

УДК 332.143+303.71

ОБ ОЦЕНКЕ МЕЖРЕГИОНАЛЬНОГО НЕРАВЕНСТВА

К.П. Глущенко

Глущенко Константин Павлович – доктор экономических наук, ведущий научный сотрудник; Институт экономики и организации промышленного производства СО РАН, пр-т Лаврентьева, 17, Новосибирск, Россия, 630090. Профессор; Новосибирский национальный исследовательский государственный университет, ул. Пирогова, 2, Новосибирск, Россия, 630090. E-mail: glu@nsu.ru.

Во многих публикациях, посвященных анализу межрегионального неравенства, при оценке величины неравенства используется взвешивание по долям численности населения регионов в населении страны. В работе показано, что такой подход, во-первых, концептуально ошибочен, поскольку дает оценку не межрегионального неравенства, а неравенства всего населения страны (причем весьма грубую), во-вторых, при некорректном взвешивании сама величина измерителя неравенства оказывается неверной. На примере неравенства по доходам между российскими регионами, измеряемого коэффициентом Джини, демонстрируется, к каким искажениям приводит применение взвешивания.

Неравенство по доходам, коэффициент Джини, кривая Лоренца, коэффициент вариации Уильямсона.

DOI: 10.14530/se.2015.4.039-058

ВВЕДЕНИЕ

Анализ межрегионального неравенства охватывает множество аспектов, и посвященная ему литература настолько обширна, что невозможно привести даже ключевые публикации на эту тему. В работе [1] предпринята попытка дать сводное описание формальных методов, применяемых в исследованиях межрегионального неравенства по доходам (отметим, что большинство этих методов пригодно для анализа неравенства и по другим экономическим и социальным показателям), а в [2] дан обзор отечественных и зарубежных публикаций, в которых с помощью этих методов изучается неравенство по доходам между регионами России. Собственно оценки неравенства работы [1; 2] не касались¹, поскольку, во-первых, по этой про-

© Глущенко К.П., 2015

¹ Точнее, в них затрагивалась лишь одна ее сторона – обеспечение межрегиональной сопоставимости стоимостных показателей, неравенство по которым рассматривается.

блематике имеется множество исчерпывающих обзоров (напр., [11; 13]), а во-вторых, вопрос о расчете величины межрегионального неравенства довольно прост и, казалось бы, не заслуживает специального рассмотрения.

Действительно, чтобы оценить величину межрегионального неравенства, нужно просто выбрать подходящий (исходя из цели исследования или даже из своих вкусов) измеритель неравенства и рассчитать по имеющимся данным его величину, пользуясь известными формулами. Наиболее широко применяется индекс Джини, достоинство которого в том, что он ограничен сверху (не превышая 1) и потому сопоставим с оценками, например, по другим странам. Однако его невозможно декомпонировать, что не позволяет определить вклад того или иного региона в общую величину неравенства. Если же такая задача стоит, то следует использовать индекс Тейла или коэффициент вариации. Есть еще довольно много измерителей неравенства, но в региональных исследованиях они применяются очень редко.

И все же, несмотря на кажущуюся ясность вопроса о расчете величины межрегионального неравенства, с ним не все просто. Нередко (как в отечественных, так и зарубежных публикациях) оценка степени неравенства производится с учетом величины регионов, т. е. долей численности их населения в населении страны. Мотивация такого подхода (если приводится) по сути одинакова и весьма лаконична, например: «Для регионов России необходимо “взвешивание” показателей по численности населения. Невзвешенные измерения искажают степень регионального неравенства вследствие очень дробного административно-территориального деления и неравноценности субъектов РФ по численности населения» [4, с. 19].

Как видно, данный тезис просто постулируется, доказательная аргументация отсутствует. А между тем он вызывает много вопросов. Раз отсутствие взвешивания приводит к искажению, значит, есть некая «истинная» величина неравенства. Что это такое? Почему именно применение взвешивания дает эту «истинную» величину?

Что значит «очень дробное» административно-территориальное деление? Если это действительно приводит к искажениям, то почему тогда не оценивать неравенство между федеральными округами, а не между субъектами федерации? Или даже объединить все в две территории: состоящую из Центрального, Северо-Западного, Южного и Северо-Кавказского федеральных округов и состоящую из Приволжского, Уральского, Сибирского и Дальневосточного федеральных округов. Тогда и деление будет не «дробное», и численность населения в двух таких «регионах» будет примерно равной.

Почему вообще нужно учитывать различия в численности населения регионов? Ведь можно же оценивать неравенство между представителями разновеликих групп населения, например, неравенство по зарплате между

работниками разных профессий – учителями, промышленными рабочими, строителями, тружениками финансового сектора и т. д., не заботясь о том, какова их доля среди занятого населения. Чем это принципиально отличается от случая, когда группа населения – жители одного региона?

И, наконец, каков способ взвешивания? Если понимать приведенную выше цитату буквально, то показатель каждого региона нужно умножить на долю численности его населения в населении страны (или рассматриваемой группы регионов), а затем полученные величины использовать для расчета степени неравенства. Однако такой способ заведомо некорректен, что показывает простой пример. Допустим, есть два региона, в которых живут 400 и 200 тыс. чел., месячные душевые доходы в этих регионах равны соответственно 15 и 30 тыс. руб. Тогда взвешенные доходы в первом регионе составляют $15 \times 2/3 = 10$ тыс. руб., во втором – $30 \times 1/3 = 10$ тыс. руб. Таким образом, получаем нелепый результат: оказывается, что неравенство по доходам между этими регионами отсутствует, какой измеритель неравенства ни используем¹.

Однако и корректное взвешивание (о нем будет сказано позже) дает удивительные исходы. Возьмем Москву и Республику Калмыкию [8; 9]. Душевые доходы в них в 2013 г. составляли 54,9 и 11,3 тыс. руб., коэффициент Джини для этих двух субъектов федерации равен 0,329. Численность населения Москвы – 12,044 млн чел., Калмыкии – 0,283 млн чел. (2,3% от московского). И если рассчитать взвешенный по долям населения коэффициент Джини, то получим 0,018, в процентном выражении – 1,8%. Это означает, что неравенства по доходам между жителями столицы и Калмыкии практически нет. Не спасет и разделение Москвы на 43 равных «кусочка» с численностью населения каждого примерно как у Калмыкии: величина взвешенного коэффициента Джини от этого не изменится.

Этот пример демонстрирует, что использование взвешивания приводит к очевидно противоречащим здравому смыслу результатам. Когда же неравенство оценивается по большому числу регионов (например, по всем регионам России), такого рода «эффекты» оказываются замаскированными и остаются незамеченными, создавая иллюзию разумности полученной оценки. Цель данной работы состоит в том, чтобы развеять заблуждение относительно «необходимости» использования взвешенных измерителей неравенства в региональных исследованиях и показать, что как раз взвешивание и искажает оценку величины *межрегионального* неравенства.

Для определенности далее будем говорить о неравенстве по доходам. Однако все сказанное равно относится к любому показателю, характеризу-

¹ Правда, есть одно исключение: индекс Джини в виде, непосредственно основанном на кривой Лоренца, – приведенная далее формула (8).

ющему некоторый аспект ситуации в регионе, причем не обязательно экономическому (уровень безработицы, зарплата на одного работающего, уровень преступности, рождаемость и т. д.).

ЧТО МЫ ИЗМЕРЯЕМ?

Прежде всего рассмотрим, в чем содержательно состоит измерение межрегионального неравенства. Хотя при этом говорится о регионах, по сути рассматриваются индивидуумы: каждый регион представлен своим «средним» (т. е. получающим доход, равный среднему по региону) жителем. И при оценке межрегионального неравенства мы сравниваем между собой представительных жителей разных регионов (подобно тому, как мы берем «среднего» учителя, рабочего и т. д., оценивая неравенство по зарплате между профессиями). Все они при этом равноправны, вне зависимости от того, живут в небольшом или крупном регионе: тот факт, что доход среднего жителя Калмыкии вчетверо меньше дохода среднего москвича, никоим образом не меняется оттого, что москвичей в 42 с лишним раза больше. При малом числе регионов можно обойтись просто набором результатов таких попарных сравнений, при большом же приходится эти результаты каким-то образом усреднить, получая одно число, характеризующее различия доходов между представительными жителями рассматриваемых регионов в среднем. Разные способы усреднения и определения различия доходов и порождают многообразные измерители неравенства – стандартное отклонение логарифмов доходов, коэффициент вариации, коэффициент Джини, индекс Тейла, индекс Аткинсона и т. п.¹ Очевидно, что при этом регионы остаются равноправными, раз они были таковыми в каждом попарном сравнении.

Обозначим через y_i доходы на душу населения в регионе i ; $i = 1, \dots, m$. Будем всегда считать, что регионы упорядочены по возрастанию душевых доходов в них: если $i < j$, то $y_i \leq y_j$. Равноправие регионов выражается в том, что все величины y_i равновероятны, т. е. вероятность того, что в случайно выбранном регионе (из рассматриваемой нами совокупности) душевой доход окажется равным y_i , есть $p(y_i) = p_i = 1/m$. Функция распределения при этом выглядит таким образом:

$$F(y_i) = \sum_{k=1}^i p_k = i/m. \quad (1)$$

¹ Речь идет только об измерителях неравенства, которые используют все межрегиональное распределение доходов (т. е. охватывают все регионы рассматриваемой совокупности). Такими не являются, например, коэффициент фондов, учитывающий информацию только о крайних квантилях (чаще всего – децилях) распределения, «размах дифференциации» (отношение доходов в самом «богатом» и самом «бедном» регионах) и ряд других.

В большинстве показателей неравенства попарные сравнения производятся косвенно: душевой доход в данном регионе сопоставляется не с доходами в каждом из других регионов, а со средним доходом во всей рассматриваемой группе регионов \bar{y} . Но это – не средний доход населения данной группы регионов (так, если она состоит из всех регионов страны, \bar{y} не совпадет со средним душевым доходом в стране). Что же тогда представляет собой величина \bar{y} ? Поскольку регионы отождествляются с их представительными жителями, совокупный доход их всех равен $y_1 + \dots + y_m$, а средний – $\bar{y} = (y_1 + \dots + y_m)/m$. Таким образом, \bar{y} – это среднее арифметическое душевых доходов в регионах, в которое доход в каждом регионе входит равноправным образом.

Возможно, отсюда и берет начало стремление учесть неравноценность регионов по численности населения: ведь ясно, что вклад Москвы в совокупные доходы россиян несравнимо больше вклада Калмыкии. Но, во-первых, это не имеет никакого отношения к оценке неравенства между Москвой и Калмыкией. Во-вторых, существующее административно-территориальное деление страны – это данность, с которой мы и должны иметь дело, оценивая межрегиональное неравенство, каким бы неудачным это деление нам ни казалось (хотя, как утверждал Гегель, все действительно разумно). Можно, конечно, обсуждать его недостатки и предлагать пути совершенствования, но это уже совсем другая история, к тому же, опять-таки, не имеющая отношения к оценке межрегионального неравенства.

Но что же мы измеряем, когда учитываем неравноценность регионов, используя в показателе неравенства взвешивание по численности их населения? Возьмем другой пример – Санкт-Петербург с доходом на душу населения в 2013 г. 34,8 тыс. руб. и численностью населения 5,08 млн чел. и Ленинградскую область – 21,2 тыс. руб. и 1,758 млн чел. Среднее арифметическое этих доходов – 28 тыс. руб., коэффициент Джини – 0,121. Средний доход жителя территории, охватывающей Санкт-Петербург и Ленинградскую область, составляет $(34,8 \times 5,08 + 21,2 \times 1,758) / (5,08 + 1,758) = 34,8 \times 0,743 + 21,2 \times 0,257 = 31,3$ тыс. руб., что представляет собой среднее взвешенное душевых доходов в этих двух субъектах федерации, где весами служат доли каждого из них в совокупной численности населения их обоих. Рассчитав коэффициент Джини с этими же весами, получаем 0,083. О чем он говорит? О том, каким было бы неравенство *всего населения* региона, объединяющего Санкт-Петербург и Ленинградскую область (предложения о таком объединении встречаются в литературе). Причем оценка эта очень грубая, поскольку, не имея данных о неравенстве по доходам внутри рассматриваемых субъектов федерации, мы приняли, что доходы всех жителей каждого из них одинаковы и равны душевому доходу в соответствующем

субъекте федерации. Таким образом, мы измерили совсем не то, что собирались. Если взять все регионы России, то значение взвешенного показателя неравенства оценивает величину неравенства не *между регионами*, а между *всеми жителями* страны (к тому же без учета внутрирегиональной дифференциации доходов).

Теперь каждый регион i представлен не одним своим средним жителем, а всеми, т. е. группой из N_i жителей, каждый из которых имеет доход y_i . Совокупный доход населения всех регионов равен $\sum_{i=1}^m y_i N_i$, а доход на душу – $\bar{y}_{(w)} = (\sum_{i=1}^m y_i N_i) / N = \sum_{i=1}^m y_i n_i$, где N – численность населения страны (или рассматриваемой совокупности регионов), n_i – вес региона i , равный его доле в общей численности населения ($n_i = N_i / N$). Очевидно, что величины y_i уже не равновероятны: вероятность получить душевой доход y_i в случайно выбранном регионе тем выше, чем больше его население: $p(y_i) = n_i$, а функция распределения приобретает вид

$$F(y_i) = \sum_{k=1}^i n_k. \quad (2)$$

По сути, это группировка населения страны по уровню дохода (y_i) с группами разного размера (N_i). Региональное деление здесь теряет значение, иллюзия, что мы имеем дело с межрегиональным неравенством, обязана только тому, что группировка произведена на основе данных по регионам. Получим же мы, следуя такой схеме, характеристику не межрегионального, а межличностного неравенства – насколько это возможно при данном способе группировки населения. Вот она, неявно предполагаемая в приведенной во введении цитате «истинная» величина неравенства – степень неравенства всего населения. Ясно, что применение взвешивания при измерении неравенства (использование $p(y_i) = n_i$ вместо $p(y_i) = 1/m$) позволяет получить лучшее приближение этой величины. Но нужна-то нам совсем не она! Да и оценить неравенство всего населения страны по доходам можно гораздо точнее, пользуясь более пригодными для этого данными.

Таким образом, оценка с неравенства с помощью взвешенного показателя оказывается искаженной как по отношению к степени межрегионального неравенства (поскольку оценивается совсем другая величина), так и по отношению к степени неравенства всего населения (из-за того, что не учитывается внутрирегиональное неравенство).

По-видимому, первым идею об учете неравноценности регионов по численности населения при оценке межрегионального неравенства выдвинул Дж. Уильямсон в 1965 г. В качестве измерителя неравенства он использовал модифицированный коэффициент вариации [15, с. 11]:

$$V_w = \frac{\sqrt{\sum_{i=1}^m (y_i - \bar{y}_{(w)})^2 n_i}}{\bar{y}_{(w)}}. \quad (3)$$

Доказательным обоснованием своей идеи он тоже не озаботился, отметив лишь, что невзвешенный коэффициент вариации «будет отчасти определяться произвольным до некоторой степени административно-политическим выделением региональных единиц» и «преимущество взвешенного показателя по сравнению невзвешенным ... неоспоримо» [15, с. 11, 34].

Чтобы понять, что представляет собой коэффициент вариации Уильямсона (иногда называемый индексом Уильямсона), оценим степень неравенства населения страны с помощью коэффициента вариации

$$V = \frac{\sqrt{\sum_{l=1}^N (y_l - \bar{y})^2 / N}}{\bar{y}} \quad (4)$$

(где y_l – индивидуальный доход l -го жителя страны), имея данные только о межрегиональном распределении доходов. Тогда мы вынуждены считать, что у всех жителей одного региона доход одинаков. При этом доход на душу населения страны, среднее арифметическое доходов всех ее жителей, оказывается равным среднему взвешенному доходов представительных жителей регионов: $\bar{y} = \sum_{l=1}^N y_l / N = \sum_{i=1}^m N_i y_i / N = \sum_{i=1}^m y_i n_i = \bar{y}_{(w)}$. Квадратичные отклонения от среднего $(y_l - \bar{y})^2$ в формуле (4) будут одинаковы для всех l , которые относятся к жителям одного и того же региона, скажем, i -го, и их сумма по всем жителям региона равна $(y_i - \bar{y})^2 N_i$; всего же таких слагаемых будет столько, сколько имеется регионов, m . В итоге приходим к формуле Уильямсона (3):

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

Таким образом, коэффициент вариации Уильямсона измеряет вовсе не межрегиональное неравенство, это грубая оценка неравенства всего населения страны (к слову, она будет тем точнее, чем мельче регионы).

КАК ВВОДИТЬ ВЕСА?

Обратимся теперь к способам взвешивания. В формуле (3) видно, что взвешивание производится отнюдь не «наивным» способом, взвешиваются не сами душевые доходы в регионах (что соответствовало бы замене всех y_i на $n_i y_i$), а квадратичные отклонения от среднего. В других измерителях неравенства веса регионов будут входить в формулы иным образом, более

того, способ взвешивания даже одного и того же показателя неравенства зависит от формы его представления.

Не будем рассматривать разные показатели неравенства (да и особого смысла в этом нет, поскольку взвешенные показатели, как выяснилось, измеряют совсем не то, что нам надо), ограничимся только коэффициентом Джини. Во-первых, потому что это самый популярный измеритель межрегионального неравенства, во-вторых, потому что он допускает удобное наглядное представление.

Коэффициент Джини имеет множество ипостасей, до удивления непохожих друг на друга. Справочные издания, например [7], чаще всего дают такую формулу (в наших обозначениях и при упорядоченности доходов по возрастанию):

$$G = \frac{m+1}{m} - \frac{2 \sum_{i=1}^m (m+1-i)y_i}{m^2 \bar{y}}. \quad (5)$$

Имеются в виду зарубежные (или переводные) издания, отечественные экономические справочники, словари и энциклопедии почему-то обходятся только словесным описанием.

Другая формула исходит из среднего абсолютного различия доходов (она была предложена самим К. Джини):

$$G = \frac{\sum_{i=1}^m \sum_{j=1}^m |y_i - y_j|}{2m^2 \bar{y}}. \quad (6)$$

В ней явно видно, что оценка неравенства с помощью коэффициента Джини основана на попарных сопоставлениях (регионов, групп населения, индивидуумов и т. п.).

Самая изящная формула выглядит таким образом:

$$G = \frac{2 \operatorname{cov}(y, r/m)}{\bar{y}}, \quad (7)$$

где y – вариационный ряд: $y = (y_1, \dots, y_m)$, а r – ранги его элементов: $r = (1, \dots, m)$, т. е. $r_i = i$. Оpoznать в этой формуле коэффициент Джини, например, ее эквивалентность выражению (6), стоит немалых усилий.

Еще одна группа формул коэффициента Джини строится из геометрических соображений, основанных на кривой Лоренца. По сути они заключаются в вычислении удвоенной разности между площадью под линией абсолютного равенства и площадью под кривой Лоренца. Наиболее известная в настоящее время форма индекса Джини такого рода была предложена в [10], получив затем название формулы Брауна. Применительно к неравенству по доходам и в наших обозначениях она имеет вид

$$G = 1 - \sum_{i=1}^m p_i \left(\sum_{j=1}^{i-1} \frac{y_j}{Y} + \sum_{j=1}^i \frac{y_j}{Y} \right), \quad (8)$$

где $Y = \sum_{i=1}^m y_i$ – совокупный доход (а p_i , напомним, – вероятность дохода y_i). Величина $\sum_{i=1}^k p_i$ – это абсцисса k -й точки кривой Лоренца («накопленная» или «кумулятивная» доля регионов или населения), а $\sum_{j=1}^k \frac{y_j}{Y}$ – ее ордината («накопленная» доля совокупного дохода); координаты начальной (нулевой) точки кривой равны (0, 0).

Приведенными формулами все разнообразие представлений коэффициента Джини далеко не исчерпывается [16; 17]. К тому же время от времени предлагаются «новые» показатели неравенства, которые на поверку оказываются просто иной формой коэффициента Джини.

Формулу (8) можно назвать универсальной: она легко модифицируется и для оценки межгруппового неравенства, и для оценки неравенства всего населения по данным о доходах в его разновеликих группах. В первом случае у нас имеется m представителей каждой группы (средних жителей регионов), тогда $p_i = 1 / m$ и $Y = \sum_{i=1}^m y_i = m\bar{y}$. Во втором случае $p_i = n_i$, место душевого дохода y_i занимает доход всей группы населения $N_j y_j$, а совокупный доход составляет $Y = \sum_{i=1}^m N_i y_i = N\bar{y}_{(w)}$. И тогда

$$G_w = 1 - \sum_{i=1}^m n_i \left(\sum_{j=1}^{i-1} N_j y_j / Y + \sum_{j=1}^i N_j y_j / Y \right) = 1 - \frac{1}{\bar{y}_{(w)}} \sum_{i=1}^m n_i \left(\sum_{j=1}^{i-1} n_j y_j + \sum_{j=1}^i n_j y_j \right). \quad (9)$$

Это единственная формула, в которой взвешивание душевых доходов в регионах по численности их населения (простая замена y_j на $n_j y_j$) дает корректную величину взвешенного коэффициента Джини¹ (так как $y_j n_j / \bar{y}_{(w)} = y_j n_j N / Y = y_j N_j / Y$). Любая другая формула, за исключением тех, которые являются просто некоторым преобразованием выражения (8), требует иного способа взвешивания.

Дело в том, что формулы (5)–(7) и им подобные исходят из равновероятности всех наблюдений y_i , т. е. в них «встроена» функция распределения вида (1). В формулу (7) она входит явным образом – r/m и есть эта функция: $r_i / m = i/m = F(y_i)$. Если переменная y непрерывна, формула (7) имеет вид $G = 2\text{cov}(y, F(y)) / \bar{y}$. И, казалось бы, нужно просто заменить $F(y)$ на

¹ Так, для приведенного во введении примера двух регионов с численностью населения 400 и 200 тыс. чел. и душевыми доходами 15 и 30 тыс. руб. формула (8) дает $G_w = 0,167$ (что совпадает с величиной невзвешенного коэффициента Джини).

функцию распределения вида (2). Однако из-за ее дискретности просто не получается, нужно расписывать выражение для ковариации дискретных величин и «вмешиваться» в его структуру. В результате взвешенный коэффициент Джини записывается с использованием знаков суммирования, теряя элегантность исходной формулы; он приведен в [17, с. 29]. Довольно хлопотно преобразовать во взвешенный вариант и коэффициент Джини в виде (5).

Обратимся поэтому к формуле (6). Оценим с ее помощью неравенство населения по доходам, предположив пока для простоты, что имеются два региона с населением N_1 и N_2 , $N_1 + N_2 = N$, душевые доходы в которых y_1 и y_2 . Средний доход на душу всего населения равен тогда среднему взвешенному душевых доходов жителей регионов $\bar{y}_{(w)} = (N_1 y_1 + N_2 y_2) / N$. Всю совокупность абсолютных различий доходов – слагаемых в числителе формулы (6) – можно записать в виде матрицы Δ размерности $N \times N$: $(\Delta_{kl}) = (|y_k - y_l|)$. Возьмем первые N_1 строк этой матрицы. В них k – это номера разных жителей первого региона, доход каждого из них равен доходу на душу населения этого региона, т. е. $y_k = y_1$. То же самое относится к первым N_1 столбцам: $y_l = y_1$. Таким образом, для $k = 1, \dots, N_1$ и $l = 1, \dots, N_1$ $|y_k - y_l| = 0$. В столбцах с $N_1 + 1$ по N (число которых равно N_2) l относится к жителям второго региона, т. е. для всех $k = 1, \dots, N_1$ и $l = N_1 + 1, \dots, N$ $|y_k - y_l| = |y_1 - y_2|$. Следовательно, первые N_1 строк содержат $N_1 N_2$ элементов, равных $|y_1 - y_2|$, а остальные элементы нулевые. Возьмем теперь N_2 строк с $k = N_1 + 1, \dots, N$. Из аналогичных рассуждений вытекает, что в них $N_2 N_1$ ненулевых элементов, равных $|y_2 - y_1|$. Если теперь сложить все ненулевые элементы матрицы Δ , получим $N_1 N_2 |y_1 - y_2| + N_2 N_1 |y_2 - y_1|$. Отсюда понятно, что если регионов m , подобные слагаемые будут иметься для каждой их пары. И взвешенный коэффициент Джини имеет вид

$$G_w = \frac{\sum_{i=1}^m \sum_{j=1}^m 2N_i N_j |y_i - y_j|}{2N^2 \bar{y}_{(w)}} = \frac{\sum_{i=1}^m \sum_{j=1}^m n_i n_j |y_i - y_j|}{\bar{y}_{(w)}}$$

Очевидно, что $n_i n_j |y_i - y_j|$ вовсе не то же самое, что $|n_i y_i - n_j y_j|$, не говоря о том, что изменился и знаменатель формулы (6).

КАКИЕ ИСКАЖЕНИЯ ДАЕТ ВЗВЕШИВАНИЕ?

Рассмотрим теперь, к каким количественным искажениям приводит оценка межрегионального неравенства с помощью взвешенного коэффициента Джини – сначала на искусственном примере двух регионов, а затем – на реальных данных. Условные примеры с малым числом регионов хороши тем, что позволяют легко увидеть направления искажений и их причины,

при большом же числе регионов картина получается «смазанной». Нужно заметить, что максимальное значение коэффициента Джини (соответствующее абсолютному неравенству) зависит от числа наблюдений, оно равно $(m - 1) / m$. И на эту величину надо бы нормировать коэффициент Джини, чтобы иметь представление о том, насколько наблюдаемое неравенство далеко от абсолютного. Однако такую нормировку не применяют, так как при достаточно большом числе регионов $(m - 1) / m$ мало отличается от единицы (равняясь, например, 0,987 для 79 российских регионов). Не применяется она и здесь. Но при двух регионах абсолютному неравенству соответствует величина 0,5, и чтобы понять, насколько действительно велики значения коэффициента Джини в случае двух регионов, их следует удваивать.

Допустим, имеются два региона с населением 700 и 300 тыс. чел. Душевой доход в первом – 20 тыс. руб. в месяц, во втором – 90 тыс. руб. Тогда величина межрегионального неравенства, измеряемая коэффициентом Джини, равна 0,318. Взвешенный коэффициент Джини составляет 0,359. Как было показано, это неравенство всего населения обоих регионов без учета того, что жители одного и того же региона имеют разный доход. Пусть доход жителей растет линейно, с 10 тыс. руб. у самого бедного до 30 тыс. руб. у самого богатого в первом регионе, и с 40 до 140 тыс. руб. во втором (такое распределение доходов дает указанные выше душевые доходы). Тогда величина внутрорегионального неравенства в первом регионе – 0,167, во втором – 0,185, а всего населения обоих регионов – 0,435. Наглядное представление о различиях оценок дают графики кривых Лоренца, приведенные на рисунке 1.

Напомним, что коэффициент Джини равен удвоенной площади фигуры, заключенной между линией абсолютного равенства и кривой Лоренца. Величины на оси абсцисс имеют разный смысл в зависимости от того, рассматриваем мы регион как «единый и неделимый» (отождествляя его с представительным жителем) или как группу населения, что соответствует трактовке региона при использовании взвешенного коэффициента Джини. В первом случае на ней откладывается доля регионов в их общем количестве, для i -й точки кривой Лоренца она равна i / m , а расстояние между соседними точками всегда одинаково и равно $1 / m$. Во втором случае на оси абсцисс – доля групп населения в его общей численности; для i -й точки имеем $n_1 + \dots + n_i$, расстояние же между соседними точками – переменное, между $(i - 1)$ -й и i -й оно составляет n_i . Если рассматривается действительно все население (упорядоченное по возрастанию дохода), без его группировки по регионам, то каждая точка оси абсцисс в идеальном случае соответствует одному человеку. На практике же все население разбивается на группы равного размера и принимается, что у всех входящих в одну группу

доход одинаков и равен среднему по группе. Кривая Лоренца для всего населения на рисунке 1 построена по разбивке населения на 1-процентные доходные группы.

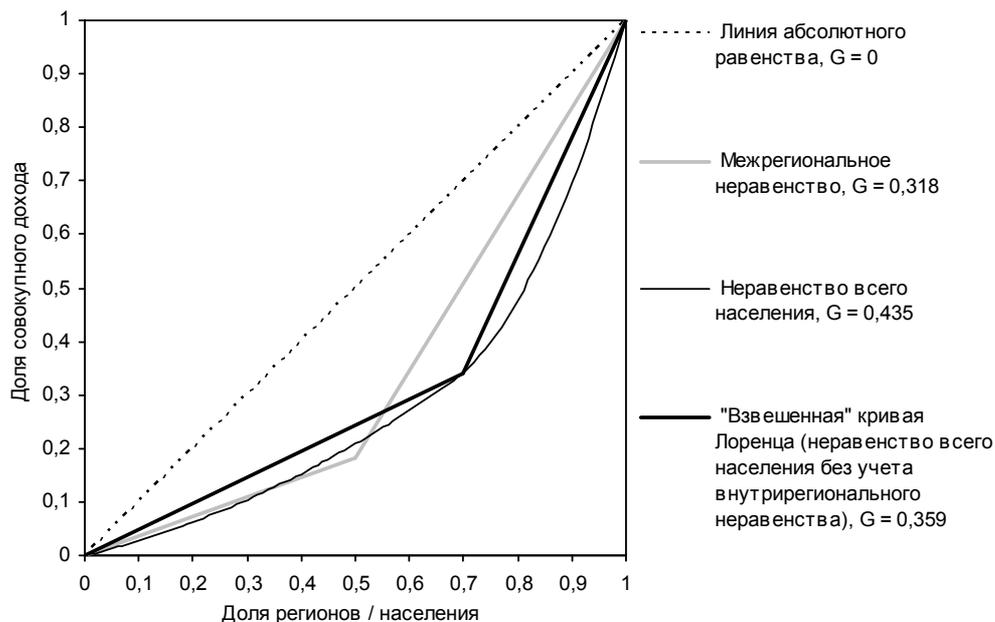


Рис. 1. Кривые Лоренца для условного примера

В нашем примере взвешенный индекс Джини оказался завышенным по сравнению с величиной межрегионального неравенства и заниженным по сравнению с неравенством всего населения. Но в общем случае взвешенный индекс Джини может быть смещен по отношению к данным оценкам в любую сторону, это будет зависеть от конкретного соотношения относительных численностей населения и душевых доходов в рассматриваемых регионах. Занижение неравенства всего населения взвешенным индексом Джини в нашем примере ожидаемо: это следствие того, что распределения доходов в регионах не пересекаются. В действительности же они пересекаются почти всегда, если не говорить о межстрановых сравнениях, и группировка населения по доходным группам будет далеко не совпадать с его группировкой по регионам (что было бы в нашем примере при диапазоне доходов во втором регионе от 20 до 160 тыс. руб.). Отметим также, что если просто взвешивать душевые доходы в регионах, заменив их значения на $30 \times 0,7 = 14$ тыс. руб. и $90 \times 0,3 = 27$ тыс. руб. в любой из формул (5)–(7), то получим сильно заниженную величину: $G_w = 0,159$. Результат тоже вполне ожидаемый, поскольку такое «взвешивание» значительно уменьшило различие в душевых доходах между регионами. Но в общем случае и тут отклонение может быть в любую сторону.

А каковы будут расхождения оценок на реальных данных? На рисунке 2 показаны результаты оценки неравенства по доходам в России в 1998–2013 гг. Для оценки межрегионального неравенства и расчета взвешенного коэффициента Джини по формуле (9) использовались данные по 79 регионам. При этом автономные округа, входящие в другие субъекты федерации, учитывались в их составе и отдельно не выделялись; из рассмотрения была исключена Чеченская Республика, данные о доходах в которой имеются только с 2010 г. Коэффициенты Джини для всего населения страны – непосредственно данные Росстата [6].

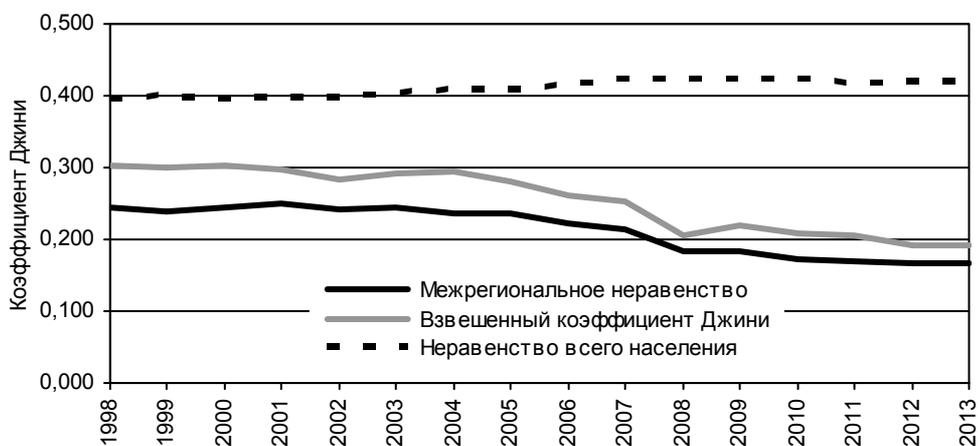


Рис. 2. Неравенство по доходам в России

Источник: расчеты автора по данным [6; 8; 9].

Как видно, взвешенный коэффициент Джини на всем рассматриваемом промежутке времени завышает межрегиональное неравенство (небольшие отличия его величин от приведенных в [4] обусловлены тем, что Росстат обновил в 2015 г. ретроспективные данные о душевых доходах в регионах; здесь использованы именно обновленные данные). По отношению же к неравенству всего населения он существенно занижен, мало того, имеет тенденцию к снижению, тогда как неравенство населения – тенденцию к росту (хотя и слабую), т. е. динамика взвешенного коэффициента Джини как грубой оценки неравенства всего населения противоположна фактической. Расхождения между невзвешенным (G) и взвешенным (G_w) коэффициентами Джини представлены в таблице.

Искажения, вносимые взвешиванием, довольно существенны: завышение составляет от 15 до 25%, при этом оно непостоянно во времени. Интересно взглянуть на кривые Лоренца, основанные на реальных данных.

Таблица

Расхождения между невзвешенными и взвешенными оценками неравенства

Показатель	1998	1999	2000	2001	2002	2003	2004	2005
G	0,245	0,240	0,245	0,249	0,241	0,244	0,237	0,235
G_w	0,303	0,301	0,303	0,298	0,282	0,293	0,294	0,280
$(G_w - G) / G, \%$	23,4	25,4	23,5	19,7	16,8	20,1	23,9	18,8
Показатель	2006	2007	2008	2009	2010	2011	2012	2013
G	0,223	0,213	0,183	0,184	0,173	0,169	0,166	0,167
G_w	0,262	0,252	0,206	0,219	0,209	0,205	0,193	0,192
$(G_w - G) / G, \%$	17,4	18,1	12,3	18,6	20,6	21,1	16,1	14,7

Источник: расчеты автора по данным [8; 9].

Они приведены для 2011 г. на рисунке 3. Кривая Лоренца для всего населения построена по распределению общего объема денежных доходов по 20-процентным группам населения [6], поэтому она огрубляет действительную кривую, заменяя ее ломаной линией.

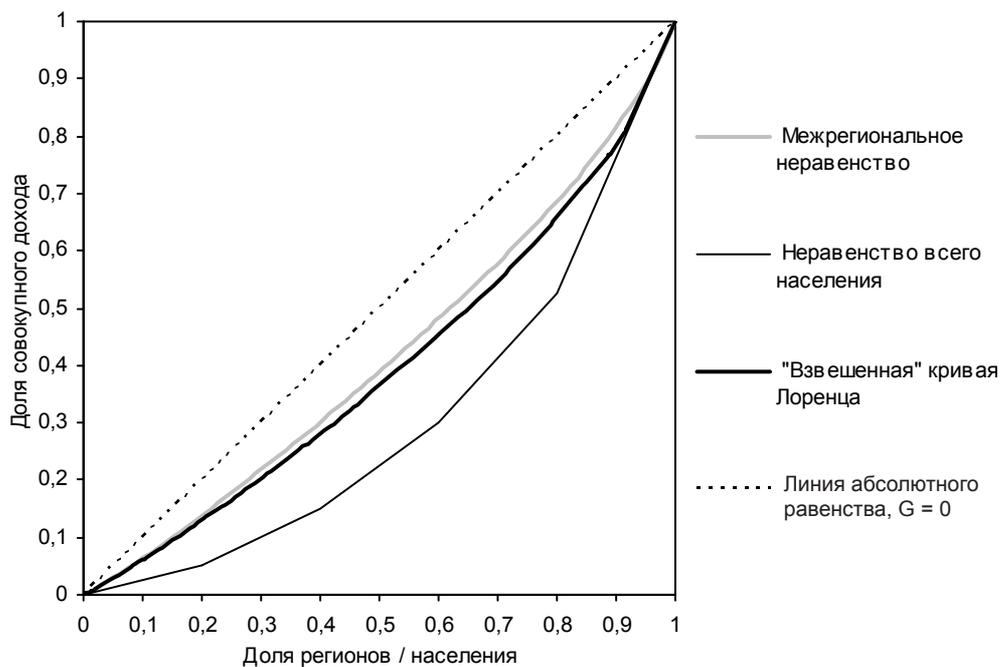


Рис. 3. Кривые Лоренца для распределения доходов в России в 2011 г.

Источник: расчеты автора по данным [6; 8; 9].

Зрительно различие кривых Лоренца, построенных на основе невзвешенных и взвешенных данных о душевых доходах в регионах, кажется небольшим (хотя площади, заключенные между ними и линией абсолютного равенства, разнятся, согласно таблице, на $21,1\% / 2 = 11\%$). Да и сама «взвешенная» кривая Лоренца и динамика неравенства, измеряемого взвешенным коэффициентом Джини на рисунке 2, выглядят вполне правдоподобно. Это и показывает, что большое количество регионов маскирует обусловленные взвешиванием несурзаицы (с точки зрения межрегионального неравенства), которые бросаются в глаза, когда регионов немного.

В работе [15, с. 12] Дж. Уильямсон представил результаты оценки межрегионального неравенства в 24 странах мира. При этом наряду со взвешенными коэффициентами вариации (которым он отдает предпочтение), приводятся и невзвешенные, что позволяет увидеть последствия взвешивания на разнообразном материале. Примерно в половине случаев оно привело к завышению межрегионального неравенства, в половине – к занижению. Диапазон смещений взвешенного показателя по отношению к невзвешенному составил от $-52,6\%$ (Индия) до $+37,6\%$ (Пуэрто-Рико), т. е. степень межрегионального неравенства в Индии оказалась заниженной при использовании взвешивания более чем вдвое.

Хотя взвешенные коэффициенты Джини для России в 1998–2013 гг. и искажают величину неравенства, их динамика, как видно на рисунке 2, качественно подобна динамике межрегионального неравенства. Поэтому качественные выводы о ее характере от этих искажений не пострадают. Так что это еще не самый плохой случай.

Бывает хуже. Б. Миланович получил, используя коэффициент Джини, динамику межстранового и глобального неравенства (т. е. неравенства всего населения мира) по доходам за период 1952–2006 гг. на основе данных о ВВП на душу населения [14]. Он рассчитывал также коэффициент Джини, взвешенный по численности населения стран, но не как показатель межстранового неравенства, а как грубую оценку неравенства населения мира, о чем прямо сказано в тексте. Сделано это не от хорошей жизни – для большей части рассматриваемого периода информация о внутривострановой дифференциации доходов во многих странах отсутствует. Тренды взвешенного и невзвешенного коэффициентов Джини оказались разнонаправленными: для первого – убывающий, для второго – возрастающий, и только с 2000 г. они стали убывающими у обоих. И если бы выводы о тенденции межстранового неравенства делались на основании значений взвешенного коэффициента Джини, они бы оказались прямо противоположными тому, что происходило в действительности.

Любопытно, что когда из выборки, по которой рассчитывался взвешен-

ный коэффициент Джини, был исключен Китай, динамика неравенства стала походить на ту, которую дает невзвешенный показатель (подобно тому, что мы наблюдаем для России на рисунке 2), хотя и при завышенной величине неравенства – в среднем, очень грубо, где-то на 10%. Но и как приближенная оценка глобального неравенства взвешенный коэффициент Джини, похоже, оказался не очень хорош. Оценки глобального неравенства за 1988–2005 гг., рассчитанные по распределениям доходов населения внутри стран, во-первых, оказались заметно выше, во-вторых, имели тенденцию к росту, а не к снижению (и тут подобие с тем, что имеем в случае России)¹.

ЗАКЛЮЧЕНИЕ

Проведенный анализ показывает, что оценка степени межрегионального неравенства с использованием взвешивания по доле численности населения региона в населении страны концептуально ошибочна, поскольку дает величину не межрегионального неравенства, а неравенства всего населения страны (причем очень грубую, можно даже сказать, малодостоверную). Приведенные примеры демонстрируют, что взвешивание приводит к существенным количественным искажениям, а в некоторых случаях может исказить и качественную картину, давая неверное представление о тенденции динамики межрегионального неравенства.

Если же взвешивание производится некорректным способом, сама величина измерителя неравенства оказывается неверной. Так, если просто заменить душевые доходы в регионах y_i «взвешенными» величинами $n_i y_i$ и затем рассчитать коэффициент Джини по одной из формул (5)–(7), то получим для России в период 1998–2013 гг. значения в диапазоне 0,547–0,584, а их динамика будет очень слабо походить на действительную динамику межрегионального неравенства. Насколько распространен такой способ «взвешивания», судить трудно, так как зачастую использованная формула взвешенного коэффициента Джини не приводится (в числе немногих исключений можно назвать, например, работу [5]), но величины оценок в некоторых публикациях наводят на мысль, что применялся именно данный способ.

Следует сказать, что и интерпретация взвешенного по численности населения регионов измерителя неравенства как приближенной оценки неравенства всего населения страны справедлива далеко не всегда. Она действительна в отношении только тех показателей, которые могут характеризовать экономическое или социальное положение отдельной личности: доход, за-

¹ Возможно, свой вклад внес и разный характер данных – все-таки душевые ВВП и денежные доходы населения даже в межстрановом анализе не вполне взаимозаменяемы, хотя и сильно коррелированы.

рабочая плата, обеспеченность жильем ($m^2/\text{чел.}$) и т. д. А есть показатели, характеризующие ситуацию в регионе в целом, но не применимые по отношению к отдельному его жителю: уровни безработицы, бедности, преступности, валовой региональный продукт (ВРП) на душу населения, инвестиции на душу населения и т. д. Нет неравенства между жителями России по уровню безработицы – есть только средний уровень безработицы в стране, как нет и неравенства между ними по душевому ВВП (точнее, по суммарному ВРП в расчете на душу населения страны). И взвешенный измеритель неравенства по такого рода показателям вообще не имеет содержательной интерпретации, он измеряет неизвестно что.

Использованные в работе реальные данные представляли собой номинальные доходы. Для примеров это приемлемо, однако в исследованиях действительного положения дел неравенство по номинальным доходам малопоказательно. Так, душевой доход в Чукотском автономном округе вдвое с лишним превышает среднероссийский. Однако если учесть стоимость жизни на Чукотке, оказывается, что покупательная способность столь высокого дохода лишь немногим больше, чем в среднем по стране. Поэтому для анализа фактического неравенства нужно использовать реальные (т. е. сопоставимые между регионами) доходы, учитывающие различия в уровнях цен между регионами.

Для их оценки можно воспользоваться данными официальной статистики: стоимостью фиксированного набора потребительских товаров и услуг для межрегиональных сопоставлений покупательной способности населения или индексом стоимости жизни. Правда, у обоих есть свои неудобства. Стоимость фиксированного набора имеется только с 2002 г., и показатель это месячный, так что приходится иметь дело с месячными данными о доходах и затем усреднять месячные реальные доходы за год. Индекс стоимости жизни публикуется лишь с 2009 г.; он дается в годовом выражении, но не по регионам, а по городам. Чтобы получить индекс стоимости жизни в регионе, нужно рассчитать среднее по входящим в него городам с весами, равными долям численности населения городов в суммарном по ним населении. Региональные же индексы потребительских цен (ИПЦ) для оценки реальных душевых доходов в регионах, как показано в [12], непригодны. Искажения при их использовании настолько велики, что даже номинальные доходы дают лучшее «приближение» реальных (главная причина в том, что в каждом регионе своя система весов, используемых при расчете ИПЦ).

Сходная ситуация и с ВРП на душу населения. Ясно, что если один и тот же продукт будет произведен на Чукотке и в Центральной России, то добавленная стоимость в первом случае будет гораздо выше (хотя бы из-за того, что в нее входит зарплата). Поэтому также возникает проблема

обеспечения межрегиональной сопоставимости ВРП. Только тут она гораздо сложнее из-за разнородности его составляющих. Ее удалось решить А.Г. Гранбергу и Ю.С. Зайцевой [3]. Однако воспользоваться их количественными результатами для получения оценок современных реальных ВРП, скорее всего, не удастся, так как приведенные в [3] оценки сильно удалены во времени – они относятся к 1999 г. А с помощью индексов физического объема ВРП (региональных дефляторов) получить сопоставимые между регионами ВРП нельзя. Причина сходна с той, которая имеет место в случае ИПЦ: при расчете индексов физического объема ВРП в каждом регионе используется специфическая для него структура ВРП.

СПИСОК ЛИТЕРАТУРЫ

1. *Глущенко К.П.* Методы анализа межрегионального неравенства по доходам // Регион: экономика и социология. 2010. № 1. С. 54–87.
2. *Глущенко К.П.* Исследования неравенства по доходам между российскими регионами // Регион: экономика и социология. 2010. № 4. С. 88–119.
3. *Гранберг А.Г., Зайцева Ю.С.* Межрегиональные сопоставления валового регионального продукта в Российской Федерации // Вопросы статистики. 2003. № 2. С. 3–17.
4. *Зубаревич Н.В., Сафронов С.Г.* Неравенство социально-экономического развития регионов и городов России 2000-х годов: рост или снижение? // Общественные науки и современность. 2013. № 6. С. 15–26.
5. *Малкина М.Ю.* Исследование взаимосвязи уровня развития и степени неравенства доходов в регионах Российской Федерации // Экономика региона. 2014. № 2. С. 238–248.
6. Распределение общего объема денежных доходов и характеристики дифференциации денежных доходов населения. 1970–2014 гг. // ФСГС России. URL: www.gks.ru/free_doc/new_site/population/bednost/tab1/1-2-2.doc (дата обращения: август 2015).
7. Словарь современной экономической теории Макмиллана. М.: ИМФРА-М, 1997. 608 с.
8. Среднедушевые денежные доходы. 2013 г. // ФСГС России. URL: www.gks.ru/dbscripts/cbsd/DBInet.cgi?pl=2340019 (дата обращения: август 2015).
9. Численность постоянного населения в среднем за год. 2013 г. // ФСГС России. URL: www.gks.ru/dbscripts/cbsd/DBInet.cgi?pl=2403024 (дата обращения: август 2015).
10. *Brown M.C.* Using Gini-Style Indices to Evaluate the Spatial Patterns of Health Practitioners: Theoretical Considerations and an Application Based on Alberta Data // Social Science & Medicine. 1994. Vol. 38. No. 9. Pp. 1243–1256. DOI: 10.1016/0277-9536(94)90189-9.
11. *Cowell F.A.* Measurement of Inequality // Handbook of Income Distribution. Vol. 1. Amsterdam: North Holland, 2000. Pp. 87–166.
12. *Gluschenko K.* Biases in Cross-Space Comparisons Through Cross-Time Price Indexes: The Case of Russia / BOFIT Discussion Papers. No 9. Helsinki: BOFIT, 2006. 51 p.
13. *Jenkins S.P., Kerm P.* The Measurement of Economic Inequality // The Oxford Handbook of Economic Inequality. Oxford: Oxford University Press, 2009. Pp. 40–67.

14. *Milanovic B.* Global Inequality Recalculated and Updated: The Effect of New PPP Estimates on Global Inequality and 2005 Estimates // *Journal of Economic Inequality*. 2012. Vol. 10. No. 1. Pp. 1–18. DOI: 10.1007/s10888-010-9155-y.

15. *Williamson J.G.* Regional Inequality and the Process of National Development: A Description of Patterns // *Economic Development and Cultural Change*. 1965. Vol. 13. No. 4. Part 2. Pp. 1–84.

16. *Yitzhaki S.* More Than a Dozen Alternative Ways of Spelling Gini // *Research on Economic Inequality*. 1998. Vol. 8. Pp. 13–30.

17. *Yitzhaki S., Schechtman E.* The Gini Methodology. A Primer on a Statistical Methodology. New York: Springer, 2013. DOI: 10.1007/978-1-4614-4720-7.

ON ESTIMATION OF INTER-REGIONAL INEQUALITY

K.P. Gluschenko

Gluschenko Konstantin Pavlovich – Doctor of Economics, Leading Researcher. Institute of Economics and Industrial Engineering of the Siberian Branch of the Russian Academy of Sciences. Professor. Novosibirsk National Research State University. 17 Lavrentyeva pr., Novosibirsk, Russia, 630090. E-mail: glu@nsu.ru.

Many publications on inter-regional inequality use weighting by region's share in the national population in estimating an inequality index. This paper demonstrates that, first, such an approach is conceptually inconsistent, yielding a (very rough) estimate of interpersonal inequality of the country's whole population rather than an estimate of inter-regional inequality, and second, an inadequate weighting leads to the incorrect value of an inequality index. The author illustrates distortions caused by the use of weighting with the help of the income inequality among Russian regions (measured by the Gini coefficient).

Keywords: income inequality, Gini coefficient, Lorentz curve, Williamson coefficient of variation.

REFERENCES

1. Gluschenko K.P. Methodologies of Analyzing Inter-Regional Income Inequalities. *Region: Ekonomika i Sotsiologiya* [Region: Economics and Sociology], 2010, no. 1, pp. 54–87. (In Russian).

2. Gluschenko K.P. Studies on Income Inequality among Russian Regions. *Regional Research of Russia*, 2011, vol. 1, no. 4, pp. 319–330. DOI: 10.1134/S2079970511040034.

3. Granberg A.G., Zaytseva Yu.S. Interregional Comparisons of Gross Regional Product in the Russian Federation. *Voprosy Statistiki* [Statistical Issues], 2003, no. 2, pp. 3–17. (In Russian).

4. Zubarevich N.V., Safronov S.G. The Inequality of Social and Economic Development of Regions and Cities of Russia of the 2000s: Growth or Decline? *Obshchestvennye Nauki i Sovremennost* [Social Sciences and Modernity], 2013, no. 6, pp. 15–26. (In Russian).

5. Malkina M.Yu. Study of the Relationship Between the Development Level and Degree of Income Inequality in the Russian Regions. *Ekonomika Regiona* [Regional Economy], 2014, no. 2, pp. 238–248. (In Russian).

6. *The Distribution of Total Money Income and Characteristics of Differentiation of Money Incomes of the Population. 1970–2014*. Federal State Statistics Service of the Russian Federation. Available at: www.gks.ru/free_doc/new_site/population/bednost/tab1/1-2-2.doc (accessed August 2015). (In Russian).

7. *Macmillan Dictionary of Modern Economics*. 4th edition. London: Palgrave Macmillan, 1992, 486 p.

8. *Average per Capita Money Income of Population. 2013*. Federal State Statistics Service of the Russian Federation. Available at: www.gks.ru/dbscripts/cbsd/DBInet.cgi?pl=2340019 (accessed August 2015). (In Russian).

9. *Resident Population, the Yearly Average. 2013*. Federal State Statistics Service of the Russian Federation. Available at: www.gks.ru/dbscripts/cbsd/DBInet.cgi?pl=2403024 (accessed August 2015). (In Russian).

10. Brown M.C. Using Gini-Style Indices to Evaluate the Spatial Patterns of Health Practitioners: Theoretical Considerations and an Application Based on Alberta Data. *Social Science & Medicine*, 1994, vol. 38, no. 9, pp. 1243–1256. DOI: 10.1016/0277-9536(94)90189-9.

11. Cowell F.A. Measurement of Inequality. *Handbook of Income Distribution. Vol. 1*. Amsterdam: North Holland, 2000, pp. 87–166.

12. Gluschenko K. Biases in Cross-Space Comparisons Through Cross-Time Price Indexes: The Case of Russia. *BOFIT Discussion Papers. No 9*. Helsinki: BOFIT, 2006, 51 p.

13. Jenkins S.P., van Kerm P. The Measurement of Economic Inequality. *The Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press, 2009, pp. 40–67.

14. Milanovic B. Global Inequality Recalculated and Updated: The Effect of New PPP Estimates on Global Inequality and 2005 Estimates. *Journal of Economic Inequality*, 2012, vol. 10, no. 1, pp. 1–18. DOI: 10.1007/s10888-010-9155-y.

15. Williamson J.G. Regional Inequality and the Process of National Development: A Description of Patterns. *Economic Development and Cultural Change*, 1965, vol. 13, no. 4, part 2, pp. 1–84.

16. Yitzhaki S. More Than a Dozen Alternative Ways of Spelling Gini. *Research on Economic Inequality*, 1998, vol. 8, pp. 13–30.

17. Yitzhaki S., Schechtman E. *The Gini Methodology. A Primer on a Statistical Methodology*. New York: Springer, 2013. DOI: 10.1007/978-1-4614-4720-7.