

УДК 339.91:339.97

А. В. Кашенов

Институт демографических исследований ФГБУН Федерального научно-исследовательского социологического центра РАН, Москва, e-mail: avkash@list.ru

ПОИСКОВАЯ ТИПОЛОГИЯ СТРАН МИРА. ЧАСТЬ 3. ФАКТОРНЫЙ АНАЛИЗ ПОКАЗАТЕЛЕЙ

Ключевые слова: страны мира, классификация, типология, статистические показатели, факторный анализ.

В данной третьей статье серии продолжается предварительное исследование показателей, исходя из которых впоследствии будет произведена типология 197 стран мира. 60 показателей разбиваются на группы посредством факторного анализа. Метод позволяет в широкой совокупности показателей выделить те, которые наиболее сильно определяют сходство и различие стран мира. Он также позволяет повторить и дополнить результаты корреляционного анализа. Латентные объединения статистических показателей (факторы), выделенные в ходе настоящего исследования, получают названия, которые одновременно содержат в себе их интерпретацию. Предлагаются следующие наименования 10 факторов: уровень социально-экономического развития, размеры экономики, населения и территории, уровень вмешательства государства в экономику, динамика населения, темпы и источники экономического роста, плотность населения и избыток трудовых ресурсов, бедность и неравенство доходов, инфляция, государственный долг, уровень человеческого развития и науки. При необходимости можно использовать список факторов в качестве нового перечня показателей, который заменит те несколько десятков, которые были актуальны в начале исследования. Но главный результат, предположительно, состоит в углублении понимания причин и механизмов социально-экономической дифференциации стран мира.

A. V. Kashenov

The Institute of Demographic Research of the Federal State Budgetary Institution of the Federal Research Sociological Center of the Russian Academy of Sciences, Moscow, e-mail: avkash@list.ru

THE SEARCH TYPOLOGY OF THE COUNTRIES OF THE WORLD. PART 3. FACTOR ANALYSIS OF INDICATORS

Keywords: world countries, classification, typology, statistical indicators, factor analysis.

This third article in the series continues the preliminary study of indicators, based on which a typology of 197 countries of the world will be subsequently produced. 60 indicators are divided into groups by means of factor analysis. The method allows us to identify in a wide range of indicators those that most strongly determine the similarity and difference of the countries of the world. It also allows you to repeat and supplement the results of correlation analysis. Latent combinations of statistical indicators (factors) identified in the course of this study are given names that simultaneously contain their interpretation. The following names of 10 factors are proposed: the level of socio-economic development, the size of the economy, population and territory, the level of government intervention in the economy, population dynamics, rates and sources of economic growth, population density and excess labor, poverty and income inequality, inflation, government debt, the level of human development and science. If necessary, you can use a list of factors as a new list of indicators that will replace the several dozen that were relevant at the beginning of the study. But the main result, presumably, is a deeper understanding of the causes and mechanisms of socio-economic differentiation of the countries of the world.

Введение

Целью серии статей «Поисковая типология стран мира» является разработка версии разделения стран и территорий мира на группы для анализа их социально-экономического развития. Поисковая типология производится путем автоматического распознавания дифференциации объектов.

В отечественной науке подобную методологию первой применила Т.И.Заславская для типологии регионов СССР [1].

В первой статье серии был предложен обзор типологий стран, разработанных ранее органами ООН и отечественными учеными. Было обосновано использование в типологическом страноведении международ-

ных баз данных и перечень экономических, демографических, социальных показателей, которые могут быть для этого использованы. Во второй статье были представлены расчеты коэффициентов корреляции и произведен их выборочный обзор. **Целью данной работы** является факторный анализ выбранных для исследования показателей и интерпретация результатов.

Материалы и методы исследования

В серии статей используются методы сравнительного, корреляционного, факторного, кластерного анализа, ориентированные на реализацию методологии типологической классификации. Статистические расчеты производятся в программах Excel, Stadia, и GRETЛ. Используются статистические данные ООН, МВФ и других международных организаций, перечень которых с подробными ссылками был представлен в первой статье.

Факторный анализ с социально-экономических показателей

Под факторным анализом обычно подразумевается перечисление и описание причин, действующих сил и механизмов определенного процесса или явления. Вот как, например, С.В.Олимпиева представляла (со ссылкой на МВФ) факторы развития экономики России в 2021 г.: «В целом экономическая ситуация в России определяется тремя основными факторами. Первый фактор – сохранение прочной макроэкономической основы...Второй фактор – это структурные проблемы, ограничивающие рост...Третий фактор – международные санкции...» [1, с.188]. Аналогичным образом трактуется данное понятие в региональной политологии [2], экономической статистике [3] и экономико-математическом моделировании [4].

В данной статье используется другое понимание: «факторный анализ – это метод статистического анализа, который используется для изучения структуры многомерных данных. Его основная цель – выявить скрытые переменные (факторы), которые объясняют взаимосвязи между наблюдаемыми явлениями»¹. При необходимости

метод может использоваться для обобщения (генерализации) и сокращения числа переменных. Методология факторного анализа разрабатывается зарубежными [5,6] и российскими статистиками [7,8,9]. Этот метод используется для поиска скрытых (латентных) структур и закономерностей в психологии и социологии, как например у Е.Е.Фоминой [10], исследования распространения инноваций в мировой экономике и построения рейтингов [11] и решения аналогичных задач в различных науках, направлениях и проектах. Метод также используется в многофакторной классификации регионов России как промежуточная ступень между корреляционным анализом данных и кластерным анализом [12].

Факторный анализ в статистическом понимании не претендует на нахождение «причин» и «следствий», «механизмов» развития социально-экономических процессов, но позволяет найти латентные группировки статистических показателей. Можно сказать, что он классифицирует – но не объекты исследования (регионы или страны) – а ряды статистических данных. Факторы, выделенные в результате такого анализа, характеризуются близкими по форме и размерности к коэффициентам корреляции «факторными нагрузками», отражающими взаимосвязь каждого отдельного показателя с латентным фактором. Каждый фактор характеризуется своим вкладом в полную дисперсию всех показателей. После осуществления расчетов исследователь может предложить наименования полученным факторам и далее волен интерпретировать их как «причины» и «следствия», но сама математическая программа не указывает на то, что является причиной или механизмом, а что следствием.

В представленных в данной статье расчетах были использованы 60 показателей, описанных в нашей первой статье данной серии и подвергнутых корреляционному анализу во второй статье. В процессе подготовки данных к факторному анализу был произведен поиск «выбросов» и «пропусков» по всему массиву исходных данных. Было принято решение не считать «выбросами» преимущество крупных держав над малыми странами (США в экономике, Индии и Китая в демографии, России по площади и т.д.) и не удалять данные по этим странам.

Затем в GRETЛ был осуществлен поиск пропущенных значений: количество наблюдений (строк, то есть стран) с пропущенными

¹ Факторный анализ: что это такое и как его проводить // Practicum.Yandex.ru. URL: <https://practicum.yandex.ru/blog/chto-takoe-faktornyj-analiz-vidy-i-metody> (дата обращения: 01.08.2025).

ми значениями = 121 (60,80%), общее количество пропущенных значений 740 (6,20% всех данных). В статистических программах GRETL и STADIA при множественных сопоставлениях пропуска обрабатываются автоматически, однако их наличие может повлиять на точность классификаций, производимых факторным или кластерным анализом. В эконометрике действует общее правило о замене пропусков средними величинами, медианами, интерполяцией и т.д. В данной работе была произведена замена пропусков на медианы. После этого показатели X1-X60 были стандартизованы в GRETL делением на стандартное отклонение выборки.

Расчеты для данной статьи производились в программе Stadia посредством корреляции Пирсона с вращением варимакс и нормализацией Кайзера, их результаты представлены в таблице. Показатели X1-X60 приводятся в той последовательности, что в таблицах предыдущих статей, только немного сокращены их наименования. Полные ссылки на источники данных были опубликованы в первой статье.

В таблице представлены 10 факторов, которые суммарно объясняют 73,031% полной дисперсии 60 показателей.

Первый из выделенных программой Stadia факторов является самым значимым – на него приходится 27,62% полной дисперсии показателей. Он сильно положительно связан с основными показателями уровня экономического развития – подушевого номинального ВВП (X5) и ВВП по паритету покупательной способности на душу населения (X6). Высокую нагрузку по этому фактору также имеют ряды данных, которые эксперты рассматривают как косвенные показатели уровня развития – доля услуг в ВВП (X18), уровень урбанизации (X30), индекс человеческого развития (X32), уровень автомобилизации (X41), потребление энергии на душу населения (X51). Другие ряды данных, имеющие положительную нагрузку по первому фактору, можно назвать индикаторами уровня социального и демографического развития. Как показывают результаты корреляционного анализа, опубликованные в предыдущей статье, они достаточно тесно связаны с показателями ВВП на душу населения.

К этой группе относятся средний возраст населения (X22) и ожидаемая продолжительность предстоящей жизни (X28).

Данные показатели показывают уровень развития медицины и социальной поддержки населения, хотя увеличение доли пожилых возрастов оказывает негативное воздействие на экономику. К группе индикаторов уровня социального развития относятся также средняя заработная плата (X38), обеспеченность населения врачами (X43), обеспеченность интернетом (X53), индекс политической стабильности и отсутствия терроризма (X55), калорийность рациона питания (X56), потребление мяса (X57), и в заключение – индекс демократии (X59) и индекс счастья (X60). Как уже отмечалось в нашей предыдущей статье, парадоксальную близость к «позитивным» показателям социально-экономического развития демонстрирует «негативный» индикатор потребления алкоголя (X45). Здесь он также имеет положительную нагрузку по первому фактору.

Высокие отрицательные нагрузки по первому фактору имеют доля сельского хозяйства в ВВП (X16) и группа демографических характеристик, отражающих прирост населения (X23-X26). Как уже указывалось в предыдущих статьях, статистические векторы демографического и социально-экономического развития в мире имеют противоположную направленность, это проявилось и в факторном анализе. Уровень бедности (X49) также оказался отрицательно связан с первым фактором. Исходя из этого, мы интерпретировали первый фактор, как *уровень социально-экономического развития*. «Узким местом» представленной здесь классификации показателей является то, что не удалось выделить «уровень демографического развития» в отдельный фактор, противоположно направленный по отношению к факторам социально-экономического развития.

Второй фактор объясняет 10,94% полной дисперсии 60 переменных. Он положительно связан с абсолютным размером номинального ВВП (X1) и ВВП по паритету покупательной способности (X4), ВВП ППС в процентах от мирового (X9). Все эти показатели коллинеарны, тем не менее они адекватно отражают величину экономики страны. Также положительную нагрузку по данному фактору имеют численность населения (X19) и коллинеарные с ней численность населения трудоспособного возраста (X34) и численность занятого населения (X35), а также площадь страны (X39).

Результаты факторного анализа (факторные нагрузки) по показателям социально-экономического развития 197 стран мира в 2023 году

	Фактор:	1	2	3	4	5	6	7	8	9	10
		Уровень социально-экономического развития	Размеры экономики, населения и территории	Экономическая недостаточность государства	Медицина и смертность населения	Рост населения и особенности жизни	Структура экономики и инвестиции	Инфляция	Экономический рост и плотность населения	Милитаризация	Безработица и социальное неравенство
Номинальный ВВП	Фактор: X1		0,8697								
Годовой прирост ВВП	X2	27,62	10,94	6,911	5,437	4,522	4,27	3,644	3,467	3,285	2,935
Прирост ВВП за 5 лет	X3								0,9338		
ВВП ППС	X4		0,9819						0,5072		
Номинальный ВВП на душу населения	X5	0,7288									
ВВП ППС на душу населения	X6	0,8119									
Дефицит ВВП в %	X7							-0,9440			
Инфляция в % к предыдущему году	X8							-0,9559			
ВВП ППС страны в % от мирового	X9		0,9819								
Общий объем инвестиций в % от ВВП	X10						-0,5206				
Валовые национальные сбережения в % от ВВП	X11						-0,6697				
Государственные доходы, в % от ВВП	X12			-0,8535							
Государственные расходы, в % от ВВП	X13			-0,8383							

54

[illegible]

[illegible]

Распространение ожирения среди взрослого населения	X46				-0,6961		0,5014					
Употребление табака, в процентах к числен- ности населения	X47						0,5006					
Расходы на образо- вание в % от ВВП	X48				-0,6368		0,6916					
Население ниже уровня бедности, %	X49		-0,5787									
Индекс Джини	X50										0,6396	
Потребление энергии на душу населения в % к среднему миро- вому уровню	X51		0,5920									
Мобильные телефоны на 100 чел населения	X52											
Пользователи интер- нета на 100 чел на- селения	X53		0,8882									
Военные расходы в % от ВВП	X54									0,7484		
Индекс политиче- ской стабильности Всемирного банка	X55		0,6113									
Энергетическая обе- спеченность рациона (ккал/суточная норма)	X56		0,7190									
Потребление мяса – кг/ чел/год	X57		0,7449									
Индекс научного цитирования Nature	X58			0,8679								
Индекс демократии The Economist	X59		0,5906									
Индекс счастья Oxford – Gallup	X60		0,6328									

Положительную нагрузку по второму фактору также имеет индекс научного цитирования Nature (X58), который коррелирует с размером экономики стран и с численностью их населения: лидерами рейтинга Nature с большим отрывом от других стран являются Китай и США. Второй фактор можно интерпретировать как «размер страны». В таблице он назван *размеры экономики, населения и территории*.

Первые два фактора объясняют 38,56% суммарной дисперсии используемых в данном исследовании данных и проявляют статистическую связь с 34 из 60 показателей. Таким образом, можно предположить, что основу дифференцирующих страны признаков в будущей типологии составят уровни развития и размеры подлежащих классификации объектов.

Третий фактор объясняет 6,911% полной дисперсии. Все показатели, связанные с этим фактором, имеют по нему отрицательную нагрузку. Это государственные доходы (X12) и расходы (X13), относительная сила валюты (X15). Как уже было сказано в первых двух статьях, сильная валюта обычно в экономически развитых государствах, которые эффективно участвуют в мировой торговле. В странах, чьи деньги не пользуются спросом на мировом валютном рынке, валюта слабее, в некоторых из них центральные банки специально ослабляют ее, чтобы стимулировать экспорт. Отрицательную нагрузку по третьему фактору имеют также расходы на здравоохранение (X42) и образование (X48), которые как правило, также ниже в менее развитых странах. Исключительное значение величины расходов на образование в развитии экономики отмечают М.Г.Колосницина [13], Е.Я.Арапова [14] и другие авторы. Так что недостаток расходов на образование в ряде стран является серьезной проблемой. В то же время в странах с пониженными государственными расходами, менее распространено ожирение населения – отмечается отрицательная нагрузка по показателю X46. Исходя из того, что основу третьего фактора составляют недостаток государственных расходов и слабость валюты, в таблице этот фактор назван *экономическая недостаточность государства*.

Четвертый фактор объясняет 5,437% полной дисперсии. Положительную нагрузку по нему имеет общий коэффици-

ент смертности (X27), который связан как с недостаточностью развития медицины в слаборазвитых странах, так и со старением населения в развитых. Также положительная нагрузка здесь отмечается по обеспеченности больничными койками (X44) и потреблением алкоголя (X45). И то, и другое свойственно развитым странам. Отрицательная нагрузка наблюдается по показателю температуры по странам (X40). Как уже было сказано в предыдущей статье, отрицательная корреляция географического идентификатора – средней температуры с показателями уровня ВВП на душу населения показывает, что страны, находящиеся дальше от экватора, как правило более экономически развиты. Таким образом, четвертый фактор может рассматриваться как дополнительный к первому в части характеристик уровня социально-экономического развития, развития медицины и демографии. В таблице он назван *медицина и смертность*.

На пятый фактор приходится 4,522% полной дисперсии исследуемых данных. Отрицательную нагрузку по нему имеют годовой прирост населения (X23) и коэффициент миграционного прироста (X31). Положительную нагрузку проявляют «вредные привычки» – распространение ожирения (X46), употребление табака (X47). Пятый фактор интерпретируется как *прирост населения и особенности образа жизни*.

Шестой фактор предопределяет 4,27% полной дисперсии данных. Он имеет положительную нагрузку по доле сферы услуг в ВВП (X18), отрицательную нагрузку по общему объему инвестиции (X10), по валовым национальным сбережениям (X11), доле промышленности в ВВП (X17). Отметим, что государственные инвестиции имеют значение для обеспечения устойчивого развития [15]. Исходя из перечисленных факторных нагрузок данный фактор назван *структура экономики и инвестиции*.

На седьмой фактор приходится 3,644% полной дисперсии. Высокую отрицательную нагрузку по этому фактору демонстрируют дефлятор ВВП (X7) и годовая инфляция (X8). В таблице он назван *инфляция*.

Восьмой фактор предопределяет 3,467% полной дисперсии всех показателей. Положительную нагрузку по этому фактору имеют годовой прирост ВВП (X2), прирост ВВП за 5 лет (X3) и плотность населения (X21),

поэтому в таблице он обозначен, как *экономический рост и плотность населения*.

Девятый фактор предопределяет 3,285% полной дисперсии и положительную нагрузку по нему имеет показатель доли военных расходов в ВВП (X54). Как было сказано в предыдущей статье, корреляционный анализ показывает, что уровень военных расходов не связан с экономическими, демографическими и социальными показателями, а также с уровнем демократии, а диктуется в основном геополитическими обстоятельствами. Данный фактор в таблице назван *милитаризация*.

Процент полной дисперсии, определяемой десятым фактором, составляет 2,935%. Положительную нагрузку по этому фактору имеют уровень безработицы (X37) и индекс Джини (X50). Поэтому фактор обозначен *безработица и социальное неравенство*.

Заключение

Таким образом, посредством факторного анализа была произведена генерализация и разделение на группы всей совокуп-

ности показателей, которые в 1 и 2 частях исследования были предназначены для разработки поисковой типологии стран мира. Были предложены следующие наименования 10 факторов (групп взаимосвязанных показателей), которые в дальнейшем могут быть полезны для различных исследований и экспертных работ в области страноведения: уровень социально-экономического развития, размеры экономики, населения и территории, экономическая слабость государства, медицина и смертность населения, рост населения и особенности образа жизни, структура экономики и инвестиции, инфляция, экономический рост и плотность населения, милитаризация, безработица и социальное неравенство. Одним из результатов произведенного факторного анализа является углубление понимания причин и механизмов социально-экономической дифференциации стран мира. В следующей статье серии будет представлен кластерный анализ и типология стран мира по 60 экономическим, демографическим и социальным показателям.

Библиографический список

1. Олимпиева С.В. Факторный анализ перспектив развития экономики российской Федерации в период посткризисного этапа восстановления от пандемии COVID-19 // Заметки ученого. 2021. № 12-2. С. 187-193. EDN: SGLSRJ.
2. Авксентьев В.А., Гриценко Г.Д. Северный Кавказ: факторный анализ и прогнозы динамики региональной ситуации // Регионология. 2023. Т. 31, № 3(124). С. 442-458. DOI: 10.15507/2413-1407.124.031.202303.442-458. EDN: NIMYPC.
3. Арженовский С.В., Орлова Ю.А., Семерикова Е.В., Сидорова Е.Е. Факторы и индикаторы теневой экономики России: эмпирический анализ региональных данных // Вопросы статистики. 2022. Т. 29, № 5. С. 17-34. DOI: 10.34023/2313-6383-2022-29-5-17-34. EDN: LUXFFH.
4. Куранов Г.О., Стрижкова Л.А., Тишина Л.И., Куранов А.Г. Факторы экономического роста в 2011-2021 годах и их отражение в макроэкономических моделях // Вопросы статистики. 2022. Т. 29, № 3. С. 5-25. DOI: 10.34023/2313-6383-2022-29-3-5-25. EDN: OWCFPJ.
5. Hill R.C., Griffiths W.E., Lim G.C. Principles of Econometrics, Wiley, 2011.
6. Mikel Aickin и Helen Gensler. Adjusting for Multiple Testing When Reporting Research Results: The Bonferroni vs Holm Methods // American Journal of Public Health. 1996. Vol. 86. No. 5.
7. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере / под ред. В.Э. Фигурнова. М.: Финансы и статистика, 1999. 380 с.
8. Малова А.С. Основы эконометрики в среде GRETL. М.: ООО «Проспект», 2016. 112 с. ISBN 978-5-392-20334-5. EDN: VMWDQF.
9. Кулаичев А.П. Методы и средства анализа данных в среде Windows: Stadia. М.: НПО «Информатика и компьютеры», 1999. 340 с.
10. Фомина Е.Е. Факторный анализ и категориальный метод главных компонент: сравнительный анализ и практическое применение для обработки результатов анкетирования // Гуманитарный вестник. 2017. № 10(60). С. 3. DOI: 10.18698/2306-8477-2017-10-473. EDN: ZRTACL.

11. Овчаров А.О., Терехов А.М. Построение странового рейтинга цифровизации экономики на основе метода главных компонент // Вопросы статистики. 2023. Т. 30, № 1. С. 58-69. DOI: 10.34023/2313-6383-2023-30-1-58-69. EDN: LPCTGI.
12. Кашепов А.В. Экономический рост, занятость и заработная плата в регионах России: динамика и типология // Экономические и социальные проблемы России. 2024. № 4(60). С. 34-62. DOI: 10.31249/espr/2024.04.02. EDN: HTXTJX.
13. Колосницына М.Г., Ермолина Ю.Е. Государственные расходы на образование и экономический рост: межстрановой анализ // Вопросы статистики. 2021. Т. 28, № 3. С. 70-85. DOI: 10.34023/2313-6383-2021-28-3-70-85. EDN: NLMPTS.
14. Арапова Е.Я. Роль государственных расходов на образование в стимулировании человеческого развития // Мировая экономика и международные отношения. 2022. Т. 66, № 6. С. 53-61. DOI: 10.20542/0131-2227-2022-66-6-53-61. EDN: XMPNLC.
15. Стрелец И.А., Чебанов С.В. Государственные инвестиции и устойчивость мировой экономики // Мировая экономика и международные отношения. 2024. Т. 68, № 6. С. 5-17. DOI: 10.20542/0131-2227-2024-68-6-5-17. EDN: DGMWHA.