

1.5. ВЛОЖЕННОСТЬ РЕГИОНАЛЬНЫХ ЭКОНОМИК И ПРОГНОЗИРОВАНИЕ НОВЫХ СИЛЬНЫХ СЕКТОРОВ

Афанасьев М. Ю., д.э.н., ЦЭМИ РАН, Москва
Кудров А. В.¹, ЦЭМИ РАН, Москва

Рассматривается задача прогнозирования появления в регионе новых сильных секторов. На основе методов вероятностного и статистического моделирования построена модель, позволяющая оценить вероятность появления в регионе нового сильного сектора с учетом характеристик структуры экономики. Возможность построения такой модели основана на предположении о том, что появление и развитие секторов в значительной степени обусловлено эволюцией прошлой экономической деятельности. Модель использует введенные авторами показатели вложенности структур сильных секторов региональных экономик. Эти показатели основаны на вероятностной интерпретации и свойствах элементов матрицы, по которой оценивается экономическая сложность в соответствии с традиционным подходом. Для каждого региона оценена вероятность возникновения в его структуре сильного сектора. На основе упорядочения секторов по значению этих вероятностей и оценок их потенциального вклада в социально-экономическое развитие может быть дана экспертная оценка целесообразности развития в регионе нового сильного сектора.

1. Введение

Существуют две основные теории, описывающие механизм создания и распространения знаний: локализованная специализация и диверсификация экономики. Теория локализованной специализации впервые была подробно представлена в работе (Marshall, 1890). Она утверждает, что компании, находящиеся в окружении других представителей той же отрасли, будут расти быстрее в силу циркуляции знаний внутри отрасли. Эта теория была развита в работах (Arrow, 1962; Romer, 1986).

Противоположная теория и эмпирические оценки представлены в работах (Blien, Wolf, 2006; Fuchs, 2011; Illy, Schwartzetal., 2011). Согласно этой теории, компании выигрывают от того, что сталкиваются с неоднородной средой, состоящей из различных отраслей, поскольку новые идеи приходят из внешней среды. Механизмы, с помощью которых разнообразие приводит к экономическому росту, обычно называют диверсификацией.

С диверсификацией производства связаны современные представления об экономической сложности. Сравнительно недавно была разработана процедура, которая при использовании данных о структуре экспорта по странам позволяет *измерять экономическую сложность* (Hartmann, 2017; Hausmann, Rodrik 2003; Hausmann, Hwang, Rodrik, 2006; Hidalgo, Hausmann, 2009). Оценки экономической сложности характеризуют уровень вовлеченности стран в международные производственные процессы. Однако эту процедуру можно применять не только для структуры экспорта, но и для структуры выпуска экономики в целом. В этой работе акцент делается на исследовании экономики регионов РФ на основе данных о налоговых поступлениях по секторам экономики, что позволяет характеризовать структуры региональных экономик, включающие сектора, ориентированные как на внешний, так и на внутренний рынки, а также обнаруживать формирование групп смежных секторов.

Представлена вероятностная интерпретация элементов матрицы, по которой оценивается экономическая сложность. Каждый элемент этой матрицы можно рассматривать как характеристику степени вложенности структуры сильных секторов одного региона в структуру другого. Введены агрегированные показатели, описывающие вложенность структуры сильных секторов каждого региона в структуры экономик других регионов. Показано, что агрегированные показатели вложенности являются статистически значимыми объясняющими переменными для экономической сложности. Высокие оценки экономической сложности регионов соответствуют большим значениям агрегированных показателей вложенности, а низкие оценки — малым значениям.

На основе методов вероятностного и статистического моделирования построена модель, позволяющая оценить вероятность появления в регионе нового сильного сектора с учетом характеристик структуры экономики. Для каждого региона оценена вероятность возникновения в его структуре конкретного сектора в качестве сильного.

Полученные результаты не противоречат предположению, что эффекты от локализации и диверсификации могут дополнять друг друга и не являются взаимоисключающими. Оценки вероятности возникновения в структуре региона конкретного сектора в качестве сильного указывают, что появление и развитие секторов в значительной степени обусловлено эволюцией прошлой экономической деятельности. Эти результаты согласуются с выводами в работе (Neffke, Henning, Boschma, 2011), где показа-

¹ Афанасьев Михаил Юрьевич — д.э.н., г.н.с. ЦЭМИ РАН, профессор ГАУГН; miafan@cemi.rssi.ru.
Кудров Александр Владимирович — к.ф.-м.н., в.н.с. ЦЭМИ РАН, Москва; kovlal@inbox.ru.

но, что регионам легче развивать новые отрасли промышленности, если они связаны с уже имеющимися в регионе.

2. Методология

Для описания структуры региональной экономики использованы данные об объемах производства в достаточно широкой номенклатуре секторов. Сначала определим показатель RCA_{cp} выявленных сравнительных преимуществ:

$$RCA_{cp} = (y_{cp} / \sum_p y_{cp}) / (\sum_c y_{cp} / \sum_{cp} y_{cp}), \quad (1)$$

где y_{cp} — объем производства сектора p экономики региона c ; RCA_{cp} — отношение доли производства от сектора p в общем объеме производства от всех секторов экономики региона c к доле производства сектора p по всем регионам в объеме производства от всех секторов экономики всех регионов. В соответствии с работой (Hausmann, Klinger, 2006), для выявления сравнительных преимуществ в экономиках используется показатель RCA_{cp} для которого проверяется условие типа ограничения снизу. А именно, если значение RCA_{cp} превышает единицу, то считается, что экономика региона c обладает выявленными сравнительными преимуществами в выпуске продукции сектора p ; в противном случае — выявленных сравнительных преимуществ не существует:

$$a_{c,p} = \begin{cases} 1, & \text{если } RCA_{cp} \geq 1; \\ 0, & \text{если } RCA_{cp} < 1. \end{cases}$$

Матрица $A = (a_{c,p})$ содержит данные о секторах экономики, которые в разных регионах развиты на уровне выявленных сравнительных преимуществ, определенных при помощи выражения (1). Строки этой матрицы соответствуют регионам, столбцы — секторам экономики. Вектор $(a_{c,p_1}, \dots, a_{c,p_m})$ будем называть *структурой сильных секторов* экономики региона c .

Понятие «*экономическая сложность региона*» рассматривается как характеристика, отражающая уровень его технологического развития, который, в свою очередь, определяется сильными секторами в структуре его экономики. Аналогично экономическая сложность сектора зависит от уровня технологического развития тех регионов, в которых этот сектор присутствует в структуре в качестве сильного. Экономическая сложность является латентной характеристикой региона ECI_c или сектора ECI_p . Оценки экономической сложности обладают следующими свойствами: экономическая сложность региона пропорциональна среднему уровню экономической сложности сильных секторов в структуре его экономики:

$$ECI_c = a_1 \sum_p r_{c,p} ECI_p, \quad r_{c,p} = a_{c,p} / k_{c,0}, \quad k_{c,0} = \sum_p a_{c,p}, \quad (2)$$

где a_1 — положительная константа. Заметим, что $k_{c,0}$ не могут быть равны нулю, поскольку для любого c существует p , для которого $a_{c,p} = 1$.

Экономическая сложность сектора пропорциональна среднему уровню экономической сложности регионов, в структуре экономик которых этот сектор является сильным:

$$ECI_p = a_2 \sum_c r_{p,c}^* ECI_c, \quad r_{p,c}^* = a_{c,p} / k_{p,0}, \quad k_{p,0} = \sum_c a_{c,p}, \quad (3)$$

где a_2 — положительная константа. Показатель $k_{c,0}$, равный числу сильных секторов в регионе c , будем называть *диверсификацией структуры экономики* региона c . Пусть $\mathbf{c} = (ECI_{c_1}, ECI_{c_2}, \dots)^T$ — вектор-столбец значений экономической сложности для регионов; $\mathbf{p} = (ECI_{p_1}, ECI_{p_2}, \dots)^T$ — вектор-столбец значений экономической сложности для секторов; $\mathbf{R}_1 = (r_{c,p})$, $\mathbf{R}_2 = (r_{p,c}^*)$ — матрицы весов.

Запишем соотношения (2) и (3) в матричном виде: $\mathbf{c} = a_1 \mathbf{R}_1 \mathbf{p}$, $\mathbf{p} = a_2 \mathbf{R}_2 \mathbf{c}$. Откуда следует, что $\mathbf{c} = a_1 a_2 \mathbf{R}_1 \mathbf{R}_2 \mathbf{c}$, $\mathbf{p} = a_1 a_2 \mathbf{R}_2 \mathbf{R}_1 \mathbf{p}$. Таким образом, экономическая сложность региона определяется как собственный вектор матрицы $\mathbf{R}_1 \mathbf{R}_2$, а экономическая сложность сектора — собственный вектор матрицы $\mathbf{R}_2 \mathbf{R}_1$. Заметим, что элемент на пересечении строки i и столбца j матрицы $\mathbf{R}_1 \mathbf{R}_2$, т.е. $(\mathbf{R}_1 \mathbf{R}_2)_{ij}$, задается формулой $\sum_t (a_{c_i,p_t} a_{c_j,p_t} / k_{p_t,0}) / k_{c_i,0}$.

Матрица $\mathbf{R}_1 \mathbf{R}_2$ является стохастической. Свойство стохастичности матрицы $\mathbf{R}_1 \mathbf{R}_2$ справедливо, так как элементы матрицы неотрицательны, а их сумма по строкам равна 1:

$$\sum_j (\mathbf{R}_1 \mathbf{R}_2)_{ij} = \sum_j \frac{1}{k_{c_i,0}} \sum_t (a_{c_i,p_t} / k_{p_t,0}) = \sum_t \sum_j (a_{c_i,p_t} a_{c_j,p_t} / k_{c_i,0} k_{p_t,0}) = 1$$

В силу стохастичности матрица $\mathbf{R}_1 \mathbf{R}_2$ имеет собственное значение, равное 1, и отвечающий ему собственный вектор, который состоит из одинаковых координат. В работах (Hausmann, Rodrik, 2003; Sciarraetal., 2020) в качестве значений оценок экономической сложности регионов и секторов предлагается использовать стандартизированное значение собственного вектора матриц $\mathbf{R}_1 \mathbf{R}_2$, которые соответствуют второму максимальному собственному значению.

Агрегированная вложенность.

Обозначим $w_{i,j} = (R_1 R_2)_{ij} = \sum_{t=1}^m \frac{a_{c_i,p_t} a_{c_j,p_t}}{k_{c_i,0} k_{p_t,0}} = \frac{1}{k_{c_i,0}} \sum_{t=1}^m \frac{a_{c_i,p_t} a_{c_j,p_t}}{k_{p_t,0}}$ и отметим некоторые свойства элементов этой матрицы.

1. Поскольку матрица $R_1 R_2$ является стохастической, то для каждого $i, j \in \{1, \dots, n\}$: $w_{i,j} \geq 0$, $\sum_{j=1}^m w_{i,j} = 1$. Таким образом, $w_{i,1}, \dots, w_{i,m}$ можно интерпретировать как распределение вероятностей.

2. Если в регионе c_i существует хотя бы один сильный регион, то $w_{i,j} > 0$. Иначе, $w_{i,j} = 0$. Справедливость этого утверждения легко показать, поскольку $w_{i,i} = (k_{c_i,0})^{-1} \sum_{t=1}^m (a_{c_i,p_t}/k_{p_t,0}) \geq 0$, а нулевое значение достигается только тогда, когда $a_{c_i,p_t} = 0$, $t = 1, \dots, m$, но таких случаев в наших данных нет.

3. Элементы $w_{i,j}$ равны нулю тогда и только тогда, когда выполняется условие $\{t: a_{c_i,p_t} = 1\} \cap \{t: a_{c_j,p_t} = 1\} = \emptyset$. Это означает отсутствие общих сильных секторов в структурах экономики регионов c_i и c_j .

4. В каждой строке матрицы $(w_{i,j})$ максимальный элемент отвечает диагональному элементу, т.е. $w_{i,j} = \max_{j \in \{1, \dots, n\}} (w_{i,j})$. Покажем это. В силу справедливости выражения

$$w_{i,j} = \frac{1}{k_{c_i,0}} \sum_{t=1}^m \frac{a_{c_i,p_t} a_{c_j,p_t}}{k_{p_t,0}} \leq \frac{1}{k_{c_i,0}} \sum_{t=1}^m \frac{a_{c_i,p_t}}{k_{p_t,0}} = \frac{1}{k_{c_i,0}} \sum_{t=1}^m \frac{a_{c_i,p_t} a_{c_i,p_t}}{k_{p_t,0}}$$

получаем, что $w_{i,j} \leq w_{i,i}$. Причем равенство в последнем неравенстве достигается только тогда, когда выполняется условие $\{t: a_{c_i,p_t} = 1\} \subseteq \{t: a_{c_j,p_t} = 1\}$. А это означает, что все сильные сектора структуры экономики региона c_i также являются сильными секторами в структуре региона c_j . Если условие не выполняется, то имеем строгое неравенство $w_{i,j}/w_{i,i} < 1$.

5. Величина $w_{i,j}$ монотонно возрастает с ростом числа элементов множества $\{t: a_{c_i,p_t} = 1\} \cap \{t: a_{c_j,p_t} = 1\}$.

6. Асимметричность матрицы $(w_{i,j})$. Легко показать, что $w_{j,i} = k_{c_i,0}/k_{c_j,0} w_{i,j}$. Если уровень диверсификации региона c_i совпадает с уровнем диверсификации региона c_j , тогда $w_{j,i} = w_{i,j}$. Если диверсификация региона c_i больше (меньше) диверсификации региона c_j , тогда $w_{j,i} > w_{i,j}$ ($w_{j,i} < w_{i,j}$). Таким образом, различные уровни диверсификации регионов гарантируют асимметричность матрицы $(w_{i,j})$.

Из свойств 1–6 для $w_{i,j}$ следует, что отношение $w_{i,j}/w_{i,i}$ можно интерпретировать как характеристику *степени вложенности* множества сильных секторов региона c_i во множество сильных секторов региона c_j . Чем ниже это отношение, тем меньше сильных секторов региона c_i входит во множество $\{t: a_{c_j,p_t} = 1\}$ сильных секторов региона c_j .

Агрегированный показатель $I_i^{(1)} = \sum_{j=1}^m (w_{i,j}/w_{i,i})^2$, $i = 1, \dots, m$, характеризует степень вложенности структуры сильных секторов региона c_i в структуры сильных секторов других регионов. Назовем этот показатель *агрегированной вложенностью структуры* экономики региона. Минимальное значение $I_i^{(1)} = 1$ возникает в условиях, когда структура сильных секторов региона c_i состоит из уникальных секторов, для которых $w_{i,j} = 0$ для всех $i \neq j$, близкое к единице — когда распределение $w_{i,j}$ мало отличается от равномерного распределения.

Показатель $I_j^{(2)} = \sum_{i=1}^m (w_{i,j}/w_{i,i})^2$, $j = 1, \dots, m$, характеризует степень вложенности в структуру сильных секторов региона c_j . Назовем его *агрегированной вложенностью в структуру* региона. При увеличении числа сложных секторов региона c_j величина $I_j^{(2)}$ не убывает. Минимальное значение $I_j^{(2)} = 1$ возникает в условиях, когда структура сложных секторов для региона c_j состоит из уникальных секторов, т.е. когда $w_{i,j} = 0$ для всех $i \neq j$.

Связь показателей агрегированной вложенности с экономическим базисом. В работе (Айвазян, Афанасьев, Кудров, 2018) было представлено понятие *экономического базиса*, компоненты которого являются характеристиками дифференциации, формируемыми с помощью теоретически обоснованных моделей регионального развития. Положение региона в базисе характеристик дифференциации определяет его экономическое своеобразие. Более подробное описание см. в работе (Айвазян, Афанасьев, Кудров, 2018). Рассмотрим регрессию логарифма показателя степени вложенности структуры экономики на компоненты экономического базиса.

Таблица 1: Регрессия логарифма показателя вложенности структуры эко-

	Оце	Стан-	t-	p-		
Константа	2,19	0,04	59,52	0		
Размер экономики	-	0,04	-4,51	0		
Техническая эффективность	-	0,04	-1,67	0,1		
Индекс отраслевой специализации	-	0,04	-4,65	0		
Индекс индустриализации	-	0,04	-1,36	0,18		
Прирост технической эффективности	0,04	0,04	1,08	0,29		
R-квадрат	0,43					
Скор.R-квадрат	0,4					

Источник: расчеты авторов.

Оценки параметров регрессии логарифма показателя вложенности структуры экономики на экономический базис показывают, что значимыми и отрицательными являются коэффициенты при размере экономики и индексе отраслевой специализации.

Таблица 2. Регрессия логарифма показателя вложенности в структуру экономики на компоненты экономического базиса.

	Оценка	Стандартное отклонение	t-статистика	p-значение
Константа	2,13	0,05	42,47	0,00
Размер экономики	-0,03	0,05	-0,54	0,59
Техническая эффективность	0,01	0,05	0,24	0,81
Индекс отраслевой специализации	-0,32	0,05	-5,97	0,00
Индекс индустриализации	0,04	0,05	0,69	0,50
Прирост технической эффективности	0,06	0,05	1,23	0,22
R-квадрат	0,36			
Скор. R-квадрат	0,31			

Источник: расчеты авторов.

Результаты оценки параметров регрессии логарифма показателя вложенности в структуру экономики на экономический базис показывают, что значимой объясняющей переменной является только индекс отраслевой специализации.

В работе (Afanasyev, Kudrov, 2020) представлены оценки экономической сложности регионов РФ, рассчитанные на основе данных о налоговых поступлениях по 82 секторам экономики. В табл.1 приведены оценки параметров регрессии экономической сложности регионов РФ, полученной в работе (Afanasyev, Kudrov, 2020), на логарифм агрегированной вложенности структуры сильных секторов экономики и логарифм агрегированного вложения в структуру.

Таблица 3. Регрессия для экономической сложности регионов

Экономическая сложность региона	Оценка	Стандартное отклонение	t-статистика	p-значение
Константа	-4,18	0,45	-9,22	0
Логарифм агрегированной вложенности структуры	0,96	0,19	5,14	0
Логарифм агрегированной вложенности в структуру	0,97	0,15	6,65	0
R ²	0,55			
Скорректированный R ²	0,54			

Примечание. Данные по оценкам экономической сложности взяты из работы (Afanasyev, Kudrov, 2020).

Источник: расчеты авторов.

Большие значения экономической сложности из (Afanasyev, Kudrov, 2020) соответствуют высокому уровню показателя вложенности структуры экономики и высокому уровню показателя вложенности в структуру. Для экономической сложности низкого уровня характерны низкие уровни показателя вложенности структуры экономики и показателя вложенности в структуру. Таким образом, показатель экономической сложности можно интерпретировать в качестве измерителя уровня развитости смежных секторов и технологических цепочек или цепочек добавленной стоимости.

Для российских регионов получены оценки экономической сложности, для которых были использованы и другие региональные данные. А именно, данные по экспорту регионов РФ. В работе (Любимов и др., 2017) по этим данным вычислялись две оценки экономической сложности регионов РФ: оценка экономической сложности по данным о структуре экспорта регионов РФ; оценка экономической сложности по данным о структуре экспорта регионов РФ и стран мира. В табл. 4 показаны результаты построения регрессии этих оценок экономической сложности на логарифмы показателей агрегированной вложенности структуры сильных секторов региона и агрегированной вложенности в структуру сильных секторов региона.

Показатели экономической сложности из (Любимов и др., 2017) по данным без стран характеризуются низким соответствием показателям вложенности. Объяснением этому может служить то, что экспортные товары вложены в международные цепочки формирования добавленной стоимости. Поэтому если мы ограничиваемся только экспортными данными по регионам, то в силу отсутствия длинных цепочек формирования добавленной стоимости для экспортных товаров на локальном российском уровне, оценка экономической сложности не отражает связей между товарами российского экспорта. Напротив, если мы подключаем к имеющимся данным о структуре экспорта регионов РФ информацию о структуре экспорта стран мира и оцениваем по ним уровень экономической сложности, то такая оценка отражает уже технологические цепочки формирования добавленной стоимости на международном уровне и позволяет количественно оценивать в них место регионов РФ. В результате, как видно из данных в табл. 4 (слева), этот показатель имеет значимую связь с показателем вложенности в структуру.

Таблица 4. Регрессии для экономической сложности регионов по данным со странами и без стран

Регрессия для экономической сложности регионов по данным со странами					Регрессия для экономической сложности регионов по данным без стран				
Экономическая сложность (по данным со странами)	Оценка	Стандартное отклонение	t-статистика	p-значение	Экономическая сложность (по данным без стран)	Оценка	Стандартное отклонение	t-статистика	p-значение
Константа	0,15	0,25	0,61	0,54	Константа	-1,03	0,67	-1,55	0,13
Логарифм агрегированной вложенности структуры	-0,19	0,10	-1,87	0,06	Логарифм агрегированной вложенности структуры	0,06	0,27	0,22	0,82
Логарифм агрегированной вложенности в структуру	0,31	0,08	3,92	0,00	Логарифм агрегированной вложенности в структуру	0,42	0,21	1,95	0,06
R ²	0,17				R ²	0,05			
Скорректированный R ²	0,15				Скорректированный R ²	0,03			

Примечание. Данные по оценкам экономической сложности взяты из работы (Lyubimov et al., 2017).
 Источник: расчеты авторов.

Графы регион-регион по вложенности структур сильных секторов.

Методом имитации проверены гипотезы о равенстве нулю показателя вложенности структуры сильных секторов региона $c_i, i = 1, \dots, n$, в структуру сильных секторов региона $c_j, i = 1, \dots, n, j \neq i$, а именно $H_0: w_{i,j}/w_{i,i} = 0$ против $H_1: w_{i,j}/w_{i,i} \neq 0$. Показано, что 95%-квантиль для функции распределения $w_{i,j}/w_{i,i}$ в условиях справедливости нулевой гипотезы равен 0,47 при заданном уровне значимости $\alpha = 0,95$.

Определим матрицу $L^{max} = (l_{i,j}^{max} | i, j \in \{c_1, \dots, c_n\})$ следующим образом:

$$l_{i,j}^{max} = \begin{cases} \frac{w_{i,j}}{w_{i,i}}, & \text{если } \frac{w_{i,j}}{w_{i,i}} > 0,47 \text{ и } \frac{w_{i,j}}{w_{i,i}} = \max_j \frac{w_{i,j}}{w_{i,i}} \\ 0, & \text{иначе} \end{cases}$$

Условие $\frac{w_{i,j}}{w_{i,i}} > 0,47$ возникает в результате требования о том, что рассматриваются только статистически отличные от нуля вложения.

Пусть L^{max} – матрица смежности для графа “максимальной вложенности” $G^{max} = (V^{max}, E^{max})$ с направленными весовыми ребрами, в котором множество узлов V^{max} отвечает множеству регионов, а E^{max} – множество узлов графа, которое задается матрицей смежности L^{max} , см. Рис.1:

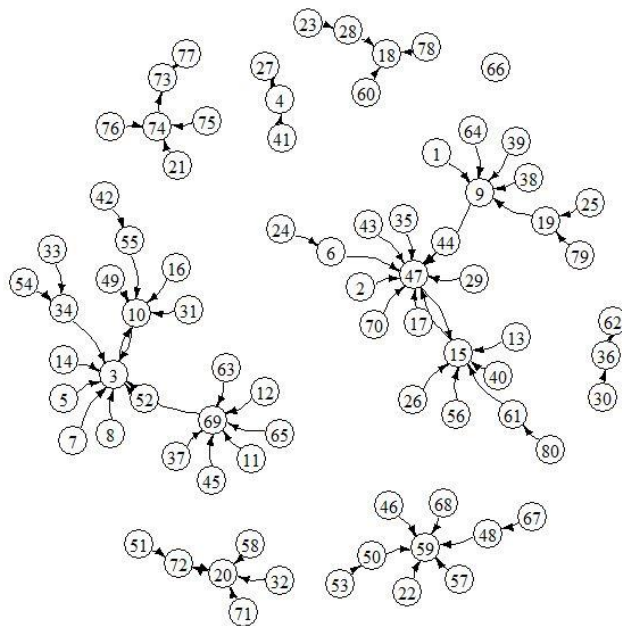


Рисунок 1. Граф “максимальной вложенности”, $G^{max} = (V^{max}, E^{max})$

Этот граф отражает наиболее сильные вложения структур сильных секторов. Отметим, что номера узлов соответствуют нумерации регионов, используемой в работах (Айвазян, Афанасьев, Кудров, 2018). Из графа видим, что наибольшее число входящих ребер соответствуют: 1. Узлу 47 (Чувашская республика) с числом входящих ребер равным 10; 2. Узлу 3 (Владимирская область) с числом входящих ребер равным 8; 3. Узлу 69 (Новосибирская область) с числом входящих ребер равным 6, который, в свою очередь, вкладывается в узел 3. Эти регионы характеризуются тем, что их экономики отличаются наиболее длинными технологическими цепочками, которые в разной степени реплицируются регионами, соответствующими либо непосредственно, либо опосредованно входящим ребрам.

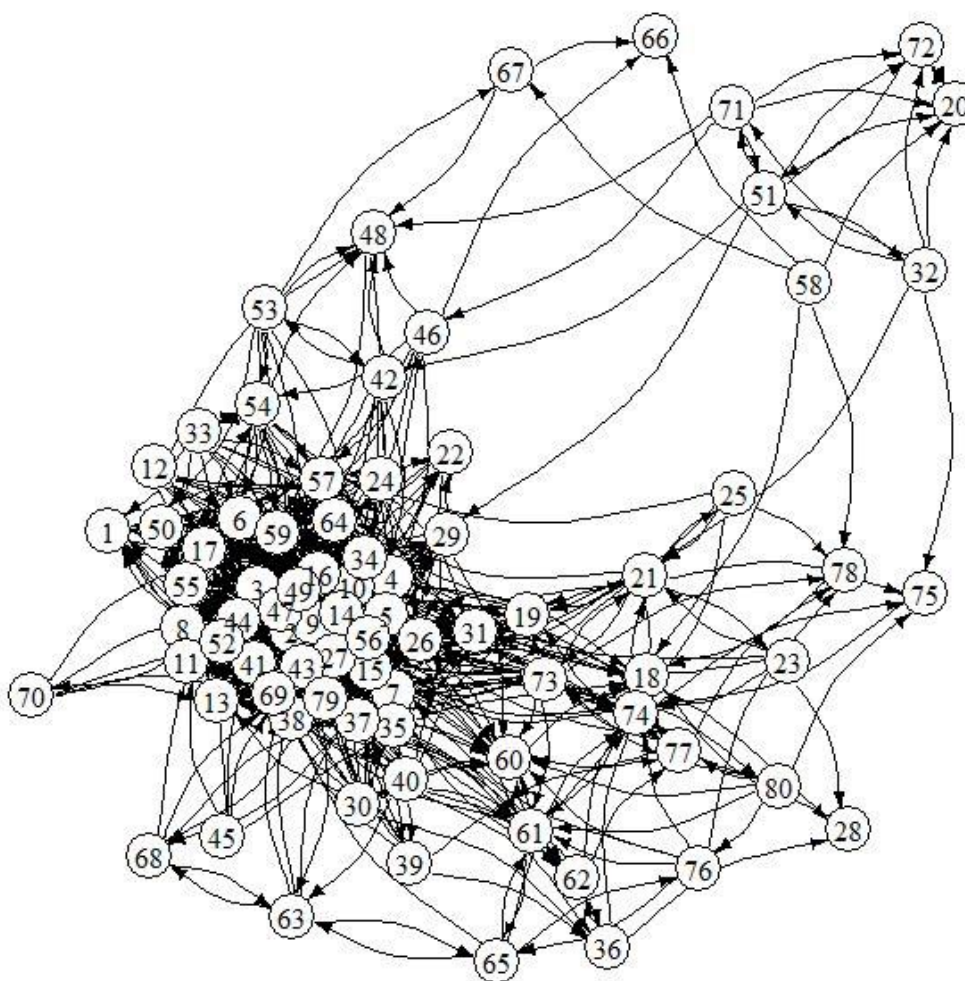
Отдельно следует отметить вложение узла 23 в узел 28, что отвечает Калининградской области и г. Санкт-

Петербургу, соответственно. Для этого вложения коэффициент вложенности равен 0,69. Коэффициенты диверсификации Калининградской области и г. Санкт-Петербурга равны 15 и 23, соответственно. А число их общих секторов равно 8. Также отметим вложение узла 28 в узел 18, что отвечает г. Санкт-Петербургу и г. Москве, соответственно. Коэффициент вложенности равен 0,60. Коэффициенты диверсификации Калининградской области и г. Санкт-Петербурга равны 23 и 24, соответственно. А число их общих секторов равно 11.

Теперь рассмотрим матрицу $L^{signif} = (l_{i,j}^{signif} | i, j \in \{c_1, \dots, c_n\})$, которая определяется следующим образом:

$$l_{i,j}^{signif} = \begin{cases} \frac{w_{i,j}}{w_{i,i}}, & \text{если } \frac{w_{i,j}}{w_{i,i}} > 0,47. \\ 0, & \text{иначе} \end{cases}$$

Пусть L^{signif} – матрица смежности для графа “значимых вложенностей” $G^{signif} = (V^{signif}, E^{signif})$ с направленными весовыми ребрами, изображенного на рис.2:



Граф на рис.2 отражает все “значимые вложенности”, которые позволяют учесть больше информации по сравнению с тем, когда используются только “максимальные вложенности”. Далее при формировании модели будет использован этот подход.

Рисунок 2. Граф “значимых вложенностей”, $G^{signif} = (V^{signif}, E^{signif})$

Вероятностная модель для прогноза появления новых сильных секторов

Пусть случайные величины

$$A_{c_i,p_k} = \begin{cases} 1, & \text{с вероятностью } \pi(f_{c_i,p_k}^1, f_{c_i,p_k}^2, \dots) \\ 0, & \text{с вероятностью } 1 - \pi(f_{c_i,p_k}^1, f_{c_i,p_k}^2, \dots) \end{cases} \quad (4)$$

где $f_{c_i,p_k}^1, f_{c_i,p_k}^2, \dots$ — факторы, влияющие на возникновение событий $A_{c_i,p_k} = 1$ и $A_{c_i,p_k} = 0$. Будем предполагать, что зависимость $\pi(f_{c_i,p_k}^1, f_{c_i,p_k}^2, \dots)$ имеет вид

$$\pi(f_{c_i,p_k}^1, f_{c_i,p_k}^2, \dots) = \frac{\exp(\beta_0 + \sum_h \beta_h f_{c_i,p_k}^h)}{1 + \exp(\beta_0 + \sum_h \beta_h f_{c_i,p_k}^h)} \quad (5)$$

где β_0 — константа.

Предположим, что последовательность $a_{c_1,p_1}, \dots, a_{c_1,p_m}, \dots, a_{c_n,p_1}, \dots, a_{c_n,p_m}$ является реализацией случайной последовательности $A_{c_1,p_1}, \dots, A_{c_1,p_m}, \dots, A_{c_n,p_1}, \dots, A_{c_n,p_m}$, элементы которой независимы при заданных $f_{c_1,p_1}^1, f_{c_1,p_2}^2, \dots$. Кроме того, будем считать, что распределение $P(A_{c_i,p_k} | f_{c_1,p_1}^1, f_{c_1,p_2}^2, \dots) = P(A_{c_i,p_k} | f_{c_i,p_k}^1, \dots, f_{c_i,p_k}^h, \dots)$, т.е. факторы, влияющие на распределение A_{c_i,p_k} , ограничиваются множеством $f_{c_i,p_k}^1, \dots, f_{c_i,p_k}^h, \dots$, у которых нижняя пара индексов (c_i, p_k) . Тогда функция правдоподобия имеет вид

$$\begin{aligned} & P(A_{c_1,p_1} = a_{c_1,p_1}, \dots, A_{c_1,p_m} = a_{c_1,p_m}, \dots, A_{c_n,p_1} = a_{c_n,p_1}, \dots, A_{c_n,p_m} = a_{c_n,p_m} | f_{c_1,p_1}^1, f_{c_1,p_2}^2, \dots) \\ &= P(A_{c_1,p_1} = a_{c_1,p_1} | f_{c_1,p_1}^1, f_{c_1,p_1}^2, \dots) \dots P(A_{c_n,p_m} = a_{c_n,p_m} | f_{c_n,p_m}^1, \dots, f_{c_n,p_m}^h, \dots) \\ &= \left[\frac{\exp(\beta_0 + \sum \beta_h f_{c_1,p_1}^h)}{1 + \exp(\beta_0 + \sum \beta_h f_{c_1,p_1}^h)} \right]^{a_{c_1,p_1}} \left[\frac{1}{1 + \exp(\beta_0 + \sum \beta_h f_{c_1,p_1}^h)} \right]^{1-a_{c_1,p_1}} \times \dots \\ &\times \left[\frac{\exp(\beta_0 + \sum \beta_h f_{c_n,p_m}^h)}{1 + \exp(\beta_0 + \sum \beta_h f_{c_n,p_m}^h)} \right]^{a_{c_n,p_m}} \left[\frac{1}{1 + \exp(\beta_0 + \sum \beta_h f_{c_n,p_m}^h)} \right]^{1-a_{c_n,p_m}} \end{aligned}$$

Максимизируя эту функцию, мы получаем оценки параметров $\beta_1, \dots, \beta_h, \dots$. Для оценки качества модели мы будем использовать критерий Акаике, а также величину $G = mn^{-1} \sum_{i,k} I(I(\pi(f_{c_1,p_k}^1, f_{c_1,p_k}^2, \dots) > 0,5) = a_{c_i,p_k})$, где $I(A)$ — показатель события A . Величина G характеризует долю правильно идентифицированных моделью случаев.

Рассмотрим способы формирования объясняющих переменных $f_{c_1,p_k}^1, \dots, f_{c_i,p_k}^h, \dots$

Способ 1. Для данного способа $h = 2$. Определим факторы $f_{c_i,p_k}^1, f_{c_i,p_k}^2$ для всех $i = 1, \dots, n$; $k = 1, \dots, m$:

$$\begin{aligned} f_{c_i,p_k}^1 &= \left(\sum_{j=1, i \neq j}^n w_{i,j} a_{c_j,p_k} \right) / (1 - w_{i,i}); \\ f_{c_i,p_k}^2 &= \left(\sum_{j=1, i \neq j}^n w_{j,i} a_{c_j,p_k} / w_{j,j} \right) / \left(\sum_{j=1, i \neq j}^n w_{j,i} / w_{j,j} \right). \end{aligned}$$

Оценки параметров модели (4)–(5) при данном способе формирования объясняющих факторов представлены в табл.5.

Таблица 5. Способ 1. Оценка параметров модели идентификации сильных секторов

Параметр	Оценка	Стандартное отклонение	t-статистика	p-значение
Константа	-2,94	0,07	-43,42	0,00
f_{c_i,p_k}^1	2,91	0,83	3,52	0,00
f_{c_i,p_k}^2	18,15	0,91	20,12	0,00
G	0,8			
1. Критерий Акаике	5681			

Источник: расчеты авторов.

Как видно из данных в табл. 5, коэффициенты при f_{c_i,p_k}^1 и f_{c_i,p_k}^2 являются значимыми и положительными. Для этого способа выбора объясняющих переменных доля правильно идентифицированных моделью случаев составляет 80%.

Способ 2. Для данного способа $h = 2$. Определим факторы $f_{c_i,p_k}^1, f_{c_i,p_k}^2$ для всех $i = 1, \dots, n$; $k = 1, \dots, m$:

$$\begin{aligned} f_{c_i,p_k}^1 &= \sum_{j=1, i \neq j}^n w_{i,j} I(w_{i,j}/w_{i,i} > 0,47) a_{c_j,p_k} / \sum_{j=1, i \neq j}^n w_{i,j} I(w_{i,j}/w_{i,i} > 0,47); \\ f_{c_i,p_k}^2 &= \sum_{j=1, i \neq j}^n \frac{w_{i,j}}{w_{j,j}} I(w_{j,i}/w_{j,j} > 0,47) a_{c_j,p_k} / \sum_{j=1, i \neq j}^n (w_{j,i}/w_{j,j}) I(w_{j,i}/w_{j,j} > 0,47). \end{aligned}$$

Оценки параметров модели (4)–(5) при данном способе формирования объясняющих факторов представлены в табл.6.

Таблица 6. Способ 2. Оценка параметров модели идентификации сильных секторов

Параметр	Оценка	Стандартное отклонение	t-статистика	p-значение
Константа	-3,11	0,07	-41,70	0,00
f_{c_i,p_k}^1	4,74	0,26	18,46	0,00
f_{c_i,p_k}^2	5,81	0,27	21,95	0,00
G	0,83			
2. Критерий Акаике	4554			

Источник: расчеты авторов.

Как видно из данных в табл. 6, коэффициенты при f_{c_i,p_k}^1 и f_{c_i,p_k}^2 являются значимыми и положительными. Для этого способа выбора объясняющих переменных доля правильно идентифицированных

моделью случаев незначительно отличается от доли для случая 1 и равна 83% общего числа случаев, но значение критерия Акаике меньше, чем для способа 1. Таким образом, способ 2 предпочтительнее.

3. Результаты моделирования

Согласно данной модели, были сформированы списки потенциально сильных секторов в регионах. В частности, для некоторых регионов, специализирующихся в сельском хозяйстве, требуется развитие сектора «Транспортировка и хранение», что дает возможность большего территориального охвата в снабжении производимой продукцией. Вследствие инфраструктурных проблем рекомендации для развития сельского хозяйства включают развитие сектора «Забор, очистка и распределение воды» (необходим для функционирования поливочных систем), сектора «Обеспечение электрической энергией, газом и паром; кондиционирование воздуха» (для развития тепличных комплексов, требующих больших энергозатрат при создании микроклимата). Кроме того, в некоторых регионах со специализацией в сельском хозяйстве недостаточно хорошо развиты сектора обрабатывающей промышленности для производимой сельхозпродукции. (Подробнее см. в табл. 7: производство напитков (Владимирская область, Псковская область, Воронежская область, Костромская область и др. регионы), переработка и консервирование мяса и мясной пищевой продукции (Оренбургская область, Пензенская область и др.), производство пищевых продуктов (Тульская область, Ярославская область), производство молочной продукции (Саратовская область, Республика Северная Осетия — Алания и др.)). Следует заметить, что развитие сельского хозяйства способствует формированию спроса на развитие сектора сельхозмашиностроения и других секторов, в том числе секторов, которые предлагают решения, способствующие автоматизации процессов на уровнях сельскохозяйственных работ и функционирования обрабатывающей промышленности. Сектора, связанные с сельским хозяйством, приведены здесь в качестве примера. Спектр выявленных потенциальных сильных секторов в структурах региональных экономик по результатам использования модели является более широким.

Таблица 7. Оценки вероятностей появления новых сильных секторов

Рекомендации построенной по способу 1		Рекомендации построенной по способу 2	
Регион	Вероятность возникновения сектора в качестве сильного	Регион	Вероятность возникновения сектора в качестве сильного
Сектор: транспортировка и хранение		Сектор: транспортировка и хранение	
Белгородская область	0,86	Белгородская область	0,89
Воронежская область	0,61	Воронежская область	0,63
Курская область	0,73	Курская область	0,57
Московская область	0,66	Московская область	0,67
г. Санкт-Петербург	0,95	г. Санкт-Петербург	0,92
Республика Адыгея	0,94	Республика Адыгея	0,91
Сектор: водоснабжение, водоотведение, организация сбора и утилизации отходов, деятельность и ликвидация загрязнений		Сектор: водоснабжение, водоотведение, организация сбора и утилизации отходов, деятельность и ликвидация загрязнений	
Орловская область	0,97	Орловская область	0,91
Рязанская область	0,98	Рязанская область	0,94
Смоленская область	0,98	Смоленская область	0,93
Тамбовская область	0,78	Тамбовская область	0,57
Республика Карелия	0,67	Республика Карелия	0,54
Республика Калмыкия	0,67	Республика Калмыкия	0,75
Краснодарский край	0,61	Краснодарский край	0,77
Астраханская область	0,83	Астраханская область	0,90
Ростовская область	0,64	Ростовская область	0,67
Карачаево-Черкесская Республика	0,60	Кабардино-Балкарская Республика	0,65
Удмуртская Республика	0,75	Карачаево-Черкесская Республика	0,87
Челябинская область	0,62	Республика Марий Эл	0,64
Красноярский край	0,82	Республика Татарстан	0,65
Иркутская область	0,71	Удмуртская Республика	0,82
Кемеровская область	0,87	Ульяновская область	0,57
Приморский край	0,59	Свердловская область	0,63
Хабаровский край	0,74	Тюменская область	0,72
Амурская область	0,50	Челябинская область	0,88
Магаданская область	0,50	Республика Бурятия	0,64
Сектор: обеспечение электрической энергией, газом и паром; кондиционирование воздуха — всего		Сектор: обеспечение электрической энергией, газом и паром; кондиционирование воздуха — всего	
Республика Алтай	0,52	Алтайский край	0,57
		Красноярский край	0,89
		Иркутская область	0,77

Сектор: производство напитков		Кемеровская область	0,89
Владимирская область	0,69	Приморский край	0,57
Воронежская область	0,94	Хабаровский край	0,69
Костромская область	0,83	Сектор: обеспечение электрической энергией, газом и паром; кондиционирование воздуха — всего	
Калининградская область	0,99	Архангельская область	0,69
Ленинградская область	0,99	Республика Калмыкия	0,81
Мурманская область	0,99	Республика Татарстан	0,81
Псковская область	0,64	Новосибирская область	0,75
Краснодарский край	0,89	Сектор: производство напитков	
Астраханская область	0,62	Владимирская область	0,70
Республика Марий Эл	0,87	Воронежская область	0,88
Пермский край	0,61	Калининградская область	0,79
Свердловская область	0,69	Ленинградская область	0,90
Тюменская область	0,92	Мурманская область	0,89
Челябинская область	0,57	Псковская область	0,56
Иркутская область	0,67	Краснодарский край	0,88
Томская область	0,75	Республика Марий Эл	0,87
Камчатский край	0,51	Пермский край	0,75
Амурская область	0,59	Свердловская область	0,80
Магаданская область	0,68	Тюменская область	0,87
Сахалинская область	0,95	Челябинская область	0,52
Еврейская автономная область	0,93	Иркутская область	0,64
Сектор: переработка и консервирование мяса и мясной пищевой продукции		Томская область	0,72
Республика Карелия	0,54	Амурская область	0,51
Мурманская область	0,84	Еврейская автономная область	0,87
Республика Ингушетия	0,85	Сектор: переработка и консервирование мяса и мясной пищевой продукции	
Кабардино-Балкарская республика	0,80	Мурманская область	0,88
Чеченская Республика	0,76	Республика Ингушетия	0,90
Республика Башкортостан	0,52	Кабардино-Балкарская Республика	0,89
Оренбургская область	0,60	Чеченская Республика	0,71
Пензенская область	0,88	Республика Башкортостан	0,75
Забайкальский край	0,71	Удмуртская Республика	0,53
Томская область	0,64	Оренбургская область	0,70
Республика Саха (Якутия)	0,93	Пензенская область	0,86
Амурская область	0,61	Забайкальский край	0,72
Сахалинская область	0,55	Иркутская область	0,70
Сектор: производство пищевых продуктов		Республика Саха (Якутия)	0,93
Тульская область	0,96	Амурская область	0,61
Ярославская область	0,96	Сахалинская область	0,71
г. Москва	0,67	Сектор: производство пищевых продуктов	
Республика Ингушетия	0,56	Тульская область	0,95
Омская область	0,52	Ярославская область	0,95
Республика Саха (Якутия)	0,59	г.Москва	0,62
Сектор: производство молочной продукции		Сектор: производство молочной продукции	
Республика Коми	0,97	Республика Северная Осетия — Алания	0,95
Республика Северная Осетия — Алания	0,78	Республика Коми	0,73
Чеченская Республика	0,92	Чеченская Республика	0,92
Самарская область	0,55	Саратовская область	0,91
Саратовская область	0,88	Забайкальский край	0,62
Забайкальский край	0,51	Красноярский край	0,54
Красноярский край	0,56	Иркутская область	0,66
Иркутская область	0,64	Приморский край	0,92
Приморский край	0,91	Хабаровский край	0,84
Хабаровский край	0,87		

4. Заключение

Свойства матрицы, по которой оценивается экономическая сложность из работы (Hausmann, Hidalgo, Bustos, Coscia, Simoes, Yildirim, 2011), позволяют интерпретировать любой (i, j) - элемент этой матрицы как характеристику степени вложенности структуры сильных секторов -го региона в структуру j -го региона. На основании свойств элементов этой матрицы вводятся агрегированные показатели, характеризующие вложенность структуры сильных секторов каждого региона в структуры экономик всех других регионов.

Агрегированные показатели вложенности оказываются статистически значимо связанными с некоторыми компонентами экономического базиса, представленного в работе (Айвазян, Афанасьев, Кудров, 2018). Наиболее сильная связь проявляется с индексом отраслевой специализации, которая

указывает на то, что большие значения показателей вложенности отвечают регионам со специализацией в обрабатывающей промышленности.

Построена модель, которая позволяет объяснить появление и отсутствие сильного сектора в структуре экономики региона. На ее основе получено условие появления определенного сильного сектора в структуре экономики конкретного региона с вероятностью, превышающей 50%. Это условие использовано для формирования списка секторов, рекомендованных для приоритетного развития в регионе. Для каждого региона оценена вероятность возникновения в его структуре конкретного сектора в качестве сильного. На основе упорядочения секторов по значению этих вероятностей и оценок их потенциального вклада в социально-экономическое развитие, может быть дана экспертная оценка целесообразности развития в регионе нового сильного сектора.

Литература

1. Любимов И.Л., Гвоздева М.А., Казакова М.В., Нестерова К.В. (2017). Сложность экономики и возможность диверсификации экспорта в российских регионах // Журнал Новой экономической ассоциации. № 2 (34). С. 94–122. [Lyubimov I.L., Gvozdeva M.A., Kazakova M.V., Nesterova K.V. (2017). Economic complexity of Russian regions and their potential to diversify. Journal of the New Economic Association, 2 (34), 94–122 (in Russian).]
2. Айвазян С.А., Афанасьев М.Ю., Курдов А.В. (2018). Индикаторы экономического развития в базе характеристик региональной дифференциации // Прикладная эконометрика, издательство Синергия (М.), том 50, № 2, с. 4–22
3. Afanasiev M.Yu., Kudrov A.V. (2020). Estimates of economic complexity in the structure of the regional economy. Montenegrin Journal of Economics, 16, 4, 43–54. DOI: 10.14254/1800-5845/2020.16-4.4
4. Arrow K.J. (1962). The economic implications of learning by doing. The Review of Economic Studies, 29 (3), 155–173.
5. Blien U., Wolf K. (2006). Local employment growth in West Germany: A dynamic panel approach. Labour Economics, 13 (4), 445–458.
6. Fuchs M. (2011). The determinants of local employment dynamics in Western Germany. Empirical Economics, 40 (1), 177–203.
7. Glodberg K., Roeder T., Gupta D., Perkins C. (2001). Eigentaste: A constant time collaborative filtering algorithm. Information Retrieval, 4, 133–151.
8. Hartmann D. (2017). Linking economic complexity, institutions, and income inequality. World Development, 93, 75–93.
9. Hausmann R., Hidalgo C., Bustos S., Coscia M., Simoes A., Yildirim M.A. (2011). The atlas of economic complexity: Mapping paths to prosperity. Cambridge: Center for International Development, Harvard University, MIT.
10. Hausmann R., Hwang J., Rodrik D. (2006). What you export matters. Journal of Economic Growth, 12 (1), 1–25.
11. Hausmann R., Klinger B. (2006). Structural transformation and patterns of comparative advantage in the product space. CID Working Paper No. 128.
12. Hausmann R., Rodrik D. (2003). Economic development as selfdiscovery. Journal of Development Economics, 72 (2), 603–633.
13. Hidalgo C.A., Hausmann R. (2009). The building blocks of economic complexity. Proceedings of the National Academy of Sciences, 106 (26), 10570–10575.
14. Hidalgo C.A., Klinger B., Barabasi A.-L., Hausmann R. (2007). The product space conditions the development of nations. Science, 317 (5837), 482–487.
15. Illy A., Schwartz M., Hornyk C., Rosenfeld M. (2011). Local economic structure and sectoral employment growth in German cities. Journal of Economic and Social Geography, 102 (5), 582–593.
16. Marshall A. (1890). Principles of Economics. London: MacMillan.
17. Neffke F., Henning M., Boschma R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Economic Geography, 87, 3, 237–265.
18. Romer P.M. (1986). Increasing returns and long-run growth. Journal of Political Economy, 94 (5), 1002–1037.
19. Sciarra C., Chiarotti G., Ridolfi L. et al. (2020). Reconciling contrasting views on economic complexity. Nat Commun, 11, 3352. DOI: 10.1038/s41467-020-16992-1.

References in Cyrillics

1. Lyubimov I.L., Gvozdeva M.A., Kazakova M.V., Nesterova K.V. (2017). Slozhnost` e`konomiki i vozmozhnost` diversifikacii e`ksporta v rossijskix regionax // Zhurnal Novoj e`konomie-skoj associacii. № 2 (34). S. 94–122. [Lyubimov I.L., Gvozdeva M.A., Kazakova M.V., Nesterova K.V. (2017). Economic complexity of Russian regions and their potential to diversify. Journal of the New Economic Association, 2 (34), 94–122 (in Russian).]

2. Ajvazyan S.A., Afanas`ev M.Yu., Kudrov A.V. (2018). Indikatory` e`konomicheskogo razvitiya v bazise karakteristik regional`noj differenciacii // Prikladnaya e`konometrika, izda-tel`stvo Sinergiya (M.), tom 50, № 2, s. 4-22

Афанасьев Михаил Юрьевич, д.э.н., ЦЭМИ РАН, Москва; miafan@cemi.rssi.ru.

Кудров Александр Владимирович, к.ф.-м.н., ЦЭМИ РАН, Москва; kovlal@inbox.ru

Mikhail Afanasiev Central Economics and Mathematics Institute of the Russian Academy of Sciences, Russia, Moscow; miafan@cemi.rssi.ru

Ключевые слова

региональная экономика, эконометрика, экономическая сложность, вложенность

Mikhail Afanasiev, Alexander Kudrov, EMBEDDING AND THE EMERGENCE FORECASTING OF NEW STRONG SECTORS OF REGIONAL ECONOMIES

Keywords

regional economics, econometrics, economic complexity, embedding

DOI: 10.34706/DE-2021-04-05

JEL classification C53 – Методы прогнозирования • Методы моделирования, D51 – Экономика товарообмена и производства

Abstract

It is considered the problem of forecasting the emergence of new strong sectors in the region. On the basis of probabilistic and statistical modeling the model which allows to estimate the probability of appearance of a new strong sector in the region taking into account characteristics of economic structure is constructed. The possibility of building such a model is based on the assumption that the emergence and development of sectors is largely determined by the evolution of past economic activity. The model uses the indicators of embedding structures of the strong sectors of regional economies introduced by the authors. These indicators are based on the probabilistic interpretation and properties of the elements of the matrix, by which economic complexity is estimated according to the traditional approach. For each region the probability of occurrence in its structure of a particular sector as a strong sector is estimated. Based on the ordering of sectors according to the value of these probabilities and assessments of their potential contribution to socio-economic development, an expert assessment of the feasibility of developing a new strong sector in the region can be made.