Аналитический центр Правительства Москвы, Москва, Россия

Классификации регионов России с учетом структуры неформальной занятости и уровня социально-экономического развития

Цель исследования. Возможности реализации общенациональных и общерегиональных стратегических задач зависят от соответствия принимаемых мер выбранной траектории развития. Комплекс мер включает в себя управленческие решения в области регулирования рынка труда и касающиеся вопросов легализации неформальной занятости. Для принятия управленческих решений по регулированию рынка труда на региональном уровне необходима объективная статистическая оценка взаимосвязи неформальной занятости с показателями социально-экономического развития.

Информационной основой регулирования неформальной занятости должны являться количественно оцененные устойчивые закономерности взаимосвязи параметров неформальной занятости и структурно-динамических характеристик экономического роста и развития. Для выявления и оценки этих статистических закономерностей на основе данных Федеральной службы государственной статистики в региональном разрезе разработана система статистических показателей и определены значимые факторы неформальной занятости и социально-экономического развития. Учитывая неоднородность субъектов Российской Федерации по масштабу и структуре неформальной занятости требуется распределение регионов на однородные группы. Статья описывает подходы классификации регионов России с учетом уровня социально-экономического развития и структуры неформальной занятости.

Материалы и методы. Статья рассматривает подход к распределению регионов на однородные группы с применением методов кластерного анализа на основе главных компонент системы статистических показателей, характеризующих занятых в неформальном секторе, который базируется на предположении о том, что показатели неформальной занятости находятся в причинно-следственной зависимости с показателями социально-экономического развития.

Результаты. Получены пять групп регионов, однородных по структурным характеристикам неформальной занятости и обобщенным факторам социально-экономического развития. Для целей дальнейшей интерпретации выделенные группы оценены и отранжированы относительно среднероссийского уровня социально-экономического развития: низкий уровень (8 регионов), ниже среднего (26 регионов), средний (41 регион), выше среднего (8 регионов), высокий уровень (2 региона).

Заключение. Полученная классификация регионов России явлется переходным этапом к построению эконометрической модели взаимосвязи неформальной занятости и показателей социально-экономического развития. Дальнейший анализ позволит оценить какие именно показатели в наибольшей степени оказывают мультипликационный эффект на региональную экономику и получить количественную оценку данного влияния на ее рост.

Ключевые слова: неформальная занятость, кластерный анализ, факторный анализ, индекс Рэнда.

Elvira I. Dubravskaya

Analytical Centre by Moscow City Government, Moscow, Russia

Classification of Russian regions taking into account the structure of informal employment and the level of socio-economic development

Purpose of the study. The possibilities for the implementation of national and regional strategic objectives depend on the compliance of the measures taken with the chosen development path. The set of measures includes managerial decisions in the field of labor market regulation and concerning the legalization of informal employment. To make managerial decisions on the regulation of the labor market at the regional level, an objective statistical assessment of the relationship between informal employment and indicators of socio-economic development is required.

The information basis for the regulation of informal employment should be quantitatively assessed stable regularities of the relationship between the parameters of informal employment and the structural and dynamic characteristics of economic growth and development. To identify and evaluate these statistical patterns, based on data from the Federal State Statistics Service in a regional context, a system of statistical indicators has been developed and significant factors of informal employment and socio-economic development have been identified. Given the heterogeneity of the constituent entities of the Russian Federation in terms of the scale and structure of informal employment, the distribution of regions into homogeneous groups is required. The article describes the approaches to the classification of Russian regions, taking into account the level of socio-economic development and the structure of informal employment.

Materials and methods. The article examines an approach to the distribution of regions into homogeneous groups using the methods of cluster analysis based on a group of indicators characterizing those employed in the informal sector, which is based on the assumption that the indicators of informal employment are causally related to indicators of socio-economic development.

Results. Five groups of regions are obtained, homogeneous in terms of the structural characteristics of informal employment and generalized factors of socio-economic development. For the purposes of further interpretation, the selected groups are assessed and ranked relative to the average Russian level of socio-economic development: low level (8 regions), below average (26 regions), average (41 regions), above average (8 regions), high level (2 regions) ...

Conclusion. The resulting classification of Russian regions is a transitional stage to the construction of an econometric model of the relationship between informal employment and indicators of socio-economic development. Further analysis will allow us to assess which indicators have the greatest multiplier effect on the regional economy and to obtain a quantitative assessment of this impact on its growth.

Keywords: informal employment, cluster analysis, factor analysis, Rand index.

Введение

Актуальность разработки методологии статистической оценки взаимного влияния показателей неформальной занятости и показателей социальразвития но-экономического является устойчивой, но она особенно возрастает в контексте пандемии коронавируса, что подтверждается в публикации Генерального директора Европейской комиссии Вальтера Дж. Радермахера, который, в контексте кризиса COVID-19, отмечает, что в «... прежде всего, относящуюся к ЦУР информацию (показатели, счета, статистику) следует рассматривать как всеобъемлющую и определяющую методологическую основу, в которую должна вписываться вся статистика» [1]. В дополнении к этому, в справке МОТ «Кризис COVID-19 и неформальная экономика» [2] отмечается, что «Умереть от голода или от вируса» — абсолютно реальная дилемма, с которой сталкиваются многие работники неформальной экономики... в отсутствие альтернативных источников дохода утрата трудовых заработков вызовет рост относительной бедности среди неформальных работников и их семей более чем на 21 процентный пункт в странах с уровнем дохода выше среднего, почти на 52 пункта в странах с высоким уровнем дохода и на 56 пунктов в странах с низким уровнем дохода и доходом ниже среднего уровня».

В совокупности эти аргументы подчеркивают важность цели данной работы как в долгосрочной, так и в нынешней экстремальной ситуации, суть которой заключается в развитии теоретических и методологических положений анализа неформальной занятости. Необходимым этапом является решение задачи выделения групп регионов, однородных по целевым характеристикам, для это в рамках данной статьи рассмотрено:

- распределение регионов на группы по показателям неформальной занятости;
- кластеризация по обобщающим компонентам показателей социально-экономического развития и неформальной занятости на основе методов интегрирования факторного и кластерного анализа;
- сравнение результатов и определение оптимального подхода к классификации регионов России определен на основе индекса Рэнда.

Предложенная автором методические рекомендации носят практический характер и применимы для оценки неформальной занятости в разрезе субъектов, что, в свою очередь, является информационной основой регулирования неформальной занятости на региональном уровне и страновом уровне.

Формирование системы статистических показателей

Неформальная занятость является самодостаточной областью изучения, привлекающей исследователей из разных сфер и дисциплин: от экономики и антропологии до гендерных исследований, политологии, социологии и городского планирования. В связи с тем, что определение размеров и состава неформальной занятости, причин её возникновения

и последствий, направления влияния и взаимосвязей с другими показателями значительно различаются, отправной точкой формирования структуры системы показателей стал анализ документов стратегического планирования и целей национального развития до 2024 года Российской Федерации.

Для формирования системы статистических показателей проанализированы современные отечественные и зарубежные исследования и фундаментальные работы по вопросам взаимного влияния неформальной занятости и показателей социально-экономического развития, сгруппированы по блокам, соответствующим направлениям национальных проектов и дополнены блоком показателей, характеризующим макроэкономическую ситуацию. Выбранная категоризация, основанная на приоритетах национального развития, способствует формулированию выводов с учетом актуальной повестки: человеческого капитала, комфортной среды для жизни и экономического роста (рис. 1).

Исследования, посвященные неформальной занятости по типу исследования, можно разделить на качественные и количественные, по источнику на отчеты международных организаций и независимые



Рис. 1. Схема системы статистических показателей для изучения взаимного влияния неформальной занятости и уровня социальноэкономического развития стран и регионов

исследования, по охвату на международные, включая группы стран и ориентированные на одну страну. Дж. Мирейя, проводившая комплексный обзор литературы по неформальной занятости из 2461 источника за последние 40 лет, пришла к выводу о том, большая часть работ (48,8%) посвящена эволюции определения неформальной занятости, определение МОТ используется в 9,3% случаев, собственное определение в 13,9% случаев, неформальная занятость как таковая не определена в 23.30% случаев [3]. Некорректное использование терминологического аппарата (под понятием неформальная занятость подразумеваются иные вещи, например, в ряде работ исследователи не дают определения неформальной занятости как таковой, оперируя доступными данными) чаще встречается в работах, дающих количественные оценки. Примером подобного подхода является работа С. Бернабе, в которой использованы семь различных способов измерения неформальной занятости для оценки явления в семи странах СНГ [4]. Несмотря на описанные выше проблемы, анализ работ отечественных и зарубежных авторов [5-7] подтверждает гипотезу о том, что неформальная занятость связана с широким кругом вопросов социально-экономического развития, что, в свою очередь, подтверждает актуальность выбранной схемы построения системы показателей.

Отбор индикаторов, составляющих наполнение блоков системы показателей осуществлялся на основе доступной статистической информации. Количество индикаторов, включаемых в систему статистических показателей, согласно Елисеевой И.И., может варьироваться «от двух-трех до нескольких сотен» [8]. Полный перечень из 141 показателя, являющийся необходимым и



Рис. 2. Распределение индикаторов в соответствии с системой статистических показателей для изучения взаимного влияния неформальной занятости и уровня социально-экономического развития регионов (Примечание: количество индикаторов приведено в круге)

достаточным для решения поставленной задачи.

Исследование показателей неформальной занятости и социально-экономического развития в России на основе метода главных компонент

Информационная база, сформированная в соответствии с разработанной система статистических показателей для изучения взаимного влияния неформальной занятости уровня социально-экономического развития регионов состоит из 141 переменной и значений признаков за 2015-2017 гг., распределение индикаторов по блокам приведено на рисунке ниже. Факторный и компонентный анализ позволяет решить проблему большой размерности путем получения интегральных показателей, число которых значительно меньше числа исходных признаков. Большое количество переменных затрудняет анализ и интерпретацию, в то время как полученные с применением метода главных компонент обобщающие характеристики являются содержательно интерпретируемой основной проведения кластерного анализа, учитывающей существенные, но не поддающиеся непосредственному наблюдению и измерению, скрытые свойства явлений. Компонентный анализ также позволит избежать мультиколлинеарности и исключения признаков, которые могли бы возникнуть при проведении корреляционно-регрессионного анализа по причине того, что некоторые из отобранных показателей в значительной степени дублируют друг друга [9, с. 201].

Проведенный факторный анализ, включает в себя решение проблем робастности, общности, факторов, вращения и оценки значения факторов [11]. Оптимальное число обобщающих факторов, отобрано на основе критерия Кайзера [10] (факторы с собственными значениями выше 1). Для интерпретации использовались признаки, факторные нагрузки которых по модулю больше 0,6, вращение varimax normolized. Ниже приведена интерпретация обобщающих факторов в соответствии с комплексной системой показателей на примере 2017 года.

Блок «Показатели неформальной занятости»

Главные компоненты блока неформальной занятости включают в себя два обобщенных фактора: первый обобщенный фактор (К1) характеризует секторальную структуру неформальной занятости (распределение участия в формальном/неформальном секторах экономики), второй обобщенный фактор (К2) отражает тип неформальной занятости (соотношение занятости по основной или дополнительной работе). Оптимальное число обобщающих факторов (главных компонент) выбрано на основе критерия Кеттера, кумулятивная (накопленная) объясненная вариация для первого фактора составляет 53,95%, для второго — 74,92%.

Блок «Показатели социально-экономического развития»

1. Показатели, характеризующие экономический рост

Для группы десяти исходных показателей, характеризующих макроэкономическую ситуацию первая компонента интерпретируется как FA1 — инвестиционная емкость экономики. Она наиболее тесно связана с показателями «Валовой региональный продукт на душу населения» и «Объем инвестиций в основной капитал на душу населения».

Вторая главная компонента FA2 — интерпретируется как «Экспортоориентированность экономики» и наиболее тесно связана с показателями: «Доля экспорта субъекта в совокупной внешней торговле со странами дальнего зарубежья» и «Доля экспорта субъекта в совокупной внешней торговле со странами СНГ».

Третья главная компонента интерпретируется как «Производительность труда». Она наиболее тесно связана с показателями «Индекс физического объема инвестиций в основной капитал» и «Индекс производительности труда относительно уровня 2011 года». В совокупности три главные компоненты объясняют 71,90% общей вариации.

Расчет первых двух главных компонент частного обобщающего показателями, характеризующего рынок труда регионов проводился на основе девяти показателей. Первая главная компонента интерпретируется как FL1 — «Вовлеченность населения в рынок труда». Ее формируют показатели: «Уровень участия в рабочей силе», «Уровень занятости населения», «Уровень безработицы».

Вторая главная компонента интерпретируется как FL2 — «Зарегистрированная безработица». Она представлена показателями: «Изменение среднегодовой численности занятых» и «Уровень зарегистрированной безработицы». Суммарный вклад двух первых компонент в общую вариацию составил 56,47%.

Интерпретация трех главных компонент частного обобщающего показателя, характеризующего блок «Предпринимательство» осуществлялась на основе тридцати семи показателей.

Первая главная компонента интерпретируется как FEn1 -«Развитие торговли, сферы услуг, транспорта и связи» Она включает семь показателей: «Число малых предприятий на 10 000 человек населения», «Оборот розничной торговли на душу населения», «Оборот оптовой торговли на душу населения», «Объем платных услуг населению на душу населения», «Объем транспортных услуг населению на душу населения», «Объем услуг связи на душу населения», «Объем коммунальных услуг населению на душу населения».

Вторая главная компонента, которая интерпретируется как FEn2 – «Развитие промышленности за счет добывающего сектора», наиболее тесно связана с показателями: «Индексы промышленного производства в соответствии с ОКВЭД2 (2015-2017 гг.)», «Индексы производства по виду экономической деятельности «Обеспечение электрической энергией, газом и паром; кондиционирование воздуха» в соответствии с ОКВЭД2 (2015-2017 гг.)», «Индексы цен производителей промышленных товаров в соответствии с ОКВЭД2» и «Индексы цен производителей «Добыча полезных ископаемых» в соответствии с ОКВЭД2».

Третья главная компонента — FEn2 «Жилищное стро-

ительство». В ее состав вошли показатели: «Ввод в действие жилых домов на 1000 человек населения (м² общей площади жилых помещений)» и «Ввод в действие квартир на 1000 человек населения (единиц)». Суммарный вклад трех первых главных компонент в общую вариацию составляет 39,51%.

Интерпретация двух главных компонент частного обобщающего показателя, характеризующего блок «Инновационное развитие экономики и цифровизация общества» осуществлялась на основе семи показателей.

Первая главная компонента - FS1 «Инновационность экономики». В ее состав вошли показатели: «Удельный вес организаций, осуществлявших технологические, маркетинговые и организационные инновации, в общем числе обследованных организаций», «Затраты на технологические инновации в процентах от общего объема отгруженных товаров, выполненных работ, услуг», «Число подключенных абонентских устройств мобильной связи на 1000 человек населения», «Объем инновационных товаров, работ, услуг».

Вторая главная компонента, которая интерпретируется как FS2 — «Цифровизация общества» наиболее тесно связана с показателем: «Удельный вес домашних хозяйств, имевших доступ к сети Интернет».

Суммарный вклад двух первых главных компонент в общую вариацию составляет 54,69%.

2. Показатели, характеризующие человеческий капитал

Для группы из двенадцати исходных показателей, характеризующих блок «Демография» получено четыре главные компоненты, объясняющие 81,93% вариации.

Первая главная компонента FD1 интерпретируется как «Устойчивость семьи (Семейное положение)», она наиболее тесно связана с такими

показателями: «Удельный вес городского населения в общей численности населения», «Общие коэффициенты разводимости на 1000 человек населения», «Общие коэффициенты рождаемости», «Общие коэффициенты смертности», «Суммарный коэффициент рождаемости» и объясняет 32,43% вариации.

Вторая главная компонента FD2 интерпретируется как «Естественный прирост населения». Ее формируют показатели: «Коэффициенты демографической нагрузки», «Общие коэффициенты брачности на 1000 человек населения» и «Общие коэффициенты разводимости на 1000 человек населения».

Третья главная компонента FD3 интерпретируется как «Ожидаемая продолжительность жизни». Она представлена показателями: «Ожидаемая продолжительность жизни при рождении», «Плотность населения», «Коэффициенты миграционного прироста на 10 000 человек населения».

Четвертая главная компонента FD4, интерпретируемая как «Демографическая нагрузка на население трудоспособного возраста» сформирована одноимённым показателем.

Для расчета трех первых главных компонент частного обобщающего показателя, характеризующего блок «Благосостояние населения» использовался тридцать один показатель.

Первая главная компонента интерпретируется как FI1 – «Оплата труда и социальная поддержка населения». Она включает в себя следующие показатели: «Реальный размер назначенных пенсий», «Доля оплаты труда в структуре денежных доходов населения», «Другие доходы (включая «скрытые», от продажи валюты, денежные переводы и пр.) в структуре денежных доходов населения», «Величина прожиточного минимума, установленная в субъектах Российской Федерации», «Потребительские расходы в среднем на душу населения», «Доля обязательных платежей и разнообразных взносов в структуре использования денежных доходов населения», «Приобретение недвижимости в структура использования денежных доходов населения», «Доля оплаты услуг в структуре потребительских расходов домашних хозяйств», «Удельный вес аварийного жилищного фонда в общей площади всего жилищного фонда», «Удельный вес семей, получивших жилые помещения, в числе семей, состоявших на учете в качестве нуждающихся в жилых помещениях», «Отношение среднемесячной номинальной начисленной заработной плате работников организаций к прожиточному минимуму», «Численность граждан, пользующихся социальной поддержкой по оплате жилого помещения и коммунальных услуг», «Задолженность по кредитам в рублях, предоставленным кредитными организациями физическим лицам по ипотечным жилищным кредитам».

Вторая главная компонента интерпретируется как FI2 -«Сбережения населения» наиболее тесно связана с показателями: «Численность населения с денежными доходами ниже величины прожиточного минимума», и такими показателями структура использования денежных доходов населения как «Покупка товаров и оплата услуг», «Прирост финансовых активов» и «Прирост (уменьшение) денег у населения за счет изменения финансовых активов».

Третья главная компонента интерпретируется как FI3 — «Уровень бедности» и тесно связана с показателем «Доля покупки непродовольственных товаров в структуре потребительских расходов домашних хозяйств».

Суммарный вклад трех первых главных компонент в общую вариацию составляет 54,23%.

Для группы из двенадцати трех исходных показателей, характеризующих блок «*Paзвитие социальной сферы*» получено четыре главные компоненты, объясняющие 71,11% вариации.

Первая главная компонента интерпретируется как FC1 -«Доступность социальных и медицинских услуг». Она включает восемь показателей: «Число больничных коек на 10 000 человек населения», «Мощность амбулаторно-поликлинических организаций на 10 000 человек населения», «Численность врачей всех специальностей на 10 000 человек населения», «Численность населения на одного врача», «Численность населения на одного работника среднего медицинского персонала», «Заболеваемость на 1000 человек населения», «Число зарегистрированных преступлений на 100 000 человек населения», «Охват детей дошкольным образованием».

Вторая главная компонента, которая интерпретируется как FC2 — «Качество медицинских услуг» наиболее тесно связана с показателями: «Численность врачей всех специальностей на 10 000 человек населения», «Численность населения на одного врача», «Отношение числа больничных коек к численности врачей всех специальностей».

Третья главная компонента интерпретируется как FC3 — «Оплата труда «указных категорий»» и тесно связана с показателем «Отношение средней заработной платы социальных работников в организациях государственной и муниципальной форм собственности к среднемесячному доходу от трудовой деятельности по субъектам Российской Федерации». Суммарный вклад трех первых главных компонент в общую вариацию составляет 62,03%.

Четвертая главная компонента, вклад которой в об-

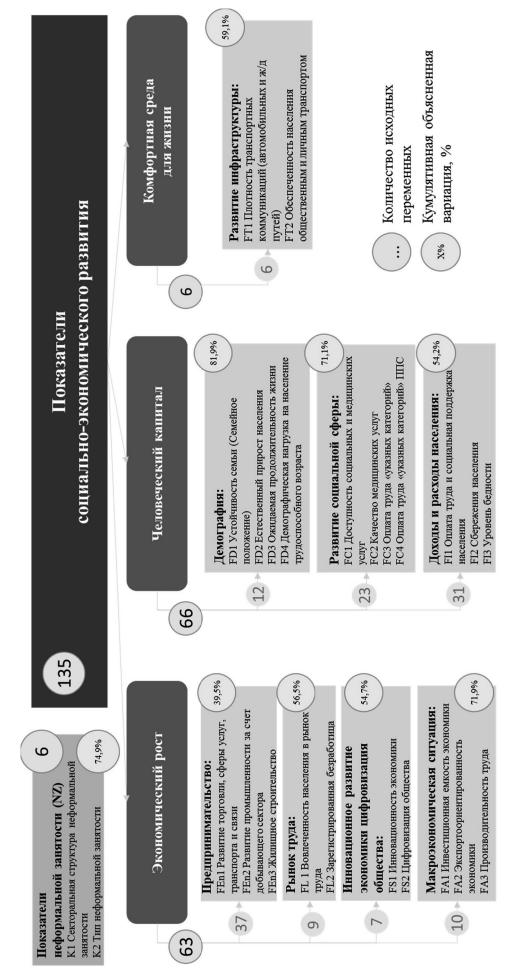


Рис. 3. Интерпретация главных компонент и кумулятивная объясненная дисперсия в соответствии с системой статистических показателей (на примере 2017 года)

ЩУЮ варианцию составляет 9,08%, интерпретируется как FC4 – «Оплата труда профессорско-преподавательского состава», она тесно связана с показателем «Отношение средней заработной платы преобразовательных подавателей организаций высшего профессионального образования государственной и муниципальной форм собственности к среднемесячному доходу от трудовой деятельности по субъектам Российской Федерации».

3. Показатели, характеризующие комфортную среду для жизни

Для группы из шести исходных показателей, характеризующих блок «Развитие инфраструктуры» получено две главные компоненты, объясняющие 59,11% вариации.

Первая главная компонента — FT1 «Плотность транспортных коммуникаций (автомобильных и ж/д путей)». В ее состав вошли показатели: «Плотность железнодорожных путей общего пользования», «Удельный вес автомобильных дорог с усовершенствованным покрытием в протяженности автомобильных дорог с твердым покрытием общего пользования», «Плотность автомобильных дорог общего пользования дорог общего пользования с твердым покрытием».

Вторая главная компонента FT2 интерпретируется как «Обеспеченность населения общественным и личным транспортом». Ее формируют показатели: «Число собственных легковых автомобилей на 1000 человек населения» и «Число автобусов общего пользования на 100 000 человек населения».

Метод главных компонент позволил выделить двадцать пять интегральных показателей (общих факторов), учитывающих разложение дисперсии через латентные факторы (обобщенные признаки) ста сорока одного исходного показателя (рис. 3).

Полученные главные компоненты позволяют провести

кластерный анализ по показателям неформальной занятости и социально-экономического развития на основе интерпретируемых интегральных показателей.

Статистическая оценка неформальной занятости и социально-экономического развития в субъектах РФ (кластерный анализ)

Формирование однородных по структурным показателям неформальной занятости обобщенным факторам социально-экономического развития групп регионов позволит провести дальнейший анализ с применением методов канонической корреляции. Для проведения корректного сравнения групп регионов необходимо зафиксировать состав анализируемых кластеров, в связи с чем кластерный анализ (агломеративный метод Уорда [14] и метод k-средних) выполнен на основе данных за 2017 год.

Кластеризация субъектов Российской Федерации по неформальной занятости осуществлена по обобщенному фактору К1 — структурная характеристика неформальной занятости. Второй обобщенный фактор (К2) отражает тип неформальной занятости (соотношение занятости по основной или дополнительной работе) не является кластерообразующим.

Характеристика социально-экономического развития представлена 21 обобщающим фактором (FA1, FA2, FA3, FD1, FD2, FD3, FD4, FL1, FL2, FEn1, FEn2, FI1, FI2, FI3, FS1, FS2, FT1, FT2, FC1, FC2, FC3). Факторы FEn3, FC4 не использованы при формировании кластеров (Р-уровень значимости > 0,05).

По определению процедура кластеризации продолжается до тех пор, пока все объекты в наборе данных не будут назначены одному или нескольким кластерам. В случае применения алгоритмов иерархического кластера все объекты в конечном итоге будут существовать в рамках едино-

го кластера — решения, которое не представляет интереса для изучения. Существует ряд внешних и внутренних критериев, помогающих определить наилучшее количество кластеров для интерпретации. Например, в работах М.Р. Резаи [12] доказано, что оптимальное число кластеров находится в диапазоне от 2 до \sqrt{n} , где n — это число всех точек данных в пространстве данных.

Одним из интересных подходов к выбору числа кластеров является алгоритм адаптивной кластеризации, основанный на максимальном расстоянии и теории байесовских решений. В его основе лежит оценка степени риска рассматриваемой схемы кластеризации с помощью предложенной функции оценки риска. Сравнивая значения риска до и после разделения, алгоритм может решить, будет ли набор данных продолжать разделяться, таким образом определяя количество кластеров и получая конечный результат кластеризации [13].

Однако до тех пор, пока в популярных статистических пакеты не интегрированы больше внутренних правил остановки процедуры кластеризации, исследователям приходится полагаться на теоретическое обоснование и субъективную проверку и для определения наилучшего числа кластеров [15].

Одним из эвристических методов является «обрезка» иерархического дерева. Визуальный анализ дендрограммы распределения регионов РФ на кластеры по компонентам, характеризующим структурные показатели неформальной занятости и обобщенные факторы социально-экономического развития с помощью дендрограммы с использованием метода Уорда, представленной на рис. 4 позволяет выделить как четыре, так и пять групп регионов. Как и в случае интерпретации результатов факторного анализа, интерпретация результатов дендрограммы

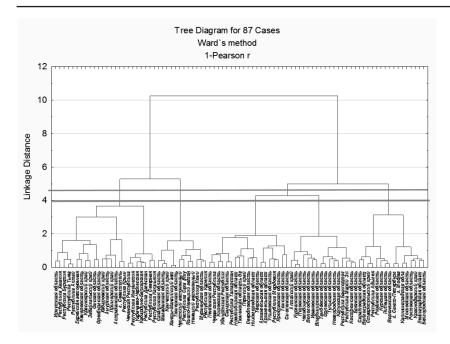


Рис. 4. Распределение регионов РФ на кластеры по компонентам, характеризующим структурные показатели неформальной занятости и обобщенные факторы социально-экономического развития с помощью дендрограммы с использованием метода Уорда

затруднена отсутствием однозначных критериев выделения кластеров. Согласно методике, предложенной Олдендерфером М.С. и Блэшфилдом Р.К. считается, что при «обрезке» дерева на отметке шкалы пять удается достигнуть 80% уровня сходства [16]. Подобный подход позволяет выделить четыре

однородных по исследуемым показателям групп субъектов совокупность регионов разделяется на кластеры состоящие из 2, 8, 11 и 66 субъектов Российской Федерации. Полученный результат затрудняет возможность содержательной интерпретации по причине большого количества единиц наблюдения в составе одного из выделенных кластеров.

Олдендерфер М.С. и Блэшфилд Р.К. допускают вариативность в выборе уровня «обрезания» «виноградной ветви» в целях получения оптимального способа кластеризации [16]. В результате группировки получено пять групп регионов, однородных по характеристикам неформальной занятости и социально-экономического развития.

Первый кластер образован 41 регионом (Белгородская, Брянская, Владимирская, Воронежская, Ивановская, Калужская, Курская, Липецкая, Московская, Орловская, Рязанская, Смоленская, Тамбовская,



Рис. 5. Картограмма распределения регионов РФ на кластеры по компонентам, характеризующим структурные показатели неформальной занятости и обобщенные факторы социально-экономического развития с использованием метода Варда

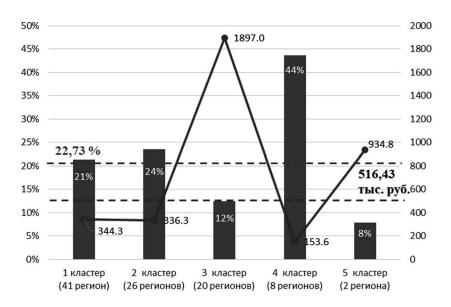
Тверская, Тульская, Ярославская, Калининградская, нинградская, Новгородская, Астраханская, Волгоградская, Ростовская, Ставропольский край, Нижегородская, Пензенская, Самарская, Саратовская, Ульяновская, Свердловская, Тюменская (без автономных округов), Челябинская, Новосибирская и Омская области, Республики Башкортостан, Мордовия, Татарстан, Адыгея и Чувашская Республика, Краснодарский, Приморский и Хабаровский край), то есть половиной субъектов РФ, что можно интерпретировать как среднестрановую ситуацию, т.е. средний уровень неформальной занятости и средний уровень социально-экономического развития. Действительно, среднее значение показателя «Занятые в неформальном секторе к общей численности занятого населения» в этой группе регионов равно 20,9%, в России 22,73%, медианное значение показателя на страновом уровне – 21,16%. Как показано на рис. 6, валовой региональный продукт на душу населения в исследуемом периоде в среднем составлял 472 049,9 рубля, медианное – 344 487,4 рубля, среднее значение данного показателя по первому кластеру 344 304,8 рублей на человека.

Второй кластер состоит из 26 регионов: Архангельская область, Архангельская область без автономного округа, Вологодская, Псковская, Кировская, Оренбургская, Курганская, Иркутская, Кемеровская, Томская, Амурская, Костромская и Еврейская автономная области, Республика Калмыкия, Марий Эл, Удмуртская, Алтай, Бурятия, Тыва, Хакасия, Карелия и Коми, Алтайский, Забайкальский, Красноярский и Пермский край. Второй кластер. В первоначальном варианте группировки регионов на четыре кластера первый и второй кластер являлись одной группой, состоящей из шестидесяти шести регионов, что объясняет то, что значения показателей, характеризующих уровень неформальной занятости и валового регионального продукта на душу населения стремится к среднестрановым значениям. Однако, первый и второй кластер во многом различны, что наглядно показывают средние значения обобщенных факторов, приведенные в таблице ниже. Среди различий можно отметить существенное различие по обобщающим характеризующим факторам, структуру неформальной занятости, устойчивость семьи, развитие промышленности за счет

добывающего сектора, естественный прирост населения, ожидаемую продолжительность жизни, сбережения населения, уровень бедности, инновационность экономики, доступность социальных и медицинских услуг и плотность транспортных коммуникаций (табл. 1). Полученная группа регионов близка к средней, однако уровень её социально-экономического развития и неформальной занятости ниже среднего.

Третий кластер объединил 20 регионов (Ненецкий автономный округ, Мурманская, Тюменская, Магаданская, Сахалинская области, Чукотский автономный округ, Ханты-Мансийский автономный округ – Югра, Ямало-Ненецкий автономный округ, Республика Саха (Якутия) и Камчатский край). Данные регионы характеризуются уровнем социально-экономического развития выше среднего за счет высокий доли ресурсной составляющей.

Четвертый кластер включает в себя 8 регионов с низким уровнем социально-экономического развития, расположенных на юге России: Республика Крым, г. Севастополь, Республика Дагестан, Республика Ингушетия, Кабардино-Балкарская Республика, Карачаево-Черкесская Республика,



- Занятые в неформальном секторе к общей численности занятого населения, % (левая шкала)
- Валовой региональный продукт на душу населения, тыс. руб (правая шкала)
- Среднее значение по РФ

Рис. 6. Характеристика кластеров по неформальной занятости и валового регионального продукта

 Таблица 1

 Распределение регионов РФ по структурным показателям неформальной занятости и обобщенным факторам социально-экономического развития

Обобщенный фактор		Номер кластера				
Блок	Код, интерпретация	1	2	3	4	5
		средний	ниже среднего	выше среднего	низкий	высокий
Неформальная занятость	К1 Структурная характеристика неформальной занятости	-0.103	0.401	0.563	-1.709	0.910
Экономический рост	FA1 Инвестиционная емкость экономики	0.209	0.350	-1.436	-0.028	-1.533
	FA2 Экспортоориентированность экономики	0.014	0.062	0.470	0.335	-4.775
	FA3 Производительность труда	-0.086	0.041	-0.694	1.160	0.067
	FL 1 Вовлеченность населения в рынок труда	0.028	-0.651	1.423	-0.433	2.502
	FL2 Зарегистрированная безработица	0.389	-0.014	-0.110	-2.046	0.942
	FEn1 Развитие торговли, сферы услуг, транспорта и связи	-0.029	-0.462	1.373	-0.978	3.640
	FEn2 Развитие промышленности за счет добывающего сектора	-0.422	0.325	0.871	0.018	-0.002
Человеческий капитал	FD1 Устойчивость семьи (Семейное положение)	0.411	-0.163	-0.168	-1.601	0.936
	FD2 Естественный прирост населения	0.389	-0.049	-1.795	0.670	-1.047
	FD3 Ожидаемая продолжительность жизни	-0.008	0.708	-0.149	-1.044	-4.121
	FD4 Демографическая нагрузка на население трудоспособного возраста	0.003	0.245	-1.036	-0.082	2.260
	FI1 Оплата труда и социальная поддержка населения	0.259	0.021	-2.071	1.538	-1.385
	FI2 Сбережения населения	-0.459	0.264	0.806	1.008	-2.075
	F13 Уровень бедности	0.109	-0.503	0.380	0.779	-0.706
	FS1 Инновационность экономики	-0.521	0.216	0.787	1.478	-1.979
	FS2 Цифровизация общества	0.135	0.351	-1.299	0.383	-2.364
	FC1 Доступность социальных и медицинских услуг	-0.291	0.434	1.409	-1.547	-0.530
	FC2 Качество медицинских услуг	-0.028	-0.540	0.343	0.597	3.488
	FC3 Оплата труда «указных категорий»	0.037	0.010	-0.081	0.508	-2.533
Комфортная среда для жизни	FT1 Плотность транспортных коммуникаций (автомобильных и ж/д путей)	-0.158	0.502	0.561	-0.137	-5.546
	FT2 Обеспеченность населения общественным и личным транспортом	0.404	-0.186	-0.433	-0.769	-0.629

блика, Республика Северная Осетия — Алания, Чеченская Республика.

Пятый кластер образован двумя городами федерального значения, которые согласно типологии Министерства экономического развития являются регионами — локомотивами роста (г. Москва, г. Санкт-Петербург), для которых характерен высокий уровень социально-экономического развития и, как следствие, низкий уровень неформальной занятости.

На основе интегрирования методов факторного и кластерного анализа получены пять групп регионов, однородных по структурным характеристикам неформальной занятости и обобщенным факторам социально-экономического развития. Для целей дальнейшей

интерпретации выделенные группы оценены и отранжированы относительно среднероссийского уровня социально-экономического развития:

- 1. Низкий уровень (8 регионов);
- 2. Ниже среднего (26 регионов);
 - 3. Средний (41 регион);
- 4. Выше среднего (8 регионов):
- 5. Высокий уровень (2 региона).

Заключение

Представленные в статье результаты имеют практическую значимость, поскольку дают убедительное подтверждение необходимости кластеризации субъектов России относительно структурных характеристик не-

формальной занятости и уровня социально-экономического развития на основе интегрирования методов факторного и кластерного анализа. Предложенная методология может быть использована при формировании информационной базой для принятия управленческих решений по регулированию рынка труда на как на федеральном, так и на региональном уровне. Переход к построению эконометрической модели взаимосвязи неформальной занятости и показателей социально-экономического развития выделить показатели, которые в наибольшей степени оказывают мультипликационный эффект на региональную экономику и получить количественную оценку данного влияния на ее рост.

Литература

- 1. Radermacher W.J. Official Statistics in the context of the COVID-19 crisis [Электрон. ресурс]. Режим доступа: https://officialstatistics.com/news-blog/crises-politics-and-statistics (Дата обращения: 19.11.2020)
- 2. ILO. Кризис COVID-19 и неформальная экономика [Электрон. pecypc]. Режим доступа: https://www.ilo.org/wcmsp5/groups/public/--ed_protect/---protrav/---travail/documents/briefingnote/wcms_745853.pdf.
- 3. Mireia J Informal employment in high-income countries for a health inequalities research: A scoping review // Work. Informal employment in high-income countries for a health inequalities research. 2016. T.53. № 2. C. 347–356.
- 4. Bernabè S. Measuring informal employment in transition countries // Note prepared for the WIEGO meeting on «Measuring Informal Employment in Developed Countries. 2008. T. 31.
- 5. Гимпельсон В.Е. «Бойцы невидимого фронта»: кто они и сколько их? История на основе данных ОНПЗ // В тени регулирования: неформальность на российском рынке труда. Под общ. ред. В.Е. Гимпельсон, Р.И. Капелюшников. М.: Издательский дом НИУ ВШЭ, 2014.
- 6. Бирюкова С.С., Синявская О.В. Возможные меры снижения неформальной занятости и скрытой оплаты труда // Журнал Новой Экономической Ассоциации. 2018. № 1(37). С. 193—203.
- 7. Binelli C. Wage inequality and informality: evidence from Mexico // IZA Journal of Labor & Development. 2016. T. 5. № 1. C. 5.

- 8. Елисеева И., Юзбашев М.М. Общая теория статистики. М.: Федеральное государственное унитарное предприятие «Издательство», 2004.
- 9. Власов М.П. Моделирование экономических систем и процессов. М.: Издательский Дом «Инфра-М», 2011. 311 с.
- 10. Kaiser H.F. The application of electronic computers to factor analysis // Educational and psychological measurement. 1960. T. 20. № 1. C. 141–151.
- 11. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. М.: Финансы и статистика, 2011. 352 с.
- 12. Rezaee M.R., Lelieveldt B.P., Reiber J.H. A new cluster validity index for the fuzzy c-mean // Pattern recognition letters. 1998. T. 19. № 3–4. C. 237–246.
- 13. Zhao F., Yang Y., Zhao W. Adaptive clustering algorithm based on max-min distance and bayesian decision theory // IAENG Int. J. Comp. Sci. IJCS. 2017. T. 44. № 2. C. 24.
- 14. Ward Jr J.H. Hierarchical grouping to optimize an objective function // Journal of the American statistical association. 1963. T. 58. № 301. C. 236–244.
- 15. Tinsley H.E., Brown S.D. Handbook of applied multivariate statistics and mathematical modeling. Academic press, 2000.
- 16. Олдендерфер М.С., Блэшфилд Р.К.. Факторный, дискриминантный и кластерный анализ. М.: Финансы и статистика, 1989. 215 с.

References

- 1. Radermacher W.J. Official Statistics in the context of the COVID-19 crisis [Interner]. Available from: https://officialstatistics.com/news-blog/crises-politics-and-statistics (cited 19.11.2020).
- 2. ILO. Krizis COVID-19 i neformal'naya ekonomika =ILO. The COVID-19 crisis and the informal economy [Internet]. Available from: https://www.ilo.org/wcmsp5/groups/public/--ed_protect/---protrav/---travail/documents/briefingnote/wcms_745853.pdf. (In Russ.)
- 3. Mireia J Informal employment in high-income countries for a health inequalities research: A scoping review. Work. Informal employment in high-income countries for a health inequalities research. 2016; 53; 2: 347-356.
- 4. Bernabè S. Measuring informal employment in transition countries. Note prepared for the WIEGO meeting on «Measuring Informal Employment in Developed Countries. 2008; 31.
- 5. Gimpel'son V.Ye. "Fighters of the invisible front": who are they and how many are there? History based on Omsk Refinery data. V teni regulirovaniya: neformal'nost' na rossiyskom

- rynke truda. Pod obshch. red. V.E. Gimpel'son, R.I. Kapelyushnikov = In the shadow of regulation: informality in the Russian labor market. Under total. ed. V.E. Gimpelson, R.I. Kapelyushnikov. Moscow: HSE Publishing House; 2014. (In Russ.)
- 6. Biryukova S.S., Sinyavskaya O.V. Sinyavskaya O.V. Possible measures to reduce informal employment and hidden wages. Zhurnal Novoy Ekonomicheskoy Assotsiatsii = Journal of the New Economic Association. 2018; 1(37): 193-203. (In Russ.)
- 7. Binelli C. Wage inequality and informality: evidence from Mexico. IZA Journal of Labor & Development. 2016; 5; 1: 5.
- 8. Yeliseyeva I., Yuzbashev M.M. Obshchaya teoriya statistiki = General theory of statistics. Moscow: Federal State Unitary Enterprise "Publishing House"; 2004. (In Russ.)
- 9. Vlasov M.P. Modelirovaniye ekonomicheskikh sistem i protsessov = Modeling of economic systems and processes. Moscow: Publishing House "Infra-M"; 2011. 311 p. (In Russ.)
- 10. Kaiser H.F. The application of electronic computers to factor analysis. Educational and psychological measurement. 1960; 20; 1: 141–151.

- 11. Dubrov A.M., Mkhitaryan V.S., Troshin L.I. Mnogomernyye statisticheskiye metody = Multivariate statistical methods. Moscow: Finance and Statistics; 2011. 352 p. (In Russ.)
- 12. Rezaee M.R., Lelieveldt B.P., Reiber J.H. A new cluster validity index for the fuzzy c-mean. Pattern recognition letters. 1998; 19; 3-4: 237–246.
- 13. Zhao F., Yang Y., Zhao W. Adaptive clustering algorithm based on max-min distance and bayesian decision theory. IAENG Int. J. Comp. Sci. IJCS. 2017; 44; 2: 24.

Сведения об авторе

Эльвира Ивановна Дубравская

Главный эксперт ГБУ «Аналитический центр» Правительства Москвы Москва, Россия Эл. noчma: elvira.dubravskaya@yandex.ru 14. Ward Jr J.H. Hierarchical grouping to optimize an objective function. Journal of the American statistical association. 1963; 58; 301: 236–244.

15. Tinsley H.E., Brown S.D. Handbook of applied multivariate statistics and mathematical modeling. Academic press; 2000.

16. Oldenderfer M.S., Bleshfild R.K. Faktornyy, diskriminantnyy i klasternyy analiz = Factor, discriminant and cluster analysis. Moscow: Finance and Statistics; 1989. 215 p. (In Russ.)

Information about the author

Elvira I. Dubravskaya

Senior Analyst at Project Office of the Analytical Centre by Moscow City Government, Moscow, Russia E-mail: elvira.dubravskaya@yandex.ru