Improving finite sample confidence intervals for inequality and poverty measures

van Garderen, K.J.; Schluter, Christian

UvA ECONOMETRICS

Discussion Paper: 2003/02

Improving finite sample confidence intervals for inequality and poverty measures

Kees Jan van Garderen and Christian Schluter

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

Department of Quantitative Economics
Faculty of Economics and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

Improving Finite Sample Con¯dence Intervals for Inequality and Poverty Measures[¤]

Kees Jan van Garderen[y]
University of Amsterdam

Christian Schluter[z]
University of Southampton

March 2003

Abstract

It is shown that standard con¯dence intervals for inequality and poverty (`welfare') measures have actual coverage errors in ¯nite samples which substantially exceed their nominal levels. The reason for this discrepancy, in some cases by a factor 4.5, is the heavy skewness of the actual distribution of the studentised welfare measure. This paper develops analytical methods to obtain reliable con¯dence intervals based on a normalising transform. This transform is surprisingly simple: $-3I^{-1/3}$ for welfare measure I. We provide analytic bias- and variance corrections and the ¯nite sample distribution of the adjusted studentised transform is much closer to the asymptotic Gaussian distribution.

Keywords: inequality and poverty measures, higher order expansions, normalising transformations, small sample inference.

JEL classi¯cation: C10, C14, D31, D63, I32

[¤]Funding from the ESRC under grant R000223640 is gratefully acknowledged. The research of van Garderen has been made possible 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.

[y]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: Garderen@fee.uva.nl

[z]Corresponding author. Department of Economics, University of Southampton, High¯eld, 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/sta®/schluter/

1 Introduction

Most attention in the statistical literature on inequality and poverty measures ('welfare measures' for short) has focused on the asymptotic properties of their estimators (see e.g. Davidson and Duclos, 1997 and Cowell, 1989). Their finite sample properties have rarely been considered, and it is this gap which we address. It is common practice to use confidence intervals based on the well-known asymptotic properties, and invert the Gaussian limit distribution in order to obtain the confidence limits irrespective of sample size. We therefore investigate first the performance of these usual confidence intervals in finite samples. It turns out that these typically grossly overstate the precision of the estimate. In some of the cases considered, the actual coverage error rate can exceed the nominal rate by a factor of 4.5. Large discrepancies persist even in samples of 500 observations. This poor performance stems from the substantial skewness of the actual finite sample distributions relative to the Gaussian limit distribution. Our goal is to develop analytical methods which remedy this poor performance.

The root cause of the problem is diagnosed to be the heavy skewness of the actual finite sample distribution. We therefore propose improved confidence intervals based on a normalising transform which seeks to annihilate this skewness. This transform of welfare measure I turns out to be surprisingly simple: $t(I) = -3I^{-1/3}$, irrespective of the income distribution considered. The resulting small sample distribution is much closer to the asymptotic Gaussian limit distribution than is the distribution of the original standardised measure. The idea underlying the normalising transform is to obtain an expansion of the third cumulant of the studentised transform, and focus on its first term, i.e. the coefficient of $n^{-1/2}$. We choose that transform which, at least approximately, reduces this coefficient to zero.

We also derive analytical bias and variance corrections that can be implemented by simply substituting sample moments into the relevant formulae. Our methods do not require any further derivations for its implementation, nor do they require computer intensive techniques or simulations. We show that they work very well for most inequality measures and all poverty indices considered: the new confidence intervals are shown to exhibit actual coverage error levels in good agreement with their nominal levels.

The issues raised in this paper are of practical relevance, for instance when evaluating policies in terms of welfare or when seeking to identify geographical areas or socio-economic groups as policy targets. These exercises require testing, such as possibly comparing inequality or poverty before and after intervention, or cross-region comparisons. Using the usual confidence intervals, which are too short and thus less precise than suggested by their nominal levels, can lead to the wrong inference, and therefore to wrong targeting or the adoption of the wrong policy. Another example is the case of decomposition analyses. A large body of applied work is frequently interested in the welfare of socio-economic subgroups and their contributions to overall welfare. Although the overall sample may be large, the partition into subgroups can result in subsamples which are comparatively small. The final example is the focus of the new macroeconomics growth literature on the distribution and inequality of incomes amongst nations. Inequality indices are frequently computed using the Summers and Heston data set contains which about 122 countries.

The paper is organized as follows. Section 2 presents the specific welfare measures, and standard 'first order' confidence intervals based on the Gaussian limit distribution. Section 3 examines their performance in practice, and reveals the great need for improved confidence intervals. The normalising transform is introduced in Section 4, and its performance is evaluated in Section 5. Section 6 concludes. All explicit derivations are collected in a technical appendix.

2 Welfare Measures and First Order Confidence Intervals

The objective is to construct confidence intervals for a given welfare measure I (inequality or poverty measure) based on a random sample of incomes X_i, $i = 1, \ldots, n$, of size n from an income distribution F_X. The measure I is a functional that maps a distribution F_X into a scalar, and the commonly used estimator $\hat{I} = I(\hat{F}_X)$, simply uses the empirical distribution function (EDF) $\hat{F}_X$ of F_X, $\hat{F}_X(x) = n^{-1}\sum_i 1(X_i \le x)$. In many cases $I(F_X)$ only depends on the moments of the distribution and the EDF estimator is then obtained by replacing the population moments by the empirical moments. The asymptotic variance $\sigma^2 = Var(n^{1/2}(\hat{I} - I))$ is obtained by the delta method, and estimated by an EDF-based estimator, denoted by $\hat{\sigma}^2$.

The standard 'first order' approach to constructing confidence intervals uses the studentised welfare measure,

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

which, under standard assumptions, has a distribution that converges asymptotically to the Gaussian distribution, denoted by Φ. This leads to the basic nominal $100\alpha\%$ percent confidence interval for the welfare measure typically used in applied research:

$$\hat{I} - \frac{\hat{\sigma}}{\sqrt{n}}\Phi^{-1}\left(\frac{1+\alpha}{2}\right) \le I \le \hat{I} - \frac{\hat{\sigma}}{\sqrt{n}}\Phi^{-1}\left(\frac{1-\alpha}{2}\right): \tag{2}$$

The case of a nominal level of 95% is examined extensively below in Section 3. Confidence intervals based on the Gaussian quantiles ± 1.96 turn out to be very poor in practice: the actual coverage failure is much larger than the nominal 5%. In short, the actual finite sample distribution of S is poorly approximated by the Gaussian distribution Φ. In particular, the distribution of S is heavily skewed. In Section 4 we propose a transformation of S which reduces this skewness and thus renders its distribution closer to the Gaussian limit.

2.1 Specific Welfare Indices

We consider two leading classes of inequality and poverty: the Generalised Entropy indices, and the poverty indices proposed in Foster, Greer and Thorbecke (FGT, 1984).

Generalised Entropy indices are defined by

$$GE_\alpha(F) = \frac{1}{\alpha^2 - \alpha}\left[\frac{\mu_\alpha(F)}{\mu_1(F)^\alpha} - 1\right] \qquad \text{for } \alpha \neq \{0; 1\}; \tag{3}$$

where α is a sensitivity parameter and $\mu_\alpha(F) = \int x^\alpha dF(x)$ is the moment functional. This inequality index is of particular interest because it is the only inequality measure that simultaneously satisfies the property of scale independence, and the principles of transfer and decomposability, and the population principle. The smaller is the sensitive parameter α, the larger is the sensitivity of the inequality index to the lower tail of the income distribution. The index, however, is not monotonic in α. If $\alpha = 2$ the index equals half the coefficient of variation squared. This value could be considered already quite large as it gives relatively large weight to the upper tail of the income ditstribution. Two popular values, the limit cases of $\alpha = 0$ and $\alpha = 1$; are better known as Theil indices. We therefore denote them by Theil_1 and Theil_0.[1] Another popular inequality measure is the Atkinson (1970) index, defined by $A_\varepsilon(F) = 1 - [\mu_{1-\varepsilon}(F)]^{1/1-\varepsilon}\mu_1(F)^{-1}$, where $\varepsilon \geq 0$ is a parameter reflecting (relative) inequality aversion. However, since the Atkinson index can be mapped into the Generalised Entropy index by setting $1 - \varepsilon = \alpha$, $GE_\alpha = [\alpha^2 - \alpha]^{-1}\left[\left(1 - A_{1-\alpha}\right)^\alpha - 1\right]$, it will not consider separately. For an extensive discussion of the properties of the Generalised Entropy index see Cowell (1980, 2000). First order methods for inequality measures have been considered in e.g. Cowell (1989) and Thistle (1990).

The FGT poverty indices are of the form

$$P_\alpha(F) = \int \left(1 - \frac{x}{z}\right)^\alpha 1(x \leq z) \quad dF(x); \qquad \text{for } \alpha > 0; \tag{4}$$

where z denotes the poverty line, which we assume to be distribution invariant, and $1(.)$ is the indicator function. α is the sensitivity parameter, which determines the weight of an income shortfall from the poverty line. This large class of indices includes the Headcount index when $\alpha = 0$ and the Poverty Gap index when $\alpha = 1$. First order methods have been proposed by Kakwani (1993).

3 The Need for Improved Confidence Intervals

In order to investigate the finite sample performance of the usual first order confidence intervals given by (2), we have carried out a simulation study varying over income distributions, sample sizes, and welfare measures. In particular, we have examined the performance for the GE_2, and the Theil indices, as well as for the Headcount, the Poverty Gap, and P_2.

[1] Rather treating these two special cases separately, we exploit the continuity of the index in α, and approximate them by $GE_{1.05}$ and $GE_{0.05}$.

3.1 Simulation Design: Income Distributions

We consider three classes of income distributions that are often used in practice.

1. The Singh-Maddala (1976) distribution $SM(a, b, c)$. Its density

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

   is a special case of the Generalized Beta distribution (McDonald, 1984). The population moments are given by $\mu_\alpha = ca^\alpha \Gamma(c - \alpha/b)\,\Gamma(1 + \alpha/b)/\Gamma(1 + c)$ where Γ denotes the Gamma function, and immediately lead to the population measure of GE_α. We note that the Singh-Maddala distribution is heavy-tailed, as its tails decay like power functions (to be precise, the (right) tail index equals bc).

2. The lognormal $LN(\mu, \sigma^2)$, which implies a population inequality index equal to $GE_\alpha = (\alpha^2 - \alpha)^{-1}\left[\exp\left((1/2)\sigma^2(\alpha - 1)\alpha\right) - 1\right]$, which is independent of μ.

3. The Gamma distribution $G(r, \lambda)$ with shape parameter r and scale parameter λ, $f(x; r, \lambda) = e^{-\lambda x}x^{r-1}\lambda^r/\Gamma(r)$. The population moments are $\mu_\alpha = \Gamma(r + \alpha)\,\lambda^{-\alpha}/\Gamma(r)$, which imply a population inequality index independent of the scale parameter λ, $GE_\alpha = (\alpha^2 - \alpha)^{-1}\left(r^{-\alpha}\Gamma(\alpha + r)/\Gamma(r) - 1\right)$.

In case of the poverty measures, given the truncation of the income distribution at the poverty line, numerical methods have to be used. We have chosen a poverty line such that 20% of the population are in poverty.

These income distributions fit real world data reasonably well. The actual distributions used in the investigations are $SM(100, 2.8, 1.7)$ as reported by Brachmann et al.(1996), and $LN(1, 0.5^2)$ as reported by Biewen (2001) for German income data. Cowell and Feser (1996) have used $G(3, 0.15)$.

3.2 Simulation Evidence

Table 1 records the incidence of coverage failures of symmetric nominal 95%-confidence intervals based on first order methods, given by equation (2). The experiment involved 100,000 repetitions and the three sample sizes 100, 250, and 500.

Consider the inequality measures first. The table makes abundantly clear that empirical coverage failures are substantially larger than the nominal value of 5 per cent: in one case (GE_2, SM, and $n = 100$) up to 4.5 times the nominal value. The coverages failures fall as the sensitivity parameter α of the Generalised Entropy Index falls (giving less weight to the upper tail of the income distribution), the right tail of the income distributions decay more rapidly, and the sample size increases, but the extent of the failure remains considerable.

Coverage failures are expected to be better for the poverty indices, given their simpler linear structure and the truncation of the income distribution at the poverty

	$SM(100;2.8;1.7)$			$LN(1;0.5^2)$			$G(3;0.15)$		
sample size	100	250	500	100	250	500	100	250	500
GE_2	22.6	18.2	15.7	16	11.9	9.5	9.9	7.5	6.5
$Theil_1$	13.1	10.1	8.7	11.1	8.3	7.1	7.7	6.2	5.7
$Theil_0$	9.4	7.2	6.4	8.6	6.6	6.0	7.6	6.1	5.5
sample size	100	250	500	100	250	500	100	250	500
P_2	10.4	6.7	6.4	9.4	7.1	6.7	8.9	6.9	5.4
P_1	7.4	6.1	5.7	7.4	5.8	5.5	7.4	5.7	5.6
P_0	6.7	5.1	5.2	6.6	5.1	5.0	7.0	5.0	5.4

Table 1: Actual coverage failure in per cent of usual nominal 95 per cent confidence intervals for the welfare indices. Based on 100,000 replications.

line (so that the speed of decay of the right tail of the income distribution becomes immaterial). In particular P_0 is expected to perform well since one is estimating only a proportion. While Table 1 shows that this is indeed the case, the coverage failure can still be twice the nominal rate (e.g. for P_2 and n=100).

The worst performer, GE_2 estimated with n=100 and incomes drawn from the heavy-tailed Singh-Maddala distribution, is further examined in Figure 1. Depicted is the simulated finite sample density of the studentised inequality measure[2], and the Gaussian limit distribution. The difference between these two densities is substantial. In particular, the finite sample distribution is heavily skewed.

Figure 1 about here.

In summary, the results show a great need for improved confidence intervals. All first order confidence intervals are too short, and in some cases the actual coverage failure rate can be as much as 4.5 times higher than the nominal rate. Inference based on such confidence intervals can therefore be seriously flawed in practice, and it is precisely the first order confidence intervals which applied researcher typically use.

4 The Normalising Transform

The large disparity between the actual and nominal confidence levels is caused by the heavy skewness of the finite sample distribution of the studentised welfare measure S. In other words, the third cumulant of S, $K_{3;n} = E(S^3) - 3E(S^2)E(S) + 2(E(S))^3$, is non-zero. Therefore, as $K_{3;n}$ admits an expansion of the form

$$K_{3;n} = n^{-1/2}k_{3;1} + O(n^{-3/2});$$

it implies that $k_{3;1}$ is substantially different from zero. Our approach is to seek a transformation that eliminates this skewness, and entails a distribution much closer

[2] The density has been estimated using kernels, whose bandwidth was chosen using a cross-validation method.

to the Gaussian limit. Denote by t the transform of the welfare measure, and by T the standardised version of t,

$$T = \sqrt{n}\frac{t(\hat{I}) - t(I)}{\hat{\sigma} t'(\hat{I})}. \tag{5}$$

In the appendix we derive the expansion of the third cumulant of T, and show that the coefficient of $n^{-1/2}$ is reduced to zero if the transform t satisfies the following differential equation[3]

$$3\,\sigma(I)\,t''(I) - k_{3,1}(I)\,t'(I) = 0. \tag{6}$$

The specific form of the differential equation depends on the particular welfare measure and on properties of the income distribution since the quantities σ, $k_{3,1}$, and I are linked through the parameters of F. In principle this differential equation can be solved exactly for a given welfare measure and income distribution. As these vary, so will the exact solution, generating a large collection of solutions. Surprisingly, however, we are able to show (see appendix) that for broad classes of income distributions an approximate solution is given by

$$t(I) = -3I^{-1/3}, \tag{7}$$

which does not depend on the specific income distribution. The simulation results of Section 5 show that this transformation also works well for income distributions where the solution to (6) is not analytically tractable.

4.1 Bias Correction

Apart from skewness, the finite sample distribution of S also suffers from a bias

$$E(S) = n^{-1/2}k_{1,2} + O(n)$$

where the coefficient $k_{1,2}$ is non-zero. The distribution of the transformed statistic T inherits such bias because

$$E(T) = n^{-1/2}\lambda_{1,2} + O(n),$$

with $\lambda_{1,2} = k_{1,2} - (1/2)\sigma t''(I)/t'(I)$. Given the particular transform in (7) and substituting the empirical moments the estimate becomes

$$\hat{\lambda}_{1,2} = \hat{k}_{1,2} + \frac{2\hat{\sigma}}{3\hat{I}}, \tag{8}$$

where $k_{1,2}$ is defined in equation (17) in the appendix. It should be noted that estimation of $k_{1,2}$ and $\lambda_{1,2}$ does not affect the order of our approximation.

[3] The general approach has been first suggested by Niki and Konishi (1984). As they examine the infeasible case $n^{-1/2}\left(t(\hat{I}) - t(I)\right)/\sigma t(I)$, the resulting differential equation is different from ours.

4.2 Variance Correction

The quality of the standardised transform (5) is, because of the studentisation, directly affected by the quality of the estimate $\hat{\sigma}$. Since the variance estimator $\hat{\sigma}^2$ is a function of the moments, just like $\hat{\lambda}$, we improve the estimate again by applying a finite sample bias correction to each raw moment appearing in the variance formula. Taking expectations of the second order stochastic expansion yields an additive correction term for the unstudentised sample moment which is estimated using the sample moments, giving

$$\hat{\mu}_a = m_a + \frac{1}{n}\frac{(m_{3a} - m_{2a}m_a) - 2m_a(m_{2a} - m_a^2)}{2(m_{2a} - m_a^2)}.$$

These are substituted in the variance estimator $\hat{\sigma}^2$, given in (11) of the appendix for the GE_α measure. For the poverty measures it turns out that the correction for the variance is not necessary, and therefore not presented, since the performance of the transform and the bias correction is already very good.

4.3 The Improved Confidence Intervals

To summarise, combining the normalising transform with the bias correction, and variance improvement leads to the improved nominal 95% percent confidence interval based on $-1.96 \le T - n^{-1/2}\widehat{\lambda_{1,2}} \le 1.96$. In terms of the original (untransformed) inequality index this equals

$$\left(\hat{I}^{-1/3} + \hat{I}^{-4/3}n^{-1/2}\frac{\hat{\sigma}}{3}(1.96 + n^{-1/2}\widehat{\lambda_{1,2}})\right)^{-3} \le I \le \left(\hat{I}^{-1/3} - \hat{I}^{-4/3}n^{-1/2}\frac{\hat{\sigma}}{3}(1.96 - n^{-1/2}\widehat{\lambda_{1,2}})\right)^{-3}; \quad (9)$$

where $\hat{\sigma}^2$ is now the improved variance estimator.

5 Improved Confidence Intervals: Simulation Evidence

The simulation design is identical to the one of Section 3 and, as before, we take the nominal level to be 95 per cent.

Table 2 reports the performance of the new confidence intervals given by (9). For the inequality measures both the bias and variance corrections are applied. They each consitute a significant improvement over using our normalising transform only. In Appendix B the individual contribution of each adjustment is quantified for the inequality measure. For the poverty indices, just using the the transform and the bias correction already yields very good results. We therefore report only results for the poverty indices without the additional variance correction.

	SM (100; 2:8; 1:7)			LN (1; 0:5^2)			G (3; 0:15)		
sample size	100	250	500	100	250	500	100	250	500
GE_2	14.8	12.4	10.7	10.1	7.7	6.7	6.2	5.5	5.3
$Theil_1$	7.7	6.9	6.3	6.8	5.6	5.3	5.1	4.9	5.0
$Theil_0$	5.3	5.0	4.9	5.0	4.9	4.9	4.7	4.8	5.0
sample size	100	250	500	100	250	500	100	250	500
P_2	6.7	6.2	5.3	6.8	5.7	5.0	6.8	5.3	6.1
P_1	5.2	4.9	5.0	5.1	5.1	4.8	4.8	5.5	4.6
P_0	3.4	4.9	5.0	3.6	4.7	5.4	3.6	5.1	5.1

Table 2: Actual coverage failure in per cent of new confidence intervals. Nominal level 5 per cent based on 100,000 replications.

Consider the poverty measures first. The actual coverage rates are now very close to their nominal 5 per cent level, irrespective of income distribution and sample size.

Turning to the inequality indices, we see that the methods also perform well for the Theil measures, and for GE_2 in the case of the Gamma and Lognormal distribution.

The three distributions used have in common their skewness but are otherwise quite different. Our new method therefore does deliver substantial improvements over first order methods for the variety of income distributions and sample sizes considered.

Room for improvement remains, however, in the case of GE_2 with the heavy-tailed Singh-Maddala distribution. Although this paper is primarily concerned with analytical methods, we did investigate alternative techniques. In particular computer intensive methods, using the bootstrap on the original statistic and on the transformed statistic. The bootstrap on the transform might be expected to work better than the untransformed statistic because it is closer to being pivotal. Neither bootstrap method provides significant improvement for the problem case. For instance the bootstrap on the transformed statistic yields a coverage failure of 15.5,12.8,10.8, for GE_2 with SM (100; 2:8; 1:7); and 100, 250 and 500 observations, which is actually worse than the results in Table 2 for our analytical methods.

5.1 The Bootstrap and Multiplicative Adjustment Factor

Although this paper focusses on analytic methods and avoid computer intensive simulation techniques, one might expect the bootstrap to yield further small sample improvements. The normalising transform does not yield zero skewness exactly in small samples and the critical values might also differ from the Gaussian ones.

Among the various bootstrap methods available, it is well known that the studentised bootstrap, based on the asymptotically pivotal quantity such as T; typically yields confidence intervals which whose coverage behaviour is better than those of other bootstrap techniques such as the BC_{α} or percentile method (Beran, 1988, Hall, 1988, Hall and Horowitz, 1996, Davison and Hinkley, 1997). We have verified that this also holds in the present case. In line with these findings our approach is to first apply the nonlinear transform to the welfare measures, and then to apply the studentised bootstrap. We have followed the recommendation of Davison and Hinkley (1997) and have used, in each of 10,000 iterations, $R=999$ bootstrap samples in order to obtain one confidence interval. Table 3 shows that some improvements in coverage behaviour, beyond those reported in Table 2, do indeed obtain.

T	$SM(100, 2.8, 1.7)$			$LN(1, 0.5^2)$			$G(3, 0.15)$		
sample size	100	250	500	100	250	500	100	250	500
GE_2	13.9	11.4	10.4	9.6	7.7	7.2	6.7	6.0	5.5
$Theil_1$	7.7	7.3	6.7	7.1	6.4	5.9	5.3	5.1	5.3
$Theil_0$	6.1	5.6	5.3	5.6	5.1	5.3	5.4	4.7	4.8

Table 3: The studentised bootstrap and the normalising transform: Actual coverage failure in per cent of nominal 95 per cent confidence intervals. Based on 10,000 replications.

We have also investigated the merit of a multiplicative factor $(1 - n^{-1/2})$, which pulls in the tails of the actual finite sample distribution of T. The implementation of this adjustment is very simple as it requires only the multiplication of the simple nonlinear transform by a constant, if the bias correction is omitted and it avoids computer intensive bootstrap replications.

	$SM(100, 2.8, 1.7)$			$LN(1, 0.5^2)$			$G(3, 0.15)$		
sample size	100	250	500	100	250	500	100	250	500
GE_2	14.3	13.2	12.0	9.48	7.93	7.19	4.96	4.97	4.91
$Theil_1$	6.97	6.67	6.45	6.07	5.54	5.37	3.86	4.19	4.3
$Theil_0$	4.86	4.74	4.79	4.46	4.46	4.57	3.82	4.05	4.21

Table 4: Actual coverage failure in per cent of confidence interval based on the adjusted transform (i.e. $(1 - n^{-0.5})T$). Nominal level 5 per cent based on 100,000 replications.

The adjustment yields an improvement for $n = 100$, but is otherwise dominated by Table 2. The following table reports the result of applying the factor to the bias corrected T. This leads to some improvements but is again dominated by Table 2.

	$SM(100, 2.8, 1.7)$			$LN(1, 0.5^2)$			$G(3, 0.15)$		
sample size	100	250	500	100	250	500	100	250	500
GE_2	12.9	12.3	11.5	8.7	7.7	6.9	5.0	4.9	4.9
$Theil_1$	6.5	6.6	6.5	5.6	5.4	5.3	3.9	4.1	4.3
$Theil_0$	4.6	4.7	4.8	4.2	4.5	4.4	3.7	4.1	4.2

Table 5: The adjusted normalising transform and the bias correction $((1 - n^{-1/2})\{T - n^{-1/2}\widehat{\lambda}_{1,2}\})$: Actual coverage failure in per cent of nominal 95 per cent confidence intervals. Based on 100,000 replications.

6 Conclusions

Standard confidence intervals for inequality and poverty measures, based on the usual first order methods, perform poorly in finite samples. Evidence from our simulation study reveals that the actual coverage failure can be as much as 4.5 times higher than the nominal rate. Inference based on such confidence intervals can therefore be seriously flawed in practice because it gives a false sense of precision. It is precisely these first order confidence intervals which applied researcher typically use.

In this paper we have developed analytical methods to obtain reliable confidence intervals. We derive a simple normalising transform which annihilates the (asymptotic) skewness, which is at the heart of the poor performance of first order methods. As we are actually dealing with small samples we further provide analytic bias and variance corrections. These corrections do not require any additional derivations or computer intensive methods or simulations, but only need the substitution of sample moments into the formulae. The new confidence intervals are shown to have actual coverage errors in close agreement with their nominal levels.

The only case in which our methods did not fully resolve the poor finite sample performance is the GE_2 measure with the Singh-Maddala distribution. This is caused by the fact that the Singh-Maddala distribution has heavy tails and the GE_2 measure gives relatively high weight to the upper tail of the income distribution. One might therefore want to be cautious when using GE_2 in practice.

References

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

Beran, R. (1988) \Prepivoting test statistics: a bootstrap view of asymptotic re-¯nements", Journal of the American Statistical Association, 83, 687-697.

Biewen, M. (2001) \Bootstrap inference for inequality, poverty and mobility measurement", Journal of Econometrics, vol. 108, 2, 317-342.

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

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.

Davison, A.C. and D.V. Hinkley (1997) Bootstrap methods and their applications, Cambridge Cambridge University Press.

Efron, B. and R.J. Tibshirani (1993), An Introduction to the Bootstrap, Chapman & Hall.

Foster, J. E., J. Greer, and E. Thorbecke (1984) \A class of decomposable poverty measures", Econometrica, 52, 761-776.

Hall, P. (1988), \Theoretical comparisons of bootstrap con¯dence intervals", Annals of Statistics, 16, 927-985.

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

Hall, P. and J.L. Horowitz, (1996) \Bootstrap critical values for tests based on generalized-method-of-moments estimators", Econometrica, 64, 891-916.

Kakwani, N. (1993) \Statistical inference in the measurement of poverty", The Review of Economics and Statistics, 75, 632-639.

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

Niki, N., and S. Konishi (1986) \E®ects of Transformations in Higher Order Asymptotic Expansions", Annals of the Institute of Statistical Mathematics, 38, 371-383.

Singh, S.K. and G.S. Maddala (1976) \A Function for Size Distribution of Incomes", Econometrica, 44, 963{970.

Thistle, P. D. (1990) \Large sample properties of two inequality indices", Econometrica, 58, 725-728.

A Technical Details.

A.1 Asymptotic Justifications

The merit of applying the normalizing transform and the bias correction becomes clear once the (second order) Edgeworth expansions of the distribution of the studentised statistics S and T are examined. The former can be written as

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

where $\Phi(x)$ and $\phi(x)$ denote the Gaussian cumulative distribution function and density (see e.g. Hall, 1992). A similar Edgeworth expansion also applies to the studentised transform,

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

where $\kappa_{3;1}$ is the coefficient of $n^{-1/2}$ of the third cumulant of T. $\kappa_{3;1}$ is shown below to be proportional to the left hand side of the differential equation (6). Annihilating the asymptotic skewness term $\kappa_{3;1}$ by an appropriate choice of the transform t and applying the bias correction yields a distribution closer to the Gaussian.

$\kappa_{3;1}$ can be derived explicitly by working out the first three moments of T and collecting all coefficients of $n^{-1/2}$. As this is fairly tedious, we consider only the first moment explicitly here. Higher moments can derived using similar arguments. Frequently, second order Taylor expansions of $t(\hat{\theta})$ will have to be taken, as well as the facts that $E(S^2) = 1 + O(n^{-1/2})$ and $E(S^4) = 3 + O(n^{-1/2})$ have to be used. Consider

$$\begin{aligned} T &= \frac{n^{1/2}}{\sigma}\frac{t(\hat{\theta}) - t(\theta)}{t'(\hat{\theta})} \\ &= \frac{n^{1/2}}{\sigma}\frac{t(\hat{\theta}) - t(\theta)}{t'(\theta)}\left[\frac{t'(\hat{\theta})}{t'(\theta)}\right]^{-1} \\ &= \frac{n^{1/2}}{\sigma}\frac{t(\hat{\theta}) - t(\theta)}{t'(\theta)}\left[1 - \frac{t''(\theta)}{t'(\theta)}\left(\hat{\theta} - \theta\right) + O_p\left(n^{-1}\right)\right]; \end{aligned}$$

the last line following after two expansions. Multiply out and take expectations. The expectation of the second term is $E\left(-\frac{n^{1/2}}{\sigma}\frac{t(\hat{\theta}) - t(\theta)}{t'(\theta)}\frac{t''(\theta)}{t'(\theta)}\left(\hat{\theta} - \theta\right)\right) = -n^{-1/2}\sigma\frac{t''(\theta)}{t'(\theta)} E(n\sigma^{-2}(\hat{\theta} - \theta)^2) + O(n^{-1}) = -n^{-1/2}\sigma t''(\theta)/t'(\theta) + O(n^{-1})$ since $E(S^2) = 1 + O(n^{-1/2})$. This fact and the expansion of $t(\hat{\theta})$ is used again to derive the expectation of the first term $E\left(\frac{n^{1/2}}{\sigma}\frac{t(\hat{\theta}) - t(\theta)}{t'(\theta)}\right) = n^{-1/2}(E(S) + 0.5\sigma\frac{t''(\theta)}{t'(\theta)}E(S^2)) + O(n^{-1}) = n^{-1/2}[k_{1;2} + 0.5\sigma t''(\theta)/t'(\theta)] + O(n^{-1})$. Putting everything together therefore yields $E(T) = n^{-1/2}\kappa_{1;2} + O(n^{-1})$ with $\kappa_{1;2} = [k_{1;2} - 0.5\sigma t''(\theta)/t'(\theta)]$ as claimed by (8). After deriving the next two moments and collecting the coefficients of $n^{-1/2}$, it follows that

$$\kappa_{3;1} = [3\sigma t''(\theta) - k_{3;1}t'(\theta)]/t'(\theta):$$

The notation $\sigma(I)$ and $k_{3;1}(I)$ has been used above to emphasize the dependence of these quantities on the welfare measure I. The standardized transform T has therefore zero asymptotic skewness, if $t(\cdot)$ solves the differential equation (6).

A.2 Derivation of the Normalising Transform

In this section we derive explicitly the normalizing transform given by equation (7). In order to solve the differential equation (6), we need to derive the asymptotic skewness term $k_{3;1}$. This in turn requires the derivation of a second order asymptotic expansion of the studentised inequality measure.

A.3 The Inequality Measure GE_α

In order to derive the key asymptotic skewness term $k_{3;1}$, we first need to obtain an asymptotic expansions of the moments of S. As a compact notation, we use S_q to denote a term of an expansion of S which is of order n^{-q}. Hence the desired stochastic expansion of S is given by

$$S = S_0 + S_{1/2} + O_p\left(n^{-1}\right); \tag{10}$$

where

$$\begin{aligned} S_0 &= n^{1/2} B^{-1/2}\left[n^{-1}\sum Y_{1;i}\right] \\ S_{1/2} &= n^{1/2} B^{-1/2}\left[\left[n^{-1}\sum_i Y_{3;i}\right]\left[n^{-1}\sum_j Y_{4;j}\right]\right] \\ &\quad - n^{1/2}\frac{1}{2} B^{-3/2}\left[n^{-1}\sum_k Z_{1;k}\right]\left[n^{-1}\sum_i Y_{1;i}\right]. \end{aligned}$$

The precise definitions of the variables B, $Z_{1;i}$ and $Y_{1;i}$ - $Y_{4;i}$ will be derived below.

We first derive explicitly the stochastic expansion for S, given by (10) and the precise variable definitions B, $Z_{1;i}$ and $Y_{1;i}$ - $Y_{4;i}$, given by equations (12) and (14) below. We then derive the asymptotic bias term $k_{1;2}$, see (17) below, and the asymptotic skewness term $k_{3;1}$, see (19) below. We are then in a position to derive the normalizing transform by solving the differential equation (6). Recall our notation for population and sample moments: $\mu_\alpha(F) = \int y^\alpha dF(y)$ and $m_\alpha = \mu_\alpha(\hat{F})$.[4]

A.3.1 The Stochastic Expansion of the Studentised Inequality Index

We derive the stochastic expansion in four steps:

[4] For expositional brevity, we treat only the case of GE_α explicitly. The case of the poverty index is much simpler, as the precise variable definitions in this case are: $Y_{1;i} = f(X_i) - E(f(X))$ and $Z_{1;i} = f(X_i)^2 - E(f(X)^2)$, with $f(x) = (1 - x/z)^\alpha 1(x \cdot z)$, and all others variables equalling zero. The truncation of the income distributions at the poverty line renders an explicit analysis of this case far less tractable.

1. Center the inequality index to obtain

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

2. Derive the asymptotic variance by applying the delta-method

$$\sigma^2 = Var(n^{1/2}(\widehat{GE}_\alpha - GE_\alpha)) = \frac{1}{(\alpha^2-\alpha)^2}\frac{1}{\mu_1^{2\alpha+2}}B, \tag{11}$$

with

$$B = \left[\alpha^2\mu_\alpha^2\mu_2 - 2\alpha\mu_1\mu_\alpha\mu_{\alpha+1} + \mu_1^2\mu_{2\alpha} - (1-\alpha)^2\mu_1^2\mu_\alpha^2\right] \tag{12}$$

or, equivalently, using the covariance function $\sigma_{a,b} = Cov\left(X^a, X^b\right) = \mu_{a+b} - \mu_a\mu_b$,

$$B = \mu_1^2\sigma_{\alpha,\alpha} + (\alpha\mu_\alpha)^2\sigma_{1,1} - 2\alpha\mu_\alpha\mu_1\sigma_{\alpha,1}.$$

This is estimated by $\hat{B}$ using the corresponding moments. The expansion of $\hat{B}^{-1/2}$ is

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

As regards $B_{1/2}$ it can be shown that, after centering and collecting terms of the same order,

$$\begin{aligned}B_{1/2} = 2&\left(\mu_1\sigma_{\alpha,\alpha} + \left(\alpha - \alpha^2\right)\mu_1\mu_\alpha^2 - 2\mu_\alpha\sigma_{\alpha,1}\right)n^{-1}\sum(X_i - \mu_1)\\ &+\alpha^2\mu_\alpha^2 n^{-1}\sum\left(X_i^2 - \mu_2\right)\\ +2&\left((\alpha-1)\mu_\alpha\mu_1^2 + \alpha^2\mu_\alpha\sigma_{1,1} - \alpha\mu_1\sigma_{\alpha,1}\right)n^{-1}\sum(X_i^\alpha - \mu_\alpha)\\ &-2\alpha\mu_1\mu_\alpha n^{-1}\sum\left(X_i^{\alpha+1} - \mu_{\alpha+1}\right) + \mu_1^2 n^{-1}\sum\left(X_i^{2\alpha} - \mu_{2\alpha}\right).\end{aligned}$$

This implies the expansion of $\hat{B}^{-1/2}$

$$\hat{B}^{-1/2} = B^{-1/2} - \frac{1}{2}B^{-3/2}\left[n^{-1}\sum_k Z_{1,k}\right] + O\left(n^{-1}\right).$$

with

$$\begin{aligned}Z_{1,i} = 2&\left(\mu_1\mu_{2\alpha} - \alpha\mu_\alpha\mu_{\alpha+1} - (1-\alpha)^2\mu_1\mu_\alpha^2\right)(X_i - \mu_1)\\ &+\alpha^2\mu_\alpha^2\left(X_i^2 - \mu_2\right)\\ +2&\left(\alpha^2\mu_\alpha\mu_2 - \alpha\mu_1\mu_{\alpha+1} - (1-\alpha)^2\mu_1^2\mu_\alpha\right)(X_i^\alpha - \mu_\alpha)\\ &-2\alpha\mu_1\mu_\alpha\left(X_i^{\alpha+1} - \mu_{\alpha+1}\right) + \mu_1^2\left(X_i^{2\alpha} - \mu_{2\alpha}\right).\end{aligned} \tag{13}$$

(obtained after multiplying out the covariance functions defining $B_{1/2}$).

3. Combine the results from steps 1 and 2 to get

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

4. Expand m_α and $m_1^{-\alpha}$ to order $O_p\left(n^{-1/2}\right)$[5] to get the stochastic expansion $S = S_0 + S_{1/2} + O_p(n^{-1})$ given by (10), the precise variables definitions being

$$Y_{1;i} = \mu_1\left(X_i^\alpha - \mu_\alpha\right) - \alpha\mu_\alpha\left(X_i - \mu_1\right) \tag{14}$$

$$Y_{3;i} = \left(X_i - \mu_1\right)$$

$$Y_{4;i} = \left(X_i^\alpha - \mu_\alpha\right) - \frac{\alpha(\alpha+1)}{2}\mu_\alpha\mu_1^{-1}\left(X_i - \mu_1\right).$$

and $Z_{1;i}$ given by (13).

Below, we will be interested in expectations of the form $E(Y_1Y_3)$ etc.:

$$\begin{aligned}
E(Y_1Y_1) &= \mu_1^2\sigma_{\alpha;\alpha} + (\alpha\mu_\alpha)^2\sigma_{1;1} - 2\alpha\mu_\alpha\mu_1\sigma_{\alpha;1} = B \qquad (15)\\
E(Y_1Y_3) &= \mu_1\sigma_{\alpha;1} - \alpha\mu_\alpha\sigma_{1;1}\\
E(Y_3Y_4) &= \sigma_{\alpha;1} - \frac{\alpha(\alpha+1)}{2}\mu_\alpha\mu_1^{-1}\sigma_{1;1}\\
E(Y_1Y_4) &= \mu_1\sigma_{\alpha;\alpha} - \frac{\alpha(\alpha+1)}{2}\mu_\alpha\sigma_{\alpha;1} - \alpha\mu_\alpha E(Y_3Y_4)\\
E(Y_1^3) &= \mu_1^3\left(\mu_{3\alpha} - 3\mu_{2\alpha}\mu_\alpha + 2\mu_\alpha^3\right)\\
&\quad - 3\alpha\mu_\alpha\mu_1^2\left(\mu_{2\alpha+1} - \mu_{2\alpha}\mu_1 - 2\mu_{\alpha+1}\mu_\alpha + 2\mu_\alpha^2\mu_1\right)\\
&\quad + 3(\alpha\mu_\alpha)^2\mu_1\left(\mu_{\alpha+2} - 2\mu_{\alpha+1}\mu_1 + 2\mu_\alpha\mu_1^2 - \mu_\alpha\mu_2\right)\\
&\quad - (\alpha\mu_\alpha)^3\left(\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3\right)\\
E(Y_1Z_1) &= 2\left(\mu_1^2\mu_{2\alpha} + \left(\alpha^2 - \alpha\right)\mu_1\mu_\alpha\mu_{\alpha+1} - \alpha^3\mu_2\mu_\alpha^2 + (\alpha-1)(1-\alpha)^2\mu_1^2\mu_\alpha^2\right)\sigma_{\alpha;1}\\
&\quad + \alpha^2\mu_1\mu_\alpha^2\sigma_{\alpha;2} + 2\mu_1\left(\alpha^2\mu_\alpha\mu_2 - \alpha\mu_1\mu_{\alpha+1} - (1-\alpha)^2\mu_1^2\mu_\alpha\right)\sigma_{\alpha;\alpha}\\
&\quad + 2\alpha^2\mu_1\mu_\alpha^2\sigma_{\alpha+1;1} - 2\alpha\mu_1^2\mu_\alpha\sigma_{\alpha+1;\alpha} - \alpha\mu_1^2\mu_\alpha\sigma_{2\alpha;1} + \mu_1^3\sigma_{2\alpha;\alpha}\\
&\quad - 2\alpha\mu_\alpha\left(\mu_1\mu_{2\alpha} - \alpha\mu_\alpha\mu_{\alpha+1} - (1-\alpha)^2\mu_1\mu_\alpha^2\right)\sigma_{1;1} - \alpha^3\mu_\alpha^3\sigma_{2;1}.
\end{aligned}$$

A.3.2 The Asymptotic Bias Term $k_{1;2}$

Taking expectations of the individual terms of (10) yields immediately $E(S_0) = n^{1/2}B^{-1/2}E\left\{n^{-1}\sum Y_{1;i}\right\} = 0$, and $E(S_{1/2}) = n^{-1/2}\left(B^{-1/2}E(Y_3Y_4) - 0.5B^{-3/2}E(Y_1Z_1)\right)$. It follows from the definition of $k_{1;2}$ that

$$E\{S\} = n^{-1/2}k_{1;2} + O\left(n^{-1}\right) \tag{16}$$

with

$$k_{1;2} = B^{-1/2}E(Y_3Y_4) - \frac{1}{2}B^{-3/2}E(Y_1Z_1). \tag{17}$$

[5] Note that we have used, for instance, the following second order expansion $m_1^{-\alpha} = \mu_1^{-\alpha} - \alpha\mu_1^{-\alpha-1}\left[n^{-1}\sum(X_i - \mu_1)\right] + \frac{\alpha(\alpha+1)}{2}\mu_1^{-\alpha-2}\left[n^{-1}\sum(X_i - \mu_1)\right]^2.$

A.3.3 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

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

by considering the constituent parts separately.

1. $E\left\{S_0^2 S_{1/2}\right\} = n^{3/2}B^{-3/2}E\{n^{-4}\sum_i\sum_j\sum_k\sum_l Y_{1;i}Y_{1;j}Y_{3;k}Y_{4;l}\} - 0.5n^{3/2}B^{-5/2}E\{n^{-4}\sum_i\sum_j\sum_k\sum_l Y_{1;i}Y_{1;j}Y_{3;k}Z_{1;l}\}$. Since we are only interested in the $O\left(n^{-1/2}\right)$ term, we conclude that

$$E\left\{S_0^2 S_{1/2}\right\} = n^{-1/2}B^{-3/2}\left[E(Y_1Y_1)E(Y_3Y_4) + 2E(Y_1Y_3)E(Y_1Y_4)\right] - n^{-1/2}\frac{3}{2}E(Y_1Z_1)$$

2. Consider $S_0^3 = n^{3/2}B^{-3/2}n^{-3}\left(\sum Y_{1;i}\right)^3$. Hence $E(S_0^3) = n^{-1/2}B^{-3/2}E(Y_1^3) + O(n^{-1})$:

In summary

$$\begin{aligned} E(S^3) &= n^{-1/2}B^{-3/2}\left(E(Y_1^3)\right. \\ &\quad \left. + 3\left[E(Y_1Y_1)E(Y_3Y_4) + 2E(Y_1Y_3)E(Y_1Y_4) - \frac{3}{2}E(Y_1Z_1)\right]\right) \\ &\quad + O\left(n^{-1}\right): \end{aligned} \tag{18}$$

Finally, since $K_{3;n} = E(S^3) - 3E(S^2)E(S) + 2(E(S))^3$, and $E(S^2) = 1 + O(n^{-1})$, using (16) and (18) we conclude that

$$k_{3;1} = B^{-3/2}\left[E(Y_1^3) + 6E(Y_1Y_3)E(Y_1Y_4) - 3E(Y_1Z_1)\right] \tag{19}$$

where $E(Y_1^3)$, $E(Y_1Y_3)$, and $E(Y_1Y_4)$ are defined in (15).

A.3.4 The Normalising Transform Derived

Equation (19) makes clear that the asymptotic skewness term depends on the parameters of the income distribution and the sensitivity parameter α of the inequality index. We will therefore only be able to derive explicit solutions to the differential equation (6) for particular income distributions and αs. We consider explicitly below the case of GE_2, and several skewed distributions. As it turns out, the approximate solution to the differential equation is the same across all income distributions considered. Other values of α can be dealt with similarly.

The Chi-squared Distribution We have included this distribution in order to examine how the skewness of the data distribution is inherited by the distribution of the studentized inequality measure. Let income be distributed as a chi-squared random variable with k degrees of freedom: $Y \sim \chi^2_k$ (we also note that χ^2_k is a particular case of the Gamma distribution, $G(k/2; 1/2)$). Its moments are given by $\mu_\alpha = 2^\alpha \Gamma(k/2 + \alpha)/\Gamma(k/2)$, implying the following quantities of interest

$$
\begin{aligned}
\iota &= k^{-1} \\
\sigma &= \frac{1}{2}\frac{1}{k^3}B^{1/2} \\
B &= 8k^4 + 16k^3 \\
E(Y_3Y_4) &= -k(2k+4) \\
E(Y_1Y_3) &= 0 \\
E(Y_1^3) &= 64k^6 + 640k^5 + 1024k^4 \\
E(Y_1Z_1) &= 64k^6 + 640k^5 + 1024k^4;
\end{aligned}
$$

which in turn imply the key quantities

$$
\begin{aligned}
k_{1;2} &= -\frac{3(k+6)}{\sqrt{2k(k+2)}} \\
k_{3;1} &= -4(k+8)\left(\frac{2}{k(k+2)}\right)^{1/2}.
\end{aligned}
\tag{20}
$$

It is of interest to note that $k_{3;1} \to -4\sqrt{2}$ as $k \to \infty$.

The differential equation thus becomes, using $k = \iota^{-1}$, and (20),

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

One round of integration yields

$$
t'(\iota) = \iota^{-4/3}(1 + O(\iota));
$$

and a subsequent term-by-term integration yields

$$
t(\iota) = -3\iota^{-1/3} + O\left(\iota^{2/3}\right);
$$

as claimed in equation (7).

The Gamma Distribution Let income be distributed as a Gamma random variable Y with shape parameter r and second parameter λ. The moments are given by

$\mu_\alpha = \frac{\Gamma(r+\alpha)}{\lambda^\alpha \Gamma(r)}$. The quantities of interest are

$$
\begin{aligned}
I &= \frac{1}{2} r^{-1} && (21)\\
\sigma &= \frac{1}{2}\left(\frac{\lambda}{r}\right)^3 B^{1/2}\\
B &= \frac{2r^3}{\lambda^6}(r+1)\\
E(Y_3Y_4) &= -\frac{r}{\lambda^3}(r+1)\\
E(Y_1Y_3) &= 0\\
E(Y_1^3) &= \frac{8r^4}{\lambda^9}\left(4+5r+r^2\right)\\
E(Y_1Z_1) &= \frac{8r^4}{\lambda^9}\left(4+5r+r^2\right);
\end{aligned}
$$

which in turn imply the key quantities

$$
\begin{aligned}
k_{1,2} &= -\frac{3}{2}\sqrt{2}\frac{r+3}{((r+1)r)^{1/2}} && (22)\\
k_{3,1} &= -4(r+4)\left(\frac{2}{r(r+1)}\right)^{1/2}.
\end{aligned}
$$

Using $r = 0.5I^{-1}$ in (21) and (22) yields the same differential equation as in the case of the χ_k^2 distribution.

The Lognormal Distribution Let income have a lognormal distribution $Y \sim LN(\mu, \sigma_y)$. The moments are given by $\mu_\alpha = \exp\left\{\alpha\mu + 0.5\alpha^2\sigma_y^2\right\}$. The quantities of interest are

$$
\begin{aligned}
I &= \frac{1}{2}\left(e^{\sigma_y^2} - 1\right) && (23)\\
\sigma &= \frac{1}{2} e^{-3\mu - \frac{3}{2}\sigma_y^2} B^{1/2}\\
B &= e^{6\mu + 5\sigma_y^2}\left(4e^{\sigma_y^2} - 4e^{2\sigma_y^2} + e^{4\sigma_y^2} - 1\right)\\
E\{Y_1Y_3\} &= e^{4\mu + 3\sigma_y^2}\left(1 - 2e^{\sigma_y^2} + e^{2\sigma_y^2}\right)\\
E\{Y_3Y_4\} &= e^{3\mu + \frac{5}{2}\sigma_y^2}\left(-3e^{\sigma_y^2} + e^{2\sigma_y^2} + 2\right)\\
E\{Y_1Y_4\} &= e^{5\mu + \frac{9}{2}\sigma_y^2}\left(6e^{\sigma_y^2} - 5e^{2\sigma_y^2} + e^{4\sigma_y^2} + 2\right)\\
E\left\{Y_1^3\right\} &= e^{9\mu + \frac{15}{2}\sigma_y^2}(12e^{\sigma_y^2} - 12e^{2\sigma_y^2} - 8e^{3\sigma_y^2} + 3e^{4\sigma_y^2}\\
&\quad + 12e^{5\sigma_y^2} - 6e^{8\sigma_y^2} + e^{12\sigma_y^2} - 2);\\
E\{Y_1Z_1\} &= e^{9\mu + \frac{15}{2}\sigma_y^2}(16e^{\sigma_y^2} - 10e^{2\sigma_y^2} - 16e^{3\sigma_y^2} + 5e^{4\sigma_y^2}\\
&\quad + 16e^{5\sigma_y^2} - 2e^{6\sigma_y^2} - 6e^{8\sigma_y^2} + e^{12\sigma_y^2} - 4).
\end{aligned}
$$

The implied key quantities are

$$
\begin{aligned}
k_{1,2} &= \frac{1}{2}\left(4e^{\sigma_y^2} - 4e^{2\sigma_y^2} + e^{3\sigma_y^2} - 1\right)^{-3/2} \times \\
&\quad \left(6e^{\sigma_y^2} - 32e^{2\sigma_y^2} + 52e^{3\sigma_y^2} - 19e^{4\sigma_y^2} - 14e^{5\sigma_y^2} + 2e^{6\sigma_y^2} + 6e^{8\sigma_y^2} - e^{12\sigma_y^2}\right);
\end{aligned}
$$

and

$$
\begin{aligned}
k_{3,1} &= -2\left(4e^{\sigma_y^2} - 4e^{2\sigma_y^2} + e^{3\sigma_y^2} - 1\right)^{-3/2} \\
&\quad \left(1 - 12e^{\sigma_y^2} + 48e^{2\sigma_y^2} - 68e^{3\sigma_y^2} + 18e^{4\sigma_y^2} + 24e^{5\sigma_y^2} - 6e^{6\sigma_y^2} - 6e^{8\sigma_y^2} + e^{12\sigma_y^2}\right).
\end{aligned} \tag{24}
$$

Using $\sigma_y^2 = \ln(1 + 2I)$, (23) and (24), yields the familiar differential equation

$$
\frac{t''(I)}{t'(I)} = -\frac{4}{3} I^{-1} (1 + O(I));
$$

and hence again the approximate solution claimed by (7).

B Contributions of Corrections

In this appendix we report addtional results to give an indication of the individual contribution each correction makes for the inequality measures. In particular we report the results for applying the normalising transform only, the normalising transform with bias correction, and the normalising transform with both the bias and variance correction.

	$SM(100; 2.8; 1.7)$			$LN(1; 0.5^2)$			$G(3; 0.15)$		
sample size	100	250	500	100	250	500	100	250	500
GE_2	18.5	15.5	13.5	13.0	9.72	8.39	7.74	6.55	5.96
$Theil_1$	10.3	8.42	7.68	8.97	7.17	6.49	6.19	5.61	5.3
$Theil_0$	7.45	6.21	5.83	6.96	5.97	5.6	6.09	5.5	5.25

Table 6: Actual coverage failure in per cent of confidence interval based on transform only . Nominal level 5 per cent based on 100,000 replications.

Table 2 in the main text contains the results for confidence intervals based on the normalising transform with both bias and variance correction. The results show that the normalising transform, as well as both bias and variance corrections make significant improvements in bringing the coverage failure probabilities closer to the nominal value.

	SM(100; 2:8; 1:7)			LN(1; 0:5^2)			G(3; 0:15)		
sample size	100	250	500	100	250	500	100	250	500
GE_2	17.1	14.6	12.9	12.1	9.8	8.1	7.7	6.6	6.1
$Theil_1$	9.7	8.4	7.5	8.6	7.0	6.4	6.2	5.6	5.4
$Theil_0$	7.2	6.3	5.9	6.7	5.8	5.6	6.1	5.5	5.1

Table 7: Actual coverage failure in per cent of con¯dence intervals based on the normalising transform and bias correction only. Nominal level 5 per cent based on 100,000 replications.

Figure 1: