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

Gluschenko, Konstantin

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

20 November 2015

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

Konstantin Gluschenko

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

pr. Lavrentieva 17, 630090 Novosibirsk, Russia.
E-mail address: glu@nsu.ru.

Abstract
Following Williamson (1965), many economists estimate inter-regional inequality with the use of indexes weighted by regions’ shares in the national population. Despite that widespread, this approach is conceptually inconsistent, yielding an estimate of interpersonal inequality among the whole population of the country rather than an estimate of inter-regional inequality. On the other hand, if one considers such an estimate as a proxy of population inequality, it is very rough and for the most part misleading.

Keywords
Inequality index   Weighting by population   Williamson coefficient of variation   Gini coefficient

JEL classifications: D31, D63, R10.

INTRODUCTION

Apparently, WILLIAMSON (1965) was the first who put forward the idea of weighting inequality indexes by regions' shares in the national population in estimation of inter-regional inequality. He was not concerned with a probative justification of his idea, merely noted that an unweighted inequality index 'will be determined in part by the somewhat arbitrary political definition of regional units' and 'The preference for an unweighted index over a weighted one, we think, is indefensible' (WILLIAMSON, 1965, pp. 11 and 34).

Since then such an approach became fairly widespread in regional studies. It hardly makes sense to cite them: their name is Legion. It seems that more often than not the authors of these studies apply Williamson's idea mechanically, not considering the use of weighting and its sense. Yet even the brief notes cited above are open to question.

First, the political division of a country is the reality which regional researchers should deal with, 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 inter-regional inequality. Therefore a desire for 'adjustment' of existing political division through weighting regional disparities seems strange.

Second, why does one need taking into account differences in regional population at all? But we can estimate inequality among representatives of population groups of different sizes without regard for these sizes. For instance, while estimating wage inequality between industrial workers, builders, teachers, lawyers, and so on, we do not care 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 inhabitants of one region?

Third, on closer inspection results of estimating inequality with the use of a weighted index look striking. Consider two Chinese regions, mainland China as a whole and Macao, the Special Administrative Region of the People's Republic of China. In hoary antiquity, when the Portuguese occupied as large part of Chinese territory as they could (or needed), Macao might be deemed 'somewhat arbitrary' regional unit. Nowadays, it is quite natural, as Macao has its own currency, and citizens of China from other regions need visa to get there. The PPP-adjusted GDP per capita (in current international dollars) for 2014 comprise 13,217 in mainland China and 139,767 in Macao, the richest territory in the world (WORLD BANK, 2015). The Gini coefficient for them equals 0.414. (Note that this figure should be doubled, as perfect inequality gives the value of 0.5 in the case of two observations.) The population of mainland China is 1,376.049 million people, and that of

Macao is 0.588 million (UNITED NATIONS, 2015, p. 13). Thus, the latter forms 0.04 percent of the total. The population-weighted Gini coefficient takes on value 0.004, suggesting that there is (almost) no income inequality between the average mainland Chinese and average inhabitant of Macao.

This result evidently contradicts common sense. However, a sufficiently great number of regions in empirical studies masks such absurdities, creating an impression that estimates of inequality with the use of weighting are reasonable. The purpose of this note is to dispel this fallacy, demonstrating that population-weighted inequality indexes measure not inter-regional inequality but something else and therefore yield distorted estimates of inter-regional inequality.

WHAT DO POPULATION-WEIGHTED INDEXES MEASURE?

Let y_i denote per capita income in region i ($i = 1, \ldots, m$), N_i its population, N the population of the country, $n_i = N_i/N$ region's share in the national population (region's weight), $\bar{y}$ the arithmetic average of regional per capita incomes ($\bar{y} = (y_1 + \ldots + y_m)/m$), and $\bar{y}_{(w)}$ their weighted (arithmetic) average ($\bar{y}_{(w)} = n_1 m_1 + \ldots + n_m y_m$) which exactly equals the national per capita income. With this notation, the Williamson coefficient of variation (WILLIAMSON, 1965, p. 11) – sometimes called Williamson index – looks like

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

To understand what it measures, let us estimate inequality among all citizens of the country, basing on cross-region income distribution. An inequality index to be used is the coefficient of variation with y_l denoting personal income of l-th citizen of the country:

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

Obviously, the population average income in this formula – national per capita income – equals the weighted average of regional per capita incomes $\bar{y}_{(w)}$. Lacking information on intra-regional income distributions, one is forced to assume that all inhabitants of a region have the same income equalling per capita income in this region. Then square deviations $(y_l - \bar{y}_{(w)})^2$ are uniform for all l relating to inhabitants of the same region, say i; hence, their sum over all inhabitants of the region is

$(y_i - \bar{y}_{(w)})^2 N_i$. Summing up such sums over all regions, the Williamson coefficient of variation is arrived at:

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

Thus, the population-weighted coefficient of variation is not a measure of inter-regional inequality; instead, it measures 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 such as the Gini coefficient, Theil index, etc. (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).

It is seen that there is a conceptual difference between an unweighted and weighted estimates of inequality. In the former case, all regions 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, representative (or ‘average’, i.e. having the region-average income) inhabitants of each region. While estimating inter-regional inequality, we compare their incomes without regard for how many people live in respective regions (like we do when compare 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 that the population of Macao is 2,340 times smaller than the population of mainland China.

Introducing regional weights means that a region is presented by all its inhabitants rather than by one representative inhabitant. That is, region i is considered as a group of N_i people, each having income y_i. Then the probability of y_i differs across regions, becoming proportional to their populations, n_i. The set $\{n_i\}$ is in fact a proxy of the personal-income distribution in the country. In other words, it is 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 is inter-regional inequality is but an illusion owing to that the grouping proceeds from the data by region. Actually one gets an estimate of interpersonal inequality in the country. As such it is very crude, since the cross-region-data-based income classes in fact overlap (not infrequently, heavily) because of uneven intra-regional income distributions.

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

The bias can have either direction depending on a particular combination of regional per capita incomes and populations. WILLIAMSON (1965, p. 12) reports values of both weighted and unweighted coefficients of variation estimated on regional data from 24 countries. Inter-regional inequality proves to be overstated in about a half of 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 a real study (covering 18 regions): the weighted index understates the extent of inter-regional inequality there by more than a half.

One more evidence is due to MILANOVIC (2012), who estimates income inequality (measured by the Gini coefficient) between counties and in the world as a whole over 1952–2006. In the latter case, he uses the index weighted by populations of the countries. However, he considers it as an approximate measure of global inequality (inequality across world individuals) rather than an estimate of international (cross-country) inequality, realizing that it is not only a rough, but possibly misleading, estimate. The only reason for 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 coefficients are found to have opposite directions, upward for the former and downward for the latter (the both become downward only since 2000). Thus, if one drew conclusions on dynamics of international inequality from the weighted estimates, they would be quite opposite to the real pattern. (Interestingly, the weighted estimates without China produce a trend similar to that for the unweigted estimates, albeit overstating international inequality by crudely some 10%.) 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, estimates obtained with the use of the weighted Gini coefficient turn out really misleading.

While the Gini coefficient, with its maximum equalling $(m-1)/m$, has the upper bound, some other inequality indexes have not. For example, the maximum is $\sqrt{m-1}$ for the coefficient of variation, and $\log(m)$ for the Theil index. To judge how great inequality is from an estimate obtained, one should know how far it is from perfect inequality. Therefore it would be desirable to normalize inequality indexes to their maxima (the Gini coefficient needs such normalization only in cases of very small number of regions, say, less than ten). In fact, WILLIAMSON'S (1965) results are not comparable across countries, as the number of regions varies in his samples from 6 to 75; thus, the perfect-inequality values differ between these extreme cases by the factor of more than 3.8. If the Theil index were applied, this ratio would equal 2.4. 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 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).

Turning to the weighted inequality indexes, we observe a striking and unpleasant feature: they have no unambiguous maxima. Now the value taken on by an index in the case of perfect inequality depends on which particular region possesses all country's income (or to which region the only person having nonzero income is placed). Denote such region by k. Then the maximum is $1 - n_k$ for the weighted Gini coefficient, $\sqrt{1/n_k - 1}$ for the Williamson coefficient of variation, and $\log(1/n_k)$ for the weighted Theil index. One could take the 'maximum of maxima', assigning the least populated region to be k. All the same, this 'global maximum' would depend on the cross-region distribution of 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 can vary over time in the same country.

The reason is the fact that the population-weighted inequality indexes do not satisfy the anonymity (symmetry) principle which requires the index to be invariant with respect to any permutation of observations. In other words, the index should depend only on the income values and not additional information such as what the region is with a particular per capita income (see, e.g. JENKINS and VAN KERM, 2009). Then an exchange of incomes between any two regions does not alter the value of the index. Obviously, this is not the case for the population-weighted indexes.

CONTRAS AND PROS

Williamson's approach to measuring inter-regional inequality was both criticized and advocated. METWALLY and JENSEN (1973, p. 134) point that the weighted coefficient of variation

> 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.

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 estimating 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 the case at hand is (inter-)regional inequality. FISCH (1984) raises the same objection: 'Williamson's coefficients of variations ignore a […] critical issue in relation to spatial inequality: the unequal regional distribution of population by income class'.

PARR (1974, p. 84) notes that

> the 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.

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. The further note (*ibid.*) is connected with the weighted index though: 'there 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'. 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 is meant, 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 indexes, their failure in providing unbiased estimates of inter-regional inequality. It is not inconceivable that such criticism exists somewhere in the literature; however, the search of it has not met with success.

Let us turn to arguments in favour of weighting inequality indexes by population. PORTNOV and FELSENSTEIN (2010) 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 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 indexes 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 $\bar{y}_{(w)}$ instead of $\bar{y}$ in calculation of these indexes. 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 indexes react to these changes, while the unweighted ones remain constant. The authors believe the latter to be a shortcoming. 'These [unweighted] indices may thus lead to spurious results when used for small countries, which are often characterized by rapid changes in population patterns', conclude PORTNOV and FELSENSTEIN (2010, p. 217). They also conclude that the population-weighted indexes – the Williamson coefficient of variation, weighted Gini coefficient, Coulter coefficient – may be considered as more or less reliable regional inequality measures (*ibid*., 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 inter-regional inequality. Therefore, their results in no way can be deemed a proof of the use of weighting.

RAVALLION (2004) defends application of the weighting by population in the international context without reference to WILLIAMSON (1965). (From all appearances, economists engaged in studies of international inequality 'reinvented' this approach.) However, Ravallion's extensive reasoning relates to estimation of inequality among individuals, the world population, and not between countries as units of observation. Under such interpretation, the weighting is reasonable, albeit its results are disappointing, as the above-reported findings by MILANOVIC (2012) suggest. Generally speaking, the authors of studies on international inequality, in contrast to authors of regional studies, do not err as to the sense of population-weighted estimates of inequality. The focus of their debate is quite different: what is an adequate characterization of inequality in the world, either inter-country inequality or interpersonal inequality of world's population? This debate seems fairly pointless. It must be agreed with FIREBAUGH (2003, p. 129), who notes that the answer depends on the goal:

> the 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?

One can understand why the population-weighted inequality indexes are used to estimate interpersonal inequalities in international studies: the lack of relevant data forces to do this. However, such roundabout way in relation to a single country does not make sense. Many (if not most) national statistical agencies report data on distributions of personal incomes in their countries. Estimating national inequality even on so rough distribution as consisting of quintile income classes, the results obtained are much more exact than those based on the cross-region income distribution. Experiments with data from a number of countries (not reported) suggest that proxies of population inequality, estimated on cross-region income distributions with the use of population-weighted indexes, prove to be severely biased relative to estimates on national distributions of personal incomes and therefore quite misleading.

CONCLUSION

Following WILLIAMSON (1965), many economists estimate inter-regional inequality with the use of indexes weighted by regions' shares in the national population. A simple analysis in this note has proved that this approach is conceptually inconsistent. Instead of an estimate of inter-regional inequality, one gets an estimate of interpersonal inequality among the whole population of the country. Therefore the population-weighted estimates of inequality are biased with regard to estimates of both inter-regional inequality (as they measure a different value) and interpersonal inequality (as they do not and cannot take account of within-region income disparities). In both cases, the result can be not only distorted, but also quite misleading.

It is worth noting that the interpretation of a population-weighted inequality index as an approximate measure of interpersonal inequality of the whole population is not always true. It holds only regarding indicators which can characterize an economic or social situation of an individual, e.g. income, wage, housing, educational level, etc. There are different indicators that characterize situation of a region, but cannot be transferred to its certain inhabitant, e.g. unemployment rate, crime rate, investment per capita, gross regional product (GRP) per capita, etc. There is no inequality in unemployment rate between citizens of a country; only the national average unemployment rate exists. Likewise, there is no inequality in GDP (as the total of GRPs) per capita between country's citizens. In such cases, the population-weighted inequality indexes have no intuitive interpretation; they measure heaven knows what.

REFERENCES

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*. Harvard University Press, Cambridge, Mass.

JENKINS S. P. and VAN KERM P. (2009) The Measurement of Economic Inequality, in SALVEDRA V., NOLAN B. and SMEEDING T.M. (Eds.) *Economic Inequality*, pp. 40–67. Oxford University Press, Oxford.

METWALLY M. M. and RODNEY C. J. (1973) A note on the measurement of regional income dispersion, *Economic Development and Cultural Change* **22**, 135–136.

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.

PARR J. B. (1974) Welfare differences within a nation: a comment, *Papers of the Regional Science Association* **32**, 83–91.

PORTNOV B. A. and 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.

RAVALLION M. (2005) Competing concepts of inequality in the globalization debate, in COLLINS S.M. and GRAHAM C.L. (Eds.), *Brookings Trade Forum: 2004: Globalization, Poverty, and Inequality*, pp. 1–38. Brookings Institution Press, Washington, DC.

THEIL H. (1967) *Economics and Information Theory*. North Holland, Amsterdam.

UNITED NATIONS (2015) *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.

WILLIAMSON J. G. (1965) Regional inequality and the process of national development: a description of patterns, *Economic Development and Cultural Change* **13**, 1–84.

WORLD BANK (2015) World Development Indicators. – GDP per capita, PPP (current international $). http://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD. Accessed Sept. 26, 2015.