A practical guide to proper estimation and inference of the Gini index by avoiding often encountered methodological pitfalls

Social Indicators Research

Full citation to this publication:

https://doi.org/10.1007/s11205-026-03831-x

Muñoz, J. F., Pavía, J.M., & Álvarez-Verdejo, E. (2026). A practical guide to proper estimation and inference of the Gini index by avoiding often encountered methodological pitfalls. *Social Indicators Research*. DOI: 10.1007/s11205-026-03831-x

***Juan F. Muñoz**, Department of Quantitative Methods for Economics and Business. University of Granada, https://orcid.org/0000-0001-7427-6630

Jose M. Pavía, Area of Quantitative Methods, Department of Applied Economics, Universitat de Valencia, https://orcid.org/0000-0002-0129-726X

Encarnación Álvarez, Department of Quantitative Methods for Economics and Business. University of Granada, https://orcid.org/0000-0002-0473-6037

***Corresponding author**

JEL Classification: C01, C13, C15, C83, D63.

Acknowledgements. This research is part of the project PID2022-136235NB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU, and has also been supported by Ministerio de Ciencia e Innovación [grant number PID2021-128228NB-I00] and the Generalitat Valenciana [grant number CIAICO/2023/031]. The authors wish to thank M. Hodkinson and E. Porritt for revising the English of the paper.

A practical guide to proper estimation and inference of the Gini index by avoiding often encountered methodological pitfalls

Abstract. The Gini index is the most widely-used measure of inequality. Unfortunately, its computation is subject to error. Researchers and practitioners often fall into common methodological pitfalls, leading to inaccurate estimates and inferences, and ultimately hindering efforts to reduce inequality and improve societal quality of life. This paper clarifies the challenges of non-parametric estimation of the Gini index more comprehensively than previous contributions, and offers robust methodological recommendations to ensure accurate estimates. Additionally, we reference a free, easy-to-use R package which, together with the clear methodological insights, enhances the real-world applicability of our findings. First, we investigate the impact of common methodological pitfalls on point estimates, providing a complete review for both infinite and finite populations. We then examine variance estimation and the performance of confidence intervals. Among other issues, the findings reveal that, when a popular regression-based variance estimator is used, the variance of the Gini index is seriously underestimated in distributions with high skewness and inequality, as often observed in real-world applications. Jackknife variance estimates and jackknife intervals, based on studentized quantiles, prove to be the most accurate approaches. The analysis employs variables with varying degrees of skewness and inequality (as both characteristics influence the potential for bias), thereby encompassing most of the situations found in empirical research.

1. Introduction

Despite wealth and income inequality being a constant throughout history, its role as a trigger for social and economic conflicts has drawn growing attention, particularly as inequality has escalated in recent decades in certain developing and OECD countries. (Filauro et al., 2025). A significant increase in income inequality directly affects poverty rates, social tensions, and economic development (Topuz, 2022), also playing a central role in shaping the population's quality of life and overall economic conditions (Banerjee and Tóth, 2025). Unequal access to education, healthcare, housing, or decent employment reduces people's chances of living healthy and fulfilling lives. Societies with greater inequality tend to have lower levels of life satisfaction, worse health outcomes, and higher rates of stress and insecurity—even among middle-income groups (OECD, 2017). These serious consequences have intensified the interest of governments and institutions in inequality, to the point that reducing it is now one of the 17 Sustainable Development Goals of the United Nations 2030 Agenda, reflecting the understanding that tackling inequality is not only a matter of fairness but also a necessary condition for improving the overall quality of life of the population.

Many measures of inequality have been proposed in the literature, with the Gini index being the most well-known and widely-used indicator—likely due to its ease of interpretation and its unique connection to Lorenz curves, which support a robust welfare interpretation and allow for graphical representations. National and regional statistical agencies (Eurostat, NBER, ONS), as well as global institutions (World Bank, IMF, UN), employ it to quantify inequality and make comparisons across time, countries, and subgroups of individuals (grouped, for instance, by gender, age, or race). Given its relevance, the Gini index has long attracted the attention of numerous researchers and continues to be an important and active area of study.

The Gini index, like many other inequality measures, is defined for non-negative values (Langel and Tillé, 2013; Giorgi and Gigliarano, 2017). Under this assumption, it ranges from 0 (perfect equality) to 1 (perfect inequality). Values outside the [0,1] interval may occur when the sample contains negative observations. Several authors have proposed transformations or normalizations to address this issue and thus make the Gini index meaningful and comparable even when data include negative values (e.g., Raffinetti et al., 2015; Lee and Suh, 2025). Additional challenges associated with the Gini index, such as the use of grouped data (Van Ourt and Clarke, 2011) or the presence of censored observations (Kattumannil et al., 2021), have also been examined in the literature. Moreover, Cowell (2011) and Ibragimova and Frants (2021) provide comparative analyses of the Gini index with alternative measures of inequality.

Giorgi and Gigliarano (2017) provide a review of the statistical properties of the Gini index and present general expressions under both discrete and continuous variables, and for various continuous probabilistic distributions. In the case of discrete variables, a bias correction is necessary to ensure an unbiased estimation (Deltas, 2003; Davidson, 2009). Bias correction is important since a large bias can lead to incorrect inferences and conclusions (Särndal et al., 2003). However, since some expressions of the Gini index are already implicitly bias-corrected, the bias correction is sometimes omitted in the other formulae, leading to biased estimates when applied to small samples. Indeed, Giorgi and Gigliarano (2017) overlooked the bias correction issue in their comprehensive review. This paper extends the existing literature by (i) providing a comprehensive overview of non-parametric formulae and techniques for both estimating the Gini index (Section 2) and conducting statistical inference (Section 4), and (ii) illustrating the relative importance of biases and the performance of estimators through Monte Carlo simulations (Sections 3-4).

Methodologically, the first contribution of this paper is to provide an updated and comprehensive review of Gini index formulations for discrete variables and simple random samples from infinite populations, including bias-corrected versions for each expression. As a second contribution, it also includes various point estimators of the Gini index for complex survey data derived from finite populations—an area not covered by Giorgi and Gigliarano (2017). The variance estimation and the construction of confidence intervals are important tools for determining the margin of error of an estimator, an issue sometimes overlooked in Gini index inferences. Hence, as a third contribution, this paper describes and examines several popular methods used to make reliable inferences for the Gini index.

Empirically, the first Monte Carlo study compares the relative biases in the point estimate of the Gini index arising from (i) neglecting small-sample bias and (ii) using the naïve rather than the smooth mid-interval empirical distribution function (Section 3). We find that both of these biases generally decline with sample size and the level of inequality, but increase with skewness. Additionally, we reveal that the bias resulting from using the naïve instead of the smooth distribution function increases with the share of ties, i.e., the proportion of observations with identical values. The second Monte Carlo study evaluates (i) the relative bias in estimating the variance of the Gini index (Subsection 4.3) using four methods—OLS, bootstrap, jackknife, and linearization—and (ii) the empirical coverage of confidence intervals derived from these four methods, along with two additional ones (Subsection 4.4). The jackknife generally outperforms the other methods. We also find that OLS tends to underestimate the variance of the Gini index in distributions with high levels of inequality and skewness, while it overestimates the variance when either inequality or skewness are lower.

Many asymptotically equivalent expressions for estimating the Gini index have been proposed in the literature, with apparently minimal differences between the bias-corrected and non-bias-corrected versions of each estimator. Therefore, a broad empirical assessment is essential to ascertain how the most commonly encountered methodological pitfalls (MPs) affect the quality of point, variance, and confidence interval estimates of the Gini index. Using Monte Carlo simulations, we empirically assess the impact of MPs across various practical scenarios, combining factors such as sample size, value repetition, skewness, and actual inequality—features of both the data and the underlying variable that strongly influence point estimates of the Gini index (Muñoz et al., 2025). In addition to studying the effect of sample size and the percentage of repeated values on the sample, we also investigate their interaction with skewness and actual inequality. This is relevant because the literature suggests an asymptotic equivalence between the jackknife and linearization methods for variance estimation, and also indicates that a regression-based variance estimator for the Gini index can yield substantial overestimations (Berger, 2008; Langel and Tillé, 2013). We demonstrate that the regression estimator may also be prone to underestimation with highly skewed variables when the inequality of the variable increases and that, in the same scenarios, the jackknife and linearization methods are not equivalent. These findings represent additional contributions that can help ensure more reliable estimates and inferences for the Gini index, thereby reducing the risk of drawing incorrect conclusions.

This paper is therefore useful in guiding applied researchers in their choice of Gini estimators for both point estimation and inference. While the paper presents some novel findings—such as the underestimation of variance by the OLS estimator in highly skewed and unequal distributions—its primary strength lies in (i) consolidating all previous results into a single, rigorously documented study using Monte Carlo simulations, and (ii) extending the analysis to a broader range of scenarios in terms of sample size, Gini values, and skewness.

Table 1 presents a graphical summary of the scenarios examined in the literature, categorized by their focus on bias (B), use of the naïve distribution function (F), and variance estimation (V). Shaded cells denote scenarios for which at least one study exists, with the number of studies reported in each cell. Each scenario is defined along three dimensions: the value of the Gini index (G); the sample size (n); and the skewness ($\gamma_{(G)}$) of the underlying distribution with Gini index G. Each dimension is categorized into three levels—Low (L), Medium (M), and High (H)— based on ranges commonly used in the literature for defining simulation scenarios; details are provided in the table note. Table SM1 in the supplementary material presents a detailed literature review, including a summary of the scenarios previously examined. In addition to analysing the 51 scenarios corresponding to the shaded cells, our study also covers the remaining cases, completing the full set of 81 possible scenarios.

Table 1. Number of studies identified in the literature review—categorized by their focus on bias (B), use of the naïve distribution function (F), and variance estimation (V)—by scenario.

Skewness: $\gamma_{(G)}$	Sample Size: n	B	F	V	B	F	V	B	F	V
H	H	1		1	1			1		
	M	3	1	2	2	1		2	1	
	L	4	1	5	4	1	2	2	1	
M	H									
	M	1	1		2	1	1	1	1	
	L	1	1		5	1	5	4	1	4
L	H							1		1
	M	2	1	1	1	1		3	1	2
	L	1	1		3	1	2	6	1	8
			L			M			H	
						Gini Index: G				

Note: Shaded cells indicate scenarios for which at least one study has been conducted, with the number of studies shown in each cell. The Gini index (G) is classified as Low (L) when $G \leq 0.25$, Medium (M) when $0.25 < G \leq 0.45$, and High (H) when $G > 0.45$. Sample size (n) is classified as Low (L) when $n \leq 100$, Medium (M) when $100 < n \leq 500$, and High (H) when $n > 500$. Given a value G, the skewness of the distribution under study ($\gamma_{(G)}$) is classified as Low (L) when $\gamma_{(G)} \leq 0.8\gamma_{L(G)}$, Medium (M) when $0.8\gamma_{L(G)} < \gamma_{(G)} \leq 1.2\gamma_{L(G)}$, and High (H) when $\gamma_{(G)} > 1.2\gamma_{L(G)}$, where $\gamma_{L(G)}$ is the skewness coefficient of the logNormal distribution with Gini index G. For real datasets, this classification is applied using the skewness coefficient of the dataset and the empirical average skewness of the logNormal distribution based on 10000 samples of the same size.

To ensure this updated review of the Gini index is accessible and practical for potential users (scholars, researchers, analysts), we refer the reader to the package *giniVarCI* (Muñoz et al., 2024). This R package implements all the methodological approaches discussed in this paper through easy-to-use and efficient (C++) functions, allowing fast computation even with large datasets. Moreover, unlike other R packages that also incorporate tools for inference—such as *laeken* (Alfons and Templ, 2013) and *DescTools* (Signorell, 2023)—*giniVarCI* not only incorporates bootstrap methods, but also jackknife and linearization, which have been shown to perform well in estimating the variance of the Gini index (Berger, 2008; Langel and Tillé, 2013). It is also important to note that most existing R packages do not offer small-sample bias correction. With the help of this tool, practitioners can easily compute any estimator and/or confidence interval for the Gini index in inequality studies, ensuring the practical applicability of this study to real-world contexts.

The rest of this paper is organized as follows. Section 2 offers an updated review of the point estimation methods for the Gini index under infinite and finite populations. We present both the bias-corrected and non-bias-corrected versions of each existing expression of the Gini index, many of them empirically equivalent. Section 3 outlines the most important MPs encountered in point estimation and examines their impact using Monte Carlo simulation. This analysis examines the effect of various factors (levels of inequality and skewness, sample size, and proportion of repeated sample values) on Gini index estimates. Section 4 focuses on assessing variance estimators and confidence intervals. Specifically, after describing a commonly-used erroneous method for estimating variances and proposing alternative methods, we evaluate the different variance estimators in terms of relative bias and assess confidence intervals regarding empirical coverage. These analyses identify the jackknife approach as the best for measuring uncertainties. Conclusions are summarized in Section 5. A file with supplementary material completes the paper. The R code required to replicate all the results and figures is available at the OSF repository: https://osf.io/4jnuz/

2. An updated review of the Gini index estimators

Consider the problem of analysing the inequality of a non-negative continuous variable X with cumulative distribution function $F_X(x) = P(X \leq \mathrm{x})$, probability density function $f(x)$ and expected value

$$\mu_X = E[X] = \int_0^{+\infty} x\, f(x) dx = \int_0^{+\infty} x\, dF_X(x).$$

A popular formula for the Gini index (Qin et al., 2010) is:

$$G = \frac{1}{2\mu_X} \int_0^{+\infty} \int_0^{+\infty} |x-y|\, dF_X(x) dF_X(y). \qquad (1)$$

Alternatively, Lerman and Yitzhaki (1984) showed that the Gini index can be computed as:

$$G = \frac{2}{\mu_X} cov[X, F_X(x)]. \qquad (2)$$

For continuous random variables, the Gini index (G) has alternative formulae that yield the same result (Giorgi and Gigliarano, 2017). Therefore, no theoretical conflict appears in its computation. An issue arises because both the distribution of the variable and the value of G are usually unknown and estimators must be used. Common practice for conducting inequality analysis involves estimating G using a sample S (of size n), drawn from the population under study. Populations can be classified as infinite or finite, and the statistical techniques applied during the estimation phase depend on the type of population.

2.1. Infinite populations

Classical statistical theory is based on infinite populations. In this context, $\{X_i: i \in S\}$ denotes a sequence of n non-negative random variables with the same distribution as X, while $\{x_i: i \in S\}$ represents the observations from a sample of n individuals independently selected from the infinite population. Given the sample, the standard procedure to estimate G has been to apply the plug-in principle in a theoretical expression for it. This approach has given rise to numerous estimators in the literature, leading to controversy due to the potential for differing results and the biases that can arise, especially in small samples. For infinite populations, the Gini index can also be estimated by fitting the sample to a parametric continuous probability distribution and then plugging the estimated parameters into the theoretical formula (Giorgi and Gigliarano, 2017). However, this approach can introduce significant biases when the chosen probabilistic model is not appropriate. Semiparametric approaches are sometimes used as alternatives (Fontanari et al., 2018).

Existing non-parametric estimators of the Gini index can be classified into two categories, referred to in this paper as bias-corrected ($\hat{G}_k^{bc}$) and non-bias-corrected ($\hat{G}_k$) estimators. Table 2 presents the 10 most common formulations, $k \in \{1, \dots, 10\}$. The classification provided in Table 2 serves as a useful reference for researchers and practitioners in inequality analysis, helping them to avoid some of the serious consequences discussed in this paper. For large sample sizes, the non-bias-corrected estimator performs similarly to the bias-corrected version. However, for small sample sizes, the non-bias-corrected estimator may exhibit non-negligible bias, so the bias-corrected estimator is strongly recommended to prevent erroneous conclusions. The consequences of the incorrect use of non-bias-corrected estimators on point estimates of the Gini index is the first methodological pitfall (MP) analysed (Section 3). $\hat{G}_k$ and $\hat{G}_k^{bc}$ estimators are related by the equation (3):

$$\hat{G}_k^{bc} = \frac{n}{n-1}\hat{G}_k. \tag{3}$$

To fully understand the expressions included in Table 2, we need to introduce some notation. The mean of the sample observations is defined as $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$. The values $x_{(i)}$, with $i \in \{1, \dots, n\}$, denote the ordered values of the sample observations in non-decreasing order, while $p_0 = q_0 = 0$ and $p_i = i/n$ and

$$q_i = \frac{\sum_{j=1}^{i} x_{(j)}}{\sum_{i=1}^{n} x_i}, \tag{4}$$

with $i = 1, \dots, n$, representing the cumulative proportions of observations and mass of the variable necessary to plot the Lorenz curve (Lorenz, 1905), which justifies the known relationship between the Gini index and the Lorenz curve.

Estimators $\hat{G}_9$ and $\hat{G}_9^{bc}$ are computed in terms of the smooth (or midpoint) empirical distribution function (Berger 2008):

$$\hat{F}_n^*(t) = \frac{1}{n}\sum_{i=1}^{n}[\delta(x_i < t) + 0.5\delta(x_i = t)], \tag{5}$$

where $\delta(\cdot)$ is the indicator function that takes the value 1 if its argument is true and 0 otherwise.

Table 2. Mathematical expressions for the non-bias-corrected ($\hat{G}_k$) and bias-corrected ($\hat{G}_k^{bc}$) estimators of the Gini index ($k = 1, \dots, 10$), based on a sample of size n drawn from an infinite population.

Non-bias-corrected formulation	Bias-corrected formulation
$\hat{G}_1 = \frac{1}{2\bar{x}n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i - x_j\rvert$	$\hat{G}_1^{bc} = \frac{1}{2\bar{x}n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i - x_j\rvert$
$\hat{G}_2 = \frac{(n-1)\sum_{i=1}^{n-1}(p_i - q_i)}{n\sum_{i=1}^{n-1}p_i}$	$\hat{G}_2^{bc} = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1}p_i}$
$\hat{G}_3 = \frac{n-1}{n} - \frac{2}{n}\sum_{i=1}^{n-1}q_i$	$\hat{G}_3^{bc} = 1 - \frac{2}{n-1}\sum_{i=1}^{n-1}q_i$
$\hat{G}_4 = 1 - \sum_{i=0}^{n-1}(q_{i+1} + q_i)(p_{i+1} - p_i)$	$\hat{G}_4^{bc} = \frac{n}{n-1}\left[1 - \sum_{i=0}^{n-1}(q_{i+1} + q_i)(p_{i+1} - p_i)\right]$
$\hat{G}_5 = \frac{2}{\bar{x}n^2}\sum_{i=1}^{n}ix_{(i)} - \frac{n+1}{n}$	$\hat{G}_5^{bc} = \frac{2}{\bar{x}n(n-1)}\sum_{i=1}^{n}ix_{(i)} - \frac{n+1}{n-1}$
$\hat{G}_6 = 2cov\left(\frac{i}{n}, \frac{x_{(i)}}{\bar{x}}\right) = \frac{2}{\bar{x}n}cov\left(i, x_{(i)}\right)$	$\hat{G}_6^{bc} = 2cov\left(\frac{i}{n-1}, \frac{x_{(i)}}{\bar{x}}\right) = \frac{2}{\bar{x}(n-1)}cov\left(i, x_{(i)}\right)$
$\hat{G}_7 = 1 - \frac{1}{\bar{x}n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}min\left(x_i, x_j\right)$	$\hat{G}_7^{bc} = 1 - \frac{1}{\bar{x}n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1,\\ j\neq i}}^{n}min\left(x_i, x_j\right)$
$\hat{G}_8 = \frac{n-1}{2\bar{x}n}\binom{n}{2}^{-1}\sum_{1\leq i<j\leq n}\lvert x_{\mathrm{i}} - x_{\mathrm{j}}\rvert$	$\hat{G}_8^{bc} = \frac{1}{2\bar{x}}\binom{n}{2}^{-1}\sum_{1\leq i<j\leq n}\lvert x_{\mathrm{i}} - x_{\mathrm{j}}\rvert$
$\hat{G}_9 = \frac{2}{\bar{x}n}\sum_{i=1}^{n}x_i\hat{F}_n^*(x_i) - 1$	$\hat{G}_9^{bc} = \frac{2}{\bar{x}(n-1)}\sum_{i=1}^{n}x_i\hat{F}_n^*(x_i) - \frac{n}{n-1}$
$\hat{G}_{10} = \frac{1}{\bar{x}n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i - x_j\rvert\lvert\hat{F}_n(x_i) - \hat{F}_n(x_j)\rvert$	$\hat{G}_{10}^{bc} = \frac{1}{\bar{x}n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert x_i - x_j\rvert\lvert\hat{F}_n(x_i) - \hat{F}_n(x_j)\rvert$

Note: x_i and x_j (for $i, j = 1, \dots, n$) denote sample observations, with $x_{(i)}$ representing the corresponding ordered values in non-decreasing order. $\bar{x} = n^{-1}\sum_{i=1}^{n}x_i$; $p_i = i/n$; $q_i = \sum_{j=1}^{i}x_{(j)} / \sum_{i=1}^{n}x_i$ (for $i = 1, \dots, n$); and $p_0 = q_0 = 0$, $\hat{F}_n^*(t)$ and $\hat{F}_n(t)$ are defined in equations (4), (5) and (6), respectively.

The second MP described in Section 3 relates to the consequences of using the naïve cumulative distribution function given by equation (6):

$$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n}\delta(x_i \leq t), \tag{6}$$

instead of $\hat{F}_n^*(t)$. This MP—using (6) instead of (5) in estimators $\hat{G}_9$ (see its derivation from $\hat{G}_{9F}$ in Section 3) or $\hat{G}_9^{bc}$—may introduce serious biases during the estimation step, especially for small sample sizes or variables with a large proportion of repeated values.

In practical terms, it is also important for researchers and practitioners to have a tool for easily calculating the various formulae of the Gini index described in Table 2. We refer the reader to the R package *giniVarCI* (Muñoz et al., 2024) to compute the estimators described in Table 2.

2.2. Finite populations

In practice, the Gini index is usually estimated from social surveys obtained through complex sampling designs from a finite population U with N individuals. Survey sampling theory (Särndal et al., 2003) must be used in this context for estimating population parameters, as the assumption of independence is not guaranteed due to the specific features of survey sampling. Survey weights must be applied during the estimation stage to ensure valid results and conclusions. These weights are generally defined as $w_i = \pi_i^{-1}$, where $\pi_i = P(i \in S)$ is the inclusion probability of the ith individual in the sample.

For finite populations, the aim is to estimate G at the population level, which can be defined by any of the formulations described in Table 2 and is assumed to be unknown. For instance, the population Gini index can be expressed as:

$$G_N = \frac{2}{\bar{X}N^2}\sum_{i=1}^{N} i x_{(i)} - \frac{N+1}{N},$$

where $\bar{X} = N^{-1}\sum_{i=1}^{N} x_i$ is the population mean. For samples derived from a finite population, some non-bias-corrected formulations for estimating G (Lerman and Yitzhaki, 1989; Berger, 2008; Langel and Tillé, 2013), all of them providing the same output, are:

$$\hat{G}_{w1} = \frac{1}{2\bar{x}_w\hat{N}^2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \left|x_i - x_j\right|;$$

$$\hat{G}_{w2} = \frac{2\sum_{i=1}^{n} w_{(i)}\,\hat{N}_{(i)}x_{(i)} - \sum_{i=1}^{n} w_i x_i^2}{\bar{x}_w\hat{N}^2} - 1;$$

$$\hat{G}_{w3} = \frac{2}{\bar{x}_w\hat{N}}\sum_{i=1}^{n} w_i x_i \hat{F}_w^*(x_i) - 1;$$

$$\hat{G}_{w4} = \frac{2}{\bar{x}_w\hat{N}}\sum_{i=1}^{n} w_{(i)}\left[x_{(i)} - \bar{x}_w\right]\left[\hat{F}_w^{LY}\left(x_{(i)}\right) - \bar{F}_w^{LY}\right],$$

where $\widehat{N} = \sum_{i=1}^{n} w_i$, $\bar{x}_w = \widehat{N}^{-1} \sum_{i=1}^{n} w_i\, x_i$ is the weighted sample mean, $w_{(i)}$ are the survey weights w_i sorted in increasing order of the sample values x_i, $\widehat{N}_{(i)} = \sum_{j=1}^{i} w_{(j)}$ and the weighted smooth (midpoint) distribution function is given by:

$$\widehat{F}_w^*(t) = \frac{1}{\widehat{N}} \sum_{i=1}^{n} w_i [\delta(x_i < t) + 0.5\delta(x_i = t)],$$

and $\widehat{F}_w^{LY}(x_{(i)}) = \widehat{N}^{-1}(\widehat{N}_{(i-1)} + w_{(i)}/2)$, where $\widehat{F}_w^{LY}$ is an alternative smooth distribution function defined by Lerman and Yitzhaki (1989), with $\bar{F}_w^{LY} = \widehat{N}^{-1} \sum_{i=1}^{n} w_{(i)} \widehat{F}_w^{LY}(x_{(i)})$. The drawback of the above estimators is that they are all are non-bias-corrected. Berger and Gedik-Balay (2020) have recently proposed the bias-corrected estimator

$$\widehat{G}_{w5}^{bc} = 1 - \frac{\bar{z}_w}{\bar{x}_w},$$

where $\bar{z}_w = \widehat{N}^{-1} \sum_{i=1}^{n} w_i z_i$ and

$$z_i = \frac{1}{\widehat{N} - w_i} \sum_{\substack{j=1 \\ j \neq i}}^{n} \min(x_i, x_j).$$

We again refer the reader to the R package *giniVarCI* (Muñoz et al., 2024) to compute the estimators described in this subsection.

3. Impact of methodological pitfalls on point estimates of the Gini index

As highlighted in Section 2, there is extensive literature on the computation of point estimates of the Gini index. However, many of these point estimators are plagued by methodological pitfalls (MPs), leading to significant biases and misleading conclusions. The extent and magnitude of the biases are not universal, as they vary depending on the estimator used as well as on the specific properties of the underlying variable's distribution and the sample data. In this section, we use Monte Carlo simulations to assess the impact of skewness, inequality, sample size, and repeated values on the bias of Gini index estimates when an MP is present.

Muñoz et al. (2025) demonstrate that highly skewed distributions can result in large biases, particularly when combined with high levels of inequality, with sample size being a key factor influencing the magnitude of the bias. Our study additionally shows that the Gini index formulae based on the naïve cumulative distribution function are also affected by the proportion of repeated values in the population (Subsection 3.3), with substantial biases emerging in some scenarios, even with large sample sizes.

Without imposing specific conditions on the collected data, Subsection 3.2 examines the effect of using the non-bias-corrected estimator ($\hat{G}_k$) and of using $\hat{G}_{9F}$, in comparison with using a bias-corrected estimator ($\hat{G}_k^{bc}$), $k = 1, \dots, 10$. As a new contribution, Subsection 3.3 evaluates the use of the naïve cumulative distribution function in cases with a large proportion of repeated data. Subsection 3.1 provides a detailed description of the simulated scenarios.

Muñoz et al. (2025) also investigate the impact of skewness, inequality and sample size on the bias when a non-bias-corrected estimator is used. Our analysis in Subsection 3.2 differs from the analysis conducted by Muñoz et al. (2025) in two key aspects: the number of sample sizes considered, and the reference value used for comparison. Muñoz et al. (2025) examine only two sample sizes (50, small; and 500, medium) and assess biases relative to the expected value of the Gini index, rather than the true Gini index corresponding to the simulated scenario.

3.1. Description of simulated scenarios and measure of bias

In Subsection 3.2, to better understand the bias associated with Gini index estimation when bias correction is ignored, we examine variables with varying degrees of skewness and inequality, since these features have a direct impact on estimator performance. We study their interaction with small, $n \in \{10, 20, \dots, 100\}$, medium, $n \in \{100, 200, \dots, 500\}$, and large, $n \in \{1000, 4000, 7000, 10000\}$, sample sizes. Pareto, logNormal, and Gamma distributions represent high, medium, and low skewness, respectively, while three Gini values ($G = 0.1, 0.3, 0.5$) cover low, medium, and high inequality. The parameters of each distribution (Pareto, logNormal, and Gamma) are selected to ensure the desired levels of inequality. This results in 162 scenarios, formed by exhaustively combining the three distributions, three inequality levels, and 18 sample sizes. The chosen ranges are consistent with those commonly used in the literature (see, e.g., Muñoz et al., 2025) and cover all the scenarios considered in Table 1, reflecting typical empirical applications and allowing for a thorough evaluation of estimator performance.

In Subsection 3.3, to empirically assess the effect of the proportion of repeated values on bias when using the naïve cumulative distribution, we form scenarios with samples drawn from populations with varying proportions of repeated values: $p = 0\%$, $p = 10\%$ and $p = 20\%$. We first generate a finite population of size $N = 10^5$ ($x_{\gamma i}$, $i = 1, \dots, N$) from a Gamma distribution (low skewness) with a low degree of inequality ($G = 0.1$) and subsequently modify its simulated values, creating derived populations with the desired proportions of repeated values. The Gamma distribution with $G = 0.1$ is chosen to minimize the confounding effects of skewness and inequality on the observed biases, allowing us to focus solely on the impact of the proportion of repeated values. As shown in Subsection 3.2, bias appears even when using bias-corrected estimators if the populations exhibit high skewness and inequality.

The original simulated population has $p = 0\%$. To obtain populations with $p = 10\%$ and $p = 20\%$ repeated values, we sort $x_{\gamma i}$ in increasing order to obtain $x_{\gamma(i)}$ and replace the central 10% and 20% of the values with the median of $x_{\gamma i}$, respectively. For each value of p, the Gini index of the generated finite populations remains close to the target index ($G = 0.1$). Samples are drawn from these populations using simple random sampling with replacement, allowing us to apply formulae for infinite populations (since N is large and samples are selected with replacement).

In both subsections, we assess bias in estimating the Gini index using relative bias, defined by

$$RB_G = 100 \times \frac{E[\hat{G}] - G}{G}. \quad (7)$$

In equation (7), $E[\hat{G}] = R^{-1} \sum_{r=1}^{R} \hat{G}(r)$ denotes the expected value of a given estimator $\hat{G}$, where $R = 10^4$ is the number of replications (simulated samples) and $\hat{G}(r)$ represents the estimate $\hat{G}$ from the rth simulated sample. Note that biases can be considered negligible when the absolute values of RB_G are smaller than 2%. For simplicity, we also classify non-negligible relative biases as mild ($2\% \leq RB_G < 5\%$), moderate ($5\% \leq RB_G < 10\%$), and severe ($RB_G \geq 10\%$).

The Monte Carlo simulations conducted in this paper have been performed using the statistical software R and are fully reproducible. The R source codes needed to replicate all results and figures presented in the paper are available at the OSF repository: https://osf.io/4jnuz/

3.2. First methodological pitfall: overlooking the bias correction in small sample sizes

The debate over the impacts of using a non-bias-corrected formula for estimating the Gini index is well-documented in the literature (Deltas, 2003; Davidson, 2009). Bias correction is recommended for small sample sizes, as the difference between bias-corrected and non-bias-corrected estimators diminishes with increasing sample size.

Figures 1 to 3 compare the biases of the non-bias-corrected estimators $\hat{G}_k$ with those obtained using the bias-corrected estimators $\hat{G}_k^{bc}$, $k = 1, \ldots, 10$. We consider small and medium sample sizes. Results for large sample sizes are available in Section SM3 of the supplementary material. Scenarios in the figures are grouped by level of skewness. The graphical representations also include the $\hat{G}_{9F}$ estimates, a variant of the $\hat{G}_9$ estimator where the midpoint cumulative distribution function is replaced by the naïve cumulative distribution function, as described in Section 2 and detailed in equation (8) (Subsection 3.3). The $\hat{G}_{9F}$ estimator is included in Figures 1 to 3 as it has been referenced in various studies on the Gini index.

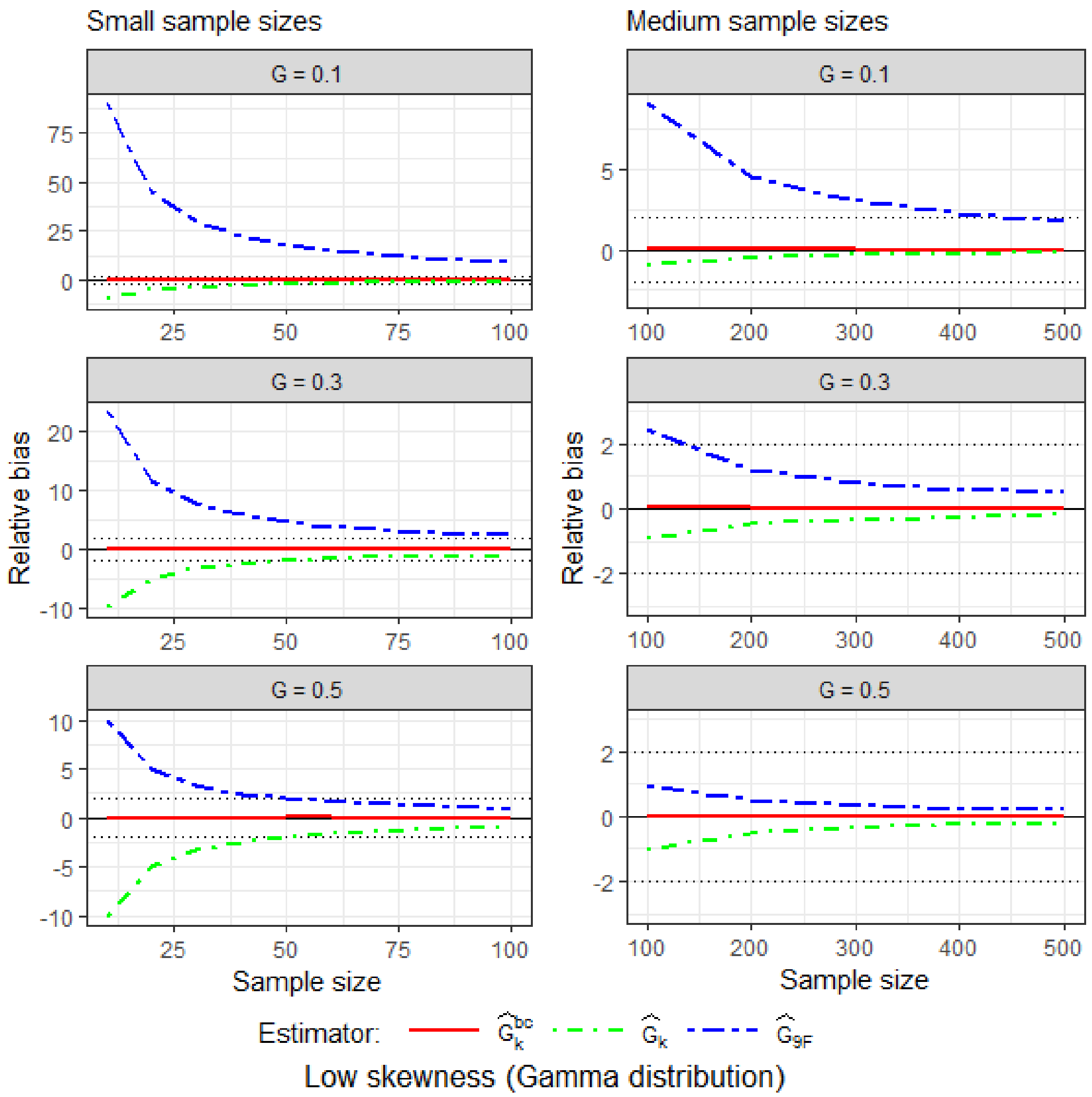

Figure 1. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, $k = 1, \dots, 10$ (Table 2), and $\hat{G}_{9F}$ (equation (8)). Samples, with sizes ranging from 10 to 500, are drawn from a Gamma distribution (low skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_G = 0$. Dashed lines at $RB_G \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_G| \leq 2$.

For the less skewed distributions (Figure 1), the relative biases (RB_G) of the bias-corrected estimator ($\hat{G}_k^{bc}$) are approximately zero, regardless of the degree of inequality or the sample size. In contrast, the non-bias-corrected estimators ($\hat{G}_k$ and $\hat{G}_{9F}$) exhibit non-negligible biases in these Gamma scenarios. Specifically, $\hat{G}_k$ systematically underestimates the true values for small sample sizes, showing moderate biases with medium and high inequality, whereas the $\hat{G}_{9F}$ estimator tends to overestimate actual inequality. In short, even with mild skewness, the biases of $\hat{G}_k$ and $\hat{G}_{9F}$ can be significant, particularly when the sample size is small.

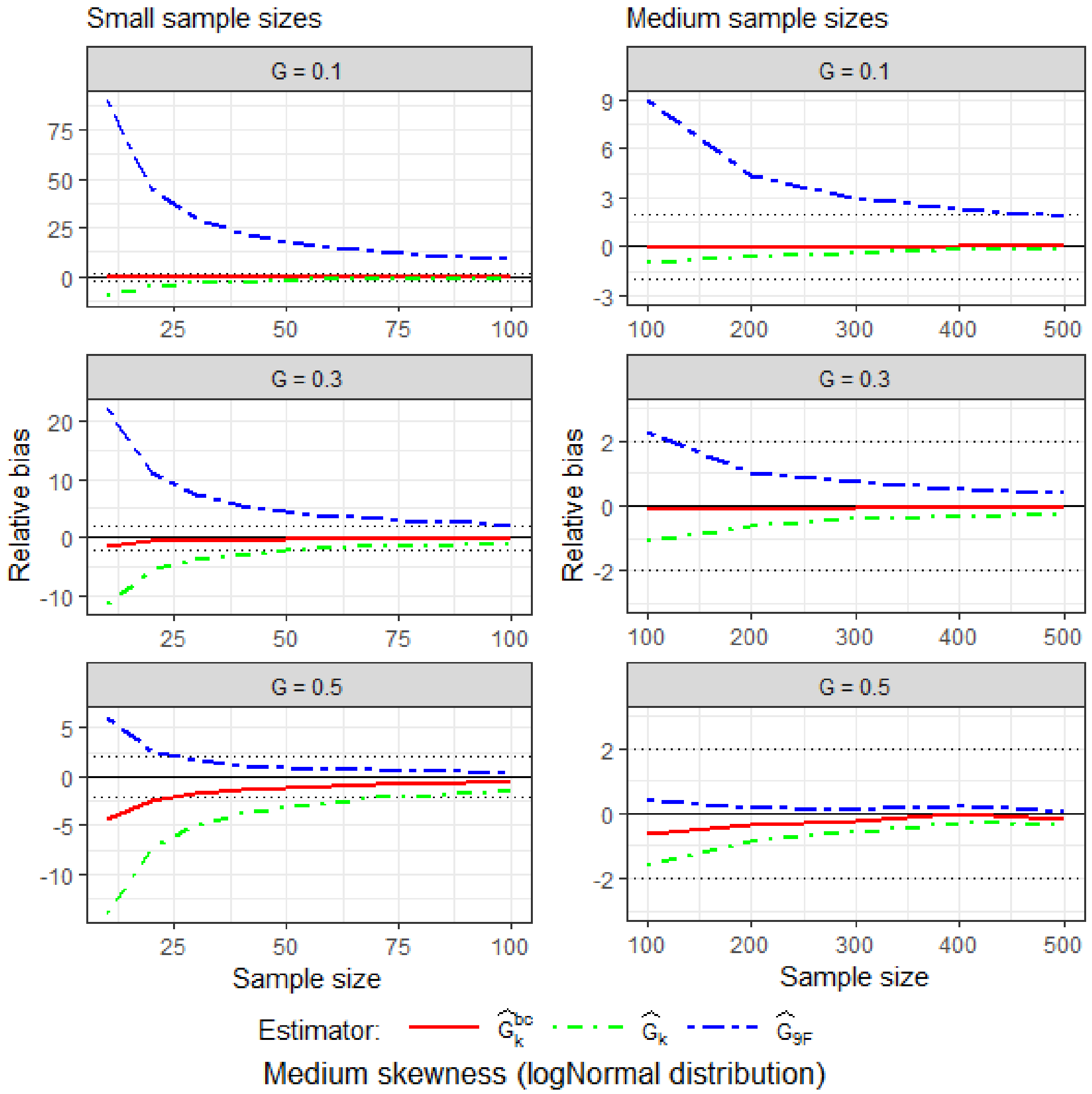

Figure 2. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, $k = 1, \ldots, 10$ (Table 2), and $\hat{G}_{9F}$ (equation (8)). Samples, with sizes ranging from 10 to 500, are drawn from a logNormal distribution (medium skewness), with standard deviation parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_G = 0$. Dashed lines at $RB_G \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_G| \leq 2$.

Figure 2 depicts the results obtained for the logNormal scenarios, where the underlying distributions exhibit medium skewness. Compared to the low skewness scenarios, the key issue here is that bias can occur with all the estimators, including the bias-corrected estimator ($\hat{G}_k^{bc}$). Although severe biases are not observed with the bias-corrected estimator, moderate biases appear with this estimator when $n \leq 25$ and $G = 0.5$. For estimators $\hat{G}_k$ and $\hat{G}_{9F}$, similar patterns to those seen in the Gamma scenarios are evident, though larger sample sizes are needed to reach the area of non-negligible biases compared to the low skewness scenarios.

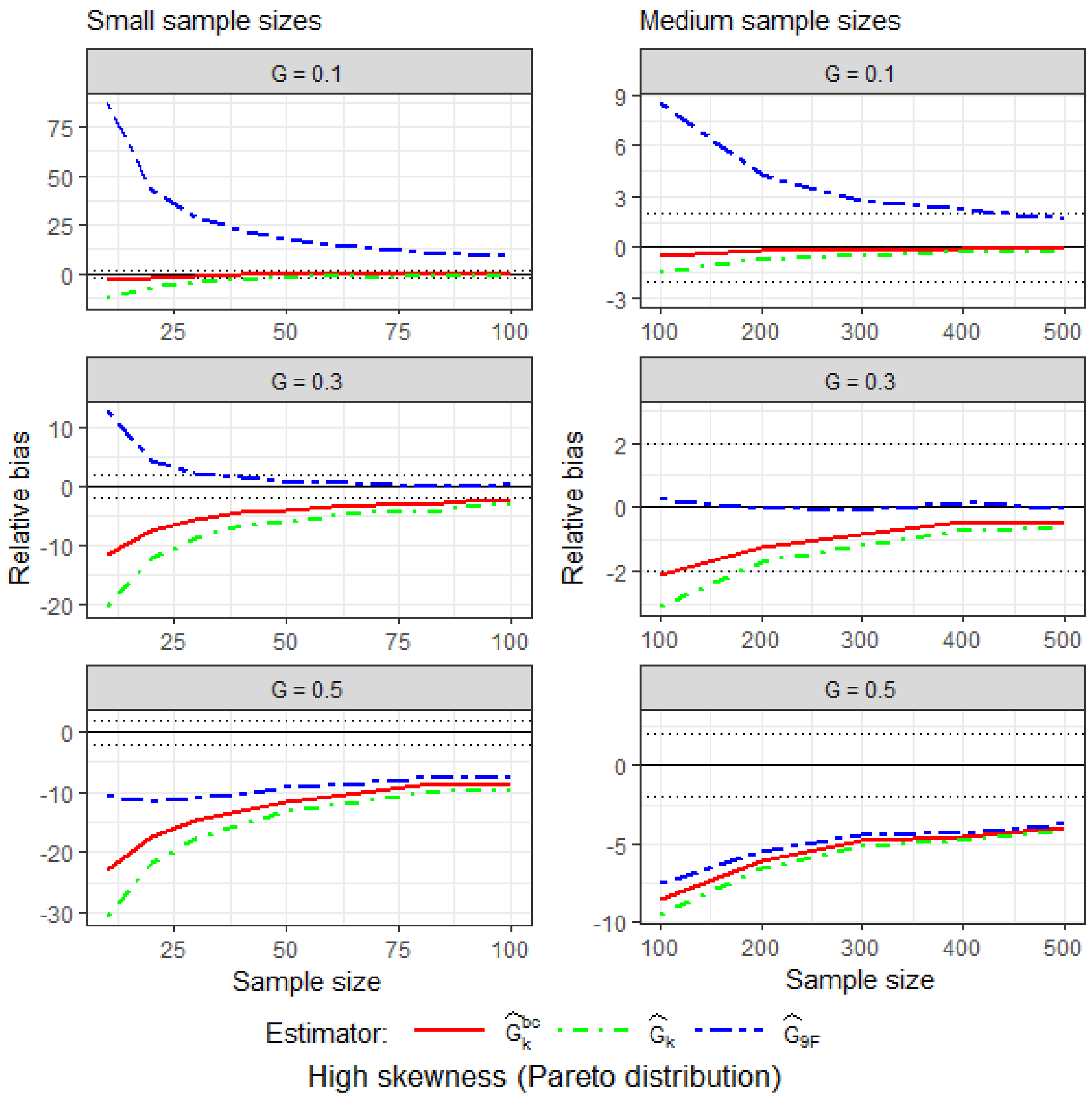

Figure 3. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, $k = 1, \dots, 10$ (Table 2), and $\hat{G}_{9F}$ (equation (8)). Samples, with sizes ranging from 10 to 500, are drawn from a Pareto distribution (high skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_G = 0$. Dashed lines at $RB_G \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_G| \leq 2$.

Figure 3 displays the relative biases for the highly skewed variables generated from Pareto distributions. The results are in line with those attained in the logNormal scenarios (Figure 2), but are significantly more pronounced. In the case of low inequality ($G = 0.1$), negligible biases are again observed for the bias-corrected estimator $\hat{G}_k^{bc}$, regardless of the sample size. For small sample sizes, moderate underestimates are recorded for the $\hat{G}_k$ estimator and severe overestimates are observed for the $\hat{G}_{9F}$ estimator. Major changes begin to appear with medium inequality ($G = 0.3$) and are consolidated with high inequality ($G = 0.5$). When $G = 0.3$, non-negligible biases are present for both $\hat{G}_k^{bc}$ and $\hat{G}_k$ when $n \leq 100$, whereas negligible biases are recorded for $\hat{G}_{9F}$ once $n > 30$.

In the Pareto scenarios (Figure 3), the most significant issues arise when $G = 0.5$. Biases due to underestimation are recorded for all small and medium sample sizes analysed. They become negligible for large sample sizes ($n \geq 4000$; see Section SM3 in the supplementary material). In contrast to other scenarios, where $\hat{G}_{9F}$ never underestimates G, $\hat{G}_{9F}$ systematically underestimates inequality when $G = 0.5$ for highly skewed variables. Slightly larger underestimates are observed in these cases for both $\hat{G}_k^{bc}$ and $\hat{G}_k$. For highly skewed variables with large inequality ($G = 0.5$), both estimators exhibit severe biases when $n \leq 75$, moderate biases when $75 < n \leq 350$, and mild biases when $350 < n \leq 500$. Sample sizes larger than 4000 are required to achieve negligible biases for highly skewed variables with large inequality.

3.3. Second methodological pitfall: use of the naïve cumulative distribution function

The estimator $\hat{G}_9$, defined in Table 2, is a plug-in estimator derived from the theoretical definition of the Gini index in equation (2). As noted by Lerman and Yitzhaki (1989) and Berger (2008), $\hat{G}_9$ must be defined in terms of the smooth (or mid interval) distribution function $\hat{F}_n^*(x)$, see equation (5), for them to correspond to $\hat{G}_k$, $k = 1, \dots, 8, 10$. However, many studies (Berger, 2008; Qin et al., 2010; Hoque and Clarke, 2015; Lv et al, 2017; Rogers et al., 2024) use the naïve cumulative distribution function $\hat{F}_n(t)$, equation (6), instead of $\hat{F}_n^*(x)$ in the definition of $\hat{G}_9$, resulting in equation (8), which is often incorrectly applied in inequality studies.

$$\hat{G}_{9F} = \frac{2}{\bar{x}n}\sum_{i=1}^{n} x_i \hat{F}_n(x_i) - 1. \tag{8}$$

In this subsection, we show that $\hat{G}_{9F}$ may provide seriously biased estimates. From Subsection 3.2, we observe very serious upward biases with small sample sizes when $G = 0.1$, with values of RB_G larger than 25% when $n < 40$, and larger than 75% for samples with a size close to 10. Sample sizes larger than 400 are required to get negligible biases when $G = 0.1$. This bias problem diminishes as G increases, although $\hat{G}_{9F}$ still has moderate and severe biases in the case of small sample sizes. The empirical performance observed for $\hat{G}_{9F}$ in Subsection 3.2 is clearly poorer than that observed for $\hat{G}_k$. This justifies the use of $\hat{F}_n^*(x)$ instead of $\hat{F}_n(t)$ in the definition of $\hat{G}_9$. Note that the use of $\hat{F}_n^*(x)$ instead of $\hat{F}_n(t)$ is also required by the estimator $\hat{G}_9^{bc}$, which is the bias-corrected version of the estimator $\hat{G}_9$.

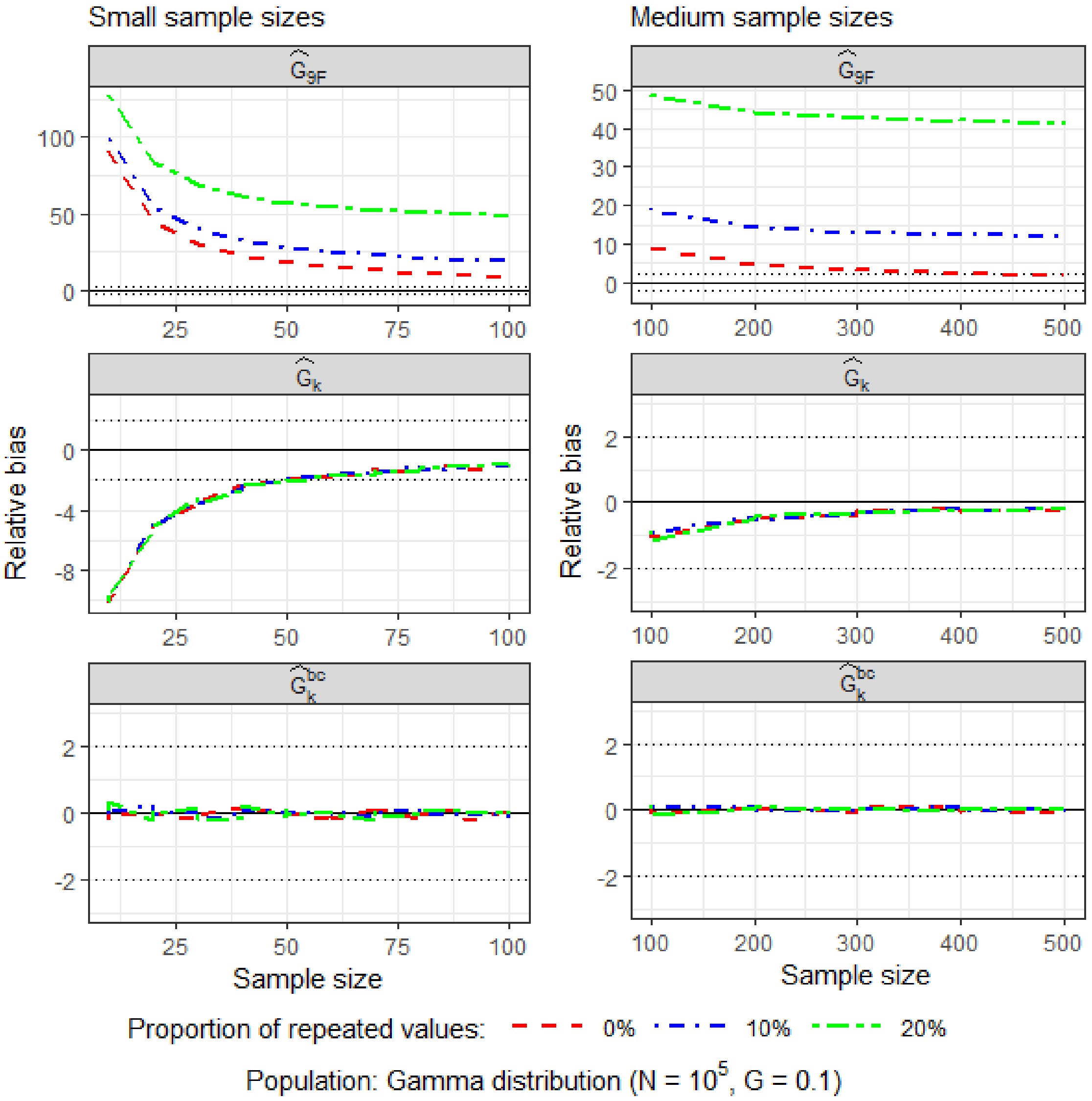

Figure 4. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, $k = 1, \ldots, 10$ (Table 2), and $\hat{G}_{9F}$ (equation (8)). Samples, with sizes ranging from 10 to 500, are drawn from a finite population of size $\mathrm{N} = 10^5$, extracted from a Gamma distribution (low skewness), with shape parameter selected such that $G = 0.1$, with varying proportions p of repeated values, with $p \in \{0, 0.1, 0.2\}$. Solid horizontal lines indicate unbiasedness: $RB_G = 0$. Dashed lines at $RB_G \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_G| \leq 2$.

On the other hand, when the proportion of repeated sample values is large, we observe, focusing on Gamma-based scenarios with low inequality, that this MP produces even larger overestimations (Figure 4), with this bias persisting even for large sample sizes. Figure 4 shows that larger values of RB_G are obtained as the proportion of repeated sample values increases, with biases being up to 30% larger for small sample sizes. Across the various sample sizes analysed in these scenarios, we observe that $\hat{G}_{9F}$ always exhibits non-negligible biases when $p = 10\%$ and severe biases when $p = 20\%$. In contrast, estimators $\hat{G}_k$ and $\hat{G}_k^{bc}$ are not affected by the presence of repeated sample observations (see the middle and bottom panels of Figure 4).

Fortunately, when all sample observations are different, i.e., when $x_i \neq x_j$ for all $i \neq j$ in the sample, the bias caused by the incorrect use of the naïve distribution function has a simple solution. Under this condition, the estimator $\hat{G}_9$ can be easily derived from $\hat{G}_{9F}$ by applying the adjustment $\hat{G}_9 = \hat{G}_{9F} - n^{-1}$.

4. Impact of a popular methodological pitfall on uncertainty estimation

Variance estimates and confidence intervals play an important role in any inferential study and should always be reported, particularly in inequality studies. Uncertainty estimates facilitate comparisons of inequality across regions and/or over time, allowing for the assessment of the statistical significance of changes. Understanding the margins of error induced by sampling in Gini index estimates provides valuable information about the precision and, when combined with expected bias, the accuracy of the point estimates. This enables the measurement of how close the estimates are to the true indices.

In this section, we describe and assess a common, well-recognized MP involving the improper use of Ordinary Least Squares (OLS) in the estimation of the variance of the Gini index. Despite the inappropriateness of this approach being well-documented (Ogwang, 2004; Langel and Tillé, 2013), it continues to be frequently used (Anderson and Thomas, 2019; Ceriani and Verme, 2022; Aspachs et al., 2021). We therefore carry out an exhaustive analysis of its consequences to warn researchers and practitioners about the risks of falling into this MP. This is particularly relevant given the availability of software that computes variance estimates for the Gini index using the regression approach (O'Donnell et al., 2016), which makes this method easily accessible to practitioners.

We contribute to the literature by evaluating the impact of this MP on the bias of variance estimates and the coverage of confidence intervals in scenarios with similar skewness and inequality to those described in Subsection 3.1. This entails studying this MP in novel situations that have not been previously investigated. Furthermore, to contextualize the results and underscore the actual impact of this pitfall, we also compute alternative (theoretically superior) variance and confidence interval estimators. The findings of this study reveal the serious consequences of making inappropriate methodological choices in this context, and serve as a useful tool for identifying the most robust and accurate approaches to estimating uncertainties under varying degrees of skewness and inequality. In fact, we additionally contribute to the literature two novel findings that emerge in highly skewed variables and inequality scenarios. We demonstrate that, in such cases, the regression estimator may also be prone to underestimation, and that the jackknife and linearization methods are not equivalent, with the jackknife method being the preferred approach.

The rest of this section is organized as follows: after outlining the pitfall (Subsection 4.1) and describing the simulated scenarios and alternative estimators used for the assessment (Subsection 4.2), we first evaluate its impact on variance estimation in terms of relative bias (Subsection 4.3), and then assess its effect on the estimated confidence intervals in terms of coverage (Subsection 4.4).

4.1. Third methodological pitfall: use of OLS for variance estimation

This MP arises from the estimator $\hat{G}_5$ defined in Table 2. In particular, Ogwang (2000) shows that $\hat{G}_5$ can be expressed as:

$$\hat{G}_5 = \frac{2\hat{\beta}}{n} - \frac{n+1}{n}, \tag{9}$$

where

$$\hat{\beta} = \frac{\sum_{i=1}^{n} i\, x_{(i)}}{\sum_{i=1}^{n} x_i}$$

coincides with the OLS estimator of the parameter β in the regression model:

$$i\sqrt{x_{(i)}} = \beta\sqrt{x_{(i)}} + u_i, \tag{10}$$

after assuming that the u_i are independent and identically distributed random variables with zero mean and variance σ^2. Giles (2004) uses the results in equations (9) and (10) to propose expression (11) as an estimator of the variance for $\hat{G}_5$:

$$\widehat{var}_{OLS}(\hat{G}_5) = \frac{4}{n^2}\widehat{var}(\hat{\beta}), \tag{11}$$

where $\widehat{var}(\hat{\beta})$ is the variance estimator for the regression coefficient $\hat{\beta}$ obtained from the regression framework.

The appropriateness of equation (11) as a variance estimator for the Gini index has been highly debated in the literature, with many authors (Ogwang, 2004; Langel and Tillé, 2013) advising against its use. They claim that it tends to result in substantial overestimation due to the violation of the independence assumption of the error term in model (10). Berger (2008) also argues that $\widehat{var}_{OLS}(\hat{G}_5)$ cannot be applied to samples derived from finite populations, as this methodology ignores the sampling design. As a novelty, we show that under specific conditions (high skewness and inequality), equation (11) can also lead to significant underestimation.

Despite warnings in the literature, this estimator is still commonly employed in inequality studies, compromising the conclusions due to its significant bias. For example, Anderson and Thomas (2019) and Aspachs et al. (2021) draw inferences based on this incorrect regression approach. Furthermore, Aspachs et al. (2021) affirm that $\widehat{var}_{OLS}(\hat{G}_5)$ is asymptotically equivalent to alternative estimators, leading to the same estimate for large sample sizes. We demonstrate that this statement is misleading. The availability of a software program to compute this estimator (O'Donnell et al., 2016) further exacerbates concerns, as it has been used by many researchers (e.g., Ceriani and Verme, 2022), who then succumb to this MP.

4.2. Description of simulated scenarios, measures of performance and alternative estimators

In this section, we examine the scenarios with low, medium, and high degrees of skewness and inequality discussed in Section 3, but with large sample sizes. Large samples are necessary to accurately estimate variances. As shown in Figure 3, Gini index estimators exhibit substantial underestimation when small sample sizes ($n \in [10, 100]$) are drawn from highly skewed populations with large inequality. This bias directly impacts the performance of confidence intervals. Therefore, to minimize the effect of bias on both the Gini index and its variance estimates, in this section we focus on scenarios with medium and large sample sizes.

To evaluate the performance of both variance and confidence intervals, we follow standard practice (Berger, 2008; Muñoz et al., 2015) and utilize relative bias and empirical coverage. We assess the bias in estimating the variance of $\hat{G}$, denoted as $var(\hat{G})$, using the relative bias, defined by

$$RB_V = 100 \times \frac{E\left[\widehat{var}(\hat{G})\right] - var(\hat{G})}{MSE(\hat{G})},$$

where (i) $\widehat{var}(\hat{G})$ denotes the estimated variance for a given estimator of $var(\hat{G})$ (see Section SM2 in the supplementary material for details), (ii) $E\left[\widehat{var}(\hat{G})\right] = R^{-1}\sum_{r=1}^{R}\widehat{var}(\hat{G})_r$, with $\widehat{var}(\hat{G})_r$ being the estimator $\widehat{var}(\hat{G})$ computed at the rth simulation run, and (iii) the variance and mean square error (MSE) of $\hat{G}$ are approximated, respectively, by:

$$var(\hat{G}) \cong \frac{1}{R-1}\sum_{r=1}^{R}\left[\hat{G}(r) - E\left[\hat{G}\right]\right]^2$$

and

$$MSE(\hat{G}) \cong \frac{1}{R-1}\sum_{r=1}^{R}\left[\hat{G}(r) - G\right]^2.$$

We gauge the empirical performance of confidence intervals using the empirical coverage (EC) statistic. For a given confidence interval $[L, U]$—where L and U are, as a rule, functions of the estimator, its variance and the level of confidence (details in section SM2 of the supplementary material)—EC is defined as:

$$EC = \frac{1}{R}\sum_{r=1}^{R}\delta(L_r \leq G \leq U_r),$$

where $[L_r, U_r]$ denotes the confidence interval computed at the rth simulation run. We set the confidence level at 95% for this simulation study, meaning that, under good performance, EC for a given confidence interval should be close to this nominal level.

To conclude this subsection, we outline the various variance and confidence interval estimators computed to illustrate, in comparative terms, the (extremely poor) performance of the variance estimator $\widehat{var}_{OLS}(\hat{G}_5)$ defined in equation (11). In addition to the regression-based estimator, we also estimate variances using the bootstrap, jackknife, and linearization methods. For confidence intervals, we additionally consider two variants of the jackknife method—one based on a Gaussian distribution (Jackk-z) and another based on the studentized bootstrap (Jackk-t)—as well as the empirical likelihood (EL) method suggested by Qin et al. (2010). A studentized bootstrap approach could also be applied to the linearization method. However, for the sake of clarity in the figures, we have omitted this analysis, as the jackknife estimator demonstrates the best performance in estimating variances in our study (Subsection 4.3).

The bootstrap method has been extensively applied to estimate variances and confidence intervals of the Gini index (Davidson, 2009; Qin et al., 2010). The jackknife method is another popular technique for variance estimation and constructing confidence intervals. In this study, we compute the jackknife estimates using the efficient algorithm suggested by Ogwang (2000). A third technique is the linearization method, which has also been widely applied in this context (Berger, 2008; Langel and Tillé, 2013).

A detailed description of the formulae for the variance estimators and confidence intervals for the Gini index assessed in this research can be found in Section SM2 of the supplementary material. There, we detail the formulae for both infinite populations (those evaluated in this paper) and also, for completeness, finite populations. We refer the reader to the R package *giniVarCI* (Muñoz et al., 2024) for computing the variance estimators and confidence intervals described in this section and in the supplementary material.

4.3. Impact of the OLS pitfall on variance estimates

Figures 5, 6 and 7 present the estimated relative biases of the various variance estimators for the Gini index. For less skewed distributions (Figure 5), we observe that the bootstrap, jackknife and linearization methods exhibit reasonable biases across all degrees of inequality, with RB_V values typically in the interval $[-2, 2]$. In contrast, the OLS method leads to a severe overestimation, which worsens as inequality decreases. Furthermore, contrary to the assertion by Aspachs et al. (2021), the overestimation problem of the OLS method does not improve with increasing sample size.

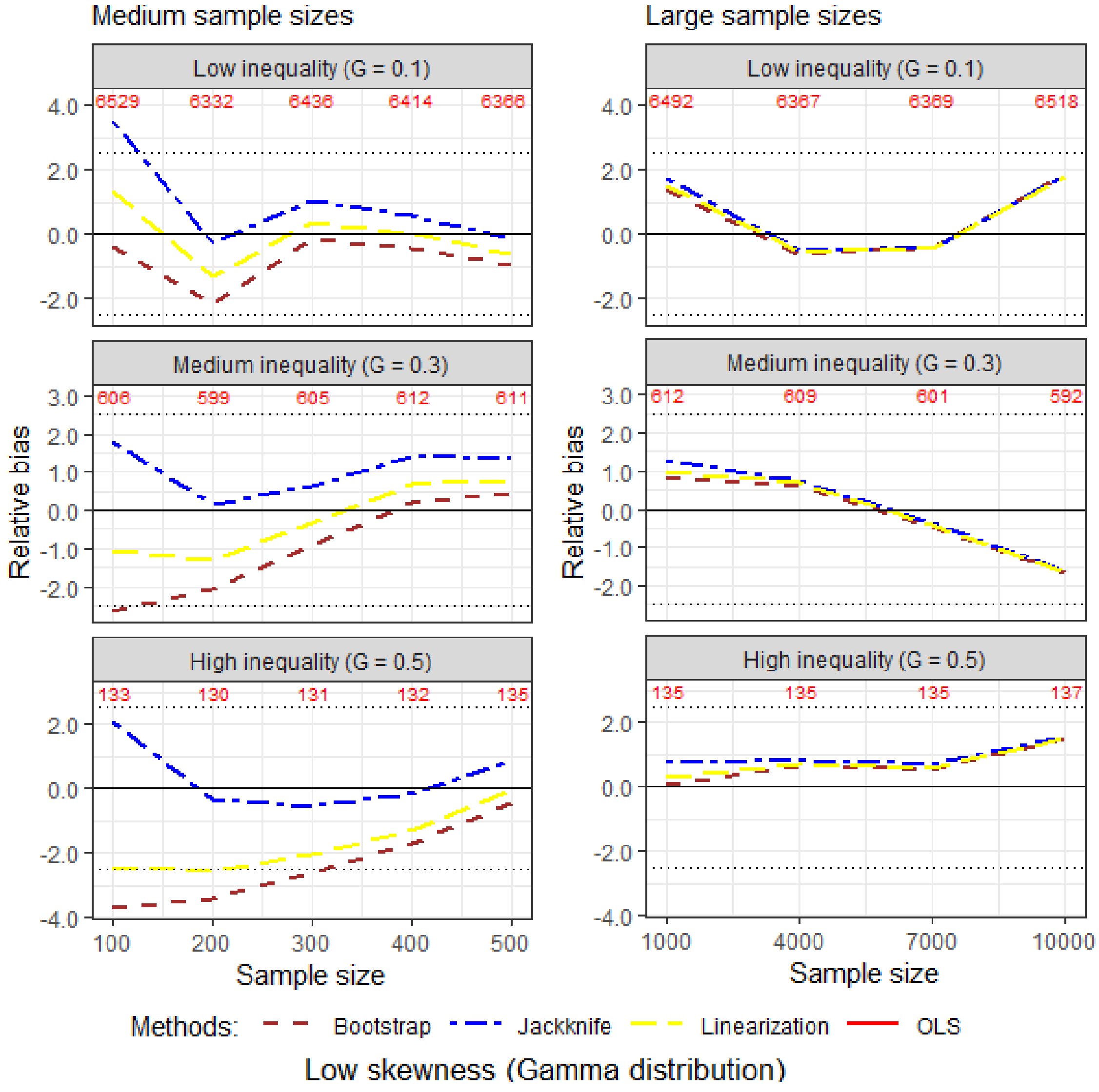

Figure 5. Relative bias (RB_V) for different variance estimators: $\widehat{var}_B(\hat{G}_k)$ (Bootstrap); $\widehat{var}_J(\hat{G}_k)$ (Jackknife); $\widehat{var}_L(\hat{G}_k)$ (Linearization); and $\widehat{var}_{OLS}(\hat{G}_k)$ (OLS), $k = 1, \dots, 10$. Details in Subsection SM2.1 in the supplementary material. The OLS method significantly overestimates the variance, and its lines have been replaced by the exact estimated values of RB_V to maintain figure interpretability and avoid scale issues. Samples, with sizes ranging from 100 to 10000, are drawn from a Gamma distribution (low skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_V = 0$. Dashed lines at $RB_V \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_V| \leq 2$.

For distributions with a medium skewness (Figure 6), the overestimation of the OLS method persists, although its RB_V values are noticeably smaller compared to less skewed distributions. The jackknife method provides the most accurate estimates across all sample sizes and inequality levels. The linearization and bootstrap methods can struggle with small sample sizes, although the linearization method still performs well when inequality levels are small or moderate. The bootstrap method is the one exhibiting the larger relative biases in small sample sizes, suffering an increasing underestimation problem as the inequality rises and the sample size decreases.

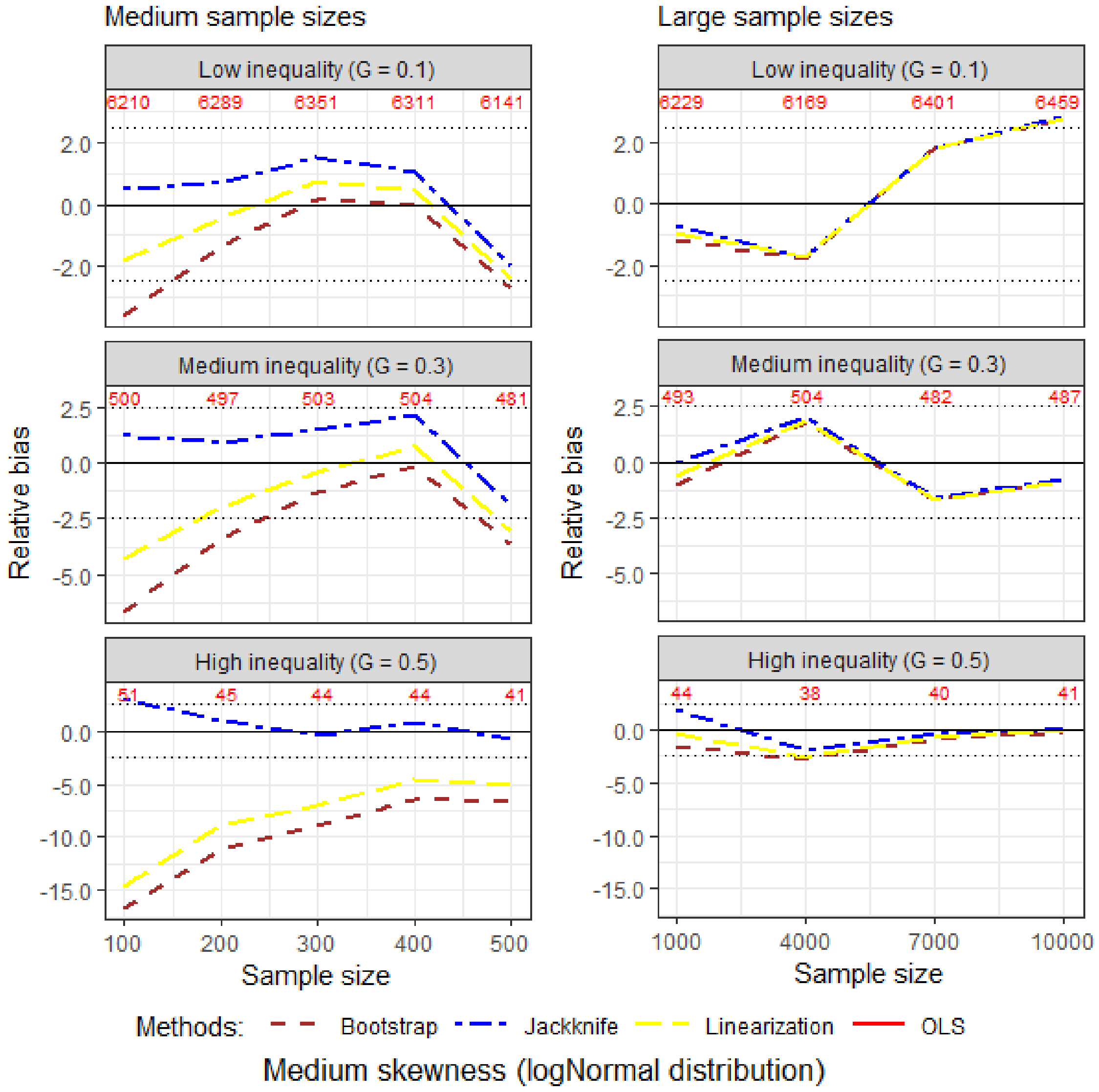

Figure 6. Relative bias (RB_V) for different variance estimators: $\widehat{var}_B(\hat{G}_k)$ (Bootstrap); $\widehat{var}_J(\hat{G}_k)$ (Jackknife); $\widehat{var}_L(\hat{G}_k)$ (Linearization); and $\widehat{var}_{OLS}(\hat{G}_k)$ (OLS), $k = 1, \ldots, 10$. Details in Subsection SM2.1 in the supplementary material). The OLS method significantly overestimates the variance, and its lines have been replaced by the exact estimated values of RB_V to maintain figure interpretability and avoid scale issues. Samples, with sizes ranging from 100 to 10000, are drawn from a logNormal distribution (medium skewness), with standard deviation parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_V = 0$. Dashed lines at $RB_V \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_V| \leq 2$.

For highly skewed distributions (Figure 7), the biases in the variance estimates obtained with the OLS method are mixed. Results show substantial overestimation when $G = 0.1$, but negative biases when $G \in \{0.3, 0.5\}$. Similar to the medium skewness scenario (Figure 6), the jackknife shows the most accurate estimates, performing well across all sample sizes and inequality levels. Both the linearization and bootstrap methods perform poorly, suffering from underestimation, though they yield good estimates when inequality is low and sample sizes are large. Overall, the jackknife method consistently delivers the best results in all the scenarios analysed.

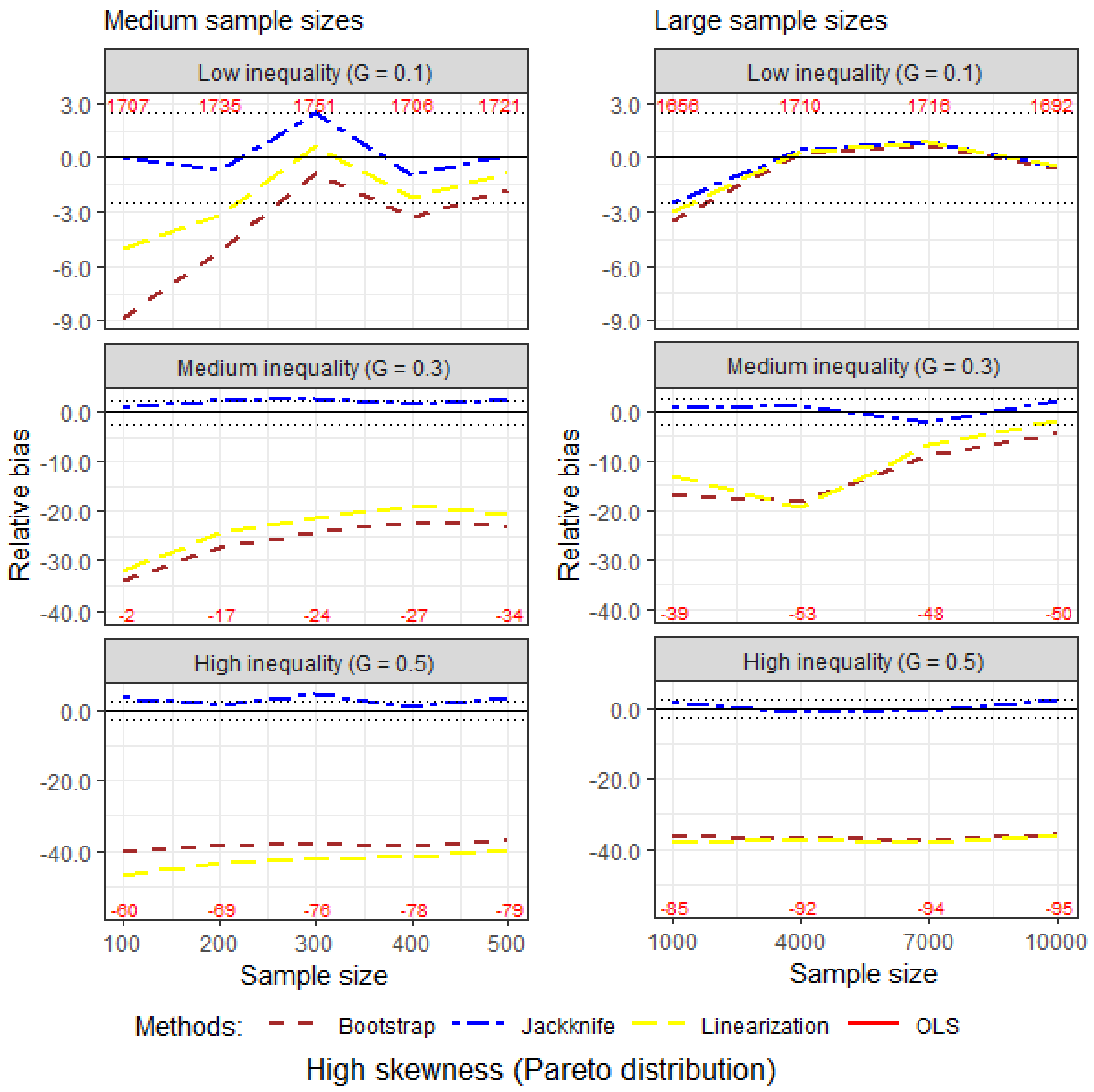

Figure 7. Relative bias (RB_V) for different variance estimators: $\widehat{var}_B(\hat{G}_k)$ (Bootstrap); $\widehat{var}_J(\hat{G}_k)$ (Jackknife); $\widehat{var}_L(\hat{G}_k)$ (Linearization); and $\widehat{var}_{OLS}(\hat{G}_k)$ (OLS), $k = 1, \dots, 10$. Details in Subsection SM2.1 in the supplementary material). The OLS method significantly overestimates the variance, and its lines have been replaced by the exact estimated values of RB_V to maintain figure interpretability and avoid scale issues. Samples, with sizes ranging from 100 to 10000, are drawn from a Pareto distribution (high skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: set at $RB_V = 0$. Dashed lines at $RB_V \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_V| \leq 2$.

Figure 7 offers a new insight into the literature on variance estimation of the Gini index. While several authors (Berger, 2008; Langel and Tillé, 2013) have claimed an asymptotic equivalence between the jackknife and linearization methods, our results suggest that this equivalence does not hold in scenarios with high skewness and moderate to large inequality. Although the two methods are equivalent in the majority of scenarios investigated in this paper, we find that the jackknife approach clearly outperforms the linearization method in terms of bias under conditions of both extreme skewness and inequality, with this advantage persisting even for large sample sizes.

4.4. Impact of the OLS pitfall on empirical coverage of confidence intervals

Figures 8, 9, and 10 depict empirical coverage (EC) for the various confidence intervals introduced in Subsection 4.2 and detailed in Subsection SM2.1 of the supplementary material. The observed coverage closely mirrors the results obtained for variance estimates in Subsection 4.3. When the variance estimator overestimates, the corresponding confidence interval exhibits an excess EC compared to the nominal level of 95%. Similarly, confidence intervals show a coverage deficit when the associated variance estimator underestimates. These patterns are seen with OLS, bootstrap, and linearization methods.

Overall, OLS confidence intervals tend to systematically overestimate coverage, except in highly skewed scenarios with moderate to high inequality, where they tend to underestimate expected coverage (Figure 10). The behaviour of bootstrap and linearization intervals is more mixed, as are the relative biases of the corresponding variance estimates. For less skewed variables (Figure 8), bootstrap and linearization intervals converge to the nominal level as sample size increases. In fact, all intervals, except the OLS ones, record EC values between 94% and 95%, approximately. In medium skewed scenarios (Figure 9), OLS intervals show near 100% coverage for low to medium inequality variables and about 97.5% coverage for high inequality ($G = 0.5$). The remaining confidence intervals provide desirable coverage for large sample sizes ($n \geq 1000$). However, for smaller samples, the coverage of bootstrap and linearization intervals worsens as inequality rises. For highly skewed variables (Figure 10), bootstrap and linearization intervals underestimate coverage similarly to how their variance estimates exhibit a significant negative bias. Likewise, the OLS interval performs poorly when both skewness and inequality are high.

For some methods, there is no isomorphism between variance estimates and confidence intervals, meaning that variance biases do not directly translate into confidence interval coverage. The EL method does not rely on variance estimation, while two confidence intervals have been constructed using the jackknife variance. Overall, in terms of coverage, EL confidence intervals are comparable to those of the linearization method, though slightly less accurate. The major drawback of EL confidence intervals is their substantial computational cost, which is more than 10 times higher than that of the jackknife method.

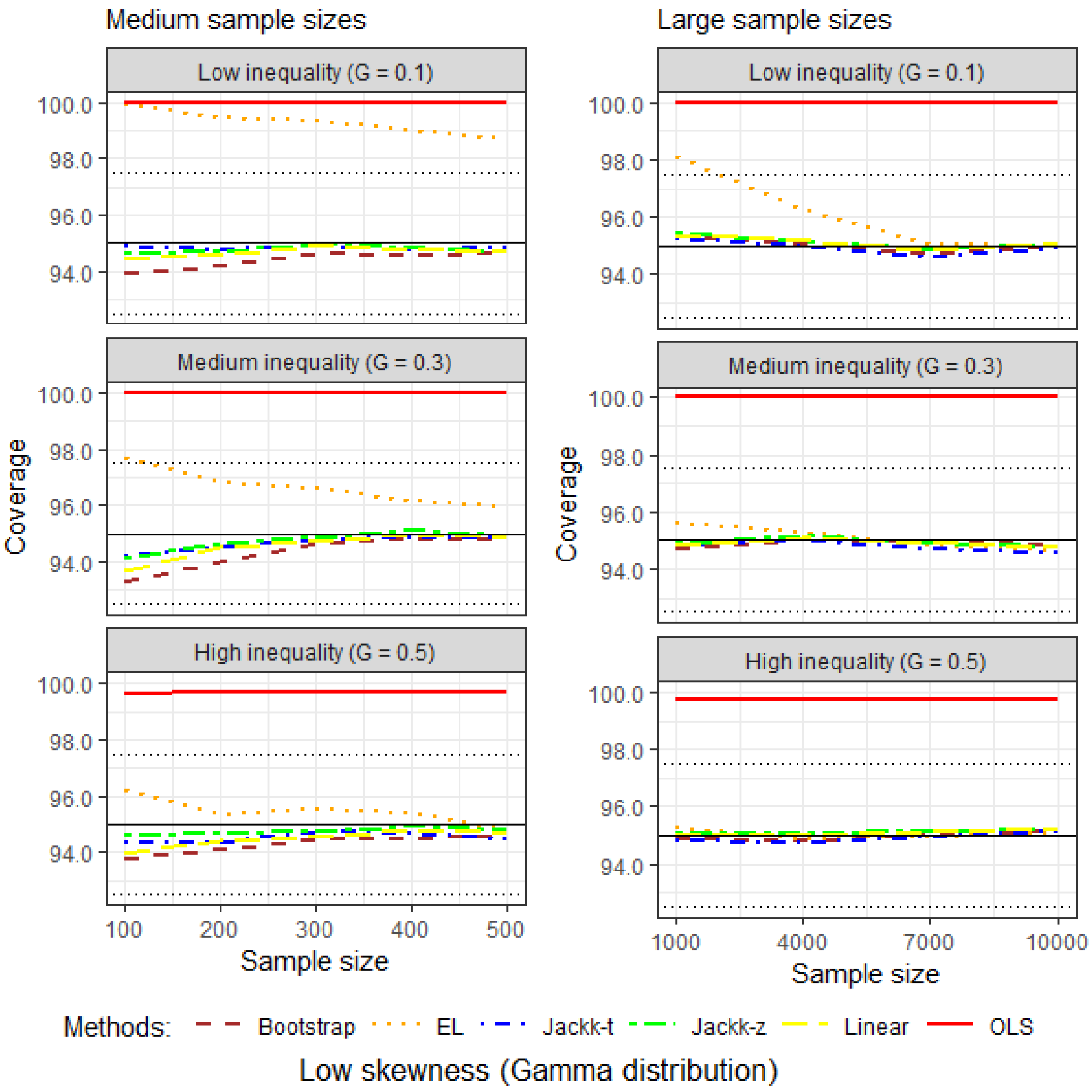

Figure 8. Empirical coverage (EC) for different 95% confidence intervals: $CI_{k;EL}$ (Empirical likelihood) $CI_{k;Bp}$ (Bootstrap); $CI_{k;Jt}$ (Jackknife-t); $CI_{k;Jz}$ (Jackknife-z); $CI_{k;L}$ (Linearization); and $CI_{k;OLS}$ (OLS), $k = 1, \dots, 10$. Details in Subsection SM2.1 in the supplementary material. Samples, with sizes ranging from 100 to 10,000, are drawn from a Gamma distribution (low skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Horizontal lines are fixed at $EC \in \{92.5\%, 95\%, 97.5\%\}$.

The best variance estimates, measured by relative bias, are obtained using the jackknife method. However, the two approaches for constructing confidence intervals based on jackknife variance are not equivalent. EC of the intervals based on the Gaussian approximation is similar to that of the linearization method but falls significantly below the nominal level for small sample sizes in medium and highly skewed scenarios (Figures 9 and 10) and for all sample sizes in highly skewed scenarios with moderate to high inequality (Figure 10). Jackknife intervals improve their coverage when studentized quantiles are used. In fact, jackknife intervals based on studentized quantiles consistently exhibit EC values much closer to the nominal level. This advantage is particularly pronounced for highly skewed variables (Figure 10), where EC values for Jackk-t intervals are clearly closer to 95% than those of competing methods.

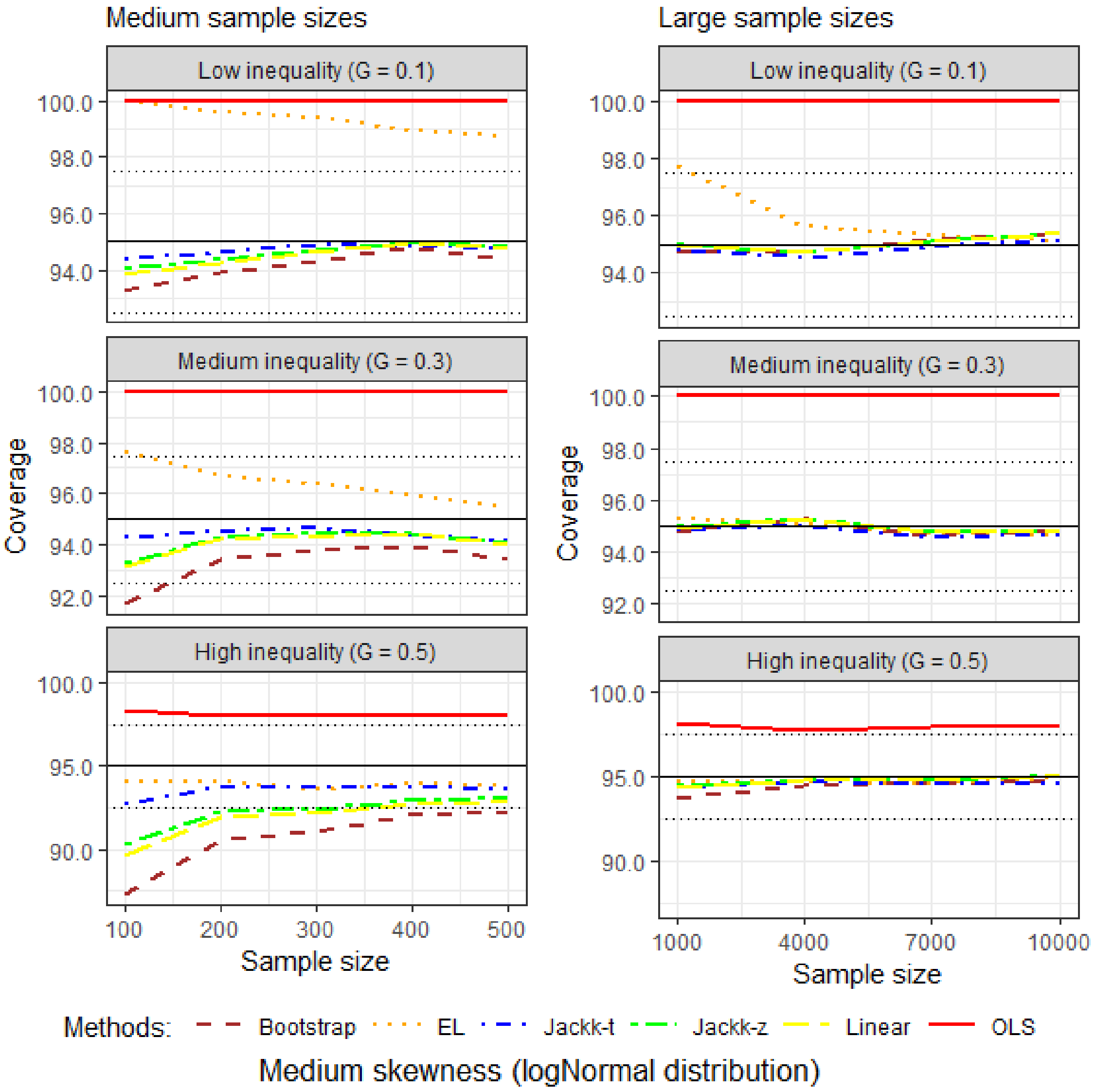

Figure 9. Empirical coverage (EC) for different 95% confidence intervals: $CI_{k;EL}$ (Empirical likelihood) $CI_{k;Bp}$ (Bootstrap); $CI_{k;Jt}$ (Jackknife-t); $CI_{k;Jz}$ (Jackknife-z); $CI_{k;L}$ (Linearization); and $CI_{k;OLS}$ (OLS), $k = 1, \dots, 10$. Details in Subsection SM2.1 in the supplementary material. Samples, with sizes ranging from 100 to 10,000, are drawn from a logNormal distribution (medium skewness), with standard deviation parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Horizontal lines are fixed at $EC \in \{92.5\%, 95\%, 97.5\%\}$.

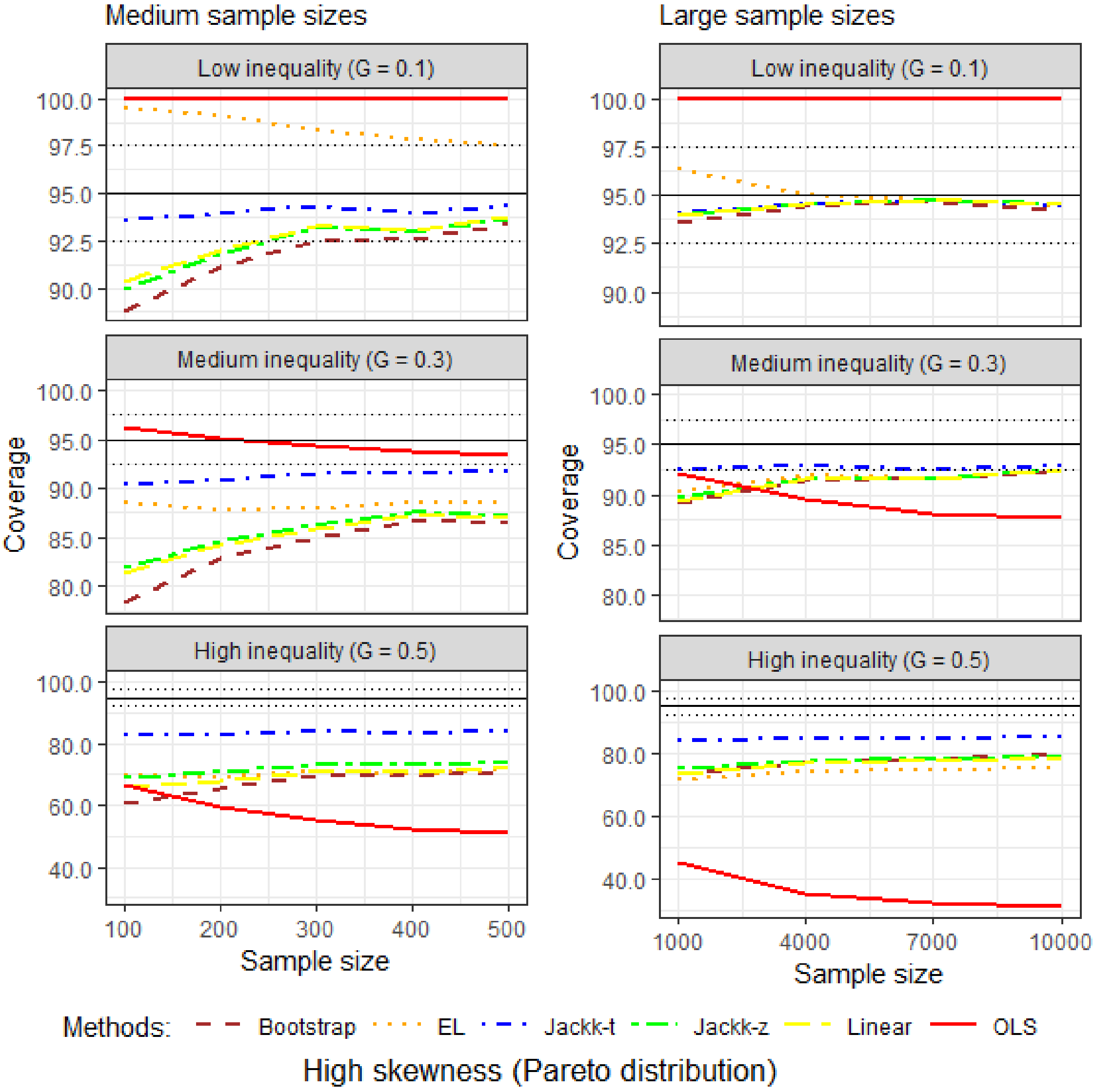

Figure 10. Empirical coverage (EC) for different 95% confidence intervals: $CI_{k;EL}$ (Empirical likelihood) $CI_{k;Bp}$ (Bootstrap); $CI_{k;Jt}$ (Jackknife-t); $CI_{k;Jz}$ (Jackknife-z); $CI_{k;L}$ (Linearization); and $CI_{k;OLS}$ (OLS), $k = 1, \dots, 10$. Details in Subsection SM2.1 in the supplementary material. Samples, with sizes ranging from 100 to 10,000, are drawn from a Pareto distribution (high skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Horizontal lines are fixed at $EC \in \{92.5\%, 95\%, 97.5\%\}$.

5. Conclusions

The Gini index has been extensively studied in the literature, with various formulations proposed for its estimation. However, some commonly-used expressions of the Gini index suffer from methodological pitfalls (MPs) that can significantly impact the accuracy of estimates. Certain MPs lead to severe biases, resulting in incorrect inferences and unsound conclusions. MPs affect not only point estimates of the Gini index but also variance estimates, producing poor-quality estimates of both variances and confidence intervals. The situation is exacerbated by the availability of software that implements these flawed approaches.

Thus, in order to conduct accurate inequality studies without incurring biases or drawing erroneous conclusions, it is essential to identify and evaluate the effects of MPs on Gini index inferences, and to reference appropriate software. We identify and analyse three main MPs: (i) using non-biased-corrected estimators, (ii) using a naïve cumulative distribution function instead of a modified one in a particular estimator, and (iii) using an OLS based estimator for sample variance. A contribution of this paper lies in empirically assessing the impact of the identified MPs across a wide range of scenarios commonly encountered in practice, defined by varying levels of skewness, inequality, and sample sizes. Our findings not only corroborate previous results, but also reveal new insights unrelated to MPs. For example, while the literature consistently reports that a regression-based variance estimator for the Gini index leads to significant overestimation, we find this does not hold for highly skewed variables. In this situation, our simulations show that this variance estimator tends to underestimate variance as the inequality increases.

This paper can serve as a powerful tool for researchers and practitioners who work with the Gini index, for various reasons. To start, it provides a complete and up-to-date review of non-parametric estimation of the Gini index (Subsections 2.1 and 2.2). In practice, bias correction is usually overlooked when estimating the Gini index, and this may lead to serious underestimation when dealing with small samples. Table 2 clearly identifies the existing expressions of the Gini index for infinite populations that take the bias correction into account. This is the first MP discussed in this paper (Subsections 3.2).

The second MP analysed consists of studying the impact of using the naïve distribution function in the definition of the Gini index estimator (Subsections 3.3). This MP can significantly overestimate the true Gini index, especially for variables with a low level of inequality and small sample sizes. Furthermore, we find that this approach also performs poorly in samples with a large proportion of repeated values.

Estimation of the variance of the Gini index has received increasing attention over recent decades. Both variance estimates and confidence intervals play an important role in inequality studies, facilitating comparisons across groups, regions and over time. The variance estimator based on the OLS method, as defined in equation (11), has been reported to lead to significant overestimation, prompting many authors to advise against its use. Nonetheless, this estimator has seen extensive application in recent years. This study examines the impact of this MP, also delving into novel scenarios not previously investigated. Within this analysis, we identify the most robust and accurate approaches for estimating variance and confidence intervals (Section 4).

First, the findings reveal that the OLS method not only leads to overestimation but can also significantly underestimate the true variance in highly skewed populations with medium to large inequality, which are common in economic studies. Second, the study identifies the jackknife method as the best approach for estimating variances and confidence intervals, consistently delivering the most accurate variance estimates across all scenarios analysed. Furthermore, jackknife intervals based on studentized quantiles show coverages that align closely with the nominal level in all examined scenarios. As a further new insight, we demonstrate that, contrary to claims based on studies of finite populations and simulations in low to moderately skewed scenarios (Berger, 2008; Langel and Tillé, 2013), the jackknife and linearization methods are not asymptotically equivalent in highly skewed populations with medium to large inequality.

Although this research has produced several important findings, it leaves ample room for future investigation. For instance, the evaluation could be extended to finite populations. This study clarifies the impact of common pitfalls on Gini index inferences, typically assuming infinite populations and within the context of simple random sampling to avoid nuisance effects of the sampling design; accordingly, a potential avenue for future research is to investigate the empirical properties of the statistical methods under finite populations and across different sampling designs. This could include assessing estimator performance when serial correlation, survey weights, clustering, or other design features are present. Furthermore, as biases can also arise from other sources such as grouped data, future work could focus on analysing how this grouping-induced bias behaves under the scenarios explored here, following, for instance, the first-order bias correction for grouped data proposed by Van Ourti and Clarke (2011). Additionally, studying the impact of all the analysed factors on Gini index decomposition (Larraz, 2016) represents another promising direction for further research.

In summary, compared to previous reviews, this paper provides a clearer and more comprehensive description of the challenges in estimating the Gini index. It provides a thorough and updated review of the literature on Gini index estimation and inference, offering a guide to practitioners. Additionally, it references user-friendly and accessible software for applying the statistical expressions presented here. The inferential properties of the Gini index are described and examined, offering readers relevant material for constructing more reliable confidence intervals and making more meaningful comparisons in inequality studies. We also analyse the serious consequences of three common MPs associated with the Gini index across a variety of practical scenarios. By highlighting these issues, we hope we can help socioeconomic researchers avoid them in the future. Moreover, future research could extend these analyses by exploring the impact of sampling design in other contexts, such as finite populations or grouped data scenarios.

Data availability Data are obtained from various probabilistic distributions and are available at: https://osf.io/4jnuz/

Declarations

Conflicts of interest The authors declare that there are no conflicts of interest.

References

Alfons, A., & Templ, M. (2013). Estimation of social exclusion indicators from complex surveys: The R package laeken. *Journal of Statistical Software*, 54, 1-25. https:// 10.18637/jss.v054.i15

Anderson, G., & Thomas, J. (2019). Measuring multi-group polarization, segmentation and ambiguity: increasingly unequal yet similar constituent Canadian income distributions. *Social Indicators Research*, 145, 1001-1032.

Aspachs, O., Durante, R., Graziano, A., Mestres, J., Reynal-Querol, M., & Montalvo, J. G. (2021). Tracking the impact of COVID-19 on economic inequality at high frequency. *PLoS One*, 16(3), e0249121.

Banerjee, B., & Tóth, P. (2025). Life satisfaction and inequality in Slovakia: The role of income, consumption and wealth. *Social Indicators Research*, 177(1), 93-126.

Berger, Y.G. (2008) A note on the asymptotic equivalence of jackknife and linearization variance estimation for the Gini Coefficient. *Journal of Official Statistics*, 24(4), 541-555.

Berger, Y.G., & Gedik-Balay, İ. (2020). Confidence intervals of Gini coefficient under unequal probability sampling. *Journal of Official Statistics*, 36(2), 237-249.

Ceriani, L., & Verme, P. (2022). Population changes and the measurement of inequality. *Social Indicators Research*, 162, 549–575.

Cowell, F. (2011). *Measuring Inequality*. Oxford University Press.

Davidson, R. (2009). Reliable inference for the Gini index. *Journal of Econometrics*, 150(1), 30-40.

Deltas, G. (2003). The small-sample bias of the Gini coefficient: Results and implications for empirical research. *The Review of Economics and Statistics*, 85(1), 226-234.

Filauro, S., Parolin, Z., & Valetto, P. (2025). What explains recent trends in income inequality in the European Union? *The Journal of Economic Inequality*, 23, 483-505.

Fontanari, A., Taleb, N.N., & Cirillo, P. (2018). Gini estimation under infinite variance. *Physica A: Statistical Mechanics and its Applications*, 502, 256-269.

Giles, D.E. (2004). Calculating a standard error for the Gini coefficient: some further results. *Oxford Bulletin of Economics and Statistics*, 66(3), 425-433.

Giorgi, G.M., & Gigliarano, C. (2017). The Gini concentration index: a review of the inference literature. *Journal of Economic Surveys*, 31(4), 1130-1148.

Hoque, A.A., & Clarke, J.A. (2015). On variance estimation for a Gini coefficient estimator obtained from complex survey data. *Communications in Statistics: Case Studies, Data Analysis and Applications*, 1(1), 39-58.

Ibragimova, Z., & Frants, M. (2021). Measuring inequality of opportunity: Does inequality index matter? *Statistika: Statistics & Economy Journal*, 101(1).

Kattumannil, S.K., & Dewan, I. (2021). Non-parametric estimation of Gini index with right censored observations. *Statistics & Probability Letters*, 175, 109113.

Langel, M., & Tillé, Y. (2013). Variance estimation of the Gini index: Revisiting a result several times published. *Journal of the Royal Statistical Society, Series A*, 176(2), 521-540.

Larraz, B. (2016). Decomposing the Gini inequality index: An expanded solution with survey data applied to analyze gender income inequality. *Sociological Methods & Research*, 44(3), 508-533.

Lee, D., & Suh, S. (2025). Measuring income and wealth inequality: A note on the Gini coefficient for samples with negative values. *Social Indicators Research*, 176(3), 947-965.

Lerman, R.I., & Yitzhaki, S. (1984). A note on the calculation and interpretation of the Gini index. *Economics Letters*, 15(3-4), 363-368.

Lerman, R.I., & Yitzhaki, S. (1989). Improving the accuracy of estimates of Gini coefficients. *Journal of Econometrics*, 42(1), 43-47.

Lorenz, M.O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70), 209-219.

Lv, X., Zhang, G., & Ren, G. (2017). Gini index estimation for lifetime data. *Lifetime Data Analysis*, 23, 275-304.

Muñoz, J.F., Álvarez-Verdejo, E., García-Fernández, R.M., & Barroso, L.J. (2015). Efficient estimation of the Headcount index. *Social Indicators Research*, 123, 713-732.

Muñoz, J. F., Moya-Fernández, P. J., & Álvarez-Verdejo, E. (2025). Exploring and correcting the bias in the estimation of the Gini measure of inequality. *Sociological Methods & Research*, 54(1), 237-274.

Muñoz, J.F., Pavía, J.M., & Álvarez-Verdejo, E. (2024). *giniVarCI: Gini Indices, Variances and Confidence Intervals for Finite and Infinite Populations*. R packages version 0.0.1-3. https://cran.r-project.org/package=giniVarCI

O'Donnell, O., O'Neill, S., Van Ourti, T., & Walsh, B. (2016). Conindex: estimation of concentration indices. *The Stata Journal*, 16(1), 112-138.

OECD (2017). *How's Life? 2017: Measuring Well-being*. OECD Publishing.

Ogwang, T. (2000). A convenient method of computing the Gini index and its standard error. *Oxford Bulletin of Economics & Statistics*, 62(1), 123-129.

Ogwang, T. (2004) Calculating a standard error for the Gini coefficient: some further results: reply. *Oxford Bulletin of Economics and Statistics*, 66, 435-437.

Qin, Y., Rao, J.N.K., & Wu, C. (2010). Empirical likelihood confidence intervals for the Gini measure of income inequality. *Economic Modelling*, 27(6), 1429-1435.

Raffinetti, E., Siletti, E., & Vernizzi, A. (2015). On the Gini coefficient normalization when attributes with negative values are considered. *Statistical Methods & Applications*, 24(3), 507-521.

Rogers, A.E., Wichman, C.S., Schenkelberg, M.A., & Dzewaltowski, D.A. (2024). Inequality in physical activity in organized group settings for children: A cross-sectional study. *Journal of Physical Activity and Health*, 1(aop), 1-11.

Särndal, C.E., Swensson, B., & Wretman, J. (2003) *Model Assisted Survey Sampling*. Springer Science & Business Media.

Signorell, A. (2023). *DescTools: Tools for Descriptive Statistics*. R package version 0.99.50. https://CRAN.R-project.org/package=DescTools

Topuz, S.G. (2022). The relationship between income inequality and economic growth: Are transmission channels effective? *Social Indicators Research*, 162, 1177-1231.

Van Ourti, T., & Clarke, P. (2011). A simple correction to remove the bias of the Gini coefficient due to grouping. *The Review of Economics and Statistics*, 93(3), 982-994.

Supplementary Material for the paper

A PRACTICAL GUIDE TO PROPER ESTIMATION AND INFERENCE OF THE GINI INDEX BY AVOINDING OFTEN ENCOUNTERED METHODOLOGICAL PITFALLS

Table of contents:

SM1. Literature review on Gini index studies and categorized by different scenarios

For various references included in our literature review, Table SM1 presents a graphical summary of the scenarios previously examines—categorized by their focus in bias (B), use of the distribution function (F), and variance estimation (V). Each scenario is defined by three dimensions: (i) the value of the Gini (G) index (Low (L) when $G \leq 0.25$, Medium (M) when $0.25 < G \leq 0.45$, and High (H) when $G > 0.45$); (ii) the sample size (Low (L) when $n \leq 100$, Medium (M) when $100 < n \leq 500$, and High (H) when $n > 500$); and the skewness ($\gamma_{(G)}$) of the underlying distribution (Low (L) when $\gamma_{(G)} \leq 0.8\gamma_{L(G)}$, Medium (M) when $0.8\gamma_{L(G)} < \gamma_{(G)} \leq 1.2\gamma_{L(G)}$, and High (H) when $\gamma_{(G)} > 1.2\gamma_{L(G)}$, where $\gamma_{L(G)}$ is the skewness coefficient of the logNormal distribution with Gini index G). For real data sets, this classification is applied using the skewness coefficient of the data set and the empirical average skewness of the logNormal distribution based on 10000 samples with the same size. A check mark symbol (✓) indicates that the corresponding reference addresses the associated scenario, and the cell is shaded in this situation. Note that Davidson (2012) uses the empirical distribution function (EDF) to compare the performance of Gini index estimators. Given that the EDF integrates both bias and variance, we believe that this reference addresses both performance dimensions.

This paper analyses all 81 scenarios presented in Table SM1, complementing the 51 scenarios previously explored in the literature (see Table 1). Although Muñoz et al. (2025) examine a significant number of scenarios, they omit the case of large sample sizes as well as the critical issues of variance estimation and confidence interval construction. This paper, therefore, fills an evident gap remaining in the literature, as most references in Table SM1 fail to explore several relevant scenarios and given that the skewness and the Gin index level markedly affect the results, as evidenced in this study. For example, Berger (2020) analyses high Gini index scenarios, but restricts the analysis to variables with a low skewness. As a result, medium and high skewness levels are neglected. Such scenarios are crucial to consider, as they can cause substantial biases in the estimation of the Gini index.

Although Nygard and Sandstrom (1985, 1989), Lerman and Yitzhaki (1989), Cowell (1989); Ogwang (2000); Karagiannis and Kovacevic (2000), Giles (2004, 2006), Ogwang (2004), Langel and Tillé (2011); Alfons et al. (2013), and Goga and Ruiz-Gazen (2014) were also considered in our literature review, they are omitted from Table SM1 due to reasons such as unavailability of datasets or the data necessary for Table SM1, or the lack of Monte Carlo simulation studies.

Table SM1*. Literature review of the Gini index with Monte Carlo simulation studies—categorized by their focus in bias (B), use of the distribution function (F), and variance estimation (V)—by scenario. A check mark symbol (✓) indicates that the corresponding reference addresses the associated scenario.*

Reference	Skewness	Sample Size	*B*	*F*	*V*	*B*	*F*	*V*	*B*	*F*	*V*
Present paper	*H*	*H*	✓	✓	✓	✓	✓	✓	✓	✓	✓
		M	✓	✓	✓	✓	✓	✓	✓	✓	✓
		L	✓	✓	✓	✓	✓	✓	✓	✓	✓
	M	*H*	✓	✓	✓	✓	✓	✓	✓	✓	✓
		M	✓	✓	✓	✓	✓	✓	✓	✓	✓
		L	✓	✓	✓	✓	✓	✓	✓	✓	✓
	L	*H*	✓	✓	✓	✓	✓	✓	✓	✓	✓
		M	✓	✓	✓	✓	✓	✓	✓	✓	✓
		L	✓	✓	✓	✓	✓	✓	✓	✓	✓
Muñoz et al. (2025)	*H*	*H*									
		M	✓	✓		✓	✓		✓	✓	
		L	✓	✓		✓	✓		✓	✓	
	M	*H*									
		M	✓	✓		✓	✓		✓	✓	
		L	✓	✓		✓	✓		✓	✓	
	L	*H*									
		M	✓	✓		✓	✓		✓	✓	
		L	✓	✓		✓	✓		✓	✓	
Berger and Gedik-Balay *(2020)*	*H*	*H*									
		M	✓		✓						
		L									
	M	*H*									
		M				✓		✓			
		L									
	L	*H*									
		M	✓		✓				✓		✓
		L									
Fontanari et al. (2018)	*H*	*H*				✓			✓		
		M				✓			✓		
		L				✓			✓		
	M	*H*									
		M									
		L									
	L	*H*									
		M									
		L									
			L			*M*			*H*		
						Gini					

***Continuation of Table SM1*.**

			Gini								
		Sample	*L*			*M*			*H*		
Reference	Skewness	Size	*B*	*F*	*V*	*B*	*F*	*V*	*B*	*F*	*V*
Wang et al. (2016)	*H*	*H*									
		M									
		L			✓						
	M	*H*									
		M									
		L						✓			
	L	*H*									
		M									
		L									✓
Langel and Tillé (2013)	*H*	*H*									
		M									
		L									
	M	*H*									
		M									
		L									
	L	*H*									
		M									
		L									✓
Davidson (2012)	*H*	*H*									
		M									
		L	✓		✓	✓		✓			
	M	*H*									
		M									
		L				✓		✓	✓		✓
	L	*H*									
		M									
		L							✓		✓
Peng (2011)	*H*	*H*									
		M									
		L			✓						
	M	*H*									
		M									
		L						✓			✓
	L	*H*									
		M									
		L									✓
			L			*M*			*H*		
			Gini								

Continuation of Table SM1.

			Gini								
		Sample	*L*			*M*			*H*		
Reference	Skewness	size	*B*	*F*	*V*	*B*	*F*	*V*	*B*	*F*	*V*
Qin et al. (2010)	*H*	*H*									
		M									
		L									
	M	*H*									
		M									
		L									✓
	L	*H*									
		M									
		L						✓			✓
Davidson (2009)	*H*	*H*									
		M									
		L	✓		✓	✓		✓			
	M	*H*									
		M									
		L				✓		✓	✓		✓
	L	*H*									
		M									
		L							✓		✓
Berger (2008)	*H*	*H*									
		M									
		L									
	M	*H*									
		M									
		L				✓		✓			
	L	*H*									
		M									
		L				✓		✓	✓		✓
Modarres and Gastwirth (2006)	*H*	*H*	✓		✓						
		M	✓		✓						
		L	✓		✓						
	M	*H*									
		M									
		L									
	L	*H*							✓		✓
		M							✓		✓
		L							✓		✓
			L			*M*			*H*		
			Gini								

***Continuation of Table SM1*.**

			Gini								
		Sample	*L*			*M*			*H*		
Reference	Skewness	size	*B*	*F*	*V*	*B*	*F*	*V*	*B*	*F*	*V*
Deltas (2003)	*H*	*H*									
		M									
		L									
	M	*H*									
		M									
		L				✓			✓		
	L	*H*									
		M									
		L				✓			✓		
			L			*M*			*H*		
			Gini								

SM2. Analytical expressions for variance estimators and confidence intervals for the Gini index

This section is divided into two subsections. Subsection SM2.1 presents the analytical expressions for the variance estimators and confidence intervals of the Gini index that are evaluated in Section 4. These expressions pertain to infinite populations. Subsection SM2.2 includes the mathematical formulae for finite populations for interested readers. Both sets of formulae are implemented in the R-package *giniVarCI* (Muñoz et al., 2024), to which we refer readers for computations.

SM2.1. Samples derived from an infinite population

For simplicity, this subsection presents the variance and confidence interval expressions for the non-bias-corrected estimator $\widehat{G}_k$, where $k = 1, \ldots, 10$. The bias-corrected expressions for the estimator $\widehat{G}_k^{bc}$ can be easily derived using the relationship between both estimators outlined in equation (3).

Bootstrap method. Let $\{x_1^*(b), \ldots, x_n^*(b)\}$ be the bth bootstrap sample taken from the original sample $\{x_i : i \in S\}$ through simple random sampling with replacement, where $b \in \{1, \ldots, B\}$ and B denotes the total number of bootstrap samples. A variance estimator of $\widehat{G}_k$ based on the bootstrap method can be defined as:

$$\widehat{var}_B(\widehat{G}_k) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{G}_k^*(b) - \bar{G}_k^*\right)^2,$$

where $\widehat{G}_k^*(b)$ represents the estimator $\widehat{G}_k$ computed from the bth bootstrap sample, and $\bar{G}_k^* = B^{-1}\sum_{b=1}^{B}\widehat{G}_k^*(b)$.

For a confidence level of $1-\alpha$, the percentile bootstrap confidence interval is defined (Qin et al., 2010) as:

$$CI_{k;Bp} = \left[\widehat{G}_{k(\alpha/2)}^*, \widehat{G}_{k(1-\alpha/2)}^*\right],$$

where $\widehat{G}_{k(\alpha)}^*$ denotes the αth quantile of the bootstrapped coefficients $\widehat{G}_k^*(b)$.

Jackknife method. The jackknife estimates suggested by Ogwang (2000) are defined by:

$$\widehat{G}_{k;-i} = \widehat{G}_k + \frac{2}{n\bar{x} - x_{(i)}}\left[\frac{x_{(i)}\sum_{j=1}^{n} j x_{(j)}}{n^2\bar{x}} + \frac{\sum_{j=1}^{n} j x_{(j)}}{n(n-1)} - \frac{n\bar{x} - \sum_{j=1}^{i} x_{(j)} + i x_{(i)}}{n-1}\right] - \frac{1}{n(n-1)},$$

from which the variance of $\widehat{G}_k$ based on the jackknife method, $\widehat{var}_J(\widehat{G}_k)$, can be estimated using the expression:

$$\widehat{var}_J(\widehat{G}_k) = \frac{n-1}{n}\sum_{i=1}^{n}\left(\widehat{G}_{k;-i} - \bar{G}_{kJ}\right)^2,$$

where $\bar{G}_{kJ} = n^{-1}\sum_{i=1}^{n}\widehat{G}_{k;-i}$.

In the case of the jackknife method, we consider two different confidence intervals: one based on the Gaussian approximation and another based on the studentized bootstrap. The $(1-\alpha)\%$ Gaussian-based confidence interval is defined by:

$$CI_{k;Jz} = \left[\hat{G}_k - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_J(\hat{G}_k)}, \hat{G}_k + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_J(\hat{G}_k)}\right],$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard Normal distribution.

The sampling distribution of the statistic $\hat{G}_k$ is skewed, so it could be theoretically improved by replacing the quantiles of this Gaussian-based confidence interval with the corresponding quantiles computed from the studentized bootstrap. This approach is expected to yield more reliable confidence intervals. The confidence interval based on the studentized bootstrap is given by:

$$CI_{k;Jt} = \left[\hat{G}_k - t^*_{J,1-\frac{\alpha}{2}}\sqrt{\widehat{var}_J(\hat{G}_k)}, \hat{G}_k - t^*_{J,\frac{\alpha}{2}}\sqrt{\widehat{var}_J(\hat{G}_k)}\right],$$

where $t^*_{J,\alpha}$ is the αth quantile of the values;

$$t^*_J(b) = \frac{\hat{G}^*_k(b) - \hat{G}_k}{\sqrt{\widehat{var}_J\left(\hat{G}^*_k(b)\right)}}.$$

Linearization method. The variance estimator of $\hat{G}_k$ based on the linearization method is given (Berger, 2008) by:

$$\widehat{var}_L(\hat{G}_k) = \frac{1}{n(n-1)}\sum_{i=1}^{n}(z_{ki} - \bar{z}_k)^2,$$

where

$$z_{ki} = \frac{1}{\bar{x}}\left[2x_i\hat{F}_n(x_i) - (\hat{G}_k + 1)(x_i + \bar{x}) + \frac{2}{n}\sum_{j=1}^{n} x_j\delta(x_j \geq x_i)\right]$$

and $\bar{z}_k = n^{-1}\sum_{i=1}^{n} z_{ki}$.

The $(1-\alpha)\%$ confidence interval is given by:

$$CI_{k;L} = \left[\hat{G}_k - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_L(\hat{G}_k)}, \hat{G}_k + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_L(\hat{G}_k)}\right],$$

OLS method. Despite the inappropriateness of using the OLS method for variance estimation, a confidence interval can be calculated in a manner similar to the other methods, using the estimate $\widehat{var}_{OLS}(\hat{G}_5)$, as defined in equation (11). Specifically, for a confidence level of $1-\alpha$, a confidence interval based on the OLS method is given by:

$$CI_{k;OLS} = \left[\hat{G}_k - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{OLS}(\hat{G}_5)}, \hat{G}_k + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{OLS}(\hat{G}_5)}\right]$$

Empirical likelihood method. Finally, we consider confidence intervals for the Gini index based on the empirical likelihood (EL) method (Qin et al, 2010), which employs a distinct conceptual approach. Specifically, EL confidence intervals are constructed by profiling the EL ratio statistic. This interval includes all values for which the null hypothesis of equality with the estimated index is not rejected at the α significance level. The theoretical EL ratio confidence interval may have under-coverage problems for small and moderate samples. For this reason, in this paper, we compute it using the bootstrap-calibrated EL ratio confidence interval, as defined by Equation (16) in Qin et al. (2010).

SM2.2. Samples derived from a finite population

Although we do not assess estimators for variances and confidence intervals in our research on finite populations (this is proposed as a future research direction in Section 5), we present estimators for that context for completeness. Specifically, for samples drawn from a finite population using a complex sampling design, various estimators have been suggested in the literature, including those based on the rescaled bootstrap, jackknife, and linearization methods.

Rescaled bootstrap method. The rescaled bootstrap (Rao et al., 1992; Rust and Rao, 1996; Berger and Gedik-Balay, 2020) can be used for estimating the variance of the Gini index as well as for constructing confidence intervals.

For the non-bias-corrected estimator $\hat{G}_{wk}$, with $k = 1, \ldots, 4$, (see Section 2.2), a variance estimator based on the rescaled bootstrap is defined as:

$$\widehat{var}_B(\hat{G}_{wk}) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{G}_{wk}^*(b) - \bar{G}_{wk}^*\right)^2,$$

where the bootstrapped coefficient $\hat{G}_{wk}^*(b)$, with $b = 1, \ldots, B$, represents the estimator $\hat{G}_{wk}$ obtained by substituting in it the original weights w_i with the bootstrap w_i^* weights, given by;

$$w_i^* = w_i \frac{r_i n}{n-1},$$

where r_i denotes the number of times the ith individual in the sample is selected by the bootstrap method.

A $(1-\alpha)\%$ confidence interval based on the rescaled bootstrap and the percentile approach is given by:

$$CI_{wk;Bp} = \left[\hat{G}_{wk(\alpha/2)}^*, \hat{G}_{wk(1-\alpha/2)}^*\right],$$

where $\hat{G}_{wk(\alpha)}^*$ is the αth quantile of the bootstrapped coefficients $\hat{G}_{wk}^*(b)$.

Jackknife method. The jackknife variance estimators can be computed in terms of pseudo-values z_{ki}, allowing for the construction of various types of variance estimators: Horvitz-Thompson, Sen-Yates-Grundy and Hartley-Rao.

The Horvitz-Thompson (HT) type variance estimator (Horvitz and Thompson, 1952) is given by:

$$\widehat{var}_{HT}\left(\hat{G}_{wk}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\breve{\Delta}_{ij} w_i w_j z_{ki} z_{kj}, \qquad \text{(i)}$$

where

$$\breve{\Delta}_{ij} = \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}},$$

and $\pi_{ij} = P(\{i,j\} \in S)$ is the second (joint) inclusion probability for the ith and jth individuals in the sample. Note that this variance estimator may give negative values. When the values π_{ij} are unknown, the Hàjek (1964) approximation can be used, which is estimated through:

$$\pi_{ij} \cong \pi_i \pi_j \left[1 - \frac{(1-\pi_i)\left(1-\pi_j\right)}{\sum_{k=1}^{n}(1-\pi_k)}\right].$$

This approximation is suggested for large-entropy sampling designs when both sample and population sizes are large (Haziza et al., 2008).

The Sen-Yates-Grundy type variance estimator, which is suitable for fixed-size sampling designs, is defined (Sen, 1953; Yates and Grundy, 1953) as:

$$\widehat{var}_{SYG}\left(\hat{G}_{wk}\right) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\breve{\Delta}_{ij}\left(w_i z_{ki} - w_j z_{kj}\right)^2.$$

Finally, the Hartley-Rao type variance estimator (Hartley and Rao, 1962) can also be computed for estimating the variance of the Gini index:

$$\widehat{var}_{HR}\left(\hat{G}_{wk}\right) = \frac{1}{n-1}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j<i}}^{n}\left(1 - \pi_i - \pi_j + \frac{1}{n}\sum_{k=1}^{N}\pi_k^2\right)\left(w_i z_{ki} - w_j z_{kj}\right)^2.$$

All these three estimators depend on the pseudo-values derived from the jackknife method, which are defined (Berger, 2008) as:

$$z_{ki} = \frac{1}{w_i}\left(1 - \frac{w_i}{\hat{N}}\right)\left(\hat{G}_{wk} - \hat{G}_{wk;-i}\right), \qquad \text{(ii)}$$

where the jackknife estimate $\hat{G}_{wk;-i}$ is the estimator $\hat{G}_{wk}$ computed from the sample observations $\left\{x_j : j \in S\right\}$ after removing the ith unit.

The $(1-\alpha)\%$ confidence interval based on the jackknife technique with the Normal approximation and the HT type variance estimator is given by:

$$CI_{wk;Jz.HT} = \left[\hat{G}_{wk} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{J.HT}\left(\hat{G}_{wk}\right)}, \hat{G}_{wk} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{J.HT}\left(\hat{G}_{wk}\right)}\right],$$

where $\widehat{var}_{J.HT}\left(\hat{G}_{wk}\right)$ is the variance estimator defined by the equation (i) with the pseudo-values described in equation (ii). Similarly, The Sen-Yates-Grundy and the Hartley-Rao type variance estimators can be used for estimating a confidence interval of the Gini index.

Linearization method. In the same manner as the jackknife variance estimators, the linearization method is based on pseudo-values. In particular, Langel and Tillé (2013) propose as pseudo-values:

$$z_{ki} = \frac{1}{\hat{N}^2 \bar{x}_w}\left[2\hat{N}_{(i)}\left(x_{(i)} - \hat{\bar{X}}_{(i)}\right) + \hat{N}\left\{\bar{x}_w - x_{(i)} - \hat{G}_{wk}\left(\bar{x}_w + x_{(i)}\right)\right\}\right], \qquad \text{(iii)}$$

where $\hat{N}_{(i)} = \sum_{j=1}^{i} w_{(j)}$, $w_{(i)}$, with $i \in S$, are the values w_i sorted according to the increasing order of the sample values x_i and

$$\hat{\bar{X}}_{(i)} = \frac{1}{\hat{N}_{(i)}}\sum_{j=1}^{i} w_{(j)} x_{(j)}.$$

A variance estimator based on the linearization method and the HT type variance estimator is defined as:

$$\widehat{var}_{L.HT}\left(\hat{G}_{wk}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \breve{\Delta}_{ij} w_i w_j z_{ki} z_{kj},$$

where z_{ki} are the pseudo-values defined in equation (iii).

The corresponding $(1-\alpha)\%$ confidence interval based on the Normal approximation is given by:

$$CI_{wk;Lz.HT} = \left[\hat{G}_{wk} - z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{L.HT}\left(\hat{G}_{wk}\right)}, \hat{G}_{wk} + z_{1-\frac{\alpha}{2}}\sqrt{\widehat{var}_{L.HT}\left(\hat{G}_{wk}\right)}\right].$$

Finally, it is important to mention the confidence intervals recently proposed by Berger and Gedik-Balay (2020), who suggest using both bootstrap and empirical likelihood methods to construct confidence intervals based on the bias-corrected estimator $\hat{G}_{w5}^{bc}$. Readers are referred to their paper for a detailed definition of these recent confidence intervals.

SM3. Additional results from the Monte Carlo simulation studies.

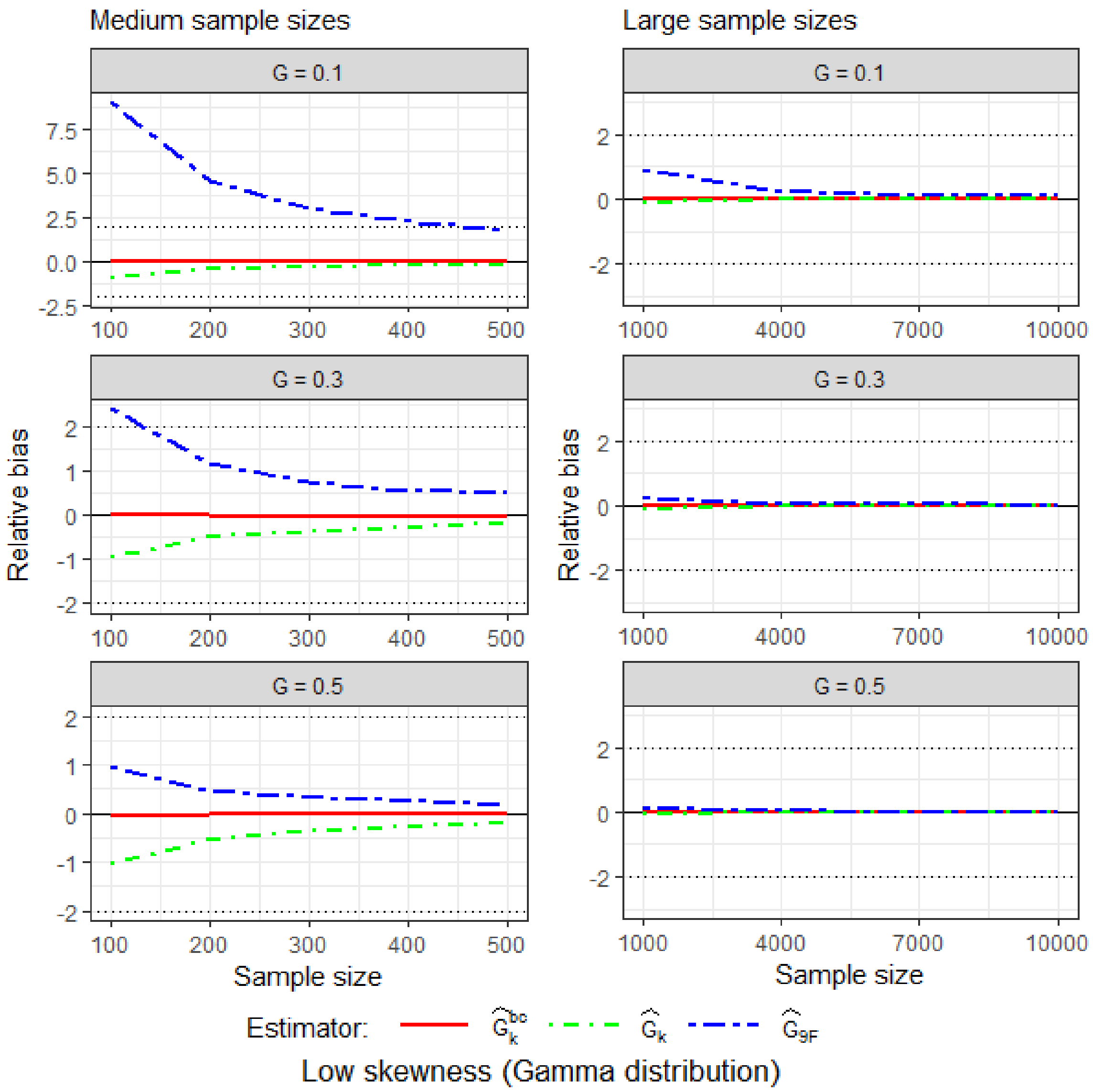

***Figure SM1**. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, with $k = 1, \dots, 10$, (see Table 2) and $\hat{G}_{9F}$ (see equation (8)). Samples, with sizes ranging from 100 to 10000, are drawn from a Gamma distribution (low skewness), with shape parameters chosen such that* $G \in \{0.1, 0.3, 0.5\}$. *Solid horizontal lines indicate unbiasedness:* $RB_G = 0$. *Dashed lines at* $RB_G \in \{-2, 2\}$ *delimit the areas of negligible bias:* $|RB_G| \leq 2$.

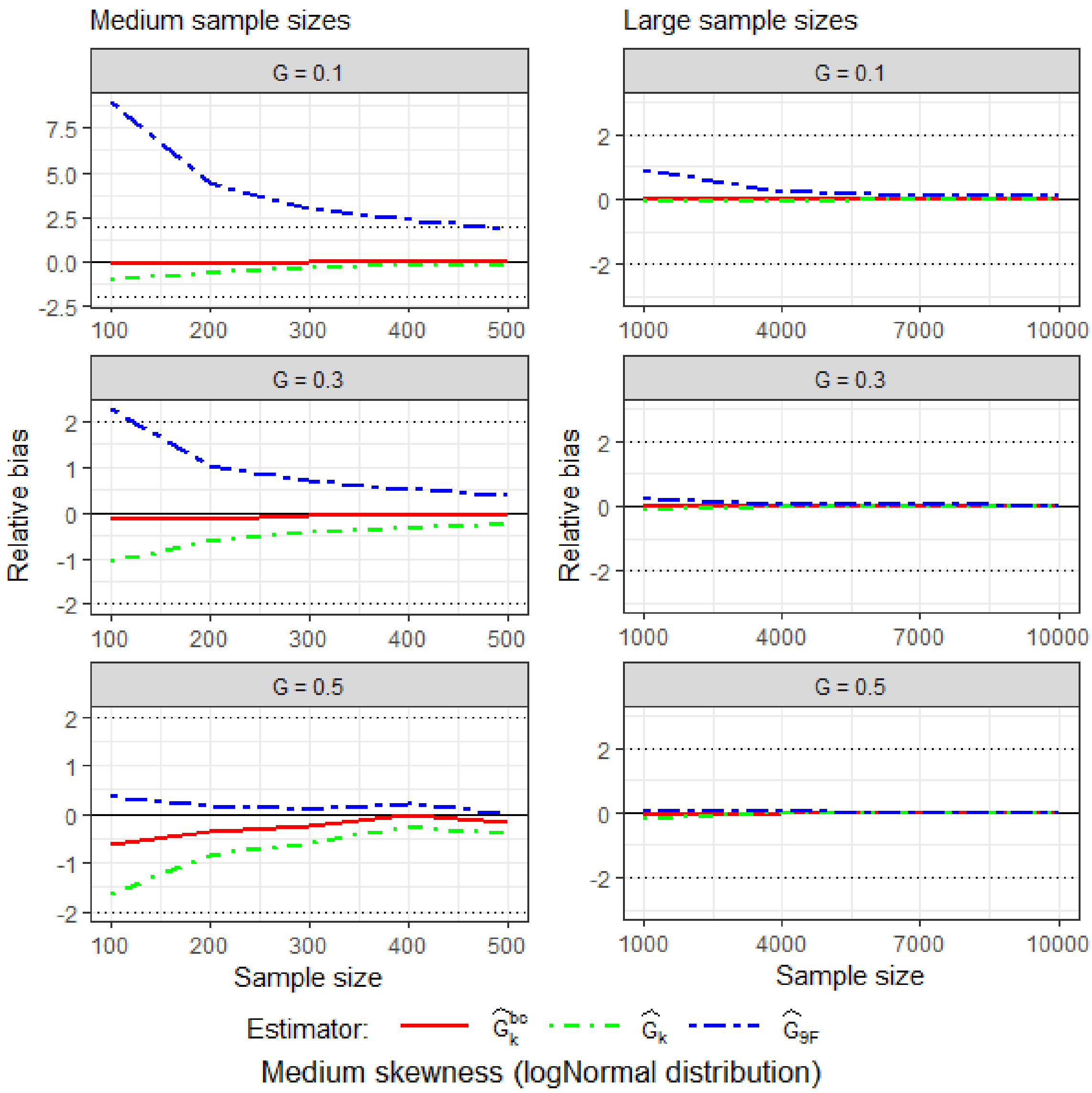

Figure SM2*. Relative biases (RB_G) for estimators $\hat{G}_k^{bc}$ and $\hat{G}_k$, with $k = 1, \dots, 10$, (see Table 2) and $\hat{G}_{9F}$ (see equation (8)). Samples, with sizes ranging from 100 to 10000, are drawn from a logNormal distribution (medium skewness), with shape parameters chosen such that $G \in \{0.1, 0.3, 0.5\}$. Solid horizontal lines indicate unbiasedness: $RB_G = 0$. Dashed lines at $RB_G \in \{-2, 2\}$ delimit the areas of negligible bias: $|RB_G| \leq 2$.*

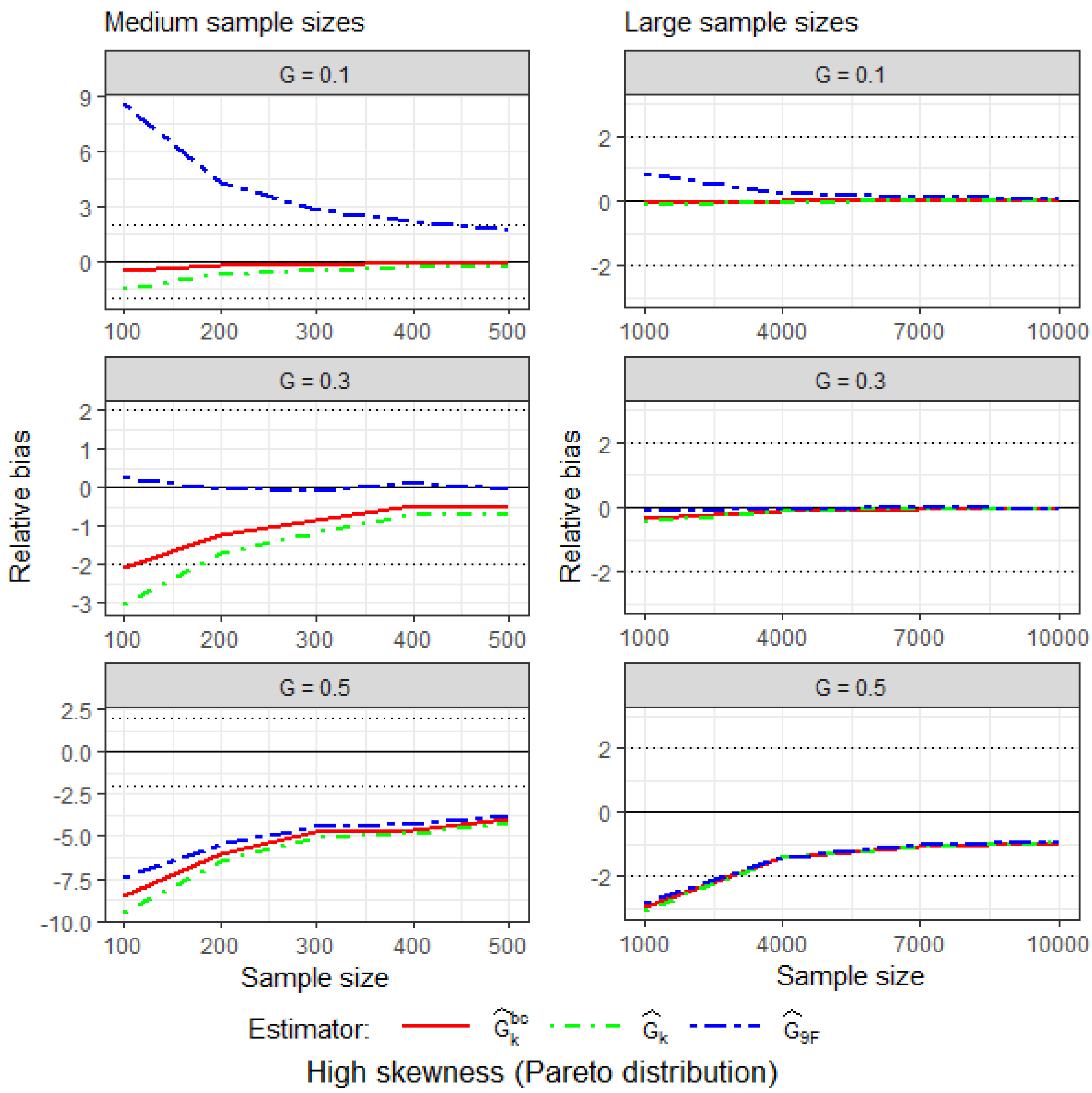

Figure SM3. *Relative biases* (RB_G) *for estimators* $\widehat{G}_k^{bc}$ *and* $\widehat{G}_k$, *with* $k = 1, \ldots, 10$, *(see Table 2) and* $\widehat{G}_{9F}$ *(see equation (8)). Samples, with sizes ranging from 100 to 10000, are drawn from a Pareto distribution (high skewness), with shape parameters chosen such that* $G \in \{0.1, 0.3, 0.5\}$. *Solid horizontal lines indicate unbiasedness:* $RB_G = 0$. *Dashed lines at* $RB_G \in \{-2, 2\}$ *delimit the areas of negligible bias:* $|RB_G| \leq 2$.

References

Alfons, A., Templ, M., & Filzmoser, P. (2013). Robust estimation of economic indicators from survey samples based on Pareto tail modelling. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 62(2), 271-286.

Berger, Y.G. (2008) A note on the asymptotic equivalence of jackknife and linearization variance estimation for the Gini Coefficient. *Journal of Official Statistics*, 24(4), 541-555.

Berger, Y. G., & Gedik-Balay, İ. (2020). Confidence intervals of Gini coefficient under unequal probability sampling. *Journal of Official Statistics*, 36(2), 237-249.

Cowell, F. A. (1989). Sampling variance and decomposable inequality measures. *Journal of Econometrics*, 42(1), 27-41.

Davidson, R. (2009). Reliable inference for the Gini index. *Journal of Econometrics*, 150(1), 30-40.

Davidson, R. (2012). Statistical inference in the presence of heavy tails. *The Econometrics Journal*, 15(1), C31-C53.

Deltas, G. (2003). The small-sample bias of the Gini coefficient: results and implications for empirical research. *Review of Economics and Statistics*, 85(1), 226-234.

Fontanari, A., Taleb, N. N., & Cirillo, P. (2018). Gini estimation under infinite variance. *Physica A: Statistical Mechanics and its Applications*, 502, 256-269.

Giles, D. E. (2004). Calculating a standard error for the Gini coefficient: some further results. *Oxford Bulletin of Economics and Statistics*, 66(3), 425-433.

Giles, D., E. (2006). A cautionary note on estimating the standard error of the Gini index of inequality: Comment. *Oxford Bulletin of Economics and Statistics*, 68(3), 395-396.

Goga, C., & Ruiz-Gazen, A. (2014). Efficient estimation of non-linear finite population parameters by using non-parametrics. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(1), 113-140.

Hàjek, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *The Annals of Mathematical Statistics*, 35(4), 1491-1523.

Hartley, H. O., & Rao, J. N. K. (1962). Sampling with unequal probabilities and without replacement. *The Annals of Mathematical Statistics*, 350-374.

Haziza, D., Mecatti, F. & Rao, J. N. K. (2008). Evaluation of some approximate variance estimators under the Rao-Sampford unequal probability sampling design. *Metron*, LXVI, 91-108.

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663-685.

Karagiannis, E., & Kovacevic, M. (2000). A method to calculate the Jackknife variance estimator for the Gini coefficient. *Oxford Bulletin of Economics & Statistics*, 62(1), 119-122.

Langel, M., & Tilé, Y. (2011). Corrado Gini, a pioneer in balanced sampling and inequality theory. *Metron*, 69, 45-65.

Langel, M., & Tillé, Y. (2013). Variance estimation of the Gini index: Revisiting a result several times published. *Journal of the Royal Statistical Society, Series A*, 176(2), 521-540.

Lerman, R. I., & Yitzhaki, S. (1989). Improving the accuracy of estimates of Gini coefficients. *Journal of Econometrics*, 42(1), 43-47.

Modarres, R., & Gastwirth, J. L. (2006). A cautionary note on estimating the standard error of the Gini index of inequality. *Oxford Bulletin of Economics and Statistics*, 68(3), 385-390.

Muñoz, J.F., Pavía, J.M., & Álvarez-Verdejo, E. (2024). *giniVarCI: Gini Indices, Variances and Confidence Intervals for Finite and Infinite Populations*. R packages version 0.0.1-3. https://CRAN.R-project.org/package=giniVarCI.

Nygård, F., & Sandström, A. (1985). The estimation of the Gini and the entropy inequality parameters in finite populations. *Journal of Official Statistics*, 1(4), 399-412.

Nygård, F., & Sandström, A. (1989). Income inequality measures based on sample surveys. *Journal of Econometrics*, 42(1), 81-95.

Ogwang, T. (2000). A convenient method of computing the Gini index and its standard error. *Oxford Bulletin of Economics & Statistics*, 62(1), 123-129.

Ogwang, T. (2004). Calculating a standard error for the Gini coefficient: Some further results: reply. *Oxford Bulletin of Economics & Statistics*, 66(3), 435-437.

Ogwang, T. (2006). A cautionary note on estimating the standard error of the Gini index of inequality: Comment. *Oxford Bulletin of Economics and Statistics*, 68(3), 391-393.

Peng, L. (2011). Empirical likelihood methods for the Gini index. *Australian & New Zealand Journal of Statistics*, 53(2), 131-139.

Qin, Y., Rao, J.N.K., & Wu, C. (2010). Empirical likelihood confidence intervals for the Gini measure of income inequality. *Economic Modelling*, 27(6), 1429-1435.

Rao, J. N. K., Wu, C. F. J., & Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology*, 18(2), 209-217.

Rust, K. F., & Rao, J. N. K. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5(3), 283-310.

Sen, A. R. (1953). On the estimate of the variance in sampling with varying probabilities. *Journal of the Indian Society of Agricultural Statistics*, 5, 119-127.

Wang, D., Zhao, Y., & Gilmore, D. W. (2016). Jackknife empirical likelihood confidence interval for the Gini index. *Statistics & Probability Letters*, 110, 289-295.

Yates, F., & Grundy, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society, Series B*, 15, 253-261.