Socio-Economic Planning Sciences 44 (2010) 212–219

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps

On the suitability of income inequality measures for regional analysis: Some evidence from simulation analysis and bootstrapping tests

Boris A. Portnov [a,*], Daniel Felsenstein [b]

[a] Department of Natural Resources and Environmental Management, Graduate School of Management, University of Haifa, Israel
[b] Department of Geography, Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel

ARTICLE INFO

Article history:
Available online 13 April 2010

Keywords:
Regions
Inequality measures
Bootstrapping
Random permutation tests

ABSTRACT

The paper looks at the sensitivity of commonly used income inequality measures to changes in the ranking, size and number of regions into which a country is divided. During the analysis, several test distributions of populations and incomes are compared with a 'reference' distribution, characterized by an even distribution of population across regional subdivisions. Random permutation tests are also run to determine whether inequality measures commonly used in regional analysis produce meaningful estimates when applied to regions of different population size. The results show that only the population weighted coefficient of variation (Williamson's index) and population-weighted Gini coefficient may be considered sufficiently reliable inequality measures, when applied to countries with a small number of regions and with varying population sizes.

1. Introduction

The study of inequality across regions is rather different to the study of inequality between individuals. This derives from the fact that regions are groups formed by individuals. This is not as obvious as it may sound. For example, a tradition exists in the regional income convergence literature that treats regions as individual observations regardless of the size of the former (cf. e.g., [1]). As such, large and small regions are assumed to carry equal weight, just as fat and thin people are treated equally when looking at inequality between them.

The computational issues associated with multi-group comparisons of income inequality were noticed (apparently for the first time) by the American economist Max Lorenz. In his seminal paper published in 1905, Lorenz highlighted several drawbacks associated with the comparison of wealth concentrations between fixed groups of individuals. In particular, he found that while an increase in the percentage of the middle class is supposed to show the diffusion of wealth, a simple comparison of percent shares of persons in each income group may often lead to the opposite conclusion. For instance, while the upper income group in a particular period may constitute a smaller proportion of the total population, the overall wealth of this group may be far larger compared to another time period under study ([2]: 210–211). The remedy he suggested was to represent the actual inter-group income distribution as a line, plotting 'along one axis cumulated percents of the population from poorest to richest, and along the other the percent of the total wealth held by these percents of the populations' (ibid. p. 217).

In an essay published in 1912, the Italian statistician Corrado Gini moved Lorenz's ideas a step further, suggesting a simple and easy comprehendible measure of inequality known as the Gini coefficient. Graphically, the calculation of this coefficient can be interpreted as follows:

$$\text{Gini coefficient} = \frac{\text{Area between Lorenz curve and the diagonal}}{\text{Total area under the diagonal}}$$

Mathematically, the Gini coefficient is calculated as the arithmetic average of the absolute value of differences between all pairs of incomes, divided by the average income (see Table 1).[1] The coefficient takes on values between 0 and 1, with zero interpreted as perfect equality [3].

A few years later, Dalton [4] carried out the first systematic attempt to compare the performance of different inequality measures against 'real world' data. As he noted, many inequality measures, though having intuitive or mathematical appeal, react to changes in income distribution in an unexpected fashion. For

* Corresponding author.
E-mail address: portnov@nrem.haifa.ac.il (B.A. Portnov).

[1] The computation includes the cases where a given income level is compared with itself.

0038-0121/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.seps.2010.04.002

Table 1
Commonly used measurements of regional inequality.

Coefficient of variation (CV) (unweighted)	Population weighted coefficient of variation (Williamson index (WI))
$CV = \frac{1}{\bar{y}}[\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2]^{1/2}$	$WI = \frac{1}{\bar{y}}[\sum_{i=1}^{n}(y_i - \bar{y})^2 \frac{A_i}{A_{tot}}]^{1/2}$
Theil index (TE(0))	Atkinson index (AT)
$TE(0) = \frac{1}{n}\sum_{u=1}^{n} \log\frac{\bar{y}}{y_i}$	$AT = 1 - [\frac{1}{n}\sum_{i=1}^{n}[\frac{y_i}{\bar{y}}]^{1-\varepsilon}]^{1/(1-\varepsilon)}$
Hoover coefficient (HC)	Coulter coefficient (CC)
$HC = \frac{1}{2}\sum_{i=1}^{n} \lvert\frac{A_i}{A_{tot}} \frac{y_i}{\bar{y}} - \frac{A_i}{A_{tot}}\rvert$	$CC = [\frac{1}{2}\sum_{i=1}^{n}(\frac{A_i}{A_{tot}} \frac{y_i}{\bar{y}} - \frac{A_i}{A_{tot}})^2]^{1/2}$
Gini (U) (unweighted)	Gini (W) (population weighted)
$Gini = \frac{1}{2n^2\bar{y}}\sum_{i=1}^{n}\sum_{j=1}^{n} \lvert y_i - y_j\rvert$	$Gini = \frac{1}{2\bar{y}}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{A_i}{A_{tot}}\frac{A_j}{A_{tot}}\lvert y_i - y_j\rvert$

Note: A_i and A_j = number of individuals in regions i and j respectively (regional populations), A_{tot} = the national population; y_i and y_j = development parameters observed respectively in region i and region j (e.g., per capita income); $\bar{y}$ is the national average (e.g., per capita national income); n = overall number of regions; ε is an inequality aversion parameter, $0 < \varepsilon < \infty$ (the higher the value of ε, the more society is concerned about inequality).

instance, if all the incomes are simply doubled, the variance quadruples the estimates of income inequality. Dalton's second observation was that some inequality measures do not comply with a basic principle of population welfare set forward by Arthur Pigou and formulated as follows: 'if there are only two income-receivers, and a transfer of income takes place from the richer to the poorer, inequality is diminished' (ibid. p. 351). After applying the 'principle of transfers' to various inequality measures, Dalton found that most measures of deviation (e.g., the mean standard deviation from the arithmetic mean, and the coefficient of variation) are perfectly sensitive to transfers and pass the 'test with distinction' (ibid. p. 352). The Gini index, commonly used in empirical studies, was also found by Dalton sufficiently sensitive to income transfers. He also found that the standard deviation is sensitive to transfers among the rich, while the standard deviation of logarithms is less sensitive to transfers among the rich than to transfers among the poor but still changes when a transfer among the rich takes place.

Two other fundamental requirements for a 'robust measure' of inequality, set forward by Dalton, are the principle of *proportional addition to incomes*, and the principle of *proportional increase in population*. According to the former, a proportional rise in all incomes diminishes inequality, while the proportional drop in all incomes increases it. According to the latter principle, termed by Dalton the 'principle of proportional additions to persons,' a robust inequality measure should be invariant to proportional increase in the population sizes of individual income groups. Dalton's calculations showed that most commonly used measures of inequality comply with these basic principles. Only the most 'simple' measures, such as absolute mean deviation, absolute standard deviations and absolute mean difference, fail to indicate any change, when proportional additions to the numbers of persons in individual income groups are applied (ibid. pp. 355–357, see also [5], pp. 87–112).[2]

Yitzhaki and Lerman [6]noted another deficiency inherent to most inequality measures, viz., insensitivity to the position which a specific population subgroup occupies within an overall distribution. Their Gini decomposition technique (see below) takes group-specific positions into account. In particular, they suggested weighting subgroups by the average rank of their members in the distribution. This is in contrast to the weighting system used more conventionally according to which between-group inequality is weighted by the rank of the average [7,8]. This latter system results in a large residual when inequality is decomposed into within and between groups. In contrast, the Yitzhaki approach results in a more concise decomposition with no residual [9].

Other empirical studies proposed and used a variety of additional inequality measurements, such as the population weighted coefficient of variation (Williamson's index), Theil index, Atkinson index, Hoover and Coulter coefficients [3,6,10–15]. However as the Gini measure ranges between 0 and 1 and is unaffected by change of scale (the population principle), it has become probably the most attractive measure for inequality in regional analysis.

While there have been numerous attempts to test the conformity of commonly used inequality measures with basic inequality criteria – e.g., principles of transfer, proportional addition to incomes, and proportional addition to population – (see *inter alia* [4,5,11]), there appears to be no systematic attempt to verify whether all of these measures are equally suitable for regional analysis, in which individual countries may be represented both by a different numbers regions and by regions of different population sizes. The lack of interest to this aspect of inequality measurement may have a simple explanation. Since commonly used inequality indices (some of which appear in Table 1) are abstract mathematical formulas, one can assume that they can be applied to both large and small countries alike and provide fully comparable results. However, it is well known that the use of different measurement indices in regional analysis gives rise to highly variable results. For example, the notion of optimal regional convergence (i.e., that point where regional convergence also reduces overall nation-level inequality) has been shown to be highly dependent on the type of inequality index used [16] as is the measurement of regional price convergence [17].

The present paper attempts to determine whether commonly used inequality measures are sufficiently sensitive to changes in the ranking, size and number of regions into which a country is divided. The paper is organized as follows. First, we look at the specificity of measuring regional inequality. Given the fact that the unit of observation (i.e., a region) is a group measure, it presumably needs some weighting as regions of a country come in different sizes. We then proceed to discuss the general principles that should govern in our view, the selection of robust inequality measures. Then we move to testing the compliance of different commonly used inequality measures against the set of criteria that should characterize a robust inequality measure. The tests are run in two phases. First, we use a number of pre-designed distributions, to verify whether a particular inequality measure meets our intuitive expectations concerning inequality estimates. Then, in the second stage of the analysis, we run more formal permutation tests to verify whether different inequality measurements respond sensibly to changes in the population distribution across the space.

2. Sizes and shapes of regions

General economic theory does not suggest *a priori* that the size and number of regions in a country should affect the distribution of inequality. Beenstock [18] investigated this issue, testing whether regional amalgamation (decreasing the number of regions) impacts on inequality between them. His analysis shows that unifying any two regions will increase the earnings of each and reduce inequality between them but the same cannot be said for the level of overall inequality between all regions. In terms of regional size, a similar conclusion is drawn from economic theory. Beenstock's work shows that regional size *per se* has little impact on the determination of regional inequality and regional social and

[2] Dalton ([4]: 352) distinguishes between measures of relative dispersion and measures of absolute dispersion. Whereas the former measures are dimensionless, the measures of absolute dispersion are estimated in units of income. The latter measures are easily transformed in the former by normalization.