

Исследование взаимосвязи производственных функций и социально-экономических показателей российских регионов методом оптимальных разбиений

И. Л. Кирилук^I, А. В. Кузнецова^{II}, О. В. Сенько^{III}

^I Федеральное государственное бюджетное учреждение науки "Институт экономики Российской академии наук", г. Москва, Россия

^{II} Федеральное государственное бюджетное учреждение науки "Институт биохимической физики им. Н.М. Эммануэля" Российской академии наук, г. Москва, Россия

^{III} Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук", г. Москва, Россия

Аннотация. Исследована зависимость вычисленной на основе производственных функций эмерджентной характеристики – отдачи от масштаба производства в регионах Российской Федерации от различных показателей, характеризующих регионы, включая макроэкономические показатели, характеристики процессов экономического воспроизводства и социально-экономических институтов, географические показатели и т.д. Вычисления произведены с использованием метода оптимальных достоверных разбиений, подразумевающего верификацию закономерностей в данных с использованием перестановочных тестов. В результате анализа статистически значимая связь с отдачей была выявлена для значительной доли показателей или их сочетаний, что указывает как на реальное существование взаимозависимости указанных показателей с отдачей, так и на достаточную точность оценок параметров модели Кобба-Дугласа по участвующим в анализе временным рядам.

Ключевые слова: интеллектуальный анализ данных, оптимально достоверные разбиения, верификация закономерностей, перестановочные тесты, производственные функции, отдача от масштаба, региональная экономика, экономическое воспроизводство.

DOI 10.14357/20718632210103

Введение

Феномен непостоянной отдачи от масштаба производства как характеристики производственных функций экономических систем разных уровней является предметом значительного интереса экономистов. В работе [1] введена простейшая модель производственных функций, сводящаяся к линейной зависимости выпускаемого продукта от факторов производства посредством логарифмирования переменных:

$$Y = AK^{\alpha}L^{\beta}\varepsilon, \quad (1)$$

где Y – «выпуск продукции», K – «капитал», L – «труд», A , α , β – вычисляемые параметры, ε – мультипликативный «шум».

В (1) отдача от масштаба производства характеризуется величиной $\alpha + \beta$. Если $\alpha + \beta = 1$, говорят о постоянной отдаче от масштаба, если $\alpha + \beta < 1$ об убывающей и если $\alpha + \beta > 1$ о возрастающей. Если отдача не равна единице, нарушается аддитивность системы, то есть,-domno-

жение труда и капитала на некоторый коэффициент приводит не к увеличению выпуска в такой же степени (к домножению на такую же величину), он меняется слабее, или сильнее, что может быть интерпретировано как результат эмерджентных явлений в системе, определяемых, например, конкуренцией или кооперацией между подсистемами.

Рядом авторов, в том числе, в наших предыдущих расчётах для экономических систем уровня стран и регионов получены значения $\alpha+\beta$, существенно отличающиеся от единицы, причём, как возрастающие, так и убывающие значения отдачи [2-5]. Однако при вычислениях производственных функций по временным рядам показателей выражен такой фактор, как мультиколлинеарность, осложняющая объективную оценку коэффициентов регрессии, приводящая к существенному увеличению дисперсии в их значениях при построении производственных функций по панелям данных (таким, как панель регионов Российской Федерации). При этом, значимое смещение среднего значения по панелям относительно единицы может говорить об объективности выявленного непостоянного характера отдачи. Также представляет интерес установление значимости, закономерности различий $\alpha+\beta$ для разных регионов.

В [5] показана возможность оценки достоверности эффекта непостоянной отдачи методами интеллектуального анализа данных. Если непостоянство отдачи не является артефактом, обусловленным эффектами типа мультиколлинеарности, это может иметь важное экономическое значение, а, следовательно, отдача может значимо зависеть от совокупности других экономических показателей.

Объект исследований

Объектом настоящего исследования являются 79 регионов Российской Федерации, напрямую (а не как часть более крупных регионов) входящие в состав России, для которых доступны наборы необходимых для расчетов отдачи временных рядов достаточной длины (то есть, за исключением Крыма, Севастополя, Чечни).

Данные брались из сборников Росстата [6-8] за 2014 год. Отдача рассчитывалась по временным рядам за 1996-2014 г.г., использовавшимся

в предыдущих работах [5; 9]. В качестве переменных в (1) использованы следующие показатели: Y – валовой региональный продукт, вместо капитала K здесь используются инвестиции в основной капитал (в литературе инвестиции обозначаются буквой I), L – среднегодовая численность занятых в экономике, помноженная на среднемесячную номинальную начисленную заработную плату работающих в экономике. Приведение величин к постоянным ценам, для используемого набора показателей имеющее значение в случае непостоянной отдачи, осуществлялось с использованием индексов потребительских цен. Использование инвестиций вместо капитала в качестве фактора производства в производственных функциях обсуждается, например, в [10].

Для исследуемых данных были рассчитаны производственные функции по формуле (1), определена их отдача для каждого из регионов, определена медиана отдачи по совокупности регионов и в зависимости от соотношений отдачи конкретного региона с медианным (больше или меньше) регионы разделены на две группы.

Далее выяснялась связь отдачи со следующим набором показателей, характеризующих условия экономического воспроизводства в исследуемых регионах (по состоянию на 2014 год): v_1 – площадь территории, тыс. км², v_2 – численность населения на 1 января 2014 г., тыс. человек, v_3 – среднегодовая численность занятых в экономике, тыс. человек, v_4 – среднедушевые денежные доходы (в месяц), руб., v_5 – потребительские расходы в среднем на душу населения (в месяц), руб., v_6 – среднемесячная номинальная начисленная заработная плата работников организаций, руб., v_7 – основные фонды в экономике (по полной учетной стоимости; на конец года), млн руб., v_8 – удельный вес городского и сельского населения в общей численности населения (оценка на конец года; в процентах), v_9 – коэффициенты миграционного прироста на 10000 человек населения, v_{10} – ожидаемая продолжительность жизни при рождении (число лет), v_{11} – уровень безработицы (в процентах), v_{12} – общая площадь жилых помещений, приходящаяся в среднем на одного жителя (на конец года; квадратных метров), v_{13} – выпуск квалифицированных рабо-

чих и служащих на 10 000 человек занятого населения, человек, v_{14} – заболеваемость на 1000 человек населения (зарегистрировано заболеваний у пациентов с диагнозом, установленным впервые в жизни), v_{15} – численность зрителей театров на 1000 человек населения, v_{16} – число посещений музеев, человек на 1000 человек населения, v_{17} – число зарегистрированных преступлений на 100 000 человек населения, v_{18} – валовой региональный продукт (миллионов рублей), v_{19} – валовой региональный продукт на душу населения (рублей), v_{20} – индекс физического объема валового регионального продукта (в постоянных ценах; в процентах к предыдущему году), v_{21} – фактическое конечное потребление домашних хозяйств на душу населения на территории субъектов Российской Федерации (в текущих рыночных ценах; рублей), v_{22} – валовое накопление основного капитала (в текущих рыночных ценах; миллионов рублей), v_{23} – стоимость основных фондов (на конец года; по полной учетной стоимости; миллионов рублей), v_{24} – степень износа основных фондов (на конец года; в процентах), v_{25} – инвестиции в основной капитал (в фактически действовавших ценах; миллионов рублей), v_{26} – инвестиции в основной капитал на душу населения (в фактически действовавших ценах; рублей), v_{27} – индекс физического объема инвестиций в основной капитал (в сопоставимых ценах; в процентах к предыдущему году), v_{28} – индексы промышленного производства (в процентах к предыдущему году), v_{29} – оборот розничной торговли (в фактически действовавших ценах; миллионов рублей), v_{30} – оборот розничной торговли на душу населения (в фактически действовавших ценах; рублей), v_{31} – плотность автомобильных дорог общего пользования с твердым покрытием (на конец года; км путей на 1000 км² территории), v_{32} – число дорожно-транспортных происшествий на 100 000 человек населения, v_{33} – число персональных компьютеров на 100 работников (штук), v_{34} – использование сети интернет населением (по данным выборочного обследования населения; в процентах от общей численности населения в возрасте 15-72 лет соответствующего субъекта Российской

Федерации), v_{35} – численность персонала, занятого научными исследованиями и разработками (человек), v_{36} – инновационная активность организаций (удельный вес организаций, осуществлявших технологические, организационные, маркетинговые инновации, в общем числе обследованных организаций, в процентах), v_{37} – затраты на технологические инновации (миллионов рублей), v_{38} – объем инновационных товаров, работ, услуг, млн руб., v_{39} – доходы консолидированных бюджетов субъектов Российской Федерации в 2014 г. (миллионов рублей), v_{40} – расходы консолидированных бюджетов субъектов Российской Федерации в 2014 г. (миллионов рублей), v_{41} – вклады (депозиты) юридических и физических лиц в рублях, привлеченные кредитными организациями (исходя из места привлечения средств; на начало года; миллионов рублей), v_{42} – сальдированный финансовый результат (прибыль минус убыток) деятельности организаций (миллионов рублей), v_{43} – индексы потребительских цен (декабрь к декабрю предыдущего года; в процентах), v_{44} – среднегодовая численность занятых (тысяч человек), v_{45} – величина прожиточного минимума, установленная в субъектах Российской Федерации за IV квартал 2014 г. (в среднем на душу населения), руб. в месяц, v_{46} – средняя месячная температура воздуха по субъектам Российской Федерации в 2014 г., С, январь, v_{47} – средняя месячная температура воздуха по субъектам Российской Федерации в 2014 г., С, июль, v_{48} – количество осадков по субъектам Российской Федерации в 2014 г., январь, среднее, мм, v_{49} – количество осадков по субъектам Российской Федерации в 2014 г., июль, среднее, мм.

Часть перечисленных показателей использовано при расчётах производственных функций, а именно: переменная Y получается из v_{18} , K из v_{25} и L из произведения v_3 на v_6 приведением к постоянным ценам с помощью v_{43} .

Расчёты показали, что 3 региона из 79 имеют возрастающую отдачу (Сахалинская область, 1.38, Тюменская область, 1.07, Чукотский автономный округ, 1.01), для остальных регионов выполнено соотношение $\alpha + \beta < 1$. Самая низкая отдача, 0.38 - у Магаданской области.

Исследование методом оптимально достоверных разбиений

После того, как значения отдачи были описанным выше образом приведены к бинарному виду, к ним был применён метод оптимально достоверных разбиений (ОДР) [11-13] с целью выявления зависимости отдачи от масштаба с показателями, или их парами. В основе метода лежит поиск параллельных координатным осям границ, оптимальным образом разделяющих объекты с разными уровнями значений прогнозируемой переменной. Качество разбиения с числом элементов (квадрантов), равным r в смысле разделения ими объектов некоторого класса K от всех остальных объектов обучающей выборки в настоящей работе оценивается с помощью функционала

$$Q = \sum_{j=1}^r m_j (v_j - v_0)^2, \quad (2)$$

где v_0 - доля класса K в обучающей выборке в целом, v_j - доля класса в квадранте j , m_j - число описаний объектов внутри квадранта j . Следует отметить, что при $r=2$ функционал Q связан со статистикой критерия χ^2 при проверке гипотезы о связи двух бинарных переменных по таблице сопряжённости размером 2×2 равенством $Q = \chi^2 v_0(1 - v_0)$.

Границы подбираются таким образом, чтобы добиться максимального значения Q . Верификация закономерностей производится с помощью перестановочных тестов [14] при использовании достигнутого максимума Q в качестве статистики критерия. Перестановочный тест заключается в многократном повторении расчётов на случайных выборках, которые получаются из исходной выборки с помощью случайных перестановок значений бинарной целевой переменной относительно фиксированных позиций векторов объясняющих переменных. В качестве меры значимости одномерных закономерностей типа изображённой на Рис. 1 используется p -значение, равное доле случайных выборок, для которых максимальное значение Q , полученное при подборе оптимальных границ, оказалось выше максимального значения Q , полученного на реальных данных. В основе перестановочного теста лежит прямая проверка нулевой гипотезы о независимости целевой переменной от объясняю-

щих переменных. Перестановочный тест не требует предположений о распределениях и не использует никаких асимптотических оценок. При верификации более сложных двумерных закономерностей типа изображённых на Рис. 2 и 3 оценивается значимость вклада каждой из двух объясняющих переменных. Двумерная закономерность, связывающая принадлежность классу K с переменными X_1 и X_2 , считается значимой на некотором уровне α , если значимым на уровне α оказывается вклад каждой из двух переменных. Оценка значимости переменной X_1 сводится к проверке нулевой гипотезы об исчерпывающем описании наблюдаемого эффекта зависимостью лишь от одной переменной X_2 , описываемой с помощью простого одномерного разбиения с одной границей. При этом, исчерпывающее описание подразумевает независимость целевой переменной от объясняющих переменных внутри квадрантов слева и справа от оптимальной границы для X_2 . Проверка данной нулевой гипотезы производится с помощью варианта перестановочного теста, аналогичного используемому при верификации простых одномерных гипотез, но с запретом при генерации случайных выборок перестановок значений целевой переменной между квадрантами. Подробно данный вариант перестановочного теста описан в работе [15]. В настоящей работе при использовании перестановочных тестов генерировалось 3000 случайных выборок. Преимущества метода ОДР - возможность выявления нелинейных эффектов и наглядность представления результатов анализа посредством двумерных диаграмм.

Одномерные разбиения

С помощью ОДР находятся границы в пространстве значений показателей, полученные посредством применения алгоритма оптимизации, разделяющие интервал значений таким образом, что с разных сторон границ преобладают значения для объектов, отнесённых к разным классам. Оценивается эффективность границ с использованием особых функционалов и значимость разбиений по p -значению.

Результаты применения ОДР для одномерных показателей приведены в Табл. 1. Приводятся коды показателей (соответствующие

Табл. 1. Наиболее информативные показатели при одномерных разбиениях (упорядочены по убыванию значимости)

№	Показатель	Граница	Ниже		Выше		χ^2	p-значение
			1 гр.	2 гр.	1 гр.	2 гр.		
1	v18	420694,5625	30 (78,9 %)	14 (34,1 %)	8 (21,1 %)	27 (65,9 %)	15,839	0,000667
2	v22	96987,796875	28 (73,7 %)	12 (29,3 %)	10 (26,3 %)	29 (70,7 %)	15,368	0,001
3	v25	95297,5	29 (76,3 %)	12 (29,3 %)	9 (23,7 %)	29 (70,7 %)	17,265	0,001333
4	v23	1455259	33 (86,8 %)	20 (48,8 %)	5 (13,2 %)	21 (51,2 %)	12,776	0,004
5	v19	273214,375	25 (65,8 %)	10 (24,4 %)	13 (34,2 %)	31 (75,6 %)	13,525	0,004333
6	v40	75419,203125	29 (76,3 %)	15 (36,6 %)	9 (23,7 %)	26 (63,4 %)	12,456	0,005
7	v34	65,150002	18 (47,4 %)	5 (12,2 %)	20 (52,6 %)	36 (87,8 %)	11,673	0,006333
8	v7	833086	26 (68,4 %)	12 (29,3 %)	12 (31,6 %)	29 (70,7 %)	11,957	0,009
9	v39	62566,648438	28 (73,7 %)	14 (34,1 %)	10 (26,3 %)	27 (65,9 %)	12,225	0,009
10	v26	69638,5	27 (71,1 %)	14 (34,1 %)	11 (28,9 %)	27 (65,9 %)	10,624	0,012667
11	v2	3385,800049	38 (100 %)	31 (75,6 %)	0 (0 %)	10 (24,4 %)	10,477	0,013667
12	v45	22713	19 (50 %)	7 (17,1 %)	19 (50 %)	34 (82,9 %)	9,561	0,020333
13	v31	401	37 (97,4 %)	29 (70,7 %)	1 (2,6 %)	12 (29,3 %)	10,049	0,022
14	v6	20986,5	19 (50 %)	7 (17,1 %)	19 (50 %)	34 (82,9 %)	9,561	0,022333
15	v3	1728	38 (100 %)	32 (78 %)	0 (0 %)	9 (22 %)	9,295	0,026
16	v5	15626,5	28 (73,7 %)	16 (39 %)	10 (26,3 %)	25 (61 %)	9,48	0,027333
17	v29	688925	38 (100 %)	32 (78 %)	0 (0 %)	9 (22 %)	9,295	0,029
18	v44	1719,800049	38 (100 %)	32 (78 %)	0 (0 %)	9 (22 %)	9,295	0,032
19	v9	42	37 (97,4 %)	30 (73,2 %)	1 (2,6 %)	11 (26,8 %)	8,85	0,034667
20	v12	25,950001	20 (52,6 %)	34 (82,9 %)	18 (47,4 %)	7 (17,1 %)	8,262	0,037333
21	v35	7501	35 (92,1 %)	25 (64,1 %)	3 (7,9 %)	14 (35,9 %)	8,659	0,048333

названия — в перечне выше), граничные значения, числа наблюдений, вошедших в первую и вторую группы ниже границ, соответствующие числа выше границ. Далее в столбце значения χ^2 , соответствующие оптимальным значениям Q и рассчитанные с помощью перестановочно-го теста p -значения.

Как видно из Табл. 1, существенную часть показателей, для которых $p < 0.05$ составляют показатели, непосредственно связанные с производством, часть из которых используется при расчёте отдачи. Из 50 исследованных показателей, для 21-го имеет место соотношение $p < 0.05$, что сви-

детельствует о высокой достоверности связи отдачи от масштаба с другими показателями.

Представим выявленную зависимость графически в наиболее значимом случае, то есть, в случае с наименьшим p -значением. Он соответствует переменной $v18$ - валовой региональный продукт (миллионов рублей).

Из Рис. 1 наглядно видно существенное преобладание регионов со значениями отдачи больше медианы справа от границы: 27 регионов со значениями отдачи выше медианы против 8 со значениями ниже медианы. Слева от границы преобладают регионы со значениями отдачи ниже медианы: 30 регионов со значениями отдачи ниже медианы против 14 со значениями выше медианы. Статистическая значимость закономерности была с помощью перестановочного теста оценена на уровне $p=0.000667$.

Перечислим регионы по группам, соответствующим Рис. 1. Здесь и далее «высокими» и «низкими» называются значения показателей в

соответствии с расположением относительно обозначенных на рисунках границ (для отдачи относительно медианы).

С высокими значениями отдачи и высоким ВРП:

Томская область, Алтайский край, Удмуртская Республика, Республика Коми, Республика Дагестан, Архангельская область, Омская область, Белгородская область, Приморский край, Республика Саха (Якутия), Ленинградская область, Волгоградская область, Оренбургская область, Кемеровская область, Сахалинская область, Новосибирская область, Пермский край, Челябинская область, Ростовская область, Республика Башкортостан, Свердловская область, Республика Татарстан, Краснодарский край, г. Санкт-Петербург, Московская область, Тюменская область, г. Москва.

С высокими значениями отдачи и низким ВРП:

Еврейская автономная область, Республика Тыва, Республика Калмыкия, Республика

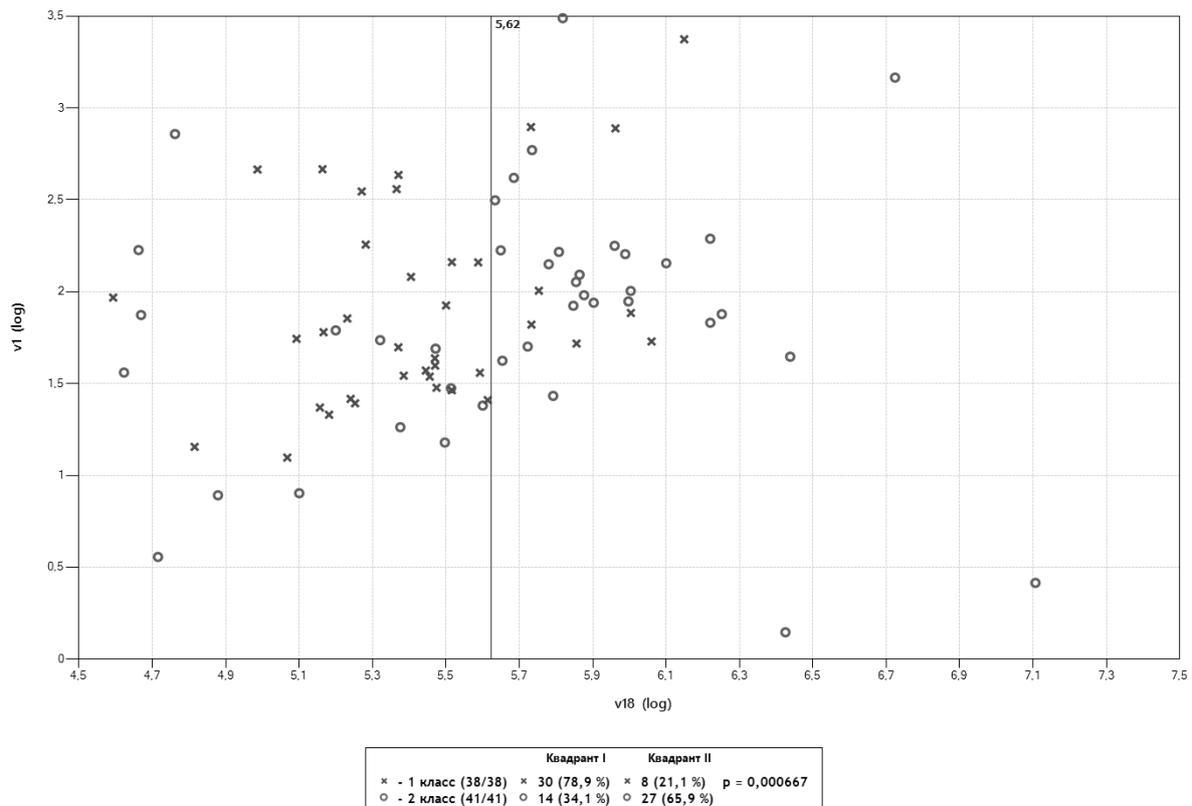


Рис. 1. По оси абсцисс — переменная $v18$ - валовой региональный продукт, по оси ординат — переменная $v1$, используемая в данном случае только для развертки. Кружки обозначают класс регионов со значениями отдачи больше медианы, крестики — со значениями меньше медианы.

Ингушетия, Чукотский автономный округ, Республика Адыгея, Республика Северная Осетия — Алания, Республика Хакасия, Новгородская область, Чувашская Республика, Астраханская область, Калининградская область, Калужская область, Липецкая область.

С низкими значениями отдачи и высоким ВРП:

Хабаровский край, Ставропольский край, Саратовская область, Воронежская область, Иркутская область, Нижегородская область, Самарская область, Красноярский край.

С низкими значениями отдачи и низким ВРП:

Республика Алтай, Карачаево-Черкесская Республика, Магаданская область, Кабардино-Балкарская Республика, Псковская область, Республика Марий Эл, Камчатский край, Костромская область, Ивановская область, Курганская область, Республика Мордовия, Орловская область, Республика Бурятия, Республика Карелия, Амурская область, Смоленская область, Забайкальский край, Брянская область, Кировская область, Ульяновская область, Тамбовская область, Пензенская область, Рязанская область, Курская область, Тверская область, Владимирская область, Мурманская область, Вологодская область, Ярославская область, Тульская область.

Двумерные разбиения

Исследование при помощи двумерных разбиений предполагает разбиение области значений двух переменных на 4 квадранта, выделенных исходя из цели, чтобы в квадрантах наиболее выражено преобладали представители одного из 2-х классов. Наряду с 20 переменными, значимость которых была выявлена с помощью одномерных разбиений, были выявлены также значимые двумерные закономерности с участием переменной v_{37} в сочетаниях с переменными v_{25} и v_{18} , переменной v_{41} в сочетании с переменной v_{25} , переменной v_{11} в сочетании с переменной v_{19} , переменной v_{46} в сочетании с переменной v_{31} , переменной v_1 в сочетании с переменной v_6 , переменной v_{16} в сочетании с переменной v_{21} , переменной v_{32} в сочетании с переменной v_{47} .

Примеры двумерных разбиений представлены на Рис. 2.

Из Рис. 2 видна высокая корреляция между переменными X и Y . При этом регионы с отдачей выше медианных значений отчётливо разбиваются на две группы. Для одной из групп значения пары переменных находятся в квадранте II, задаваемом неравенствами $X > 90056$ млн руб., $Y > 188.3$ млн руб. В этом квадранте находится 30 регионов с высокой отдачей и только 10 регионов с низкой отдачей. Другая подгруппа регионов с высокой отдачей почти целиком сосредоточена в квадранте IV, задаваемом неравенствами $X < 90056$ млн руб., $Y < 188.3$ млн руб. В этом квадранте находится 8 регионов с высокой отдачей и только 2 региона с низкой отдачей. Регионы с отдачей ниже медианного значения преобладают в квадранте II, задаваемом неравенствами $X < 90056$ млн руб., $Y > 188.3$ млн руб. В этом квадранте находится 26 регионов с низкой отдачей и только 3 региона с высокой отдачей. Статистическая значимость этой закономерности оценивается на уровне $p < 0.0005$ по инвестициям в основной капитал и на уровне $p = 0.0143$ по затратам на технологические инновации.

Далее приводится соотнесение конкретных регионов с квадрантами:

Первый квадрант, низкая отдача:

Ярославская область, Рязанская область, Пензенская область, Республика Мордовия, Владимирская область, Курская область, Амурская область, Ульяновская область, Тверская область, Кировская область, Вологодская область, Мурманская область, Смоленская область, Брянская область, Забайкальский край, Республика Марий Эл, Республика Бурятия, Орловская область, Курганская область, Камчатский край, Костромская область, Республика Алтай, Магаданская область, Ивановская область, Кабардино-Балкарская Республика, Псковская область.

Первый квадрант, высокая отдача:

Чувашская Республика, Новгородская область, Калининградская область.

Второй квадрант, низкая отдача:

Красноярский край, Самарская область, Нижегородская область, Иркутская область, Тульская область, Ставропольский край, Воронеж-

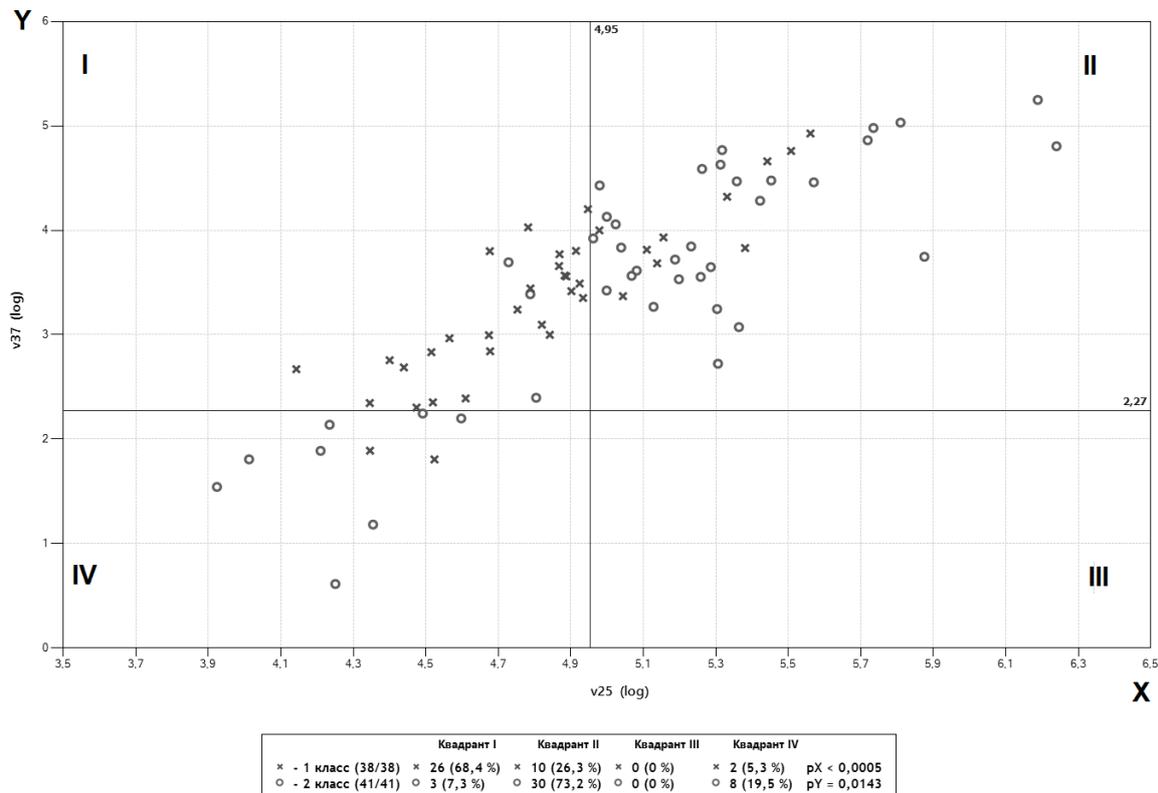


Рис. 2. Вдоль оси X отложены значения переменной v25- инвестиции в основной капитал, вдоль оси Y отложены значения переменной v37- затраты на технологические инновации. Для удобства переменные представлены в логарифмической шкале.

ская область, Хабаровский край, Саратовская область, Тамбовская область.

Второй квадрант, высокая отдача:

г. Москва, Московская область, Республика Татарстан, г. Санкт-Петербург, Тюменская область, Пермский край, Сахалинская область, Волгоградская область, Республика Башкортостан, Челябинская область, Свердловская область, Омская область, Ростовская область, Калужская область, Липецкая область, Удмуртская Республика, Ленинградская область, Томская область, Краснодарский край, Оренбургская область, Новосибирская область, Белгородская область, Астраханская область, Республика Саха (Якутия), Архангельская область, Алтайский край, Приморский край, Республика Коми, Кемеровская область, Республика Дагестан.

Четвертый квадрант, низкая отдача:

Карачаево-Черкесская Республика, Республика Карелия.

Четвертый квадрант, высокая отдача:

Республика Северная Осетия — Алания, Республика Хакасия, Республика Адыгея, Республика Ингушетия, Еврейская автономная область, Чукотский автономный округ, Республика Калмыкия, Республика Тыва.

Из Рис. 3 видно, что низкой отдаче соответствуют регионы, для которых пары значений уровня безработицы и валового регионального продукта на душу населения расположены в левом нижнем квадранте, границы которого задаются неравенствами $X < 5.85\%$, $Y < 308836.438$ руб. В этом квадранте существенно преобладают регионы с низкой отдачей: 22 региона с отдачей выше медианы и только 2 региона с высокой отдачей. Регионы с отдачей выше медианной преобладают в остальных квадрантах. Статистическая значимость этой закономерности оценивается как $p=0.0107$ по уровню безработицы, $p=0.0007$ по валовому национальному продукту на душу населения.

Далее приводится соотнесение конкретных регионов с квадрантами:

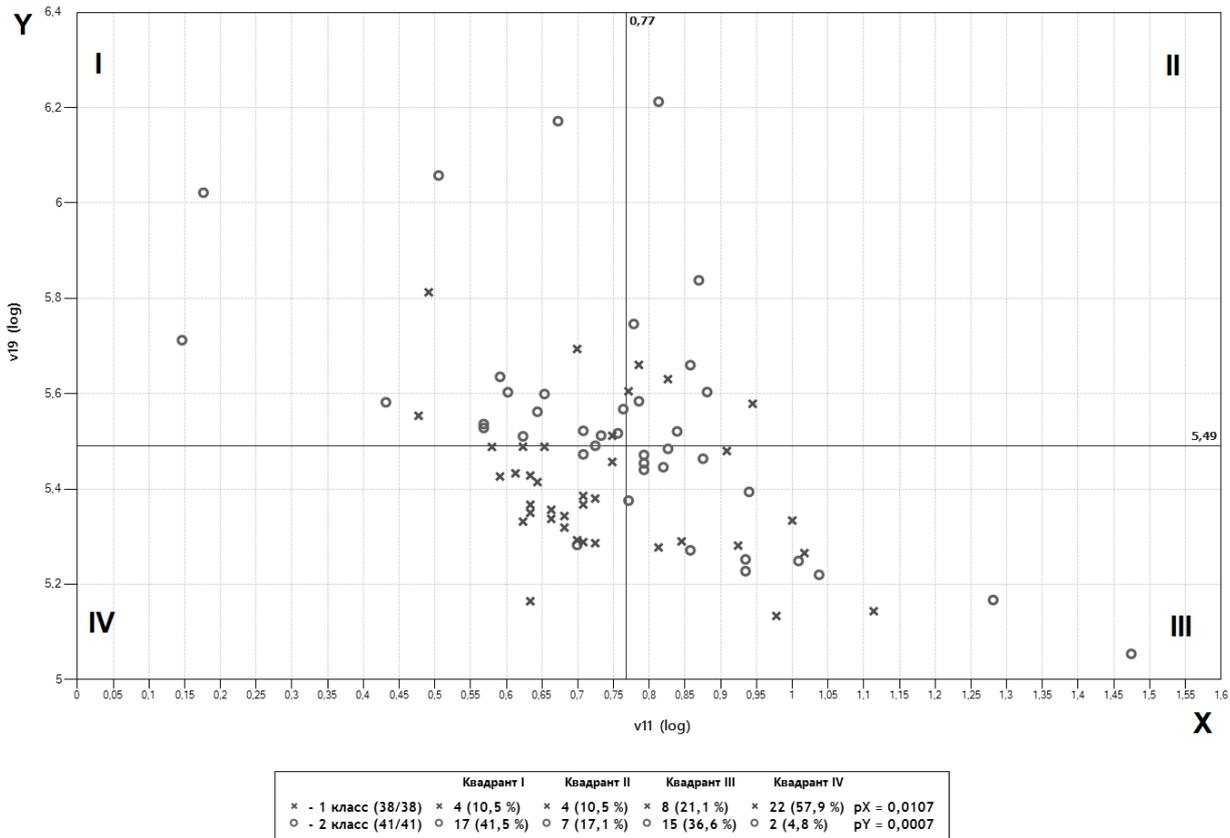


Рис. 3. Вдоль оси X отложены значения переменной v11- уровень безработицы (в процентах), вдоль оси Y отложены значения переменной v19- валовой региональный продукт на душу населения (рублей). Для удобства переменные представлены в логарифмической шкале

Первый квадрант, низкая отдача:

Магаданская область, Красноярский край, Самарская область, Липецкая область, Вологодская область.

Первый квадрант, высокая отдача:

Тюменская область, Чукотский автономный округ, г. Москва, г. Санкт-Петербург, Республика Татарстан, Белгородская область, Ленинградская область, Московская область, Пермский край, Оренбургская область, Новгородская область, Новосибирская область, Краснодарский край, Калининградская область, Калужская область, Республика Башкортостан.

Второй квадрант, низкая отдача:

Камчатский край, Мурманская область, Хабаровский край, Иркутская область.

Второй квадрант, высокая отдача:

Сахалинская область, Республика Саха (Якутия), Республика Коми, Архангельская область, Томская область, Свердловская область, Приморский край.

Третий квадрант, низкая отдача:

Республика Карелия, Забайкальский край, Курганская область, Республика Бурятия, Псковская область, Республика Алтай, Карачаево-Черкесская Республика, Кабардино-Балкарская Республика.

Третий квадрант, высокая отдача:

Омская область, Республика Хакасия, Астраханская область, Челябинская область, Волгоградская область, Кемеровская область, Еврейская автономная область, Ростовская область, Алтайский край, Республика Северная Осетия — Алания, Республика Дагестан, Республика Адыгея, Республика Калмыкия, Республика Тыва, Республика Ингушетия.

Четвертый квадрант, низкая отдача:

Нижегородская область, Воронежская область, Ярославская область, Амурская область, Тульская область, Тамбовская область, Курская область, Рязанская область, Смоленская область, Тверская область, Орловская область,

Владимирская область, Саратовская область, Костромская область, Ульяновская область, Пензенская область, Республика Мордовия, Республика Марий Эл, Брянская область, Кировская область, Ставропольский край, Ивановская область.

Четвертый квадрант, высокая отдача:

Удмуртская Республика, Чувашская Республика.

Заключение

Из проведенного анализа можно сделать вывод, что существенная часть показателей экономического воспроизводства в регионах демонстрирует значимую связь с отдачей от масштаба производства. Результаты показывают, что наиболее значительное влияние на отдачу имеет, во-первых, размер экономики (хотя можно выделить группу регионов с небольшими значениями валовых показателей, но относительно высокими значениями отдачи производства). Важную, но не меньшую роль играют показатели инфраструктуры, включая плотность автомобильных дорог. Фактором, способствующим увеличению отдачи, является высокий средний уровень доходов населения и, одновременно, высокая безработица. Таким образом, одним из интересных результатов является выявленная связь сочетаний признаков с факторами отдачи. Это свидетельствует о том, что необходимо анализировать особенности экономического воспроизводства в каждом регионе в целом, включая и специфику действующей институциональной среды. Проведенные ранее исследования других авторов подтверждают важность такого подхода [16–17]. В этом авторы видят перспективы дальнейшей работы, связанные с использованием предложенной методики. Кроме того, анализируя отдачу регионов, можно заметить, что она в целом выше при наличии в регионах промышленных центров, кластеров, агломераций, что позволяет использовать эту методику для обоснования и проверки принципов моделей популярной в последние десятилетия новой экономической географии, где основополагающую роль в формировании агломераций играет возрастающая отдача в некоторых отраслях, а, следовательно, может иметь более высокие

значения, чем в регионах с менее выраженной агломерацией агрегированная по отраслям отдача [18–19].

Литература

1. Cobb C.W., Douglas P.H. A Theory of Production // American Economic Review. 1928. V. 18. No 1, Supplement, Papers and Proceedings of the Fortieth Annual Meeting of the American Economic Association. P. 139–165.
2. Чубрик А.С. Отдача от масштаба производственной функции и общефакторная производительность: пример Польши и Белоруссии // ЭКОВЕСТ. 2002. Т. 2. №2. С. 252–275.
3. Гафарова Е.А., Моделирование регионального развития на основе производственных функций // Интернет-журнал "Наукоедение". 2013. № 3. С. 1–7. <https://naukovedenie.ru/PDF/39evn313.pdf>.
4. Кирилук И.Л. Модели производственных функций для российской экономики // Компьютерные исследования и моделирование. 2013. Т. 5. № 2. С. 293–312.
5. Kirilyuk I., Senko O. Verification of the Returns to Scale of Production Type for the Russian Federation Regions // EPJ Web of Conferences. 2019. 224. 06011. p. 1–6.
6. Регионы России. Социально-экономические показатели. 2014: Стат. сб. / М.: Росстат, 2014.
7. Регионы России. Социально-экономические показатели. 2017: Стат. сб. / М.: Росстат, 2017.
8. Российский статистический ежегодник. 2015: Стат.сб./ М.: Росстат, 2015.
9. Кирилук И.Л., Сенько О.В. Выбор моделей оптимальной сложности методами Монте-Карло (на примере моделей производственных функций регионов Российской Федерации) // Информатика и ее применения. 2020. Т. 14. Выпуск 2. С. 111–118.
10. Гребнев М.И. Построение производственных функций регионов России // ВУЗ. XXI век. 2015. № 2. С. 50–56.
11. Сенько О.В. Перестановочный тест в методе оптимальных разбиений // Журнал вычислительной математики и математической физики. 2003. Т. 43. № 9. С. 1438–1447.
12. Kirilyuk I.L., Kuznetsova A.V., Sen'ko O.V., Morozov A.M. Method for detecting significant patterns in panel data analysis // Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications). 2017. V. 27. № 1. P. 94–104.
13. Борисова Л.Р., Кузнецова А.В., Сергеева Н.В., Сенько О.В. Применение методов машинного обучения для сравнения компаний Арктической зоны РФ по экономическим критериям в соответствии с рейтингом Полярного индекса // Компьютерные исследования и моделирование. 2020. Т. 12. № 1. С. 201–215.
14. Pesarin F., Salmaso L. Permutation Tests for Complex Data. Chichester, UK: Wiley. 2010.
15. Senko O.V., Kuznetsova A.V. A recognition method based on collective decision making using systems of regularities of various types // Pattern Recognition and Image Analysis. 2010. V. 20. № 2. P. 152–162.

16. Кирдина С.Г., Малков С.Ю. Два механизма самоорганизации экономики: модельная и эмпирическая верификация (научный доклад). М.: Институт экономики РАН, 2010. — 69 с.
17. Маевский В.И., Малков С.Ю. Перспективы макроэкономической теории воспроизводства // Вопросы экономики. 2014. №. 4. С. 137-155.
18. Dixit A.K., Stiglitz J.E. Monopolistic Competition and Optimum Product Diversity // *The American Economic Review*. 1977. V. 67. № 3. P. 297–308.
19. Krugman P. Increasing Returns and Economic Geography // *Journal of Political Economy*. 1991. V. 99, № 3. P. 483–499.

Кирилук Игорь Леонидович. Федеральное государственное бюджетное учреждение науки "Институт экономики Российской академии наук" (ИЭ РАН), г. Москва, Россия. Научный сотрудник. Количество печатных работ: более 60. Область научных интересов: Эконометрика, эволюционная экономика, data mining, синергетика, эконофизика, геронтология, программирование. E-mail: igokir@rambler.ru

Кузнецова Анна Викторовна. Федеральное государственное бюджетное учреждение науки "Институт биохимической физики им. Н.М. Эммануэля" Российской академии наук (ИБХФ РАН), г. Москва, Россия. Старший научный сотрудник, к.б.н. Количество печатных работ: 111. Область научных интересов: Data Science, анализ данных, методы машинного обучения, доказательная медицина, геронтология. E-mail: azforus@yandex.ru

Сенько Олег Валентинович. Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук" (ФИЦ ИУ РАН), г. Москва, Россия. Ведущий научный сотрудник, д.ф.-м.н. Количество печатных работ: 136. Область научных интересов: машинное обучение и анализ данных. E-mail: senkoov@mail.ru

Investigation of the Relationship Between Production Functions and Socio-Economic Indicators of Russian Regions by the Method of Optimal Partitioning

I. L. Kirilyuk^I, A. V. Kuznetsova^{II}, O. V. Senko^{III}

^IInstitute of Economics of Russian Academy of Sciences, Moscow, Russia

^{II}N.M. Emanuel Institute of Biochemical Physics of Russian Academy of Sciences, Moscow, Russia

^{III}Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

Abstract. The dependence of the emergent characteristic- returns to scale, calculated on the basis of production functions in the regions of the Russian Federation, on various indicators characterizing the regions, including macroeconomic indicators, characteristics of economic reproduction processes and socio-economic institutions, geographical indicators, etc. was investigated. The calculations were carried out using the method of optimal valid partitioning, which implies verification of patterns in the data using permutation tests. As a result of the analysis, a statistically significant relationship with returns to scale was revealed for a significant percentage of indicators or their combinations, which indicates both the real existence of the interdependence of these indicators with the returns to scale, and the sufficient accuracy of estimates of the parameters of the Cobb-Douglas model for the time series involved in the analysis.

Keywords: data mining, optimal valid partitioning, pattern verification, permutation tests, production functions, returns to scale, regional economics, economic reproduction.

DOI 10.14357/20718632210103

References

1. Cobb, C.W., and P.H. Douglas. 1928. A Theory of Production. *American Economic Review*, Supplement, Papers and Proceedings of the Fortieth Annual Meeting of the American Economic Association 18(1):139–165.
2. Chubrik, A.S. 2002. Otdacha ot masshtaba proizvodstvennoi funktsii i obshchefaktornaya proizvoditel'nost': primer Pol'shi i Belorussii [Returns to scale of production function and general factor productivity: the example of Poland and Belarus]. *EKOVEST* 2(2):252–275.
3. Gafarova, E.A. 2013. Modelirovanie regional'nogo razvitiya na osnove proizvodstvennykh funktsii [Modeling of regional development based on production functions]. *Internet-zhurnal "Naukovedenie"* [Internet-journal "Science and Innovation"]

- ence of science”] (3):1–7. <https://naukovedenie.ru/PDF/39evn313.pdf>.
4. Kirilyuk, I.L. 2013. Modeli proizvodstvennykh funktsii dlya rossiiskoi ekonomiki [Models of production functions for the Russian economy]. *Komp'yuternye issledovaniya i modelirovaniye*. [Computer Research Modeling] 5(2): 293–312.
 5. Kirilyuk, I., and O. Senko. 2019. Verification of the Returns to Scale of Production Type for the Russian Federation Regions. *EPJ Web of Conferences*. 224(06011):1-6.
 6. Regiony Rossii. Sotsial'no-ekonomicheskie pokazateli [Regions of Russia. Socio-economic indicators] 2014: Stat. sb. / M.: Rosstat, 2014.
 7. Regiony Rossii. Sotsial'no-ekonomicheskie pokazateli [Regions of Russia. Socio-economic indicators] 2017: Stat. sb. / M.: Rosstat, 2017.
 8. Rossiiskii statisticheskii ezhegodnik [Russian Statistical Yearbook] 2015: Stat.sb./ M.: Rosstat, 2015.
 9. Kirilyuk, I.L., and O.V. Sen'ko. 2020. Vybór modelei optimal'noi slozhnosti metodami Monte-Karlo (na primere modelei proizvodstvennykh funktsii regionov Rossiiskoi Federatsii) [Assessing the validity of clustering of panel data by Monte Carlo methods (using as example the data of the Russian regional economy)]. *Informatika i ee primeneniya*. [Informatics and Applications] 14(2):111-118.
 10. Grebnev, M.I. 2015. Postroenie proizvodstvennykh funktsii regionov Rossii [Construction of production functions of Russian regions]. *VUZ. XXI vek* [High School. XXI Century] 2:50–56.
 11. Sen'ko, O.V. 2003. Perestanovochnyi test v metode optimal'nykh razbienií [Permutation test in the optimal partitioning method]. *Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki* [Computational Mathematics and Mathematical Physics] 43(9):1438-1447.
 12. Kirilyuk, I.L., Kuznetsova, A.V., Sen'ko, O.V., and A.M. Morozov. 2017. Method for detecting significant patterns in panel data analysis. *Pattern Recognition and Image Analysis (Advances in Mathematical Theory and Applications)* 27(1):94-104.
 13. Borisova, L.R., Kuznetsova, A.V., Sergeeva, N.V., and O.V. Sen'ko. 2020. Primenenie metodov mashinnogo obucheniya dlya sravneniya kompanii Arkticheskoi zony RF po ekonomicheskim kriteriyam v sootvetstvii s reitingom Polyarnogo indeksa [Application of machine learning methods for comparing companies in the Arctic zone of the Russian Federation by economic criteria in accordance with the Polar Index rating]. *Komp'yuternye issledovaniya i modelirovaniye* [Computer Research Modeling] 12(1):201-215.
 14. Pesarin F., and L. Salmaso. 2010. *Permutation Tests for Complex Data*. Chichester, UK: Wiley.
 15. Senko, O.V., and A.V. Kuznetsova. 2010. A recognition method based on collective decision making using systems of regularities of various types. *Pattern Recognition and Image Analysis* 20(2):152–162.
 16. Kirdina, S.G., and S.Yu. Malkov. 2010. Dva mekhanizma samoorganizatsii ekonomiki: model'naya i empiricheskaya verifikatsiya (nauchnyi doklad) [Two mechanisms of self-organization of the economy: model and empirical verification (scientific report)]. M.: Institut ekonomiki RAN, 69 s.
 17. Maevskii, V.I., and S.Yu. Malkov. 2014. Perspektivy makroekonomicheskoi teorii vosproizvodstva [Perspectives of the macroeconomic Reproduction Theory]. *Voprosy ekonomiki* [Economic issues] 4:137-155.
 18. Dixit, A.K., and J.E. Stiglitz. 1977. Monopolistic Competition and Optimum Product Diversity. *The American Economic Review* 67(3):297–308.
 19. Krugman, P. 1991. Increasing Returns and Economic Geography. *Journal of Political Economy* 99(3):483-499.

Kirilyuk I. L. Institute of Economics, Russian Academy of Sciences, 32 Nakhimovskii pr., Moscow, 117218, Russia, e-mail: igokir@rambler.ru.

Kuznetsova A. V. PhD, Institute for Biochemical Physics, Russian Academy of Sciences, 4, Kosygina st., Moscow, 119334, Russia, e-mail: azforus@yandex.ru.

Senko O. V. Doctor of Physical and Mathematical Sciences, Federal Research Center Computer Science and Control, Russian Academy of Sciences, 44/2 Vavilova st., Moscow, 119333, Russia, e-mail: senkoov@mail.ru