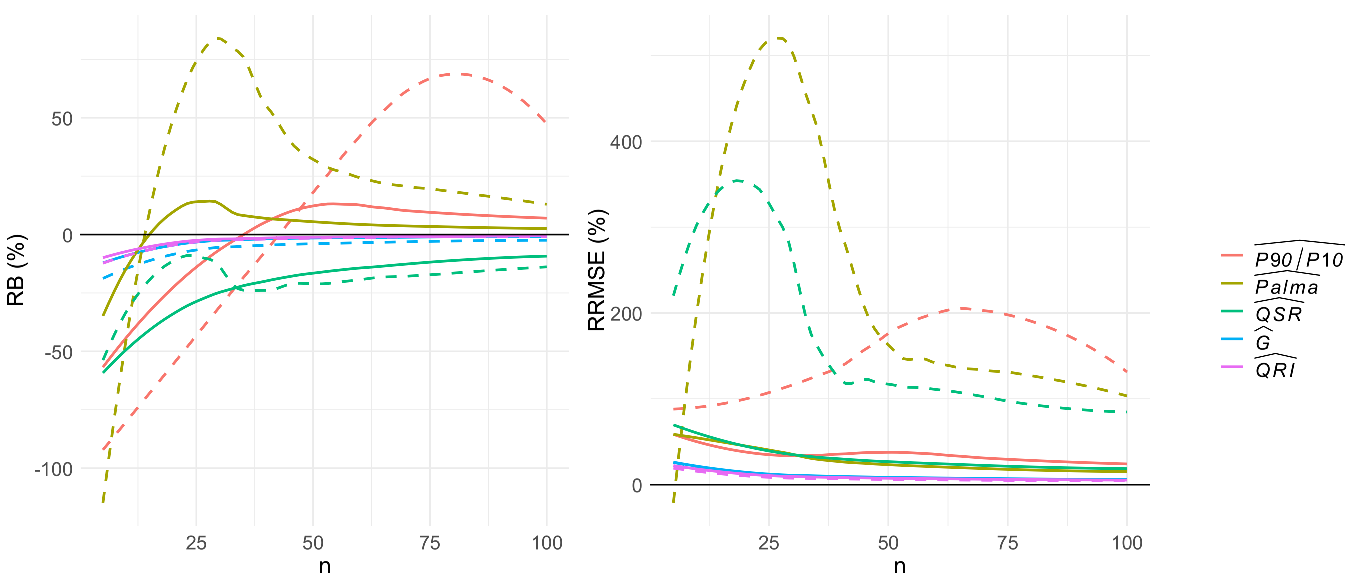

FIGURE 1. *RB and RRMSE (%) of the estimators by sample size n when the finite population is generated from* $\mathcal{LN}(10.27,\ 0.6)$ *(solid line) and from* $\mathcal{LN}(10.27,\ 4)$ *(dashed line).*

$$I\left(\frac{\widehat{Q}(p_n)}{\widehat{Q}(p_d)}\right)_j = \frac{\left(\frac{p_n - \mathbb{1}(y_i \leq \widehat{Q}(p_n))}{\hat{F}'(\hat{Q}(p_n))\hat{N}}\right)\widehat{Q}(p_d) - \left(\frac{p_d - \mathbf{1}(y_i \leq \widehat{Q}(p_d))}{\hat{F}'(\hat{Q}(p_d))\hat{N}}\right)\widehat{Q}(p_n)}{\hat{Q}(p_d)^2}$$

to then obtain the QRI influence function by integrating over $p \in (0,1)$:

$$I(\widehat{QRI})_j = -\int_0^1 \frac{\left(\frac{\frac{p}{2} - \mathbb{1}(y_j \leq \widehat{Q}(p/2))}{\hat{F}'(\hat{Q}(p/2))\hat{N}}\right)\widehat{Q}(1-p/2) - \left(\frac{\left(1-\frac{p}{2}\right) - \mathbb{1}(y_j \leq \widehat{Q}(1-p/2))}{\hat{F}'(\hat{Q}(1-p/2))\hat{N}}\right)\widehat{Q}(p/2)}{\hat{Q}(1-p/2)^2} dp,$$

where $\widehat{N} = \sum_{j \in s} w_j$; and $\widehat{F}'(y)$ is (in this study) a Gaussian kernel density estimate of the income distribution, with bandwidth $h = 0.79 \cdot \left(\widehat{Q}(0.75) - \widehat{Q}(0.25)\right) \cdot \widehat{N}^{-1/5}$. The QSR estimator influence function, as demonstrated by Langel & Tillé (2011), can be computed as follows:

$$I(\widehat{QSR})_j = \frac{y_j - \left\{0.8\widehat{Q}(0.8) - \left(\widehat{Q}(0.8) - y_j\right)\mathbf{1}\left[y_j \leq \widehat{Q}(0.8)\right]\right\}}{\widehat{Y}_{0.2}} - \frac{\left(\widehat{Y} - \widehat{Y}_{0.8}\right)\left\{0.2\widehat{Q}(0.2) - \left(\widehat{Q}(0.2) - y_j\right)\mathbf{1}\left[y_j \leq \widehat{Q}(0.2)\right]\right\}}{\widehat{Y}_{0.2}^2}$$

The Palma index shares the same structure as the QSR and its influence function can be derived analogously, replacing the relevant quantiles accordingly.

The influence function of the Gini coefficient is approximated in Langel & Tillé (2013) by

$$I(\widehat{G})_j = \frac{1}{\hat{N}\hat{Y}}\left\{2W_j\left(y_j - \widehat{\overline{Y}}_j\right) + \widehat{Y} - \widehat{N}y_j - \widehat{G}\left(\widehat{Y} + y_j\widehat{N}\right)\right\}$$

$$\text{with} \qquad \widehat{Y} = \sum_{j \in s} w_j y_j; \quad \widehat{\overline{Y}}_j = \frac{\sum_{l \in S} w_l y_l 1\left(W_l \nleq W_j\right)}{W_k}.$$

The influence functions show how much each observation affects the final estimate, no matter if it is a true value or a measurement error; robustness protects the estimator from negative impacts in both situations in the same way. In Section 3.1, we already got a smell of it in Table 1 and Figure 1. Figure 2 displays the values of the influence functions for one of the above simulated SILC samples with 3% sampling rate. Clearly, $I(\widehat{QRI})_j$ lies within a narrow range, has a discontinuous change at the median and flattens out for higher incomes. $I(\widehat{G})_j$, $I(\widehat{QSR})_j$, $I(\widehat{Palma})_j$ are growing unboundedly, reaching substantially larger values for increasing income. The influence functions of the share ratios (QSR and Palma) have discontinuous changes at their respective defining quantiles and are flat in-between them. $I\left(\widehat{P90/P10}\right)$ is a step function admitting only three distinct values.

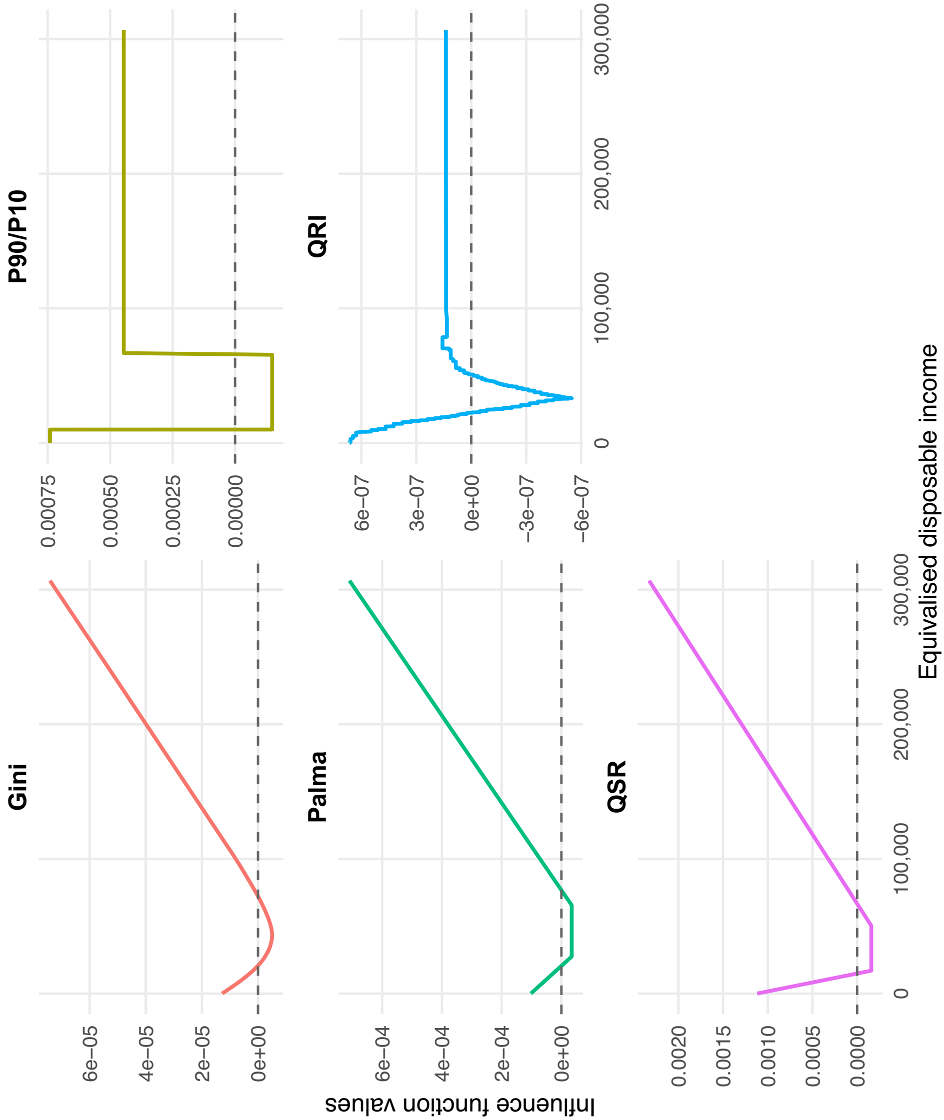

FIGURE 2. *Quantile-based indicators and Gini influence function estimates on a 3% IT-SILC sample.*

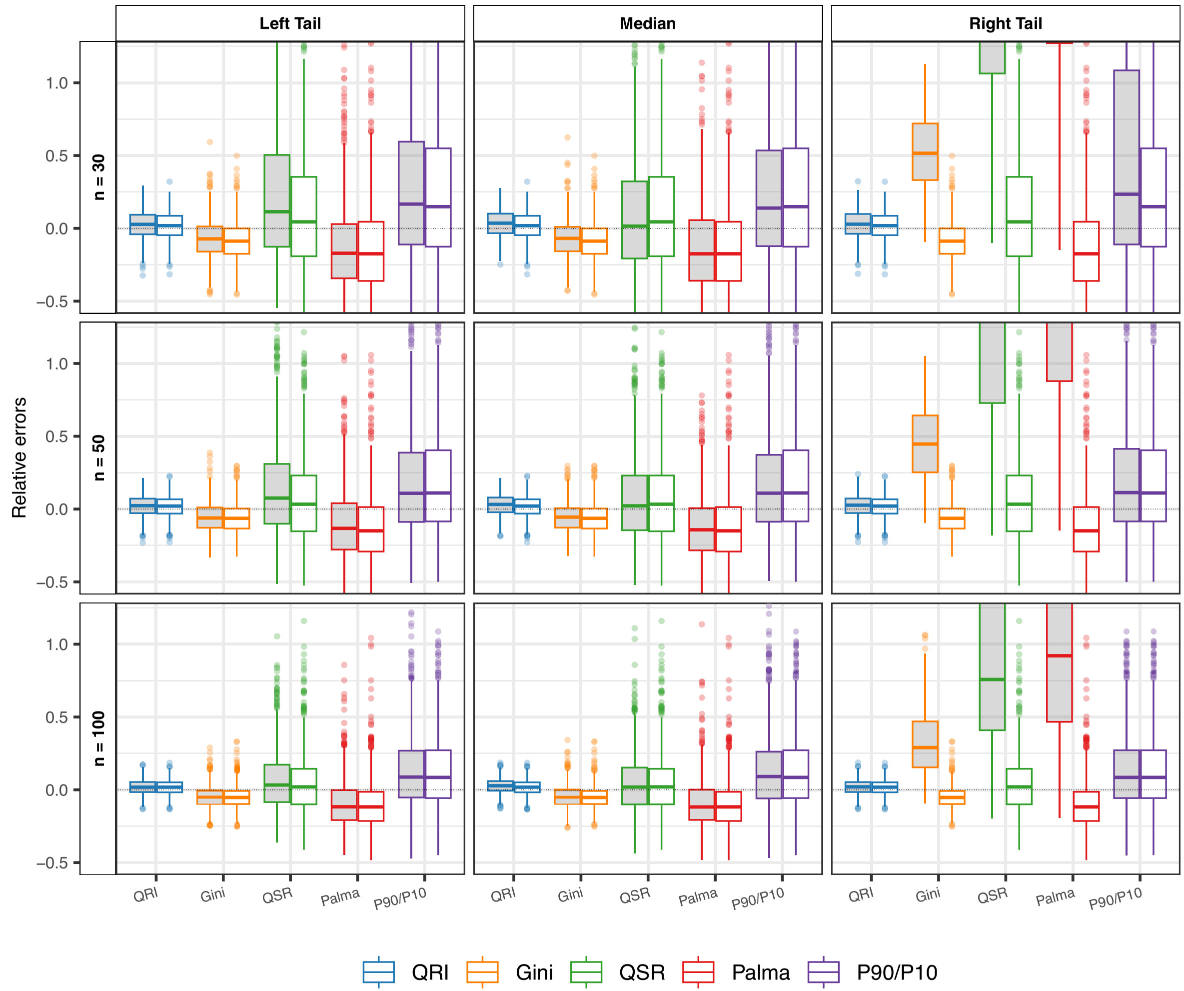

FIGURE 3. *Distributions of relative errors of the inequality estimates for samples of sizes 30, 50 and 100 from a* $\mathcal{LN}(10.27,\ 0.6)$ *population without (empty boxplots) and with contamination (filled boxplots).*

We now study the impact of measurement errors on our estimators. Consider the $R = 1\,000$ samples of size $n = 30, 50, 100$ drawn within the simulation framework on synthetic data described in Section 3.1, with $\mathcal{LN}(10.27,\ 0.6)$. Similarly to Cowell & Victoria-Feser (1996), we first estimate the parameters, and then contaminate the sample observations as follows: In one case, we multiply the largest observation by 10; in another, we divide the smallest by 10—a so-called "decimal point error" occurring during survey recording. In a third case, we take the three observations closest to the observed median and subtract the value of 10 000 from each. Figure 3 shows boxplots of relative errors $(\widehat{\theta}_r - \theta)/\theta$, with $r = 1, \ldots, R$, with respect to the population parameter θ where empty boxplots indicate the (relative) error distribution when data were not contaminated, coloured boxplots those under contamination. Both $\widehat{QSR}$ and $\widehat{\text{Palma}}$ are based on shares, and their sums in the denominator and numerator are strongly affected by changes in the tails; right-tail contamination in particular induces a pronounced upward bias. The two indicators behave similarly, reflecting their shared structural dependence on tail shares. $\widehat{P90/P10}$ is also strongly affected by contamination, especially in small samples; notably, its sensitivity decreases as n increases. None of them is sensitive to median contamination. $\widehat{G}$ shows moderate sensitivity, although more accentuated for right-tail contamination, while remaining comparatively stable under left-tail and median contamination. Clearly, $\widehat{QRI}$ is the most robust and reliable estimator: its distribution remains tightly concentrated around zero under all contamination scenarios and sample sizes, with no visible asymmetry between left- and right-tail perturbations, even for $n = 30$.

4 Conclusion

This study shows that the QRI combines the advantages of quantile-based inequality indicators, such as robustness, with those of the Gini coefficient, namely, its intuitive interpretation and the accounting for the entire distribution. In practice, the performance of inequality estimators obtained from survey data under complex sampling designs is of particular interest. A direct comparison of them, including other indicators based on quantile ratios or shares, like the popular P90/P10, QSR and Palma index, reveals that the QRI estimator consistently outperforms the other indices' estimators in terms of minimal bias, smallest mean squared error and bounded influence function in all considered situations. Indeed, the Gini estimator was only competitive (though still worse) for synthetic data generated from a smooth Log-Normal population under moderate skewness and without extreme values. This is in line with observations made so far in the literature. One may argue that in practice one often faces much larger samples. However, even for them, both the literature and our simulations found a serious underestimation of the Gini coefficient. Moreover, the present trend (fostered by the SDGs) in official statistics goes towards high levels of disaggregation, which requires the study of increasingly small samples.

ACKNOWLEDGEMENTS

Funder Si.Sc.: PNRR - Missione 4 Componente 2 Investimento 1.2 - D.D. n. 47/2025 PNRR - M4C2 INV.1.2, CUP: E93C25000610006; Funder St.Sp.: Grants PID2024-156871NB-I00 and PID2023-147127OB-I00 funded by MICIU/AEI /10.13039/501100011033, España and by FEDER, UE; Funder M.R.F.: European Union–NextGenerationEU, Mission 4, Component 2, in the framework of the GRINS–Growing Resilient, INclusive and Sustainable project (PNRR–M4C2–I1.3–PE00000018–CUP J33C22002910001).

Open access publishing facilitated by Universita degli Studi di Modena e Reggio Emilia, as part of the Wiley - CRUI-CARE agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data underlying this article were provided by Eurostat under licence. Data will be shared on request to the corresponding author with permission of Eurostat.

Notes

[1]The SILC sampling weights were ignored in this step as these are unrelated to income such that their inclusion hardly changed the Log-Normal parameter estimates.

REFERENCES

Aaberge, R. & Atkinson, A.B. 2013. The Median as Watershed. Discussion Papers.
Alfons, A. & Templ, M. (2012). Estimation of Social Exclusion Indicators From Complex Surveys: The R package laeken. *J. Stat. Softw.*, **54**, 1–25.
Alfons, A., Templ, M. & Filzmoser, P. (2013). Robust Estimation of Economic Indicators From Survey Samples based on Pareto tail modelling. *J. R. Stat. Soc. Ser. C*, **62**, 271–286.
Atkinson, A.B. 2007. Measuring Top Incomes: Methodological Issues. Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries, 18–42.
Atkinson, A.B., Piketty, T. & Saez, E. (2011). Top Incomes in the Long Run of History. *J. Econ. Literat.*, **49**(1), 3–71.
Brazauskas, V., Greselin, F. & Zitikis, R. (2024). Measuring Income Inequality via Percentile Relativities. *Qual. Quant.*, **58**, 4859–4896.
Cobham, A. & Sumner, A. 2013. Is It All About the Tails? The Palma Measure of Income Inequality. Center for Global Development Working Paper.
Cowell, F.A. (2011). *Measuring Inequality.* Oxford University Press.
Cowell, F.A. & Flachaire, E. (2007). Income Distribution and Inequality Measurement: The Problem of Extreme Values. *J. Econometr.*, **141**, 1044–1072.
Cowell, F.A. & Flachaire, E. (2024). Inequality Measurement and the Rich: Why Inequality Increased More Than We Thought. *Rev. Income Wealth*, **70**(2), 254–277.
Cowell, F.A. & Victoria-Feser, M.P. (1996). Robustness Properties of Inequality Measures. *Econometrica*, **64**, 77–101.
Davidson, R. & Flachaire, E. (2007). Asymptotic and Bootstrap Inference for Inequality and Poverty Measures. *J. Econometr.*, **141**, 141–166.
Davydov, Y. & Greselin, F. (2020). Comparisons Between Poorest and Richest to Measure Inequality. *Sociol. Methods Res.*, **49**, 526–561.
De Nicolò, S., Ferrante, M.R. & Pacei, S. (2024). Small-Sample Bias Correction of Inequality Estimators in Complex Surveys. *J. Off. Stat.*, **40**, 238–261.
Deltas, G. (2003). The Small-Sample Bias of the Gini Coefficient: Results and Implications for Empirical Research. *Rev. Econ. Stat.*, **85**, 226–234.
Deville, J.-C. (1999). Variance Estimation for Complex Statistics and Estimators: linearization and residual techniques. *Surv. Methodol.*, **25**, 193–204.
Gastwirth, J.L. (2014). Median-Based Measures of Inequality: Reassessing the Increase in Income Inequality in the US and Sweden. *Stat. J. IAOS*, **30**, 311–320.
Gastwirth, J.L. (2017). Is the Gini Index of Inequality Overly Sensitive to Changes in the Middle of the Income Distribution?. *Stat. Publ. Policy*, **4**, 1–11.
Gini, C. (1914). Sulla Misura Della Concentrazione e Della Variabilità dei Caratteri. *Atti del Reale Istituto veneto di scienze, lettere ed arti*, **73**, 1203–1248.

Giorgi, G.M. & Gigliarano, C. (2017). The Gini Concentration Index: A Review of the Inference literature. *J. Econ. Surv.*, **31**, 1130–1148.
Hampel, F.R., Ronchetti, E., Rousseeuw, P. & Stahel, W. (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons.
Hlasny, V. & Verme, P. (2018). Top Incomes and Inequality Measurement: A Comparative Analysis of Correction Methods Using the EU SILC data. *Econometrics*, **6**(2), 30.
Jenkins, S.P. (2017). Pareto Models, Top Incomes and Recent Trends in UK Income inequality. *Economica*, **84**(334), 261–289.
Jenkins, S.P. & Van Kerm, P. 2009. The Measurement of Economic Inequality. In *The oxford handbook of economic inequality*, Eds. Salverda, W., Nolan, B. & Smeeding, T.M., Oxford University Press, pp. 40–67.
Jokiel-Rokita, A. & Piątek, S. (2025). Nonparametric Estimators of Inequality Curves and Inequality Measures. *J. Stat. Plan. Inference*, **237**, 106251.
Jorda, V. & Niño-Zarazúa, M. (2019). Global Inequality: How Large Is the Effect of Top Incomes?. *World Develop.*, **123**, 104593.
Langel, M. & Tillé, Y. (2011). Statistical Inference for the Quintile Share Ratio. *J. Stat. Plan. Inference*, **141**, 2976–2985.
Langel, M. & Tillé, Y. (2013). Variance Estimation of the Gini Index: Revisiting a Result Several Times Published. *J. R. Stat. Soc. Ser. A*, **176**, 521–540.
Palma, J.G. 2006. Globalizing Inequality: “Centrifugal” and “Centripetal” Forces at Work. United Nations, Department of Economics and Social Affairs.
Palma, J.G. (2011). Homogeneous Middles vs. Heterogeneous Tails, and the End of the “Inverted-U”: It’s All About the Share of the Rich. *Develop. Change*, **42**, 87–153.
Pfeffermann, D. (1993). The Role of Sampling Weights When Modeling Survey Data. *Int. Stat. Rev.*, **61**, 317–337.
Prendergast, L.A. & Staudte, R.G. (2016). Quantile Versions of the Lorenz Curve. *Electron. J. Stat.*, **10**, 1896–1926.
Prendergast, L.A. & Staudte, R.G. (2018). A Simple and Effective Inequality Measure. *Am. Stat.*, **72**, 328–343.
Scarpa, S., Ferrante, M.R. & Sperlich, S. (2025). Inference for the Quantile Ratio Inequality Index in the Context of Survey Data. *J. Surv. Stat. Methodol.* https://doi.org/10.1093/jssam/smaf024
Scarpa, S. & Sperlich, S. 2026. Inequantiles: Quantile-Based Inequality Indicators for Complex Survey Data. R package version 0.1.0.
Sen, A. & Foster, J. (1997). *On Economic Inequality.* Oxford university press.
Van Kerm, P. 2007. Extreme Incomes and the Estimation of Poverty and Inequality Indicators From EU-SILC. In *Comparative EU statistics on income and living conditions: Issues and challenges*, Proceedings of the EU-SILC Conference, Helsinki, 6-8 November 2006, Eurostat, European Commission, pp. 83–94.
Zenga, M. (2007). Inequality Curve and Inequality Index Based on the Ratios Between Lower and Upper Arithmetic Means. *Stat. Appl.*, **V**, 3–7.

[*Received September 2025; accepted June 2026*]