Confidence sets for inequality measures: Fieller-type methods [a]

Jean-Marie Dufour [b]
McGill University

Emmanuel Flachaire [c]
Aix-Marseille Université

Lynda Khalaf [d]
Carleton University

Abdallah Zalghout [e]
Carleton University

April 6, 2017

[a] This work was supported by the William Dow Chair of Political Economy (McGill University), the Bank of Canada (Research Fellowship), the Toulouse School of Economics (Pierre-de-Fermat Chair of excellence), the Universitad Carlos III de Madrid (Banco Santander de Madrid Chair of excellence), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, the Fonds de recherche sur la société et la culture (Québec), and by project ANR-16-CE41-0005 managed by the French National Research Agency (ANR).

[b] William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill Uniersity, Leacock Building, Room 414, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 6071; FAX: (1) 514 398 4800; email: jean-marie.dufour@mcgill.ca. Web page: http://www.jeanmariedufour.com

[c] Aix-Marseille Université, CNRS, EHESS, Centrale Marseille, AMSE. Mailing address: GREQAM-EHESS, Aix-Marseille Université, 2 rue de la charité, 13002, Marseille, France. Email: emmanuel.flachaire@univ-amu.fr.

[d] Economics Department, Carleton University, Centre interuniversitaire de recherche en économie quantitative (CIREQ), and Groupe de recherche en économie de l'énergie, de l'environnement et des ressources naturelles (GREEN), Université Laval. Email: Lynda_Khalaf@carleton.ca.

[e] Economics Department, Carleton University. abdullahzalghout@cmail.carleton.ca.

ABSTRACT

Asymptotic and bootstrap inference methods for inequality indices are for the most part unreliable due to the complex empirical features of the underlying distributions. In this paper, we introduce a Fieller-type method for the Theil Index and assess its finite-sample properties by a Monte Carlo simulation study. The fact that almost all inequality indices can be written as a ratio of functions of moments and that a Fieller-type method does not suffer from weak identification as the denominator approaches zero, makes it an appealing alternative to the available inference methods. Our simulation results exhibit several cases where a Fieller-type method improves coverage. This occurs in particular when the Data Generating Process (DGP) follows a finite mixture of distributions, which reflects irregularities arising from low observations (close to zero) as opposed to large (right-tail) observations. Designs that forgo the interconnected effects of both boundaries provide possibly misleading finite-sample evidence. This suggests a useful prescription for simulation studies in this literature.
KEYWORDS: Inequality measures, Fieller-type confidence set, Delta method, Singh-Maddala distribution, Gamma distribution, mixture.

Contents

List of Figures

1 Introduction

Asymptotic inference methods for inequality indices are for the most part unreliable due to the complex empirical features of the underlying distributions, particularly in the case of income. Typically, the presence of heavy tails invalidates standard parametric and nonparametric inference methods based on central limit theory (CLT), leading to spurious conclusions with samples of realistic size. This problem persists even with very large samples. Moreover, for some parameter values, the moments of widely used distributions in this literature, such as the Singh-Maddala and Pareto distributions, do not exist. Early references can be traced back to Maasoumi (1997) or Mills and Zandvakili (1997); for a survey, see Cowell and Flachaire (2015).

Bootstrap inference methods emerge as an appealing alternative, since observations can often be viewed as independent random draws from the population. The first study to use and recommend bootstrap methods for inequality indices is the one of Mills and Zandvakili (1997). Biewen (2002) studied the performance of standard bootstrap methods in the context of inequality measures assuming a lognormal distribution as the Data Generating Process (DGP). Although his results suggest that the bootstrap performs well in finite samples, the lognormal distribution he used does not capture the thick tails typically observed in empirical work (Davidson and Flachaire, 2007). Other simulation studies based on heavy-tailed distributions, such as the Singh-Maddala distribution, confirm that bootstrapping fails – often by far – to control coverage rates, despite the fact that they lead to higher-order refinements relative to asymptotic methods (Davidson and Flachaire, 2007; Cowell and Flachaire, 2007).

Non-standard inference methods have recently been suggested in an attempt to improve the quality of inference for inequality measures. Two notable approaches are permutation tests (Dufour, Flachaire and Khalaf; 2010) and semi-parametric methods (Davidson and Flachaire, 2007; Cowell and Flachaire, 2007). The permutational approach focuses on testing the equality of two indices, and the authors show that it performs very well when the two indices come from similar distributions. The semi-parametric bootstrap approach assumes a parametric distribution for the right tail and a nonparametric empirical distribution function (EDF) for the rest. This method leads to considerable refinement over their asymptotic and bootstrap counterparts, provided the probability of the tail (p) and the ordered statistics defining the upper tail (k) are well chosen, which is usually not an easy task. Thus except for very specific cases, accurate inference methods on inequality measures are not available.

In this paper, we introduce the Fieller method for the Theil Index, and we assess its finite-sample properties through a Monte Carlo simulation study. Fieller's method was originally introduced for inference on the ratio of two means of normal variates. It is based on inverting a t-test of a linear restriction associated with the ratio, and allows one to get exact confidence sets for this ratio. This holds promise relative to the standard Delta method especially when the the denominator of the ratio approaches zero, since the implicit linear reformulation addresses the underlying weak identification. Most inequality indices can be written as a ratio of functions of moments; so a Fieller-type method may plausibly lead to more reliable inference on these indices. However, given the non-linear dependence between the numerator and the denominator of the indices along with the typically positive support of the underlying distributions, the advantages from employing a Fieller-type method should not be taken for granted. This motivates the present work.

The method first introduced by Fieller (1940, 1954) was extended to independent samples of different sizes (Bennett, 1953), multivariate models (Bennett, 1959), general exponential regression models (Cox, 1967), general linear regression models (Zerbe, 1978; Dufour, 1997), and dynamic models with possibly persistent covariates (Bernard, Idoudi, Khalaf and Yélou, 2007; Bernard, Chu, Khalaf and Voia, 2015; Stock and Lazarus, 2016). Bolduc, Khalaf and Yélou (2010) used several variants of Fieller's approach to build simultaneous confidence sets for multiple (possibly weakly identified) ratios and they showed in a simulation study that a Fieller-type method outperforms the Delta method and controls level globally. Empirically, Fieller's approach has been routinely applied in medical research and to a lesser extent in economics (Srivastava, 1986; Willan and O'Brien, 1996; Johannesson, Jonsson and Karlsson, 1996; Laska, Meisner, and Siegel, 1997; Stock and Lasaruz, 2016).

Fieller-type confidence sets may be perceived as counter-intuitive, because they can produce unbounded regions including the whole real line.[1] This perhaps gives reason for their unpopularity in applied work relative to Delta method-based confidence sets (DCS), despite their solid theoretical foundation. However, the geometric interpretation of Fieller's method is quite intuitive (see von Luxburg and Franz, 2004). More to the point here, in the presence of identification problems, valid coverage requires possibly unbounded outcomes (Gleser and Hwang, 1987; Dufour, 1997), which is allowed by a Fieller-type solution as opposed to the Delta method.

Our simulation results provide evidence on the superiority of a Fieller-type method in terms of reducing size distortions in many useful cases. In particular, a Fieller-type method improves coverage over the Delta method when the distribution under the null allows for bunching of low observations (close to zero) in addition to a thick right tail. For such cases, the denominator of the Theil index is small relative to the numerator and inequality is high. Methodologically, our findings suggest that studies focusing only on the upper tail may misrepresent finite-sample

[1] See Scheffé (1970) for a modified version of Fieller's method that avoid the confidence set $\mathbb{R}$.

distortions with positive support distributions. In contrast, our design does allow us to assess further irregularities arising from low observations. As illustrated by Cowell and Victoria-Feser (1996), both boundaries may matter for general entropy class of indices, although not necessarily for the Theil index.

The paper is organized as follows. Section 2 presents the Fieller-type method for the Theil index. In section 3, Monte Carlo results are provided. Section 4 concludes.

2 Fieller-type inference for inequality measures

Most income inequality indices depend solely on the underlying distribution of income. Technically speaking, they can be typically written as a functional which maps the space of cumulative distribution functions (CDFs) of income to the positive real line.

2.1 General functional ratios

Denote by Y the random variable representing income, and by $F_Y(y)$ its CDF. The class of indices considered in this paper can be written as the ratio of functions of two moments, namely the mean μ and another moment $\nu = \mathbb{E}[\phi(Y)]$, where $\phi(\cdot)$ is a given function. In particular, for the Theil index $\phi(Y) = Y\log(Y)$. In general, most inequality indices can be written as

$$I = \psi(\mu;\,\nu) = \frac{\psi_1(\mu;\,\nu)}{\psi_2(\mu;\,\nu)}. \tag{1}$$

The index I can be estimated using sample moments:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i\,, \quad \hat{\nu} = \frac{1}{n}\sum_{i=1}^{n}\phi(Y_i)\,, \tag{2}$$

where $Y_1, \ldots, Y_n$ is a sample of observations on Y, and $\phi(Y_i)$ is a function that takes different forms for different inequality indices. If we assume that the estimator is asymptotically normal, then the asymptotic covariance matrix can be estimated by

$$V(\hat{I}) = \frac{1}{n}\begin{bmatrix}\frac{\partial\psi}{\partial\mu} & \frac{\partial\psi}{\partial\nu}\end{bmatrix}\begin{bmatrix}\hat{\sigma}_\mu^2 & \hat{\sigma}_{\mu\nu}\\ \hat{\sigma}_{\mu\nu} & \hat{\sigma}_\mu^2\end{bmatrix}\begin{bmatrix}\frac{\partial\psi}{\partial\mu}\\ \frac{\partial\psi}{\partial\nu}\end{bmatrix} \tag{3}$$

where $\hat{V}(\hat{I}) \equiv V(\hat{I})|_{\mu=\hat{\mu}}$. Here, $\hat{\sigma}_\mu^2$, $\hat{\sigma}_\mu^2$ and $\hat{\sigma}_{\mu\nu}$ are, respectively, estimates of the variance of Y, the variance of $\phi(Y)$, and the covariance of Y and $\phi(Y)$.[2]

[2]Note that the variance of $\hat{\mu}$, the variance of $\hat{\nu}$ and covariance of the $(\hat{\mu},\,\hat{\nu})$ are equal to $\hat{\sigma}_\mu^2/n$, $\hat{\sigma}_\nu^2/n$ and $\hat{\sigma}_{\mu\nu}/n$.

In this paper, we consider the problem of building Fieller-type confidence sets (FCS) and Delta method confidence sets (DCS) for an index of the form I in (1). In general, this can be viewed as equivalent to finding the values of I_0 which are not rejected when one tests null hypotheses of the form

$$H_0(I_0) : \frac{\psi_1(\mu,\, \nu)}{\psi_2(\mu,\, \nu)} = I_0 \tag{4}$$

where I_0 is any admissible value of I. Here, this can be achieved by inverting the absolute value or the square of the relevant t-type statistic. To invert a t-test with respect to the parameter tested, we collect all the values of this parameter for which the test is not significant at a given level.

Following the Delta method, we invert the test statistic

$$t^2(I_0) = \frac{(\hat{I} - I_0)^2}{\hat{V}(\hat{I})} \tag{5}$$

which leads to the confidence set

$$\text{DCS}(I; 1-\alpha) = \left[\hat{I} - z_{\alpha/2}[\hat{V}(\hat{I})]^{1/2}\,;\; \hat{I} + z_{\alpha/2}[\hat{V}(\hat{I})^{1/2}]\right] \tag{6}$$

where $z_{\alpha/2}$ is the usual α critical point based on the normal distribution (*i.e.*, $\mathbb{P}[Z \geq z_{\alpha/2}] = \alpha/2$ for $Z \sim N[0,\, 1]$).

By contrast, the Fieller approach can be applied as follows. For each possible value I_0, the Fieller-type approach consists in considering the equivalent linear hypothesis

$$H_L(I_0) : \theta(I_0) = 0\,, \quad \text{where } \theta(I_0) = \psi_1(\mu,\, \nu) - I_0\, \psi_2(\mu,\, \nu) \tag{7}$$

where the superscript L is added to differentiate the original null hypothesis from its linear reformulation. Thorough this exact linearization, the Fieller-type method avoids possible (weak) identification problems when the denominator $\psi_2(\mu, \nu)$ is close to zero. To construct the FCS, we consider the square of the t-statistic associated with $H_L(I_0)$ in (7):

$$t(I_0)^2 = \frac{\hat{\theta}(I_0)^2}{\hat{V}[\hat{\theta}(I_0)]} \tag{8}$$

where $\hat{V}[\hat{\theta}(I_0)]$ is an estimate of the variance of $\hat{\theta}(I_0)$. If the statistic follows asymptotically a standard normal distribution, then a confidence set with level $1-\alpha$ for the index I can be built by noting that

$$t(I_0)^2 \leq z^2_{\alpha/2} \quad \Leftrightarrow \quad \frac{\hat{\theta}(I_0)^2}{\hat{V}[\hat{\theta}(I_0)]} \leq z^2_{\alpha/2} \quad \Leftrightarrow \quad \hat{\theta}^2(I_0) - z^2_{\alpha/2}\, \hat{V}[\hat{\theta}(I_0)] \leq 0\,. \tag{9}$$

This yields the confidence set

$$\text{FCS}(I;1-\alpha)=\left\{I_0:\hat{\theta}^2(I_0)-z^2_{\alpha/2}\,\hat{V}[\hat{\theta}(I_0)]\leq 0\right\}. \tag{10}$$

Since $\hat{\theta}(I_0)$ is linear in I_0, $\theta(I_0)^2$ and $\hat{V}[\hat{\theta}(I_0)]$ are quadratic functions of I_0:

$$\hat{\theta}^2(I_0)=A_1I_0^2+B_1I_0+C_1\,,\quad V[\hat{\theta}(I_0)]=A_2I_0^2+B_2I_0+C_2\,, \tag{11}$$

where the coefficients (defined below) depend on the data and the Gaussian critical point. On substituting (11) into (10), we get the quadratic inequality

$$AI_0^2+BI_0+C\leq 0 \tag{12}$$

where

$$A=A_1-z^2_{\alpha/2}A_2\,,\quad B=B_1-z^2_{\alpha/2}B_2\,,\quad C=C_1-z^2_{\alpha/2}C_2\,. \tag{13}$$

The coefficients, A_1, B_1, C_1, A_2, B_2, C_2 are functions of the sample moments and their variance estimates. The FCS solve the second degree polynomial inequality in (12) for I_0. Let $\Delta=B^2-4AC$, then the $(1-\alpha)$-level Fieller-type confidence set is characterized as follows:

1. if $\Delta>0$ and $A>0$, then $\text{FC}(I;1-\alpha)=\left[\frac{-B-\sqrt{\Delta}}{2A},\,\frac{-B+\sqrt{\Delta}}{2A}\right]$,
2. if $\Delta>0$ and $A<0$, then $\text{FC}(I;1-\alpha)=\left]-\infty,\,\frac{-B+\sqrt{\Delta}}{2A}\right]\cup\left[\frac{-B-\sqrt{\Delta}}{2A},\,+\infty\right[$,
3. if $\Delta<0$, then $A<0$ and $\text{FC}(I;1-\alpha)=\mathbb{R}$.

For more details, see Bolduc, Khalaf and Yélou (2010) and the references therein.

2.2 Fieller-type inference for the Theil Index

The Theil index belongs to the family of GE indices and can be written as a function of two moments $\mu=\mathbb{E}(Y)$ and $\nu=\mathbb{E}[Y\log(Y)]$, where μ and ν can be estimated using their sample counterparts. In this paper, we will use the following expression for the Theil index:

$$I_T=\frac{\nu}{\mu}-\log(\mu). \tag{14}$$

For the Theil index (I_T), the null hypothesis defined in (4) can be written as:

$$H_0(I_{T0}):\frac{\nu}{\mu}-\log(\mu)=I_{T0}\,. \tag{15}$$

The variance of the estimated Theil index can be derived using the Delta method and it is defined by (3) where the expressions of the derivatives in this context are:

$$\frac{\partial \psi}{\partial \mu} = -\frac{(\nu + \mu)}{\mu^2}\,, \quad \frac{\partial \psi}{\partial \nu} = \frac{1}{\mu}\,. \tag{16}$$

The Fieller-type method for the Theil index starts by considering the equivalent linear hypothesis as shown in (7):

$$H_0(I_{T0}) : \nu - \mu\, \log(\mu) - \mu\, I_{T0} = 0\,, \tag{17}$$

along with the corresponding t-statistics (squared). The confidence set for I_T is then obtained by solving the quadratic inequality described by (10) - (12). For this, we derive the parameters $A1$, $B1$, $C1$, $A2$, $B2$ and $C2$ in equation (11) for the Theil index:

$$A_1 = \hat{\mu}^2\,, \quad B_1 = -2\hat{\mu}\,[\hat{\nu} - \hat{\mu}\log(\hat{\mu})]\,, \quad C_1 = [\hat{\nu} - \hat{\mu}\log(\hat{\mu})]^2\,. \tag{18}$$

To get the variance of $\hat{\theta}(I_0)$, we apply the Delta method to $\theta(I_0)$ in (7):

$$A_2 = \hat{\sigma}^2_{\mu}/n\,, \quad B_2 = \big(2\hat{\sigma}^2_{\mu}\,[\log(\hat{\mu}) + 1] - 2\hat{\sigma}_{\mu\nu}\big)/n\,, \tag{19}$$

$$C_2 = \big(\hat{\sigma}^2_{\mu}[\log(\hat{\mu}) + 1]^2 - 2\hat{\sigma}_{\mu\nu}[\log(\hat{\mu}) + 1] + \hat{\sigma}^2_{\mu}\big)/n\,, \tag{20}$$

where $\hat{\sigma}^2_{\mu}$, $\hat{\sigma}^2_{\mu}$ and $\hat{\sigma}_{\mu\nu}$ are defined in (3).

3 Simulation results

In this section, we provide Monte Carlo evidence on the finite-sample properties of the Fieller-type method for the Theil index. We conduct several simulation studies focusing on the behaviour of the Fieller method when the hypothesized income distribution under the null is characterized by thick tails. To this end, we simulate data sets from the Singh-Maddala distribution [$Y_i \sim \mathrm{SM}(a, b, q)$], the Gamma distribution [$\mathrm{Gamma}(k, \theta)$], and finite mixtures of the latter. These distributions have been used in the literature in the context of income inequality measures (Brachman, Stich and Trede, 1996; McDonald, 1984; Kleiber and Kotz, 2003; Cowell and Victoria-Feser, 1996).

The CDF of the Singh-Maddala distribution can be written as

$$F(y) = 1 - \left[1 + \left(\frac{y}{b}\right)^a\right]^{-q}, \tag{21}$$

where a is a shape parameter which affects both tails, q is another shape parameter which affects only the right tail, and b is a scale parameter which has no impact on our analysis (for the indices in question are scale invariant. For this distribution, the k-th moment exists for $-a < k < aq$.

The expectation of the Singh-Maddala distribution can be expressed as

$$\mu = \frac{q\, b\, \Gamma\left(a^{-1}+1\right)\Gamma\left(q-a^{-1}\right)}{\Gamma\left(q+1\right)}\,. \tag{22}$$

In this case, a closed-form expression for $\nu = \mathbb{E}[Y \log(Y)]$ is also available:

$$\nu = \mu\, a^{-1}[\psi(a^{-1}+1) - \psi(q-a^{-1}) + a\log(b)] \tag{23}$$

where $\Gamma(\cdot)$ is the Gamma function and $\psi(\cdot) \equiv \Gamma'(\cdot)/\Gamma(\cdot)$ is the digamma function.

The other distribution we consider in the simulations is the Gamma distribution with density function

$$f(y) = \frac{y^{k-1}e^{-(k/\theta)}}{\theta^k \Gamma(k)}\,,\ y > 0\,, \tag{24}$$

where k is a shape parameter and θ is a scale parameter. The expectation of this distribution (μ) is the scale multiplied by the shape parameter ($\mu = k\theta$). The value of ν for the Gamma distribution was computed by numerical methods.

The number of replications was set to $N = 10000$. For each sample, we compute the Theil inequality, and the underlying estimated variance, and the t-type statistics associated with the Delta and Fieller-type methods. Because of the duality between tests and confidence sets, the coverage rate of the confidence sets can be evaluated by computing the rejection probabilities of these tests. The coverage error rate (or equivalently the rejection probability) is computed as the proportion of times the relevant t-statistic rejects the null hypothesis. For a significance level α, we say the test approaches the nominal level when the rejection rate approaches α.

The main results of the simulation experiments are presented in the form of plots where the numbers of observations are on the x-axis and the coverage error rates on the y-axis. The 5% nominal level is maintained for all tests. The horizontal solid lines in the graphs represent the nominal level 0.05.

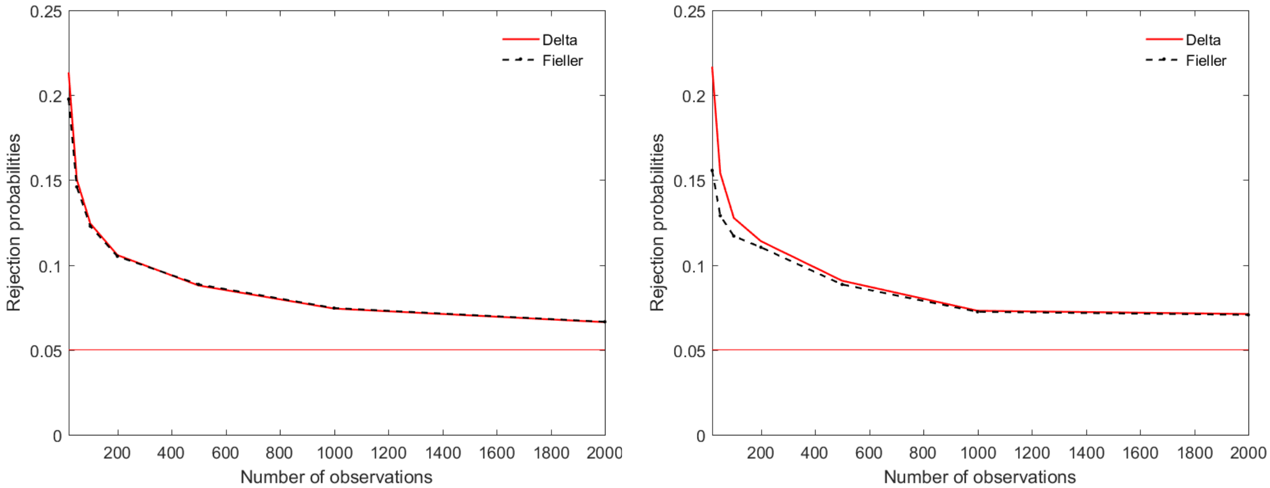
Our simulation results show that the Fieller-type method has better coverage than the Delta method in several cases, especially when the underlying distribution involves heavy lower and upper tails.

Figure 1 plots the rejection probabilities of the Fieller-type and Delta methods under Singh-Maddala distributions. In the left panel, the distribution is Singh-Maddala with parameters $a = 2.8$ and $q = 1.7$. The Fieller-type and Delta methods have similar coverage. However, the other designs considered reveal important improvement with the Fieller method. In the right panel, the distribution is Singh-Maddala with parameters $a = 1.1$ and $q = 5$. For this choice of parameters, the distribution exhibits bunching of low observations. In this context, the Fieller-type method outperforms the Delta method for relatively small samples up to 400 observations.

Figure 1: Singh-Maddala distributions

Rejection probabilities for Delta and Fieller methods

Left panel: $SM(a = 2.8,\ q = 1.7)$; Right panel: $SM(a = 1.1,\ q = 5)$

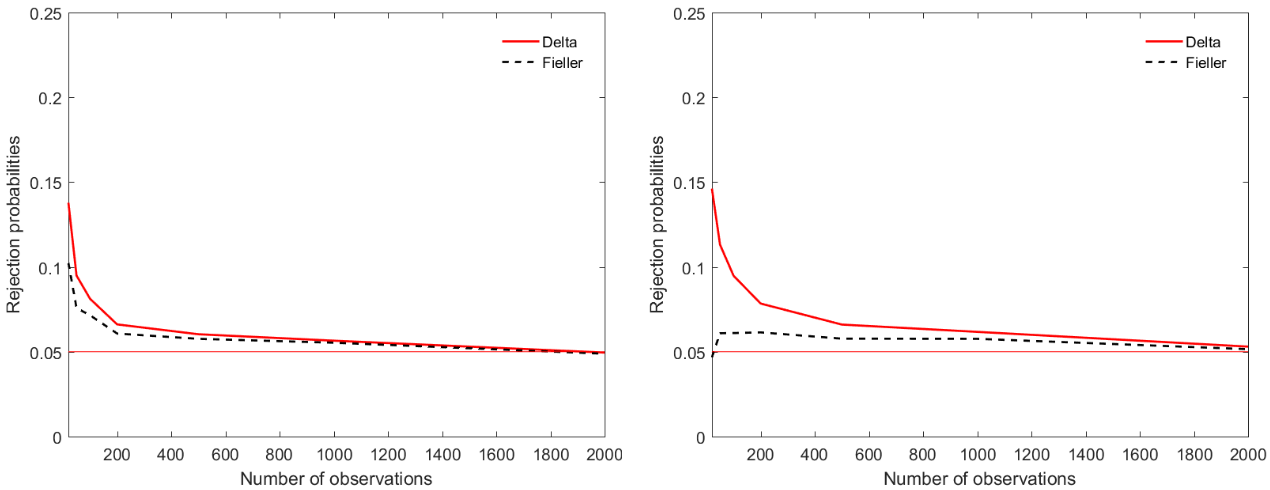


Note - The Delta method and Fieller method statistics are defined by (5) and (8) respectively. The null hypothesis tested is $H(I_{T0}) : I = I_0$ where I_0 is computed analytically.

Figure 2: Gamma distributions

Rejection probabilities for Delta and Fieller methods

Left panel: $\text{Gamma}(k = 1,\ \theta = 1)$; Right panel: $\text{Gamma}(k = 0.3,\ \theta = 1)$



Note - The Delta method and Fieller method statistics are defined by (5) and (8) respectively. The null hypothesis tested is $H(I_{T0}) : I = I_0$ where I_0 is computed analytically.

The same conclusion can be drawn from the case where we assume a Gamma distribution under the null. The size improvements we find with the Fieller-type method increase as the left tail of the distribution gets thicker. In the left panel of Figure 2, we plot the rejection probabilities under both methods under a $\mathrm{Gamma}(k = 1,\ \theta = 1)$ distribution. The differences in the rejection probabilities for samples of size 20 is around 4%, and around 2% with 100 observations. As we increase the proportion of low observations, the Fieller-type method provides remarkable size improvements, and in some cases it approaches the 5% nominal level for sample sizes as small as 200. In the right panel of this figure, the Fieller-type method coverage error is less than that of Delta method by around 9% going down from almost 15% to 6%.

The shape of the distributions underlying the aforementioned results represent populations where most of the individuals are poor and few are rich. The choice of these distributions was made to study the performance of the two methods when tails are fat both near the zero boundary and to the right. As we will discuss shortly our findings conform with the theoretical work of Cowell and Victoria-Feser (1996).

In Figures 3 and 4, we consider mixed designs with bimodal distributions under the null. These distributions, their parameters and the associated mixture weights are chosen to capture represent tail thickness at both ends of the distribution. Since the analytical expression for the Theil under the null of mixtures does not exist, we used an estimate of the Theil index based on a very large sample (n=1000000000). This approach of computing the true Theil index under the null is justified by the consistency of the Theil Index.

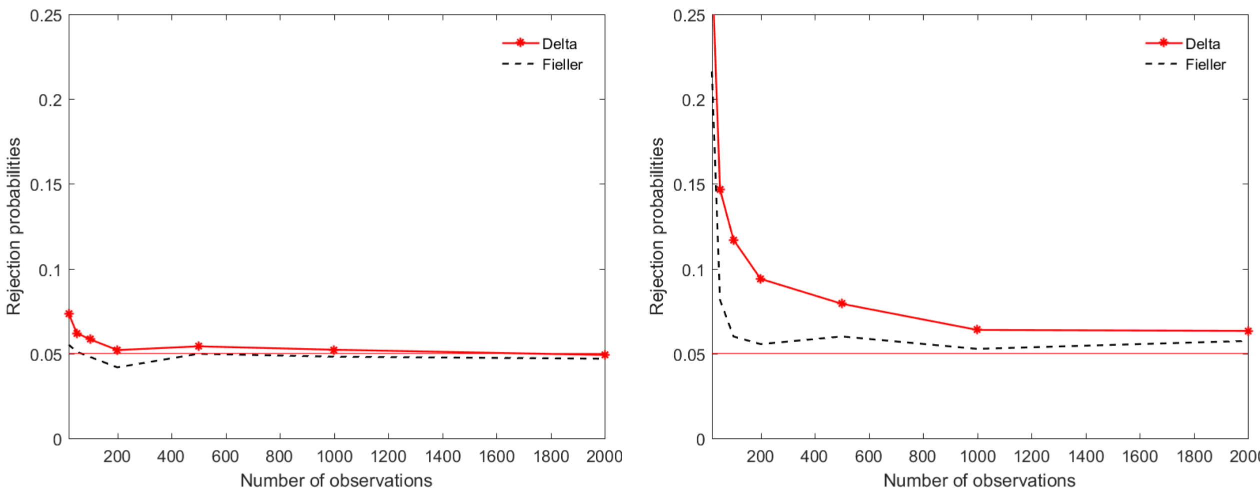
Figure 3 plots the rejection probabilities for mixtures of two Singh-Maddala distributions. In the left panel, the mixture combines a $\mathrm{SM}(1.1,\ 5)$ with probability 0.7 weight and $\mathrm{SM}(2.8, 1.7)$ with probability 0.3: *i.e.* on average, we draw 70% of the sample from a distribution with a peak near low incomes, and 30% from a distribution characterized by a thick right tail. For this design, the Fieller-type method approaches the nominal significance level for samples as small as 50 observations.

In the right panel, we increase the weight for the first distribution from 0.7 to 0.9. Thus we are giving more weight to the distribution with irregularities on the left tail rather than the one characterized with right tail thickness. The Fieller-type method dominates the Delta method by wide margins. This confirms our previous conclusion that the Fieller-type method is superior to the Delta method, especially when the distribution exhibits bunching of low observations.

Further evidence appears in Figure 4 where we consider mixtures of $\mathrm{Gamma}(0.3,\ 1)$ and $\mathrm{SM}(2.8,\ 1.7)$ distributions. Again the left panel gives to the first distribution (the Gamma distribution) a weight of 0.7, while the right one increases this weight to 0.9. Again, the Fieller-type method improves coverage, especially for small samples.

Figure 3: Mixtures of Singh-Maddala distributions

Rejection probabilities for Delta and Fieller methods
Left panel: $0.7{*}SM(1.1,\ 5) + 0.3{*}SM(2.8,\ 1.7)$
Right panel: $0.9{*}SM(1.1,\ 5) + 0.1{*}SM(2.8,\ 1.7)$



Note - The Delta method and Fieller method statistics are defined by (5) and (8) respectively. The null hypothesis tested is $H(I_{T0}) : I = I_0$ where I_0 is calibrated via a separate simulation.

The designs we considered can be interpreted through the work of Cowell and Victoria-Feser (1996). This paper views the underlying distribution as a mixture of a finite number of other distributions where bunching and tail behaviour can be formally modelled. Results for cases of "extreme" behaviour are pointed out, including the zero and infinite boundaries. In particular, the Theil index can have an unbounded influence function when some of the data approaches ∞, although the zero boundary may matter for other inequality indices.

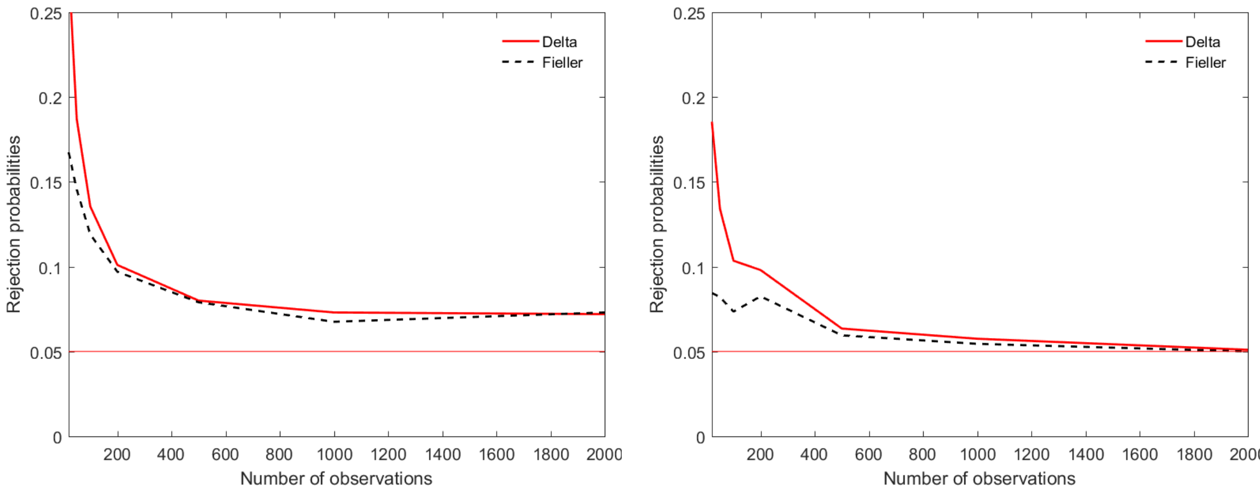
The influence function measures the change in the estimator for a small perturbation of the data. It is related to the bias of the estimator in the sense that when the IF is unbounded the bias can be infinite. Cowell and Victoria-Feser (1996) show that any decomposable scale invariant index for which the mean is estimated from the sample has an unbounded IF. We find that the Fieller-type method improves coverage, especially when small (near zero) or large observations are highly probable.

Figure 4: Mixtures of Gamma and Singh-Maddala distributions

Rejection probabilities for Delta and Fieller methods

Left panel: $0.7 * \text{Gamma}(0.3, 1) + 0.3{*}SM(2.8, 1.7)$

Right panel: $0.9 * \text{Gamma}(0.3, 1) + 0.1{*}SM(2.8, 1.7)$



Note - The Delta method and Fieller method statistics are defined by (5) and (8) respectively. The null hypothesis tested is $H(I_{T0}) : I = I_0$ where I_0 is calibrated via a separate simulation.

4 Conclusion

This paper proposes Fieller-type procedures for inference on the Theil inequality index and illustrates its superiority relative to its standard Delta method counterpart, using various empirically relevant simulation designs. Our results confirm that, in contrast with the Delta method, the proposed procedures can capture some of the distributional irregularities arising from the concentration of low observations and the thickness of the right tail. More broadly, our findings suggest that the Fieller-type approach holds concrete promise for many other inequality measures, as well as for inference on differences between measures.

References

[1] Bennett, B. (1962), 'On combining estimates of relative potency in bioassay', *Journal of Hygiene* **60,** 379–385.

[2] Bennett, B. (1959), 'On a multivariate version of Fieller's theorem', *Journal of the Royal Statistical Society* Series B **21** (1)**,** 59–62.

[3] Bernard, J. T., Idoudi, N., Khalaf, L. & Yélou, C. (2007), 'Finite sample inference methods for dynamic energy demand models', *Journal of Applied Econometrics* **22(7)**, 1211–1226.

[4] Bernard, B., Chu, B., Khalaf, L. & Voia, M. (2015), 'Non-standard Confidence Sets for Ratios and Tipping Points with Applications to Dynamic Panel Data', Paper presented at the International Panel Data Conference, Budapest, Hungary.

[5] Biewen, M. (2002), 'Bootstrap inference for inequality, mobility and poverty measurement', *Journal of Econometrics* **108**, 317–342.

[6] Bolduc, D., Khalaf, L. & Yélou, C. (2010), 'Identification robust confidence set methods for inference on parameter ratios with application to discrete choice models', *Journal of Econometrics* **157**, 317–327.

[7] Brachmann, K., Stich, A. & Trede, M. (1996), 'Evaluating parametric income distribution models', *Allgemeines Statistisches Archiv* **80**, 285–298.

[8] Cowell, F. & Flachaire, E. (2007), 'Income distribution and inequality measurement: The problem of extreme values', *Journal of Econometrics* **141**, 1044–1072.

[9] Cowell, F. & Flachaire, E. (2015), 'Statistical methods for distributional analysis', *Handbook of Income Distribution* **2**, 359–465.

[10] Cowell, F. A. & Victoria-Feser, M. P. (1996), 'Robustness properties of inequality measures', *Econometrica* **64** (1), 77–101.

[11] Cox, D. (1967), 'Fieller's theorem and a generalization', *Biometrika* **54**, 567–572.

[12] Davidson, R. & Flachaire, E. (2007), 'Asymptotic and bootstrap inference for inequality and poverty measures', *Journal of Econometrics* **141**, 141–166.

[13] Dufour, J.-M. (1997), 'Some impossibility theorems in econometrics with applications to structural and dynamic models', *Econometrica* **65**, 1365–1387.

[14] Dufour, J.-M., Flachaire, E.. & Khalaf, L. (2015), 'Permutation tests for comparing inequality measures', Department of Economics, McGill University.

[15] Fieller E. C. (1940), 'The Biological Standardization of Insulin', *Journal of the Royal Statistical Society (Supplement)* **7**, 1–64.

[16] Fieller E. C. (1954), 'Some Problems in Interval Estimation', *Journal of the Royal Statistical Society B* **16**, 175–185.

[17] Gleser, L. J. & Hwang, J. T. (1987), 'The nonexistence of $100(1-\alpha)$ confidence sets of finite expected diameter in errors-in-variables and related models', *The Annals of Statistics* **15**, 1351–1362.

[18] Johannesson, M., Jönsson, B. & Karlsson, G. (1996), 'Outcome measurement in economic evaluation', *Health Economics* **5**, 279–296.

[19] Kleiber, C. & Kotz, S. (2003), 'Statistical Size Distributions in Economics and Actuarial Sciences', *Hoboken, N.J.*: John Wiley.

[20] Laska, E., Meisner, M. & Siegel, C. (1997), 'Statistical inference for cost effectiveness ratios', *Health Economics* **6**, 229–242.

[21] Maasoumi, E. (1997), 'Empirical analyses of inequality and welfare', *Handbook of applied econometrics: Microeconomics* **2**, 202–245.

[22] McDonald, J. B. (1984), 'Some generalized functions for the size distribution of income', *Econometrica* **52**, 647–664.

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

[24] Scheffé, H. (1970), 'Multiple testing versus multiple estimation. Improper confidence sets. Estimation of directions and ratios', *The Annals of Mathematical Statistics* **41** (1), 1-29.

[25] Stock, J., & Lazarus. E. (2016), 'Identification of Factor-Augmenting Technical Growth and the Decline of the Labor Share', Paper presented at the CIREQ Econometrics Conference in Honor of Jean-Marie Dufour, Montréal, Canada.

[26] Srivastava, M. S. (1986), 'Multivariate bioassay, combination of bioassays, and Fieller's theorem', *Biometrics* **42**, 131–141.

[27] von Luxburg, U. & Franz, V. (2004), ‘Confidence sets for ratios: A purely geometric approach to Fieller’s theorem’, Tech. Rep. TR-133, Max Planck Institute for Biological Cybernetics, Germany.

[28] Willan, A. & O’Brien, B. (1996), ‘Confidence intervals for cost effectiveness ratios: An application of Fieller’s theorem’, *Health Economics* **5**, 297–305.

[29] Zerbe, G. (1978), ‘On Fieller’s theorem and the general linear model’, *The American Statistician* **32** 103–105.

[30] Zerbe, G. O., Laska, E., Meisner, M. & Kushner, H. B. (1982). ‘On multivariate confidence regions and simultaneous confidence limits for ratios’, *Communications in Statistics - Theory and Methods* **11**, 2401–2425