

# Выявление факторов неравенства в России с использованием регрессионной декомпозиции

**Антон Игоревич Вотинов**, младший научный сотрудник Центра макроэкономических исследований НИФИ Минфина России (г. Москва)

E-mail: avotinov@nifi.ru, ORCID 0000-0002-2972-8498

**Самвел Сергеевич Лазарян**, руководитель Центра макроэкономических исследований НИФИ Минфина России (г. Москва)

E-mail: lazaryan@nifi.ru, ORCID 0000-0001-6829-3979

**Вячеслав Николаевич Овчинников**, ассистент Института экономики и предпринимательства Нижегородского государственного университета им. Н. И. Лобачевского (г. Нижний Новгород)

E-mail: vyacheslav\_ovchinnikov\_1993@mail.ru, ORCID 0000-0001-9786-3299

## Аннотация

Настоящее исследование посвящено выявлению детерминант уровня неравенства в России на основе данных РМЭЗ НИУ ВШЭ, охватывающих период 2009–2016 гг. Под неравенством подразумевалась дифференциация в эквивалентном доходе домохозяйств с учетом стоимости потребительской корзины. Для решения поставленной задачи была применена регрессионная декомпозиция индекса неравенства. Согласно полученным результатам переменные «доля детей в д/х», «размер д/х» и «доля неработающих взрослых» привносили наибольший положительный вклад в индекс неравенства. Переменная «регион проживания», которая во многих исследованиях определяется основным фактором неравенства, оказывала выравнивающее влияние. Связано это с тем фактом, что в работе анализировалась дифференциация в уровне жизни, то есть зависимая переменная дохода корректировалась на уровень цен в соответствующем регионе. Таким образом, на уровень неравенства влиял не столько региональный признак сам по себе, сколько дифференциация других значимых факторов по регионам. Полученные результаты могут быть впоследствии использованы для более глубокого изучения причин изменения неравенства.

**Ключевые слова:** доходное неравенство, домохозяйство, регрессионная декомпозиция, квартиль

**JEL:** D31, C51

**Для цитирования:** Вотинов А. И., Лазарян С. С., Овчинников В. Н. Выявление факторов неравенства в России с использованием регрессионной декомпозиции // Финансовый журнал. 2019. № 5. С. 74–89. DOI: 10.31107/2075-1990-2019-5-74-89

## **ВВЕДЕНИЕ**

Проблема социально-экономического неравенства обсуждается уже не одним поколением экономистов. В России данный вопрос стал особенно острым ввиду кризисных явлений последних лет, при которых происходило реальное ухудшение уровня жизни населения и перераспределение доходов в пользу верхних квантилей.

Снижение уровня неравенства важно с той точки зрения, что оно имеет достаточно прочную связь с уровнем экономического роста. Исторически наблюдалась отрицательная зависимость между уровнем неравенства и долгосрочным уровнем конкурентоспособности [Джомо К. С., Попов В. В., 2016]. По мнению авторов, неравенство приводит к росту социальных рисков, что в конечном счете имеет отрицательное влияние на экономическое развитие. Хотя в недавнем обобщенном анализе эмпирических работ на предмет корреляции неравенства и экономического роста [Лазарян С. С., Черноталова М. А., 2017] показано, что эта зависимость неоднозначна, отрицательные связи все же преобладают.

Вопрос декомпозиции факторов неравенства доходов России является важным. Выявление причин неравенства позволит сформировать «группы риска», что в свою очередь поможет разработать оптимальную и комплексную стратегию по снижению расслоения общества, избегая социально чувствительных реформ фискального блока. Основная задача настоящего исследования — использовать количественные методы анализа (регрессионную факторную декомпозицию), которые позволят выявить наиболее значимые факторы неравенства. В ходе анализа особое внимание будет уделено методологии исследования и его преимуществам по отношению к групповым методам декомпозиции неравенства.

Работа построена следующим образом. Сначала будет дан краткий обзор существующих исследований о факторах доходного неравенства в России и за рубежом. Далее будет описана используемая база данных и методологические подходы к измерению и декомпозиции неравенства. На следующем этапе работы будет реализована факторная декомпозиция индекса неравенства на основе регрессионного анализа.

## **ОБЗОР ИССЛЕДОВАНИЙ**

Обращаясь к иностранным исследованиям, посвященным изучению причин доходной дифференциации, отметим широту методологического инструментария, положенного в их основание. Это может быть функциональное разложение индексов неравенства (коэффициент Джини, коэффициент вариации, энтропийные индексы) по источникам совокупного дохода (заработная плата, социальные трансферты и др.). В частности, речь идет о декомпозиции [Shorrocks A. F., 1982]. Впоследствии декомпозиция [Shorrocks A. F., 1982] нашла свое естественное «эконометрическое» продолжение в работах [Morduch J., Sicular T., 2002; Fields G. S., 2003]. В этих работах был предложен эконометрический подход к декомпозиции доходных различий, позволяющий оценить совместный вклад разнообразных факторов в дифференциацию доходов населения. Остается академически востребованной и декомпозиция по групповому признаку [Shorrocks A. F., 1984]. Суть ее сводится к выделению некоторого признака (к примеру, уровня образования главы домохозяйства), согласно которому происходит разбиение выборки на строго определенные группы, затем общее неравенство раскладывается на внутригрупповую и межгрупповую компоненты. Здесь же следует сразу упомянуть и об ограничениях групповой декомпозиции, в частности, речь идет о применимости только в отношении категориальных переменных и невозможности оценки изолированного вклада того или иного фактора.

Отметим, что применение эконометрического подхода к декомпозиции доходных различий не является новым. Ранее схожая методология была предложена Р. Оаксакой [Oaxaca R., 1973]. Ее особенность сводится к изучению различий в доходах двух групп населения (в работе [Oaxaca R., 1973] речь шла о мужчинах и женщинах). На первом шаге оцениваются уравнения дохода для каждой из групп, а на втором шаге находится их разность. Таким образом, удается ответить на вопрос, какие факторы влияют на дифференциацию доходов в большей или меньшей степени. Причем факторы собираются в агрегированные эффекты состава и рыночной оценки. Если эффект состава описывает различия в наделенности каждой из групп тем или иным фактором, то эффект рыночной оценки свидетельствует о различиях в отдаче на тот или иной фактор (так называемая дискриминация). Подход Оаксаки и Блайндера [Blinder A., 1973]. нашел продолжение в работах [DiNardo J. et al., 1996; Machado J., Mata J., 2005; Firpo S. P. et al., 2018], когда декомпозиция осуществлялась для каждой доходной квантили и других выборочных статистик.

Говорить о взаимозаменяемости описанных выше регрессионных типов декомпозиции не было бы справедливым. Скорее напротив, каждый из них позволяет исследовать проблему дифференциации доходов с разных сторон, формируя более глубокое представление о происходящих процессах. В настоящем исследовании будет использована первая группа подходов к декомпозиции по [Shorrocks A. F., 1982; Shorrocks A. F., 1984], так как она позволяет сделать наиболее релевантные относительно заданной проблемы выводы.

В первую очередь необходимо остановиться на двух работах, предложивших эконометрическое расширение классической функциональной декомпозиции [Shorrocks A. F., 1982]. Во-первых, это работа [Fields G. S., 2003], где автор попытался понять причины резкого роста неравенства заработных плат в американской экономике в период 1980–2000 гг., придя к заключению, что ключевым фактором такого роста стало стремительное увеличение отдачи на отдельные категории образования трудящихся. Во-вторых, это исследование [Morduch J., Sicular T., 2002], где речь шла о декомпозиции доходных различий жителей сельских поселений Китая. Авторы пришли к несколько неоднозначным результатам, поскольку разложение разных мер неравенства приводило к разным практическим выводам. Так, образование и структурные особенности домохозяйств оказывали понижающий эффект на неравенство, если речь шла о декомпозиции индекса Тейла, и противоположный эффект, если речь шла о декомпозиции индекса Джини. Единственным фактором, который во всех случаях существенно увеличивал доходную дифференциацию, стал географический фактор (принадлежность к разным сельским поселениям).

Выводы, полученные в работе [Fields G. S., 2003], перекликаются с результатами по другим развитым экономикам. В частности, в работе [Brewer M., Wren-Lewis L., 2016] исследовались причины интенсивного роста доходного расслоения британских домохозяйств в период 1978–1991 гг. и его последующей стагнации (1991–2008 гг.). По мнению авторов, поводом для ускорения дифференциации в доходах стали растущие различия в заработной плате, произошедшее же затем замедление связывалось ими с более выраженным выравнивающим эффектом налогов на трудовые доходы и снижением различий в остальных источниках дохода (к примеру, в инвестиционном доходе). Помимо этого авторы подробно остановились на ключевых факторах дифференциации доходов (если быть точнее, дифференциации в располагаемых эквивалентных доходах — *OECD-modified scale*) на протяжении всех последних 40 лет (1978–2018 гг.). Если обратиться к первому временному интервалу (1978–1991 гг.), то неравенство здесь происходило по причине различий в уровне образования и статусе занятости взрослых членов домохозяйства. Отдача на образование претерпела значительный рост, а доходы занятых и самозанятых росли опережающими темпами относительно остальных групп граждан.

В последующие годы наиболее влиятельным фактором стала демографическая компонента, описывающая структурные особенности домохозяйства.

У. Рани и М. Фуррер попытались разобраться в причинах дифференциации доходов на страновой выборке большей размерности, а именно в 13 странах G20 [Rani U., Furrer M., 2016]. Авторами было продемонстрировано, что именно дифференциация трудовых доходов (под ними подразумевалась сумма их заработной платы и доходов от самозанятости) порождала общую дифференциацию доходов домохозяйств, в то время как налоги на труд и выплаты на социальное страхование оказывали выравнивающее воздействие. На следующем шаге авторы обратились к индивидуальным опросникам и задействовали декомпозицию, основанную на уравнении регрессии. Это позволило им выделить ключевые причины неравенства в самих трудовых доходах, которыми стали уровень образования и интенсивность занятости индивида (постоянная/временная). В рамках исследования было к тому же показано, что в отдельных странах существенный вклад в неравенство приходился на половозрастной фактор.

Ранее о влиянии фактора гендера на дифференциацию доходов уже упоминалось в работе [Manna R., Regoli A., 2012] по итальянской экономике (1998–2008 гг.). Авторы, используя уравнение регрессии для разложения различных мер неравенства (коэффициент Джини, индекс Тейла и др.), пришли к выводу о значительной дискриминации женщин на рынке труда Италии. Кроме того, важными факторами неравенства были названы уровень образования и накопленное богатство.

Анализ опыта исследования проблемы дифференциации доходов в развивающихся странах также демонстрирует значимую роль образовательного фактора в ее формировании. Об этом упоминалось в работе Э. Викаксоно. Применяя регрессионную методологию, автор пришел к заключению, что дифференциация доходов в Индонезии становится следствием различий в продолжительности обучения главы домохозяйства. В частности, это объясняло около 20 % мер неравенства (индексы Джини, Тейла, Аткинсона). Связывалось происходящее с тем, что в обеспеченных домохозяйствах у детей имелось сравнительно больше шансов на получение качественного образования и, соответственно, более высокой величины трудового дохода в будущем [Wicaksono E. et al., 2017]. В других работах по китайской экономике авторы подробно останавливались на постижении причин разнящихся доходов жителей городских и сельских поселений. К примеру, в работе [Sicular et al., 2007] было показано, что дифференциация доходов становится следствием различий в уровне образования главы домохозяйства.

В целом работы российских исследователей на предмет декомпозиции доходных различий домохозяйств описывают схожий контур факторов, но имеются и расхождения в оценках. Особенно это касается переходного периода или периода рыночных реформ для российской экономики. Любопытным и довольно ранним исследованием, посвященным анализу детерминант доходного неравенства населения России, является работа О. Кислицыной [Кислицына О. А., 2003]. В частности, автор, используя данные РМЭЗ НИУ ВШЭ за период с 1994 по 2000 г., пришла к заключению, что на протяжении всех исследуемых лет ключевыми характеристиками уровня дифференциации доходов российских домохозяйств (речь шла об эквивалентных совокупных доходах) стали доля занятых в его структуре, регион проживания и уровень образования главы домохозяйства. При этом начиная с 1998 г. вес региона проживания постепенно снижался, роль образовательного фактора на протяжении всех лет только усиливалась, а влияние фактора доли занятых в структуре домохозяйства не имело выраженного тренда. Автор остановилась и на проблеме дифференциации доходов российских домохозяйств по источнику их формирования. Как и следовало ожидать, оплата труда играла доминирующую роль в дифференциации доходов, что связывалось с ее подавляющей долей в общей композиции доходов.

Еще одной работой, посвященной изучению причин неравенства в переходной экономике России, стало исследование [Jansen W. et al., 2013]. Авторы, отталкиваясь от выборочных данных 15 различных опросов домохозяйств, пришли к несколько иному пониманию ключевых факторов доходного неравенства. В частности, ими было показано, что продолжительность обучения главы домохозяйства не оказывала значимого влияния на дифференциацию доходов, в то время как переменные «статус занятости (работающий/безработный)» и «принадлежность к самозанятым» ощутимо влияли на доходное расслоение. Вероятно, такая рассогласованность результатов могла быть связана с различием используемых методологических приемов и положенных в основу исследования статистических данных. Все дело в том, что в числе широкого многообразия опросов не нашлось места волнам опросов РМЭЗ НИУ ВШЭ, ставшим уже традиционным массивом статистических данных для отечественного исследователя. С другой стороны, авторы использовали несколько нестандартный регрессионный подход, моделируя не значение дохода (и потом переходя к декомпозиции), а изначально меру неравенства — дисперсию лог-дохода.

Важной работой, посвященной проблеме дифференциации в оплате труда в российской экономике, стало исследование А. Лукьяновой. Во-первых, автором был зафиксирован тренд на сокращение различий в оплате труда за счет подтягивания заработков низкоквалифицированных работников. Во-вторых, используя декомпозицию [Fields G. S., 2003], автор пришел к выводу, что на протяжении всего временного периода (1998–2005 гг.) в основе дифференциации оплаты труда лежали ее межрегиональные диспропорции, а также различия в обладании отдельными характеристиками человеческого капитала (профессия и образование) [Лукьянова А. Л., 2007].

В более современной работе Л. Н. Овчаровой и соавторов [Овчарова Л. Н. и др., 2016] оценивался одновременный вклад факторов внешнего окружения (регион проживания), характеристик человеческого капитала и иждивенческой нагрузки в общую дисперсию среднедушевых денежных расходов населения России с помощью регрессионного подхода. Использование категории расходов связывалось с тем, что она в большей степени свидетельствует о реальном уровне жизни домохозяйств, а также позволяет учесть высокую долю теневого сектора в российской экономике. Ключевым фактором был назван уровень образования членов домохозяйства, вес которого составлял более 30 % дисперсии душевых расходов на протяжении всех анализируемых лет (1994–2014 гг.).

Таким образом, можно попытаться выделить некоторую общность факторов, лежащих в основе доходного расслоения как в развитых, так и развивающихся странах. В первой группе стран оно становится следствием различий в обладании человеческим капиталом, а также может быть отнесено к дифференциации в особенностях трудовой занятости. Во второй группе стран природа факторов несколько шире, важную роль играют в том числе географические детерминанты и композиционные особенности домохозяйств.

## **ДАННЫЕ, МЕТОДОЛОГИЯ И РЕЗУЛЬТАТЫ**

Для нужд настоящего исследования были использованы панельные данные РМЭЗ ВШЭ по индивидам и домохозяйствам с 13-й по 25-ю волн, охватывающие 2004–2016 гг. Выбор волн связан с отсутствием некоторых ключевых переменных в более ранних выборках. Необходимо отметить, что используемая для анализа выборка является несбалансированной, что необходимо учитывать при описании полученных результатов. Под домохозяйством понимается множество «всех людей, живущих вместе... и имеющих общие доходы и расходы». Также в домохозяйства включены дети до 18 лет, не живущие с семьей из-за учебы в другом городе.

Практически в каждом домохозяйстве проводились как семейные, так и индивидуальные опросы. Наличие индивидуальных опросников позволяет выявить «главу домохозяйства» (далее также — глава д/х), который определяется в настоящей работе как индивид с максимальным доходом. В том случае, если выявить такого индивида не представляется возможным, под «главой домохозяйства» понимается тот индивид, который заполнял опросник по домохозяйствам. Наблюдения, для которых «глава д/х» не принимал участия в индивидуальном опросе, исключены из выборки (менее 0,1 % от первоначального массива данных).

Итоговая база данных имеет около 38 тыс. наблюдений, чуть менее 3 тыс. наблюдений на каждую волну. Часть данных была исключена из итоговой выборки из-за пропущенных значений по тем или иным переменным. В целом размер анализируемого массива наблюдений позволяет получить достаточно точные оценки с точки зрения статистических характеристик.

Также отметим, что поскольку в настоящей работе исследуется неравенство на уровне домохозяйств, используется понятие неравенства по эквивалентному доходу, который находится через отношение совокупного дохода домохозяйства к квадратному корню из количества его членов (*square root scale*). Данный подход предполагает, что при увеличении размера домашнего хозяйства необходимый для поддержания такого же уровня жизни доход должен также вырасти, но из-за экономии на масштабе в меньшей степени. Более того, для исключения влияния различий в уровнях цен делается поправка на прожиточный минимум в регионе (с учетом структуры домохозяйства). Таким образом, в настоящем исследовании оцениваются факторы, влияющие на дифференциацию в уровне жизни населения.

### Показатели неравенства, их динамика

Дифференциация доходов населения характеризуется специальными показателями. К настоящему времени насчитывается несколько десятков показателей неравенства. Одни показатели стали широко популярны благодаря своей наглядности и относительной простоте вычисления, а другие — благодаря разнообразию математических свойств. При этом простые и наглядные показатели часто оказываются неудобными с точки зрения математических свойств.

Помимо базовых показателей оценки уровня неравенства населения, коэффициента фондов и коэффициента Джини, которые не будут описаны в этой работе, существует семейство энтропийных индексов. В работе [Hill M. O., 1973] было предложено использование понятия обобщенной энтропии для оценки неравенства. Класс обобщенных энтропийных мер неравенства описывается следующей формулой:

$$I_{\theta} = \frac{1}{\theta^2 - \theta} \left[ \frac{1}{n} \sum_{i=1}^n \left( \frac{y_i}{\bar{y}} \right)^{\theta} - 1 \right], \quad (1)$$

где  $n$  — количество индивидов в выборке;  $y_i$  — доход индивида  $i$ ; некоторый параметр  $\theta$ ;  $\bar{y}$  — средний доход. Меры  $I_{\theta}$  с параметрами 0 и 1 являются средним логарифмическим отклонением  $I_0$  и индексом Тейла  $I_1$ .

Параметр  $\theta$  определяет вес различных хвостов распределения доходов населения. Например, среднее логарифмическое отклонение придает больший вес различиям изучаемого признака в нижней части распределения, а индекс Тейла придает больший вес правому хвосту распределения. Основной недостаток энтропийных индексов состоит

в неочевидной интерпретации полученных абсолютных значений. С другой стороны, данные индексы обладают удобными декомпозиционными свойствами.

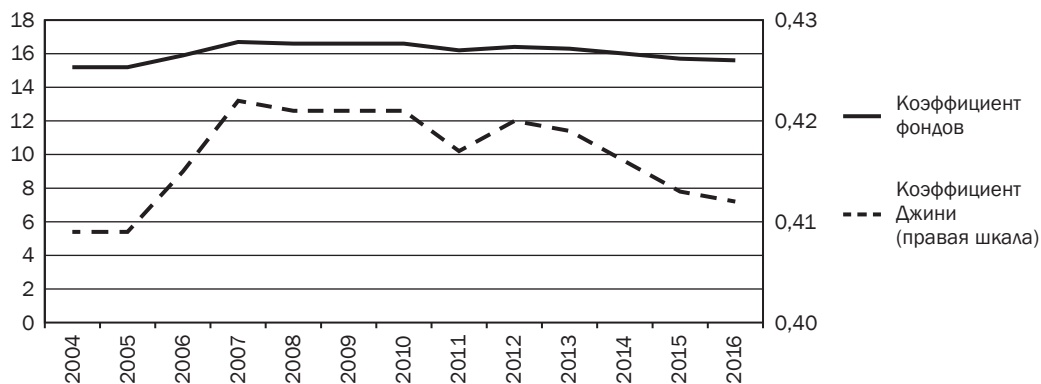
Индекс Аткинсона — еще один способ измерения уровня неравенства. Данный индекс отражает, насколько среднее значение доходов отклоняется от некоторого эквивалентного дохода, при котором уровень общественного благосостояния сохранился бы на прежнем уровне. В каком-то смысле данный показатель отражает, сколько в относительных величинах общество могло бы сэкономить для получения того же благосостояния.

В данной работе акцент сделан на индексе Тейла, который равен индексу обобщенной энтропии с параметром единица. Выбор данного индекса связан с тем, что он обладает рядом удачных декомпозиционных свойств. Более подробно про свойства различных показателей неравенства можно найти в работе [Cowell F. A., 1989; Cowell F. A., Victoria-Feser M. P., 1996].

В целом динамика индекса Джини и коэффициента фондов не характеризовалась наличием преобладающего тренда на протяжении всех анализируемых лет — 2004–2016 гг. Тренд скорее был кусочно-линейный. Так, в периоды интенсивного роста экономики страны (с 2004 по 2008 г.) и посткризисного восстановления (с 2011 по 2013 г.) дифференциация доходов населения увеличивалась. Вместе с тем сокращение доходных различий отмечалось в период мирового экономического кризиса 2008 года. Стагнация последних лет также привела к снижению дифференциации доходов. С одной стороны, это было связано с «подтягиванием» доходов нижних квантилей и перераспределительной ролью государства. С другой стороны, неравенство сокращалось и по причине негативных изменений в компонентах совокупного дохода обеспеченных граждан страны — номинальное «сжатие» доходов от предпринимательской деятельности и оплаты труда (рис. 1).

Рисунок 1

Динамика индексов неравенства /  
Dynamics of inequality indices



Источник: данные Федеральной службы государственной статистики / Source: data from the Federal State Statistics Service.

### Декомпозиция неравенства с использованием регрессионного подхода

При использовании групповой декомпозиции невозможно изолировать влияние одного фактора на неравенство от другого. Для этих целей предлагается использовать регрессионный подход, изложенный в работе [Morduch J., Sicular T., 2002], который является расширением стандартной декомпозиции [Shorrocks A. F., 1982]. В общем виде декомпозиция доходов по [Shorrocks A. F., 1982] может быть описана следующим образом:

$$Y = \sum_k Y^k, \quad (2)$$

где  $Y$  — совокупный доход индивида, а  $Y^k$  —  $k$ -я компонента/фактор.

Основным предположением декомпозиции неравенства является тот факт, что каждый отдельный фактор может иметь разное влияние на общее неравенство. Так, например, доход от капитала распределен гораздо менее равномерно, чем доход от трудовой деятельности, что приводит к тому, что именно гетерогенность капитальных доходов в большей степени влияет на общее неравенство. В общем виде декомпозиция индекса неравенства по факторам может быть представлена как:

$$I(Y) = \sum_k S^*(Y^k), \quad (3)$$

где функция  $S^*(Y^k)$  отражает абсолютный вклад  $k$ -го фактора в общее неравенство. В работе [Shorrocks A. F., 1982] автор продемонстрировал, что такое разложение индекса неравенства можно выразить через специальным образом взвешенные компоненты дохода:

$$s^k(Y) = \frac{\sum_{i=1}^n a(y_i) y_i^k}{I(Y)}, \quad (4)$$

где  $s^k(Y)$  — относительный (нормированный на величину индекса) вклад  $k$ -го фактора в общее неравенство. В данном выражении функция  $a(\cdot)$  зависит от выбора того или иного индекса неравенства. Например, для индекса Тейла  $I_1$  разложение принимает вид:

$$s^k(Y) = \frac{\frac{1}{n} \sum_{i=1}^n \frac{y_i^k}{\mu} \ln\left(\frac{y_i}{\mu}\right)}{\frac{1}{n} \sum_{i=1}^n \frac{y_i}{\mu} \ln\left(\frac{y_i}{\mu}\right)}. \quad (5)$$

Описанный подход к декомпозиции неравенства предполагает наличие очевидной факторизации изучаемого показателя. Обобщением факторной декомпозиции индекса неравенства является подход, основывающийся на регрессионном анализе. В настоящем исследовании будет использован подход, который был детально описан в работе [Morduch J., Sicular T., 2002]. Регрессионный подход предполагает, что изучаемый показатель может быть разложен на некоторые факторы с помощью стандартной эконометрической модели:

$$Y_i = X_i \beta + \varepsilon_i. \quad (6)$$

Между приведенной эконометрической моделью (6) и описанной выше декомпозицией по факторам (2) существует прямая зависимость, которая может быть выражена следующим образом:

$$\widehat{\beta}_k X_{k,i} = Y^k, k = \overline{1, K}, \varepsilon_i = Y^{k+1}. \quad (7)$$

Таким образом, после оценки эконометрической модели можно применить «натуральную декомпозицию» с использованием индекса неравенства  $I(Y)$ . Тогда нормированный вклад  $k$ -го фактора в общее неравенство может быть описан с помощью формулы:

$$s^k(Y) = \widehat{\beta}_k \left( \frac{\sum_{i=1}^n a(y_i) x_i^k}{I(Y)} \right). \quad (8)$$



Можно выделить как минимум три существенных преимущества регрессионного подхода к определению факторов неравенства. Во-первых, можно оценивать изолированное влияние факторов на неравенство, чего невозможно качественно сделать в рамках группового подхода к декомпозиции. Во-вторых, регрессионный анализ позволяет использовать ряд процедур, которые в значительной степени улучшают качество получаемых результатов. Например, возможно использование поправок при нарушениях теоремы Гаусса — Маркова. Наконец, эконометрический подход дает возможность оценить вклад факторов, не связанных прямой зависимостью с изучаемым показателем. Так, можно оценить влияние фактора образования на неравенство по доходам, чего нельзя сделать, используя предложенный в [Shorrocks A. F., 1982] подход.

В рамках настоящего исследования будет проведена регрессионная декомпозиция индекса неравенства Тейла по доходам. Для этих целей будет оценена модель зависимости логарифма эквивалентного дохода от целого набора вспомогательных характеристик домохозяйства: территориальной принадлежности, демографических характеристик и характеристики главы домохозяйства (в соответствии с работой [Wicaksono E. et al., 2017], где в качестве факторов используются характеристики именно главы, как наиболее влияющего на благополучие домохозяйства в развивающихся странах). Последний набор включает в себя уровень образования, возраст, пол, официальный статус трудоустройства (0 = официальное трудоустройство) и прочие характеристики, которые обычно связаны с доходами домохозяйства.

Существенным вопросом является интерпретация полученных при декомпозиции результатов. При используемой регрессионной декомпозиции естественнее под вкладом  $k$ -го фактора подразумевать изменение индекса неравенства при исключении этого фактора из показателя, то есть  $S^k(Y) = I(Y) - I(Y - Y^k)$ . Отметим, что если некоторый показатель распределен равномерно, то его вклад в неравенство будет отрицательный, так как исключение уравнивающего всех фактора приведет к росту дифференциации.

В табл. 1 приведены результаты оценки уравнения дохода и производные величины. Эталонные значения для переменных «отрасль трудоустройства главы д/х», «регион проживания» и «тип населенного пункта» выбраны таким образом, чтобы все коэффициенты имели положительное значение. Данная мера была принята для корректного расчета поквартильного распределения показателей, представленного в столбце «Квартили» (методология расчета будет представлена далее). Последние два столбца таблицы отражают степень неравномерности<sup>1</sup> распределения доходов домохозяйств по соответствующим показателям.

Таблица 1

**Регрессионная декомпозиция индекса доходного неравенства домохозяйств /  
Regression-based decomposition of the household income inequality index**

Показатель	Коефф.	t-статистика	Квартили				Отношение верхнего и нижнего квартилей	Степень неравномерности
			1	2	3	4		
Отрасль трудоустройства главы д/х	>0	-	23,8	25,0	25,5	25,6	1,07	0,08
Регион проживания	>0	-	21,6	24,8	25,7	27,8	1,29	0,80
Тип населенного пункта	>0	-	19,8	24,6	27,0	28,7	1,45	1,78

<sup>1</sup> Для простоты степень неравномерности рассчитывается как статистика критерия согласия Пирсона,  $q_j$  — значение для соответствующего квартиля:  $\chi^2 = \sum_j \left( \frac{q_j - 25}{25} \right)^2$ .

Показатель	Коэфф.	t-статистика	Квартили				Отношение верхнего и нижнего квартилей	Степень неравномерности
			1	2	3	4		
Размер д/х	-0,25	-89,4	31,8	27,5	23,3	17,5	0,55	4,43
Доля детей	-0,25	-16,9	26,5	26,9	24,7	21,9	0,82	0,64
Доля пенсионеров	0,08	2,6	36,0	28,2	19,7	16,2	0,45	9,47
Доля инвалидов	-0,17	-7,3	36,5	26,8	19,0	17,7	0,49	8,96
Уровень образования	0,04	41,1	23,8	24,6	25,1	26,4	1,11	0,15
Пол	0,10	12,7	24,5	27,6	27,0	20,9	0,86	1,09
Возраст	-0,004	-2,0	24,6	24,5	25,0	25,9	1,05	0,04
Возраст (квадрат)	-3E-5	-1,7						
Официальный статус трудоустройства	-0,14	-10,2	33,1	26,9	23,1	17,0	0,51	5,49
Доля неработающих взрослых членов д/х	-1,10	-63,7	41,4	29,3	19,5	9,9	0,24	21,85

Примечание: серым закрашены ячейки коэффициентов, значимых на 5-процентном уровне значимости. Также сумма по строкам для квартилей может не равняться 100 из-за округления. / Note: coefficients in grey cells are significant at a 5 percent level. Also, the row amount for quartiles may not equal 100 due to rounding errors.

Источник: расчеты авторов на основе данных РМЭЗ НИУ ВШЭ / Source: authors' calculations based on data from the RLMS of the Higher School of Economics.

Используя результаты регрессии, можно рассчитать дополнительную описательную статистику, отражающую распределение изучаемых факторов среди квартилей с учетом их влияния на уровень доходов. Столбцы с квартилями были рассчитаны с применением следующего алгоритма. На первом шаге все домохозяйства ранжируются по уровню эквивалентного дохода. Далее для квартиля  $G$  и показателя  $j$  рассчитывается величина:

$$\frac{\sum_{i \in G} \hat{\beta}_j x_{j,i}}{\sum_i \hat{\beta}_j x_{j,i}} \quad (9)$$

где  $x_{j,i}$  — значение  $j$ -го показателя для  $i$ -го домохозяйства;

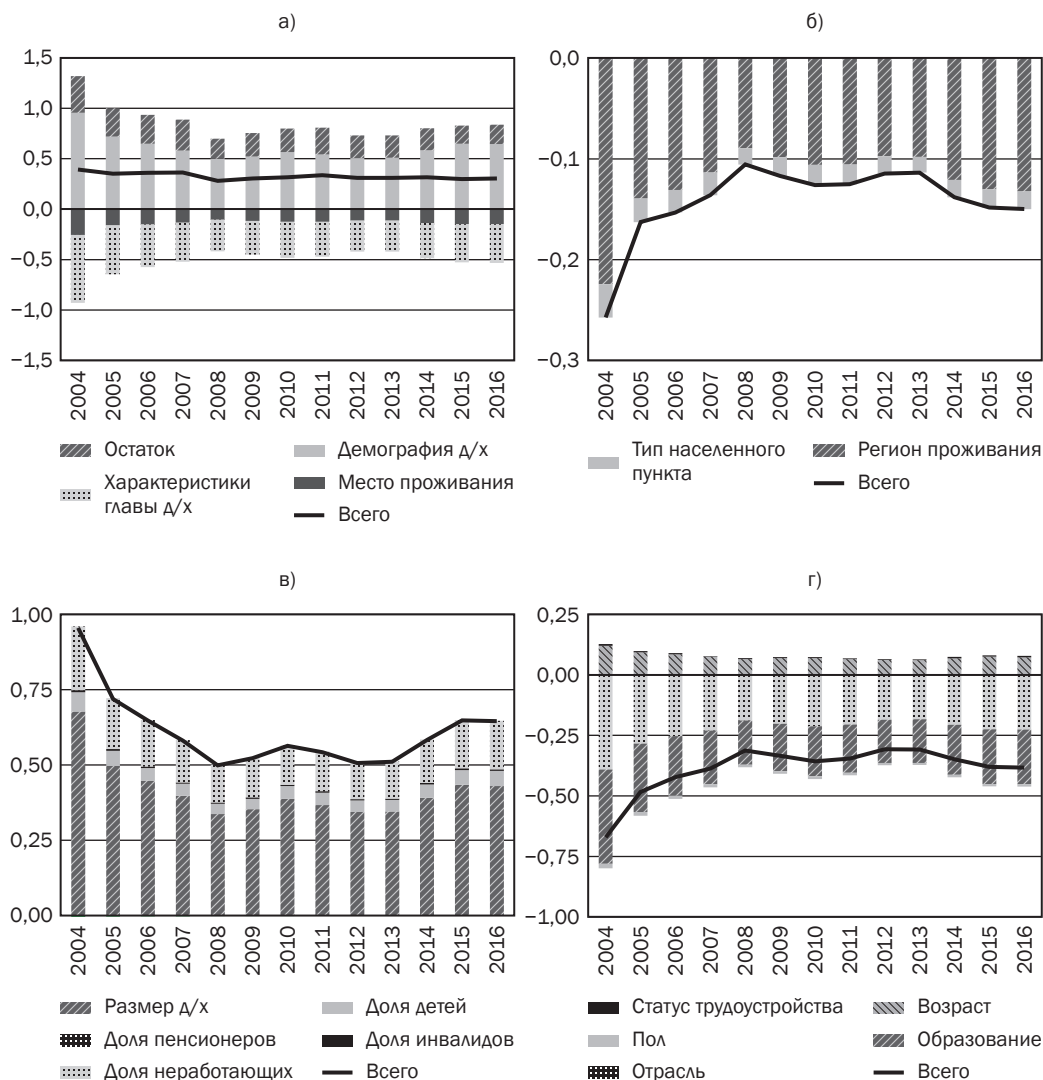
$\hat{\beta}_j$  — оценка коэффициента перед соответствующим показателем.

Данная формула может быть интерпретирована как доля дохода, которая приходится на соответствующий квартиль и объясняется исключительно фактором  $j$ . Так, наиболее простая интерпретация данного показателя характерна для бинарных переменных. Для переменной «пол», которая принимает значение 1, если глава домохозяйства мужчина, изучаемый показатель является долей мужчин, приходящейся на соответствующий квартиль. Для первого квартиля характерна наименьшая доля мужчин, а для третьего квартиля — наибольшая доля мужчин. В целом неравенство по этому показателю находится на достаточно низком уровне.

Интерпретация данных показателей для небинарных и непрерывных переменных несколько сложнее. Интуиция данного показателя такова — чем выше значение, приходящееся на квартиль, тем больше «концентрация» этого фактора с учетом его влияния на доход. Такое представление не является итоговым результатом регрессионной декомпозиции, но может дать дополнительное представление о влиянии факторов на неравенство.

На рис. 2 приведен результат применения регрессионной декомпозиции неравенства для индекса Тейла (ненормированный, на основе формул (3) и (8)). Наибольший вклад в неравенство характерен для фактора «демография д/х». Другие компоненты, а именно «место проживания» и «характеристика главы д/х», имеют сглаживающее неравенство влияние.

**Результаты регрессионной декомпозиции уровня неравенства по (а) укрупненным группам переменных, (б) по характеристике места проживания д/х, (в) по демографическим характеристикам д/х, (г) по характеристикам главы д/х / Results of regression-based decomposition of inequality index for (a) aggregated groups of variables, (b) according to the characteristics of the place of household residence, (c) according to the demographic characteristics of household, (d) according to the characteristics of the household head**



Примечание: некоторые факторы имеют незначительный вклад в неравенство, поэтому они не отображаются на рисунках. / Note: some factors are not shown in the figures as they have a negligible contribution to inequality. Источник: расчеты авторов на основе данных РМЭЗ НИУ ВШЭ / Source: authors' calculations based on data from the RLMS of the Higher School of Economics.

Динамика компонент индекса неравенства Тейла, приведенная на рисунке, свидетельствует о том, что наибольший вклад в неравенство характерен для группы переменных

«демография  $d/x$ », причем он оставался в некоторой степени стабильным во времени. Выравнивающий эффект на дифференциацию в доходах оказывали переменные человеческого капитала и места проживания домохозяйства. Также отметим наличие определенного вклада переменных группы «остаток», наличие которого связано с техническими особенностями.

Для более глубокого понимания тех или иных процессов, влияющих на неравенство, необходимо проанализировать отдельные компоненты индекса.

Декомпозиция компоненты «демографическая нагрузка» позволяет сделать вывод о том, что наиболее значимый вклад в рост неравенства характерен для показателя «размер  $d/x$ », который объясняет более половины всего неравенства, связанного с демографией. Также существенный вклад в неравенство характерен для показателей «доля неработающих взрослых членов  $d/x$ » и «доля детей». Все остальные факторы, хотя и вносят определенный вклад в неравенство, не оказывают на него значительного влияния.

Переменные, отражающие характеристики главы домашнего хозяйства, снижают уровень неравенства. Связано это с тем, что данные характеристики, несмотря на значимое влияние на уровень дохода домашнего хозяйства, равномерно распределены между всех доходных квартилей, что подтверждается расчетами из табл. 2. При этом прочие характеристики оказывают значительно менее ошутимое влияние на уровень неравенства. Наконец, переменная «регион проживания» оказывает понижающее воздействие на неравенство. Связано это с тем фактом, что зависимая величина нормирована на прожиточный минимум, то есть показывает реальный уровень дохода. Несмотря на то что с точки зрения номинальных значений уровень неравенства по регионам значителен, территориальная дифференциация не увеличивает неравенство с точки зрения реальных уровней дохода.

Необходимо оговориться, что полученные результаты отражают статистическую зависимость между уровнем неравенства и микропоказателями домохозяйств. Так, домохозяйств из нижних децильных групп в среднем больше, именно поэтому вклад показателя «размер  $d/x$ » в неравенство является положительным. Та же статистическая зависимость прослеживается и для показателя «доля неработающих взрослых». Развитие адресной социальной поддержки населения России, направленной на домохозяйства с обозначенными выше характеристиками, снижает уровень неравенства.

Наконец, необходимо оговориться, что все приведенные выше выводы основаны на выборочных обследованиях домохозяйств, которые не являются в полной мере репрезентативными. Полученные выводы необходимо экстраполировать с определенным уровнем осторожности.

### **Проверка устойчивости результатов на сбалансированность выборки**

Для проверки полученных результатов на устойчивость регрессионная декомпозиция была проведена на основе двух дополнительных наборов данных: первый охватывает волны РМЭЗ НИУ ВШЭ с 13-й по 18-ю, второй — с 19-й по 25-ю. При составлении данных были использованы наблюдения только по тем домохозяйствам, которые входили во все волны, то есть эти две выборки являются полностью сбалансированными.

Для сопоставления полученных результатов по основной выборке (базовой несбалансированной) и комбинации из двух сбалансированных выборок была применена формула (8), которая отражает нормированный вклад фактора в неравенство. В табл. 2 представлены средние значения этих вкладов для наборов данных.

**Сравнение нормированного вклада факторов в индекс неравенства по двум наборам данных — базовой несбалансированной выборке и комбинации из двух сбалансированных наборов / Comparison of the normalized contribution of factors to the inequality index for two data sets — a basic unbalanced sample and a combination of two balanced sets**

Выборка	Регион проживания	Тип населенного пункта	Размер д/х	Доля детей
Базовая	-0,371	-0,061	1,270	0,134
Сбалансированная	-0,344	-0,018	1,198	0,206
Выборка	Доля пенсионеров	Доля инвалидов	Доля неработающих	Отрасль
Базовая	-0,005	0,016	0,448	-0,699
Сбалансированная	-0,002	0,021	0,384	0,089
Выборка	Образование	Пол	Возраст	Статус трудоустройства
Базовая	-0,689	-0,035	0,230	0,014
Сбалансированная	-0,765	-0,066	0,253	0,016

*Источник: расчеты авторов на основе данных РМЭЗ НИУ ВШЭ / Source: authors' calculations based on data from the RLMS of the Higher School of Economics.*

Из табл. 2 видно, что наиболее существенная разница наблюдается для показателя «отрасль занятости главы д/х», что может быть связано с изменением количества доступных для ответа отраслей. В целом по остальным переменным можно утверждать, что результаты являются качественно устойчивыми к используемому набору данных. Соответственно, проблема истощаемости выборки, которая приводит к несбалансированности данных, не оказывает качественного влияния на полученные выводы.

## **ЗАКЛЮЧЕНИЕ**

В рамках настоящего анализа была выполнена регрессионная декомпозиция дифференциации доходов домохозяйств на основе данных РМЭЗ НИУ ВШЭ за 2004–2016 гг. В частности, были получены следующие результаты. Если рассматривать не номинальное неравенство, а неравенство в уровне жизни (выраженное в виде номинального эквивалентного дохода, поделенного на прожиточный минимум), то существенной региональной дифференциации не наблюдается. Так, вклад переменной «регион проживания» в индекс неравенства сам по себе отрицателен. Наибольший положительный вклад в индекс неравенства наблюдается для демографических переменных — размера д/х, доли детей и доли неработающих взрослых. Переменные, отражающие человеческий капитал главы домохозяйства, оказывают выравнивающий эффект на общий уровень неравенства.

Необходимо отметить, что наличие статистической взаимосвязи между переменными не говорит о наличии причинно-следственной. Так, положительный вклад переменной «размер д/х» в неравенство не означает, что дробление домашнего хозяйства на два независимых приведет к снижению расслоения в обществе. С другой стороны, выявленные переменные являются важными индикаторами при проведении направленной на снижение уровня неравенства социально-экономической политики.

**Список источников**

- Джомо К. С., Попов В. В. Долгосрочные тенденции в распределении доходов // Журнал Новой экономической ассоциации. 2016. № 3 (31). С. 146–160.
- Кислицына О. А. Неравенство доходов в России в переходный период: научный доклад. М.: EERC, 2003. 79 с.
- Лазарян С. С., Черноталова М. А. Глобальная угроза роста неравенства // Научно-исследовательский финансовый институт. Финансовый журнал. 2017. № 4 (38). С. 34–46.
- Лукьянова А. Л. Динамика и структура неравенства по заработной плате (1998–2005). М.: ГУ ВШЭ, 2007. 68 с.
- Овчарова Л. Н., Попова Д. О., Рудберг А. М. Декомпозиция факторов неравенства доходов в современной России // Журнал Новой экономической ассоциации. 2016. № 3 (31). С. 170–186.
- Blinder A. S. Wage Discrimination: Reduced Form and Structural Estimates // Journal of Human Resources. 1973. Vol. 8. No. 4. P. 436–455. URL: <https://doi.org/10.2307/144855>.
- Brewer M., Wren-Lewis L. Accounting for Changes in Income Inequality: Decomposition Analyses for the UK, 1978–2008 // Oxford Bulletin of Economics and Statistics. 2016. Vol. 78. Iss. 3. P. 289–322. URL: <https://doi.org/10.1111/obes.12113>.
- Cowell F. A. Sampling Variance and Decomposable Inequality Measures // Journal of Econometrics. 1989. Vol. 42. Iss. 1. P. 27–41.
- Cowell F. A., Victoria-Feser M. P. Robustness Properties of Inequality Measures // Econometrica. 1996. Vol. 64. Iss. 1. P. 77–101. URL: <https://doi.org/10.2307/2171925>.
- DiNardo J., Fortin N., Lemieux T. Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach // Econometrica. 1996. Vol. 64. No. 5. P. 1001–1044. URL: <https://doi.org/10.2307/2171954>.
- Fields G. S. Accounting for Income Inequality and Its Changes: A New Method with Application to the Distribution of Earnings in the United States // Research in Labor Economics. 2003. No. 22. P. 1–38.
- Firpo S. P., Fortin N. M., Lemieux T. Decomposing Wage Distributions Using Recentered Influence Function Regressions // Econometrics. 2018. Vol. 6 (2). P. 1–40. URL: <https://doi.org/10.3390/econometrics6020028>.
- Hill M. O. Diversity and Evenness: a Unifying Notation and Its Consequences // Ecology. 1973. Vol. 54. No. 2. P. 427–432. URL: <http://dx.doi.org/10.2307/1934352>.
- Jansen W., Dessens J., Verhoeven W-J. Income Inequality Decomposition, Russia 1992–2002: Method and Application // Studies of Transition States and Societies. 2013. Vol. 5. Iss. 2. P. 21–34.
- Machado J., Mata J. Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression // Journal of Applied Econometrics. 2005. Vol. 20. Iss. 4. P. 444–465. URL: <https://doi.org/10.1002/jae.788>.
- Manna R., Regoli A. Regression-based Approaches for the Decomposition of Income Inequality in Italy, 1998–2008 // Rivista di statistica ufficiale. 2012. Vol. 14. Iss. 1. P. 5–18.
- Mincer J. (1974). The Human Capital Earnings Function. NBER, pp. 83–96.
- Morduch J., Sicular T. Rethinking Inequality Decomposition, with Evidence from Rural China // The Economic Journal. 2002. Vol. 112. No. 476. P. 93–106. URL: <https://doi.org/10.1111/1468-0297.0j674>.
- Oaxaca R. Male-Female Wage Differentials in Urban Labour Markets // International Economic Review. 1973. Vol. 14. Iss. 3. P. 693–709.
- Rani U., Furrer M. Decomposing Income Inequality into Factor Income Components: Evidence from Selected G-20 Countries // ILO Research Paper. No. 15. 2016. P. 1–28.
- Shorrocks A. F. The Class of Additively Decomposable Inequality Measures // Econometrica. 1980. Vol. 48. No. 3. P. 613–625.
- Shorrocks A. F. Inequality Decomposition by Factor Components // Econometrica, Econometric Society. 1982. Vol. 50. Iss. 1. P. 193–211.
- Shorrocks A. F. Inequality Decomposition by Population Subgroups // Econometrica. 1984. Vol. 52. Iss. 6. P. 1369–1385.
- Sicular T., Ximing Y., Gustafsson B. et al. The Urban-Rural Income Gap and Inequality in China // Review of Income and Wealth. 2007. Vol. 53. No. 1. P. 93–126.
- Wicaksono E., Amir H., Anda N. The Sources of Income Inequality in Indonesia: A Regression-Based Inequality Decomposition / Asian Development Bank Working Paper Series No. 667. 2017.

*Поступила в редакцию 2 сентября 2019 г.  
Принята к публикации 21 октября 2019 г.*

DOI: 10.31107/2075-1990-2019-5-74-89

## Regression-Based Decomposition of Income Inequality Factors in Russia

**Anton I. Votinov**<sup>1</sup> (e-mail: avotinov@nifi.ru), ORCID 0000-0002-2972-8498

<sup>1</sup> Financial Research Institute, Moscow 127006, Russian Federation

**Samvel S. Lazaryan**<sup>1</sup> (e-mail: lazaryan@nifi.ru), ORCID 0000-0001-6829-3979

<sup>1</sup> Financial Research Institute, Moscow, 127006, Russian Federation.

**Vyacheslav N. Ovchinnikov**<sup>2</sup>, (e-mail: vyacheslav\_ovchinnikov\_1993@mail.ru), ORCID 0000-0001-9786-3299

<sup>2</sup> Lobachevsky State University, Nizhny Novgorod 603022, Russian Federation

### Abstract

The purpose of this study was to identify the factors of income inequality in Russia using RLMS-data (2009–2016). In this paper inequality refers to differentiation in equivalent household income, adjusted for the cost of the consumer basket in the regions of Russia. To solve the problem, we used regression decomposition of the inequality index. According to the results, such factors as “the share of children in households”, “the size of households” and “the share of non-working adults” had the largest positive contribution to the inequality index. The “region of residence” factor, which is usually considered as the key factor of inequality in Russia, had an equalizing effect. This is due to the fact that we analyzed differentiation in the standard of living. Indeed, we calculated the dependent variable of income, adjusted for the cost of the consumer basket in the relevant region. Thus, the level of inequality was not affected by the regional attribute, but rather by the differentiation of other significant factors of the regions. The obtained results can be used for a deeper study of the causes of changes in inequality.

**Keywords:** income inequality, household, regression decomposition, quartile

**JEL:** D31, C51

**For citation:** Votinov A.I., Lazaryan S.S., Ovchinnikov V.N. Regression-Based Decomposition of Income Inequality Factors in Russia. *Financial Journal*, 2019, no. 5, pp. 74–89 (In Russ.). DOI: 10.31107/2075-1990-2019-5-74-89.

### References

- Blinder A.S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, vol. 8, no. 4, pp. 436–455. Available at: <https://doi.org/10.2307/144855>.
- Brewer M., Wren-Lewis L. (2016). Accounting for Changes in Income Inequality: Decomposition Analyses for the UK, 1978–2008. *Oxford Bulletin of Economics and Statistics*, vol. 78, iss. 3, pp. 289–322. Available at: <https://doi.org/10.1111/obes.12113>.
- Cowell F.A. (1989). Sampling Variance and Decomposable Inequality Measures. *Journal of Econometrics*, vol. 42, iss. 1, pp. 27–41.
- Cowell F.A., Victoria-Feser M.P. (1996). Robustness Properties of Inequality Measures. *Econometrica*, vol. 64, iss. 1, pp. 77–101. Available at: <https://doi.org/10.2307/2171925>.
- DiNardo J., Fortin N., Lemieux T. (1996). Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach. *Econometrica*, vol. 64, no. 5, pp. 1001–1044. Available at: <https://doi.org/10.2307/2171954>.
- Fields G.S. (2003). Accounting for Income Inequality and Its Changes: A New Method with Application to the Distribution of Earnings in the United States. *Research in Labor Economics*, vol. 22, pp. 1–38.
- Firpo S.P., Fortin N.M., Lemieux T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, vol. 6 (2), pp. 1–40. Available at: <https://doi.org/10.3390/econometrics6020028>.
- Hill M.O. (1973). Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, vol. 54, no. 2, pp. 427–432. Available at: <http://dx.doi.org/10.2307/1934352>.
- Jansen W., Dessens J., Verhoeven W.-J. (2013). Income Inequality Decomposition, Russia 1992–2002: Method and Application. *Studies of Transition States and Societies*, vol. 5, iss. 2, pp. 21–34.
- Jomo K.S., Popov V.V. (2016). Long-Term Trends in Income Distribution. *Journal of the New Economic Association*, iss. 3 (31), pp. 146–160 (In Russ.).
- Kislitsyna O.A. (2003). Income Inequality in Russia: A Scientific Report. Moscow: EERC, 79 p. (In Russ.).
- Lazaryan S.S., Chernotalova M.A. (2017). Global Risk of Rising Inequality. *Finansovyy zhurnal – Financial Journal*, no. 4 (38), pp. 34–46 (In Russ.).
- Lukiyanova A.L. (2007). Changes in the Wage Structure and Wage Inequality in Russia (1998–2005). Moscow: State University – Higher School of Economics, 68 p. (In Russ.).

Machado J., Mata J. (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics*, vol. 20. iss. 4, pp. 444–465. Available at: <https://doi.org/10.1002/jae.788>.

Manna R., Regoli A. (2012). Regression-Based Approaches for the Decomposition of Income Inequality in Italy, 1998–2008. *Rivista di statistica ufficiale*, vol. 14, iss. 1, pp. 5–18.

Morduch J., Sicular T. (2002). Rethinking Inequality Decomposition, with Evidence from Rural China. *The Economic Journal*, vol. 112, no. 476, pp. 93–106. Available at: <https://doi.org/10.1111/1468-0297.0j674>.

Ovcharova L.N., Popova D.O., Rudberg A.M. (2016). Decomposition of Income Inequality in Contemporary Russia. *Journal of the New Economic Association*, iss. 3 (31), pp. 170–186 (In Russ.).

Oaxaca R. (1973). Male-Female Wage Differentials in Urban Labour Markets. *International Economic Review*, vol. 14, iss. 3, pp. 693–709.

Rani U., Furrer M. (2016). Decomposing Income Inequality Into Factor Income Components: Evidence from Selected G-20 Countries. *ILO Research Paper*, no. 15, pp. 1–28.

Shorrocks A.F. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica*, vol. 48, no. 3, pp. 613–625.

Shorrocks A.F. (1982). Inequality Decomposition by Factor Components. *Econometrica*, vol. 50, iss. 1, pp. 193–211.

Shorrocks A.F. (1984). Inequality Decomposition by Population Subgroups. *Econometrica*, vol. 52, iss. 6, pp. 1369–1385.

Sicular T., Ximing Y., Gustafsson B. et al. (2007). The Urban-Rural Income Gap and Inequality in China. *The Review of Income and Wealth*, vol. 53, no. 1, pp. 93–126.

Wicaksono E., Amir H., Anda N. (2017). The Sources of Income Inequality in Indonesia: A Regression-Based Inequality Decomposition. Asian Development Bank Working Paper Series No. 667.

*Received 02.09.2019*

*Accepted for publication 21.10.2019*