

Методика формирования сбалансированного набора устойчивых индикаторов оценки регионального профессионального образования*

Григорьев В. Ю.

Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации, Москва, Российская Федерация; grigorev-vy@ranepa.ru

РЕФЕРАТ

Для оценки состояния региональной системы профессионального образования авторами разработана открытая пятиуровневая модель оценки, на базе которой выполнен комплекс исследований, включающий кластерный и факторный анализ. На основе использования метода главных компонент предложена методика понижения размерности массива индикаторов для последующей кластеризации регионов.

Впервые предложено и апробировано совместное использование регрессионного анализа и метода главных компонент для получения устойчивого набора индикаторов для оценки региональной системы профессионального образования.

Обнаружена зависимость между устойчивостью модели оценки системы профессионального образования и направленностью действия индикаторов, как внутри одной индикаторной группы, так и в разных индикаторных группах.

Рекомендуется использовать в интересах Министерства просвещения Российской Федерации; органов управления образованием субъектов Российской Федерации для оценки основных характеристик текущего состояния региональных систем среднего профессионального образования; корректировки и обоснования принимаемых управленческих решений.

Ключевые слова: индикаторы состояния системы профессионального образования, модель оценки, кластерный анализ, метод главных компонент, регрессионный анализ, факторы влияния

Для цитирования: Григорьев В. Ю. Методика формирования сбалансированного набора устойчивых индикаторов оценки регионального профессионального образования // Управленческое консультирование. 2023. № 6. С. 90–104.

Methodology for the Formation of a Balanced Set of Sustainable Indicators for Assessing Regional Vocational Education

Valery Yu. Grigorev

Russian Presidential Academy of National Economy and Public Administration, Moscow, Russian Federation; grigorev-vy@ranepa.ru

ABSTRACT

To assess the state of the regional vocational education system, the authors developed an open five-level assessment model, on the basis of which a set of studies was performed, including cluster and factor analysis. Based on the use of the main components method, a methodology for reducing the dimension of an array of indicators for subsequent clustering of regions is proposed.

For the first time, the joint use of regression analysis and the method of main components for obtaining a stable set of indicators for assessing the regional vocational education system was proposed and tested.

* Статья подготовлена в рамках выполнения НИР госзадания РАНХиГС 11.3-2023-1 «Анализ организационно-экономических механизмов повышения эффективности развития региональных систем СПО».

A relationship was found between the stability of the assessment model of the vocational education system and the directions of action of indicators, both within one indicator group and in different indicator groups.

Keywords: indicators of the state of the vocational education system, assessment model, cluster analysis, principal component analysis, regression analysis, influence factors

For citing: Grigorev V. Yu. Methodology for the Formation of a Balanced Set of Sustainable Indicators for Assessing Regional Vocational Education // Administrative consulting. 2023. N 6. P. 90–104.

Решение задачи адекватного распределения регионов Российской Федерации по кластерам в качестве первого шага для определения групповых управленческих политик в тех или иных сферах социально-экономической жизни находит широкое применение в самых различных исследованиях [4; 5; 6; 10; 11]. При этом часто основной акцент делается на проработку «технических» аспектов собственно кластеризации, связанных с выбором методов и алгоритмов, и на «несодержательные» свойства показателей, такие как требования их относительности [13] (в ряде работ эта процедура называется «переход к удельным величинам» [9]), некоррелированности, стандартизация данных и т. д. Обращается внимание на важность статистической значимости показателей, под которой фактически понимается большая дисперсия. Такой подход может привести к слишком малому числу получаемых кластеров с большим числом регионов в самом крупном из них, например, более 70% от общего количества [13], что делает затруднительной выработку сколь угодно разумной групповой политики.

Применяется также подход постобработки результатов кластеризации, заключающийся в выявлении и временном исключении из объектного пространства «нетипичных регионов», искажающих общую картину, с последующим их возвратом уже на основе анализа профилей кластеров [3]. Однако, вероятно, самым распространенным приемом анализа является подключение отраслевой дифференциации внутри региона — либо занятости населения по отраслям [12; 15], либо отраслевой структуры регионального ВРП [1; 9].

Однако ключевая специфика результатов кластеризации обусловлена прежде всего перечнем показателей, на основе которых она реализуется. При увеличении числа показателей начинают негативно сказываться два фактора — трудность интерпретации полученных результатов и излишний уклон в сторону «автоматизации процесса». Нам представляется, что саму кластеризацию следует рассматривать как некий компромисс между работой эксперта и искусственным интеллектом в виде Machine Learning. С одной стороны — существует естественное стремление эксперта на основе своих знаний предмета исследований подобрать «правильные» показатели, которые следует использовать при разбиении на группы. Частным случаем такого подхода является использование одного-двух отдельных показателей, на основании которых и осуществляется разделение. Но в таком случае сложно говорить о комплексном подходе и дальнейшей выработке эффективных управленческих решений. На обратной стороне находится соблазн отдать процесс на откуп информационным технологиям, предоставив в распоряжение компьютера максимально возможный набор исходных данных. А там уже пусть «умная машина» сама вытянет из набора максимальную дисперсию.

Поэтому в процессе кластеризации приходится так или иначе решать задачу понижения размерности исходного массива показателей, и это практически всегда делается с помощью метода главных компонент [1; 9; 16]. Определение необходимого количества индикаторов, позволяющих дифференцировать регионы, может осуществляться при этом на основании зависимости доли объясненной дисперсии

от числа индикаторов [9]. Второй подход связан с включением в базовый набор характеристик первых двух главных компонент уже на старте, когда индексы отраслевой специализации и индустриализации, как раз и представляющие собой первые две главных компоненты, вбирают в себя 78% общей дисперсии количественных характеристик структуры ВРП [1]. Кроме того, авторы делают очень важное замечание о необходимости контроля устойчивости используемых показателей во времени, и мы вернемся к нему чуть позже.

Сугубо машинно-ориентированный подход может привести к проблемам другого свойства, которые можно проиллюстрировать на примере спорта высоких достижений. Есть такой олимпийский вид спорта в составе легкой атлетики, как семиборье. Из самого названия легко понять, что в него входят семь дисциплин, которые сами по себе являются отдельными видами легкой атлетики, например, прыжки в длину или бег с барьерами. Как определить, какие отдельные дисциплины имело смысл объединить в семиборье, чтобы получить сбалансированный набор. Учитывая, что легкоатлетические дисциплины содержат десятки наименований, составить программу семиборья можно было триллионами различных комбинаций. При этом, простое использование метода главных компонент, обеспечивающего эффективный учет разброса данных, безусловно, позволило бы определить семь видов, максимально разделяющих спортсменов. Однако такой подход мог бы привести к тому, что в состав семиборья одновременно попали бы, например, и толкание ядра, и метание молота, и метание диска, и метание копья, т. е. все виды из силовой части легкой атлетики. А на долю всех остальных беговых, прыжковых и других дисциплин осталось бы всего три места. Ясно, что так поступать нельзя. Необходимо сохранить представительство от содержательно разных по сути групп дисциплин.

Аналогично обстоят дела с кластеризацией регионов по различным индикаторам профессионального образования. Так или иначе экспертами в области образования всегда выделяются группы показателей [2; 7; 14]. В нашем предыдущем исследовании было показано [8], что для оценки состояния региональных систем профессионального образования, таких групп уже выявлено не менее семи, а верхний предел еще предстоит определить. Поэтому целью данного исследования являлась выработка «объективной» методики сокращения массива потенциальных индикаторов оценки профобразования с сохранением осмысленности оставшегося набора без предварительного отбрасывания кандидатов на основании «экспертного мнения», а также проверка — является ли полученный набор индикаторов устойчивым при переходе к другому календарному году (не следует забывать, что любая кластеризация — это распределение по группам на какой-то конкретный момент времени). Таким образом, мы хотим сохранить присутствие всех экспертных индикаторных групп и сохранить их таким образом, чтобы отобранные для кластеризации в финальный массив индикаторы максимально вбирали в себя дисперсию региональных различий.

В развитие данного подхода авторами для оценки состояния региональной системы профессионального образования разработана открытая модель оценки, содержащая пять уровней (табл. 1).

На нулевом уровне модели задается уровень профессионального образования — среднее профессиональное образование (далее СПО), высшее образование или дополнительное профессиональное образование. Поскольку данный уровень модели не участвует в собственно процедуре оценки, он выведен за рамки нумерации. Обратим внимание дополнительно на следующий аспект: в отличие от классического социально-экономического группирования регионов при анализе профессионального образования мы разделяем показатели на индикаторы и факторы. «Индикатор» в явном виде содержит показатель, связанный с образованием, «фактор»

Открытая обобщенная модель оценки состояния региональной системы образования

Table 1. Open generalized model for assessing the state of the regional education system

Номер уровня	Наименование уровня
0-й уровень	Уровень образования
1-й уровень	Уровень индикаторных групп
2-й уровень	Уровень индикаторов
3-й уровень	Уровень главных компонент
4-й уровень	Уровень изменений
5-й уровень	Уровень основных факторов

в явном виде не содержит показатель, связанный с образованием, и, как правило, влияет на индикатор. При этом кластеризацию в конечном счете мы проводим на основе индикаторов, а факторы используются как некие аналоги контрольных переменных в регрессионном анализе.

В общем случае будем считать, что мы имеем N групп потенциальных индикаторов для оценки профобразования. Тогда отдельный индикатор можно обозначить как X_{ij} , где i — номер индикаторной группы, лежащий в диапазоне от 1 до N , а j — номер индикатора в индикаторной группе. Применяя метод главных компонент к общему массиву индикаторов, мы на первом шаге получаем массив вкладов индикаторов в первые две главные компоненты (loadings) — $\{P1X_{ij}, P2X_{ij}\}$, где $P1X_{ij}$ — вклад индикатора X_{ij} в первую главную компоненту, $P2X_{ij}$ — вклад индикатора X_{ij} во вторую главную компоненту.

Существенно, что вклад каждого индикатора в главную компоненту характеризуется двумя параметрами — весом и направлением (геометрически — длиной и направлением вектора на диаграмме). Обозначим «единицей» вклад со знаком плюс, «минус один» — вклад со знаком минус и нулем — нулевой вклад. Очевидно, что может быть восемь комбинаций направлений вклада любого показателя в две главные компоненты (теоретически возможный девятый вариант — два нуля — означает, что индикатор никак не присутствует в составе первых главных компонент) — $\{(1, 1); (1, 0); (1, -1); (0, 1); (0, -1); (-1, 1); (-1, 0); (-1, -1)\}$.

То есть, образно говоря, индикаторы «растягивают» регионы в разные стороны, при этом еще и с разной силой (см., например, рис. 1).

На этом же шаге мы получаем относительные вклады самих главных компонент в разброс данных (importance of components), показывающих, какую долю общей дисперсии вобрали в себя главные компоненты. Обозначим их «P1Comp» и «P2Comp» соответственно. В таком случае численный вклад индикатора в разброс данных можно определить следующим образом:

$$PX_{ij} = |P1X_{ij}| * P1Comp + |P2X_{ij}| * P2Comp.$$

А общий вклад индикатора в разброс данных можно представить как совокупность его численного вклада и направления его действия. Таким образом, первый этап заканчивается построением таблицы вкладов индикаторов примерно следующего содержания (табл. 2; все направления и численные значения вкладов индикаторов определены случайным образом).

Далее осуществляется итерационная процедура понижения размерности массива индикаторов за счет последовательного применения следующих правил:

- от каждой индикаторной группы оставляется как минимум один индикатор;

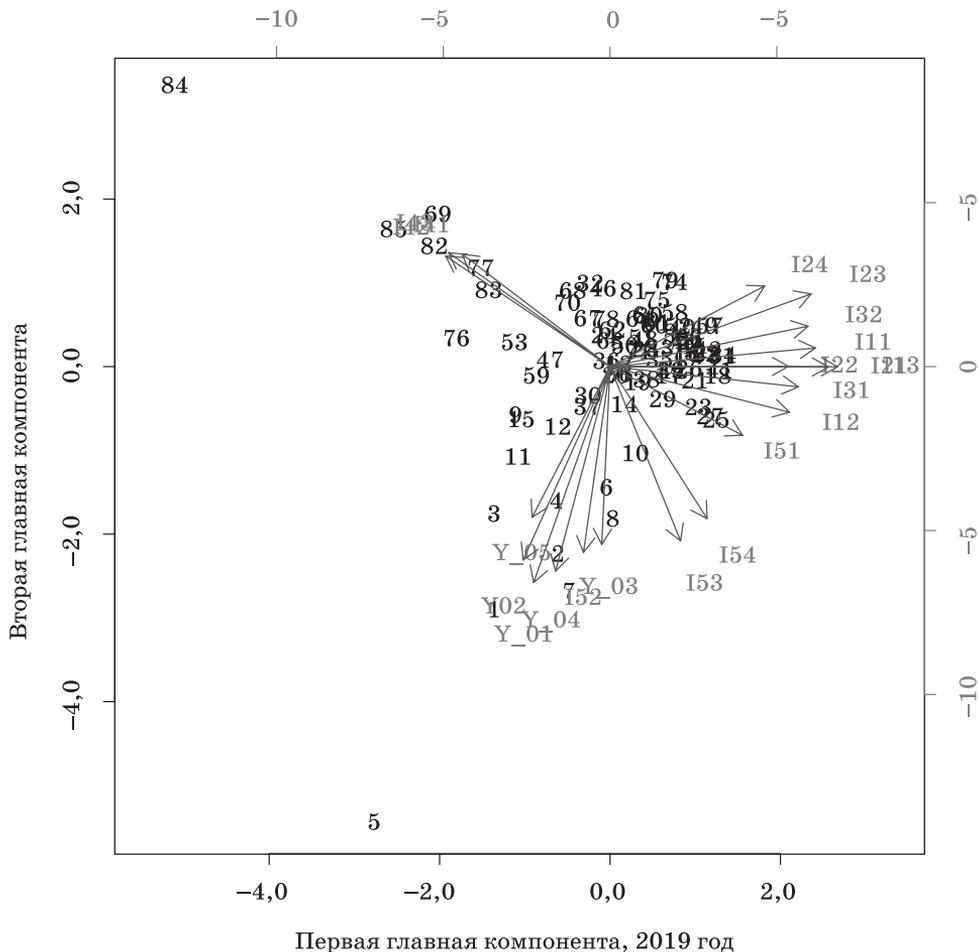


Рис. 1. Образец развертки индикаторов по двум первым главным компонентам
 Fig. 1. Sample drilldown of indicators for the first two main components

- если в рамках одной индикаторной группы присутствуют индикаторы, имеющие разную направленность своего действия, то для следующего этапа анализа оставляются все различные индикаторы;
- из индикаторов одной индикаторной группы, обладающих одинаковым направлением действия, выбирается индикатор, обладающий максимальным численным вкладом в главные компоненты.

После исключения части индикаторов к оставшемуся массиву индикаторов снова применяется метод главных компонент и т. д. Данная процедура повторяется до тех пор, пока не будет исчерпан потенциал понижения размерности — либо останется по одному индикатору в каждой индикаторной группе, либо эти индикаторы будут иметь разную направленность своего действия.

Первичная апробация методики в части проверки устойчивости получаемых индикаторов, что являлось второй задачей исследования, осуществлялась на основе предварительно выявленных пяти индикаторных групп для среднего профессионального образования — «уровень участия населения в СПО», «доступность СПО», «баланс

Пример вкладов индикаторов в первые две главные компоненты
 Table 2. Example of indicator contributions to the first two main components

Индикатор	Направление действия индикатора	PX_{ij}
X11	(1, 1)	0,2070507
X12	(1, 1)	0,2343202
X21	(1, 1)	0,1888490
X22	(1, 0)	0,1843559
X23	(1, 0)	0,1734084
X31	(-1, -1)	0,2027619
X32	(-1, 1)	0,2057174
X41	(0, 1)	0,1272108
X42	(-1, 1)	0,2405748
X43	(-1, 1)	0,2115340

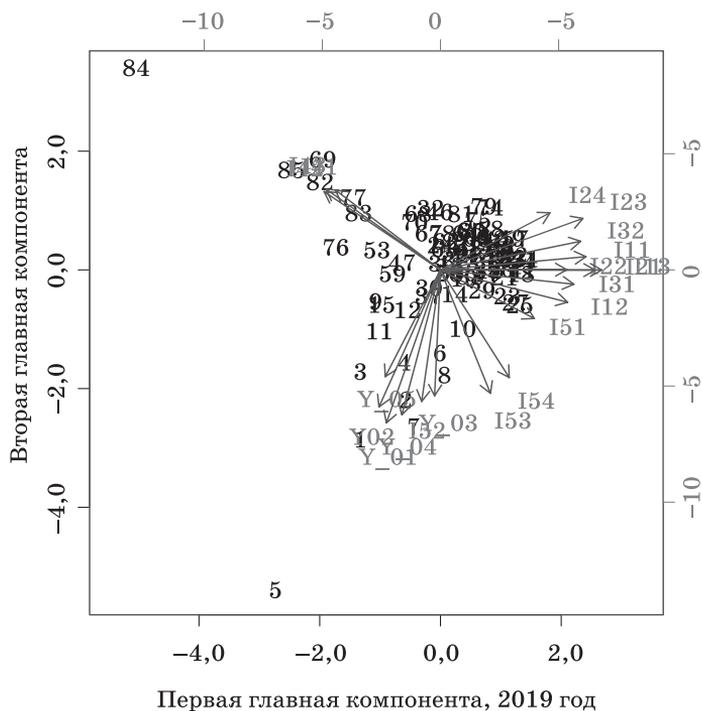
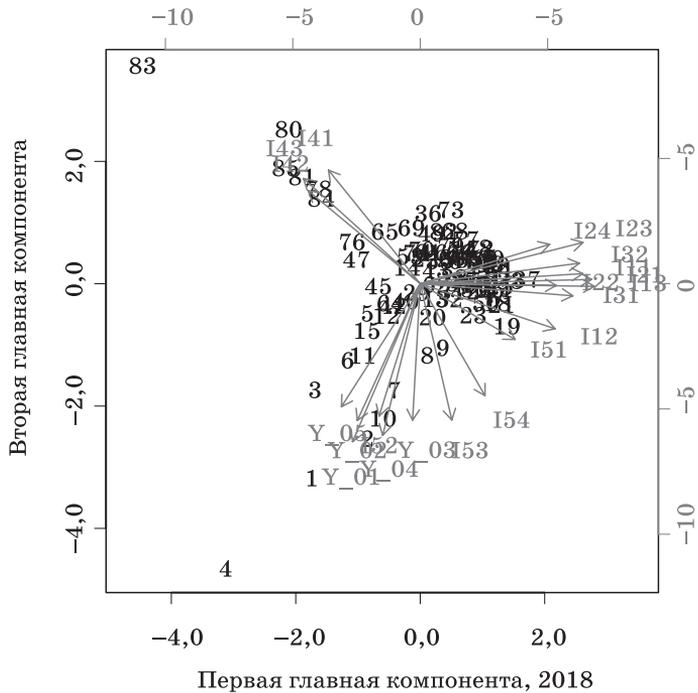
между средним профессиональным и общим образованием», «влияние миграции на поступление учащихся в организации СПО» и «баланс между СПО и рынком труда». Общий первоначальный массив показателей включал семнадцать индикаторов и пять контрольных показателей. Для расчета индикаторов было использовано восемьдесят семь первичных показателей из различных форм статистической отчетности за 2017–2020 гг. Результаты разложения индикаторов и контрольных показателей на главные компоненты для 2018, 2019 и 2020 гг. показаны на рис. 2.

Уже предварительный визуальный анализ показывает, что раскладка индикаторов и контрольных показателей по главным компонентам оставалась неизменной на интервале 2017–2019 гг. (диаграмма за 2017 г. не представлена исключительно в целях экономии места) и кардинально изменилась при переходе от 2019 к 2020 г., что очевидно связано с эффектом пандемии. Для проверки гипотезы о том, что именно и исключительно 2020 г. вносит индивидуальные эффекты по годам, был выполнен отдельный регрессионный анализ на двух различных панелях данных. В качестве образца зависимости был использован следующий сценарный вариант «Зависимая переменная — уровень безработицы граждан со средним специальным или высшим образованием, исследуемая переменная — отношение численности выпускников из организаций среднего профессионального и высшего образования региона к общей численности занятых в регионе». В качестве конкретной формы модели использовался сценарный подвариант линейно-логарифмической модели

$$Y \sim \ln(X) + \mu(i) + \lambda(t) + \varepsilon(it),$$

где $\mu(i)$ — фиксированные эффекты по регионам, $\lambda(t)$ — фиксированные эффекты по годам.

Было подготовлено две панели данных, одна включала данные с 2017 по 2020 г. включительно, вторая — те же данные, но с 2017 по 2019 г. включительно, т.е. без данных за 2020 г. В обоих случаях оценивалась сквозная регрессия (pooled) и модель с индивидуальными эффектами по годам (fixed.time), а затем выполнялся тест Фишера на линейные ограничения — pooled vs fixed.time. В первом случае полученное p-value составило 9.261E-14, т.е. нулевая гипотеза об отсутствии индивидуальных эффектов по годам отвергается. Во втором случае — панель данных 2017–2019 гг. —



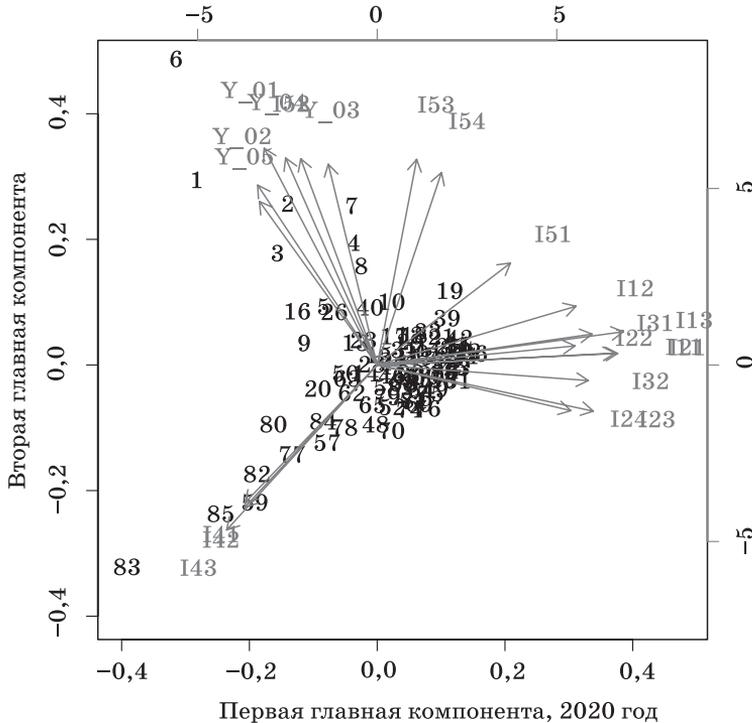


Рис. 2. Развертки индикаторов и факторов по двум первым главным компонентам за 2018–2020 гг. (исходный набор показателей)

Fig. 2. Drills of indicators and factors for the first two main components for 2018–2020 (initial set of indicators)

p-value составило 0,1665, что не позволяет отвергнуть нулевую гипотезу об отсутствии индивидуальных эффектов по годам. Таким образом, исключение из панели показателей за 2020 г. привело к утрате индивидуальных эффектов по годам. При этом, если сравнить диаграммы главных компонент за 2019 и 2020 гг., собранные на основе исключительно полученных по разработанной методике сокращения размерности индикаторов (рис. 3 и 4), то видно, что действие отобранных индикаторов устойчиво. Таким образом, регрессионный анализ позволил нам определить года, на которых следует тестировать устойчивость индикаторной модели.

На следующем этапе был выполнен анализ устойчивости полученной модели при включении в рассмотрение дополнительных индикаторов. Для этой цели были дополнительно рассчитаны два индикатора: «доступность платного СПО — отношение средней стоимости одного семестра обучения в образовательных организациях СПО к средней заработной плате по региону» (идентификатор индикатора *I25*) и «условная эффективность СПО с точки зрения бюджетных расходов — отношение расходов консолидированного бюджета на СПО, скорректированного на индекс бюджетных расходов, к общей численности обучающихся в системе СПО региона, приведенной к очной форме обучения» (идентификатор индикатора *I61*). Существенным моментом, различающим данные индикаторы помимо содержательной составляющей, является то, что первый из них *I25* принадлежит к уже рассмотренной ранее индикаторной группе, а именно к группе 2 — «доступность СПО», а индикатор *I61* начинает собой новую шестую индикаторную группу, которая называется «Экономическая эффективность профобразования».

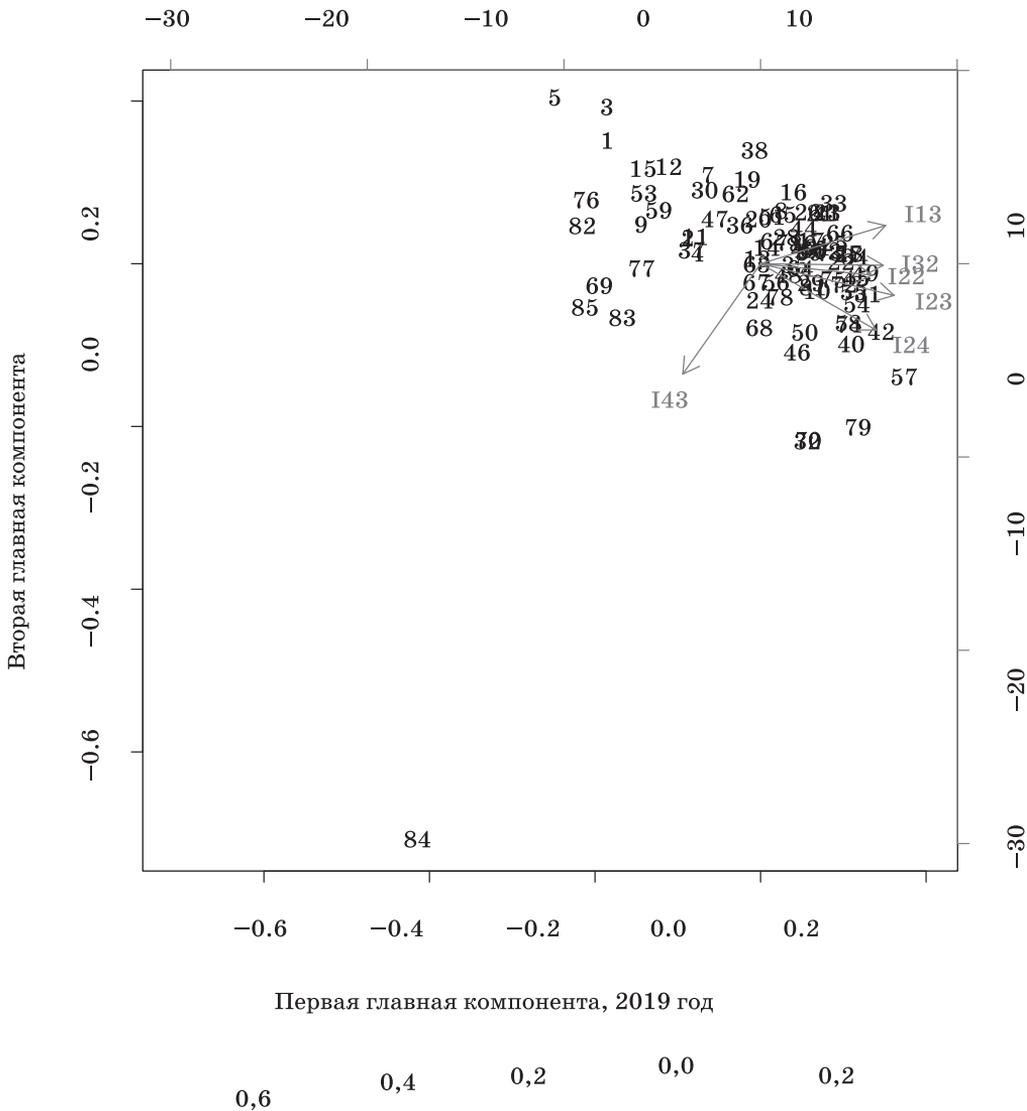


Рис. 3. Развертка индикаторов по двум первым главным компонентам за 2019 г. (финальный набор показателей)
 Fig. 3. Breakdown of indicators for the first two main components for 2019 (final set of indicators)

После добавления индикатора *I25* и выполнения процедуры понижения размерности массива показателей оказалось, что устойчивость модели пропала. Нами была высказана гипотеза, что индикатор *I25* подобран недостаточно корректно, что значение имеет направление действия индикатора, и, что все индикаторы в пределах индикаторной группы должны быть «направлены в одну сторону». Обратим внимание, какие индикаторы входят во вторую группу — «доступность СПО». Первый показатель группы — «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения на базе основного общего обра-

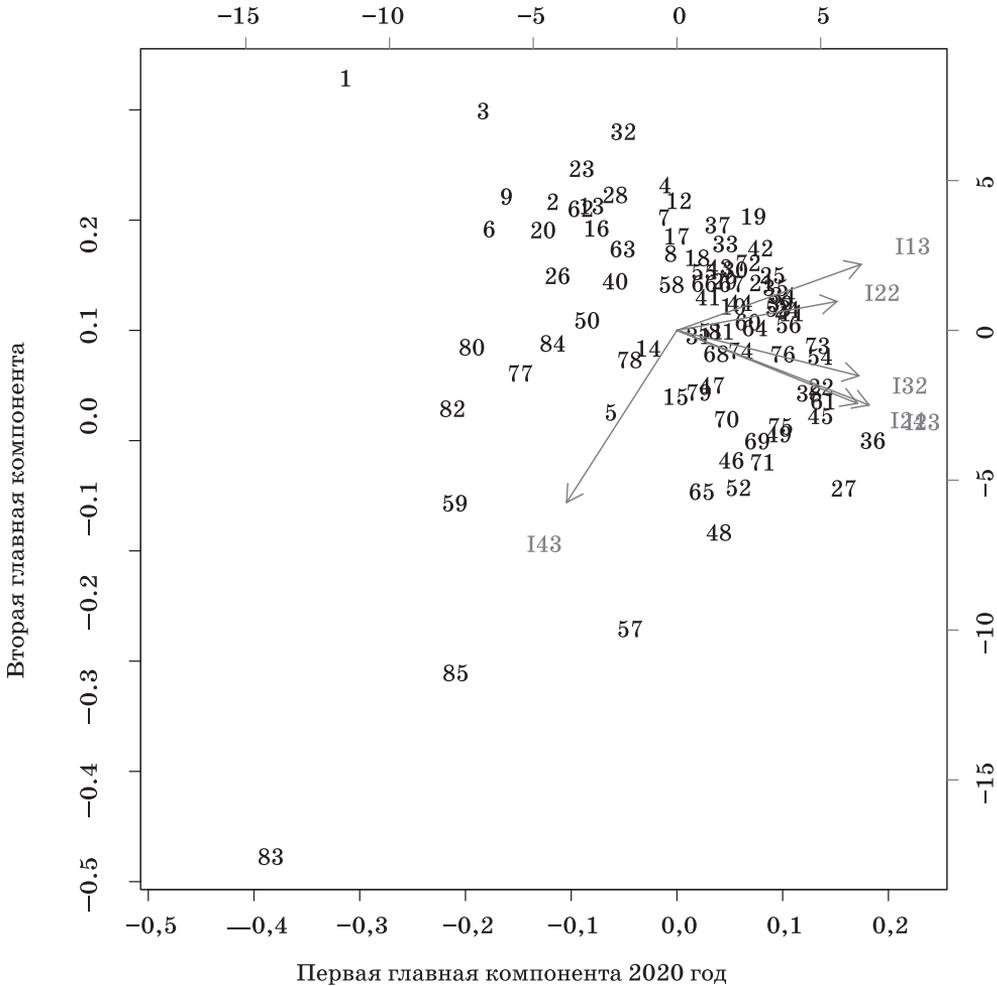


Рис. 4. Развертка индикаторов по двум первым главным компонентам за 2020 г. (финальный набор показателей)

Fig. 4. Breakdown of indicators for the first two main components for 2020 (final set of indicators)

зования к общей численности выпуска из девярых классов школ региона». Очевидно, что чем больше значение данного показателя, тем выше доступность среднего профобразования. Второй показатель группы — «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения на базе основного общего образования к общей численности выпуска из девярых классов школ региона за вычетом численности продолживших обучение в десятых классах школы». Здесь мы имеем аналогичную ситуацию — чем больше значение данного показателя, тем выше доступность среднего профобразования. Такая же ситуация с третьим и четвертым индикаторами — «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения за счет бюджетных ассигнований на базе основного общего образования к общей численности выпуска из девярых классов школ региона» и «отношение общей численности

приема в организации СПО всех форм собственности и всех форм обучения за счет бюджетных ассигнований на базе основного общего образования к общей численности выпуска из девярых классов школ региона за вычетом численности продолживших обучение в десятых классах школы» — рост значений индикаторов приводит к росту доступности СПО в регионе. При этом добавленный индикатор «отношение средней стоимости одного семестра обучения в образовательных организациях среднего профессионального образования к средней заработной плате по региону» показывает скорее недоступность среднего профобразования. Для проверки высказанной гипотезы была произведена замена индикатора *I25* на обратный индикатор — «отношение средней заработной платы по региону к средней стоимости одного семестра обучения в образовательных организациях среднего профессионального образования» (идентификатор индикатора *I25_1*). И результаты расчетов подтвердили гипотезу — устойчивость индикаторной модели восстановилась.

И на последнем этапе было исследовано влияние добавления в модель новой индикаторной группы. Для этого была создана индикаторная группа «Экономическая эффективность профобразования», в которую был помещен индикатор «отношение расходов консолидированного бюджета на СПО, скорректированного на индекс бюджетных расходов, к общей численности обучающихся в системе СПО региона, приведенной к очной форме обучения» (идентификатор показателя *I61*). После добавления индикатора *I61* и выполнения процедуры понижения размерности массива показателей устойчивость модели снова нарушилась. По аналогии с индикатором *I25* нами было сделано предположение, что и в случае с индикатором *I61* также может иметь значение направление действия индикатора. Для проверки сформулированной гипотезы была произведена замена индикатора *I61* на обратный индикатор — «численность студентов, обучающихся в системе СПО региона, приведенной к очной форме обучения, которые условно обучены за один миллион рублей расходов консолидированного бюджета на СПО, скорректированного на индекс бюджетных расходов» (идентификатор индикатора *I61_1*). Повторный расчет и анализ полученных результатов показал, что устойчивость модели восстановилась.

Применив описанную выше методику, мы получили финальный устойчивый массив индикаторов для кластеризации регионов, содержащий следующие показатели:

- «уровень участия населения в возрасте пятнадцати — девятнадцати лет в СПО» (*I13*);
- «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения на базе основного общего образования к общей численности выпуска из девярых классов школ региона за вычетом численности продолживших обучение в десятых классах школы» (*I22*);
- «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения за счет бюджетных ассигнований на базе основного общего образования к общей численности выпуска из девярых классов школ региона» (*I23*);
- «отношение общей численности приема в организации СПО всех форм собственности и всех форм обучения за счет бюджетных ассигнований на базе основного общего образования к общей численности учащихся в десятых классах школ региона» (*I32*);
- «отношение числа прибывших в регион в возрасте пятнадцати или шестнадцати лет к общей численности приема в организации СПО всех форм собственности и всех форм обучения в возрасте пятнадцати или шестнадцати лет» (*I43*);
- «отношение общей численности выпускников всех форм собственности и всех форм обучения из системы СПО к общей численности занятых в экономике региона, имеющих среднее профессиональное образование» (*I52*);

- «отношение общей численности выпускников всех форм собственности и всех форм обучения из систем СПО или ВО к общей численности занятых в экономике региона (I54);
- «численность студентов, обучающихся в системе СПО региона, приведенной к очной форме обучения, которые условно обучены за один миллион рублей расходов консолидированного бюджета на СПО, скорректированного на индекс бюджетных расходов» (I61).

Для определения числа кластеров использован анализ графика изменения межкластерного расстояния в процессе собирания регионов в кластеры, который показал, что оптимальным в данном случае представляется разбиение множества на десять групп, в результате которого получены следующие кластеры (табл. 3).

Таблица 3

Распределение регионов Российской Федерации по кластерам на основе восьми индикаторов оценки системы профессионального образования, полученных по предлагаемой методике

Table 3. Distribution of regions of the Russian Federation by clusters based on eight indicators of assessment of the vocational education system obtained according to the proposed methodology

Кла-стер	Число регионов в кластере	Регионы
1	1	Чеченская Республика
2	6	Республика Тыва, Республика Дагестан, Республика Ингушетия, Республика Саха (Якутия), Кабардино-Балкарская Республика и Республика Карелия
3	8	Республика Калмыкия, Республика Северная Осетия — Алания, Забайкальский край, Республика Адыгея, Новосибирская область, Республика Крым, Тюменская область без автономных округов и Томская область
4	10	...
5	22	...
6	10	...
7	21	...
8	5	Ямало-Ненецкий автономный округ, Еврейская автономная область, Ханты-Мансийский автономный округ, Московская область и Ленинградская область
9	2	г. Санкт-Петербург и г. Москва
10	1	Чукотский автономный округ

Заключение

Для оценки состояния региональной системы профессионального образования разработана открытая пятиуровневая модель оценки, на базе которой выполнен комплекс исследований, включающий кластерный и факторный анализ. На основе использования метода главных компонент предложена методика понижения размерности массива индикаторов для последующей кластеризации регионов.

Впервые предложено и апробировано совместное использование регрессионного анализа и метода главных компонент для получения устойчивого на-

бора индикаторов для оценки региональной системы профессионального образования.

Обнаружена зависимость между устойчивостью модели оценки системы профессионального образования и направлениями действия индикаторов, как внутри одной индикаторной группы, так и в разных индикаторных группах.

Литература

1. Айвазян С. А., Афанасьев М. Ю., Кудров А. В. Индикаторы основных направлений социально-экономического развития и их агрегаты в пространстве характеристик региональной дифференциации // Прикладная эконометрика. 2019. Т. 54. С. 51–69. DOI: 10.24411/1993-7601-2019-10003.
2. Алашеев С. Ю., Кутейницына Т. Г., Посталюк Н. Ю., Прудникова В. А. Управленческая фокусировка региональной системы профессионального образования на потребности социально-экономического развития // Образование и наука. 2021. № 10. С. 44–77. DOI: 10.17853/1994-5639-2021-10-44-77
3. Гавриленко Ю. Е. Методы устойчивой кластеризации регионов России по занятости населения // Федерализм. 2022. Т. 27. № 3 (107). С. 160–177. DOI: 10.21686/2073-1051-2022-3-160-177
4. Гичиев Н. С. Кластерный анализ в экономике: теоретический аспект // Региональные проблемы преобразования экономики. 2020. № 8 (118). С. 176–186.
5. Голова И. М., Суховой А. Ф. Дифференциация стратегий инновационного развития с учетом специфики российских регионов // Экономика региона. 2019. Т. 15, вып. 4. С. 1294–1308. DOI: 10.17059/2019-4-25
6. Гордополов Ю. В., Лукашевич Н. С. Кластеризация регионов по уровню социально-экономического развития на основе самоорганизующихся карт Кохонена // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Экономические науки. 2010. Вып. 3. С. 27–33.
7. Григорьев В. Ю. Оценка и анализ образования: что оцениваем? // Управленческое консультирование. 2019. № 11. С. 145–154. DOI: 10.22394/1726-1139-2019-11-145-154.
8. Григорьев В. Ю. Подходы к оценке состояния системы регионального профобразования // Управленческое консультирование. 2022. № 10. С. 84–98. DOI: 10.22394/1726-1139-2022-10-84-98.
9. Кетова К. В., Касаткина Е. В., Вавилова Д. Д. Кластеризация регионов Российской Федерации по уровню социально-экономического развития с использованием методов машинного обучения // Экономические и социальные перемены: факты, тенденции, прогноз. 2021. Т. 14. № 6. С. 70–85. DOI: 10.15838/esc.2021.6.78.4.
10. Костина С. Н., Трынов А. В. Кластерный анализ динамики рождаемости четвертых и последующих детей в регионах Российской Федерации // Экономические и социальные перемены: факты, тенденции, прогноз. 2021. Т. 14. № 3. С. 232–245.
11. Кузнецов В. Н. Применение кластерного анализа для оценки транспортной системы регионов РФ // Университетская наука. 2020. № 1 (9). С. 71–73.
12. Орлов Д. А., Постников Е. А. Кластеризация региональных рынков труда с учетом оценок NAIRU // Вестник Южно-Уральского государственного университета. Сер. «Экономика и менеджмент». 2021. Т. 15. № 3. С. 34–44.
13. Протасов Ю. М., Юров В. М. Кластеризация регионов РФ по уровню их социально-экономического развития // Вестник Московского государственного областного университета. Сер.: Экономика. 2022. № 2. С. 95–103. DOI: 10.18384/2310-6646-2022-2-95-103.
14. Сирая Г. С. Ключевые аспекты развития системы среднего профессионального образования региона (на примере Ростовской области) // Вестник Воронежского государственного университета. Сер.: Экономика и управление. 2021. № 3. С. 109–122. DOI: 10.17308/escp.2021.3/3625
15. Шубат О. М., Караева А. П. Кластерный анализ в исследовании социально-экономических процессов: опыт критического // Проблемы моделирования социальных процессов: Россия и страны АТР: материалы Второй всероссийской научно-практической конференции с международным участием. Владивосток, 2016. С. 325–328.

16. *Omuya E. O., Okeyo G. O., Kimwele M. W.* Feature selection for classification using principal component analysis and information gain. *Expert Systems with Applications*, 2021, vol. 174, 114765. DOI: 10.1016/j.eswa.2021.114765

Об авторе:

Григорьев Валерий Юрьевич, ведущий научный сотрудник Центра экономики непрерывного образования Института прикладных экономических исследований РАНХиГС (Москва, Российская Федерация), кандидат технических наук, доцент, лауреат Премии Президента России в области образования; grigorev-vy@ranepa.ru

References

1. Ayvazyan S.A., Afanasyev M.Yu., Kudrov A.V. Indicators of the main directions of socio-economic development and their aggregates in the space of characteristics of regional differentiation // *Applied econometrics [Prikladnaya ekonometrika]*. 2019. Vol. 54. P. 51–69. DOI: 10.24411/1993-7601-2019-10003 (in Rus).
2. Alasheev S.Yu., Kuteinitsyna T.G., Postalyuk N.Yu., Prudnikova V.A. Managerial focus of the regional vocational education system on the needs of socio-economic development // *Education and science [Obrazovanie i nauka]*. 2021. N. 10. P. 44–77. DOI: 10.17853/1994-5639-2021-10-44-77 (in Rus).
3. Gavrilenko Yu.E. Methods of sustainable clustering of Russian regions by employment // *Federalism [Federalizm]*. 2022. Vol. 27. N. 3 (107). P. 160–177. DOI: 10.21686/2073-1051-2022-3-160-177 (in Rus).
4. Gichiev N.S. Cluster analysis in economics: theoretical aspect // *Regional problems of economic transformation [Regional'nye problemy preobrazovaniya ekonomiki]*. 2020. N. 8 (118). C. 176–186 (in Rus).
5. Head I.M., Sukhoi A.F. Differentiation of innovative development strategies taking into account the specifics of Russian regions // *Regional economy [Ekonomika regiona]*. 2019. Vol. 15, issue 4. P. 1294–1308. DOI: 10.17059/2019-4-25 (in Rus).
6. Gordopolov Yu.V., Lukashevich N.S. Clustering regions in terms of socio-economic development on the basis of self-organizing Kohonen maps // *Scientific and technical statements of St. Petersburg State Polytechnic University. Economic sciences [Nauchno-tekhnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politekhnicheskogo universiteta. Ekonomicheskie nauki]*. 2010. N 3. P. 27–33 (in Rus).
7. Grigoriev V.Yu. Education assessment and analysis: what do we evaluate? // *Administrative consulting [Upravlencheskoe konsul'tirovanie]*. 2019. N. 11. P. 145–154. DOI: 10.22394/1726-1139-2019-11-145-154 (in Rus).
8. Grigoriev V.Yu. Approaches to assessing the state of the regional vocational education system // *Administrative consulting [Upravlencheskoe konsul'tirovanie]*. 2022. N. 10. P. 84–98. DOI: 10.22394/1726-1139-2022-10-84-98 (in Rus).
9. Ketova K.V., Kasatkina E.V., Vavilova D.D. Clustering of regions of the Russian Federation by the level of socio-economic development using machine learning methods // *Economic and social changes: facts, trends, forecast [Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz]*. 2021. Vol. 14. N. 6. P. 70–85. DOI: 10.15838/esc.2021.6.78.4 (in Rus).
10. Kostina S.N., Trynov A.V. Cluster analysis of the dynamics of the birth rate of fourth and subsequent children in the regions of the Russian Federation // *Economic and social changes: facts, trends, forecast [Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz]*. 2021. Vol. 14. N. 3. P. 232–245 (in Rus).
11. Kuznetsov V.N. Application of cluster analysis to assess the transport system of the regions of the Russian Federation // *University Science [Universitetskaya nauka]*. 2020. N. 1 (9). P. 71–73 (in Rus).
12. Orlov D.A., Postnikov E.A. Clustering of regional labor markets taking into account NAIRU estimates // *Bulletin of South Ural State University. Economics and Management Ser. [Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Ser. «Ekonomika i menedzhment»]*. 2021. Vol. 15. N. 3. P. 34–44 (in Rus).
13. Protasov Yu.M., Yurov V.M. Clustering of regions of the Russian Federation by the level of their socio-economic development // *Bulletin of Moscow State Regional University. Ser.: Economics [Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Ser.: Ekonomika]*. 2022. N. 2. P. 95–103. DOI: 10.18384/2310-6646-2022-2-95-103 (in Rus).

14. Siraya G. S. Key aspects of the development of the system of secondary vocational education in the region (on the example of the Rostov region) // Bulletin of Voronezh State University. Ser.: Economics and Governance [Vestnik Voronezhskogo gosudarstvennogo universiteta. Ser.: Ekonomika i upravlenie]. 2021. N. 3. P. 109–122. DOI: DOI: 10.17308/econ.2021.3/3625 (in Rus).
15. Shubat O. M., Karaeva A. P. Cluster analysis in the study of socio-economic processes: experience of critical // Problems of modeling social processes: Russia and the Asia-Pacific countries: Materials of the Second All-Russian Scientific and Practical Conference with international participation. Vladivostok, 2016. P. 325–328 (in Rus).
16. Omuya E. O., Okeyo G. O., Kimwele M. W. Feature selection for classification using principal component analysis and information gain. *Expert Systems with Applications*, 2021, vol. 174, 114765. DOI: 10.1016/j.eswa.2021.114765

About the author:

Valery Yu. Grigorev, Leading researcher of Institute for Applied Economic Research of RANEPА (Moscow, Russian Federation), PhD in Technical Sciences, Associate Professor, Laureate of the President of the Russian Federation Award in the Field of Education; grigorev-vy@ranepa.ru