'Williamson's Fallacy' in Estimation of Inter-Regional Inequality

Gluschenko, Konstantin

Novosibirsk State University

20 November 2015

‘Williamson’s Fallacy’ in Estimation of Inter-Regional Inequality

[In the journal version, **Measuring Regional Inequality: To Weight or not to Weight?**]

KONSTANTIN GLUSCHENKO
Institute of Economics and Industrial Engineering, Siberian Branch of the Russian Academy of Sciences, Russia;
and
Novosibirsk State University, Russia

pr. Lavrentieva 17, 630090 Novosibirsk, Russia.
Tel.: +7 383 330 2548;
fax: +7 383 330 2580.
E-mail address: glu@nsu.ru

First version: November 2015

This version: May 2017

ABSTRACT *When estimating regional inequality, many economists use inequality indices weighted by the regions’ shares in the national population. Although this approach is widespread, its adequacy has not received attention in the regional science literature. This paper proves that such approach is conceptually inconsistent, yielding an estimate of interpersonal inequality among the whole population of the country rather than an estimate of regional inequality. Nevertheless, as a measure of interpersonal inequality, such an estimate is very rough (up to misleading) and does not always have an intuitive interpretation. Moreover, the population-weighted inequality indices do not meet the requirements for an adequate inequality measure.*

KEYWORDS: *inequality index; weighting by population; Williamson coefficient of variation; inequality axioms*

JEL CLASSIFICATION: D31; D63; R10

1. Introduction

When studying economic inequality in a country, an economist may consider the distribution of income between individuals or between country's regions. The latter not only introduces spatial dimension in studies of inequality, but it can also reveal important links remaining overlooked with treating the country as a whole. For example, while the literature on civil war has found little support for a link between individual-level economic inequality and civil war, Deiwiks et al. (2012) find strong evidence that regional inequality affects the risk of secessionist conflict. In both cases, the same statistical methodology and inequality indices (which amount to a few tens) are applied, with the difference that regions rather than individuals are taken as observations while estimating regional inequality. However, there is a modification of the inequality indices that is applied in order to measure regional inequality.

Apparently, Williamson (1965) was the first who put forward the idea of weighting indices that measure the inequality between the regions of a country by the regions' shares in the national population. Since then, such an approach became fairly widespread in regional studies. Publications that use it number in hundreds. Therefore I am able to cite only a small part of them, using a dozen of recent journal articles as a 'sample'. Table 1 tabulates them, reporting the population-weighted inequality indices applied as well as the geographical and temporal coverage of the respective studies. In this table, CV = coefficient of variation, G = Gini index, Th = Theil index, MLD = mean logarithmic deviation, σ = standard deviation of logarithms and RMD = relative mean deviation. Subscript w indicates that the index is population-weighted.

Most studies from Table 1 use the regional GDP per capita as a well-being indicator. An exception is Doran & Jordan (2013) who exploit the regional gross value added per capita; a few studies consider some additional indicators. The table shows that the application of the population-weighted inequality indices is greatly varied both in geographical terms and time spans (note that if different countries are involved in a study, the case at hand is not international inequality; the study deals with regional inequalities in a respective set of countries). The inequality indices employed are also manifold. The most popular ones are the coefficient of variation, and the Gini and Theil indices (many other, 'out of sample', papers confirm this). Therefore, only these three indices will be dealt with in what follows. It should be noted that the population-weighted indices are present not only in the literature on economic inequality; they find use in studies of inequality in the areas of health care, education, energy policy, etc.

Table 1. Selected recent studies that use the population-weighted inequality indices.

Author(s)	Weighted index(es) employed	Geographical coverage	Time span
Doran & Jordan (2013)	Th_w	14 EU countries (NUTS 2 regions)	1980–2009
Enflo & Rosés (2015)	MLD_w	Sweden	1860–2000
Ezcurra & Rodríguez-Pose (2014)	Th_w, CV_w, MLD_w, σ_w	22 emerging countries	1990–2006
Kyriacou & Roca-Sagalés (2014)	CV_w, MLD_w, σ_w	22 OECD countries	1990–2005
Lessmann (2014)	CV_w	56 countries	1980–2009
Li & Gibson (2013)	G_w, CV_w, Th_w	China	1990–2010
Martínez-Galarraga et al. (2015)	Th_w	Spain	1860–2000
Mussini (2015)	G_w	28 EU countries (NUTS 3 regions)	2003–2011
Petrakos & Psycharis (2016)	CV_w	Greece	2000–2012
Sacchi & Salotti (2014)	CV_w, σ_w, MLD_w	21 OECD countries	1981–2005
Wijerathna et al. (2014)	G_w, CV_w, RMD_w	Sri Lanka	1996–2011
Zubarevich & Safronov (2011)	G_w, CV_w	Russia, Ukraine, Kazakhstan	1998–2009

Williamson did not provide a more or less detailed substantiation of his idea, instead he merely noted that an unweighted inequality index “will be determined in part by the somewhat arbitrary political definition of regional units” and “[t]he preference for an unweighted index over a weighted one, we think, is indefensible” (Williamson, 1965, pp. 11, 34). Nor such substantiations appeared within next 50 years. Even a handbook chapter on measuring regional divides only asserts that the use of unweighted inequality indices “may lead to unrealistic results in certain cases, affecting our perception of convergence or divergence trends” (Ezcurra & Rodríguez-Pose, 2009, p. 332), providing no proof or example. A sole attempt to explore the properties of the population-weighted indices is made by Portnov & Felsenstein (2010), and this will be discussed in Section 5. Yet even Williamson’s cited notes are open to question.

First, the political division of a country is the reality which regional researchers should address, irrespective of whether they believe it to be ‘somewhat arbitrary’ or ‘natural’. Certainly, they may discuss its shortcomings and find ways of improvement, but it is a quite different story unrelated to the issue of regional inequality. Therefore the desire for ‘adjustment’ of the existing political division by assigning less importance to lesser populated regions seems strange.

Second, why do we need to take into account the differences in regional population at all? But we can estimate the inequality among the groups in a country's population without regard for the sizes of these groups. For instance, while estimating wage inequality between industrial workers, builders, teachers, lawyers and so on, we do not focus on what shares of these occupational groups in the total population (or employees) are. What is a fundamental difference between this and the case when each population group consists of the inhabitants of one region?

Third, upon closer inspection the results of the estimating inequality with the use of the population-weighted indices look striking; they may prove to be evidently unrealistic. The next section gives an impressive example.

The purpose of this paper is to show that the application of population-weighted indices for measuring regional inequality is nothing but a fallacy. The main point is that they measure not the inequality between regions but something else and therefore yield distorted estimates of regional inequality. In other words, the unweighted and weighted indices **measure different phenomena**. Albeit Williamson's approach has received some criticism (which will be discussed in Section 5), the literature has overlooked this point. Moreover, this paper proves that these indices do not meet the requirements for an adequate inequality measure.

The statement above seemingly contradicts the fact that the approach under consideration is commonly employed in the literature. However, this fact in no way evidences adequacy of the approach. For instance, analyzing β-convergence is even more widespread (in the literature on economic growth and inequality); relevant publications number in thousands. Nonetheless, a number of authors, e.g. Friedman (1992), Quah (1993) and Wodon & Yitzhaki (2006), proved invalidity of this methodology.

The rest of the paper is organized as follows. Section 2 reveals the true sense of the inequality estimates obtained with the use of the population-weighted indices. Section 3 considers the issues of bias in the weighted indices and the interpretability of these indices. Section 4 analyzes the properties of the population-weighted indices, providing proofs that they violate three important axioms. Section 5 discusses the arguments against and in favour of population weighting found in the literature. Section 6 summarizes conclusions drawn in the paper.

2. What Do Population-Weighted Indices Measure?

Consider the cross-region income distribution $y = (y_i)$, $i = 1, \ldots, m$; y_i = per capita income in region i and $\bar{y}$ = the arithmetic average of the regional per capita incomes ($\bar{y} = (y_1 + ... + y_m)/m$). Then the coefficient of variation measuring regional inequality has the form

$$CV = \frac{\sqrt{\sum_{i=1}^{m}(y_i - \bar{y})^2 / m}}{\bar{y}}. \tag{1}$$

Now let N_i = population of region i; N = population of the country; $n_i = N_i/N$ = region's share in the national population (region's weight); $n = (n_i)$ will be called population distribution. The weighted average of the regional per capita incomes ($\bar{y}_{(w)} = n_1 y_1 + ... + n_m y_m$) is denoted by $\bar{y}_{(w)}$. It equals the national per capita income: $\bar{y}_{(w)} = (Y_1 + ... + Y_m)/N = Y/N$, where Y_i stands for the region's total income ($Y_i = N_i y_i$) and Y represents the national total income. Under this notation, the Williamson coefficient of variation (Williamson, 1965, p. 11) – sometimes referred to as the Williamson index – looks like

$$CV_w = \frac{\sqrt{\sum_{i=1}^{m}(y_i - \bar{y}_{(w)})^2 n_i}}{\bar{y}_{(w)}}. \tag{2}$$

The Gini and Theil indices can be respectively written as

$$G = \frac{\sum_{i=1}^{m}\sum_{k=1}^{m}|y_i - y_k|}{2m^2\bar{y}}; \tag{3}$$

$$Th = \frac{1}{m}\sum_{i=1}^{m}\frac{y_i}{\bar{y}}\ln(\frac{y_i}{\bar{y}}). \tag{4}$$

Their population-weighted counterparts take the forms

$$G_w = \frac{\sum_{i=1}^{m}\sum_{k=1}^{m}n_i n_k |y_i - y_k|}{2\bar{y}_{(w)}}; \tag{5}$$

$$Th_w = \sum_{i=1}^{m} n_i \frac{y_i}{\bar{y}_{(w)}}\ln(\frac{y_i}{\bar{y}_{(w)}}). \tag{6}$$

In some cases, the weighting by population is present in the Theil index implicitly. For example, Doran & Jordan (2013, pp. 25–26) construct the index from the regions' contributions to the total income, Y_i/Y, and regions' shares of total population, N_i/N. Martínez-Galarraga et al. (2015, p. 510) use a similar method. It is easily seen that such an index is equivalent to the one represented by Formula (6):

$$\sum_{i=1}^{m} \frac{Y_i}{Y} \ln(\frac{Y_i / Y}{N_i / N}) = \sum_{i=1}^{m} \frac{N_i y_i}{N\bar{y}_{(w)}} \ln(\frac{Y_i / N_i}{Y / N}) = \sum_{i=1}^{m} n_i \frac{y_i}{\bar{y}_{(w)}} \ln(\frac{y_i}{\bar{y}_{(w)}}) = Th_w .$$

Let us perform a simple test, applying the population-weighted indices in order to estimate regional inequality in a two-region case. Consider two Chinese regions, namely mainland China as a whole (in Chinese, 大陆 – Dàlù) and Macao, the Special Administrative Region of the People's Republic of China (and the richest territory of the world). In hoary antiquity, when the Portuguese occupied as large a part of the Chinese territory as they could (or needed), Macao might be deemed a 'somewhat arbitrary' regional unit. Nowadays, it is quite natural, as Macao has its own currency, and citizens of China from other regions need a visa to get there. Table 2 reports data on these regions.

Table 2. Per capita income and population in mainland China and Macao in 2014.

Region	PPP-adjusted GDP per capita (y_i), current international dollars[1]	Population (N_i), million people[2]	Region weight (n_i)
Mainland China	13,217	1,376.049	0.999573
Macao	139,767	0.588	0.000427

Upon estimating the income inequality between mainland China and Macao, we get the results listed in Table 3. It reports the values of the population-weighted coefficient of variation and Gini and Theil indices defined by Formulae (2), (5) and (6) as well as the values of the unweighted indices according to Formulae (1), (3) and (4).

Table 3. Estimates of income inequality between mainland China and Macao.

Index	Population-weighted		Unweighted	
	Raw	Standardized	Raw	Standardized
Coefficient of variation	0.197	0.004	0.827	0.827
Gini index	0.004	0.004	0.414	0.827
Theil index	0.007	0.001	0.399	0.576
Average income	$\bar{y}_{(w)}$ = 13,721		$\bar{y}$ = 76,492	

For comparability, the table also reports these indices in a standardized manner so that

[1] Data source: World development indicators. – GDP per capita, PPP (current international $). http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD (Accessed Sept. 26, 2015).

[2] Data source: World population prospects: the 2015 revision, key findings and advance tables, Working Paper No. ESA/P/WP.241, United Nations, Department of Economic and Social Affairs, Population Division, New York, 2015, p. 13.

they range from 0 to 1. That is, an index is divided by its maximum corresponding to perfect inequality. For our case of two observations, the maxima of *CV*, *G* and *Th* are respectively 1, 0.5 and log(2). The maxima of CV_w, G_w and Th_w approximately equal 1, 48.4 and 7.8; the way of computing these maxima will be explained in Section 4 and summarized in its Table 8.

While the unweighted indices indicate a high degree of inequality, the population-weighted ones yield the reverse pattern. The standardized values of CV_w and G_w are equal to 0.4% in percentage terms; and the standardized Th_w is even less than 0.1%. This suggests that there is (almost) no income inequality between the average mainland Chinese and the average inhabitant of Macao. Indeed, our perception of spatial inequality is greatly distorted, but in the sense opposite to the above-cited view of Ezcurra & Rodríguez-Pose (2009, p. 332): it is the population weighting that gives rise to distortions.

In the two-region case, the result evidently contradicts common sense. A sufficiently great number of regions in the empirical studies masks, as a rule, such absurdities, thus creating the impression that the estimates of inequality with the use of population weighting are reasonable.

Then what is the reason for that low inequality suggested by the population-weighted inequality indices in the above example? What is the sense of the estimates obtained? In order to understand what the weighted indices measure, let us estimate inequality among all the citizens of a country on the basis of the cross-region income distribution. The 'national' coefficient of variation (CV_{nat}) with y_l standing for personal income of *l*-th citizen of the country looks like

$$CV_{nat} = \frac{\sqrt{\sum_{l=1}^{N}(y_l - \bar{y})^2 / N}}{\bar{y}}.$$

Obviously, the population-average income in this formula – national per capita income, $\bar{y} = (y_1 + ... + y_N)/N$ – equals the weighted average of the regional per capita incomes, $\bar{y}_{(w)}$. Lacking information on intra-regional income distributions, we are forced to assume that all the inhabitants of a region have the same income equalling per capita income in this region. Then the square deviations $(y_l - \bar{y}_{(w)})^2$ are uniform for all *l* relating to the inhabitants of the same region, say *i*. Hence, their sum over all the inhabitants of the region is $(y_i - \bar{y}_{(w)})^2 N_i$. Summing up such sums over all regions, we come to the Williamson coefficient of variation:

$$CV_{nat} = \frac{\sqrt{\sum_{i=1}^{m} (y_i - \bar{y}_{(w)})^2 N_i / N}}{\bar{y}_{(w)}} = \frac{\sqrt{\sum_{i=1}^{m} (y_i - \bar{y}_{(w)})^2 n_i}}{\bar{y}_{(w)}} = CV_w .$$

Thus, the population-weighted coefficient of variation **is not a measure of inequality between regions**; instead, **it measures national inequality**, i.e. interpersonal inequality in the whole population of the country. In doing so, it does not (and cannot) take into account intra-regional inequalities. Certainly, this relates not only to the coefficient of variation but to any other inequality index (maybe, except for those based on partial information from cross-region income distributions, e.g. the relative range of disparities, $R = \max_i y_i / \min_i y_i$, interquartile range, and the like; however, it seems that the weighting is hardly applicable to them); the proof is simple and similar to the one above. (Of course, we need not necessarily take an entire country; the conclusion still holds if we consider any subset of regions as a 'country'.)

This explains the sense of results obtained with the population-weighted indices in Table 3. These measure the inequality between the inhabitants of united mainland China and Macao. Provided that the inequality within mainland China is zero (as all its inhabitants are supposed to have the same income), adding less than one million people – even with extremely high income – to its 1.4-billion population can increase the degree of the overall inequality only slightly.

It is seen that there is a conceptual distinction between the unweighted and population-weighted estimates of inequality; **they measure different phenomena**. The unweighted index measures the inequality between regions (considered as a whole), while the weighted one measures the inequality between all the country's citizens.

Considering the inequality between regions, all of them enjoy equal rights in the sense that all y_i are equiprobable (i.e. the probability of finding income y_i in a randomly chosen region is the same for all i and equals $1/m$). Albeit speaking of regions, we actually deal with individuals, or in other words the representative (or 'average', i.e. having the region-average income) inhabitant of each region. While estimating regional inequality, we compare their incomes without considering how many people live in the respective regions (similarly to comparing the wages across occupations). Indeed, the fact that the average inhabitant of Macao is almost 11 times richer than the average mainland Chinese in no way changes because of the fact that the population of Macao is 2,340 times smaller than the population of mainland China.

Introducing regional weights implies that a region is represented by all its inhabitants rather than by one 'average' inhabitant. That is, we consider region i as a group of N_i people, each individual within the group having an income y_i. Then the probability of y_i differs across regions, becoming proportional to their populations, n_i. Thus, (n_i) is a proxy of the personal-income distribution in the country. In fact, we 'split' the regions into their individual inhabitants so that their aggregate represents the whole population of the country, as Figure 1 illustrates, and estimate the inequality between these N persons, so inevitably substituting regional inequality for an interpersonal one. However, lacking information on income differences within regions, we consider the inhabitants of each region as identical and the regions as internally homogeneous groups of people. Thus, we arrive at a grouping of the whole country's population into income classes (y_i) of different sizes (N_i). The regional division matters no more; the impression that the case at hand constitutes inequality between regions is but an illusion owing to that the grouping proceeds from the data by region. An estimate of the national inequality obtained with such grouping is very crude, since it neglects inequality within regions and – what is much more important – the income classes y_i (constructed from cross-region data) in fact heavily overlap because of the overlapping of the intra-regional income distributions. (This issue is considered in more detail in the next section.)

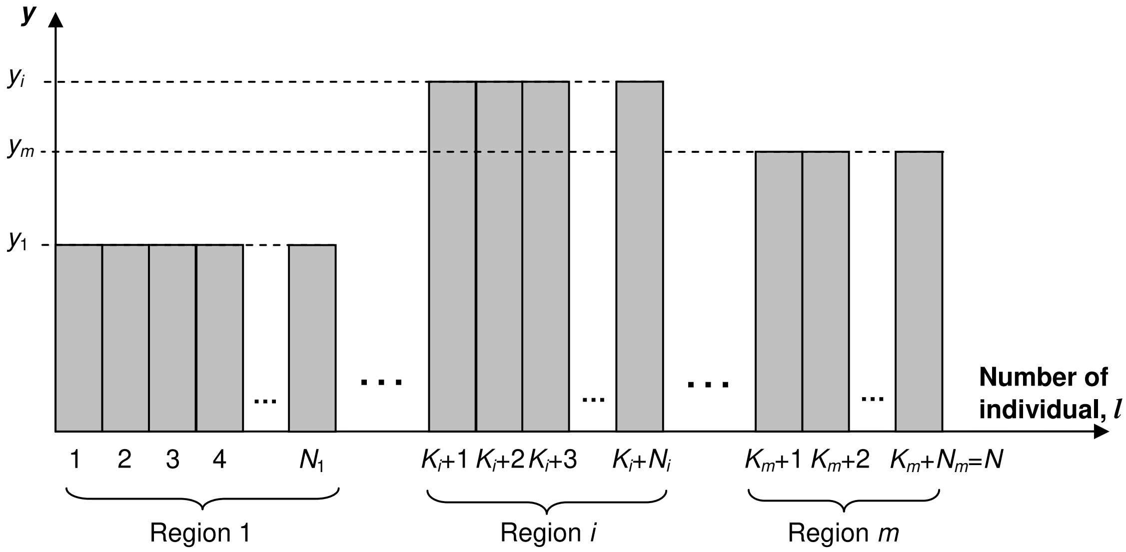


Figure 1. Population of the country as a set of regional populations (assuming the population of each region to be income homogeneous, $y_l = y_i$ for $K_i + 1 \le l \le K_i + N_i$).

Note: $K_i = \sum_{j=1}^{i-1} N_j$; $K_1 = 0$..

As it is known, inequality in the country, being measured by the Theil index, can be decomposed into two components: within-region inequality and between-region inequality. Under notation of this paper, the decomposition of national inequality looks like

$$Th_{nat} = \sum_{i=1}^{m} n_i \frac{y_i}{\bar{y}_{(w)}} Th_i + \sum_{i=1}^{m} n_i \frac{y_i}{\bar{y}_{(w)}} \log(\frac{y_i}{\bar{y}_{(w)}}) = \sum_{i=1}^{m} n_i \frac{y_i}{\bar{y}_{(w)}} Th_i + Th_w,$$

where Th_i = Theil index for the population of i-th region. Thus, the population-weighted Theil index represents only a part of national inequality, namely, between-region inequality. It answers to the counterfactual question: 'how much inequality would be observed [in the country] if there was no inequality within regions?' (Shorrocks & Wan, 2005, p. 60).

It follows herefrom that a population-weighted estimate of inequality is biased with regard to the estimates of both regional inequality (as it measures a different value) and interpersonal inequality (as it does not take account of the within-region income disparities). In both cases, the result can be misleading as the example of the two Chinese regions demonstrates.

The bias can have either direction depending on a particular combination of the regional per capita incomes and populations. Williamson (1965, p. 12) reports the values of both the weighted and unweighted coefficient of variation estimated from the regional data in 24 countries. The values CV_w prove to be overstated in about a half of the countries, and understated in another half. The biases (relative to the unweighted estimates) range from –52.6% (in India) to +37.6% (in Puerto Fico). The case of India is an example of quite misleading result in an actual study (covering 18 regions): the population-weighted index understates the extent of regional inequality there by more than a half.

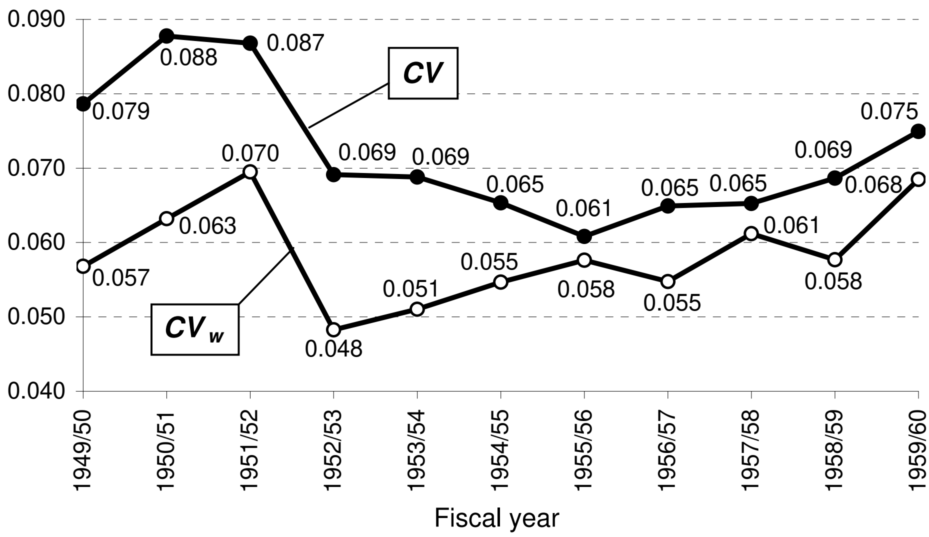


Figure 2. Paths of the weighted and unweighted coefficient of variation in Australia.

Differences in the trends also can occur. Figure 2 depicts the evolution of inequality in

Australia over 11 years according to the unweighted and population-weighted coefficient of variation. The estimates are computed from Williamson's (1965, p. 48) data. It is seen that the trends of CV and CV_w are sometimes opposite, e.g. in the whole period of 1952/53 to 1958/59. Regional inequality, measured by CV, fell by 4.7% in 1959/60 as compared to 1949/50, and increased by 20.6% according to the weighted estimates. And so we come to contradicting conclusions depending on the use of the unweighted or population-weighted measures.

One more piece of evidence is provided by Petrakos & Psycharis (2016). They estimate the evolution of regional inequality in Greece across its NUTS 2 and NUTS 3 regions over 2000–2012, using both the population-weighted and unweighted coefficient of variation. The trend of CV_w is upward, while CV has either a downward trend (in the case of NUTS 3 regions) or is stable (for NUTS 2 regions). Thus, if one considered the weighted estimates, the conclusion would be that regional inequality rises, whereas actually it remains unchanged or even decreases.

There is a prominent example in the international context. Milanovic (2012) estimates the income inequality (measured by the Gini index) between counties and in the world as a whole over the period 1952–2006. In the latter case, he uses the index weighted by the populations of the countries. However, unlike most (if not all) regional studies applying population-weighted measures, he explicitly interprets it as an approximate measure of global inequality (inequality across world individuals) rather than as an estimate of international (cross-country) inequality, realizing that it is not only a rough, but possibly misleading, estimate. The sole reason for the application of the weighted index is that household survey data for a sufficient number of countries are not available for the period prior to 1980s (Milanovic, 2012, p. 8). The trends of the unweighted and weighted Gini indices are found to have opposite directions, upward for the former and downward for the latter (the both become downward only starting with 2000). Thus, if one drew conclusions on dynamics of international inequality from the weighted estimates, they would be quite opposite to the real pattern. With reference to interpersonal inequality, Milanovic (2012, p. 14) also reports estimates of global inequality for 1988–2005 based on household survey data (i.e. taking into account income distributions within countries). These prove to be, first, much higher that the weighted estimates, and, second, sliding upward (although only slightly) rather than downward. Thus, the estimates obtained with the use of the weighted Gini index turn out to be really misleading with respect to inequality both between countries and between world individuals.

3. Population-Weighted Indices as Measures of Interpersonal Inequality

As it has been mentioned in the previous section, an estimate of interpersonal inequality in the country or any subset of regions with the use of the population-weighted indices is biased because the inequality within regions is neglected. An actual 'splitting' regions into their individual inhabitants would yield something like the pattern depicted in Figure 3 (individuals within each region are arranged according to their personal incomes), which fundamentally differs from the pattern supposed in Figure 1.

Individuals in a region may have incomes that are similar to the incomes of the inhabitants of other regions, which implies that individuals from the same region in fact fall into different income classes and individuals from different regions may fall into the same income classes. In other words, regional income distributions overlap with one another. Because of this overlapping the division of the country's population into income classes according to the regional per capita income – as shown in Figure 1 – turns out to be improper, thus resulting in an inadequate estimation of the inequality in the country. In order to correctly estimate the inequality between N persons making up the population of the country, they should all be rearranged by income within the whole country and then grouped (irrespective of their regions of origin) into some actual income classes.

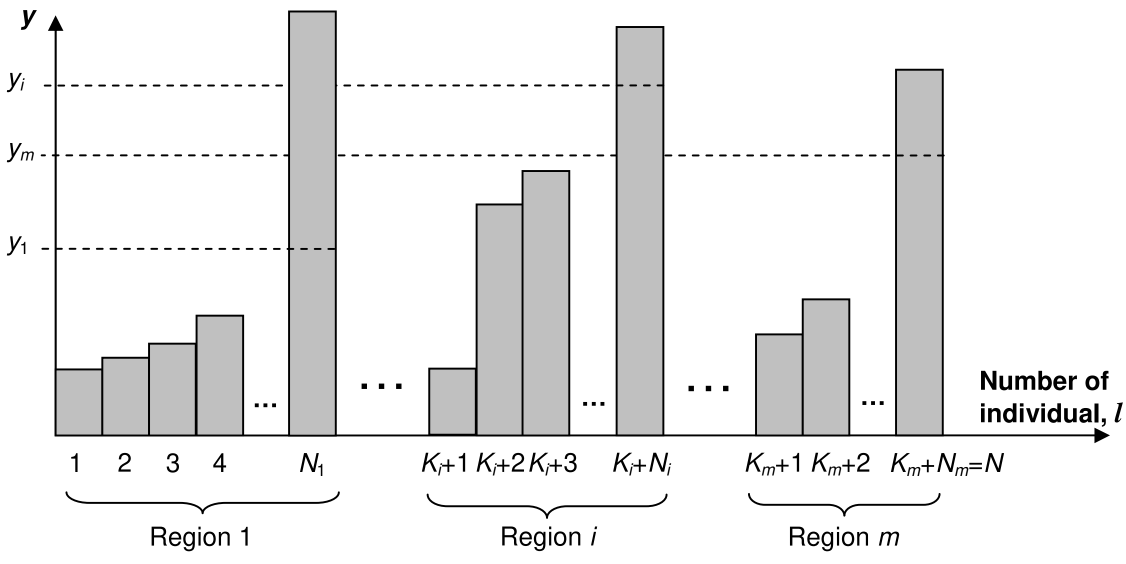


Figure 3. Population of the country as a set of regional populations (actual pattern).

Note: $K_i = \sum_{j=1}^{i-1} N_j$; $K_1 = 0$.

Mussini (2015) estimates inequality between the NUTS 3 regions in the EU-28 over

2003–2011 (applying the population-weighted Gini index) and decomposes its changes into those caused by the population change, re-ranking of regions and growth of regional per capita incomes. In light of the above, the intuitive sense of the first component becomes absolutely obscure. In fact, inequality within the population of the whole geographical entity consisting of the NUTS 3 regions is measured. Imagine that the cross-individual income distribution in this entity remains invariant while the cross-region population distribution changes. Then the effect of population change in the decomposition of inequality change reflects nothing but a result of replacing one improper division of the population into the income classes by another, also improper, one.

Table 4. Income and population in the Arkhangelsk and Tyumen Oblasts in 2014.

i	Region/subregion	Personal income, Russian rubles (RUR) per month [3]			Gini index[4]	Population, thousand people, annual average[5]	Subregion weight (n_i)
		Per capita (y_i)	Median (Md_i)	Modal (Mo_i)			
0	Arkhangelsk Oblast	29,432	23,125	14,276	0.378	1,187.6	
1	Nenets AO	66,491	48,281	25,457	0.429	43.2	0.036
2	Southern part	28,033	22,354	14,213	0.368	1,144.4	0.964
0	Tyumen Oblast	38,523	27,508	14,026	0.439	3,563.8	
1	Khanty-Mansi AO	41,503	30,440	16,375	0.423	1,604.7	0.450
2	Yamalo-Nenets AO	61,252	44,517	23,515	0.429	539.8	0.151
3	Southern part	26,509	20,052	11,473	0.404	1,419.3	0.398

Data drawn from the Russian statistics provide convenient real examples with small numbers of regions that make it possible to judge the extent of the distortions in the estimates of the interpersonal inequality caused by the application of the population-weighted indices. At present, there are two regions in Russia, the Arkhangelsk Oblast and Tyumen Oblast, which include national entities, namely the so-called autonomous *okrugs* (hereafter, AO). The Arkhangelsk Oblast includes the Nenets AO, and the Tyumen Oblast includes Khanty-Mansi

[3] Data source: Mean, median and modal level of population's money incomes in Russia as a whole and subjects of the Russian Federation for 2014. Web site of the Federal State Statistics Service of the Russian Federation. http://www.gks.ru/free_doc/new_site/population/bednost/tabl/1-2-6.doc (Accessed Sept. 15, 2016). [In Russian.]

[4] Data source: Distribution of total money incomes and description of differentiation of population's money incomes in Russia as a whole and subjects of the Russian Federation for 2014. Web site of the Federal State Statistics Service of the Russian Federation. http://www.gks.ru/free_doc/new_site/population/bednost/tabl/1-2-4.doc (Accessed Sept. 15, 2016). [In Russian.]

[5] Data source: *Regions of Russia. Socio-Economic Indicators. 2015*. Moscow, Rosstat, 2015. Pp. 39–40. [In Russian.]

AO and Yamalo-Nenets AO. Statistical data on the personal income distribution and inequality are available for each *oblast* as a whole and all its parts ('subregions'), namely, AO(s) and the *oblast* excluding AO(s); for brevity, the latter will be called the Southern part. Based on such data, we can compare the actual estimates of inequality in the whole region with those obtained with the use of the population-weighted index for these two-subregion and three-subregion cases. Table 4 tabulates the relevant data.

The Russian statistical agency, Rosstat (formerly, Goskomstat), models the income distributions in regions and the whole country as log-normal ones (Goskomstat of Russia, 1996, p. 79). The distribution parameters from Table 4 make it possible to restore the log-normal income distributions for the subregions of the regions under consideration:

$f_i(y) = \frac{1}{y\sigma_i\sqrt{2\pi}}\exp(-\frac{(\log(y)-\mu_i)^2}{2\sigma_i^2})$, where $\mu_i = \log(Md_i)$ and

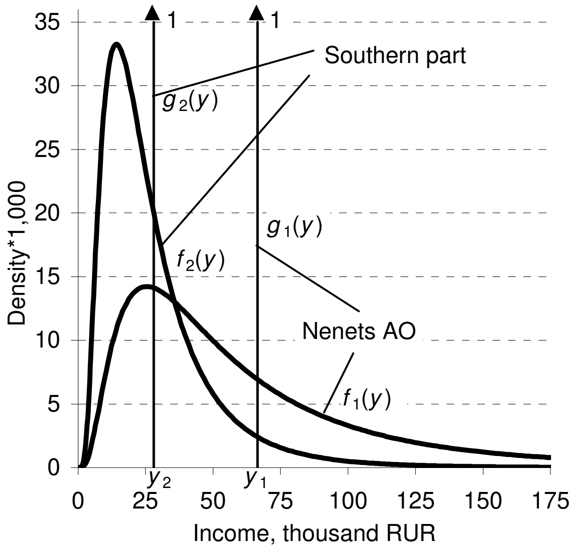
$\sigma_i^2 = \log(Md_i / Mo_i) = 2\log(y_i / Md_i)$. Figure 4, (a) and (c), depicts these distributions.

Estimating inequality in the whole country or – as in our case – a multi-regional entity from per capita incomes only (like the population-weighted indices do), a within-(sub)region income distribution is in fact represented as the delta function $\delta(x)$ which is zero everywhere except at zero and $\delta(0) = \infty$ so that $\int_{-\infty}^{\infty}\delta(x)dx = 1$ (see, e.g., Kanwal, 2004). The delta function can be viewed as a limit: $\delta(x) = \lim_{\sigma\to 0}\frac{1}{\sigma\sqrt{2\pi}}\exp(-x^2/2\sigma^2)$. Denote such distribution in (sub)region i by $g_i(\cdot)$; then $g_i(y) = \delta(y - y_i)$. These distributions are represented in Figure 4 by vertical arrows starting at y_i, a number near the arrowhead specifying the area under the function.

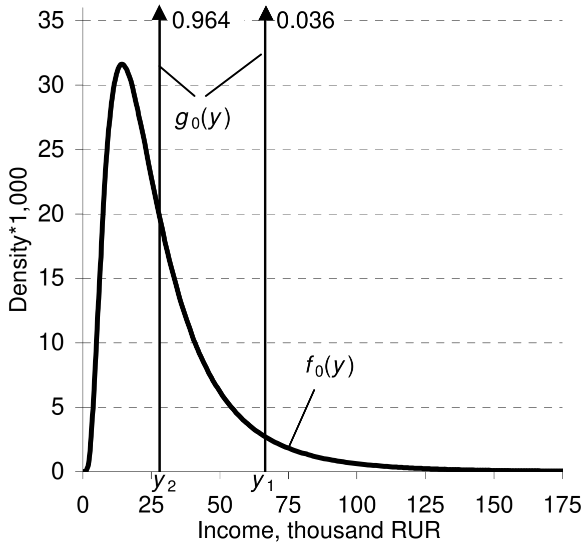
The income distribution in the whole region can be computed either in the same manner as the subregional distributions (from the parameters of the distribution) or equally well as the weighted sum of subregional distributions, $f_0(y) = \sum_{i=1}^{m} n_i f_i(y)$. Similarly,

$$g_0(y) = \sum_{i=1}^{m} n_i g_i(y) = \sum_{i=1}^{m} n_i \delta(y - y_i). \tag{7}$$

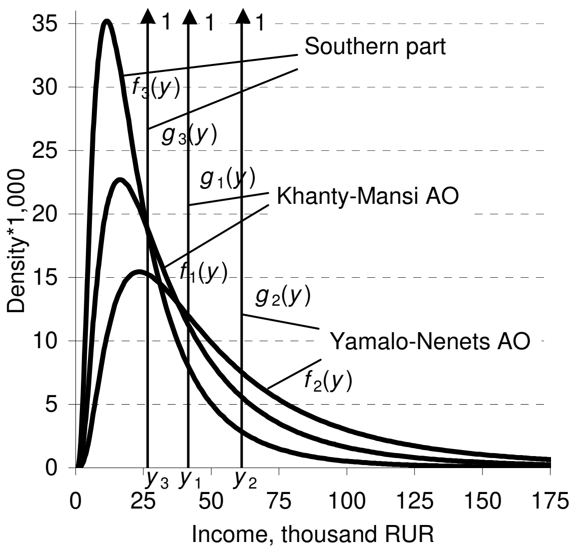
Figure 4 (b) and (d) shows both $f_0(y)$ and $g_0(y)$.



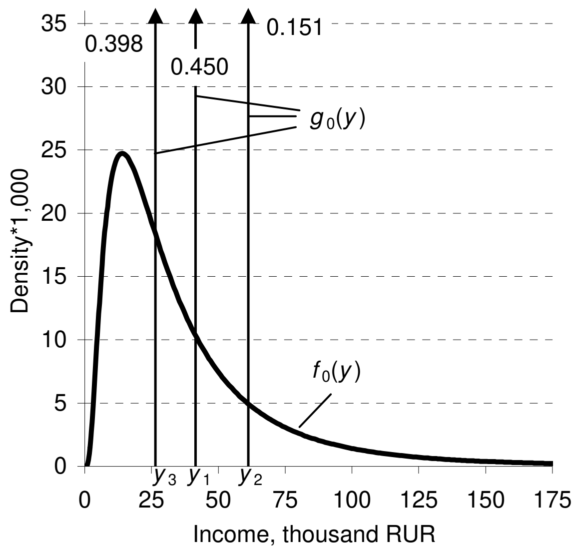
(a) Subregions of the Arkhangelsk Oblast



(b) Arkhangelsk Oblast as a whole



(c) Subregions of the Tyumen Oblast



(d) Tyumen Oblast as a whole

Figure 4. Income distributions in the Arkhangelsk and Tyumen Oblasts.

Given $g_0(y)$, the expectation of y is $E_0(y) = \int_0^\infty y \sum_{i=1}^m n_i \delta(y - y_i) dy = \sum_{i=1}^m n_i y_i = \bar{y}_{(w)}$; the variance is $Var_0(y) = \int_0^\infty (y - E_0(y))^2 \sum_{i=1}^m n_i \delta(y - y_i) dy = \sum_{i=1}^m n_i (y_i - \bar{y}_{(w)})^2$. Then the coefficient of variation coincides with that given by Formula (2), $Var_0(y)^{1/2}/E_0(y) = CV_w$. Computing the Theil index for continuous distribution, we obtain its weighted version,

$\int_0^\infty \frac{y}{E_0(y)} \ln(\frac{y}{E_0(y)}) g_0(y)dy = Th_w$. Expressing the Gini index as $\int_0^\infty F_0(y)(1-F_0(y))dy / E_0(y)$ (Yitzhaki & Schechtman, 2013, pp. 15–16 and 26), where $F_0(y)$ is the cumulative distribution function, $F_0(y) = \int g_0(y)dy$, we arrive at G_w. (The derivation needs cumbersome mathematics and is therefore not reported.) Using $f_0(y)$ instead of $g_0(y)$ in the above calculations, we would obtain the unweighted inequality indices, *CV*, *Th* and *G*, that measure the inequality of the whole population of a region for the case of continuous income distribution. (This is one more proof of the fact that the population-weighted indices measure inequality between all individuals, and not between regions.)

It is obvious – and clearly seen in Figure 4 (b) and (d) – that the approximation of the actual income distribution $f_0(y)$ by the weighted sum of the delta functions, $g_0(y)$, is overly rough and therefore will never yield correct estimates of the population's inequality. In Table 5, the population-weighted estimates, G_w, are compared with the estimates of inequality between the subregions, *G*, and the estimates of the region's population inequality labelled G_{pop}. Because of the small numbers of observations, *G* and G_w are standardized to range from 0 to 1 in order to render them comparable across regions. For the Arkhangelsk Oblast, the normalizing factor equals 2 for *G* and $1/(1 - n_1) = 1/0.964$ for G_w; for the Tyumen Oblast, it is equal to 3/2 and $1/(1 - n_2) = 1/0.849$, respectively (for explanation, see Table 8 in the next section).

Table 5. Estimates of inequality in the Arkhangelsk and Tyumen Oblasts.

Region	Measure	Gini index, raw / standardized	Average income, RUR
Arkhangelsk Oblast	Inter-subregional inequality (*G*)	0.203 / 0.407	$\bar{y}$ = 47,262
	Population-weighted estimate (G_w)	0.046 / 0.048	$\bar{y}_{(w)}$ = 29,432
	Population's inequality (G_{pop})*	0.378	
Tyumen Oblast	Inter-subregional inequality (*G*)	0.179 / 0.269	$\bar{y}$ = 43,088
	Population-weighted estimate (G_w)	0.159 / 0.188	$\bar{y}_{(w)}$ = 38,523
	Population's inequality (G_{pop})*	0.439	

* The official estimate from Table 4.

The case of the Arkhangelsk Oblast resembles the example of China in the previous section. Like in that example, there are two territorial units, one with a large population and a relatively small income per capita, and one with a small population (3.6% of the total) and a high income per capita (about 2.4 times higher than in the first unit). Although the difference between these subregions is not that dramatic as between mainland China and Macao, the estimation results are qualitatively similar. The weighted Gini index suggests low inequality, 4.8% in percentage terms, while the inequality between the Nenets AO and the Southern part of the Arkhangelsk Oblast – measured by the unweighted index – is rather high, 40.7%. The latter reflects the fact that the average inhabitant of the Nenets AO is 2.4 times richer than the average inhabitant of the southern part of the *oblast*. As for the population's inequality, it equals 37.8%, only one percent point higher than the inequality in the Southern part of the Arkhangelsk Oblast (see Table 4). It is the small population of the Nenets AO that is responsible for a minor contribution of this subregion to the inequality in the whole *oblast*. As it is seen in Figure 4, the overall income distribution in the whole Arkhangelsk Oblast, Figure 4 (b), differs from that in its southern part in Figure 4 (a) only slightly. The weighted Gini index – equalling 4.8% – fails to provide a more or less adequate approximation of the population's inequality in the whole *oblast* as well. The weighted index severely understates inequality between both the subregions and the inhabitants of the whole Arkhangelsk Oblast.

The case of the Tyumen Oblast involves three territorial units. They are closer to one another, both in incomes per capita and weights, than in the previous case. A smaller difference in the incomes per capita results in smaller inequality between the subregions (measured by G). The weighted index again is understated as compared to the inequality between the subregions and the inhabitants of the whole Tyumen Oblast.

The patterns provided by G and G_{pop} and the differences between them in the Arkhangelsk and Tyumen Oblasts can be easily explained. The high inequality between subregions of the Arkhangelsk Oblast is due to the great difference in the income per capita between them. Interpersonal inequality is smaller than the regional one, remaining approximately close to the inequality in the southern part of the *oblast*, since adding the rich (on average) but small population of the Nenets AO only slightly changes the income distribution. In the Tyumen Oblast, the inequality between subregions is lower because of the lesser differences in the incomes per capita. At the same time, the inequality of the whole population of the *oblast*, G_{pop}, is higher than the regional one, G, and rises as compared to the inequality in each subregion. The reason is the unification of the poor (on average) population of the southern part of the *oblast* with the richer (on average) population of AOs, the population sizes of the subregions being comparable. With reference to the results suggested by G_w, they hardly can be intuitively explained.

It is worth noting that even the interpretation of a population-weighted inequality index as an approximate measure of interpersonal inequality of the whole country's population is not always true. It holds only regarding indicators which can be applied to an individual, e.g. personal income, wage, housing, education, etc. Otherwise the meaning of the population-weighted index is obscure. Estimating regional income inequality, many authors use the regional GDP per capita in order to characterize incomes in regions. However, there is no inequality in the national GDP (as the total of regional GDPs) per capita between the country's citizens. There are many other indicators that characterize the situation of a region, but cannot be applied to its certain inhabitants, e.g. birth rate, investment per capita, crime rate, etc. In such cases, the population-weighted inequality indices have no intuitive interpretation at all; it is totally incomprehensible what they measure.

For example, Zubarevich & Safronov (2011) estimate, in addition to the income inequality, the regional inequalities in the investment per capita, unemployment rate and poverty rate. Again, there is no, e.g. unemployment inequality between the country's inhabitants; only the national average unemployment rate exists. Consider a simple example. A country consists of two regions. Labouring population numbers 15 million people in the first region and 5 million in the second; the unemployment rates are 40% and 20%, respectively. Then the unemployment rate in the country is 35%. With some trick, we can measure the 'unemployment inequality' of the total labouring population. A person can be either employed, y_i =1, or unemployed, y_i =0. Measuring the unemployment inequality of the given 20 million persons with the Gini index over so quantified $\{y_i\}$, we get G_{nat} = 0.35, representing the exact country-average unemployment rate. It can be easily checked that this is not a coincidence. Provided that the variable is binary, the Gini index always gives the percentage of zeros. At the same time, the population-weighted estimate yields G_w = 0.107; being standardized, it equals 0.107/(1 – 0.25) = 0.143. Both figures are far from the overall inequality G_{nat}. (As well as they are far from the inequality between regions, G_w = 0.167, the standardized value equalling 0.167/0.5 = 0.333.)

4. Some Properties of the Population-Weighted Indices

An adequate inequality index should satisfy a number of axioms, i.e. the desirable properties of an inequality measure (see, e.g. Cowell, 2000). Ezcurra & Rodríguez-Pose (2009, pp. 332–333) argue – with no proof – that a number of the population-weighted inequality indices, including the coefficient of variation and the Gini and Theil indices, fulfil the basic axioms, namely, scale invariance, population principle, anonymity principle and the principle of transfers (the Pigou-Dalton principle). Similar assertions can be found in Kyriacou & Roca-Sagalés (2014, p. 188–189), Lessmann (2014, p. 37), Sacchi & Salotti (2014, p. 148–149) and elsewhere.

Indeed, these indices are scale-invariant; the check is easy and straightforward. As regards the population principle, anonymity (symmetry) principle and principle of transfers, the population-weighted inequality indices violate them (while their unweighted counterparts satisfy them).

The population principle (or replication invariance) states that a simple replication of the sample under consideration should not change the value of the inequality index. Let us replicate the income distribution (y_i), along with the population distribution (n_i), R times, indicating new values of the variables by the superscript (R). The population-weighted coefficient of variation takes the form $CV_w^{(R)} = \sqrt{(CV_w^2 + 1)/R + R - 2}$; it increases with the rising R. The weighted Gini index becomes R times greater: $G_w^{(R)} = RG_w$. The weighted Theil index, contrastingly, diminishes: $Th_w^{(R)} = Th_w - \log(R)$, taking on negative values. (Note that the weighted average also changes because of replication: $\overline{y}_{(w)}^{(R)} = R\overline{y}_{(w)}$).

The violation of the anonymity and transfer principles will be proved below for the population-weighted coefficient of variation. Such proofs for the population-weighted Gini and Theil indices need more cumbersome mathematics; therefore only numerical examples will illustrate the violations of these axioms by them.

Adjusting Jenkins & van Kerm's (2009, p. 52) definition to the case of regions, the anonymity principle requires the inequality index to depend only on the per capita income values used to construct it and not on additional information such as what the region is with a particular per capita income or what the regional populations are. In other words, the index must be invariant to any permutation of the income observations.

Consider a cross-region income distribution $y = (y_1,\dots, y_N)$ and its permutation y_*, i.e. $y = (\dots, y_i, \dots, y_k, \dots)$ and $y_* = (\dots, y_k, \dots, y_i, \dots)$; the other elements in y_* remain the same as in y; hereafter $y_k > y_i$. One can expect the value of the population-weighted inequality index to change under such a transformation if for no other reason than changing the weighted average:

$$\Delta\overline{y}_{(w)} = \overline{y}_{*(w)} - \overline{y}_{(w)} = (n_i - n_k)(y_k - y_i)\,. \tag{8}$$

It is seen that the weighted average remains intact only in the trivial case of $n_i = n_k$.

The change in the population-weighted coefficient of variation is characterized by the following equation:

$$\Delta CV_w^2 = CV_w^2(y_*) - CV_w^2(y) = \frac{\Delta \bar{y}_{(w)}}{(\bar{y}_{(w)} + \Delta \bar{y}_{(w)})^2}(y_i + y_k - (2\bar{y}_{(w)} + \Delta \bar{y}_{(w)})\frac{\overline{y^2}_{(w)}}{\bar{y}_{(w)}^2}), \tag{9}$$

where $\overline{y^2}_{(w)}$ is the weighted average of the squared incomes and $\bar{y}_{(w)}^2$ is the square of the weighted average; $\Delta \bar{y}_{(w)}$ is defined by Formula (8). Note that $\overline{y^2}_{(w)} / \bar{y}_{(w)}^2 = CV_w^2(y) + 1$; hence, it always (given that $y_k \neq y_i$) exceeds unity. Thus, ΔCV_W^2 depends on six variables: y_i, y_k, n_i, n_k, $\bar{y}_{(w)}$, and $\overline{y^2}_{(w)}$. (This number may be reduced by one, replacing the latter two variables with $CV_w(y)$.) The signs of the relationship

$$\frac{y_i + y_k}{2\bar{y}_{(w)} + \Delta \bar{y}_{(w)}} \cdot \frac{\bar{y}_{(w)}^2}{\overline{y^2}_{(w)}} - 1 \equiv H(y_i, y_k, n_i, n_k, \bar{y}_{(w)}, \overline{y^2}_{(w)}) - 1$$

and $\Delta \bar{y}_{(w)}$ determine the sign of ΔCV_W^2, hence the direction of change in the inequality measure: $\mathrm{sgn}(\Delta CV_w^2) = \mathrm{sgn}(H(\cdot) - 1) \cdot \mathrm{sgn}(\Delta \bar{y}_{(w)})$. Table 6 shows different possible cases.

Table 6. Permutation-induced changes in the population-weighted coefficient of variation.

	$n_i > n_k$ ($\Delta \bar{y}_{(w)} > 0$)	$n_i < n_k$ ($\Delta \bar{y}_{(w)} < 0$)
$H(\cdot) > 1$	CV_w increases	CV_w decreases
$H(\cdot) < 1$	CV_w decreases	CV_w increases

Given too many variables in $H(\cdot)$, its behaviour is not amenable to a more or less comprehensive formal analysis. It is possible for some particular cases only. For instance, if both y_i and y_k are less than the weighted average and $n_i > n_k$, then $H(\cdot) < 1$ knowingly holds and CV_w diminishes.

In principle, the case of $\Delta CV_W^2 = 0$ is possible as well. Let all regions except i and k have the same per capita income y_r. Then we can aggregate them into a single 'region' r with income y_r and weight $n_r = 1 - (n_i + n_k)$. (Such a 'region' will be used elsewhere below.) In this instance $H(\cdot) = F(y_i, y_k, n_i, n_k; y_r)$. Keeping all variables except y_r constant, we can find the value of y_r such that $H(y_r) = 1$. Equation $H(y_r) = 1$ is a cubic one with respect to y_r; its closed-form solution is very cumbersome and therefore is not reported. (In fact, we can dispense with it, solving the equation numerically.) This equation may have a real positive root, albeit not always. However, no significance should be attached to this fact. First, probability of finding an actual cross-region income distribution (along with the population distribution) that satisfies $H(\cdot) = 1$ even for some single pair of i and k seems to be close to zero. Second, particular cases of satisfying the anonymity principle do not matter at all, while the only (non-degenerate) case – even a single numerical example – of its violation would evidence that the inequality index under consideration does have this unpleasant property.

Table 7 provides numerical examples that illustrate four cases listed in Table 4 and the case of no change in the population-weighted coefficient of variation. It tabulates three income distributions and their permutations – (A), (B) and (C), the population distribution $n = (n_j)$ being uniform across these. Therefore, $\Delta\bar{y}_{(w)} > 0$ holds for all the three cases of transition from y to y_*. However, we can also consider the reverse transitions from y_* to y, exchanging the indices i and k; in these transitions, $\Delta\bar{y}_{(w)} < 0$. Along with the coefficient of variation, the table reports the population-weighted Gini and Theil indices as well as the unweighted inequality indices.

Table 7. Permutation-induced changes in the population-weighted inequality indices.

Region index	n	(A)		(B)		(C)	
		y	y_*	y	y_*	y	y_*
i	0.15	150	300	150	300	150	300
k	0.05	300	150	300	150	300	150
r	0.80	400	400	100	100	218.9	218.9
$\bar{y}_{(w)}$		357.5	372.5	117.5	132.5	212.62	227.62
CV_w		0.251	0.167	0.387	0.537	0.149	0.149
G_w		0.098	0.062	0.129	0.205	0.059	0.060
Th_w		0.039	0.017	0.057	0.115	0.011	0.011
CV		0.363		0.464		0.275	
G		0.196		0.242		0.149	
Th		0.070		0.104		0.038	

Case (A) is that of the diminishing values of the population-weighted inequality measures caused by the exchange of incomes between two regions; $H(\cdot) < 1$ here. The decrease is fairly sizeable, equalling to more than one third for CV_w and more than a half for Th_w. Considering the reverse transition, we have $\Delta\bar{y}_{(w)} < 0$ and $H(\cdot) > 1$; the permutation of regional incomes causes the weighted inequality indices to rise. In case (B), the effect of the permutation in y is an increase in the weighted indices, as $H(\cdot) > 1$; the reverse permutation has an adverse effect. At last, the weighted coefficient of variation does not change under the permutation in case (C). Interestingly, the weighted Gini and Theil indices are also near-invariant in this case: $\Delta G_w = 3.7 \cdot 10^{-4}$ and $\Delta Th_w = 5.5 \cdot 10^{-4}$. Comparing values of the respective weighted and unweighted indices in Table 3, we can see that the weighting leads to significant undervaluation of inequality, except for $Th_w(y_*)$ and $CV_w(y_*)$ in case (B).

The indices under consideration range from 0 (perfect equality) to some index-specific maximum (perfect inequality). Perfect inequality implies that the income is nonzero in a sole region, say, k. The second column of Table 8 lists the maxima of CV, G and Th. They depend on the number of the country's regions, m, only. The violation of the anonymity principle by the population-weighted inequality indices has a crucial corollary: they have no unambiguous maxima. The value taken on by such an index in the case of perfect inequality depends on which specific region k possesses all the country's income. The relevant maxima are listed in the third column of Table 8.

Table 8. Maxima of the unweighted and weighted inequality indices.

Index	Unweighted	Population-weighted
Coefficient of variation	$\sqrt{m-1}$	$\sqrt{1/n_k - 1}$
Gini index	$(m-1)/m$	$1 - n_k$
Theil index	$\log(m)$	$\log(1/n_k)$

The variability of the upper bounds of the inequality indices matters in at least two cases. First, in order to judge how great the inequality is from an obtained estimate, we should know how far it is from the perfect inequality. Therefore it would be desirable to standardize the inequality indices, i.e., to normalize them to their maxima so that they range from 0 to 1 (the Gini index needs such normalization only if the number of regions is small, when $(m-1)/m$ is not sufficiently close to 1).

Second, the differences in the ranges of the inequality indices render inequalities incomparable across countries. Lessmann (2014, p. 37) notes that the Theil index is not applicable for the cross-country comparison for this reason. However, as it is seen from Table 7, this all the more holds for the coefficient of variation. For example, Williamson's (1965) results are not comparable across countries, as the number of regions varies in his sample from

6 to 75. Thus, the respective maxima of CV differ by the factor of more than 3.8.

The normalization of the inequality indices would solve this problem. However, Theil (1967, p. 92) objects to normalization, giving an example of two situations. The first society consists of two individuals, only one of them having nonzero income; in the second society, all income belongs to the only of two million persons. The second society is evidently much more unequal. Nonetheless, considerations of cross-country comparability and a uniform 'benchmark' of perfect inequality seem more important than Theil's argument (the more so as the number of regions does not differ that dramatically across countries).

In the case of the population-weighted indices, the normalization turns out ambiguous. We could take the 'maximum of maxima', assigning k to the least populated region. (It is such maxima that have been used to standardize the population-weighted indices in Tables 3 and 5.) All the same, this 'global maximum' would depend on the cross-region distribution of the country's population. Then the values of a weighted inequality index are not comparable even between countries with an equal number of regions. Moreover, such 'benchmark' of perfect inequality may vary over time in the same country with varying n_k (or even k, if some other region becomes the least populated one).

Let us turn to the principle of transfers which "is usually taken to be indispensable in most of the inequality literature" (Cowell, 2000, p. 98). Let the cross-region income distribution $y = (\ldots, y_i, \ldots, y_k, \ldots)$ be transformed into $y_* = (\ldots, y_{*i} = y_i + \theta, \ldots, y_{*k} = y_k - \theta, \ldots)$, where $y_{*j} = y_j$ for $j \neq i, k$, and $0 < \theta < \theta_{\max} = (y_k - y_i)/2$, thus keeping region k still richer than i. The principle of transfers requires the inequality index to decrease under such a transformation. This requirement for the weighted coefficient of variation (denoting $CV_{w*} \equiv CV_w(y_*)$) can be represented as

$$\frac{dCV_{w*}}{d\theta} = \frac{1}{\bar{y}_{*(w)} CV_{w*}} \left(\frac{n_i y_{*i} - n_k y_{*k}}{\bar{y}_{*(w)}} - (CV_{w*}^2 + 1)(n_i - n_k) \right) < 0 . \qquad (10)$$

Condition (10) unambiguously holds only if $n_i y_{*i} < n_k y_{*k}$ and $n_i > n_k$, as both summands in the right-hand side of the equation have negative sign. However, as θ rises, y_{*i} and y_{*k} become progressively closer to each other, which inevitably causes $n_i y_{*i} - n_k y_{*k}$ to change its sign to positive. When the signs of summands in the right-hand side of Equation (10) are different (in the case of $n_i < n_k$ they always are), the resulting sign of their sum depends on particular combination of y, n and the value of θ. Then it is not inconceivable that the derivative of CV_{w*} is positive somewhere in the definitional domain of θ, so violating the principle of transfers.

In order to show that $dCV_{w*}/d\theta > 0$ is possible, consider the case when the transfer is close to the right bound of its domain, $\theta \approx (y_k - y_i)/2$. Then $y_{*i} \approx y_{*k} \approx (y_k + y_i)/2$. In this instance, provided that $n_i > n_k$, $dCV_{w*}/d\theta > 0$ if $(y_i + y_k)/2 > (CV_{w*}^2 + 1)\overline{y}_{*(w)}$. Let $y_i = (1+\alpha)\overline{y}_{*(w)}$ and $y_k = (1+\beta)\overline{y}_{*(w)}$ (note that α may be negative), then the latter inequality looks like $(\alpha + \beta)/2 > CV_{w*}^2$. Such a relationship is fairly realistic. Usually $CV_{w*} < 1$, therefore α and β should not be too great. For example, if $CV_{w*} = 0.7$, the principle of transfers will be violated with, say, $y_i = 1.2\overline{y}_{*(w)}$ and $y_k = 1.8\overline{y}_{*(w)}$ in the neighbourhood of $\theta = 0.3\overline{y}_{*(w)}$, or with $y_i = 0.9\overline{y}_{*(w)}$ and $y_k = 2.1\overline{y}_{*(w)}$ near $\theta = 0.6\overline{y}_{*(w)}$. Note that with $n_i > n_k$, a necessary condition for $dCV_{w*}/d\theta > 0$ is the exceedance of the weighted average by y_k, $y_k > \overline{y}_{*(w)} = \overline{y}_{(w)} + (n_i - n_k)\theta > \overline{y}_{(w)}$.

Provided that $n_i < n_k$, $dCV_{w*}/d\theta > 0$ if $(y_i + y_k)/2 < (CV_{w*}^2 + 1)\overline{y}_{*(w)}$. This inequality obviously holds when both y_i and y_k are below the weighted average $\overline{y}_{*(w)}$, or when $\alpha \leq -\beta$. It also may be true if both variables are above $\overline{y}_{*(w)}$, e.g. with $y_i = 1.1\overline{y}_{*(w)}$ and $y_k = 1.8\overline{y}_{*(w)}$ near $\theta = 0.35\overline{y}_{*(w)}$, given that $CV_{w*} = 0.7$.

Considering CV_{w*} as a function of transfer, $CV_w(y_*) = CV_w(\theta)$ (then $CV_w(y) = CV_w(0)$), we can distinguish four types of its behaviour (depending on particular y and n). They are depicted in Figure 5, with $CV_w(\theta)$ normalized to $CV_w(0)$ and θ normalized to $\theta_{\max}$.

Type 1 is a monotonic rise in the weighted coefficient of variation everywhere in the definitional domain of θ. In type 2, $CV_w(\theta)$ decreases at first and then begins to rise (i.e. $dCV_{w*}/d\theta$ changes its sign from negative to positive). Starting with some θ, it reaches the initial value, $CV_w(0)$, and then exceeds it more and more. Type 3 is qualitatively similar to type 2, except for the fact that $CV_w(\theta)$ does not reach the initial value by the end of the domain of θ. At last, type 4 is a monotonic decreasing $CV_w(\theta)$.

The weighted Gini and Theil indices have the same four types of behaviour. A peculiarity of the Gini index is a break on curve $G_w(\theta)$ in some point (instead of a smooth inflection) in the case of the behaviour of types 2 and 3. However, given the same y and n, $G_w(\theta)$ and $Th_w(\theta)$ may differ from $CV_w(\theta)$ in the type of behaviour. For instance, the curves of the weighted Gini index corresponding to curves 1, 2 and 3 in Figure 4 behave according to type 1; the behaviour is similar only in the case of curve 4. The curves of the weighted Theil index corresponding to curves 2, 3 and 4 in Figure 4 have the same type of behaviour, while behaviour of type 2 corresponds to curve 1of CV_w.

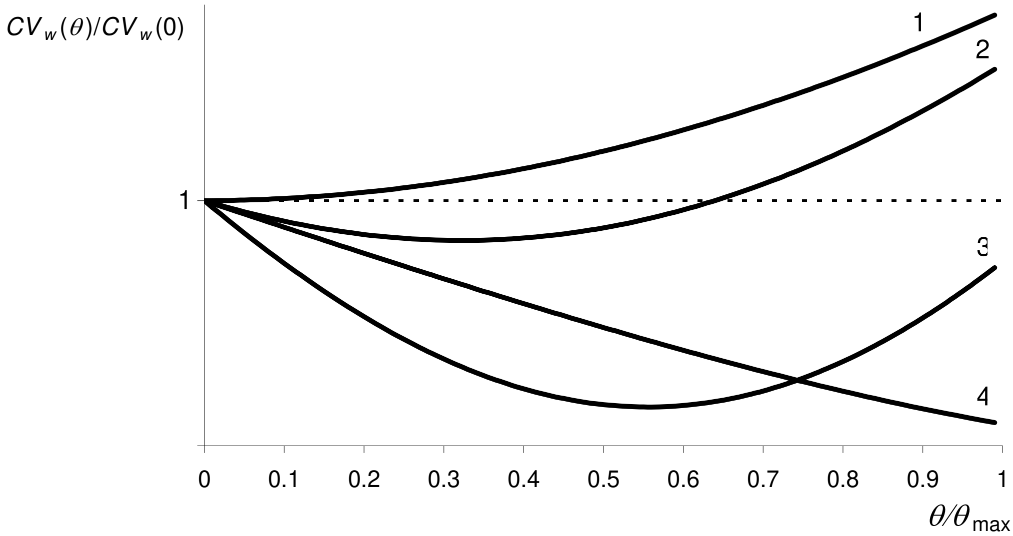


Figure 5. Different types of behaviour of $CV_w(\theta)$.

Note: for all curves, n = (0.15, 0.05, 0,8); for curve 1, $y_{(1)} = (100 + \theta, 300 - \theta, 420)$; for curve 2, $y_{(2)} = (100 + \theta, 300 - \theta, 350)$; for curve 3, $y_{(3)} = (100 + \theta, 300 - \theta, 300)$; for curve 4, $y_{(4)} = (100 + \theta, 300 - \theta, 30)$.

The violations of the principle of transfers have serious implications for empirical studies. Consider the evolution of income inequality in some country (assume that the population distribution remains invariant). Provided that the behaviour of the population-weighted inequality measure entails type 1, we would observe an increasing inequality, while the income gaps between the regions of the country become progressively smaller over time. In the case of behaviour of types 2 and 3, the results will be even more striking and unaccountable. At first, inequality falls with the decreasing income gaps, as could be expected; but then from some point on, a further decrease in the income gaps leads to a rise in inequality.

Certainly, the situation is much more involved in actual empirical studies. For example, the population-weighted inequality measure may have varied types of behaviour for different region pairs (i, k); besides, an increase in the per capita income in the poorer region of a pair is not equal, as a rule, to decrease in the richer region. But the above results indicate that in any case these features of the population-weighted inequality measures will produce (unpredictable) distortions in the pattern of the evolution of inequality (as well as in the perception of the convergence or divergence trends).

Usually (albeit not always), the dynamics of inequality obtained with the use of different unweighted inequality measures, say, the coefficient of variation, Gini and Theil indices, are qualitatively similar, having the same directions of change in inequality and their turning points. Since different population-weighted indices computed on the same data may have different types of behaviour, they can provide quite diverse patterns in the evolution of inequality in a country, depending on a particular index applied.

Table 9 gives numerical examples of violating the transfer principle for cases (A) $n_i < n_k$ and (B) $n_i > n_k$. It tabulates the results for the baseline distribution y and its transformations $y_*(\theta)$ with $\theta = 10$ and $\theta = 90$ ($\theta_{\max} = 100$).

Table 9. Transfer-induced changes in the population-weighted inequality indices.

Region	(A)				(B)			
index	n	y	$y_*(10)$	$y_*(90)$	n	y	$y_*(10)$	$y_*(90)$
i	0.05	100	110	190	0.18	100	110	190
k	0.15	300	290	210	0.02	300	290	210
r	0.80	370	370	370	0.80	110	110	110
$\bar{y}_{(w)}$		346.0	345.0	337.0		112.0	120.0	126.4
CV_w		0.178	0.177	0.196		0.242	0.222	0.260
G_w		0.060	0.062	0.079		0.046	0.031	0.104
Th_w		0.021	0.020	0.022		0.020	0.017	0.030
CV		0.446	0.424	0.314		0.541	0.499	0.254
G		0.234	0.225	0.156		0.261	0.235	0.131
Th		0.114	0.101	0.047		0.136	0.116	0.035

In case (A), the population-weighted coefficient of variation and the Theil index have a behaviour of type 2. Their values decrease with the small transfer $\theta = 10$ and increase with the greater transfer $\theta = 90$. The weighted Gini index behaves according to type 1, its value rising with both transfers. In case (B), all three weighted indices have a behaviour of type 2, falling with $\theta = 10$ and rising with $\theta = 90$. Figure 6 illustrates this case graphically for the whole domain of θ.

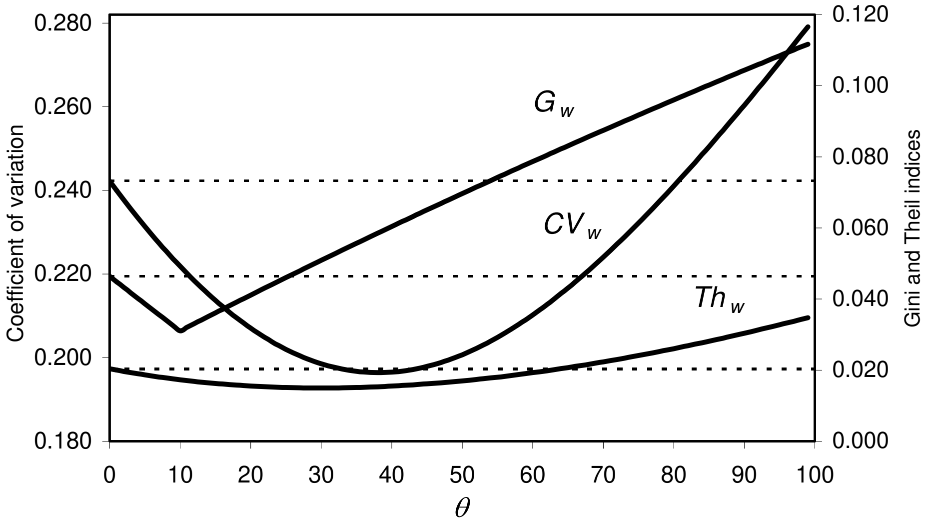


Figure 6. Population-weighted indices as functions of transfer.

Note: the dashed lines correspond to the initial levels (with $\theta = 0$) of the indices.

Transfers apart, let us consider case (B) in Table 6 as a pattern of income evolution over three periods, $t = 0, 1, 2$: $y = y_0$, $y_*(10) = y_1$ and $y_*(10) = y_2$. The cross-region income distribution y_1 is evidently more even than y_0; the poorest and richer regions converge to each other; both weighted and unweghted indices indicate diminishing inequality. Then convergence continues; these regions become further closer to each other in y_2. The unweighted indices show further decrease in inequality. However, the weighted indices rise (becoming even greater than for y_0), indicating divergence. The same takes place in similarly interpreted case (A). Thus, contrary to Ezcurra & Rodríguez-Pose's (2009, p. 332) assertion, it is the weighted inequality measures, and not the unweighted ones, that may lead to unrealistic results, affecting our perception of convergence or divergence.

5. Contras and Pros

Williamson's approach to measuring regional inequality did receive some criticism in the literature. Metwally & Jensen (1973) point out:

> Williamson's coefficient […] fails to take into account either the dispersion of incomes nationally, or what is more important in a spatial context, the dispersion of incomes within regions. […] It is possible for this coefficient to decrease over time, suggesting a convergence in regional mean incomes, while dispersion in actual incomes could show

> an opposite trend. (Metwally & Jensen, 1973, p. 135)

As it is seen, the authors mean measuring national (interpersonal) inequality; therefore their criticism is beside the point. But Williamson (1965) in no way intended to estimate inequality among countries' populations. There is not a grain of evidence of such purpose in his paper; quite the contrary, he highlights throughout the paper that he deals with regional inequality.

Fisch (1984) raises a similar objection:

> Williamson's coefficients of variations ignore a [...] critical issue in relation to spatial inequality: the unequal regional distribution of population by income class. (Fisch, 1984, p. 91)

Again, the case in point is inability of the population-weighted coefficient of variation to adequately approximate interpersonal income inequality in the whole country.

In fact, objections due to Metwally & Jensen (1973) and Fisch (1984) are not those to the population weighting. The essence is in that they believe the national inequality rather than regional one to be more proper for Williamson's (1965) research.

Parr (1974) considers a different aspect; he notes:

> [T]he value of the [Williamson] index is likely to be influenced by the regionalization scheme employed, and there will be a tendency for the value of the index to be high when the regionalization involves a relatively large number of regions. (Parr, 1974, p. 84)

This is so indeed concerning the unweighted coefficient of variation with its maximum rising as the square root of the number of regions, but it is not true for the population-weighted index in the general case (as it has been shown in the previous section). The further Parr's note is connected with the weighted index though:

> [T]here is no way of knowing whether the official statistical regions on which the index is based reflect the extent of spatial income differentiation, given the particular number of regions involved. (Parr, 1974, p. 84)

To manage with this problem, the author suggests a bootstrap procedure of placing a number of points, corresponding to the number of official regions, at random over the territory of the country, thus obtaining a standard of spatial income differentiation against which the original index could be compared. It is not entirely clear what Parr means, but it seems that this procedure would yield something like an approximation of the maximum of $\sqrt{1/n_i - 1}$.

Thus, the above considerations do not concern the main sin of the population-weighted indices, their failure in providing unbiased estimates of regional inequality (as well as their unpleasant properties as inequality measures at all). It is not inconceivable that such criticism exists somewhere in the literature; however, I failed in finding it.

Let us turn to arguments in favour of the weighting inequality indices by population. Almost all of them are based on intuitive considerations, being in fact expanded versions of Williamson's (1965) statement cited in the Introduction. Lessmann's (2014) reasoning is typical for arguments of such kind:

> [The unweighted inequality measures] cannot account for the heterogeneity of regions with respect to (population) size. This is a very important issue […] due to the lack of a uniform territorial classification for all countries […]. In countries with large economic differences and a very unequally distributed population, an unweighted inequality measure might be difficult to interpret. An example should illustrate the problem. The northern Canadian Territories are much poorer than the provinces to the south, so that an inequality measure might indicate large economic differences, although very few people are actually poor (note that the Territories are inhabited by only 100,000 people in total). (Lessmann, 2014, p. 37)

The example in this quotation evidently relates to the inequality of the whole population of Canada, and not to the inequality between regions. Indeed, adding 'very few people' living in the Canadian territories to the large population of the rest of Canada, the overall inequality changes only slightly. But this does not imply that there is no inequality between the 'average' inhabitants of the Territories and the provinces to the south. An analogy may be, e.g. earnings inequality between generals/admirals and other military personnel in the US Armed Forces. The 'per capita' salary of the former is al least three times higher than that of the latter, which implies a rather high inequality. Comparing these 'per capita' salaries, why should we care about the percentage of generals/admirals in the Armed Forces? Provided that this percentage is very small, 0.069% (Kapp, 2016, p. 5), the value of any weighted inequality index will be close to zero, thus suggesting no (significant) inequality between generals/admirals and other servicemen. The example of Canada is worth considering in more detail with the use of actual data reported in Table 10.

Table 10. Personal income and the population in Canada, 2013.

Region	Total income, million CAD[6]	Population, thousand people[7]	Income per capita, CAD	Region weight	Region weigh among provinces
Canada	1,222,216	35,102	34,819		
Provinces	1,217,972	34,986	34,813	0.997	
Newfoundland and Labrador	18,027	528	34,158	0.015	0.015
Prince Edward Island	4,241	145	29,228	0.004	0.004
Nova Scotia	29,378	944	31,125	0.027	0.027
New Brunswick	22,693	756	30,025	0.022	0.022
Quebec	257,579	8,144	31,626	0.232	0.233
Ontario	468,655	13,538	34,618	0.386	0.387
Manitoba	38,445	1,264	30,419	0.036	0.036
Saskatchewan	39,114	1,102	35,487	0.031	0.032
Alberta	181,359	3,979	45,577	0.113	0.114
British Columbia	158,481	4,586	34,556	0.131	0.131
Territories	4,244	115	36,832	0.003	
Northwest Territories	1,816	36	50,160	0.001	
Yukon	1,439	44	32,890	0.001	
Nunavut	989	35	28,042	0.001	

Note: CAD = Canadian dollar

The population of three Canadian territories comprise only 0,33% of the total country's population. Here are the poorest and richest regions of Canada, the difference in incomes per capita between them equalling 79%. In relation to the richest region among the provinces, Alberta, the income per capita there is 63% higher than in Nunavut. Contributing to the total population one order of magnitude smaller than the Nenets AO in the example from Section 3 (see Table 4), all three territories are much less able to change the overall income inequality in the country. This notwithstanding, the average inhabitant of Nunavut remains 1.6 times poorer than the average inhabitant of Alberta. It does not matter a hoot for this fact that they are only 35 thousand in number. Table 11 presents the estimates of the inequality measures in two spatial samples: for all Canadian regions and for the provinces only (i.e., excluding the northern territories). Note that only the standardised values are comparable across the samples, since these differ in the number of regions as well as in the least populated regions. For

[6] Data source: Canada Revenue Agency. Income Statistics 2015 (2013 tax year). Final Table 1. General statement by province and territory of taxation. http://www.cra-arc.gc.ca/gncy/stts/t1fnl/2013/tbl1-eng.pdf (Accessed Nov. 21, 2016). Returns from outside Canada are excluded.

[7] Data source: Statistics Canada. Table 051-0005 – Estimates of population, Canada, provinces and territories, quarterly (persons). CANSIM database, http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=0510005 (Accessed Nov. 21, 2016). Annual average (the arithmetic mean of the quarterly estimates).

instance, the normalising factor for CV_w, $1/\sqrt{1/n_k - 1}$, equals to 0.032 for all Canadian regions and to 0.065 for provinces only.

Table 11. Estimates of income inequality in Canada.

Index	Unweighted		Population-weighted	
	Raw	Standardized	Raw	Standardized
	All regions			
Coefficient of variation	0.180	0.052	0.120	0.004
Gini index	0.089	0.096	0.054	0.054
Theil index	0.015	0.006	0.007	0.001
	Provinces only			
Coefficient of variation	0.133	0.044	0.119	0.008
Gini index	0.065	0.072	0.054	0.054
Theil index	0.008	0.004	0.007	0.001

The unweighted indices indicate a decrease in the inequality between the regions when the territories are deleted from the spatial sample. It is quite understandable, as both the richest and poorest regions are excluded. The weighted Gini (as well as Theil) index remains invariant, as if measuring the overall population's inequality. But it is by no means close to the Canadian Gini index for 2013 equalling 0.358 (varying across provinces from 0.319 to 0.368).[8] Thus the weighted Gini index underestimates both the regional and interpersonal inequalities. The (standardised) weighted coefficient of variation behaves strikingly; it doubles when the northern territories are eliminated, suggesting a rise in inequality. Then what is difficult to interpret, the unweighted coefficient of variation or the weighted one?

Gisbert (2003) provides reasons similar to Lessmann's (2014) to defend the relevance of the weighting by population in the context of constructing a kernel density of the world income distribution. As he points out,

> [Unweighted kernel density] abstracts from the 'size' of the different countries. […] [T]he world income distribution in terms of countries […] can be highly misleading, for example if we drew national borders differently this would affect the shape of the densities […]. The natural alternative is to attach a weight to the observations where the weights reflects the contribution of each observation in the sample. In our example, per capita GDP, the obvious weight is the population (POB) of each country. […]

[8] Data source: Statistics Canada. Table 206-0033 – Gini coefficients of adjusted market, total and after-tax income, Canada and provinces, annual. CANSIM database, http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=2060033 (Accessed Nov. 21, 2016).

[P]opulation is very unevenly distributed among countries; for example China and India, two of the poorest countries, account for more than one third of the total population in the world, on the other side some of the richest countries, like Iceland or Luxembourg, only account for 0.01% of the world population. It does not seem fair to treat all these countries equally in estimation. (Gisbert, 2003, p. 337–338)

This reasoning again relates to the whole population, this time, of the world. Returning to the example of the US Armed Forces, let us draw the 'border' in such a way as to add colonels/navy captains to generals/admirals. Certainly, the 'cross-rank' earnings distribution as well as inequality between this group and the group of other servicemen changes, while the earning distribution and inequality in the whole US Armed Forces remains intact. However, the case at hands is two different phenomena, first, inequality between an ('average') high-rank officer and ('average') serviceman with no high rank and the relevant earnings distribution, and second, inequality in the whole military personnel and cross-person earnings distribution in the US army.

A standard way of constructing continuous distribution from a set of m discrete observations is a kernel density estimator (Silverman, 1986). It is defined as (in notation of this paper):

$$f(y) = \frac{1}{m}\sum\nolimits_{i=1}^{m} K(\frac{y - y_i}{h}) / h, \tag{11}$$

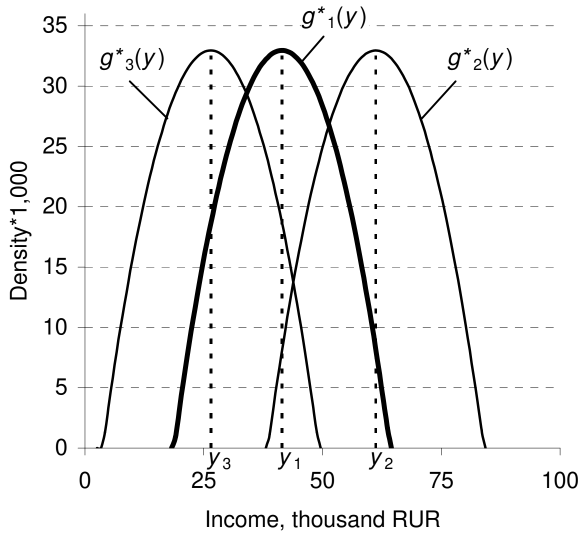
where h is the smoothing bandwidth which depends on the number of observations, m, as well as on parameters of the source distribution (y_i); $K(\cdot)$ is a kernel function. Considering regions instead of countries, $f(y)$ is an estimate of the cross-region income distribution (to be exact, the probability density). Based on his reasoning, Gisbert (2003) modifies Formula (11) in the following way:

$$g_0^*(y) = \sum\nolimits_{i=1}^{m} n_i K(\frac{y - y_i}{h}) / h \equiv \sum\nolimits_{i=1}^{m} n_i g_i^*(y). \tag{12}$$

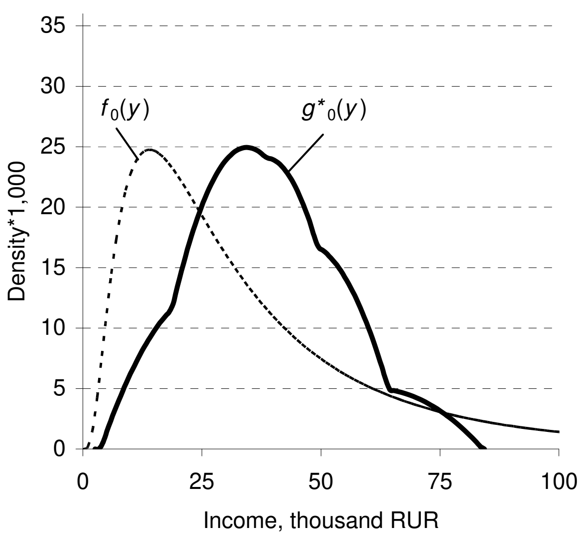
This formula resembles Formula (7) in Section 3. The similarity is not formal; the essence of Formulae (7) and (12) is the same. Both approximate cross-person income distributions in the whole territory that consists of m territorial units. The difference is in that the delta function representing the within-region income distribution $g_i(y)$ in Formula (7) is replaced by an arbitrary – with respect to the actual within-region distribution – function $g_i^*(y)$ in Formula (12). A number of functions can serve as the kernel in Formulae (11) and

(12). To be specific, employ the Epanechnikov kernel:

$K(\frac{y-y_i}{h}) = K_i^E(y) = 0.75(1-\frac{(y-y_i)^2}{h^2})$ if $y \in [y_i + h, y_i - h]$, otherwise $K_i^E(y) = 0$; $h = 0.9(4\pi)^{0.1}(15/m)^{-0.2}\sigma$, where σ = standard deviation of $\{y_i\}$. Then $g_i^*(y) = K_i^E(y)/h$. Figure 7 shows such proxies of the regional income distributions (in the left panel) and income distribution of the whole population (in the right panel) as applied to the example of the Tyumen Oblast from Section 3.



(a) Subregions of the Tyumen Oblast



(b) The Tyumen Oblast as a whole

Figure 7. Proxies of income distributions in the Tyumen Oblasts.
Note: $f_0(y)$ is drawn from Figure 4 (d).

In this example, $h = 22.756$. It artificially 'imputes' the income dispersion to regions; as a result, $CV(g_0^*(y))$, $G(g_0^*(y))$ and $Th(g_0^*(y))$ are not equal to CV_w, G_w and Th_w, respectively. Nonetheless, the former also do not provide adequate estimates of population's income inequality. Comparing Figure 7 (a) with Figure 4 (c), it is seen that the 'imputed' regional distributions $g_i^*(y)$ are far from being similar to the actual distributions $f_i(y)$. Owing to this, their weighted sum (weighted kernel density estimate) $g_0^*(y)$ is a severely distorted proxy of the actual population income distribution $f_0(y)$ as Figure 7 (b) evidences.

Petrakos et al. (2005, p. 1839–1840) derive the need for the population weighting from a critique of the β-convergence methodology. According to them, the analysis of β-convergence can distort the perception of the convergence trends, since it neglects relative sizes of regions. In order to illustrate this statement, the authors offer a simple three-region example. Table 12 tabulates this example (Petrakos et al., 2005, p. 1840), supplementing it with estimates of different inequality measures, both unweighted and population-weighted. Among them, σ stands for the standard deviation of log income and σ_w is its population-weighted counterpart.

Table 12. Inequality estimates in Petrakos' et al. (2005) example.

Region	Population	n	$y(t)$	$y(t+\tau)$, scenario 1	$y(t+\tau)$, scenario 2
A	4.0	0.714	20	25	25
B	1.5	0.268	14	15	15
C	0.1	0.018	6	7	8
$\bar{y}_{(w)}$			18.143	22.000	22.018
CV / CV_w			0.430 / 0.172	0.470 / 0.221	0.436 / 0.218
G / G_w			0.233 / 0.075	0.255 / 0.099	0.236 / 0.098
σ / σ_w			0.505 / 0.215	0.523 / 0.271	0.466 / 0.261
Th / Th_w			0.100 / 0.017	0.115 / 0.027	0.097 / 0.026

The initial state, $y(t)$, is compared with a final state, $y(t+\tau)$, under two scenarios. Regarding scenario 1, both β-convergence analysis and all inequality indices unambiguously indicate income divergence. However, β-convergence occurs under scenario 2, while CV_w suggests divergence. Petrakos et al. (2005) assign this to the fact that fast growth of small region C (by 33%) blurs the picture when all regions are treated as equal, whereas CV_w accounts properly for the relative importance of region C and therefore adequately indicates divergence. However, the unweighted indices CV and G also indicate divergence under scenario 2. At the same time, σ and Th suggest convergence. Hence the weighting is not the case; the point is that specific inequality measures differ in sensitivity to changes in income distribution (Lambert, 2001). As for β-convergence, it results from diminishing σ under scenario 2. Wodon & Yitzhaki (2006) prove that from σ-convergence follows β-convergence (but the converse is not true: β-convergence does not necessary implies σ-convergence). All weighted inequality indices, indeed, indicate divergence under scenario 2. However, this is a particular case. For example, if population of region C were 1 instead 0.1, σ_w would suggest convergence, being equal to 0.451 in the initial state and 0.439 under scenario 2.

A sole attempt to justify the need for weighting by population on the basis of quantitative analysis is performed by Portnov & Felsenstein (2010). They explore the sensitivity of four unweighted and four population-weighted inequality measures to changes in the ranking, size and number of regions into which a country is divided, explicitly treating the regions as groups of people. One of their tests consists in comparison between two situations that differ in the cross-region population distribution and national per capita income, keeping the cross-region income distribution invariant. Surprisingly, the values of the unweighted indices change across the situations, although they should not, being independent of the population distribution. A closer look shows that this is due to the mistaken use of $\overline{y}_{(w)}$ instead of $\overline{y}$ in calculation of these indices. In one more test, the population distribution randomly changes, the cross-region income distribution and national per capita income being kept constant. As one would expect, the weighted inequality indices react to these changes, while the unweighted ones remain constant. The authors believe the latter to be a shortcoming. They conclude:

> These [unweighted] indices may thus lead to spurious results when used for small countries, which are often characterized by rapid changes in population patterns. (Portnov & Felsenstein, 2010, p. 217)

They also conclude that the population-weighted indices – the Williamson coefficient of variation, Gini index and Coulter coefficient – may be considered as more or less reliable regional inequality measures (Portnov & Felsenstein, 2010, pp. 217–218). Both conclusions are fallacious. Explicitly treating regions as groups of people, the authors implicitly deal with the estimation of interpersonal inequality in the country, misinterpreting it as the estimation of regional inequality. Therefore, their results in no way can be deemed as a proof of the use of weighting.

The above discussion shows that the supporters of weighting by population confuse the inequality between regions (i.e., between the representative inhabitants of the regions and the overall interpersonal inequality. According to them, the population weights should reflect the contribution of each territorial unit. But, contribution to what? They interpret it as a contribution to the inequality between regions, while in fact it is a contribution to the inequality between all inhabitants of the set of territorial units under consideration.

Studies on international inequality also widely use the population-weighted indices. From all appearances, economists engaged in studies of international inequality 'reinvented' Williamson's approach. In contrast to regional researchers (who sometimes perform international studies as well), they are aware of the conceptual distinction between the unweighted and population-weighted inequality indices, explicitly interpreting the latter as approximate measures of the inequality among the world population, and not between nations. A surprising thing is that there appears to be a barrier between the literature on regional inequality and that on international inequality. The former almost never references the latter (Akita et al., 2011, can be mentioned as one of the extremely rare examples). The conversance with the literature on international inequality would surely prevent regional researchers from misinterpreting the population-weighted indices as measures of regional inequality.

While the literature on regional inequality does not discuss need for the population weighting in inequality indices, getting by short notes like those cited throughout this paper, the literature on international inequality widely debates the question 'To weight or not to weight?'. Both viewpoints are considered in detail by e.g. Firebaugh (2003) and Ravallion (2005). Under the interpretation of the population-weighted estimates as proxies of inequality among the world population, the arguments in favour of weighting look reasonable; at least, they are seriously substantiated.

However, the results of applying the population-weighted indices to the estimation of global inequality are disappointing as, e.g., the findings of Milanovic (2012) cited in the end of Section 2 suggest. This is of no surprise in light of the above exposition. As Milanovic (2005, p. 10) notices, population-weighted inequality "deals neither only with nations nor individuals but falls somewhere in between" (in fact, this is not always true; it may fall below the both as the examples of Russia in Table 5 and Canada in Table 11 evidence). He also accepts that it may be misleading (Milanovic, 2012, p. 8). Worse yet, this is the prevailing situation as it is proved in Section 3: the estimates of interpersonal inequality with the use of population weighting are always severely distorted.

The debate regarding the population weighting in the literature on international inequality focuses on the issue of what an adequate characterization of inequality in the world is, either inter-country inequality or interpersonal inequality among the world population. In my view, this debate is fairly pointless. It must be agreed with Firebaugh (2003), who notes that the answer depends on the goal:

> [T]he issue of unweighted versus weighted between-nation inequality reduces to this question: are we interested in between-nation income inequality because of what it tells us about the average difference between nations' income ratios, or because of what it

tells us about the average difference between individuals' income ratios? (Firebaugh, 2003, p. 129)

At last, one more issue needs to be touched upon. Exploring the determinants of regional inequality with the use of population-weighted inequality indices, some authors, e.g. Kyriacou & Roca-Sagalés (2014) and Lessmann (2014), also employ unweighted indices for robustness checks. Such a way seems contradictory. On the one hand, if the authors believe unweighted measures to distort the perception of inequality, then why should these measures confirm the results obtained with the use of 'adequate' measures? On the other hand, if they do confirm, then why do we need the weighting?

6. Conclusions

Following Williamson (1965), many economists estimate regional inequality with the use of the indices weighted by the regions' proportions of the national population. The analysis in this paper shows that this approach is conceptually inconsistent. Instead of an estimate of regional inequality, we get a rough estimate of interpersonal inequality among the whole population of the country (and this estimate makes sense only if it deals with indicators applicable to an individual). Therefore the population-weighted estimates of inequality are biased with respect to the estimates of both regional inequality (as they measure a different value) and interpersonal inequality (as they do not and cannot take account of the within-region income disparities). In both cases, the result may be not only distorted, but also quite misleading. Hence the population-weighted inequality indices **never** give adequate results.

Moreover, the population-weighted inequality indices do not satisfy the requirements for an adequate inequality measure. They violate three of the four basic axioms, namely the population, anonymity and transfer principles. This may lead to estimates of inequality evolution that contradict common sense. One more consequence is the absence of unambiguous maxima of the population-weighted inequality indices. This makes it impossible to standardize the estimates of inequality with the aim of cross-time or cross-country comparability.

Thus, it can be concluded that the application of the population-weighted indices to measuring regional inequality is nothing but a fallacy.

References

Akita, T., Kurniawan, P. A. & Miyata, S. (2011) Structural changes and regional income inequality in Indonesia: a bidimensional decomposition analysis, *Asian Economic Journal* 25, 55–77.

Cowell, F. A. (2000) Measurement of inequality, in: A.B. Atkinson & F. Bourguignon (eds), *Handbook of Income Distribution*, vol. 1, pp. 86–166, Amsterdam, North Holland.

Deiwiks, C., Cederman, L. & Gleditsch, K. S. (2012) Inequality and conflict in federations, *Journal of Peace Research*, 49, 289–304.

Doran, J. & Jordan, D. (2013) Decomposing European NUTS2 regional inequality from 1980 to 2009: national and European policy implications, *Journal of Economic Studies*, 40, 22–38.

Enflo, K. & Rosés, J. R. (2015) Coping with regional inequality in Sweden: structural change, migrations, and policy, 1860–2000, *Economic History Review* 68, 191–217.

Ezcurra, R. & Rodríguez-Pose, A. (2009) Measuring the regional divide, in: R. Capello & P. Nijkamp (eds) *Handbook of Regional Growth and Development Theories*, pp. 329–353, Cheltenham, Edward Elgar.

Ezcurra, R. & Rodríguez-Pose, A. (2014) Trade openness and spatial inequality in emerging countries, *Spatial Economic Analysis*, 9, 162–201.

Fisch, O. (1984) Regional income inequality and economic development, *Regional Science and Urban Economics*, 14, 89–111.

Firebaugh, G. (2003) *The New Geography of Global Income Inequality*, Cambridge, MA, Harvard University Press.

Friedman, M. (1992) Do old fallacies ever die? *Journal of Economic Literature*, 30, 2129–2132.

Gisbert F.J.G. (2003) Weighted samples, kernel density estimators and convergence, *Empirical Economics*, 28, 335–351.

Goskomstat of Russia (1996) *Methodological Regulations on Statistics*, Moscow. [In Russian.]

Jenkins, S. P. & van Kerm, P. (2009) The measurement of economic inequality, in: V. Salvedra, B. Nolan & T.M. Smeeding (eds), *The Oxford Handbook of Economic Inequality*, pp. 40–67, Oxford, Oxford University Press.

Kanwal**,** R.P. (2004) *Generalized Functions: Theory and applications*, Boston, Birkhäuser.

Kapp, L. (2016) General and flag officers in the U.S. Armed Forces: background and considerations for Congress. Congressional Research Service Report R44389, February 18, 2016. http://www.fas.org/sgp/crs/natsec/R44389.pdf

Kyriacou, A. P. & Roca-Sagalés O. (2014) Regional disparities and government quality: redistributive conflict crowds out good government, *Spatial Economic Analysis*, 9, 183–201.

Lambert, P. (2001) *The Distribution and Redistribution of Income*, Manchester, Manchester University Press.

Lessmann, C. (2014) Spatial inequality and development – Is there an inverted-U relationship? *Journal of Development Economics*, 106, 35–51.

Li, C. & Gibson, J. (2013) Rising regional inequality in China: fact or artifact? *World Development*, 2013, 47, 16–29.

Martínez-Galarraga, J., Rosés, J. R. & Tirado, D. A. (2015) The long-term patterns of regional income inequality in Spain, 1860–2000, *Regional Studies*, 49, 502–517.

Metwally, M. M. & Jensen R. C. (1973) A note on the measurement of regional income dispersion, *Economic Development and Cultural Change*, 22, 135–136.

Milanovic, B. (2005) *Worlds apart: measuring international and global inequality*, Princeton, MA, Princeton University Press.
Milanovic, B. (2012) Global inequality recalculated and updated: the effect of new PPP estimates on global inequality and 2005 estimates, *Journal of Economic Inequality*, 10, 1–18.
Mussini, M. (2015) Decomposing changes in inequality and welfare between EU regions: the roles of population change, re-ranking and income growth, *Social Indicators Research*, DOI 10.1007/s11205-015-1184-2
Parr, J. B. (1974) Welfare differences within a nation: a comment, *Papers of the Regional Science Association*, 32, 83–91.
Petrakos, G. & Psycharis, Y. (2016) The spatial aspects of economic crisis in Greece, *Cambridge Journal of Regions, Economy and Society*, 9,137–152.
Petrakos, G., Rodríguez-Pose, A. & Rovolis, A. (2005) Growth, integration, and regional disparities in the European Union, *Environment and Planning A*, 37,1837–1855.
Portnov, B.A. & Felsenstein D. (2010) On the suitability of income inequality measures for regional analysis: some evidence from simulation analysis and bootstrapping tests, *Socio-Economic Planning Sciences*, 44, 212–219.
Quah, D. (1993) Galton's fallacy and tests of the convergence hypothesis, *Scandinavian Journal of Economics*, 95, 427–443.
Ravallion, M. (2005) Competing concepts of inequality in the globalization debate, in: S.M .Collins & C.L. Graham (eds), *Brookings Trade Forum: 2004: Globalization, Poverty, and Inequality*, pp. 1–38, Washington, D.C., Brookings Institution Press.
Sacchi, A. & Salotti, S. (2014) How regional inequality affects fiscal decentralisation: accounting for the autonomy of subcentral governments, *Environment and Planning C: Government and Policy* 32, 144–162.
Shorrocks, A. & Wan, G. (2005) Spatial decomposition of inequality, *Journal of Economic Geography*, 2005, 5, 59–81.
Silverman, B.W. (1986) *Density estimation for statistics and data analysis*, London, Chapman & Hall.
Theil, H. (1967) *Economics and Information Theory*, Amsterdam, North Holland.
Wijerathna, D., Bandara, J. S., Smith, C. & Naranpanawa, A. (2014) Regional disparities in Sri Lanka: an empirical analysis, *Asia-Pacific Development Journal* 21, 77–102.
Williamson, J. G. (1965) Regional inequality and the process of national development: a description of patterns, *Economic Development and Cultural Change*, 13, 1–84.
Wodon, Q. & Yitzhaki, S. (2006) Convergence forward and backward? *Economics Letters*, 92, 47–51.
Yitzhaki, S. and Schechtman, E. (2013) *The Gini Methodology. A Primer on a Statistical Methodology*, New York, Springer.
Zubarevich, N.V. & Safronov, S.G. (2011) Regional inequality in large post-Soviet countries, *Regional Research of Russia*, 1, 15–26.