UvA ECONOMETRICS

Discussion Paper: 2006/06

Edgeworth expansions and normalizing transforms for inequality measures

Kees Jan van Garderen and Christian Schluter

www.fee.uva.nl/ke/UvA-Econometrics

Amsterdam School of Economics
Department of Quantitative Economics
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Edgeworth Expansions and Normalizing Transforms for Inequality Measures*

Kees Jan van Garderen† University of Amsterdam

Christian Schluter‡ University of Southampton

November 2006

Abstract

Finite sample distributions of studentized inequality measures differ substantially from their asymptotic normal distribution in terms of location and skewness. We study these aspects formally by deriving the second order expansion of the first and third cumulant of the studentized inequality measure. We state distribution-free expressions for the bias and skewness coefficients. In the second part we improve over first-order theory by deriving Edgeworth expansions and normalizing transforms. These are designed to eliminate the second order term in the distributional expansion of the studentized transform. The resulting finite sample distributions are shown to be much closer to the Gaussian limit distribution than the distributions of the studentized inequality measure.

Keywords: Generalized Entropy inequality measures, higher order expansions, normalizing transformations.

JEL classification: C10, C14, D31, D63, I32

*Funding from the ESRC under grant R000223640 is gratefully acknowledged. The research of van Garderen has been supported by a research fellowship of the Royal Dutch Academy of Arts and Sciences. Helpful comments from Andrew Chesher, Russell Davidson, Grant Hillier, Patrick Marsh, and Richard Smith are gratefully acknowledged.

†Corresponding author. Department of Quantitative Economics, Faculty of Economics and Econometrics, University of Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, The Netherlands, Tel. +31-20-525 4220, Fax: +31-20-525 4349, Email: K.J.vanGarderen@uva.nl

‡Department of Economics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. Tel. +44 (0)2380 59 5909, Fax. +44 (0)2380 59 3858. Email: C.Schluter@soton.ac.uk. http://www.economics.soton.ac.uk/staff/schluter/

1 Introduction

Most attention in the statistical literature on inequality measures has focused on the asymptotic properties of their estimators (see e.g. Cowell, 1989, Thistle, 1990, Davidson and Duclos, 1997). Their finite sample properties have rarely been considered. Exceptions are, for instance, Mills and Zandvakili (1997) and Biewen (2001) who investigate bootstrap inference, and Maasoumi and Theil (1979) who develop small-sigma approximations. We consider finite sample properties of Generalized Entropy (GE) indices of inequality, which constitute a leading class of inequality indices since it is the only class that simultaneously satisfies the key properties of anonymity and scale independence, and the principles of transfer, decomposability, and population (see e.g. Maasoumi 1999 or Cowell 2000). Studies of industrial concentration, or income studies after decomposition into population subgroups, or cross country comparisons based on macro data, can easily yield samples of the sizes considered here.

Even for relatively large samples we show that standard first order theory provides poor guidance for actual behavior. The distribution of the studentized inequality measure differs substantially from the Gaussian limit in terms of location and skewness. We study the bias (average deviation from zero) and skewness formally in the first part of this paper by deriving the second order expansions of the first three cumulants. We refer to the resulting coefficients of $n^{-1/2}$ as bias and skewness coefficients. This is the first key contribution of this paper. Moreover, the bias and skewness coefficients can be estimated non-parametrically using sample moments without affecting the order of the approximation. In all applications considered below, it is shown that the bias and skewness coefficients times $n^{-1/2}$ are substantial compared to the limit values of zero.

Having analyzed these departures from normality of the finite sample distribution of the studentized inequality index, we turn to potential corrections in the second part of the paper. These corrections are based on considering the second order term in the distributional expansion, which is a function of the bias and skewness coefficient derived in the first part of the paper. We consider two approaches. Edgeworth expansions directly adjust the asymptotic approximation by including the $O\left(n^{-1/2}\right)$ term, whereas normalizing transformations of the inequality measure are nonlinear transformations designed to annihilate this term asymptotically. Edgeworth expansions can suffer from negativity of the density and oscillations in the tails and we show that this is indeed a problem for standardized inequality measures.

The focus of the second part of the paper is therefore on normalizing transforms. The second key contribution of this paper is the derivation of normalizing transforms for GE inequality measures. First, we show that the skewness coefficient of a standardized non-linear transform of the inequality measure is zero if the transform satisfies a crucial differential equation. We further derive the bias coefficient of this transform, so that we obtain a bias-corrected transform which yields the desired asymptotic refinement. Second, we use this general result to compute the normalizing transform for various income distributions and sensitivity parameters of the inequality index, and study their finite sample distributions. We show that these are indeed closer to the Gaussian limit distribution.

The organization of this paper is as follows. Section 2 states the class of inequality measures, considers estimation, and states the first order (Normal) approximation. The quality of this Normal approximation in finite samples is studied via simulations in Section 3. We consider quantiles of the actual density of the studentized inequality index, and we illustrate the consequences of the departure from normality for inference. In Section 4 we study the problems of bias and skewness formally by deriving the bias and skewness coefficients. These enable us to give the Edgeworth expansion for the GE indices. We proceed to study bias and skewness coefficients for specific income distributions and sensitivity parameters of the inequality index. The behavior of the Edgeworth expansion is illustrated and we derive the normalizing transforms in Section 5, and study their behavior. Section 6 concludes and the proofs are collected in the Appendix.

2 Generalized Entropy Indices of Inequality

We consider the popular and leading class of inequality indices, the GE indices. These are of particular interest because it is the only class of inequality measures that simultaneously satisfies the key properties of anonymity and scale independence, the principles of transfer and decomposability, and the population principle. For an extensive discussion of the properties of the GE index see Cowell (1977, 1980, 2000).

The class of indices is defined for any real α by

$$I(\alpha; F) = \begin{cases} \frac{1}{\alpha^2-\alpha}\left[\frac{\mu_\alpha(F)}{\mu_1(F)^\alpha} - 1\right] & \text{for } \alpha \neq \{0,1\} \\ -\int log(\frac{x}{\mu_1(F)})\ dF(x) & \text{for } \alpha = 0 \\ \int \frac{x}{\mu_1(F)} log(\frac{x}{\mu_1(F)})\ dF(x) & \text{for } \alpha = 1 \end{cases} \tag{1}$$

where α is a sensitivity parameter, F is the income distribution, and $\mu_\alpha(F) = \int x^\alpha dF(x)$ is the moment functional, and we will assume incomes to be positive. The index is continuous in α. The larger the parameter α, the larger is the sensitivity of the inequality index to the upper tail of the income distribution. It is not monotonic in α, however.

GE indices constitute a large class which nests some popular inequality measures popular as special cases. If $\alpha = 2$ the index is known as the (Hirschman-)Herfindahl index and equals half the coefficient of variation squared. Herfindahl's index plays an important role as measure of concentration in industrial organization and merger decisions (see e.g. Hart, 1971). In empirical work on income distributions this value of α is considered large. Two other popular inequality measures are the so-called Theil indices, which are the limiting cases $\alpha = 0$ and $\alpha = 1$ (Theil, 1967). Finally, the Atkinson (1970) index is ordinally equivalent to the GE index.

Although the index is defined for any real value of α, in practice only values between 0 and 2 are used and we confine our examination to this range. The limiting cases 0 and 1 are treated implicitly below since all key quantities are continuous in α.

2.1 Estimation and Normal Approximations

In empirical work the inequality measure I needs to be estimated from a sample of incomes denoted X_i, $i = 1, \ldots, n$. We follow standard practice and assume that incomes are independently and identically distributed with distribution F and are positive. The measure I is a functional, mapping income distributions into scalars. The commonly used estimator simply uses the empirical distribution function (EDF) $\widehat{F}(x) = n^{-1} \sum_i \mathbf{1}_{(-\infty,x)}\left(X_i\right)$, where $\mathbf{1}_{(-\infty,x]}\left(.\right)$ denotes the indicator function on the open interval smaller than or equal to x,

$$\widehat{I} = I(\widehat{F}).$$

Since I is a function of moments, the EDF-estimator is also referred to in the literature as the method of moments estimator. It is standard practice to obtain the asymptotic variance $\sigma^2 = aVar(n^{1/2}(\widehat{I} - I))$ by the delta method, yielding

$$\sigma^2 = \frac{1}{\left(\alpha^2 - \alpha\right)^2} \frac{1}{\mu_1^{2\alpha+2}} \left[\mu_1^2 \mu_{2\alpha} + \alpha^2 \mu_\alpha^2 \mu_2 - 2\alpha \mu_\alpha \mu_1 \mu_{\alpha+1} - \left(1-\alpha\right)^2 \mu_\alpha^2 \mu_1^2\right],$$

and to estimate it by an EDF-based estimator, denoted $\widehat{\sigma}^2$.

Inference about the population value I is then based on the studentized measure, defined as

$$S = n^{1/2} \left(\frac{\widehat{I} - I}{\widehat{\sigma}} \right). \tag{2}$$

By standard central limit arguments, S has a distribution that converges asymptotically to the Gaussian distribution (see, inter alia, Cowell, 1989, or Thistle, 1990). Denoting the Gaussian distribution and density by Φ and ϕ respectively, then using order notation, we have

$$\Pr\left(S \leq x\right) = \Phi\left(x\right) + O\left(n^{-1/2}\right). \tag{3}$$

Standard, first order, inference methods only use the first term $\Phi\left(x\right)$ in this approximation. In the next section we examine in experiments how well $\Pr\left(S \leq x\right)$ is approximated by $\Phi\left(x\right)$ in finite samples. In the light of our negative findings, we will then consider the $O\left(n^{-1/2}\right)$ term explicitly.

3 Quality of the Normal Approximation

In this section we investigate how well the Normal approximation of $\Pr\left(S \leq x\right)$ performs in realistic settings for samples of varying sizes and income distributions that we will now describe.

3.1 Income Distributions

The experiments are based on three parametric income distributions which are regularly used to fit real real-world income data: the Gamma, the Lognormal, and the Singh-Maddala distribution. We use the common shorthand notation $G(r, \lambda)$, $LN(\mu, v^{1/2})$, and $SM(a, b, c)$ to refer to them. Generalized Entropy indices are scale invariant, and thus independent of the scale parameters λ, μ, and a for the G, LN, and SM distributions respectively. For notational convenience, we suppress the irrelevant scale parameters.

McDonald (1984) has shown that these three distributions are special cases of the Generalized Beta distribution of the second kind ($GB2$), whose density is given by

$$f(x; a, b, c, d) = \frac{bx^{bd-1}}{a^{bd} B(c, d) \left[1 + (x/a)^b\right]^{d+c}},$$

where $B(\cdot, \cdot)$ denotes the Beta function. In particular, SM has density $f(x; a, b, c, 1)$, G has density $\lim_{c \to \infty} f\left(x; c\lambda^{-1}, 1, c, r\right)$, and LN is a special case involving $c \to \infty$ and $b \to 0$. All three distributions are skewed to the right, but differ in other ways, such as their tail behavior. Schluter and Trede (2002), for instance, show that the right tail of the generalized beta distribution can be written as $1 - F(x; a, b, c, d) = g_1 x^{-bc}\left(1 + g_2 x^{-b} + O\left(x^{-2b}\right)\right)$ for some constants g_1 and g_2 and x large. It follows that SM has a heavy right tail which decays like a power function (with right tail index equal to bc). G and LN decay exponentially fast. The left tail of $GB2$ can be written as $F(x, a, b, c, d) = g_3 x^{bd}\left(1 + g_4 x^b + O\left(x^{2b}\right)\right)$ for some constants g_3 and g_4 and x small. The moments of the distributions are stated in McDonald (1984).

The population inequality index specializes for the different income distributions to:

$$\begin{array}{ll}
\text{Gamma :} & I(\alpha; r) = (\alpha^2 - \alpha)^{-1}\left[r^{-\alpha}\Gamma(\alpha + r)/\Gamma(r) - 1\right], \\
\text{Lognormal:} & I(\alpha; v) = (\alpha^2 - \alpha)^{-1}\left[\exp\left(\frac{1}{2}v(\alpha - 1)\alpha\right) - 1\right], \\
\text{Singh-Maddala:} & I(\alpha; b, c) = (\alpha^2 - \alpha)^{-1}\dfrac{c^{-(\alpha-1)}B(1 + \alpha/b, c - \alpha/b)}{B(1 + 1/b, c - 1/b)^{\alpha} - 1}, \quad bc > \alpha.
\end{array}$$

We focus on these three income distributions not only because they are quite different, but more importantly because they are regularly used to fit actual real-world income data.

For instance, Brachmann et al. (1996) estimate the distributional parameters on German income data for the 1980s and early 1990s. For G they report point estimates $r \in [3.4, 4]$, for LN $v \in [0.28, 0.31]$, and for SM $b \in [2.7, 2.9]$ and $c \in [1.6, 2.1]$. Singh and Maddala (1976) report point estimates of $b \in [1.9, 2.1]$, $c \in [2.5, 3]$ for US income data from the 1960s. For US income data from the 1970s McDonald (1984) reports $r = 2.3$, $v \in [0.48, 0.51]$, $b \in [2.9, 3.76]$ and $c \in [1.8, 2.9]$. For the Lognormal Kloek and van Dijk (1978) find $v \in [0.21, 0.54]$ for different groups of income earners using 1973 Dutch data. Across these studies, we have $r \in [2.3, 4]$, $v \in [0.28, 0.54]$, $b \in [1.9, 3.76]$, and $c \in [1.6, 3]$. Throughout this paper, we use parameter values in similar ranges. For example, we use $r = 3$, $\nu = 0.1$ and 0.49, $b = 3.5$ and 2.7 and $c = 1.7$ and 3.

For further recent examples see Bandourian et al. (2003) who fit these income distribution models for a large number of countries, including the USA, Canada, Taiwan and most European countries for the period 1969-1997.

3.2 Quantiles and Confidence Levels for Studentized Inequality Indices

In the first experiment, we simulate the quantiles of the finite sample distribution $\Pr(S \leq x)$ and compare these to the Gaussian quantiles. We find substantial bias and skewness. In order to assess the consequences of the departure from normality, we turn to inferential procedures in the second experiment, and examine the extent to which the actual finite sample behavior of standard confidence intervals deviates from its nominal behavior.

In the first experiment we focus on two tail quantiles (0.025, 0.975) because of their role in inferential procedures, and three middle quantiles (0.25, 0.5, 0.75). We simulate these quantiles of I for the each income distribution, different sample sizes, and also vary the sensitivity parameter α of the inequality index in the range $[0, 2]$. The top panel of Figure 1 depicts the results. The Gaussian limit values are indicated on the right axes.

From the figure we can observe the following. First, the discrepancy between actual and Gaussian quantiles varies substantially across the distributions. It is the smallest for the Gamma case, and the worst for the Lognormal case. Second, the performance worsens across all income distributions as α increases. This is to be expected, since as α increases, the moments entering S and the expected finite sample bias increases. Third, across the five depicted quantiles, the extreme 0.025 quantile exhibits the largest deviation from the corresponding Gaussian quantile. All empirical quantiles lie below the corresponding Gaussian quantiles. This suggests that the actual distribution is biased, skewed to the left, and that the skewness increases in α. We focus extensively on this skewness in the next section. Fourth, the deviations decrease naturally as sample size increases, but the improvements are slow. The worst performer is the lower extreme quantile, relative to corresponding Gaussian quantile. Hence skewness is persistent even in fairly large samples.

In order to illustrate the consequences for inference of this substantial departure we have also examined the actual coverage behavior of standard confidence intervals derived from approximation (3). Setting $x = \Phi^{-1}(p)$, yields the standard one-sided confidence interval for I with lower confidence bound $\widehat{I} - \widehat{\sigma} n^{-1/2} \Phi^{-1}(p)$, and the usual nominal $p \times 100\%$ symmetric confidence intervals are given by

$$\widehat{I} - \frac{\widehat{\sigma}}{\sqrt{n}} \Phi^{-1}\left(\frac{1+p}{2}\right) \leq I \leq \widehat{I} - \frac{\widehat{\sigma}}{\sqrt{n}} \Phi^{-1}\left(\frac{1-p}{2}\right).$$

Note that the asymptotic coverage rate for one-sided confidence intervals is $\Pr(S \leq \Phi^{-1}(p)) = p + O\left(n^{-1/2}\right)$, and for symmetric confidence intervals, based on a standard symmetry argument, equals $p + O\left(n^{-1}\right)$.

The lower panel of Figure 1 depicts the actual coverage errors of symmetric two-sided confidence intervals for samples from the three income distributions of sizes 100,

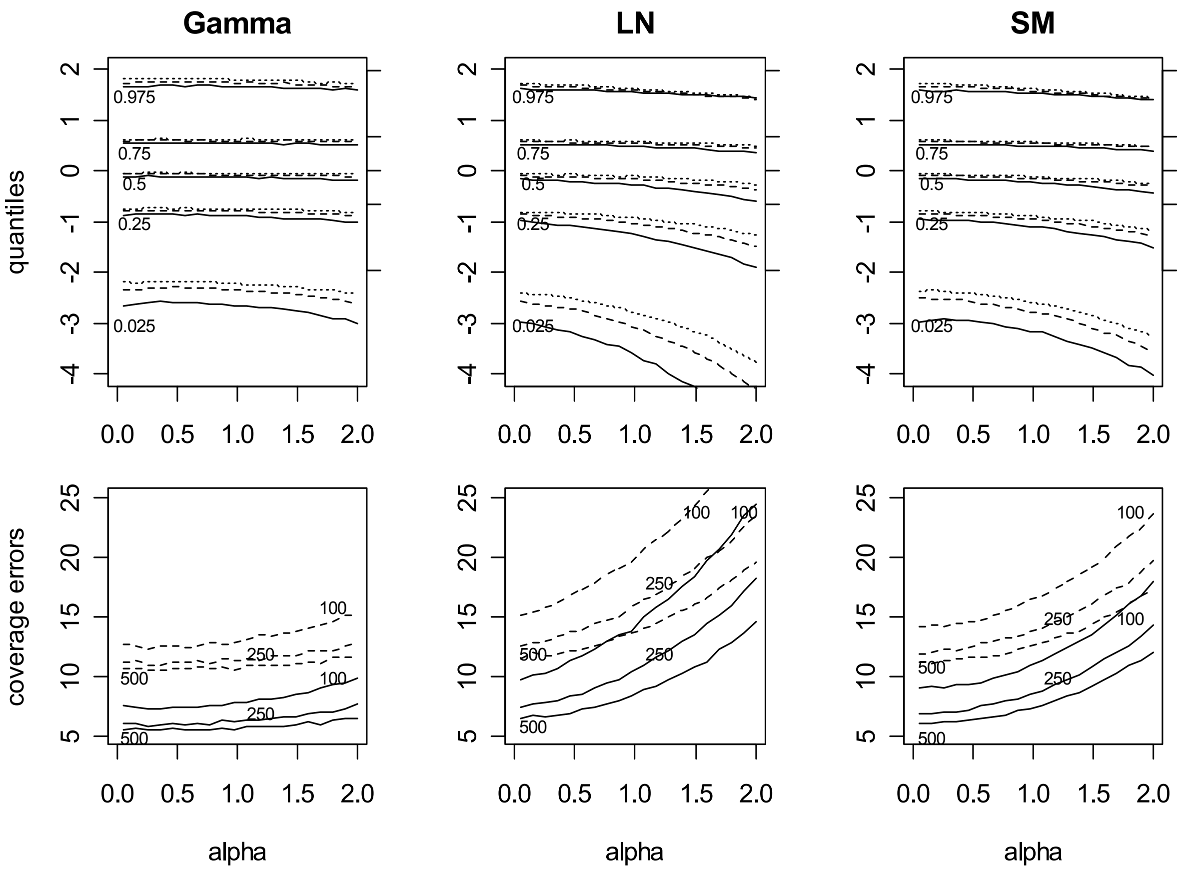

Figure 1: Quantiles and Coverage errors. Notes: The top panel depicts selected quantiles of the actual finite sample distributions, as a function of the sensitivity parameter α of the inequality measure. The ticks on the right axis are the Gaussian limit values. The solid line refers to samples of size 100, the dashed lined to sizes 250, and the dotted line to sizes 500. The bottom panel depicts the actual coverage errors of the standard two sided symmetric confidence intervals (CIs). The solid line refers to CIs with nominal 5%, the broken line to CIs with nominal 10% coverage errors. The sample sizes are again 100, 250, and 500. The distributions are $G(3,.)$, $LN(.,0.49)$, $SM(.,2.9,1.9)$. All simulations are based on $R = 10^6$ replications.

250, and 500 as α varies. The solid line refers to confidence intervals with nominal coverage errors of 5%, the broken line to nominal rates of 10%. We see that the actual error rates can be much larger than the nominal ones and the discrepancy increases with increasing α.

Performance also varies substantially across income distributions. For clarity, Table 1 shows explicit coverage error values for $\alpha = 2$. The worst performance here is for the Lognormal case where, for a sample size of 100, the coverage failure is almost 5 times larger than the nominal 5%. The best results obtain in the Gamma case, but even there the coverage rates are two times the nominal 5% when $n = 100$.

	$n = 50$	100	250	500
$G(3, \cdot)$	12.8	10.0	7.6	6.3
$LN(\cdot; 0.49)$	30.5	24.5	18.1	14.7
$SM(\cdot, 2.9, 1.9)$	21.4	17.8	14.2	12.3

Table 1: Actual coverage failure in percent of nominal 95% symmetric first order confidence intervals for the GE index with $\alpha = 2$. Based on 100,000 replications.

Changing parameter values of the income distribution does change the absolute and relative performance, however. Keeping $\alpha = 2$, and $n = 100$, we found from simulations not reported in this paper, that the performance improves in the Lognormal case as σ^2 decreases (although for $\sigma^2 = 0.25$ coverage failure is still 15.7%), worsens in the Singh-Maddala case as b and c decrease, (e.g. when they are decreased to $b = 2.8$ and $c = 1.7$, coverage failure worsens to 22.6%), and does not change very much in the Gamma case as r varies over the range 2.3 to 4, with coverage rates from 10.3% to 9.3% respectively. See Garderen and Schluter (2003) for further results and also Biewen (2001). The results in the next section help to understand these empirical findings.

4 Cumulants and Edgeworth Expansions

The simulation study in the previous section has shown that the Normal approximation suffers from substantial bias and skewness problems. In this section we study bias and skewness formally by considering expansions to second order of the first and third cumulant (assuming they exist) of the studentized inequality measures S. These expansions are given by

$$\begin{aligned} K_{S,1} &= n^{-1/2} k_{1,2} + O\left(n^{-3/2}\right), \\ K_{S,3} &= n^{-1/2} k_{3,1} + O\left(n^{-3/2}\right). \end{aligned} \tag{4}$$

The expansion of the second cumulant is $K_{S,2} = 1 + O(n^{-1})$.[1] A key contribution of our paper is the derivation of the bias and skewness coefficients $k_{1,2}$ and $k_{3,1}$ for the

[1]Due to the studentization of S, $K_{S,3}/K_{S,2}^{3/2} = n^{-1/2} k_{3,1} + O\left(n^{-1}\right)$, and $k_{3,1}$ is therefore also the coefficient of skewness.

examined inequality indices. These coefficients are the critical factors in the second order terms in the expansion of the cumulant generating function of S,

$$\frac{1}{2}t^2 + n^{-1/2}\left[t\ k_{1,2} + \frac{1}{6}t^3 k_{3,1}\right] + O\left(n^{-1}\right),$$

and the Edgeworth expansion

$$\Pr(S \leq x) = \Phi(x) - n^{-1/2}\left(k_{1,2} + \frac{1}{6}k_{3,1}(x^2-1)\right)\phi(x) + O(n^{-1}). \tag{5}$$

See e.g. Hall (1992) for an extensive discussion of Edgeworth expansions, who observes, for instance, that the right hand side of equation (5) does not necessarily converge as an infinite series. Regularity conditions for the validity of the expansion are also stated in Hall (1992, Section 2.4). The GE index is a smooth function of the moments with continuous third derivatives and $\mu_1 > 0$ since we assume incomes to be positive. This implies that Theorem 2.2 in Hall (1992) applies and hence we require that the income distribution for X satisfies the moment conditions $E\left(X^3\right) < \infty$ and $E\left(X^{3\alpha}\right) < \infty$ and that X has a proper density function (implying that Cramér's condition is satisfied). Note that these moment conditions restrict the admissible parameter values for the Singh-Maddala distribution ($bc > max(3, 3\alpha)$).

Normal approximations only consider the first order term, i.e. $\frac{1}{2}t^2$ or $\Phi(x)$, so the higher order term indicates deviation from normality. The coefficients $k_{1,2}$ and $k_{3,1}$ can also be used to predict when first order inference will be poor. In Section 5 they are used in the derivation and construction of the normalizing transform.

We first state explicitly the bias and skewness coefficients as a function of the sensitivity parameter α of the inequality measures nonparametrically, and proceed to study their behavior for the income distributions considered in Section 3.

4.1 Bias and Skewness Coefficients for the Studentized Inequality Measures

Proposition 1. *Assuming the expectations $E\left(X^3\right)$ and $E\left(X^{3\alpha}\right)$ exist, then the bias and skewness coefficients for the studentized inequality measures are given by*

$$\begin{aligned} k_{1,2} &= \left(B^{-1/2}M_2 - \frac{1}{2}B^{-3/2}M_5\right)\cdot\left(1 - 2\cdot\mathbf{1}_{(0,1)}\left(\alpha\right)\right), \\ k_{3,1} &= B^{-3/2}\left(M_4 + 6M_1M_3 - 3M_5\right)\cdot\left(1 - 2\cdot\mathbf{1}_{(0,1)}\left(\alpha\right)\right), \end{aligned}$$

where $\mathbf{1}_{(0,1)}$ *is the indicator function on the interval* $(0,1)$ *and*

$$
\begin{aligned}
B &= \mu_1^2\mu_{2\alpha} + \alpha^2\mu_\alpha^2\mu_2 - 2\alpha\mu_\alpha\mu_1\mu_{\alpha+1} - (1-\alpha)^2\mu_\alpha^2\mu_1^2, \\
M_1 &= \mu_1\mu_{\alpha+1} - \mu_1^2\mu_\alpha - \alpha\mu_\alpha\mu_2 + \alpha\mu_\alpha\mu_1^2, \\
M_2 &= \mu_{\alpha+1} - \mu_1\mu_\alpha - \frac{1}{2}\alpha(\alpha+1)\frac{\mu_\alpha}{\mu_1}\left(\mu_2 - \mu_1^2\right), \\
M_3 &= \mu_1\left(\mu_{2\alpha} - \mu_\alpha^2\right) - \frac{1}{2}\alpha(\alpha+1)\mu_\alpha\left(\mu_{\alpha+1} - \mu_1\mu_\alpha\right) - \alpha\mu_\alpha M_2, \\
M_4 &= \mu_1^3\left(\mu_{3a} - 3\mu_{2\alpha}\mu_\alpha + 2\mu_\alpha^3\right) \\
&\quad -3\alpha\mu_\alpha\mu_1^2\left(\mu_{2\alpha+1} - \mu_{2\alpha}\mu_1 - 2\mu_{\alpha+1}\mu_\alpha + 2\mu_\alpha^2\mu_1\right) \\
&\quad +3(\alpha\mu_\alpha)^2\mu_1\left(\mu_{\alpha+2} - 2\mu_{\alpha+1}\mu_1 + 2\mu_\alpha\mu_1^2 - \mu_\alpha\mu_2\right) \\
&\quad -(\alpha\mu_\alpha)^3\left(\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3\right), \\
M_5 &= 2\left(\mu_1^2\mu_{2\alpha} + \left(\alpha^2-\alpha\right)\mu_1\mu_\alpha\mu_{\alpha+1} - \alpha^3\mu_2\mu_\alpha^2 + (\alpha-1)(1-\alpha)^2\mu_1^2\mu_\alpha^2\right)\left(\mu_{\alpha+1} - \mu_1\mu_\alpha\right) \\
&\quad +\alpha^2\mu_1\mu_\alpha^2\left(\mu_{\alpha+2} - \mu_2\mu_\alpha\right) \\
&\quad +2\mu_1\left(\alpha^2\mu_\alpha\mu_2 - \alpha\mu_1\mu_{\alpha+1} - (1-\alpha)^2\mu_1^2\mu_\alpha\right)\left(\mu_{2\alpha} - \mu_\alpha^2\right) \\
&\quad +2\alpha^2\mu_1\mu_\alpha^2\left(\mu_{\alpha+2} - \mu_1\mu_{\alpha+1}\right) - 2\alpha\mu_1^2\mu_\alpha\left(\mu_{2\alpha+1} - \mu_{a+1}\mu_\alpha\right) \\
&\quad -\alpha\mu_1^2\mu_\alpha\left(\mu_{2\alpha+1} - \mu_1\mu_{2\alpha}\right) + \mu_1^3\left(\mu_{3\alpha} - \mu_\alpha\mu_{2\alpha}\right) \\
&\quad -2\alpha\mu_\alpha\left(\mu_1\mu_{2\alpha} - \alpha\mu_\alpha\mu_{\alpha+1} - (1-\alpha)^2\mu_1\mu_\alpha^2\right)\left(\mu_2 - \mu_1^2\right) \\
&\quad -\alpha^3\mu_\alpha^3\left(\mu_3 - \mu_1\mu_2\right),
\end{aligned}
$$

and μ_α *is the* α*-th moment of the income distribution* F*.*[2]

All moments μ_α exist under the regularity assumptions we made for the existence of the Edgeworth expansion above.

Appendix B presents simulation support for the expressions in the proposition.

The second order Edgeworth expansion for the studentized inequality index is given by (5) with $k_{1,2}$ and $k_{3,1}$ given in Proposition 1. Bias and skewness coefficients can be estimated non-parametrically by the sample analogues of the formulas in Proposition 1 (we refer to these estimators by $\widehat{k}_{1,2}$ and $\widehat{k}_{3,1}$). Moreover, such estimation does not affect the order of the approximation of the Edgeworth expansion since the estimators are $\sqrt{n}$-consistent.

4.2 Examples

The bias and skewness coefficients are functions of the sensitivity parameter α of the inequality index and the relevant parameters of the income distribution. We proceed to examine the behavior of the coefficients for the income distributions studied in Section 3, the Gamma, Lognormal, and Singh-Maddala distributions.

[2] The R and Mathematica computer code for these expressions are available from the authors upon request.

The interpretation of the individual contributions is made plain in the derivation contained in Appendix A.1.

Figure 2 depicts the contour plots of $k_{1,2}$ and $k_{3,1}$. The coefficients share important features across all three income distributions. (i) All bias and skewness coefficients are negative (ii) except for small values of α, bias and skewness increase in magnitude as α increases (iii) bias and skewness decrease with r in the Gamma case and with b in the Singh-Maddala case (holding c constant and chosen so that all relevant moments exist), and they increase with $sd = v^{1/2}$ in the Lognormal case. (iv) the numeric values of the coefficients are of quantitative importance, as dividing them by the square root of the sample sizes considered in Section 3 yields values which are large relative to the Gaussian limit values of zero. The correction in the Edgeworth expansion is large and is this often regarded as a negative indication of the reliability of the approximation.

Despite the rather lengthy expressions for the coefficients, it is possible to simplify these considerably in some special cases.

As a first example, consider the Gamma case with $\alpha = 2$ fixed and r varying. In terms of the contour plot of 3, this is the top-most horizontal section. We have

$$k_{1,2} = -\frac{3}{\sqrt{2}}\frac{r+3}{\sqrt{r(r+1)}} \quad \text{and} \quad k_{3,1} = -\frac{8}{\sqrt{2}}\frac{r+4}{\sqrt{r(r+1)}}. \tag{6}$$

Consistent with observation (iii) above, these decrease in magnitude as r increases. For the case of fixing $r = 3$ and varying α instead, we have

$$\begin{aligned} k_{3,1} &= -\frac{2}{\sqrt{3}}\left(-\left(\alpha^2+3\right)\Gamma^2(3+\alpha)+6\Gamma(3+2\alpha)\right)^{-3/2}\times \\ &\quad \left(\Gamma^3(3+\alpha)\left(18\alpha^2+\alpha^3+3\alpha^4+18\right)+36\Gamma(3+3\alpha)\right. \\ &\quad \left.-18\Gamma(3+2\alpha)\Gamma(3+\alpha)\left(2\alpha^2+3\right)\right)\times\left(1-2\cdot\mathbf{1}_{(0,1)}(\alpha)\right). \end{aligned} \tag{7}$$

Consistent with (ii) skewness increases in magnitude for $\alpha > 0.5$. Obviously, (6) and (7) coincide for $\alpha = 2$ and $r = 3$.

As another example consider the skewness coefficient in the Lognormal case with $sd = v^{1/2}$ and $\alpha = 2$ fixed. We have

$$\begin{aligned} k_{3,1} &= \left(4e^v - 4e^{2v} + e^{4v} - 1\right)^{-3/2}\times \\ &\quad \left(-2e^{12v}+12e^{8v}+12e^{6v}-48e^{5v}-36e^{4v}+136e^{3v}-96e^{2v}+24e^v-2\right). \end{aligned}$$

Consistent with (iii) that skewness increases in magnitude with sd.

4.3 Edgeworth density

The Edgeworth expansion follows immediately from the derived coefficients, and the implied second order expansion of the density is $(1+n^{-1/2}x[\frac{1}{6}k_{3,1}(x^2-3)+k_{1,2}])\phi(x)$.

For instance, in the Gamma case with $\alpha = 2$ fixed and r varying, using (6), we have

$$pdf(x) = \phi(x)\left[1 - n^{-1/2}x\frac{\sqrt{2}}{6\sqrt{r(r+1)}}\left(4x^2(4+r)-3r-21\right)\right] + O\left(n^{-1}\right)$$

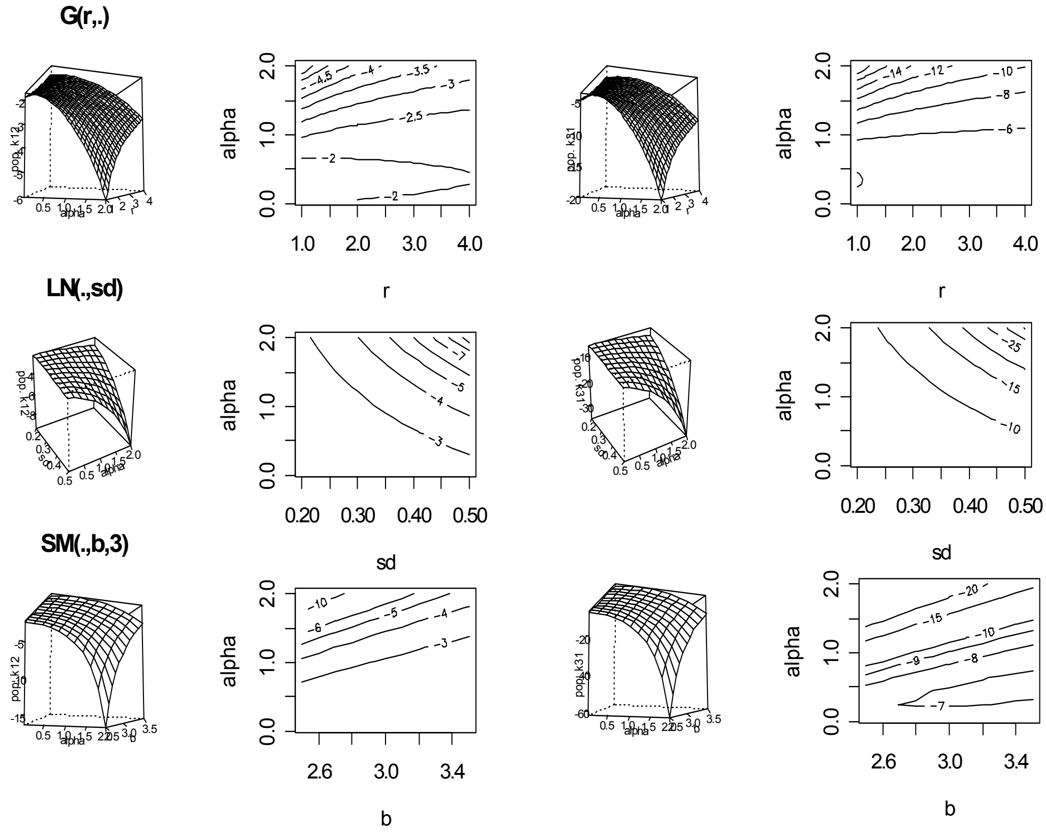

Figure 2: Bias and skewness coefficients $k_{1,2}$ and $k_{3,1}$ as functions of relevant parameters of the income distributions, and of the sensitivity parameter α of the inequality index.

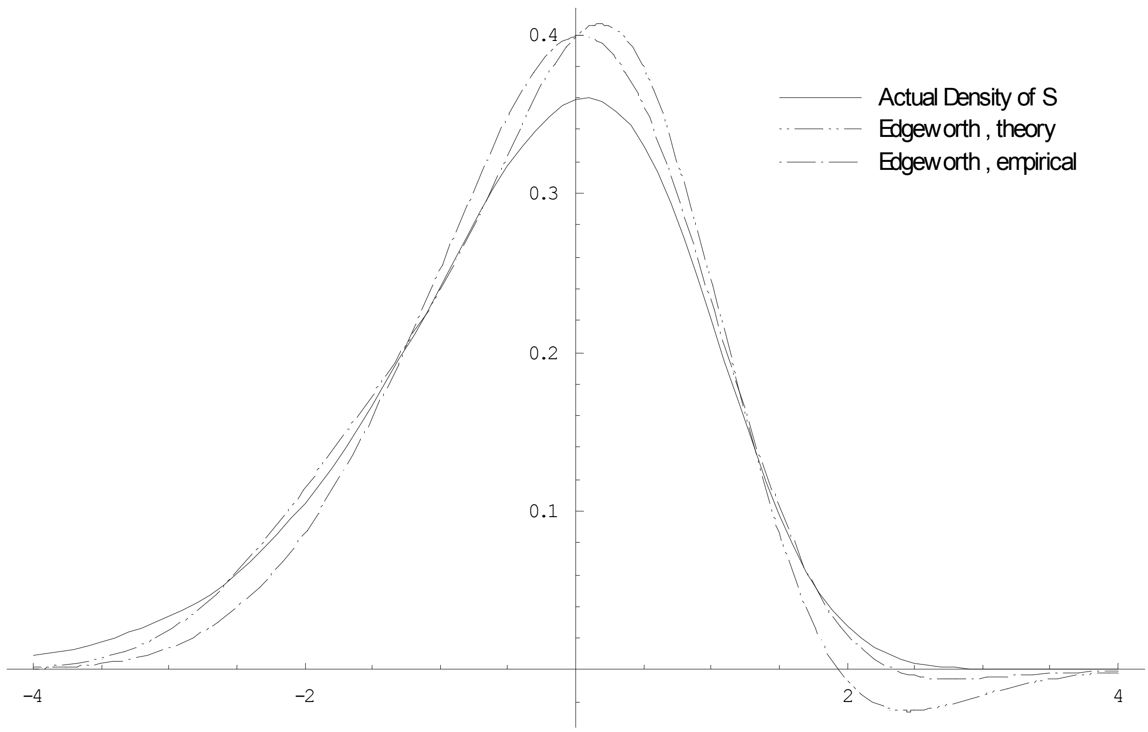

Figure 3: Density estimates. Notes: The income distribution is G(3,.), the sensitivity parameter of the inequality index is α=2, and the sample size is n =100. The solid line depicts the simulated density of S, the first dashed line $(- \cdot \cdot -)$ depicts the Edgeworth density based on the theoretical $k_{1,2}$ and $k_{3,1}$, and the second dashed line $(- \cdot -)$ is the Edgeworth density based on the estimated $k_{1,2}$ and $k_{3,1}$. Kernel density estimates based on 10^5 replications.

Figure 3 depicts two versions of this Edgeworth density for the case $r = 3$ and $n = 100$, one with the theoretical coefficients, and one empirical version. This empirical version can be thought of as the mean Edgeworth expansions when averaged over simulation iterations. By the linearity of the density functions in terms of the $k_{1,2}$ and $k_{3,1}$, this simply equals the Edgeworth density evaluated at the averaged estimates of $k_{1,2}$ and $k_{3,1}$. The graph also shows the simulated density of the studentized inequality index.

Both approximations are an improvement over the Normal approximation in that they capture the skewness of the distribution. Several features are noteworthy, however. First, the Edgeworth density is not guaranteed to be positive and we see that the right hand tail actually goes negative, although less so for the empirical than for the theoretical Edgeworth approximation. Second, the right tail of the theoretical Edgeworth density decays too quickly. As r decreases, we know from the contour plots that both bias and skewness increase, and the actual density departs further from the Gaussian density.

For $r = 0.6$, not shown in Figure 3, the Edgeworth density exhibits the third problematic feature more clearly: oscillations in the tails of the approximation. This problem of oscillation, which is also present for the other income distributions studied here, is well known (see e.g. Niki and Konishi, 1986), and motivates the search for a normalizing transform.

5 Normalizing Transforms

Rather than directly adjusting the asymptotic approximation for S by including the $O\left(n^{-1/2}\right)$ term in the approximating density, normalizing transformations of the inequality measure are designed to annihilate this term asymptotically. The resulting distribution of the studentized transformed and bias corrected inequality measure then satisfies $\Phi(x)+O(n^{-1})$, so that, compared to (3) or (5), the order of the approximation has improved.

In our derivations of the required refinement we essentially follow an approach proposed in Niki and Konishi (1986). See also Marsh (2004) for a multivariate extension. However, there are four important differences. First, we standardize using the empirical quantities, whereas Niki and Konishi use theoretical versions. Second, we deal explicitly with the open issue highlighted by Niki and Konishi (1986, p.377) that the cumulants depend on the true quantity I being estimated. Third, we deal with the complication that income distributions and the inequality measure I often depends on more than one parameter, as in the case of the Singh-Maddala distribution. This implies that I is not an invertible mapping of the parameters and as a consequence a whole family of solutions exist. One can choose any member of this family and we show what difference this choice makes. Finally, only in the simplest of cases can we find an analytic solution as in Niki and Konishi (1986) or our Gamma example with $\alpha = 2$ below. In general we need numerical techniques, which we develop and implement.

We investigate the properties of the transformations and derive the associated bias and skewness coefficients. The skewness coefficient is zero if the transform is the solution to the key differential equation given below. We then derive the solutions to this differential equation for different inequality measures and income distributions.

We will leave the dependence of σ, $k_{1,2}$, and $k_{3,1}$ and other quantities on α and the income distribution and its parameters implicit for notational simplicity throughout this section.

5.1 Transformations

Let t denote a transformation of the inequality measure I with continuous first and second derivatives t' and t'', satisfying $t'(\widehat{I}) \neq 0$. The standardized transform defined by

$$T = n^{1/2}\frac{t(\widehat{I}) - t(I)}{\widehat{\sigma}t'(\widehat{I})}, \tag{8}$$

will also be asymptotically Normal, but its cumulants will have changed and depend on the nonlinear transformation t. We want to relate the cumulants of T to the cumulants of S and determine T such that the third cumulant vanishes. In order to do so, we first state the basic relation between S and T in the following lemma.

Lemma 2.

$$T = S - \frac{1}{2}\frac{t''(I)}{t'(I)}n^{-1/2}\sigma S^2 + O_p\left(n^{-1}\right) \tag{9}$$

Assuming that the distribution of T also admits a valid Edgeworth expansion, it will be of the form

$$\Pr(T \leq x) = \Phi(x) - n^{-1/2}\left(\lambda_{1,2} + \frac{1}{6}\lambda_{3,1}(x^2-1)\right)\phi(x) + O(n^{-1}), \tag{10}$$

where $\lambda_{1,2}$ and $\lambda_{3,1}$ are the coefficients for $n^{-1/2}$ of the first and third cumulant of T respectively. The cumulants of T are naturally related to the cumulants of S since t is a smooth function of I. The next lemma states this relationship.

Lemma 3.

$$\lambda_{1,2} = k_{1,2} - \frac{1}{2}\sigma\frac{t''(I)}{t'(I)}, \tag{11}$$

$$\lambda_{3,1} = -\frac{3\sigma t''(I) - k_{3,1}t'(I)}{t'(I)}. \tag{12}$$

Our results differ form those stated in Niki and Konishi (1986) because our definition of the standardized transform (8) has $\widehat{\sigma}t'(\widehat{I})$ in the denominator instead of $\sigma t'(I)$ used by Niki and Konishi. The consequence of this is that in expression (9) the second term on the right has a coefficient of -1. Finally, the differences between $\lambda_{1,2}$ and $\lambda_{3,1}$ of Lemma 3, and the results of Niki and Konishi are the negative signs of the second right hand terms. This again is a result of $\sigma t'(I)$ being estimated.

The normalizing transform we seek is a function t that reduces the skewness of T to zero up to second order. It follows from equation (12) that the skewness term $\lambda_{3,1}$ is reduced to zero if the transform t satisfies the differential equation

$$3\sigma t''(I) - k_{3,1}t'(I) = 0, \tag{13}$$

or, assuming $t'(I) \neq 0$,

$$\frac{t''(I)}{t'(I)} = \frac{1}{3}\frac{k_{3,1}}{\sigma}.$$

The formal solution to the differential equation is

$$t(I) = \int \exp\left(\int \frac{1}{3}\frac{k_{3,1}}{\sigma}dI\right)dI.$$

The asymptotic refinement we seek is found by solving the differential equation (13), and making a subsequent direct bias correction based on (11). Note that the differential equation is invariant to affine transformations.

Proposition 4. *If the transform t satisfies the differential equation (13), then*

$$\Pr\left(T - n^{-1/2}\lambda_{1,2} \le x\right) = \Phi(x) + O(n^{-1}). \tag{14}$$

The availability of the refinement in practice depends on whether the differential equation (13) can be solved for particular inequality measures and income distributions. The bias correction can be applied using the $\sqrt{n}$-consistent estimator $\widehat{\lambda}_{1,2}$ based on the EDF, giving rise to the following result.

Corollary 5. $\Pr\left(T - n^{-1/2}\widehat{\lambda}_{1,2} \le x\right) = \Phi(x) + O(n^{-1})$.

5.2 Transforms for Inequality Measures

We turn to deriving the normalizing transforms for the three distributional cases discussed above. The relative simplicity of the Gamma case yields an explicit analytical solution. Typically, however, the transform is computed using numerical techniques which we develop and implement below. We then provide a systematic discussion of the properties of the transforms across income distributions and sensitivity parameters.

5.2.1 The Gamma Distribution

Assume that incomes follow the Gamma distribution $G(r, .)$, and consider the case of varying r and α fixed at 2. The formulas for $k_{1,2}$ and $k_{3,1}$ specialize to expression (6), and using the fact that $I(2) = (2r)^{-1}$, the differential equation (13) yields

$$\frac{t''(I)}{t'(I)} = -\frac{4}{3}\frac{1}{I}\frac{1+8I}{1+2I}.$$

Integration yields the exact solution

$$\begin{aligned} t(I) &= -3I^{-1/3} + \frac{140}{81}2^{1/3}\ln\left(I^{1/3} + 2^{-1/3}\right) \\ &\quad -\frac{70}{81}2^{1/3}\ln\left(I^{2/3} - I^{1/3}2^{-1/3} + 2^{-2/3}\right) - \frac{140}{81}3^{1/2}2^{1/3}\arctan 3^{-1/2}\left(2^{4/3}I^{1/3} - 1\right) \\ &\quad -\frac{118}{27}\frac{I^{2/3}}{1+2I} - \frac{16}{9}\frac{I^{2/3}}{(1+2I)^2} - \frac{2}{3}\frac{I^{2/3}}{(1+2I)^3}. \end{aligned} \tag{15}$$

The transformation for this particular case is depicted in the second panel of Figure 4 together with the transformations for other values of the sensitivity parameter α, and other income distributions. We discuss this figure below. We also discuss the distribution of the transformed statistic following Figure 6 below.

5.2.2 Numerical Solutions

In general we cannot obtain analytic solutions. The Gamma example with $\alpha = 2$ is special for two reasons. First, because the simple form of the differential equation, and second because there is a simple invertible relation between the inequality index and the parameters of the distribution. This relation is no longer trivial when $\alpha \neq 2$, and in general there is no analytically tractable relation for other distributions. It is possible however to obtain a numerical inverse and to solve the differential equation (13) numerically. This involves three steps:

1. Using the general formulas for the cumulants $k_{1,2}$, $k_{3,1}$, and σ^2 we can use the theoretical moments from a specific income distributions to express the cumulants $k_{1,2}$, $k_{3,1}$, and σ^2 in terms of parameters from the income distribution.

2. We express the cumulants in terms of I. This requires the inverse of I which we calculate numerically. The inverse can be determined if I depends on one parameter only. This obviously holds for one parameter families, but also for the Log-Normal and Gamma distributions, because I is scale invariant. This invariance property is inherited by the cumulants, so that $k_{1,2}$ and $k_{3,1}$ depend only on the shape parameter and we can express them as functions of I.
   For other distributions we can choose a one dimensional path for the parameters such that I becomes an invertible function. This can be achieved by imposing the correct number of restrictions ($d-1$, if I depends on d income parameters), or by making the d income distribution parameters a function of only one parameter. In our Singh-Maddala distribution we show what difference the restriction makes by first holding b fixed, such that I is an invertible function of c only, and then holding c fixed.

3. We solve the differential equation (13) numerically. This solution is then available for use in practice and simulations.[3]

The solutions depend, in general, on the sensitivity parameter α, which is a fixed known constant, and the true underlying distribution. If I depends on more than one parameter of the income distribution, then the solution will furthermore depend on the chosen restriction, or parametrization. This applies, for instance, to the Singh-Maddala distribution, and we compare the two solutions associated with the restriction that b is fixed (such that the inequality measure I is a function of c only), and that c is fixed.

Since no analytical solutions are available in general, we display the transforms graphically in Figures 4 and 5. Figure 4 shows the transforms for different values of the sensitivity parameter α and, in each panel, we consider a different income distribution. In Figure 5 we compare the transforms across income distributions when α is fixed at 2. In order to relate the curvature of the transforms to the untransformed case, we have also depicted the 45-degree lines, which represents the identity transforms. The figures have been generated by setting the initial conditions of the differential equation such that the solutions cross the horizontal axis at the same point at an angle of 45 degrees. Consequently, if no transformation was required, the solution would coincide with this 45 degree line. For the Singh-Maddala income distribution,

[3]The Mathematica code is available from the authors upon request.

we display the solutions for the two restrictions separately. Finally, for the Gamma case with $\alpha = 2$, we cannot visually distinguish the numerical solution from the exact solution given by Equation (15).

Consider Figure 4 first. We see that for all the distributions the transforms change substantially as we vary α. The transform for $a = 2$ is the one most curved in all four cases. This implies that $\alpha = 2$ requires, informally speaking, the biggest amount of transformation to obtain a standard Normal distribution. Recall from the previous simulations that $\alpha = 2$ is the most troublesome case. For $\alpha = 0$, the least amount of adaptation is required in the Lognormal and Singh-Maddala cases, whereas for the Gamma case, $\alpha = 0.5$ is slightly flatter than $\alpha = 0$. The behaviour of the transform is therefore not monotonic in α. For the Singh-Maddala distribution we see quite a difference between the case where we hold b constant and where we hold c fixed. This is related to the fact that I, and the gradient of I with respect to b change much faster than with respect to c. This has a further consequence that the domain of definition of the numerical transform is much broader for c fixed than for b fixed. This can be seen from the graph e.g. for $\alpha = 2$ the numerical transform is only calculated for I between approximately 0.055 and 0.095 for b fixed and between 0.03 and 0.28 for c fixed.

Next, we compare the transforms in Figure 5 across income distributions when $\alpha = 2$. One might have hoped that the normalizing transforms were similar, so that a "consensus" transform could be applied by practitioners across income distributions. The figure shows, however, that this is not the case: the transforms vary substantially between income distributions and there is no single transform that is appropriate for this variety of cases.

The transforms, which we derive as numerical objects in Mathematica, can be used in simulations to show that the transformed statistics are distributed closer to a standard Normal distribution. We examine the extent to which this transform achieves skewness reduction in samples of size $n = 100$. We simulate the densities of the studentized inequality measure S, and the studentized bias corrected transform T based on Corollary 5. The top panel in Figure 6 considers the Gamma case for which we derived an exact solution for the transform in Equation (15). The actual finite sample density of S departs substantially from the limiting density (the same distributional case has been depicted in Figure 3). In particular, the density is skewed to the left, and biased. The transform succeeds in reducing substantially the skewness, and the bias correction shifts the density to the right. The resulting density is much closer to the Gaussian density.

Figure 6 also shows the resulting distribution for the Lognormal and Singh-Maddala cases which require the application of our numerical solutions.

The graph shows that the distribution of T is indeed closer to a standard Normal distribution than that of S, with the exception of the transform based on the Singh-Maddala distribution holding b fixed. The transforms seem to overcompensate the skewness as the distribution of T is now skewed to the right.

When constructing 95% confidence intervals using the standard Normal critical values, the transformed statistic gives coverage rates much closer to the nominal one than the simple standardized statistic S. For the Gamma case the coverage failure

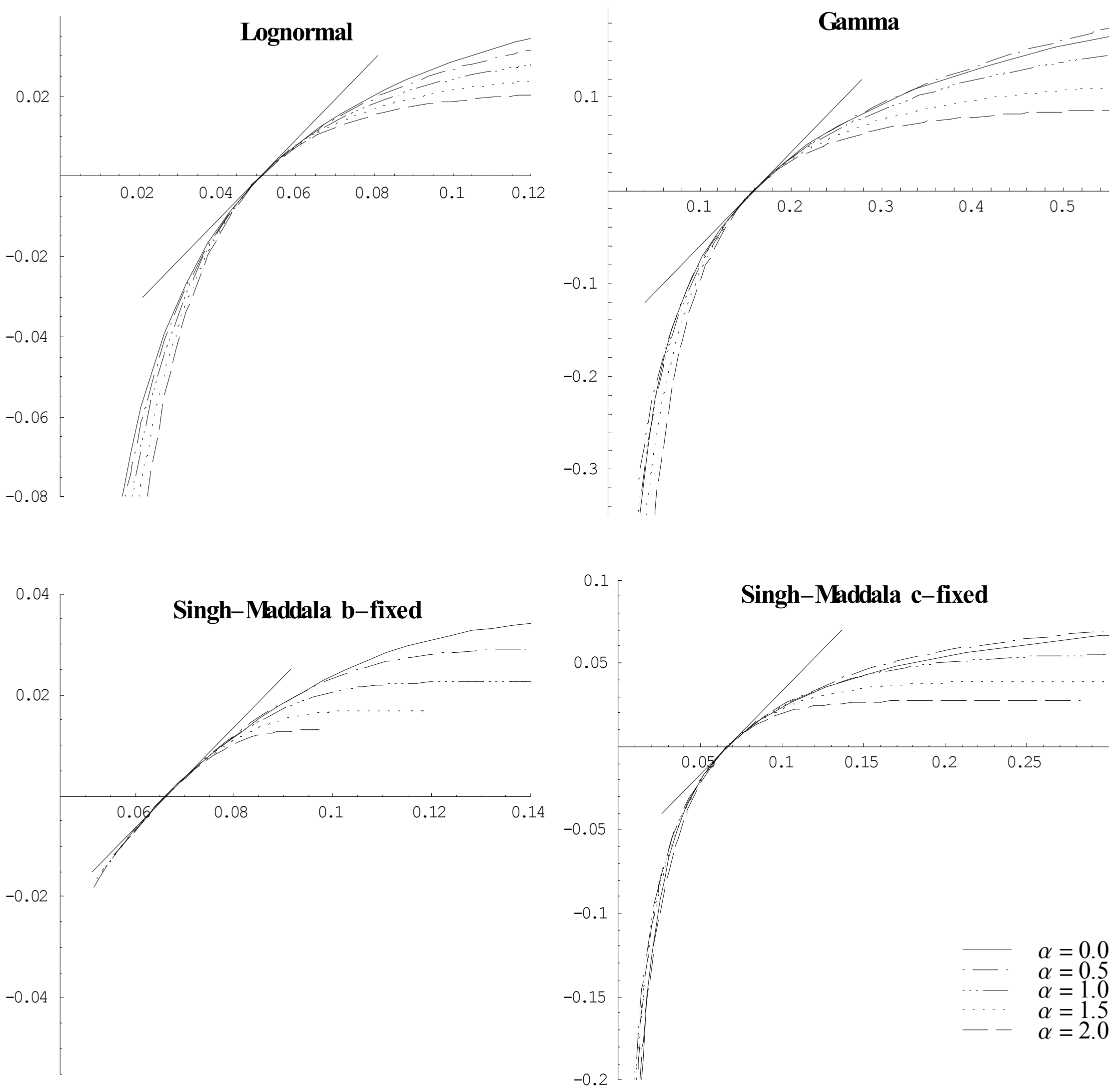

Figure 4: Transforms for different distributions and αs. Notes: On the horizontal axis is I (itself a function of parameters), and on the vertical axis is $t(I)$. The Singh-Maddala case at the bottom requires a restriction and $b = 3.5$ is chosen in the left hand panel and $c = 3.5$ in the right hand panel.

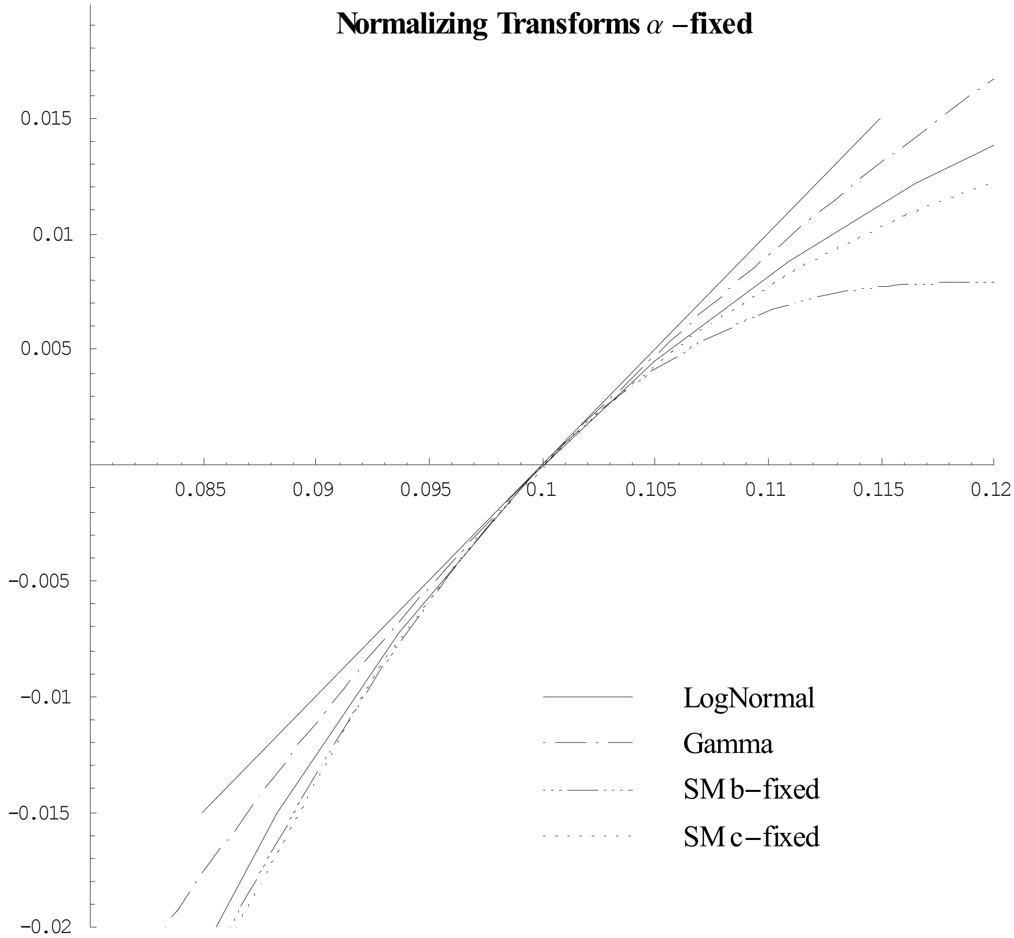

Figure 5: Normalizing Transforms for different distributions and fixed $\alpha = 2$. Notes: On the horizontal axis is I (itself a function of parameters), and on the vertical axis is $t(I)$. For the two Singh-Maddala cases $b = 3$ and $c = 3$ fixed are chosen.

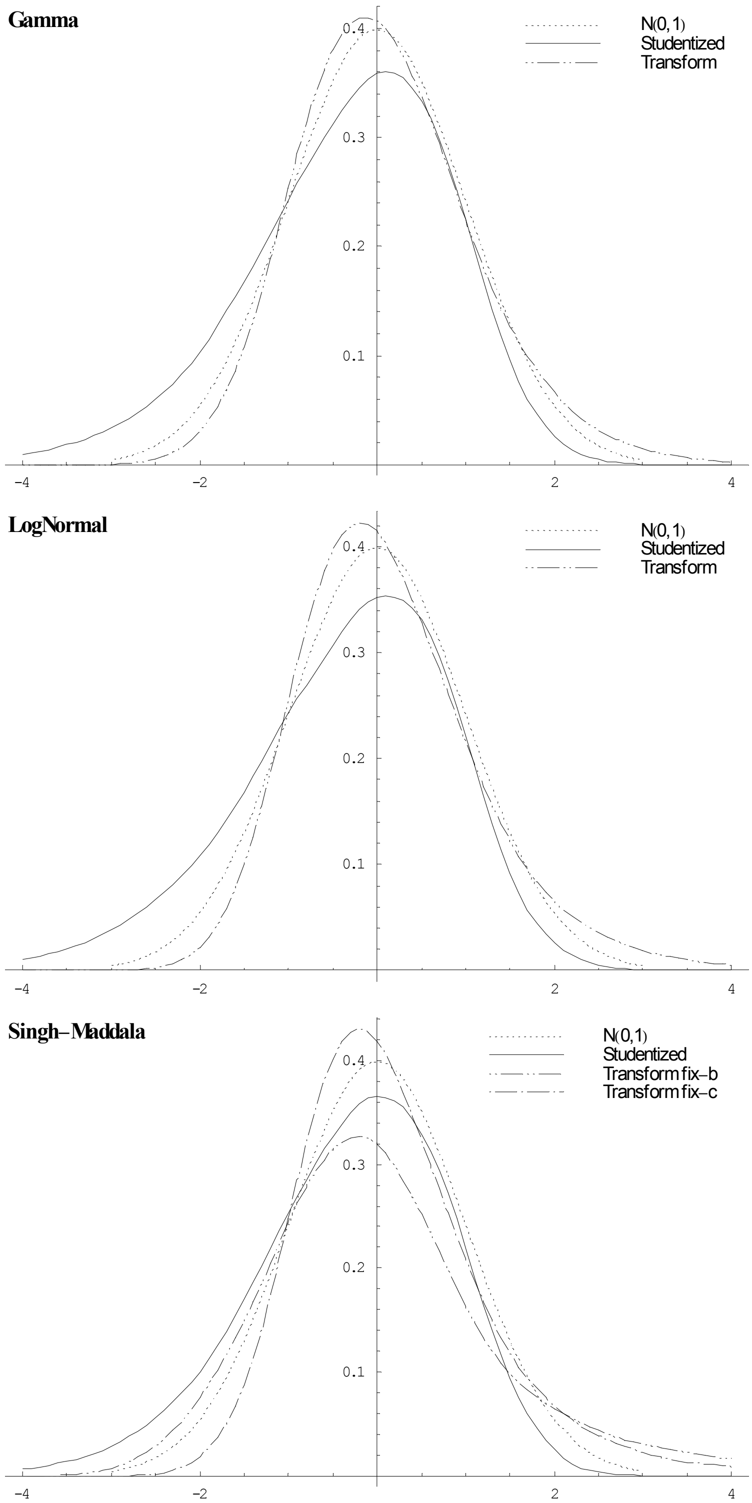

Figure 6: Simulated densities of S and T for various income distributions. Notes: Sample size is $n = 100$. Parameter values: Gamma: $r = 3$, Lognormal: $\nu^2 = 0.1$, Singh-Maddala: $b = 3.5$, $c = 3$. Kernel density estimates based on 10^5 replications.

rate based on S is 10.0% and based on T it is 5.3%. For the Lognormal case these rates are 10.6% and 5.8% respectively and for the Singh-Maddala case they are 8.9% and 7.5% when holding c fixed.

These similar gains are not unexpected since the cumulants are fairly close for the three cases with $k_{1,2} = -3.67, -4.32, -4.60$ and $k_{3,1} = -11.4, -14.1, -15.8$ for the Gamma, Lognormal and Singh-Maddala cases respectively.

The choice of restriction used to establish an invertible relation between the index and parameters is not innocuous, however, since holding b fixed instead of c leads to a coverage that is far worse with 19% coverage failure (of which 2% points are caused by the fact that $\hat{I}$ falls outside the domain of definition of the numerical transform). This implies that care needs to be taken with this choice. One option would be to use a path in the parameter space for which the change in the index I is maximized. This requires a further solution to a differential equation, which is not difficult, but would complicate the exposition. The path holding c fixed is closer to this direction than holding b fixed. The optimal direction is (0.97,0.23) when $b = 3.5$ and $c = 3.0$ and (0.99,0.16) when $b = 3.5$ and $c = 3.5$, hence close to holding c fixed and little gain is therefore expected from determining the optimal restriction.

6 Conclusions

The finite sample distribution of the studentized inequality measure is not located at zero and is substantially skewed. In the first part of the paper we have derived general nonparametric bias and skewness coefficients based on cumulant expansions. Edgeworth expansions directly adjust the asymptotic approximation by including the $O\left(n^{-1/2}\right)$ term, a function of the bias and skewness coefficients. In contrast, normalizing transformations of the inequality measure are designed to annihilate this term asymptotically. The observed problems for the Edgeworth expansion of negativity of the density and tail oscillation have led us to derive and construct the normalizing transforms in the second part of this paper. We have shown that the finite sample distributions of these transforms are much closer to the Gaussian distribution. However, we show that the transforms of the inequality measure vary with the sensitivity parameter and between income distributions. This implies that there is no universal transform that works well across all cases.

Although this paper is a second order investigation into the problems associated with GE inequality indices, we have shown that the nonlinear transform can also be used to improve inference. For standard symmetric two-sided 95% confidence intervals using the standard Normal critical values, the transformed statistic gives coverage rates much closer to the nominal one than the simple standardized statistic S. The construction of good finite sample confidence intervals across various income distributions is the subject of current work.

References

Atkinson, A. B. (1970), "On the Measurement of Inequality", *Journal of Economic Theory*, 2, 244–263.

Biewen, M. (2001), "Bootstrap Inference for Inequality, Poverty and Mobility Measurement", *Journal of Econometrics,* 108, 2, 317-342.

Bandourian, R. and J.B. McDonald, and R.S. Turley (2003), "A Comparison of Parametric Models of Income Distribution Across Countries and Over Time", *Estadistica*, 55, 135-152, Luxembourg Income Study Working Paper No. 305. , www.lisproject.org/publications/liswps/305.pdf.

Brachmann, K. and A. Stich, and M. Trede (1996) "Evaluating Parametric Income Distribution Models", *Allgemeines Statistisches Archiv*, 80, 285–298.

Cowell, F. A. (1977), Measuring Inequality, (First. Ed.), Oxford: Phillip Allan.

Cowell, F. A. (1980), "On the Structure of Additive Inequality Measures", *Review of Economic Studies*, 47, 521-531.

Cowell, F. A. (1989), "Sampling Variance and Decomposable Inequality Measures", *Journal of Econometrics* 42, 27-41.

Cowell, F. A. (2000), "Measurement of Inequality", In A.B. Atkinson and F. Bourguignon (Eds.), Handbook of Income Distribution, Chapter 2. Amsterdam: North Holland.

Davidson, R. and J.-Y. Duclos (1997), "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers", *Econometrica*, 65, 1453-1465.

Garderen, K.J.van, and C. Schluter (2003), "Improving Finite Sample Confidence Intervals for Inequality and Poverty Measures", *UvA-Econometrics Working Paper*, 2003/02, http://www.fee.uva.nl/ke/UvA-Econometrics.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansions*, Springer Verlag.

Hart, P.E. (1971), "Entropy and Other Measures of Concentration", *Journal of the Royal Statistical Society, Ser.A* 134, 73-85.

Kloek, T. and H. van Dijk (1978), "Efficient Estimation of Income Distribution Parameters", *Journal of Econometrics* 8, 61-74.

Maasoumi, E. (1997), "Empirical Analyses of Inequality and Welfare", in M.H.Pesaran and P.Schmidt (Eds.) Handbook of Applied Econometrics, Chapter 5. Oxford: Blackwell.

Maasoumi, E. and H. Theil (1979), "The Effects of the Shape of the Income Distribution on Two Inequality Measures", *Economics Letters*, 4, 289-291.

Marsh, P. (2004), "Transformations for Multivariate Statistics", *Econometric Theory,* 20, 963-987.

McDonald, J.B. (1984), "Some Generalized Functions for the Size Distribution of Income", *Econometrica*, 52, 647–663.

Mills, A.M. and S. Zandvakili (1997), "Statistical Inference Via Bootstrapping for Measures of Inequality", *Journal of Applied Econometrics,* 12, 133-150.

Niki, N., and S. Konishi (1986), “Effects of Transformations in Higher Order Asymptotic Expansions”, *Annals of the Institute of Statistical Mathematics,* 38, 371-383.

Schluter, C. and M. Trede (2002), “Tails of Lorenz Curves", *Journal of Econometrics,*109, 1, 151-166.

Singh, S.K. and G.S. Maddala (1976), “A Function for Size Distribution of Incomes”, *Econometrica,* 44, 963–970.

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland.

Thistle, P.D. (1990), “Large Sample Properties of Two Inequality Indices”, *Econometrica,* 58, 725-728.

A Proofs

We first prove lemmas 2, 3, and Proposition 4 and postpone the proof of Proposition 1 until Section A.1.

Proof of Lemma 2.

Expand $T(\widehat{I})$ about I to second order and use the definition of S to obtain

$$\frac{n^{1/2}}{\widehat{\sigma}}\frac{t(\widehat{I})-t\left(I\right)}{t'\left(\widehat{I}\right)}=S-\frac{1}{2}\frac{t''\left(I\right)}{t'\left(I\right)}n^{-1/2}\hat{\sigma}S^{2}+O\left(\left\|\hat{I}-I\right\|^{2}\right),$$

Now using the fact that $\hat{\sigma}$ and $\hat{I}$ are $\sqrt{n}$ consistent estimators the result follows.

Proof of Lemma 3.

Taking expectations of (9) using $E\left(S\right)=n^{-1/2}k_{1,2}+O\left(n^{-3/2}\right)$, and $E\left(S^{2}\right)=1+O\left(n^{-1}\right)$ leads to

$$\lambda_{1,2}=k_{1,2}-\frac{1}{2}\frac{t''\left(I\right)}{t'\left(I\right)}\sigma.$$

Also $E\left(T^{2}\right)=1+O\left(n^{-1}\right)$. We have

$$T^{3}=S^{3}-\frac{3}{2}\frac{t''\left(I\right)}{t'\left(I\right)}n^{-1/2}\sigma S^{4}+O_{p}\left(n^{-1}\right).$$

Taking expectations, noting that $E\left(S^{4}\right)=3+O_{p}\left(n^{-1/2}\right)$, yields

$$E\left(T^{3}\right)=E\left(S^{3}\right)-\frac{9}{2}\frac{t''\left(I\right)}{t'\left(I\right)}n^{-1/2}\sigma+O\left(n^{-1}\right),$$

with $E\left(S^{3}\right)=n^{-1/2}\left[k_{3,1}+3k_{1,2}\right]$. Therefore $E(T^{3})-3E(T^{2})E(T)+2\left(E(T)\right)^{3}=n^{-1/2}\lambda_{3,1}+O\left(n^{-1}\right)$ with

$$\lambda_{3,1}=k_{3,1}-3\frac{t''\left(I\right)}{t'\left(I\right)}\sigma.$$

Proof of Proposition 4.

Note that $\lambda_{3,1}=0$ by construction. Considering the Edgeworth expansion Equation (10) for T at $x+n^{-1/2}\lambda_{1,2}$, expanding it about x and collecting terms of the same order yields the stated result.

A.1 Proof of Proposition 1

Proposition 1 is derived in several steps. First, we derive an asymptotic expansions of the studentized inequality measure S. As a compact notation, we use S_{q} to denote a term of an expansion of S which is of order in probability n^{-q}. The desired stochastic expansion of S is of the form

$$S=S_{0}+S_{1/2}+O_{p}\left(n^{-1}\right). \tag{16}$$

We determine the terms S_{0} and $S_{1/2}$. We then derive the bias and skewness coefficients $k_{1,2}$ and $k_{3,1}$ by considering expectations of powers of S. We only consider the case $|\alpha|>1$ explicitly. For $|\alpha|<1$, the coefficients of the expansions need to be multiplied by -1 since $\alpha\left(\alpha-1\right)<0$ but $(\alpha^{2}\left(\alpha-1\right)^{2})^{1/2}>0$.

A.1.1 The Stochastic Expansion of S

Recall our notation for population and sample moments, $\mu_\alpha(F) = \int y^\alpha dF(y)$ and $m_\alpha = \mu_\alpha(\widehat{F})$. The basic technique in the derivation is to center and expand sample moments. For instance, we have $m_1^{-\alpha} = \left(\mu_1 + n^{-1}\sum(X_i - \mu_1)\right)^{-\alpha} = \mu_1^{-\alpha} - \alpha\mu_1^{-\alpha-1}\left(n^{-1}\sum(X_i-\mu_1)\right) + O_p\left(n^{-1}\right)$. For this technique it is convenient to define the following stochastic quantities:

$$
\begin{aligned}
Y_1 &= (X-\mu_1), \\
Y_2 &= \mu_1\left(X^\alpha - \mu_\alpha\right) - \alpha\mu_\alpha\left(X-\mu_1\right), \\
Y_3 &= \left(X^\alpha-\mu_\alpha\right) - \alpha\left(\alpha+1\right)\mu_\alpha\mu_1^{-1}\left(X-\mu_1\right)/2, \\
Y_4 &= 2\left(\mu_1\mu_{2\alpha} - \alpha\mu_\alpha\mu_{\alpha+1} - (1-\alpha)^2\mu_1\mu_\alpha^2\right)\left(X-\mu_1\right) \\
&\quad + \alpha^2\mu_\alpha^2\left(X^2-\mu_2\right) \\
&\quad + 2\left(\alpha^2\mu_\alpha\mu_2 - \alpha\mu_1\mu_{\alpha+1} - (1-\alpha)^2\mu_1^2\mu_\alpha\right)\left(X^\alpha-\mu_\alpha\right) \\
&\quad - 2\alpha\mu_1\mu_\alpha\left(X^{\alpha+1}-\mu_{\alpha+1}\right) + \mu_1^2\left(X^{2\alpha}-\mu_{2\alpha}\right).
\end{aligned}
\tag{17}
$$

with specific elements for observation i written like $Y_{1,i} = (X_i - \mu_1)$, etc.

We derive the stochastic expansion of $S = n^{1/2}\left(\widehat{I}-I\right)/\widehat{\sigma}$ in four steps.

First, write out the numerator

$$
n^{1/2}(\widehat{I}-I) = n^{1/2}\left[\alpha^2-\alpha\right]^{-1}\mu_1^{-\alpha}m_1^{-\alpha}\left[\mu_1^\alpha m_\alpha - \mu_\alpha m_1^\alpha\right].
$$

Second, consider the asymptotic variance by applying the delta-method

$$
\sigma^2 = aVar(n^{1/2}(\widehat{I}-I)) = \frac{1}{\left(\alpha^2-\alpha\right)^2}\frac{1}{\mu_1^{2\alpha+2}}B_0,
$$

with

$$
B_0 = \left[\alpha^2\mu_\alpha^2\mu_2 - 2\alpha\mu_1\mu_\alpha\mu_{\alpha+1} + \mu_1^2\mu_{2\alpha} - (1-\alpha)^2\mu_1^2\mu_\alpha^2\right]. \tag{18}
$$

The variance is estimated by using the corresponding sample moments. Denote the estimate of B_0 by $\widehat{B}_0$. Then combining the results from steps 1 and 2 yields

$$
S = n^{1/2}\widehat{B}_0^{-1/2}\left[m_\alpha m_1 - \mu_1^{-\alpha}\mu_\alpha m_1^{\alpha+1}\right].
$$

Third, consider the expansion $\widehat{B}_0 = B_0 + B_{1/2} + O_p\left(n^{-1}\right)$. We have

$$
\begin{aligned}
\widehat{B}^{-1/2} &= \left[B_0 + B_{1/2} + O_p\left(n^{-1}\right)\right]^{-1/2}, \\
&= B_0^{-1/2} - \frac{1}{2}B_0^{-3/2}B_{1/2} + O_p\left(n^{-1}\right).
\end{aligned}
$$

The term $B_{1/2}$ is derived by centering and collecting terms of the same order. It then follows that $B_{1/2} = \left[n^{-1}\sum_i Y_{4,i}\right]$

Fourth, consider the term $\left[m_\alpha m_1 - \mu_1^{-\alpha}\mu_\alpha m_1^{\alpha+1}\right]$ by expanding the functions of the sample moments. Putting everything together and collecting terms of the same order, it follows that $S = S_0 + S_{1/2} + O_p\left(n^{-1}\right)$ with

$$\begin{aligned} S_0 &= n^{1/2}B_0^{-1/2}\left[n^{-1}\sum Y_{2,i}\right], \qquad (19) \\ S_{1/2} &= n^{1/2}B_0^{-1/2}\left[n^{-1}\sum_i Y_{1,i}\right]\left[n^{-1}\sum_j Y_{3,j}\right] \\ &\quad -n^{1/2}\frac{1}{2}B_0^{-3/2}\left[n^{-1}\sum_i Y_{2,i}\right]\left[n^{-1}\sum_k Y_{4,k}\right]. \end{aligned}$$

A.1.2 The Asymptotic Bias Term $k_{1,2}$

Taking expectations of the individual terms of (16) immediately yields, because of centering, $E(S_0) = 0$, and $E(S_{1/2}) = n^{-1/2}(B^{-1/2}E(Y_1Y_3) - 0.5B^{-3/2}E\left(Y_2Y_4\right))$. Since $E\left(S\right) = n^{-1/2}k_{1,2} + O\left(n^{-1}\right)$ it follows immediately that

$$k_{1,2} = B_0^{-1/2}E(Y_1Y_3) - \frac{1}{2}B_0^{-3/2}E\left(Y_2Y_4\right), \qquad (20)$$

with $E(Y_1Y_3) = M_2$ and $E\left(Y_2Y_4\right) = M_5$ stated explicitly in Proposition 1.

A.1.3 The Asymptotic Skewness Term $k_{3,1}$

In order to derive the asymptotic skewness term, we first need to obtain an expansion of the third moment of S. We take expectations of

$$S^3 = \left(S_0 + S_{1/2} + O_p\left(n^{-1}\right)\right)^3 = S_0^3 + 3S_0^2S_{1/2} + O_p\left(n^{-1}\right).$$

by considering the constituent parts separately.

1. $E\left(S_0^2S_{1/2}\right) = n^{3/2}B_0^{-3/2}E\left(n^{-4}\sum_i\sum_j\sum_k\sum_l Y_{2,i}Y_{2,j}Y_{1,k}Y_{3,l}\right) +$ $-0.5n^{3/2}B_0^{-5/2}\times E\left(n^{-4}\sum_i\sum_j\sum_k\sum_l Y_{2,i}Y_{2,j}Y_{2,k}Y_{4,l}\right)$.

   Since we are only interested in the $O\left(n^{-1/2}\right)$ term, we conclude that

$$\begin{aligned} E\left(S_0^2S_{1/2}\right) &= n^{-1/2}B_0^{-3/2}\times \\ &\quad \left[E\left(Y_2^2\right)E\left(Y_1Y_3\right) + 2E\left(Y_1Y_2\right)E\left(Y_2Y_3\right) - \frac{3}{2}E\left(Y_2Y_4\right)\right] \\ &\quad +O\left(n^{-1}\right), \end{aligned}$$

   after noting that $E\left(Y_2^2\right) = B_0$.

2. Consider $S_0^3 = n^{3/2}B_0^{-3/2}n^{-3}\left(\sum Y_{2,i}\right)^3$. Hence $E(S_0^3) = n^{-1/2}B_0^{-3/2}E(Y_2^3) + O\left(n^{-1}\right)$.

In summary

$$\begin{aligned} E(S^3) &= n^{-1/2}B^{-3/2} \times && (21) \\ & \left(E(Y_2^3) + 3\left[E\left(Y_2Y_2\right)E\left(Y_1Y_3\right) + 2E\left(Y_1Y_2\right)E\left(Y_2Y_3\right) - \frac{3}{2}E\left(Y_2Y_4\right)\right]\right) \\ & +O\left(n^{-1}\right). \end{aligned}$$

Finally, since $K = n^{-1/2}k_{31} + O\left(n^{-3/2}\right)$, and $K_3 = E\left(S^3\right) - 3E\left(S^2\right)E\left(S\right) + 2\left(E\left(S\right)\right)^3$, using $E\left(S^2\right) = 1 + O\left(n^{-1}\right)$, (20) and (21) we conclude that

$$k_{3,1} = B_0^{-3/2}\left[E(Y_2^3) + 6E\left(Y_1Y_2\right)E\left(Y_2Y_3\right) - 3E\left(Y_2Y_4\right)\right], \qquad (22)$$

where $E(Y_2^3) = M_4$, $E\left(Y_1Y_2\right) = M_1$, and $E\left(Y_2Y_3\right) = M_3$ are stated explicitly in Proposition 1.

B Simulation Evidence for $k_{1,2}$ and $k_{3,1}$

This section provides a comparison of the population bias and skewness coefficients $k_{1,2}$ and $k_{3,1}$ as defined in Proposition 1 and simulated k-statistics. The experiments are designed as follows. We draw R independent samples of size n from income distribution F and index the iteration by subscript r with $r = 1, 2, ..., R$. The resulting studentized inequality measure is denoted S_r.

The scaled k-statistics are defined as follows. Consider the first cumulant of S for which we have $K_1 = n^{-1/2}k_{1,2} + O\left(n^{-3/2}\right)$ or

$$k_{1,2} = n^{1/2}K_1 + O\left(n^{-1}\right).$$

The cumulant K_1 is simulated using the k-statistic $\widehat{K}_1 = R^{-1}\sum_r S_r$ with $\widehat{K}_1 = K_1 + O_p\left(R^{-1/2}\right)$. Therefore

$$\begin{aligned} k_{1,2}^{sim} &\equiv n^{1/2}\widehat{K}_1, \\ &= k_{1,2} + O\left(n^{-1}\right) + O_p\left(n^{1/2}R^{-1/2}\right). \end{aligned}$$

Similarly

$$\begin{aligned} k_{3,1}^{sim} &\equiv n^{1/2}\widehat{K}_3, \\ &= k_{3,1} + O\left(n^{-1}\right) + O_p\left(n^{1/2}R^{-1/2}\right), \end{aligned}$$

where $\widehat{K}_3 = rR^{-1}\sum_r\left(S_r - R^{-1}\sum_r S_r\right)^3$ with correction factor $r = R^2/[(R-1)(R-2)] \rightarrow 1$ which ensures unbiasedness of this k-statistic.

Figure 7 depicts both $k_{1,2}$ and $k_{1,2}^{sim}$, and $k_{3,1}$ and $k_{3,1}^{sim}$ as functions of the sensitivity parameter α of the inequality measure for the Singh-Maddala SM(.,3.5, 3.5) income distribution. The simulated values are based on n=10^3 and R=10^6 replications. The simulated values are in good agreement with the theoretical values. We have repeated these experiment for various income distributions and arrive at similar conclusions.

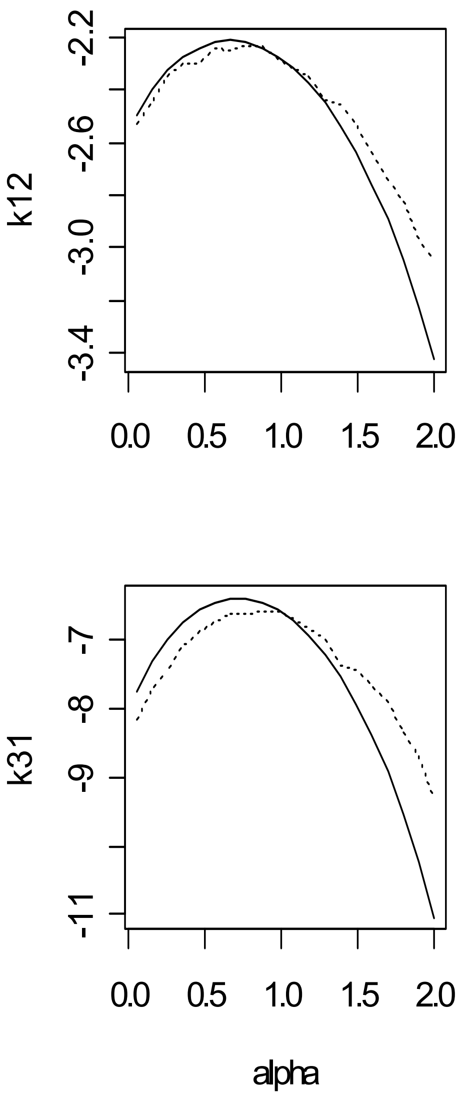

Figure 7: Theoretical bias and skewness coefficients and simulated k-statistics for the SM(.,3.5,3.5) income distribution. Notes: The solid lines depict the population coefficients $k_{1,2}$ and $k_{3,1}$ as a function of α. The dashed lines are the simulated k-statistics with $n = 1000$ and $R = 10^6$ repetitions.