

Рабочий доклад 6
29 October 2018

**ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ
КОМИССИЯ ООН
КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ**

**Экспертное заседание по измерению бедности и неравенства
29-30 ноября 2018 года, Вена, Австрия
Тема В: Увеличение процента откликнувшихся на обследования и улучшение
точности выборки**

Отрывок из доклада

**Сложности эмпирического характера при сравнении неравенства в разных странах:
кейс стран со средним уровнем дохода из Межнационального центра хранения и
обработки данных LIS**

Доклад подготовили: Daniele Checchi ^a, Andrej Cupak ^b, Janet Gornick ^c, and Teresa Munzi ^d
(Предварительная версия: 11/10/2018)

^a Corresponding author. LIS Cross-National Data Center in Luxembourg, Esch-Belval, Luxembourg. E-mail: checchi@lisdatacenter.org.

^b LIS Cross-National Data Center in Luxembourg, Esch-Belval, Luxembourg. E-mail: cupak@lisdatacenter.org.

^c Stone Center on Socio-Economic Inequality, The City University of New York, New York, USA. E-mail: jgornick@gc.cuny.edu.

^d LIS Cross-National Data Center in Luxembourg, Esch-Belval, Luxembourg. E-mail: munzi@lisdatacenter.org.

2 Трудности в гармонизации данных из стран со средним уровнем дохода

С момента своего возникновения в 1980-ых гг LIS по правилу уделял наибольшее внимание странам с высоким уровнем дохода. В 2007 году при сотрудничестве с группой из Всемирного Банка был проведен пилотный проект с целью изучения осуществимости включения стран со средним уровнем дохода в базу данных LIS. Вслед за решением провести данное расширение LIS произвел некоторые концептуальные коррекции и изменения в своем списке гармонизируемых переменных, что позволило бы включить более разнообразные характеристики трудового рынка, структуры социальных льгот, схемы потребления, переменные показатели потоков доходов внутри страны и между странами.

Среди основных произведенных изменений в результате основательного пересмотра формуляра в 2011 году, следующие из них были в основном нацелены на максимальное повышение его применения к базам данных стран как с высоким, так и со средним уровнем дохода:

- Скорректировали понятие «располагаемого дохода домашнего хозяйства» таким образом, что включили туда немонетарные доходы труда, а также доходы полученные от общественных и частных третьих лиц. Основной причиной такого расширения концепции РДДХ послужил факт, что во многих странах со средними доходами доля неденежных доходов в их собственном потреблении и социальных и/или получаемых от частных лиц/организаций трансфертов была слишком важна, чтобы ею пренебречь, и кстати, эти суммы намного чаще предоставляются в странах со средним уровнем дохода, чем с высоким, где многие организации, занимающиеся сбором данных, даже не учитывают их из-за их незначительности (См. Рис. 1, где включены только страны LIS, по которым идет сбор немонетарных доходов).
- Скорректировано понятие «члена домашнего хозяйства» таким образом, чтобы лица, физически находящиеся в домашнем хозяйстве, но чьи доходы не включены в доход домашнего хозяйства (а именно, проживающая в дома прислуга, постояльцы и снимающие жилплощадь квартиранты) не учитывались в формировании общего дохода домашнего хозяйства или в расчетах эквивалентной шкалы.
- Было включено несколько переменных по организации проживания, что позволило проводить лучший анализ многоквартирных/многочисленных домашних хозяйств таким образом, чтобы была учтена доступная информация о партнерстве и родительских обязанностях взрослых, не входящих в ядро семьи.
- Были включены переменные, содержащие информацию по целому ряду новых тем, в том числе сельские/городские показатели, показатель сельскохозяйственной деятельности, типы жилых объектов, участие в маргинальной/неформальной рабочей деятельности, характеристики второстепенной работы.

По прошествии десяти лет гармонизации данных из стран со средним доходом с данными стран с высокими доходами, LIS приобрел определенный опыт решения основных трудностей, с которыми обычно сталкиваются при работе с микроданными доходов, полученных из этих источников, как об этом говорится ниже.

Благодаря отличиям сельских и городских районов во многих странах со средними доходами обследования доходов проводятся или только в городских районах¹, где легче учитывать доходы, или используется сильно отличающийся инструментарий (включая различную выборку и вопросники) для сельских и городских районов. Эти особенности вызвали особые сложности в китайском обследовании, где интеграция трех разных выборок (городских, сельских и мигрантов из сельских в городские районы) в единую национальную выборку потребовала определенной коррекции веса и самих переменных, что поставило под угрозу качество окончательных результатов (до той меры, что в 2007 году, где было особенно много трудностей, LIS принял решение не публиковать данные по Китаю)².

Определение *членства домашнего хозяйства* (с вытекающим из него методом учета индивидуальных доходов при формировании уровня доходов домашнего хозяйства) имеет особое значение в странах со средним уровнем дохода. Проживающая в домашнем хозяйстве прислуга, явление намного чаще встречающееся, и эти лица или учитываются как члены домашнего хозяйства (но тогда возникает риск двойного учета доходов, которые они получают от своего работодателя – домашнего хозяйства), или не учитываются как члены домашнего хозяйства (тогда возникает риск, что они вообще не будут представлены в выборке). Временно отсутствующие члены семьи иногда считаются членами домашнего хозяйства, а иногда нет, и часто бывает очень сложно разделить две такие ситуации. Особенно это сложно, когда взрослые временно отсутствуют с целью заработков в другом месте (в случае отсутствия главы семьи или супругов), поскольку это в значительной мере меняет способ учета их доходов (следует ли включить весь его доход в категорию доходов от трудовой деятельности, или только ту часть, которую он пересылает своей семье в виде денежных переводов). В некоторых случаях также собирается информация на индивидуальном уровне об увеличившемся количестве членов семьи, и тогда может быть сложно решить, являются ли эти лица членами домашнего хозяйства или нет (что опять-таки имеет свои очевидные последствия на метод учета их доходов).

Домашние хозяйства, где представлено много поколений, а также многочисленные или *сложные домашние хозяйства* встречаются намного чаще. В зависимости от того, кто считается главой (представитель старшего или среднего поколения в случае трех поколений членов), характеристики домашнего хозяйства, часто основывающиеся на главе домашнего хозяйства и его

¹ Например, Непрерывное обследование домашних хозяйств в Уругвае в период до 2005 года охватывало только городское население.

² Еще один хороший пример это корейские данные, которые берутся из двух разных обследований с различной структурой и вопросниками. Первое это Обследование доходов и расходов домашних хозяйств, которое охватывает вне неведомственные домашние хозяйства на территории Южной Кореи, исключая 1) сельскохозяйственные домашние хозяйства, 2) рыболовные домашние хозяйства, 3) домашние хозяйства, чьи обычные доходы и расходы трудно отделить от доходов/расходов бизнеса (например домашние хозяйства имеющие рестораны, мотели или жилье внаем на базе своей жилой площади, а также домашние хозяйства с двумя или больше сотрудниками, проживающими тут же), 4) иностранные домашние хозяйства. А второе это Обследование доходов и расходов сельскохозяйственных домашних хозяйств, которое охватывает сельскохозяйственные домашние хозяйства (за некоторыми исключениями, например, домашние хозяйства с одним членом).

непосредственной семье, будут отличаться³. В случае полигамии многие обычные показатели, которые основываются на принципе главы, имеющего одну супругу/а, может быть сложно сформировать.

(вставить Рисунок 1 здесь)

При анализе получения образования в странах с высоким уровнем доходов используемая информация обычно относится к наивысшему уровню полученной квалификации (то есть какой наивысший диплом об образовании получен). Тем не менее, в странах со средним уровнем образования эту информацию не так легко получить, так как информация об образовании чаще всего имеет отношение к самому высокому классу (или уровню) прослушанного курса/программы, а не обязательно оконченного уровня образования; информация о грамотности также часто является относительной, поэтому трансформация такого рода информации в показатель наивысшего уровня полученной квалификации часто является затруднительной.

Если обратиться к информации с рынка труда, а особенно к *определению трудовой занятости*, следует отметить, что во многих обследованиях в странах со средним уровнем доходов модуль/кластер по рынку труда создан по методу «собираемых» вопросов о трудоустройстве. То есть после получения ответа респондента, что у него нет работы или он нигде не работает, ему задается ряд вопросов с целью выявить, занимается ли он или они какой-либо деятельностью в целях поддержки семьи (например выращивание фруктов и овощей, продажа товаров на улице, предоставление каких-либо услуг другим лицам, помощь по уходу за домом и т.п.). Поскольку во многих странах со средним уровнем дохода большая часть такой деятельности выполняется женщинами и детьми, что рассматривается как трудовая деятельность согласно МОТ (с оговоркой, что они попадают или в разряд оплачиваемой работы в течение минимум одного часа или в разряд неоплачиваемого домашнего труда), это значительно увеличивает уровень трудовой занятости, что создает большой разрыв между населением с чистой выручкой и трудозанятым населением, а также ведет к потенциальной предубежденности по отношению к странам, которые не включают подобные вопросы в свои обследования⁴.

³ Обследования, в которых главой является представитель старшего поколения, покажет непропорциональное количество домашних хозяйств с преимущественно старшим поколением, и наоборот, с очевидными последствиями, в том числе при предоставлении информации о (среднем) уровне образования, на условия трудоустройства и т.п.

⁴ Очень яркий пример этого существует в Перу (по данным Национального обследования домашних хозяйств – ENAHO), которая, наряду со Швейцарией, является страной с самым высоким уровнем трудовой занятости среди всех стран LIS. Как и в обследованиях, проводимых в большинстве стран Латинской Америки, респонденту сначала задается вопрос, есть ли у него работа, и если они отвечают, что нет, его спрашивают о том, выполняет ли он какую-либо деятельность для помощи семье, причем вопрос включает длинный список возможной второстепенной/маргинальной деятельности. Очень большое количество лиц, ответивших отрицательно на первый вопрос, и положительно – на второй, показывает что трудовая занятость в основном состоит из маргинальной трудовой деятельности. Хотя весьма вероятно, что вопрос сформулирован таким образом, чтобы помочь людям, ответившим бы на подобный вопрос отрицательно, ответить на него положительно.

Похожая ситуация касается детского труда. В большинстве стран со средним уровнем доходов информация по рынку труда собирается также и по детям, чтобы проанализировать его распространение. В некоторых обследованиях отдельные разделы посвящены конкретно детям; в других случаях (некоторые из) те же самые вопросы задаются и взрослым и детям. В обоих случаях формирование полностью сравнимого уровня участия в рынке труда (или уровня трудовой занятости) между этими разными странами становится весьма проблематичным. В дополнение к этому обычно существует много вопросов по неофициальной работе, внештатной деятельности, домашнему производству, незаконному труду (не зарегистрированному, не охваченному социальным страхованием, с которого не оплачиваются налоги) с целью хоть в какой-то мере охватить *неформальный труд*. Обычно в зависимости от обследования эти вопросы могут сильно отличаться (часто это зависит от ведомственной организации в стране), и их практически невозможно гармонизировать. Поскольку очень сложно охватить и сравнить размеры неформальной трудовой занятости в рамках разных обследований, один из способов выразить этот показатель, это рассчитать уровень трудоузанности для старшего поколения (предполагая, что старшее поколение в основном работает неформально).

(вставить Рисунок 2 здесь)

И наконец, большое количество лиц, работающих на нескольких работах, дополнительно усложняет гармонизацию (а таким образом и следующее из этого сравнение) характеристик работы, особенно когда в вопроснике разные виды работы разнесены по разным разделам (например, работа в рамках семейного бизнеса отдельно от сельскохозяйственных работ и отдельно от другой работы). Это вызывает беспокойство по нескольким причинам: i) часто не ясно, сообщает ли человек об одной и той же работе в разных разделах вопросника, что создает риск двойного учета; и ii) становится сложно определить, какая работа является основной, а какая второстепенной.

(вставить Рисунок 3 здесь)

Рассматривая переменные дохода, возникает несколько рисков, когда речь идет о странах со средним уровнем доходов. Прежде всего показатели неравенства, бедности и благосостояния все еще в большей своей мере основываются на данных о *потреблении*, а не о *доходах*, что часто ведет к предположению, что микроданные о доходах либо отсутствуют, либо являются недостаточными для того, чтобы рассчитать обоснованные показатели дохода (не собираются; собираются, но не предоставляются; собираются, но в недостаточной мере, чтобы полностью охватить доход домашнего хозяйства)⁵.

⁵ Ситуация с доступностью данных о полном доходе часто неясна, пока не обратишься к микроданным, и это часто приводило к ситуациям, где потенциально длительный процесс приобретения микроданных из вновь включенной страны со средним уровнем доходов приводил к избавлению от данных. Это произошло, например, с данными из Палестины, где данные о доходах существовали на бумаге, но не распространялись из-за их низкого качества, а также с данными из Египта, где информация о доходах от самозанятости попросту не собиралась вообще.

Как уже говорилось выше, включение в учет стран со средним уровнем доходов привело к необходимым коррекциям концепции общего располагаемого дохода и включению в него, в том числе, и немонетарных доходов труда (в частности, собственное потребление продукции, полученной от сельскохозяйственной деятельности), а также общественных и частных трансфертов (конкретно общественная и частная помощь в качестве льгот, предоставляемых в натуре).

(вставить Рисунок 4 здесь)

Поскольку коррекция была необходима для получения более беспристрастной картины условий жизни домашних хозяйств в этих странах, включение этих доходов в данные часто оказывалось особенно проблематичным.

- Первая проблема возникла из-за того, что степень охвата немонетарных доходов, собранных в рамках разных обследований, сильно отличается в разных странах, что приводит к усложнению их сравнения. Например, в обследованиях, которые в основном сосредоточены на потреблении, информация по стоимости большинства товаров и услуг потребленных, но не оплаченных (поскольку или они самостоятельно произведены, или получены от работодателя, государства, благотворительных организаций или других частных домашних хозяйств) собирается очень подробно, в то время как в других видах обследований доступность информации по такой продукции намного более скудная.
- Другая проблема возникла ввиду немонетизации количества товаров и услуг. На данной стадии LIS решил включать только те доходы, которые были уже монетизированы учреждением, предоставляющим эти данные, что снизило потенциальную беспристрастность, поскольку в некоторых странах, ввиду чисто практических, а не концептуальных, причин, определение конечного дохода включает больше немонетарных видов доходов, чем каких-либо других.
- В некотором смысле произвольные предположения должны быть сделаны в случае, когда информация о немонетарных доходах собирается в разных разделах вопросника (с помощью переменных потребления, уровня доходов домашних хозяйств от домашней деятельности, и индивидуального уровня трудовых доходов), из чего становится ясно, что эти суммы в некоторой мере будут пересекаться, а вычисление окончательной суммы, в которой не будет каких-либо задвоений или недосчетов тех или иных источников доходов, оказывается очень сложным.

Независимо от (но в некотором смысле связано с ней) проблемы немонетарных доходов другой проблематичной областью является *доход от самозанятости* в общем, а особенно от сельскохозяйственной деятельности и неформального труда. Поскольку по своей природе такие доходы более нерегулярны, и их сложно измерить, становится весьма сложно обеспечить надежность показателя общего дохода домашнего хозяйства, который по большей части состоит из такого вида доходов. В дополнение к этому, когда эта информация собирается только на уровне домашних хозяйств (что чаще всего и происходит в странах со средним уровнем доходов, где в обследованиях выделен специальный раздел для деятельности домашних хозяйств),

формирование всеобъемлющего измерения общего индивидуального трудового дохода становится невозможным, что в свою очередь ограничивает использование такой важной переменной во многих анализах.

(вставить Рисунок 5 здесь)

Кроме измерения самого дохода, его классификация по различным компонентам дохода также может быть проблематичной в странах со средним уровнем дохода. Одна конкретная проблема имеет отношение к классификации *пенсии и льгот, предоставленных работодателем в рамках трудового дохода в отличие от социального обеспечения*. В то время как льготы, предоставляемые работодателем (такие как пособия и субсидии, выплачиваемые вместе с основной зарплатой), обычно рассматривались как часть трудового дохода в странах с высоким уровнем дохода, в странах со средним уровнем доходов становится ясно, что некоторые из этих льгот на самом деле замещают практически несуществующее социальное обеспечение, и по сути своей намного больше похожи на льготы соцобеспечения, чем на трудовой доход. Подобным же образом, привычное разделение на *социальное страхование, предоставление помощи и универсальную систему пособий* часто теряет свой смысл в странах со средним уровнем дохода, где льготы трудоустройства полностью зависят от желания работодателя, а не от щедрости государства, а большинство, строго говоря, общественных льгот предназначаются наиболее бедным.

Способ учета налогов и взносов по социальному страхованию также отличается в странах с высоким и средним уровнем доходов. Трудности в странах с высоким уровнем дохода сосредоточены на различиях между странами/обследованиями, где предоставляются данные по брутто доходам, то есть вместе с налогами и взносами (так называемый «набор данных брутто»), и теми, где предоставляются данные после вычета налогов и взносов («набор данных нетто»). Более конкретно, в первом случае все доходы предоставляются в брутто, а общая сумма налогов и взносов вычитается из общего брутто дохода, и получается располагаемый доход. С другой стороны в странах, которые предоставляют информацию по доходам уже с вычетом налогов и взносов, сумма всех компонентов дохода уже дана без налогов и взносов, и таким образом соответствует определению располагаемого дохода. В результате этого, хотя на уровне общего располагаемого дохода переменные прекрасно сравниваются, возможность их сравнения снижается тем фактом, что на уровне компонентов некоторые наборы данных предоставляют брутто доходы, а некоторые – нетто доходы. В странах со средним уровнем доходов сложность в учете налогов совсем другая. Низкая зависимость от прямых налогов в большинстве стран со средним уровнем доходов практически лишает вышеуказанные особенности какого-либо смысла, поскольку разница между «брутто» и «нетто» наборами данных очень мала. Даже становится проблематичным просто разделить информацию обследований на брутто и нетто, поскольку в некоторых случаях вместе предоставляется та и другая информация (часто только зарплата указана с налогами и взносами, а все остальное в нетто), а иногда провайдер данных не

предоставляет разделения на брутто и нетто⁶. Несколько стран со средним уровнем дохода даже предоставляют данные по доходам только в брутто выражении, без указания размеров налогов и взносов, выплачиваемых с них, что приводит к необходимости симулятивно рассчитывать налоги и взносы, чтобы дойти до располагаемого дохода, который можно было бы сравнить с другими странами⁷. В любом случае, даже при наличии полной информации по налогам и взносам, низкая зависимость от прямых налогов по сравнению с непрямыми в странах со средним уровнем дохода увеличивает предвзятость к сравнительности показателей благосостояния на основании располагаемого дохода домашнего хозяйства. Если бы в расчет принимались и не прямые налоги, то реальное различие между неравенством в странах с высоким и средним уровнем доходов могло бы быть еще более разительным, чем то, что нам сейчас показывают цифры.

(вставить Рисунок 6 здесь)

С более технической точки зрения, сложности, которые часто могут становиться серьезными, особенно что касается данных из стран со средним уровнем дохода (но не обязательно только этих стран) заключается в наличии большого количества наблюдений, где отсутствуют (или разнородные) данные. Когда процент домашних хозяйств, где не хватает (или отсутствует) информации по общему располагаемому доходу, перейдет определенный порог⁸, и особенно когда учреждения, предоставляющие данные, не принимают этого во внимание при расчете весов, потенциальное предубеждение, вызванное неслучайным распределением таких домашних хозяйств, становится достаточно большим, и может поставить под угрозу любой анализ распределения доходов на уровне целой страны⁹.

(вставить Рисунок 7 здесь)

В общем, не смотря на усилия, приложенные на разных уровнях цепочки производства данных (концепция обследования, внедрение, обработка данных и гармонизация данных), все еще остаются серьезные пробелы, которые не позволяют обеспечить идеальной взаимосогласованности наборов микроданных по доходам, поступающих из стран с высоким и средним уровнем доходов. Весьма важным остается вопрос, возможно ли проанализировать эти два набора данных в рамках одной и той же концептуальной схемы, или их следует рассматривать отдельно. LIS занял позицию, что общая концептуальная схема возможна, но при этом не может

⁶ Такова например ситуация в Индии, где, принимая во внимание практически незначительность налогов, учреждения, предоставляющие данные, не уточняют следует ли учитывать доход до или после отчисления налогов, поэтому окончательный результат представляет собой смесь и того и другого в зависимости от наблюдения, без какого-либо указания на то, как должно проводиться каждое наблюдение.

⁷ Такова ситуация в Национальном выборочном обследовании домашних хозяйств (RNAD) в Бразилии, в Большом колумбийском интегрированном обследовании домашних хозяйств (GEIH) и Непрерывающемся обследовании домашних хозяйств (ECH) в Панаме.

⁸ В LIS обычно используется порог в 10%, который требует внимательного учета, а 20% это максимально приемлемый порог для надежных расчетов доходов.

⁹ Например в Тунисе. LIS получил данные из единственного существующего источника микроданных о доходах (Тунисское панельное обследование трудового рынка – TLMPS). После детального рассмотрения этих данных LIS принял решение не включать их в свою базу данных, так как очень большая доля выборки не содержала информации о доходе домашнего хозяйства (примерно половина всех домашних хозяйств).

не подчеркнуть важность того внимания, которое необходимо уделить всем предупреждениям, сопутствующим такому подходу.