GRÉDI

Groupe de Recherche en Économie et Développement International

Cahier de recherche / Working Paper
06-13

Measuring Significance of Inequalities with Heterogeneous Groups and Income Sources

Stéphane Mussard

María Noel Pi Alperin

UNIVERSITÉ DE SHERBROOKE

Measuring Significance of Inequalities with Heterogeneous Groups and Income Sources*

Stéphane Mussard† Université de Perpignan

María Noel Pi Alperin‡ Université de Montpellier I

March 2006

Abstract

The aim of this paper is threefold. First, we propose a bidimensional decomposition of the Gini ratio that combines two decomposition techniques of inequality measurement in a bivariate distribution context. Usually decomposed measures of inequality are gauged to be significant or not, using non-stratified bootstrap techniques. Second, we show, with an illustrative example on the Italian survey on household's income and wealth, that the stratified bootstrap yields less distortion in estimated parameters. Finally, a brief review of the well-known confidence intervals follows.

Keywords: Inequalities, Complex surveys, Multi-decomposition, Stratified bootstrap.

JEL Codes: C14, D31.

*We acknowledge Camilo Dagum for all the things he taught us. The usual disclaimer applies.

†Corresponding auhor: GRÉDI and GEREM, Département des Sciences Economiques, Université de Perpignan, 52 Avenue Paul Alduy, 66860 Perpignan Cedex, France, E-mail: smussard@adm.usherbrooke.ca

‡LAMETA, Université de Montpellier I, UFR des Sciences Economiques, E-mail: pialperin@lameta.univ-montp1.fr

1 Introduction

Mills and Zandvakili (1997) proposed to bootstrap the Gini coefficient. Different reasons incite researchers to develop variance estimator or bootstrap methods to gauge the significance of inequality indices. When inequality measures are normalized and belong to the close interval $[0, 1)$, such as the Gini ratio, changes in inequality over time concern very small values. Then, wondering about the significance of the index variation is of interest.

Another reason that stimulates researchers with inference is concerned with inequality decompositions. Two decomposition techniques are itemized. On the one hand, there is the subgroup decomposition that enables one to share the overall inequality into within-group and between-group inequalities. On the other hand, there is the source (factor) decomposition: when incomes are desegregated into various income sources such as labour income, child support benefits, and social taxes, the source decomposition yields the contribution of each factor to the global inequality. Whatever the chosen decomposition technique, confidence intervals are needed since the contributions (within-group, between-group or factor components) represent very small quantities.

Biewen (2002) investigates bootstrap inference for generalized entropy inequality measures and their decompositions, with the distinction of standard *i.i.d.* distributions and complex survey cases. Biewen and Jenkins (2006) pursue this idea and develop variance estimators for the generalized entropy and Atkinson indices. But derivation of variance estimators can be challenging when the studied measure is an overcomplicated function such as the Gini ratio. Bootstrap inference is then appropriated in those cases.

The aim of this article is to suggest a bootstrap procedure to deal with a particular technique of decomposition. Mussard (2004, 2006) proposed to work within an unified decomposition framework. Indeed, the Gini coefficient is liable to a bidimensional law of decomposition. Traditionally, inequality measures such as entropy yield "marginal" decompositions. Following Table 1, these marginal decompositions only address economic conclusions based on the margins. For instance, when the population is divided in two groups, the subgroup decomposition enables one to state that the first group contributes with a 70% to the overall inequality, the second one with a 20%, and that the between-group inequality represents 10% of the global inequality. On the other hand, when incomes are fractionated into three sources, the factor decomposition ascribes a part of the overall inequality to each factor: 40% to labour income, 30% to social benefits, and 30% to social taxes.

Table 1: Structure of Inequality Decomposition Possibilities

Sources → Groups ↓	Labour Income	Social Benefits	Social Taxes	Total
Inequalities within group 1	×	×	×	70%
Inequalities within group 2	×	×	×	20%
Inequalities between groups 1 and 2	×	×	×	10%
Total	40%	30%	30%	100%

The bidimensional decomposition of the Gini ratio (Gini multi-decomposition) allows one to combine the two marginal techniques and to determine the missing values " × ". Bootstrapping each couple or combination " × " (implying the bootstrap of the margins) appears as a crucial feature when the number of groups and sources is important. Nevertheless, following Athanasopoulos and Vahid (2003), bootstrapping the Gini index (without decomposition) is a complex problem. They advance critical insights to Mills and Zandvakili (1997). Athanasopoulos and Vahid (2003) underline the use of the weighted Gini index in order to formulate a correction of the probabilities because stratified sample data implicitly suppose that individuals are not drawn randomly from the population.

In the light of this discussion, we try to answer to the following interrogations. Is it possible to formulate a weighted Gini multi-decomposition? And, which bootstrap technique can be appropriated for this weighted Gini multi-decomposition?

2 The bidimensional decomposition of the weighted Gini index

Do the inequalities have to be reduced? Given libertarian theories, State interventions are not welcomed since they cause perturbations in market mechanisms, which normally lead to equilibrium and optimality. On the contrary, Public Economics can recommend State interventions to perform some policies of inequality reduction. In this way, Sen's (1973) interrogation "Inequality of what?" is fundamental. In our study, combining source and subgroup decompositions induces us to focus essentially on incomes and their various income components or consumption expenditures.

Another question is central. How measuring inequalities? Two regular inequality measures are principally used: the Gini index and the generalized entropy measures.

Both are symmetric (in order to guarantee individual's anonymity), homogeneous of degree zero (distribution units can be changed without changing the index value), population invariant (distributions with heterogeneous sizes can be measured and compared), sensitive to Pigou-Dalton transfers (the index deceases after income transfers from a higher-income individual to a lower-income one), and normalized (taking the zero value when distributions are egalitarian).

Cowell (1989) derives variance estimators for univariate subgroup decomposition of the generalized entropy indices, using the method of moments. In the same paper, using U-statistics, he finds the variance estimator of the Gini index.

Biewen (2002) demonstrates that inequality, poverty and mobility measures can be bootstrapped. He also demonstrates how it is possible to bootstrap the entropy measures when univariate subgroup decomposition, univariate source decomposition, and time variation are taken into account.

Biewen and Jenkins (2006) determine the variance estimator of entropy and Atkinson indices, which depend on aversion inequality parameters. Their study focuses essentially on complex survey data. In these cases, the population is divided into many strata that involve many clusters, which again involve several individuals.

Bootstrap methods especially concern the overall Gini index. Xu (2000) applies the iterated-bootstrap method on the generalized Gini indices that depend on an aversion parameter of inequality. Athanasopoulous and Vahid (2003) use Biewen's (2002) method, which consists in drawing B bootstrap samples ($b = 1, \dots, B$) from a sample empirical distribution $\hat{F}$ with a bivariate nature, that is, a distribution of couples (income, weight):

$$\hat{F} \longrightarrow \left[(y_1^{*b}, w_1^{*b}), (y_2^{*b}, w_2^{*b}), \dots, (y_i^{*b}, w_i^{*b}), \dots, (y_n^{*b}, w_n^{*b})\right], \tag{1}$$

where y_i and w_i stands respectively for the income of the ith household and its corresponding weight. Hence, computing measures of inequality on each b sample helps to derive confidence intervals.

In this way, let us now expose the specification of the weighted Gini multi-decomposition. Formally, the well-known Gini coefficient is:

$$G = \sum_{i=1}^{n} \sum_{r=1}^{n} \frac{|y_i - y_r|}{2n^2 \overline{y}}, \tag{2}$$

where $\overline{y}$ is the arithmetic mean of incomes. As explained *supra*, we have the possibility to weight the Gini index to make a correction of the probability for which the household is supposed to be randomly drawn from the overall population, whereas a percentage of census data has been used. Then, the weighted Gini coefficient can be formulated as follows:

$$\mathcal{G} = \sum_{i=1}^{n}\sum_{r=1}^{n} \frac{|y_i - y_r| w_i w_r}{2\overline{y}_w}, \tag{3}$$

where $\overline{y}_w$ is the weighted arithmetic mean of incomes and where w_i and w_r are respectively the weights attached to the ith and the rth observations. Indeed, if stratified samples are used, which include representative sub-samples, then to each ith sample observation corresponds a weight equal to the relative frequency of the ith sample observation in the total population.[1]

Let P be a population of n individuals of mean $\overline{y}$ divided into k subgroups P_j $(j, h = 1, \ldots, k)$. The mean and the size of each subpopulation are respectively $\overline{y}_j$ and n_j. Mussard (2004, 2006) demonstrates that the Gini index G (Eq.2) is multi-decomposable, so that, the elements of Table 1 are determinable. Mussard and Terraza (2006) generalize this result and prove that it exists a class multi-decomposable Gini indices:[2]

$$g(y) = \sum_{i=1}^{n}\sum_{r=1}^{n} \frac{|y_i - y_r|}{f(y_i, y_r)}, \text{ for all } f(y_i, y_r) \neq 0. \tag{4}$$

Proposition 2.1 *The weighted Gini index is multi-decomposable.*

Proof: Suppose that $f(y_i, y_r) := 2\overline{y}_w \frac{1}{w_i}\frac{1}{w_r}$. This entails immediately that $\mathcal{G}$ belongs to $g(y)$ provided that $\frac{1}{w_i}, \frac{1}{w_r}, \overline{y}_w \neq 0$. Now, invoking the subgroup decomposition structure based on income pairs, that is, the subgroup decomposition methodology that brings out the inequality within the groups by computing the within-group income differences and the inequality between the groups by gauging the between-group

[1] These values are calculated considering population and region characteristics.

[2] They address this family of measures to show that the Gini mean difference and the Gini mean ratio belongs to this class of indices, without demonstrating the possibility to work with weighted Gini indices.

income differences, we have:

$$\begin{aligned}\mathcal{G} &= \sum_{i=1}^{n}\sum_{r=1}^{n}\frac{|y_i - y_r|}{2\overline{y}_w \frac{1}{w_i}\frac{1}{w_r}} \\ &= \underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\frac{|y_{ij} - y_{rj}|}{2\overline{y}_w \frac{1}{w_{ij}}\frac{1}{w_{rj}}}}_{\mathcal{G}_w} + \underbrace{2\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}\frac{|y_{ij} - y_{rh}|}{2\overline{y}_w \frac{1}{w_{ij}}\frac{1}{w_{rh}}}}_{\mathcal{G}_{gb}},\end{aligned} \tag{5}$$

where $\mathcal{G}_w$ and $\mathcal{G}_{gb}$ stands for the within-group weighted Gini index and for the between-group weighted Gini index, respectively. The weight w_{ij} is related to the ith individual of group P_j. Now, suppose that incomes are fractionated into q income sources (labour income, transfers, taxes, etc.) such as: $y_i = y_i^1 + \ldots + y_i^m + \ldots + y_i^q$ for the ith individual or $y_{ij} = y_{ij}^1 + \ldots + y_{ij}^m + \ldots + y_{ij}^q$ since the ith individual belongs to group P_j. Consider that $y^*_{j,ir}$ is an operator that takes the minimum value between the ith and the rth individual's income, and that $y^{m*}_{j,ir}$ is the mth source of this minimum. Hence, the weighted Gini ratio is multi-decomposable:

$$\begin{aligned}\mathcal{G} &= \sum_{m=1}^{q}\sum_{j=1}^{k}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\frac{y_{ij}^m + y_{rj}^m - 2y_{j,ir}^{*m}}{2\overline{y}_w \frac{1}{w_{ij}}\frac{1}{w_{rj}}} \\ &\quad + 2\sum_{m=1}^{q}\sum_{j=2}^{k}\sum_{h=1}^{j-1}\sum_{i=1}^{n_j}\sum_{r=1}^{n_h}\frac{y_{ij}^m + y_{rh}^m - 2y_{jh,ir}^{*m}}{2\overline{y}_w \frac{1}{w_{ij}}\frac{1}{w_{rh}}}. \ \square\end{aligned} \tag{6}$$

3 The income-receiving unit

Dealing with inequality indices implies that the analyst must choose between individuals or households as income-receiving units. Each one of these units has advantages and inconvenient. Although using individuals brings in higher variability in incomes than using household incomes and avoids the necessity of adjusting the size and the age composition of the household, household studies are preferable in order to obtain more realistic inequality comparisons. It is then necessary to include an equivalence scale, but there is no consensus in the choice of an appropriated one.

For instance, it is possible to measure the inequality of income distribution amongst persons but considering the information on family composition. As Cowell (1984) wrote:

"[...] *what we might be interested in is the income that an adult would have to have in order to enjoy the same standard of living as he presumably enjoys in the* *particular family within which he actually lives*". Cowell, (1984), p. 358

Some years ago, it was impossible to find any information, but nowadays the micro-data deliver indications both on household compositions and family incomes. Afterward, Cowell (1984) proposed a system of random weights treating the income-receiving units as a random sample of households issued from a bivariate distribution (household sizes, incomes).

Another eventuality would be the study of households as income units. In this case, it is possible to consider the total income family, but this indicator does not take in consideration the demographic characteristics of the households.

Finally, there is the use of incomes per head that represent total income family divided by the number of household members. But this methodology does not consider household characteristics any more. Consequently, working with households as income-receiving units requires an appropriated equivalent scale to introduce socio-demographic characteristics such as household size, composition by age, etc.

4 The bootstrap principle

The analysis and the measurement of inequalities are based on information obtained from sample surveys. These samples are subject to sampling and non-sampling errors. Sampling errors are due to the fact that we observe only a sub-set of the total population. Statistical inference deals with this problem and allows us to determine whether the estimated inequality measures represent the true population parameters.

The problem can be formulated as follows. Consider a random sample $X = (x_1, \ldots, x_n)$ of unknown probability distribution F from which we want to estimate a parameter ξ. We opt for an estimator $\hat{\xi} = s(X)$ using X. The question is: how accurate is $\hat{\xi}$? One possibility to answer this question is the use of bootstrap methods that allow to carry out the statistical inference for inequality measures.

The bootstrap technique is based on re-sampling with replacement. Each bootstrap sample, X^*, is an independent random sample of size n from the empirical distribution $\hat{F}$ which is formed by attaching probability $1/n$ to each observation x_i, $i = 1, \ldots, n$.[3] The elements of each bootstrap sample are the same than those of the original data set, but some observations will appear once, twice or zero times. To each bootstrap sample corresponds a bootstrap replication of $\hat{\xi}$: $\hat{\xi}^* = s(X^*)$. Following

[3]Notice that independence of observations is a sinequanon condition for the bootstrap method to be valid.

Efron and Tibshiriani (1993) the bootstrap algorithm providing the standard error of a parameter is summarized as follows.

(i) Select B independent bootstrap samples, $X^{*1}, \ldots, X^{*b}, \ldots, X^{*B}$ of n data values each one drawn with replacement from X.

(ii) Evaluate the bootstrap replication corresponding to each bootstrap sample:

$$\hat{\xi}^*(b) = s(X^{*b}) \text{ for all } b = 1, 2, \ldots, B. \tag{7}$$

(iii) Estimate the standard error by the sample standard deviation of the B replications,

$$s\hat{e}_{\hat{\xi}^*} = \left[\frac{\sum_{b=1}^{B} \left(\hat{\xi}^*(b) - \hat{\xi}^*(\cdot) \right)^2}{B-1} \right]^{1/2} \quad \text{for all } b = 1, 2, \ldots, B, \tag{8}$$

where

$$\hat{\xi}^*(\cdot) = \frac{\sum_{b=1}^{B} \hat{\xi}^*(b)}{B}. \tag{9}$$

It can be shown that the empirical standard deviation tends towards the population standard deviation as far as the number of bootstrap replications grows large, $B \longrightarrow \infty$.[4]

5 The stratified bootstrap principle and the weighted bidimensional decomposition

When we apply the bootstrap method to the global Gini index we must consider a density of probability F with an unknown associated law. We also consider a random sample of n individuals. The $\hat{F}$ sample is constructed by attaching a probability of $1/n$ to each observation. Each one of the individuals in the sample is affected with a j vector containing all the different studied income sources:

$$\hat{F} \longrightarrow \hat{Y} := \left[(y_1^{1*b} + \ldots + y_1^{q*b}, w_1^{*b}), (y_2^{1*b} + \ldots + y_2^{q*b}, w_2^{*b}), \ldots, (y_n^{1*b} + \ldots + y_n^{q*b}, w_n^{*b}) \right]. \tag{10}$$

From $\hat{F}$, we withhold B random samples with replacement from the initial sample, allowing us to construct the confidence interval for the global Gini index.

When bootstrap and subgroup decomposition are analyzed together, many prob-

[4]See the appendix, for a review of the different confidence interval that can be deduced from this method.

lems arise if we draw B samples from the studied sample. Taking households from the overall sample empirical distribution leads inevitably to introduce distortion in mean incomes and in number of individuals within each group. Indeed, in the case of a simple or a multilevel Gini decomposition[5], we immediately conclude (see Eq. (5)) that the within-group Gini $[G_w]$ converges towards 0 when the number of groups is equal to the number of households, and that the between-group Gini $[G_{gb}]$ converges towards 0 when the number of groups tends towards one. This remark holds in both cases of weighted Gini decompositions and unweighed Gini decompositions. Consequently, in order to neutralize these distortions in estimation within each group, it is possible to use the stratified bootstrap sample (see Bickel and Freedman (1984) that studied the validity of this method for many statistics). The stratified method consists in drawing samples within each group of the empirical distribution instead of drawing samples from the overall empirical distribution. We then draw B bootstrap samples from the empirical distributions of each group $\hat{F}_j$ for all $j = 1, \ldots, k$, by attaching the probability $1/n_j$ to each observation:

$$\hat{F}_j \longrightarrow \hat{Y}_j := \left[(y_{1j}^{1*b} + \ldots + y_{1j}^{q*b}, w_{1j}^{*b}), (y_{2j}^{1*b} + \ldots + y_{2j}^{q*b}, w_{2j}^{*b}), \ldots, (y_{n_j j}^{1*b} + \ldots + y_{n_j j}^{q*b}, w_{n_j j}^{*b}) \right]. \tag{11}$$

Consequently, for all b samples, the number of households within each group remains equal to that of the empirical distribution. Then, it is possible to measure B weighted Gini multi-decompositions using the $\hat{Y}_j$'s concatenations and derive the well-known confidence intervals (see the appendix). Indeed, if the empirical distribution involves k groups, we draw k samples from each group in a first stage, and in a second stage, we estimate the multi-decomposition. We reiterate these stages B times.

6 Empirical illustration

We illustrate the technique on Italian data using the Survey on Household's Income and Wealth (Bank of Italy, 2000). The empirical distribution involves 4,007 households partitioned into three areas: North, Centre, and South and Islands. Households's income is composed of three expenditure sources plus saving: (A) expenditure for transport and equipment, (B) expenditure for furniture, (C) non-durable

[5] In this paper we only focus on simple subgroup Gini decomposition. The multilevel ones concern the groups of the empirical sample P_j, which are partitioned into many subgroups, which are themselves partitioned into many subgroups, etc.

expenditure, and (D) saving.[6]

As can be seen in Table 2, the main generating expenditures of inequalities are transport and equipment (0,1819) and saving (0,1553). The generating groups of inequalities are characterized by income differences between North and South (0,1298) and within Centre (0,1010). The advantage of the multi-decomposition technique is to provide a precise determination of the combined "group/source" inequalities, showing that the most important combinations are produced by the differences between North and South in transport and equipment, and within North in transport and equipment, with 0,0661 and 0,0408 respectively.[7]

Using the simple percentile method (see A2. in appendix), we compute confidence intervals based on stratified bootstrap. The combinations " Inequalities within group j / from Source m " and " Inequalities between groups j and h / from Source m " are determined with the possibility to verify that the couples are significant.[8]

Table 2: Estimation of the multi-decomposition and stratified confidence intervals**

Sources → Group Inequalities ↓	Source A	Source B	Source C	Source D	Total Groups
Inequalities in North	0.0408 [0.0368 , 0.0448]	0.0037 [0.0019 , 0.0055]	0.0012 [0.0008 , 0.0017]	0.0364 [0.0311 , 0.0430]	0.0821
Inequalities in Centre	0.0051 [0.0042 , 0.0062]	0.0007 [0.0004 , 0.0011]	0.0002 [0.0001 , 0.0002]	0.0042 [0.0032 , 0.0053]	0.101
Inequalities in South and Islands	0.018 [0.0157 , 0.0208]	0.0016 [0.0008 , 0.0026]	0.0005 [0.0003 , 0.0007]	0.0182 [0.0146 , 0.0219]	0.0383
Inequalities between North/Centre	0.029 [0.0264 , 0.0319]	0.0033 [0.0022 , 0.0045]	0.0009 [0.0007 , 0.0012]	0.0251 [0.0219 , 0.0287]	0.0583
Inequalities between North/South	0.0661 [0.0615 , 0.0709]	0.0074 [0.0057, 0.0093]	0.0023 [0.0019 , 0.0029]	0.054 [0.0484 , 0.0599]	0.1298
Inequalities between Centre/South	0.0227 [0.0201 , 0.0256]	0.0028 [0.0019 , 0.0039]	0.0006 [0.0004 , 0.0009]	0.0176 [0.0149, 0.0204]	0.0437
Total Sources	0.1819	0.0195	0.0056	0.1553	$\mathcal{G} \cong 0.3623$ [0.3506 , 0.3756]

* Source: Bank of Italy "Survey on Household's Income and Wealth 2000"

**[.]: Confidence intervals based on stratified bootstrap

[6]The income units are divided by the corresponding value of the equivalence scale built by Carbonaro (1985, 2002) for the Commissione di indagine sulla povertà created by the Presidency of the Council of Ministers of the Italian Government.

[7]Another advantage of the multi-decomposition technique is to reveal, in some sample cases, independent marginal decompositions. Indeed, the sub-group and income sources decompositions can be independent. For instance the higher group inequalities can be Group 1, the higher source inequalities Source 1, whereas the higher generating couple can be "Group 2/Source 1". Note that in our study this is not the case, but the computation of the multi-decomposition is necessary to provide a conclusion on the independence and on the intensity of the couples.

[8]As we concatenate the sampling distributions of the k groups for all b, it is possible to compute the bootstrap intervals of the margins, that is, the confidence intervals of each consumption expenditure, of the within- and the between-group inequalities (Eq. 5). See the literature for this subject, e.g. Biewen (2002).

The non-stratified bootstrap method also shows that all the couples are actually significant. However, a slight difference is noted, as can be shown in Table 3. Indeed, computing the amplitude of the confidence interval (upper bound minus lower bound) with stratified bootstrap indicates that 21 couples out of 24 correspond to narrowest confidence intervals. Consequently, attention must be paid on the use of the well-known non-stratified bootstrap techniques, which lead to the conclusion that within- and between-group inequalities are more frequently significant than in stratified cases, for a given probability level.

Table 3: Amplitude of the stratified and non-stratified confidence intervals**

Sources → Group Inequalities ↓	Source A	Source B	Source C	Source D
Inequalities in North	0.00802549 [0.00989246]	0.00352234 [0.00360647]	0.00093683 [0.00098175]	0.01189225 [0.01206459]
Inequalities in Centre	0.00201616 [0.00231459]	0.00074438 [0.00079004]	0.00016133 [0.00016549]	0.00211704 [0.00235857]
Inequalities in South and Islands	0.00514012 [0.0064381]	0.00174991 [0.00182902]	0.00041548 [0.00043244]	0.00730249 [0.00791613]
Inequalities between North/Centre	0.0055121 [0.00633355]	0.00229132 [0.00241112]	0.00053835 [0.0005619]	0.00674351 [0.00738303]
Inequalities between North/South	0.00940961 [0.01073636]	0.0035091 [0.00339919]	0.00098108 [0.00096588]	0.01146824 [0.01165437]
Inequalities between Centre/South	0.00557332 [0.00616949]	0.00197787 [0.00210644]	0.00042161 [0.0003909]	0.00546796 [0.00616646]

* Source: Bank of Italy ”Survey on Household’s Income and Wealth 2000”

** []: Values of the confidence interval amplitudes based on non-stratified bootstrap

7 Conclusion

The combination of groups and sources when measuring inequalities deliver thorough information on the structure of income inequality but does not provide any precise indices if one does not recourse to weighted decompositions. These weights ensure that when a percentage of census data is used, individuals can reasonably be considered as drawn at random from the population. As this probability correction is possible in the context of the Gini multi-decomposition, the underlying couples ”Inequalities within group j / from Source m” and ”Inequalities between groups j and h / from Source m” are determinable with a higher precision.

But in order to know wether the couples are significant, the bootstrap principle appears as less challenging than the derivation of the variance estimators of the couples. Indeed, it would be possible to derive the within- and the between-group Gini indices with U-statistics, but it becomes overcomplicated for group and source combinations. Then, the bootstrap technique is welcomed especially as the stratified technique introduces less distortion in estimated Gini values, provided the size of each group is sufficiently large.

Appendix

Applying the stratified bootstrap method, whatever the following chosen confidence intervals, necessitates to draw within-group samples for all b (Eq. 11), and to compute the required lower and upper bounds.

A1. Standard bootstrap confidence interval

This method is constructed by the standard error method:

$$\hat{\xi} \pm \mu_{1-\alpha/2} s\hat{e}_{\hat{\xi}^*},$$

where $\mu_{1-\alpha/2}$ is the 100.$(1-\alpha/2)$-th percentile of the reduced normal distribution, $1-\alpha$ the selected significance level, and $s\hat{e}_{\hat{\xi}^*}$ the standard error computed on the B replications.

To make this approach efficient, the sample distribution of the studied parameter must be approximately normal, the estimator has not to possess any bias, and $s\hat{e}_{\hat{\xi}^*}$ must be a good estimation of the empirical distribution standard error.

A2. Simple percentile confidence interval

Suppose the bootstrap data set X^* are generated according to $\hat{F} \longrightarrow X^*$ and the bootstrap replications $\hat{\xi}^* = s(X^*)$ are computable. Let $\hat{\psi}$ be the cumulative distribution of $\hat{\xi}^*$. The $100-\alpha$ per cent confidence interval is defined by the 100.$\alpha/2$-th and the 100.$(1-\alpha/2)$-th percentiles of $\hat{\psi}(\cdot)$:

$$\left[\hat{\xi}_{inf}, \hat{\xi}_{sup}\right] \equiv \left[\hat{\psi}^{-1}(\alpha/2), \hat{\psi}^{-1}(1-\alpha/2)\right].$$

By definition $\hat{\psi}^{-1}(\alpha/2) := \hat{\xi}^*_{(\alpha/2)}$ is the 100.$\alpha/2$-th percentile of the bootstrap distribution. Thus:

$$\left[\hat{\xi}_{inf}, \hat{\xi}_{sup}\right] \equiv \left[\hat{\xi}^*_{(\alpha/2)}, \hat{\xi}^*_{(1-\alpha/2)}\right]$$

is the simple percentile confidence interval.

A3. Bias corrected percentile confidence interval

The bias corrected percentile bootstrap confidence interval for a given α significance level is defined by:

$$\left[\hat{\xi}_{inf}, \hat{\xi}_{sup}\right] \equiv \left[\hat{\xi}^*_{(\alpha_1)}, \hat{\xi}^*_{(\alpha_2)}\right],$$

where

$$\alpha_1 = \Phi\left(2\hat{z}_0 + \hat{z}^{\alpha/2}\right) \text{ and } \alpha_2 = \Phi\left(2\hat{z}_0 + \hat{z}^{1-\alpha/2}\right),$$

where Φ is the normal cumulative distribution function, and where $z^{\alpha/2}$ and $z^{1-\alpha/2}$ are the 100.$(\alpha/2)$-th and 100.$(1-\alpha/2)$-th percentiles, respectively. The value of $\hat{z}_0$ (bias correction) can be obtained directly auditioning the number of bootstrap replications ($\hat{\xi}^{*(r)}$) inferior to the initial estimator value ($\hat{\xi}$):

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\left\{\hat{\xi}^{*(r)} < \hat{\xi}\right\}}{B}\right),$$

where Φ^{-1} is the inverse of the normal cumulative distribution function.

A4. Bias corrected and accelerated confidence interval

Let BC_α be this confidence interval. The bounds of BC_α are given by the percentiles of the bootstrap distribution, which depend on $\hat{a}$ and $\hat{z}_0$ called acceleration and bias correction, respectively. The BC_α interval with $(1-\alpha)$ significance level is expressed as:

$$\mathrm{BC}_\alpha\colon = \left[\hat{\xi}^*_{(\alpha 1)}, \hat{\xi}^*_{(\alpha 2)}\right],$$

where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha/2)})}\right), \; \alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha/2)})}\right),$$

Φ the normal cumulative distribution function, $z^{\alpha/2}$ and $z^{1-\alpha/2}$ the 100.$\alpha/2$-th and 100.$(1-\alpha/2)$-th percentiles, respectively, and $\hat{z}_0$ the bias correction (see *supra*). The value of $\hat{a}$ is defined in jackknife values[9] of the ξ^* statistics. Suppose x_i is the brut

[9]The jackknife is another re-sampling method that allows one to estimate the standard error and the bias of a parameter. This method necessities to calculate n times the parameter's value with a different sample of $n-1$ observations. The n samples of size $n-1$ are always different since in each

sample to which the i-th observation has been taken, then:

$$\hat{\xi}_{(\cdot)} = \frac{\sum_{i=1}^{n} \hat{\xi}_{(x_i)}}{n}.$$

The acceleration can be expressed as:

$$\hat{a} = \frac{\sum_{i=1}^{n} \left(\hat{\xi}_{(\cdot)} - \hat{\xi}_{(x_i)} \right)^3}{6 \left[\sum_{i=1}^{n} \left(\hat{\xi}_{(\cdot)} - \hat{\xi}_{(x_i)} \right)^2 \right]^{3/2}}.$$

This equation is called acceleration because it obeys the rule of the standard error transformation of $\hat{\xi}$ respecting the true value of the ξ parameter.

A5. Hall's (1994) percentile method to construct confidence intervals

Hall's bias corrected confidence interval can be specified as:

$$\mathrm{BC}_{\alpha}^{\mathrm{H}} := \left[2\hat{\xi} - \xi_{sup}^{*} \, , 2\hat{\xi} - \xi_{inf}^{*} \right],$$

where ξ_{inf}^{*} and ξ_{sup}^{*} denote the $100.\alpha/2$-th and the $100.(1-\alpha/2)$-th lower and upper percentiles of the bootstrap distribution of ξ^{*}, respectively. This method is more appropriated than the simple percentile confidence interval. Obviously, the two methods produce the same interval if $\hat{\xi}$ is the midpoint between ξ_{inf}^{*} and ξ_{sup}^{*}.

A6. Hall's (1994) percentile confidence interval for the difference between two samples

Suppose we have two samples of sizes n_1 and n_2, drawn from completely unspecified probability distribution F_1 and F_2 and let $\Delta\hat{\xi}$ denote the difference between the estimated parameters, i.e, $\Delta\hat{\xi} := \hat{\xi}_1 - \hat{\xi}_2$. The bootstrap procedure simulates the distribution of the difference of these parameters. This is done by drawing bootstrap samples of size n_1 ad n_2 from the empirical distributions F_1 and F_2 respectively, and recording the difference between the estimated parameters. Repeating this routine B

re-sampling stage, a new observation is eliminated.

times (B bootstrap samples) produces:

$$\Delta\xi^{*b} = \xi_1^{*b} - \xi_2^{*b}, \text{ for all } b = 1, \ldots, B.$$

Hall's percentile confidence interval for the difference between the estimated values computed on two samples is deduced from the bootstrap distribution:

$$\Pr\left(2\Delta\hat{\xi} - \xi^*_{inf} \leq \Delta\xi \leq 2\Delta\hat{\xi} - \xi^*_{sup}\right) = \frac{100-\alpha}{100},$$

where $\Delta\xi^*_{sup}$ and $\Delta\xi^*_{inf}$ are the $100.\alpha/2$-th and $100.(1-\alpha/2)$-th lower and upper percentiles of the bootstrap distribution for the difference of two parameters, respectively. If this confidence interval does not include zero, we can conclude that the difference between the two parameters estimated from the empirical distributions F_1 and F_2 is statistically significant.

References

Athanasopoulos, G. and F. Vahid (2003), "Statistical Inference on Changes in Income Inequality in Australia", *The Economic Record* 79(247), 412-424.

Bickel, P. J. and D. A. Freedman (1984), "Asymptotic normality and the bootstrap in stratified sampling", *The Annals of Statistics* 12, 470-482.

Biewen, M. (2002), "Bootstrap inference for inequality, mobility and poverty measurement", *Journal of Econometrics* 108(2), 317-342.

Biewen, M. and S.P. Jenkins (2006), "Variance estimation for Generalized Entropy and Atkinson inequality indices: the complex survey case", forthcoming in: *Oxford Bulletin of Economics and Statistics.*

Carbonaro, G. (1985), "Nota sulla scala di equivalenza", in Commissione di indagine sulla povertà, Study di base, Presidenza del Consiglio dei Ministri, Roma.

Carbonaro, G. (2002), *Studi sulla povertà: problemi di misura e analisi comparative*, Franco Angeli, Milano.

Cowell, F. A. (1984), "The Structure of American Income Inequality", *Review of Income and Wealth* 30(3), 351-75.

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

Efron, B. (1995), *Le Bootstrap et ses Applications. Discrimination et Régression*, C.I.S.I.A.

Efron, B. and Tibshirani R. (1993), *Introduction to the Bootstrap*, New York, Springer-Verlag.

Hall, P. (1992), *The Bootstrap and Edgeworth expansion*, Springer Series in Statistics.

Hall, P. (1994), ”Methodology and Theory for the Bootstrap”, in Engle and McFadden (eds.), *Handbook of Applied Econometrics* 4, North Holland.

Mills, J. A. and S. Zandvakily (1997), ”Statistical Inference via Bootstrapping for Measures of Inequality”, *Journal of Applied Econometrics* 12, 133-150.

Mussard, S. (2004), ”The bidimensional Decomposition of the Gini Ratio. A Case Study: Italy”, *Applied Economics Letters* 11(8), 503-505.

Mussard, S. (2006), ”Une nouvelle décomposition de la mesure de Gini en sources de revenu, et la décomposition en sous-populations : une réconciliation”, forthcoming in: *Annales d'Economie et de Statistique.*

Mussard, S. and M. Terraza (2006), ”Gini ou Entropie ? Un choix basé sur le critère de multi-décomposabilité des mesures d'inégalité”, forthcoming in: *Journal de la Société Française de Statistique.*

Sen, A. K. (1973), *On Economic Inequality*, Clarendon Press, Oxford.

Xu, K. (2000), ”Inference for Generalized Gini Indices Using the Iterated-Bootstrap Method”, *Journal of Business and Economic Statistics* 18(2), 223-27.