
РЕГИОНАЛЬНЫЕ ПРОБЛЕМЫ

Опыт статистического и энтропийного анализа взаимосвязей информационных и производственных отраслей региональной экономики России

© 2023 г. Ю.Н. Гаврилец, И.В. Тараканова

Ю.Н. Гаврилец,
ЦЭМИ РАН, Москва; e-mail: yurkag@mail.ru

И.В. Тараканова,
ЦЭМИ РАН, Москва; e-mail: itar40@mail.ru

Поступила в редакцию 05.07.2023

Аннотация. В работе предлагаются специальные методы статистического подхода к анализу региональных межотраслевых взаимосвязей. В качестве реальной системы, по статистическим данным которой проводятся все количественные расчеты, рассматривается Российская Федерация без трех самых больших административных единиц (Москвы, Санкт-Петербурга, Московской области). В отличие от материальных связей модели Леонтьева анализируются информационно-статистические связи между отраслями. Для этого используются понятия энтропии и информации с целью выявления особенностей взаимодействия между шестью производственными отраслями и шестью информационными секторами экономики. Эти взаимодействия рассматриваются с позиций закона необходимого разнообразия, а сложность по Колмогорову, которой в работе придается особое значение и которая оценивается энтропией распределения показателей выпуска региональных подсистем. Вводятся понятия и произведена количественная оценка сложности регионов и отраслей. Структура информационных связей между показателями экономической системы находится как с помощью регрессионного анализа, так и на основе предлагаемых коэффициентов информативности. Особенно подчеркивается роль условных энтропий при статистическом анализе межотраслевых взаимодействий. Структуры взаимодействия представлены в виде соответствующих графов связи. Обосновывается, что рассчитанные коэффициенты информативности и условные энтропии дают возможность сделать качественный анализ и прогноз. Произведенные расчеты по статистическим данным показали, что в регионах с преобладанием энтропии управления над энтропией производства выпуск на одного занятого существенно выше. На основании общего рассмотрения региональных энтропий представляется целесообразным рекомендовать регионам России повысить сложность отраслей связи и научно-профессиональной деятельности хотя бы до уровня сложности отрасли обрабатывающего производства.

Ключевые слова: производственные и информационные отрасли, энтропия и информация, сложность по Колмогорову, граф взаимосвязей, закон необходимого разнообразия.

Классификация JEL: C01, C02.

Для цитирования: Гаврилец Ю.Н., Тараканова И.В. (2023). Опыт статистического и энтропийного анализа взаимосвязей информационных и производственных отраслей региональной экономики России // *Экономика и математические методы*. № 4. С. 71–85. DOI: 10.31857/S042473880028218-8

1. ВВЕДЕНИЕ

В статье делается попытка провести анализ экономической системы с позиций теории управления сложными системами, т.е. кибернетики; используется язык и методология теории информации. В последнее время этот подход при анализе экономики получает более широкое распространение. Так, в работе (Рыжкова, 2010) используется энтропийный подход к анализу инвестиционной привлекательности компании, а в (Рыжкова, 2012) исследуется финансовая деятельность предприятия. В работе (Боженко, Махина, 2018) описана процедура проверки гипотезы об альтернативных поставках на основе информационного критерия. В работах (Ebeling et al., 2002; Mollaei, Darooneh, Karimi, 2019) развивается техника использования энтропийного анализа для измерения уровня постоянства во временных рядах произвольной природы. В работе (Jakimowicz, 2020) показано, что использование энтропии является ценным дополнением к традиционной эконометрике, поскольку она объясняет явления, основанные на степенном распределении вероятностей и позволяет эконометрически оценивать нужные параметры.

К сожалению, нам не удалось найти работ экономико-математического направления, в которых исследовалась бы именно взаимозависимость между характеристиками самого управления и экономики. В нашей работе нет данных о конкретной экономической системе с явным разделением на пассивную и активную подсистему, типа — заводоуправление и цеха завода. Подобные особенности и различия определяются весьма опосредованно.

Мы рассмотрим региональную экономику (без Москвы, Санкт-Петербурга и Московской области, все показатели которых резко отличаются от остальных) как целостную систему. Заметим, что наряду с традиционным межотраслевым балансом, отражающим связи между выпуском и затратами отраслей, существуют и информационно-статистические связи между отраслями.

Считая, что совокупность регионов образует систему, элементы которой в определенном смысле *однородны*, можно формально рассматривать выпуски отраслей как отдельные реализации некой многомерной случайной величины — экономики региона России. Фактически поступают подобным образом, когда строят модели производственных функций и т.п. При отсутствии для регионов России данных о межотраслевых взаимодействиях, представляемых обычно схемами межотраслевого баланса (МОБ), все же имеется возможность статистической оценки размеров выпуска одних отраслей по выпуску других на основе исключительно информационных связей. В настоящей работе этот методологический подход используется для выделения и описания кибернетических особенностей взаимосвязей между отраслями экономики, в частности факторов информации и управления.

Являясь целостной системой, экономика включает принципиально различные по характеру их выпуска отрасли, а именно — сферу материального производства и *сферу производства и использования информации*. И это различие должно приводить к определенным особенностям соотношений между показателями конкретных отраслей. Цель данной работы — выявить проявление кибернетического закона «необходимого разнообразия», иногда называемого законом Эшби, требующего, чтобы сложность управляющей подсистемы превосходила некоторый порог, зависящий от сложности управляемой подсистемы (Бир, 1965; Гаврилец, 2016). К сожалению, в настоящий момент мы не имеем данных, характеризующих систему непосредственного управления производством, но у нас есть данные, относящиеся к производству и использованию информации, без которой невозможно никакое управление. Сейчас мы не можем рассмотреть сколь-нибудь глубоко совокупную проблему сложности, производства, информации и управления, а тем более — получить конкретные выводы практического характера. Укажем лишь простейшие процедуры измерения и сопоставления сложности разных отраслей и регионов, чтобы наметить направление, по которому желательно идти дальше.

В качестве условно управляемой, «пассивной» подсистемы нами взяты шесть отраслей материального производства, суммарный выпуск которых составляет более 57% общего выпуска страны. В качестве условно управляющей подсистемы было выбрано шесть отраслей, деятельность которых не связана напрямую с материальным производством, но явно имеет, в широком смысле слова, *управляющий характер*. К ним мы отнесли отрасли, характеризующиеся информационной деятельностью. Конечно, термины «*управляемые*» отрасли экономики и «*управляющие*» на уровне простых статистических показателей используются нами весьма условно, но по характеру деятельности эти отрасли вполне могут так называться в макроэкономическом смысле.

Далее мы сначала проанализируем информационно-статистические взаимозависимости между всеми переменными системы, а затем рассмотрим их кибернетические особенности. Отметим также, что используемый нами термин *информационные* связи не обязательно означает *причинные зависимости*, описываемые некоторой содержательной моделью, как например между затратами и выпуском. Они лишь отражают сам факт, выражают наблюдаемые статистические соответствия между частями в целом, по которым, зная одно, можно узнать и другое, хотя, разумеется, некоторая весьма опосредованная по времени причинная связь может быть.

2. СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ И ОСОБЕННОСТИ ЭКОНОМИЧЕСКОЙ ПОДСИСТЕМЫ

Введем обозначение для годовых выпусков: U_1 — деятельность в области информации и связи; U_2 — деятельность финансовая и страховая; U_3 — деятельность по операциям с недвижимым имуществом; U_4 — деятельность профессиональная, научная и техническая; U_5 — деятельность административная и сопутствующие дополнительные услуги; U_6 — деятельность государственного управления и обеспечения военной безопасности, социальное обеспечение; E_1 — сельское, лесное

Таблица 1. Описательные статистики показателей, млн руб.

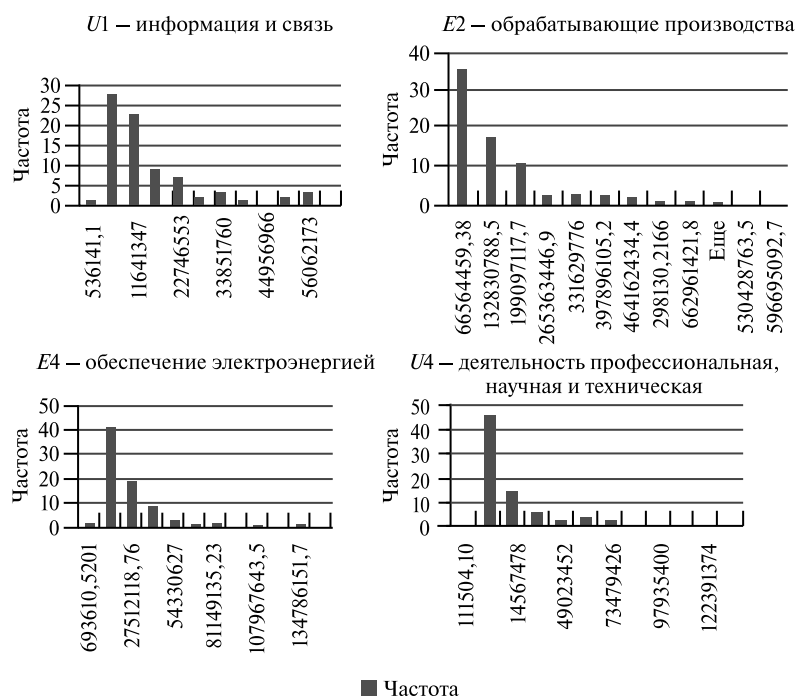
Признак	$U1$	$U2$	$U3$	$U4$	$U5$	$U6$
Среднее	13873647,91	1907257,08	64868272,25	18477986,12	15073224,23	39944678,1
Стандартное отклонение	15156103,66	2547681,532	63366864,69	23845890,75	20537085,39	29595444,95
Минимум	327207,9881	744,2	2240128,095	111504,1024	179662,5247	8643570,923
Максимум	67385771,42	17233396,51	359801577,9	122391374	129575542,8	154401406,2

Признак	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$
Среднее	49230438,5	102134668	124737397	21226556,7	43473001.Е18	52115080,5
Стандартное отклонение	47676501,1	415356304	138134593	22494342	87527502,4	66623673,8
Минимум	2326990,39	24160,438	224752,55	693610,52	2671982,8	637429,357
Максимум	282344625	3633887662	659659896707	131334786152	7533870265	33707071408

хозяйство, охота, рыболовство и рыбоводство; $E2$ — добыча полезных ископаемых; $E3$ — обрабатывающие производства; $E4$ — обеспечение электрической энергией, газом и паром, кондиционирование воздуха; $E5$ — строительство; $E6$ — транспортировка и хранение.

Для расчетов значения всех перечисленных выше показателей были взяты по статистическим данным о выпусках этих отраслей в 2020 г. Описательные статистики (табл. 1) и гистограммы (рис. 1) показывают, что наблюдаемые распределения значений этих показателей не могут считаться гауссовскими.

Разброс значений свидетельствует также о существенной неоднородности регионов. Тем не менее, можно убедиться, что выделенные две группы по шести переменным образуют явно сформированные подсистемы региональной экономики. Были найдены первые главные компоненты для этих групп переменных и установлено, что доля объясняющей дисперсии для производственных отраслей равна 62%, а для информационных секторов — 87%. Коэффициент корреляции между

**Рис. 1.** Гистограммы используемых показателей

этими компонентами оказался равным 0,92. Наличие сильной линейной связи между ними означает возможность прогноза между соответствующими переменными.

Для более детального анализа желательно рассмотреть характер взаимного влияния между всеми переменными. Анализ системы, описываемой большим числом статистических показателей, удобно начинать с выявления структуры *непосредственных связей* между переменными (Гаврилец, 1974). По причине негауссового распределения показателей будет весьма трудно получить адекватное представление о непосредственных связях между переменными с помощью коэффициентов частной корреляции. Тем не менее, выводы о характере статистических взаимосвязей отраслей региональной экономики могут быть получены, если найти зависимости типа непосредственных связей другими способами. Конкретно была поставлена задача выявить *ближайшие* непосредственные связи. А именно для каждой переменной специальным перебором (Тараканова, 2019) были найдены такие небольшие группы других переменных, которые в наибольшей степени оказываются связанными линейно с каждой из них. Были получены коэффициенты регрессии, отражающие наиболее сильные линейные связи. Регрессионные уравнения по этим трем переменным позволяют определять значения зависимой переменной (табл. 2).

Таблица 2. Регрессионные зависимости переменных по управлению U и производству E

$U1$			$E1$		
$R^2 = 0,892$	b^*	$t(75)$	$R^2 = 0,656$	b^*	$t(75)$
$E4$	-0,24	-3,81	$U3$	1,50	9,95
$U3$	0,61	7,24	$U4$	-0,59	-2,48
$U4$	0,55	5,63	$U5$	-0,33	-1,77
$U2$			$E2$		
$R^2 = 0,898$	b^*	$t(75)$	$R^2 = 0,952$	b^*	$t(75)$
$E2$	0,56	14,79	$E5$	0,97	24,48
$E3$	0,23	3,95	$U3$	-0,49	-10,83
$U1$	0,43	7,47	$U5$	0,32	5,64
$U3$			$E3$		
$R^2 = 0,924$	b^*	$t(75)$	$R^2 = 0,744$	b^*	$t(75)$
$E1$	0,30	7,29	$E6$	-0,60	-5,21
$U1$	0,28	3,39	$U3$	0,61	4,22
$U4$	0,51	6,85	$U4$	0,75	5,56
$U4$			$E4$		
$R^2 = 0,94$	b^*	$t(75)$	$R^2 = 0,771$	b^*	$t(75)$
$U1$	0,27	4,71	$E2$	0,37	5,90
$U5$	0,57	11,71	$E3$	0,38	5,29
$E3$	0,21	4,85	$U6$	0,36	4,48
$U5$			$E5$		
$R^2 = 0,936$	b^*	$t(75)$	$R^2 = 0,954$	b^*	$t(75)$
$E2$	0,19	5,28	$E2$	0,92	24,48
$E6$	0,30	5,56	$U3$	0,43	8,54
$U4$	0,59	11,40	$U5$	-0,19	-3,09
$U6$			$E6$		
$R^2 = 0,848$	b^*	$t(75)$	$R^2 = 0,882$	b^*	$t(75)$
$E4$	0,19	2,77	$E3$	-0,18	-3,33
$E6$	0,50	5,48	$U5$	0,56	7,40
$U3$	0,30	3,34	$U6$	0,54	7,52

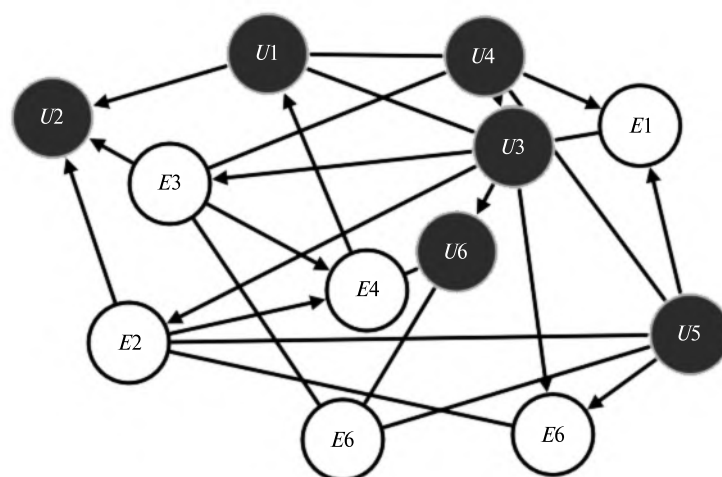


Рис. 2. Граф регрессионных связей между отраслями производства и «условно управляющими сферами деятельности». Отсутствие стрелок на ребрах означает взаимное влияние

В табл. 2 приведены коэффициенты регрессий, отражающих аналог непосредственных связей и их статистические характеристики. В блоке для каждого показателя указаны коэффициенты его зависимости от трех других переменных, а также статистические характеристики данной регрессии R^2 и t -статистики.

По этим связям для большей наглядности был построен ориентированный граф (рис. 2), отражающий некий аналог структуры непосредственной связи. По крайней мере для каждой переменной он указывает ближайшие переменные по информационной связи.

Как видно из данных в табл. 1 и графа связей на рис. 2, изменения значений показателей в регионе двух отраслей — $U1$ и $U2$ — не влияют ни на одну производственную отрасль. Связь переменной $U1$ с другими производствами опосредуется отраслями с показателями $U3$, $U4$. О росте выпуска сельского и лесного хозяйства можно судить по уровню обрабатывающего производства и уменьшению научно-профессиональной и административной деятельности в регионе. При этом зависимость не очень серьезная. В то же время росту выпуска добывающей отрасли способствует строительство и операции с недвижимостью. О росте выпуска добывающего производства свидетельствует рост научно-профессиональной деятельности и деятельности по операциям с недвижимостью. Рост регионального строительства также сопровождается ростом рынка недвижимости и добычи полезных ископаемых.

В то же время показатель $U3$ информационно влияет на первые три отрасли и на пятую, оставляя свободными только энергетику и транспорт. Интересно, что отрасль $U4$ информативен только для обрабатывающего производства $E3$ и сельского хозяйства $E1$. $U6$ не влияет только на $E4$ и $E3$, а $U5$ связана только с $E4$ и $E6$. Показатель $E1$ влияет только на $U3$. $E4$ влияет только на $U1$ и $U6$. Отрасль $E5$ не проявила влияния ни на какой информационный сектор. В целом на *производство* направлено 12 информационных влияний из 18 возможных, в то же время для *информационных отраслей* число тех и других влияющих отраслей оказалось одинаковым, т.е. управляющих влияний — больше, чем производственных. Как видим, роль информационных отраслей оказывается весьма существенной.

3. СЛОЖНОСТЬ ЭКОНОМИКИ РЕГИОНОВ РОССИИ И СЛОЖНОСТЬ ОТРАСЛЕЙ РФ

В данной работе для анализа статистических показателей экономики регионов России вводится и используется термин *сложность*. Проблеме сложности как таковой уделялось и уделяется много внимания, особенно в работах исследователей кибернетических связей, биологов и философов. Мы считаем, что главное в проблеме сложности для науки об обществе — ее значение и роль в решении задач прогнозирования и управления. Выражаясь предельно категорично, можно утверждать, что *все ошибки и неудачи принятых где-либо экономических и социально-экономических решений вызваны односторонним несоответствием между сложностью решаемой проблемы и компетенцией лиц, принимавших решения (их компетенция и есть сложность управляющей системы)*.

В середине прошлого века, когда бурно развивались методы оптимального управления сложными системами, преодолеть «проклятие размерности» призывал известный математик Р. Беллман. В отличие от концепции так называемой «экономической сложности» (Афанасьев, Кудров, 2021; Моисеев, Бондаренко, 2020) мы, исследуя экономические системы, понимаем *сложность по Колмогорову* (Колмогоров, 1987), которую иногда называют *K-сложностью* (Буданов, Аршинов, 2018) и содержание которой как раз соответствует пониманию классиков науки Н. Винера, У.Р. Эшби, С. Бира и др.

Переход от обсуждений к конструктивному решению требует умения *измерять сложность* реальных систем, и эта проблема пока что далека от решения. Известный советский экономист В.В. Новожилов (Новожилов, 1959) полагал, что сложность системы можно измерять «числом степеней свободы». Рост сложности системы означает *развитие системы*, мы считаем, что слово *прогресс* следует употреблять для обозначения роста некоего этически окрашенного параметра (Щепанский, 1969; Гаврилец, 2016). Кибернетики 1950–1960-х годов (например, Бир, 1965)) считали, что такая важная характеристика сложных систем, как *разнообразие*, и есть характеристика сложности, и она может быть описана числом различных состояний или свойств. Формально для заданного множества различных элементов под этим термином надо понимать число классов эквивалентности по какому-то признаку. В.Л. Макаров в статье (Макаров, 2015) показывает важность для экономических систем характеристики *разнообразия*.

Согласно определению Колмогорова сложность — минимальное число информационных сигналов, декодирующих сообщение для получающего их субъекта. Естественной *мерой разнообразия*, особенно когда разные характеристики имеют разные вероятности проявления, служит *энтропия* по К. Шеннону. Энтропия как раз представляет собой минимальное среднее число дихотомических вопросов для получения знания об исходе случайного опыта с равномерным распределением вероятностей. По этой причине мерой сложности систем, характеризующихся законами дискретного распределения вероятностей, естественно считать энтропию. Известно, что максимум выпуклой функции величины энтропии — $\sum X_k \log(X_k)$ (при ограничениях $X_k \geq 0$, $\sum X_k = S$) — достигается равномерным распределением $X_k = S/N$, где N — число переменных X_k . Минимум энтропии достигается при отсутствии ненулевых значений переменных, кроме одной, равной S .

Рассуждая о сложности некоторого объекта, необходимо иметь набор характеристик этого объекта. В данной работе рассматривается два объекта, сложности которых мы хотим исследовать: отдельная отрасль всей экономики (т.е. ее предприятия, разнесенные по всем регионам России) и экономика отрасли внутри отдельного региона России. Характеристики отрасли представляют собой статистические данные о ее выпусках в регионах, характеристики региона — объемы выпусков каждой отрасли данного региона.

В соответствии с этим показателем *K-сложности* экономики отдельного региона России рассматривается величина энтропии, согласно наблюдаемому распределению регионального выпуска по отраслям народного хозяйства России. Тем самым энтропия подсистем экономики *конкретного региона* определяется структурой (разнообразием) отраслевых выпусков подсистем. С точки зрения здравого смысла, чем больше различных отраслей в регионе, тем сложнее экономика этого региона, тем сложнее ее прогнозировать.

Одновременно с этим *энтропию конкретной отрасли всей экономики* определяем по разнообразию объемов выпуска этой отрасли в регионах. Очевидно также, чем больше выпуск отрасли распределен по разным регионам страны, тем сложнее представить отрасль как целое и тем труднее прогнозировать (или управлять) таким выпуском. Заметим, что мы называем энтропией отрасли энтропию распределения показателя выпуска этой отрасли в стране. Мы получаем минимальную *неопределенность отрасли*, когда у всех регионов оказывается одинаковый выпуск, а максимальную неопределенность — когда каждому диапазону распределенного выпуска (значений на дискретной шкале выпуска) будет соответствовать одинаковое число регионов.

В этом случае *информационная сложность* отрасли экономики отражает степень непохожести регионов по величине отраслевого выпуска. А энтропия региона состоит в разнообразии объемов всех отраслевых выпусков. Можно считать *сложность региона* минимальной, когда весь выпуск определяется одной отраслью, максимальной — когда выпуски всех отраслей составляют одинаковую долю совокупного выпуска региона. Заметим, что при таком подходе сложность определяется только *относительными* величинами выпусков, что при необходимости может корректироваться дополнительно, например множителем.

Ясно, что для внешнего наблюдателя прогнозировать или управлять отраслью экономики с большим числом регионов с одинаковыми выпусками проще, чем при их большем разнообразии. Аналогично: прогнозировать экономику самого региона проще, когда в нем мало отраслей.

4. СЛОЖНОСТИ РЕГИОНАЛЬНЫХ ПОДСИСТЕМ ПРОИЗВОДСТВА И УПРАВЛЕНИЯ

Наша цель состоит не просто в определении сложности отдельных частей экономики, а в рассмотрении взаимоотношения между сложностями подсистем производства и управления. Поэтому прежде чем анализировать *сложности* всех отраслей региональной экономики, были измерены сложности подсистемы управления и подсистемы производства в каждом регионе. Конкретно были рассчитаны величины энтропий для шести отраслей производственной сферы и для шести условно управляющих информационных отраслей (рис. 3). Энтропии каждой подсистемы вычислялись по долям общего выпуска.

На рис. 3 можно различить регионы, у которых энтропия сферы «информация — управление» больше энтропии сферы производства, и те, у которых она меньше. Для всех регионов была рассчитана величина валового продукта, приходящаяся на одного занятого (что в некоторой степени характеризует эффективность выпуска в регионе). Как оказалось, общая эффективность экономики была выше там, где энтропия управления выше энтропии производства (табл. 3, рис. 4).

В первой строке приведены данные для 48 регионов, у которых энтропия информационных отраслей меньше энтропии производственных. Во второй строке указаны данные для 30 регионов, у которых, напротив, энтропия производственных отраслей меньше.

Этот факт, по-видимому, можно интерпретировать как проявление требований закона необходимого разнообразия Эшби, поскольку энтропия непосредственно выражает сложность управляемой и управляющей подсистем по Колмогорову. Для экономической практики вывод, по-видимому, такой: росту эффективности общественного труда непосредственно способствует рост К-сложности управляющей подсистемы экономики, которая не должна быть меньше сложности системы производства. В отношении других реальных экономических систем, конечно, остаются

Таблица 3. Суммарный выпуск на одного занятого в регионах

Число регионов	ВРП, млн руб.	Число занятых	ВРП / Число занятых
Для 48 регионов	27 035 243 884	31 438 400	859,9433777
Для 30 регионов	26 687 882 661	23 611 300	1130,301282

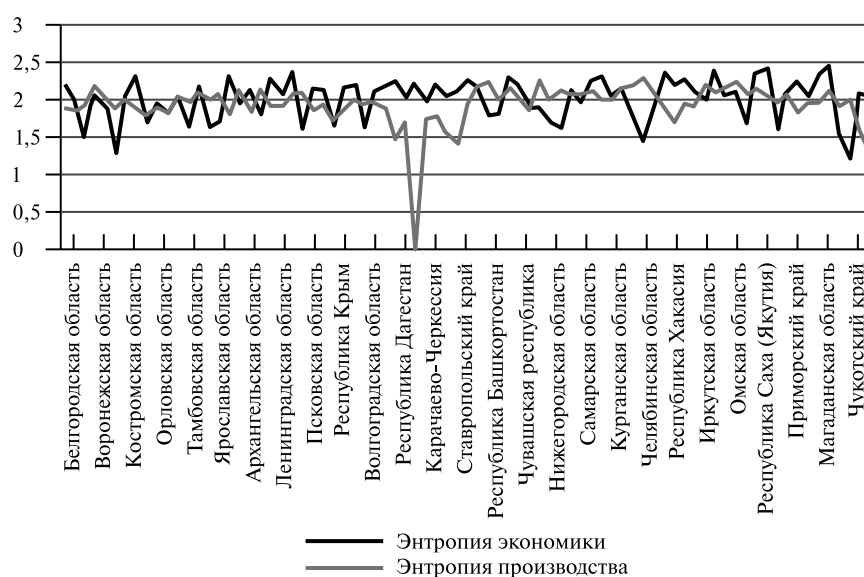


Рис. 3. Графики энтропий производства и управления по основным регионам РФ

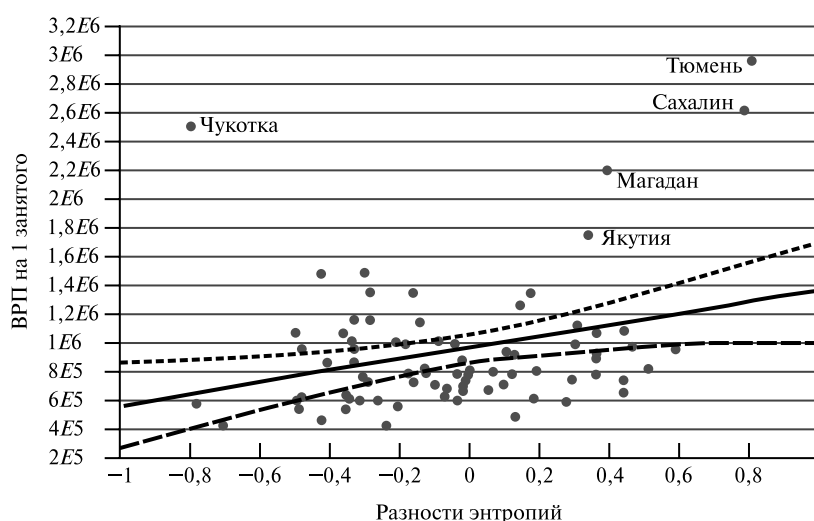


Рис. 4. Зависимость эффективности производства от разности энтропий управления и производства

вопросы, какими показателями их описывать и как измерять и сопоставлять эти сложности в конкретных случаях.

Как видно из графиков на рис. 4, регионы нельзя считать однородной выборкой, хотя в среднем по мере роста разности между энтропиями управления и производства происходит возрастание эффективности выпуска на одного занятого, что подтверждает вывод из данных в табл. 3. Если отбросить данные Чукотского АО, для которого характерна минимальная разность между энтропиями, то положительная корреляция становится больше 0,4.

На рис. 4 также можно заметить, как резко возрастает производительность одного занятого при увеличении энтропии управления для высокоэффективных регионов по ВРП. Начиная с Республики Башкортостан, после Сахалинской области к Тюменской идут Красноярский край, Республики Саха-Якутия и Магаданская область. Все они наиболее эффективны с точки зрения выпуска на одного занятого — среди тех регионов, в которых сложность управления выше сложности производства. Конечно, для этих регионов характерна и малая энтропия производства в силу его специфики. Но, к сожалению, и у многих регионов, имеющих высокую степень управления, эффективность производства остается низкой.

5. ИЗМЕРЕНИЕ СЛОЖНОСТИ ОТРАСЛЕЙ ЭКОНОМИКИ

Задачей этого раздела статьи является сравнение сложностей отраслей управляемой подсистемы экономики и управляющей. Для статистической оценки «сложности — энтропии» отрасли необходимо определенное число наблюдений. Сама отрасль представляется здесь величиной своего выпуска, поэтому данные региональной официальной статистики можно считать выборочным наблюдением отраслевых выпусков. В соответствии с определением сложности отрасли, приведенном в разд. 3, необходимо единообразно вычислять величины энтропий отраслей. Так как все показатели отраслевых выпусков измеряются в непрерывных шкалах, дискретные шкалы для них были составлены с одинаковым числом градаций, несмотря на различные диапазоны изменения показателей. Поскольку рассчитываемая величина энтропии отрасли зависит от числа градаций дискретной шкалы значений, то для проверки стабильности характера соотношений между энтропиями отраслей были проведены расчеты их величин при разном числе градаций — от 5 до 15 (рис. 5). Как видно из графиков, для разного числа градаций упорядоченность значений энтропий не меняется.

Рост энтропии при увеличении числа градаций представляется весьма естественным для обычных распределений, а уменьшение ее значений в каких-то случаях означает специфику групп регионов, попадающих в градацию с конкретными значениями показателя. Необходимо заметить, что выбор числа градаций, и даже входящих в каждую градацию значения признака, определяется исследователем и в этом смысле имеет условно-относительный характер. Принятое нами представление в виде равномерной шкалы кажется наиболее простым, а все дальнейшие расчеты

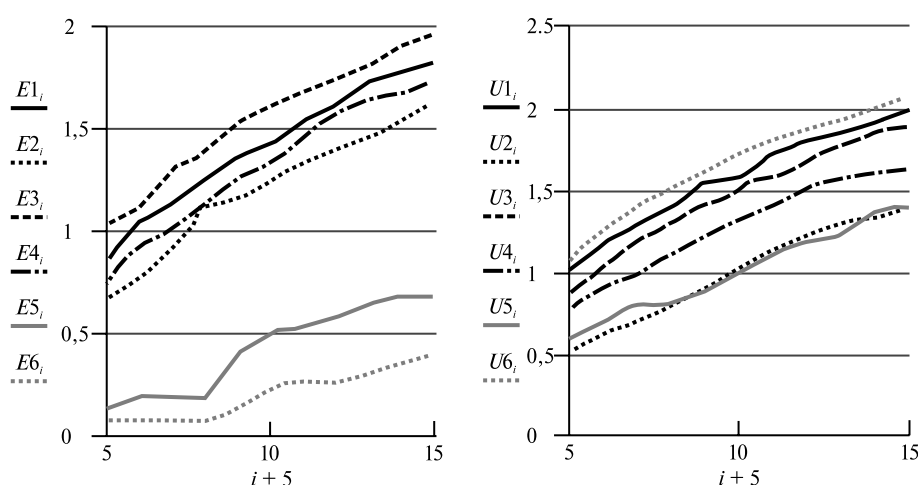


Рис. 5. Зависимость энтропии отраслевых выпусков от числа градаций шкалы измерения выпуска

имеют методологический характер, иллюстрируя возможности анализа, которые, тем не менее, отражают объективную реальность.

Рассмотрим теперь величины энтропии отраслей в 2020 г. (табл. 4). Но прежде обратимся к табл. 1 описательных статистик.

По регионам наибольший разброс (по коэффициентам вариации) имеют показатели управления (в порядке убывания) U_6 , U_3 , U_1 , как и их энтропии. Наибольший разброс — у показателей производства E_3 , E_4 , E_1 . С энтропиями порядок несколько иной — E_1 , E_4 , E_3 . Наибольшая информационная сложность оказывается у показателей государственного управления (U_6) и сектора связи (U_1). Можно отметить, что здесь отражается общее требование закона необходимого разнообразия Эшби на уровне всего народного хозяйства. Далее идет торговля недвижимостью (U_3). Административная (U_5) и финансово-страховая (U_2) виды деятельности, концентрируясь в малом числе регионов, имеют небольшую энтропию. Порядок для показателей E_6 , E_5 , E_2 оказывается одинаковым и для энтропий, и для коэффициентов вариации. Таким образом, энтропии и коэффициенты вариации отраслевых выпусков упорядочены похожим образом, но не одинаково. Различие может указывать на сравнительное отставание одной отрасли от другой по уровню управляемости. Энергетика больше разбросана по отраслям, чем обработка, но энтропия гораздо больше у обрабатывающих отраслей.

Наибольшая сложность для производственных отраслей оказывается у обрабатывающего производства (E_3), и даже превосходит энтропию всех информационных отраслей, кроме государственной деятельности (U_6). То, что энтропии услуг связи (U_1) и профессиональной деятельности (U_4) меньше энтропии показателя E_3 , на наш взгляд, свидетельствует об определенном несоответствии необходимым требованиям закона Эшби — недостаточной степени сложности общей системы управления для такой важной отрасли, как обрабатывающее производство. В регионах России не хватает соответствующего уровня развития систем связи и подготовки профессиональных кадров при наличии сложных отраслей обрабатывающего производства. Заметим, что в работе (Гаврилец, Кудров, Тараканова, 2022) было установлено влияние уровня научных исследований на ВРП и тесная связь между ними и отраслями обрабатывающего производства. Там же приводились

Таблица 4. Энтропии отраслей по 10 градациям

Показатель управления	2018 г.	2020 г.	Показатель производства	2018 г.	2020 г.
U_1	1,671	1,586	E_1	1,525	1,531
U_2	1,399	1,036	E_2	0,268	0,229
U_3	1,641	1,522	E_3	1,581	1,613
U_4	1,024	1,323	E_4	1,27	1,347
U_5	0,814	1,005	E_5	0,48	0,512
U_6	1,726	1,721	E_6	1,208	1,26

статистические данные об определенном отставании финансирования российской науки по сравнению с финансированием в ряде западных стран.

Любопытно и требует какого-то объяснения также тот факт, что информационная сложность *сельского и лесного хозяйств* превосходит сложности всех информационных секторов, кроме *услуг государства и связи*. Возможно, это связано с природными факторами и географическими особенностями нашей страны, но, во всяком случае, подчеркивает важность этой отрасли в экономической системе.

Динамика энтропий за два года показана в табл. 4. Как легко видеть, во всех производственных отраслях, кроме добывающих, произошло повышение энтропии. Разнообразие региональных выпусков увеличилось, особенно в энергетике и обрабатывающем производстве. Возможно, это произошло вследствие сокращения импорта. В целом же, как это ни странно, ситуация означает определенное развитие системы производства, потому что, с нашей точки зрения (Гаврилец, 2016), развитие есть не что иное, как усложнение (в отличие от *прогресса*, означающего рост определенного социально-этического совершенства общества). Что же касается информационных отраслей, то рост разнообразия наблюдался только в административной и профессионально-научной отраслях деятельности. Для отраслей *связь, финансы, операции с недвижимостью и государственная деятельность* характерно сужение разнообразия, т.е. отсутствие развития.

6. ЭНТРОПИЙНЫЙ АНАЛИЗ И ИНФОРМАЦИОННАЯ СТРУКТУРА РЕГИОНАЛЬНОЙ ЭКОНОМИКИ

В упрощенной постановке прогнозирование представляет собой определение значений переменной Y по значениям переменной X в предположениях, что эти переменные связаны. Рассуждая об использовании регрессионных уравнений, полученных по данным одного и того же момента времени для прогнозирования, мы имеем дело с условным прогнозом. В детерминированном случае это означает, что существует и может использоваться некоторая известная функция $Y = f(X)$, определяющая для каждого значения X соответствующее ему значение Y .

В простейшем дискретном статистическом случае предполагается, что можно использовать условную вероятность $P(Y/X)$. Очевидно, что более точный прогноз переменной Y по переменной X означает меньшую неопределенность исхода, т.е. меньшую величину условной энтропии $H(Y/X)$, которая равна разности

$$H(Y/X) = H(X, Y) - H(X), \quad (1)$$

где $H(X, Y)$ — совместная энтропия величин X, Y , а $H(X)$ — энтропия переменной X . Совместная энтропия рассчитывается стандартно по формуле $H(X, Y) = -\sum_X \sum_Y P(X, Y) \log_2 P(X, Y)$.

В свою очередь, *информацией по Шеннону* (Яглом А., Яглом И., 1973) о величине Y с энтропией $H(Y)$, заключенной в переменной X , называют величину изменения неопределенности Y при известных значениях X :

$$I(X, Y) = H(Y) - H(Y/X), \quad (2)$$

где $H(Y/X)$ — средняя условная энтропия величины Y , когда значения X известны. Важным фактом также являются равенства:

$$I(X, Y) = I(Y, X) = H(X) + H(Y) - H(X, Y), \quad (3)$$

Определенный интерес представляет величина информации по Шеннону между показателями информационных и производственных отраслей.

В первом столбце табл. 5 приведены значения энтропий информационных отраслей, а в первой строке — значения энтропий производственных отраслей. Как видно из данных в табл. 5, сектор государственной деятельности с максимальной энтропией имеет наибольшую информационную связь с отраслью транспортировки и хранения, у которой энтропия не столь значительная. Наибольшее значение информационной связи у обрабатывающего производства с отраслью связи и торговлей недвижимостью ($I = 0,987; 0,985$). Наименьшая информационная связь с остальными — у добывающей отрасли, которая сама имеет минимальную энтропию.

Несмотря на симметричность формального определения информации (3), условные энтропии $H(Y/X)$ и $H(X/Y)$, вообще говоря, различаются по величине (1) и вследствие этого каждая переменная устраняет у энтропии пары разные количества неопределенности. Лучше всего устраняет

Таблица 5. Парная информация между информационными и производственными отраслями на 2020 г. (для 15 градаций)

Показатель управления	Производство энтропии	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$
		1,832	0,394	1,968	1,749	0,681	1,631
$U1$	2,005	0,799	0,139	0,987	0,792	0,424	0,909
$U2$	1,413	0,47	0,129	0,833	0,624	0,329	0,683
$U3$	1,899	0,686	0,119	0,985	0,766	0,414	0,823
$U4$	1,642	0,592	0,141	0,906	0,741	0,379	0,783
$U5$	1,396	0,424	0,14	0,66	0,65	0,328	0,767
$U6$	2,056	0,689	0,196	0,785	0,802	0,471	0,959

неопределенность та переменная, энтропия которой больше. При стремлении $H(X)$ к величине $H(X, Y)$ условная энтропия стремится к нулю, поэтому для улучшения прогноза желательно из множества возможных наборов $\{X\}$ выбирать те связанные с Y переменные, которые минимизируют условную энтропию Y , устраняя из пары максимальное количество неопределенности. Управление величиной Y или ее прогнозирование может происходить с помощью выбора значений *перемежных* X , наиболее сильно на нее влияющих. Для этого в разделе 2 были рассчитаны регрессионные зависимости всех показателей от трех наилучших по R^2 (см. табл. 2).

Для анализа информационной связи и взаимных влияний между всеми отраслевыми выпусками дополнительно был использован метод, который не определяется характером зависимости между переменными. Для основы прогноза величины Y предлагается выбирать те показатели, для которых условная энтропия $H(Y/X)$ максимальная. Так как условная энтропия показателя Y при известном X : $H(Y/X) = H(X, Y) - H(X)$, — то дробь $H(Y/X)/H(Y)$ показывает, какую долю неопределенности переменной Y составляет ее условная неопределенность при известных значениях X . Тогда, согласно определению информации, разность (1) составляет часть неопределенности в Y , устраняемую переменной X , т.е. коэффициент

$$\gamma(X, Y) = I(X, Y) / H(Y) \quad (4)$$

выражает силу информационного влияния X на Y . В табл. 4 приводится эта величина, она отражает силу информационной связи между ними, а по главной диагонали показаны энтропии всех переменных. Коэффициент информативности (табл. 6) характеризует парную связь, и он несимметричен; помним также, что знак влияния энтропийные меры не отражают. Тем не менее, мы покажем, каким бы был граф непосредственных связей (см. рис. 5) по частным корреляциям, если бы переменные были распределены согласно гауссовскому закону с корреляциями, равными усредненным коэффициентам информативности $\gamma(U, E)$.

Таблица 6. Коэффициенты информативности γ

Показатель	$U1$	$U2$	$U3$	$U4$	$U5$	$U6$	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$
$U1$	2,005	0,582	0,602	0,585	0,212	0,731	0,436	0,353	0,502	0,453	0,623	0,557
$U2$	0,41	1,413	0,454	0,481	0,429	0,372	0,257	0,327	0,423	0,357	0,483	0,419
$U3$	0,571	0,61	1,899	0,544	0,528	0,473	0,374	0,302	0,501	0,438	0,608	0,505
$U4$	0,479	0,558	0,47	1,642	0,559	0,411	0,323	0,358	0,46	0,424	0,558	0,48
$U5$	0,148	0,424	0,388	0,476	1,396	0,357	0,231	0,355	0,335	0,371	0,482	0,47
$U6$	0,75	0,541	0,512	0,515	0,526	2,056	0,377	0,497	0,399	0,459	0,692	0,589
$E1$	0,399	0,333	0,361	0,361	0,303	0,336	1,832	0,221	0,336	0,309	0,461	0,286
$E2$	0,069	0,091	0,063	0,086	0,1	0,095	0,047	0,394	0,069	0,099	0,195	0,108
$E3$	0,492	0,59	0,519	0,552	0,473	0,382	0,361	0,343	1,968	0,433	0,523	0,479
$E4$	0,396	0,442	0,403	0,451	0,465	0,391	0,295	0,439	0,385	1,749	0,64	0,437
$E5$	0,211	0,233	0,218	0,231	0,235	0,229	0,171	0,338	0,181	0,249	0,681	0,249
$E6$	0,453	0,483	0,434	0,477	0,549	0,467	0,255	0,447	0,397	0,407	0,596	1,631

Таблица 7. Коэффициенты частной корреляции по усредненным коэффициентам информативности

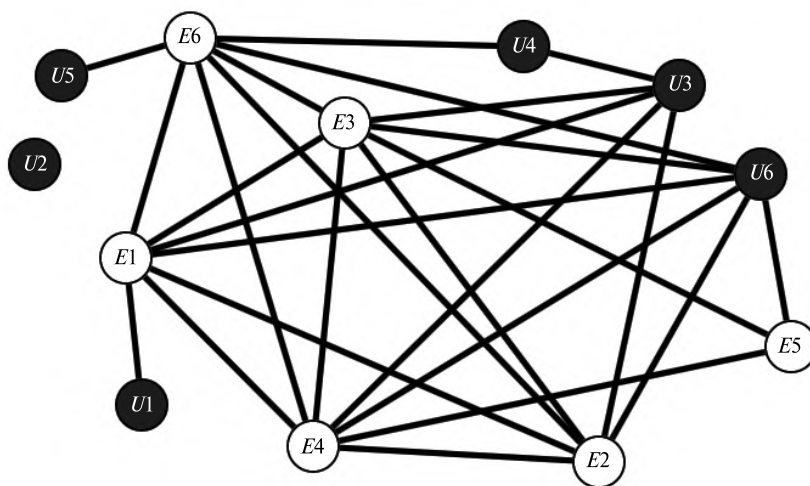
Показатель	$U1$	$U2$	$U3$	$U4$	$U5$	$U6$	$E1$	$E2$	$E3$	$E4$	$E5$	$E6$
$U1$	1	0,496	0,586	0,532	0,18	0,74	0,417	0,211	0,497	0,424	0,417	0,505
$U2$	0,496	1	0,532	0,519	0,427	0,457	0,295	0,209	0,506	0,399	0,358	0,451
$U3$	0,586	0,532	1	0,507	0,458	0,492	0,368	0,182	0,51	0,421	0,413	0,47
$U4$	0,532	0,519	0,507	1	0,518	0,463	0,342	0,222	0,506	0,437	0,395	0,478
$U5$	0,18	0,427	0,458	0,518	1	0,441	0,267	0,228	0,404	0,418	0,358	0,51
$U6$	0,74	0,457	0,492	0,463	0,441	1	0,356	0,296	0,391	0,425	0,46	0,528
$E1$	0,417	0,295	0,368	0,342	0,267	0,356	1	0,134	0,348	0,302	0,316	0,271
$E2$	0,211	0,209	0,182	0,222	0,228	0,296	0,134	1	0,206	0,269	0,266	0,277
$E3$	0,497	0,506	0,51	0,506	0,404	0,391	0,348	0,206	1	0,409	0,352	0,438
$E4$	0,424	0,399	0,421	0,437	0,418	0,425	0,302	0,269	0,409	1	0,445	0,422
$E5$	0,417	0,358	0,413	0,395	0,358	0,46	0,316	0,266	0,352	0,445	1	0,423
$E6$	0,505	0,451	0,47	0,478	0,51	0,528	0,271	0,277	0,438	0,422	0,423	1

Каждая строка табл. 6 (относится либо к E , либо к U) характеризует информационную силу данной переменной (кроме диагонального элемента с энтропиями) и указывает, какую долю неопределенности переменной столбца (также — либо E , либо U) она снимает в среднем, когда известны ее значения.

Данные табл. 7 были получены в предположении, что коэффициенты корреляции между переменными U , E равны средним значениям для каждой пары соответствующих коэффициентов информативности.

Сопоставление двух графов (рис. 2, 6) показывает их различие, поскольку один из них отражает возможности взаимных прогнозов, а другой — соотношения между условными энтропиями и общими неопределенностями. Однако роль отдельных показателей в уменьшении неопределенности видна и там и там, например для связей показателя $U1$. Наличие изолированных узлов вызвано тем, что для графа использовались только самые большие коэффициенты. Существенно также, что энтропийные связи могут быть несимметричными.

Дополнительно по наибольшим коэффициентам информативности в качестве независимых были отобраны переменные для стандартной линейной регрессии каждого показателя. Оказалось,

**Рис. 6.** Граф информационных взаимодействий построен по частным корреляциям для усредненных коэффициентов информативности

что качество полученной зависимости для показателей с большой энтропией практически совпадает с качеством регрессий в табл. 2, например для обрабатывающего производства. Для показателя, имеющего малую энтропию, линейная регрессия оказывается хуже (по значениям R^2). Для показателей информационных отраслей регрессии были достаточно хорошими, хотя и несколько слабее прямого отбора из табл. 2. Это представляется естественным, поскольку парные влияния слабее влияния сразу двух или более переменных

Парные коэффициенты информативности $\gamma_{i,j}$ аналогичны коэффициентам Пирсона, но не зависят от шкалы измерения. В нашем случае *сильных линейных взаимозависимостей их преимущество не может проявить себя в явном виде*. Однако сама идея энтропии позволяет нам предложить некий аналог коэффициента R^2 стандартного регрессионного анализа. Можно использовать отношение величины информации по Шеннону о показателе Y , заключенной, например, в паре показателей X, Z к самой энтропии Y :

$$\gamma(Y/X, Z) = I(Y, (X, Z)) / H(Y). \quad (5)$$

Для примера были рассчитаны коэффициенты информационного влияния пары $(U2, U3)$ и пары $(U3, U4)$ на показатели $U1, E3$, обладающие большой энтропией. Тем не менее, были получены коэффициенты $\gamma(U1/U2, U3) = 0,777$ и $\gamma(E3/U3, U4) = 0,772$, существенно большие, чем γ -коэффициенты влияния на $U1$ и $E3$ в табл. 4.

Поэтому для приближения к истинному графу непосредственных связей необходимо переходить к условным энтропиям более высокой размерности и находить марковские тройки типа $(Y, (X, (U, V)))$ с информацией по Шеннону, когда равенство $I(Y, (X, U, V)) = I(Y, X) \forall U, V$ означает, что среди переменных X, U, V только X непосредственно связан с Y . Однако в нашем случае весьма малой выборки это не представляет интереса.

ЗАКЛЮЧЕНИЕ: НЕКОТОРЫЕ ОБЩИЕ ВЫВОODY

1. В данной работе экономика РФ представлена как совокупность регионов со своими отраслевыми выпусками. Статистические данные о региональных выпусках рассматривались как характеристики элементов целостности экономической системы. Именно целостность системы приводит к наличию информационной связи между ее элементами, как это вытекает из общей концепции системного подхода. Необходимо заметить, что все расчеты и, соответственно, содержательные выводы были сделаны таким образом, как будто региональная экономика РФ без трех самых значительных регионов представляет собой изолированную самостоятельную экономическую систему. Неучет влияния и роли таких экономических центров, как Москва и Санкт-Петербург, конечно, частично искажает полученные содержательные выводы, но ничто не мешает использовать эту методологию для анализа экономик с учетом неоднородности объемов региональных выпусков. Надеемся в дальнейшем расширить наше исследование в этой области.

2. В работе предложены специальные методы статистического подхода к анализу региональных межотраслевых взаимосвязей. Одновременно предлагается рассматривать экономику в разрезе информационных взаимодействий между управляющими и управляемыми подсистемами. Для этого отраслевые региональные выпуски были разграничены по двум группам отраслей: производящих материальный продукт и отражающих деятельность информационного характера. Вместо симметричных коэффициентов корреляции линейной связи Пирсона предложено также измерять направленную силу информационных связей, позволяющих иногда делать взаимные прогнозы. Возможности анализа и прогноза были оценены двумя способами: через коэффициенты частных регрессий и через предложенные «коэффициенты информативности», отражающие степень взаимной статистической неопределенности между показателями при любых видах вероятностных распределений. Оба вида коэффициентов могут дополнять один другого при качественном анализе системы переменных. Так, полученные наилучшие частные регрессии позволяют осуществлять ориентировочный линейный прогноз, хотя распределение выпусков по величине не является Гауссовым. Аналогично, построенные по усредненным информационным коэффициентам частные корреляции указывают возможности прогнозов.

3. В работе было установлено, что энтропии показателей являются удобными характеристиками состояний системы и взаимодействий между отраслями экономики. К тому же их можно понимать как сложность и уровень собственного развития подсистем. К сожалению, измерение сложности объекта зависит от внешнего наблюдателя, т.е. исследователя. Понятие сложности по существу является не

онтологическим, а гносеологическим. Для демонстрации предложенного подхода авторы ограничились выбором нескольких вариантов градаций. Определение наиболее подходящих шкал измерения для разных случаев — отдельная проблема. Предполагается продолжить подобные исследования в дальнейшем.

В данной работе показано, что отрасль обрабатывающего производства имеет сложность, превосходящую по сложности и сектор связи, и сектор профессионально-научной деятельности, что, по-видимому, можно интерпретировать как отставание этих двух секторов от желательного уровня развития. Развитие всей экономики в значительной степени определяется уровнем обрабатывающего производства, которое, конечно, должно быть обеспечено развитием научно-профессиональных кадров и современными средствами связи. В то же время энтропия сектора государственного управления превосходит энтропию всех остальных отраслей и секторов, что можно интерпретировать как достаточный уровень управления региональной экономикой в целом.

4. Сравнение регионов по уровням развития информационных и производственных секторов выявило, что региональный выпуск на одного занятого больше в тех регионах, где энтропия управления превосходит энтропию производства. Этот факт означает, что развитие производственных отраслей в регионах должно сопровождаться развитием отраслей, производящих информацию.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Афанасьев М.Ю., Кудров А.В.** (2021). Экономическая сложность и вложенность структур региональных экономик // *Экономика и математические методы*. Т. 57. № 3. С. 67–78. [Afanasyev M.Yu., Kudrov A.V. (2021). Economic complexity and nesting of structures of regional economies. *Economics and Mathematical Methods*, 57, 3, 67–78 (in Russian).]
- Буданов В.Г., Аршинов В.И.** (2018). Сложность и проблема знания. М.: Институт философии. [Budanov V.G., Arshinov V.I. (2018). *Complexity and the problem of knowledge*. Moscow: Institute of Philosophy (in Russian).]
- Бир С.** (1965). Кибернетика и управление производством. М. Наука. [Bir S. (1965). *Cybernetics and production management*. Moscow: Nauka (in Russian).]
- Боженко Е.В., Махина Т.Ю.** (2018). Теория информации в экономико-математическом моделировании // *Журнал экономических исследований*. Т. 4. № 3. Режим доступа: <https://naukaru.ru/ru/nauka/article/20404/view> [Bozhenko E.V., Makhina T.Yu. (2018). Information theory in economic and mathematical modeling. *Journal of Economic Studies*, 4, 3. Available at: <https://naukaru.ru/ru/nauka/article/20404/view> (in Russian).]
- Гаврилец Ю.Н.** (1974). Социально-экономическое планирование. Системы и модели. М.: Экономика. [Gavrilets Yu.N. (1974). *Socio-economic planning. Systems and models*. Moscow: Ekonomika (in Russian).]
- Гаврилец Ю.Н., Кудров А.В., Тараканова И.В.** (2022). Статистический анализ и моделирование взаимосвязи региональной экономики и науки // *Экономика и математические методы*. Т. 58. № 4. С. 56–70. [Gavrilets Yu.N., Kudrov A.V., Tarakanova I.V. (2022). Statistical analysis and modeling of the relationship between regional economy and science. *Economics and Mathematical Methods*, 58, 4, 56–70 (in Russian).]
- Гаврилец Ю.Н.** (2016). Сложность, управление и информация в социально-экономических системах // *Международный журнал «Проблемы теории и практики управления»*. № 12. С. 39–49. [Gavrilets Yu.N. (2016). Complexity, management and information in socio-economic systems. *International journal "Problems of Theory and Practice of Management"*, 12, 39–49 (in Russian).]
- Колмогоров А.Н.** (1987). Избранные труды. Теория информации и теория алгоритмов. М.: Наука. [Kolmogorov A.N. (1987). *Selected works. Information theory and theory of algorithms*. Moscow: Nauka (in Russian).]
- Макаров В.Л.** (2015). О разнообразии экономического и политического устройства в мире. (Планирование возвращается?) // *Философия, методология и история науки*. Т. 1. № 1. С. 55–67. [Makarov V.L. (2015). On the diversity of the economic and political structure in the world. (Is planning back?) *Philosophy, Methodology and History of Science*, 1, 1, 55–67 (in Russian).]
- Моисеев А.К., Бондаренко П.А.** (2020). Применение индекса экономической сложности в макрофинансовых моделях // *Проблемы прогнозирования*. № 3. С. 101–112. [Moiseev A.K., Bondarenko P.A. (2020). Application of the economic complexity index in macrofinancial models. *Studies on Russian Economic Development*, 3, 101–112 (in Russian).]
- Новожилов В.В.** (1959). Измерение затрат и их результатов в социалистическом хозяйстве. В сб.: «*Применение математики в экономических исследованиях*». Т. 1. М.: Соцгиз. [Novozhilov V.V. (1959). Measurement of costs and their results in a socialist economy. In: *Application of mathematics in economic research*. Vol. 1. Moscow: Socegiz (in Russian).]
- Рыжкова Т.В.** (2012). Энтропийные показатели финансовой системы предприятия // *Вестник МГТУ*. № 2. С. 124–127. [Ryzhkova T.V. (2012). Entropy indicators of the financial system of an enterprise. *Herald of the Bauman Moscow State Technical University*, 2, 124–127 (in Russian).]

- Рыжкова Т.В.** (2010). Энтропийный анализ инвестиционной привлекательности компании // *Вестник Российского экономического университета им. Г.В. Плеханова*. № 6. Режим доступа: <https://cyberleninka.ru/article/n/entropiynyy-analiz-investitsionnoy-privlekatelnosti-kompanii/viewer> [Ryzhkova T.V. (2010). Entropy analysis of the investment attractiveness of the company. *Vestnik of the Plekhanov Russian University of Economics*, 6. Available at: <https://cyberleninka.ru/article/n/entropiynyy-analiz-investitsionnoy-privlekatelnosti-kompanii/viewer> (in Russian).]
- Тараканова И.В.** (2019). Выявление структуры статистических взаимозависимостей в системе переменных в ППК MATHCAD // *Вестник ЦЭМИ*. Т. 2. Вып. 1. DOI: 10.33276/S265838870004835-1 [Tarakanova I.V. (2019). Revealing the structure of statistical interdependencies in the system of variables in the MATHCAD software package. *Vestnik CEMI*, 1, 2. DOI: 10.33276/S265838870004835-1 (in Russian).]
- Щепанский Я.** (1969). Элементарные понятия социологии. М.: Прогресс. [Shchepansky Ya. (1969). *Elementary concepts of sociology*. Moscow: Progress (in Russian).]
- Яглом А.М., Яглом И.М.** (1973). Вероятность и информация. М.: Наука. [Yaglom A.M., Yaglom I.M. (1973). *Probability and Information*. Moscow: Nauka (in Russian).]
- Ebeling W., Molgedey L., Kurths J., Schwarz U.** (2002). Entropy, complexity, predictability, and data analysis of time series and letter sequences. In: *The science of disasters*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-56257-0
- Jakimowicz A.** (2020). The role of entropy in the development of economics. *Polish Academy of Sciences, Entropy*, 22 (4), 452.
- Mollaie S., Darooneh A.H., Karimi S.** (2019). Multi-scale entropy analysis and Hurst exponent. *Physica A. Statistical Mechanics and its Applications*, 528, 121292. DOI: 10.1016/j.physa.2019.121292

Experience of statistical and entropy analysis of relationships between the information and production sectors of the regional economy of Russia

© 2023 Yu.N. Gavrilets, I.V. Tarakanova

Yu.N. Gavrilets,

*Central Economics and Mathematics Institute, Russian Academy of Sciences (CEMI RAS), Moscow, Russia;
e-mail: yurkag@mail.ru*

I.V. Tarakanova,

*Central Economics and Mathematics Institute, Russian Academy of Sciences (CEMI RAS), Moscow, Russia;
e-mail: itar40@mail.ru*

Received 05.07.2023

Abstract. The work proposes special methods for a statistical approach to the analysis of regional intersectoral relationships. The real system, according to the statistics of which all quantitative calculations are carried out. The Russian Federation without the three largest administrative units (Moscow, St. Petersburg, Moscow Region) is considered. In contrast to the material ties of the Leontyev model, information and statistical links between industries are analyzed. For this, the concepts of entropy and information are used in order to identify the peculiarities of interaction between six production sectors and six information sectors of the economy. These interactions are considered from the standpoint of the law of the necessary diversity, and the complexity according to Kolmogorov, which is given special importance in the work and which is estimated by the entropy of the distributions of the output indicators of regional subsystems. Concepts are introduced and a quantitative assessment of the complexity of regions and industries is made. The structure of information links between indicators of the economic system is found both using regression analysis and based on the proposed information content coefficients. The role of conditional entropies in the statistical analysis of intersectoral interactions is especially emphasized. Both structures are represented as corresponding link graphs. It is justified that calculated coefficients of information links and conditional entropies provide the possibility of qualitative analysis and forecast. The calculations made according to statistics showed that in regions with a predominance of entropy of control over entropy of production, the output is significantly higher by employee. Based on the general consideration, it seems appropriate to recommend that the regions of Russia increase the complexity of the communication sectors and scientific and professional activities at least to the level of complexity of the manufacturing industry.

Keywords: production and information industries, entropy and information, Kolmogorov complexity, interconnection graph, law of necessary diversity.

JEL Classification: C01, C02.

For reference: **Gavrilets Yu.N., Tarakanova I.V.** (2023). Experience of statistical and entropy analysis of relationships between the information and production sectors of the regional economy of Russia. *Economics and Mathematical Methods*, 59, 4, 71–85. DOI: 10.31857/S042473880028218-8 (in Russian).