

Прикладная эконометрика, 2016, т. 41, с. 24–46.  
Applied Econometrics, 2016, v. 41, pp. 24–46.

С. А. Айвазян, М. Ю. Афанасьев, А. В. Кудров<sup>1</sup>

## Метод кластеризации регионов РФ с учетом отраслевой структуры ВРП

*Представлена методология выявления однородных групп регионов РФ, каждой из которых соответствует своя модель производственного потенциала, определяющая зависимость ВРП от стоимости основных фондов и численности занятых. Проводится проверка гипотезы о том, что параметры функций, описывающих производственный потенциал регионов из разных групп, отличаются в силу особенностей отраслевой структуры ВРП. Предложена основанная на компонентном анализе процедура построения вспомогательных интегральных индикаторов, отражающих специализацию регионов. Разработан и апробирован на данных за период 2009–2013 гг. алгоритм формирования однородной группы регионов. Выявлено наличие большей чувствительности ВРП к изменению стоимости основных фондов для групп регионов со специализацией в добывающей промышленности и специализацией смешанного типа, чем для регионов со специализацией в обрабатывающей промышленности и сельском хозяйстве. В то же время, две последние группы регионов и группа регионов с формирующейся экономикой имеют большую чувствительность к изменению численности занятых.*

**Ключевые слова:** региональная экономика; производственный потенциал; эконометрическое моделирование; проверка гипотез.

**JEL classification:** C12; C51; R15.

### Введение

Оценка эффективности регионов РФ с позиции рационального использования ресурсов представляет собой важную задачу, которая еще более актуализируется в условиях, когда экономика РФ испытывает проблемы, а у почти половины регионов дефицит бюджета превышает средний дефицит по всем регионам (Зубаревич, Горина, 2015).

В качестве измерителя эффективности регионов РФ естественно использовать соотношение между фактическим и потенциальным размером экономики (в терминах ВРП), рассчитанным в предположении рационального использования имеющихся факторов производства, т. е. труда и капитала. В рамках такой концепции «эффективность отражает степень успеха, который достигают производители при распределении имеющихся ресурсов и формировании выпуска, следуя своим целям, ... а именно ... для достижения высокой степени эффективности затрат, выручки или чистой прибыли» (Kumbhakar, Lovell, 2000). В работе (Kumbhakar, Lovell, 2000) указано, что под производственными возможностями, отвечающими макси-

<sup>1</sup> Айвазян Сергей Артемьевич — ЦЭМИ РАН, Москва; aivazian@cemi.rssi.ru.  
Афанасьев Михаил Юрьевич — ЦЭМИ РАН, Москва; miafan@cemi.rssi.ru.  
Кудров Александр Владимирович — ЦЭМИ РАН, Москва; kovlal@inbox.ru.

мальной эффективности, следует понимать две возможности: а) максимальный выпуск при заданном профиле ресурсов и уровне технологий (производственная граница); б) заданный уровень выпуска, обеспечиваемый минимальными ресурсами (граница затрат).

В (Farrell, 1957) были впервые проведены эмпирические измерения неэффективности по выпуску, рассчитанные как отклонения от идеальной производственной границы, а также было предложено различать две составляющие экономической эффективности:

- *техническую эффективность* (technical efficiency), которая позволяет измерять отклонения фактического выпуска от максимально возможного при заданном профиле ресурсов, где максимально возможный выпуск при каждом возможном профиле ресурсов отвечает производственной границе;

- *эффективность распределения ресурсов* (allocative efficiency), которая позволяет измерить степень оптимальности распределения ресурсов.

В настоящем исследовании внимание сосредоточено на оценке технической эффективности использования регионами РФ таких факторов производства, как трудовые ресурсы и физический капитал. Поскольку производственная граница не наблюдаема, для ее оценки обычно применяется модель граничного производственного потенциала (SFA-модель). Впервые эта модель была представлена и подробно описана в работах (Aigner et al., 1977) и (Meeusen, van den Broeck, 1977). Основная идея модели граничного производственного потенциала состоит в использовании аддитивной ошибки, включающей две составляющие: 1) шум, сформированный случайными факторами; 2) случайная составляющая, сформированная в результате влияния факторов неэффективности. Для каждой из них делаются предположения относительно вида функции распределения. Кроме того, они полагаются независимыми. Чаще всего для случайной составляющей, отвечающей неэффективности, рассматривается полунормальное распределение (Tsionas, Kumbhakar, 2004), но встречаются и работы, где используется экспоненциальное, логнормальное (Tsionas, 2006) или гамма-распределение (Bera, Sharma, 1999; van den Broeck et al., 1994). Распределения с большим числом параметров являются более гибкими, но проблемы, возникающие с оценкой параметров, часто перевешивают преимущества гибкости.

Оценке технической эффективности регионов РФ посвящено совсем немного работ, среди которых следует выделить (Айвазян и др., 2014). В ней представлены модели граничного производственного потенциала регионов РФ, где в производственной функции, наряду с классическими факторами, трудовыми ресурсами и физическим капиталом, используется оценка интеллектуального капитала, а среди факторов неэффективности рассматриваются характеристики уровня благосостояния, качества жизни и доля инновационно-активных предприятий в общем числе предприятий региона.

Для проведения успешной бюджетной и социально-экономической политики необходимо учитывать региональные особенности, и на основании этого оценивать нереализованный потенциал. В данной статье проверяется гипотеза о зависимости параметров модели граничного производственного потенциала регионов РФ от характеристик структуры экономики. Естественно предполагать, что производственные функции для некоторых регионов отличаются в силу особенностей устройства региональных экономик и разных конкурентных преимуществ, но при этом имеются однородные группы похожих регионов, которым соответствует единая производственная функция.

Например, следует ожидать, что производственная функция регионов с преобладанием в ВРП добывающей промышленности должна быть более чувствительной к изменению фи-

зического капитала и менее чувствительной к изменению трудовых ресурсов, чем производственная функция регионов с преобладанием обрабатывающей промышленности, поскольку организация добычи полезных ископаемых требует больших капитальных затрат и сравнительно небольших трудовых ресурсов. Отметим, что ВРП включает расходы регионально-го бюджета на здравоохранение, образование, социальную защиту и др., поэтому регионам, получающим значительные трансферты из федерального бюджета, должна соответствовать производственная функция, которая менее чувствительна к изменениям трудовых ресурсов. В периоды экономического роста в более сильных регионах создается больше рабочих мест, а в периоды спада сильные регионы чувствительнее реагируют на изменения, что проявляется в более резком сжатии рынка труда. В слабых регионах ситуация в экономике плохая вне зависимости от экономической конъюнктуры и поэтому следует ожидать, что изменение ВРП в таких регионах менее чувствительно к изменению рыночных факторов производства. Таким регионам будет требоваться поддержка из федерального бюджета на протяжении длительного времени, поскольку собственных ресурсов для развития у них явно недостаточно, что должно отражаться в характерной для этой группы производственной функции. В связи с гипотезой о влиянии структуры экономики на вид производственной функции возникает обобщение модели граничного производственного потенциала, которое включает в себя, как частный случай, модель, отвечающую отсутствию влияния структуры экономики на производственную функцию. В данной статье будет представлена методология выявления однородных групп регионов РФ, в каждой из которых есть своя производственная функция типа Кобба–Дугласа, учитывающая структурные особенности экономики регионов, что выражается в различной чувствительности ВРП к изменению трудовых ресурсов и величины капитала.

Статья имеет следующую структуру. В разделе 1 представлено описание модели граничного производственного потенциала с локально-постоянными коэффициентами. В разделе 2 описана методология выявления однородных групп и оценки параметров модели. В разделе 3 перечислены виды данных, используемых для эмпирического анализа. В разделе 4 излагается процедура построения вспомогательных интегральных индикаторов, характеризующих специализацию регионов РФ. В разделе 5 продемонстрированы результаты эмпирического анализа предлагаемой модификации модели граничного производственного потенциала на примере данных по регионам РФ за период 2009–2013 гг. Общие выводы по результатам работы приводятся в заключительном разделе 6.

## 1. Модель граничного производственного потенциала с локально-постоянными коэффициентами для ВРП регионов РФ

Предполагается, что модель зависимости ВРП регионов РФ от основных факторов производства имеет вид:

$$P_i(t) = e^{\beta_0(q_i(t))} [L_i(t)]^{\beta_1(q_i(t))} [K_i(t)]^{\beta_2(q_i(t))} \exp(\varepsilon_i(t)), \quad (1)$$

$$\varepsilon_i(t) = v_i(t) - u_i(t), \quad v_i(t) \sim N\left(0, [\sigma_v(q_i(t))]^2\right), \quad u_i(t) \sim N^+\left(0, [\sigma_u(q_i(t))]^2\right),$$

$$i = 1, \dots, I, \quad t = 1, \dots, T,$$

где  $P_i(t)$  — валовый региональный продукт  $i$ -го региона в год  $t$ ;  $L_i(t)$  — численность экономически активного населения  $i$ -го региона в год  $t$ ;  $K_i(t)$  — объем основных фондов для  $i$ -го

региона в год  $t$ ;  $q_i(t)$  — значение вектор-характеристики структуры экономики для  $i$ -го региона в год  $t$ ;  $\varepsilon_i(t)$  — случайная величина, сформированная в результате воздействия различных источников неопределенности для  $i$ -го региона в год  $t$ ;  $\beta_0(q_i(t))$ ,  $\beta_1(q_i(t))$ ,  $\beta_2(q_i(t))$  — параметры модели;  $v_i(t)$  — случайная величина, характеризующая влияние случайных факторов;  $u_i(t)$  — неотрицательная случайная величина, характеризующая влияние факторов неэффективности;  $I$  — число регионов;  $T$  — число рассматриваемых лет.

Заметим, что параметры модели  $\beta_0(q_i(t))$ ,  $\beta_1(q_i(t))$ ,  $\beta_2(q_i(t))$  предполагаются зависимыми от вектор-характеристик  $q_i(t)$ , что отражает возможность изменения параметров модели в зависимости от специализации регионов. Каждый регион имеет определенный тип специализации. Обозначим число таких специализаций через  $m$ , которое также является параметром модели. Предполагается также, что существует такое разбиение области возможных значений  $q(\cdot)$  на  $m$  подмножеств  $H_1, \dots, H_m$ , для которого при любых  $q_i(t) \in H_j$ ,  $j = 1, \dots, m$ ,  $i = 1, \dots, I$ , функциональные параметры  $\beta_0(q_i(t))$ ,  $\beta_1(q_i(t))$ ,  $\beta_2(q_i(t))$  принимают постоянные значения. Таким образом:

$$\beta_n(q_i(t)) = \sum_{k=1}^m \gamma_k(n) I(q_i(t) \in H_k), \quad n = 0, 1, 2, \quad (2)$$

$$\sigma_v(q_i(t)) = \sum_{k=1}^m \delta_k(1) I(q_i(t) \in H_k), \quad (3)$$

$$\sigma_u(q_i(t)) = \sum_{k=1}^m \delta_k(2) I(q_i(t) \in H_k), \quad (4)$$

где  $\gamma_k(n)$  — возможные значения для  $\beta_n(q_i(t))$ ;  $\delta_k(1)$  и  $\delta_k(2)$  — возможные значения для  $\sigma_v(q_i(t))$  и  $\sigma_u(q_i(t))$  соответственно ( $k = 1, \dots, m$ );  $I(A)$  — индикатор множества  $A$ .

Логарифмируя выражение (1) и учитывая высказанные выше предположения, получим:

$$\begin{aligned} \ln P_i(t) = & \sum_{k=1}^m \gamma_k(0) I(q_i(t) \in H_k) + \left[ \sum_{k=1}^m \gamma_k(1) I(q_i(t) \in H_k) \right] \ln L_i(t) + \\ & + \left[ \sum_{k=1}^m \gamma_k(2) I(q_i(t) \in H_k) \right] \ln K_i(t) + \varepsilon_i(t). \end{aligned}$$

## 2. Методология выявления однородных групп и оценки параметров модели граничного производственного потенциала с локально-постоянными коэффициентами

Каждому региону соответствует вектор-характеристика  $q(\cdot)$ , которая позволяет описать особенности структуры экономики рассматриваемого региона (см. раздел 4). Заданному региону отвечают  $T$  значений вектор-характеристики  $q(\cdot)$ , каждое из которых, в свою очередь, соответствует определенному (из рассматриваемых) году. Далее будем называть их  $q$ -объектами. При идентификации однородных групп используются понятия расстояния между вектор-характеристиками, в качестве которого берется евклидово расстояние

$d^2(\cdot, \cdot)$ , и понятие расстояния между классами, в качестве которого будет использоваться  $K$ -расстояние  $\rho_\tau^{(K)}$  (обобщенное по Колмогорову расстояние) с параметром  $\tau = 2$  (см. (Айвазян и др., 1989)). Также экспертно задается минимальное число объектов, имеющееся в однородной группе ( $M = 15$ ).

На первом шаге выбирается  $q$ -объект, находящийся на границе области возможных значений  $H_1 \cup \dots \cup H_m$ , хотя бы одна из координат которого принимает экстремальное (максимальное или минимальное) значение. Такие объекты, как будет показано в разделе 5, отвечают регионам со специализацией только в одном из следующих секторов:

- добывающей промышленности;
- обрабатывающей промышленности;
- сельском хозяйстве.

Далее для этого выбранного объекта выбираются  $M$  ближайших по значению  $q(\cdot)$ . По соответствующим им наблюдениям оцениваются параметры модели (1) в предположении, что  $m = 1$ , и рассчитывается среднее значение логарифма функции правдоподобия (см. Приложение). Далее последовательно добавляются следующие ближайшие  $q$ -объекты, а по совокупности соответствующих им наблюдений переоцениваются параметры модели (1) при  $m = 1$ , и рассчитывается среднее значение логарифма функции правдоподобия. Из совокупности средних значений логарифма функции правдоподобия, упорядоченных по числу добавленных ближайших относительно значений вектора  $q(\cdot)$ , выбирается максимальное, а из соответствующих ему наблюдений формируется «базовая» выборка. Для соответствующей «базовой» выборке модели (1) при  $m = 1$  строится доверительный интервал средних значений логарифма функции правдоподобия при добавлении следующих наблюдений, сгенерированных из этой модели. К «базовой» выборке добавляются наблюдения, которые отвечают следующим ближайшим  $q$ -объектам, до тех пор пока не происходит первое пересечение средних значений логарифма функции правдоподобия нижней границы доверительного интервала.

Описанная выше процедура повторяется до тех пор, пока все наблюдения не будут распределены по группам. Более подробное описание представлено в Приложении.

### 3. Используемые для эмпирического анализа виды данных по регионам РФ

Эмпирической основой являлись региональные данные за период 2008–2012 гг., полученные из базы данных Росстата:

- ВРП в стоимостном выражении;
- отраслевая структура ВРП в процентах от ВРП регионов;
- численность экономически активного населения;
- объем основных фондов регионов в стоимостном выражении.

### 4. Построение интегральных индикаторов, характеризующих специализацию регионов РФ

Важной особенностью, которую следует учитывать при оценивании модели граничного производственного потенциала, является специализация региональной экономики. Но для количественного анализа необходимо выработать специальные интегральные индексы, позво-

ляющие отражать специализацию экономик регионов РФ. При построении такого индекса следует учитывать, что сектора, представленные в региональной экономике, могут либо генерировать основные доходы ВРП и отвечать определенной специализации экономики региона (например, добывающая или обрабатывающая промышленность, сельское хозяйство), либо формироваться для обеспечения определенного уровня социально-экономических отношений внутри региона (например, торговля, строительство, здравоохранение и др.).

При построении интегрального индикатора возникает проблема выбора характеристик, по которым можно судить о специализации регионов. В качестве таких данных в работе использовалась отраслевая структура валовой добавленной стоимости.

Как представляется, наиболее естественным способом построения интегрального индикатора является компонентный анализ (см. (Айвазян и др., 1989)).

При построении индикаторов, характеризующих специализацию региональных экономик, по данным об отраслевой структуре ВРП для обеспечения невырожденности эмпирической корреляционной матрицы исключались данные, отвечающие гостиницам и ресторанам, транспорту и связи, а также производству и распределению электроэнергии, газа, воды. С экспертной точки зрения для большинства регионов исключенные отрасли нельзя назвать специализацией экономики. Кроме того, к доле сельского хозяйства в ВРП была прибавлена доля рыболовства.

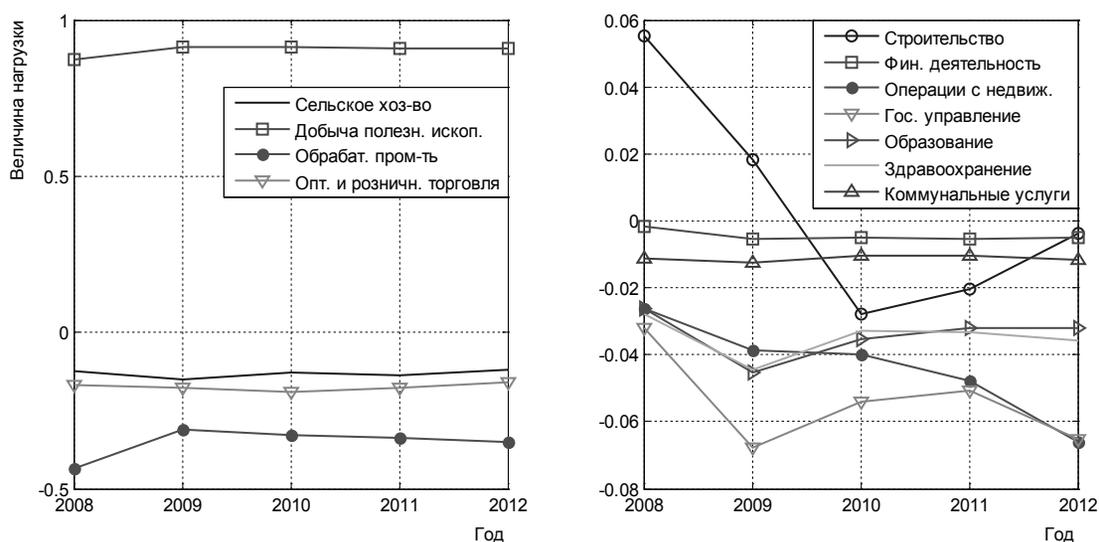
Для того чтобы исследовать устойчивость нагрузок и информативности главных компонент для отраслевой структуры ВРП по каждому году в отдельности за период с 2008 по 2012 г., рассмотрим результаты компонентного анализа. Как видно из табл. 1, доля объясненной дисперсии (информативность) для первой и второй главных компонент за каждый из годов за период 2008–2012 гг. превышает 78%. Однако информативность первой и второй главных компонент для 2008 г. отличается от информативности для 2009–2012 гг. Это может объясняться тем, что в 2008 г. регионы подверглись экономическому шоку, связанному с международным кризисом.

**Таблица 1.** Доля объясненной дисперсии (в %), приходящаяся на первую и вторую главные компоненты за период 2008–2012 гг.

	Год				
	2008	2009	2010	2011	2012
Первая главная компонента	54.00	63.28	65.41	65.89	64.98
Вторая главная компонента	24.16	19.36	19.60	19.60	20.14

Графики нагрузок на первую главную компоненту, рассчитанную по каждому из рассматриваемых годов, представлен на рис. 1.

Из рисунка 1 видно, что только нагрузка, отвечающая строительству, меняла знак с положительного в 2008–2009 гг. на отрицательный в 2010–2012 гг. Остальные нагрузки оставались знакопостоянными. Отметим также, что в 2008–2012 гг. положительный знак имеет только нагрузка, приходящаяся на добычу полезных ископаемых. В кризисном 2008 г. наблюдается меньшая, чем в последующие 2009–2012 гг., корреляция первой главной компоненты с долей в ВРП добычи полезных ископаемых и большая корреляция с долей в ВРП обрабатывающей промышленности. Кроме того, доля добычи полезных ископаемых в отраслевой структуре ВРП имеет максимальную корреляцию с первой главной компонентой,

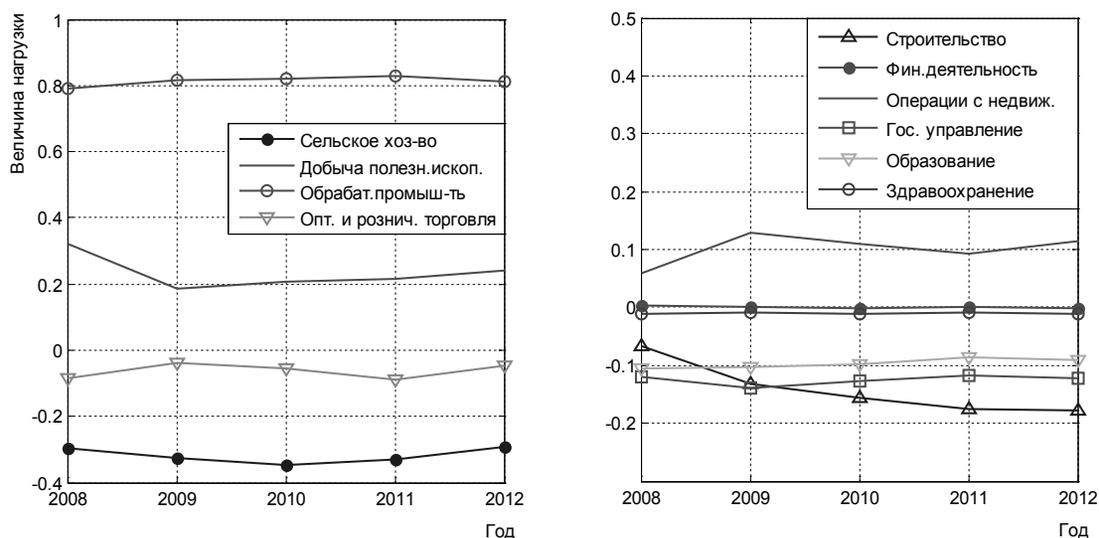


**Рис. 1.** Динамика нагрузок на первую главную компоненту, рассчитанную по данным о структуре ВРП по отраслям для регионов РФ для каждого года за период 2008–2012 гг.

которая достаточно устойчива на временном интервале с 2008 по 2012 г. Следующая максимальная корреляция с первой главной компонентой приходится на долю обрабатывающей промышленности в ВРП. Ее абсолютное значение в 2008 г. превышает абсолютные значения в 2009–2012 гг.

Графики нагрузок на вторую главную компоненту, рассчитанную для каждого года с 2008 по 2012 г., представлены на рис. 2.

Из результатов, представленных на рис. 2, видно, что нагрузки на вторую главную компоненту являются знакопостоянными. Кризисному 2008 году отвечает большая корреляция



**Рис. 2.** Динамика нагрузок на вторую главную компоненту

второй главной компоненты с долей в ВРП добычи полезных ископаемых и сельского хозяйства. Доли для остальных отраслей либо имеют устойчивую корреляцию со второй главной компонентой, построенной по данным за 2008 г., либо достаточно малы.

Рассмотрим проекции данных об отраслевой структуре ВРП за 2008–2011 гг. на первую и вторую главные компоненты, построенные по данным за 2012 г., и рассчитаем доли объясненных дисперсий. Результаты представлены в табл. 2.

**Таблица 2.** Доля объясненной дисперсии (в %), приходящаяся на первую и вторую главные компоненты, рассчитанные по данным за 2012 г.

	Год				
	2008	2009	2010	2011	2012
Первая главная компонента за 2012 год	53.38	63.01	65.22	65.79	64.98
Вторая главная компонента за 2012 год	23.92	19.40	19.54	19.50	20.14

Сравнивая результаты из табл. 2 с результатами из табл. 1, можно видеть, что потери информации при переходе к двум главным компонентам для данных за 2012 г., измеряемые величиной, на которую уменьшается доля объясненной дисперсии, не превышают 0.7% для двух главных компонент. При этом максимальная потеря информации приходится на 2008 г., для всех других годов потери не превышают 0.3%. Это указывает на устойчивость первой и второй главных компонент, рассчитанных по данным за каждый год на интервале с 2008 по 2012 г.

Назовем первую компоненту, построенную по данным отраслевой структуры, *индексом промышленной специализации*; а вторую главную компоненту — *индексом индустриализации*.

Определившись с устойчивостью весов в формуле для расчета индекса промышленной специализации (отраслевая структура ВРП) и индекса индустриализации (отраслевая структура ВРП), рассмотрим регионы с точки зрения динамики значений этих индексов. В таблице 3 приведена корреляционная матрица рассчитанных значений индекса промышленной специализации (отраслевая структура ВРП) для регионов РФ по каждому году за период 2008–2012 гг.

**Таблица 3.** Корреляционная матрица значений индекса промышленной специализации для регионов РФ за 2008–2012 гг.

	2008 г.	2009 г.	2010 г.	2011 г.	2012 г.
2008 г.	1.00	0.97	0.98	0.98	0.97
2009 г.	0.97	1.00	0.99	0.98	0.98
2010 г.	0.98	0.99	1.00	1.00	0.98
2011 г.	0.98	0.98	1.00	1.00	0.99
2012 г.	0.97	0.98	0.98	0.99	1.00

Как видно из табл. 3, между значениями индекса для регионов РФ, рассчитанными за соответствующий год, имеется значимая корреляция. Это означает, что большинство регионов РФ на протяжении 2008–2012 гг. имеют сравнительно стабильные значения индекса промышленной специализации для большинства регионов РФ. Однако существуют регионы, для которых индекс промышленной специализации за период с 2008 по 2012 г. значительно менялся. Максимальные изменения индекса соответствуют регионам, приведенным в табл. 4.

**Таблица 4.** Регионы с максимальным изменением индекса промышленной специализации за период с 2008 по 2012 г.

	Год	Индекс промышленной специализации	Доля добывающей промышленности в ВРП (в %)	Доля обрабатывающей промышленности в ВРП (в %)
Липецкая область	2008	-21.32	0.70	53.50
	2012	-13.68	0.90	32.10
Красноярский край	2008	-11.17	4.40	37.60
	2012	1.31	15.20	30.70
Астраханская область	2008	-9.50	2.00	30.70
	2012	12.69	19.20	5.40
Пермский край	2008	-3.49	12.10	34.90
	2012	2.48	17.30	31.40
Иркутская область	2008	-3.13	4.50	15.70
	2012	7.10	17.40	13.60
Республика Башкортостан	2008	-2.01	11.80	28.50
	2012	-13.30	2.90	37.20
Калининградская область	2008	-0.75	8.60	18.00
	2012	-6.99	4.80	22.20
Сахалинская область	2008	44.41	49.90	4.40
	2012	54.10	61.60	3.60

Как можно видеть из табл. 4, в некоторых из перечисленных регионов доля добывающей промышленности в ВРП в 2008 г. была меньше, чем в 2012 г. Это можно объяснить, в частности, низкими ценами на сырьевые товары в период кризиса 2008 г. и более быстрым снижением прибылей в добывающей промышленности и строительстве, чем в других секторах региона. Имеются и регионы, для которых наблюдается уменьшение доли добывающей промышленности в ВРП с одновременным увеличением доли обрабатывающей промышленности. Отдельно следует отметить, что в Республике Башкортостан доля добывающей промышленности в ВРП уменьшилась с 11.8% в 2008 г. до 2.9% в 2012 г., тогда как доля обрабатывающей промышленности в ВРП увеличилась с 28.5% в 2005 г. до 37.2% в 2012 г. Причиной этому, в частности, является интенсивное развитие машиностроения и нефтеперерабатывающих мощностей в республике.

В таблице 5 приведена корреляционная матрица значений индекса индустриализации для регионов РФ по каждому году за период 2008–2012 гг.

Значения индекса индустриализации для регионов РФ, рассчитанные за соответствующий год, также имеют значимую корреляцию, что также указывает на сравнительно стабильные значения индекса для большинства регионов РФ. Последнее может являться следствием слабой диверсификации экономики РФ в целом, поскольку анализируемый интервал включает в себя кризис 2008 г. Перечислим регионы, для которых значение индекса индустриализации максимально изменилось с 2008 по 2012 г. Среди них имеются те, у которых индекс индустриализации увеличивался (см. табл. 6).

Регионы, у которых индекс индустриализации максимально снижается, перечислены в табл. 7.

Из таблицы 7 видно, что только в Белгородской области уменьшение доли в ВРП обрабатывающей промышленности сопровождалось увеличением доли сельского хозяйства. Во всех остальных регионах, представленных в табл. 7, доля сельского хозяйства в ВРП менялась незначительно.

**Таблица 5.** Корреляционная матрица значений индекса индустриализации для регионов РФ с 2008 по 2012 г.

	2008 г.	2009 г.	2010 г.	2011 г.	2012 г.
2008 г.	1.00	0.96	0.96	0.95	0.93
2009 г.	0.96	1.00	0.98	0.97	0.95
2010 г.	0.96	0.98	1.00	0.99	0.96
2011 г.	0.95	0.97	0.99	1.00	0.98
2012 г.	0.93	0.95	0.96	0.98	1.00

**Таблица 6.** Регионы с максимальным увеличением индекса индустриализации за период с 2008 по 2012 г.

	Год	Индекс индустриализации	Доля сельского хозяйства в ВРП (%)	Доля добывающей промышленности в ВРП (%)	Доля обрабатывающей промышленности в ВРП (%)
Калужская область	2008	8.44	8.30	0.70	31.90
	2012	18.13	6.60	0.40	40.20
Республика Башкортостан	2008	10.82	8.30	11.80	28.50
	2012	17.13	5.60	2.90	37.20
Республика Марий Эл	2008	-1.69	15.30	0.10	23.40
	2012	6.51	13.40	0.10	29.80
Республика Адыгея	2008	-11.66	17.90	3.10	12.60
	2012	-6.34	13.40	1.20	14.80
Кабардино-Балкарская Республика	2008	-13.89	21.80	0.30	11.30
	2012	-6.18	17.40	0.10	18.30
Республика Ингушетия	2008	-28.25	8.70	2.10	1.60
	2012	-21.81	6.30	1.90	4.40

**Таблица 7.** Регионы с максимальным уменьшением индекса индустриализации за период с 2008 по 2012 г.

	Год	Индекс индустриализации	Доля сельского хозяйства в ВРП (%)	Доля добывающей промышленности в ВРП (%)	Доля обрабатывающей промышленности в ВРП (%)
Липецкая область	2008	27.59	7.80	0.70	53.50
	2012	10.63	9.30	0.90	32.10
Вологодская область	2008	25.52	5.30	0.00	50.50
	2012	14.44	3.90	0.10	36.10
Кемеровская область	2008	10.69	3.40	28.40	19.90
	2012	6.41	2.90	26.90	15.80
Астраханская область	2008	8.94	7.10	2.70	30.70
	2012	-8.22	7.40	19.20	5.40
Белгородская область	2008	7.33	12.10	18.50	22.90
	2012	1.30	17.30	16.00	18.10

Согласно выводам, полученным выше, для компонентного анализа отраслевой структуры ВРП можно использовать данные любого года за период 2008–2012 гг., получая схожие нагрузки на главные компоненты. В качестве такого года возьмем 2012 г. В таблице 8 приведены факторные нагрузки для отраслевой структуры ВРП за 2012 г.

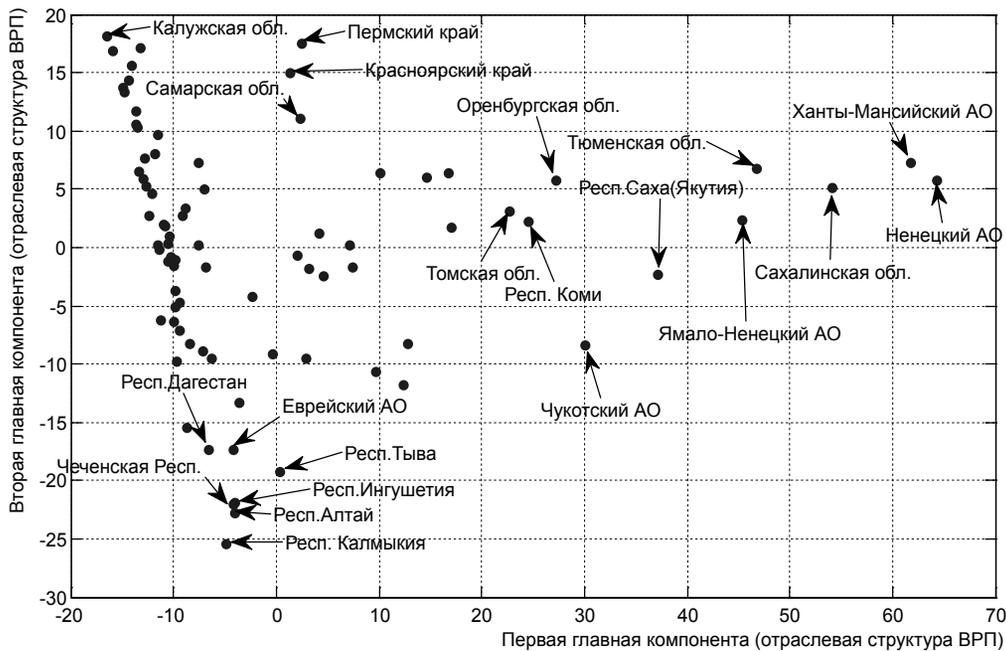
**Таблица 8.** Нагрузки на главные компоненты для отраслевой структуры ВРП за 2012 г.

	Номер главной компоненты					
	1	2	3	4	5	6
Сельское хозяйство	-0.12	-0.29	-0.54	0.72	0.15	0.06
Добыча полезных ископаемых	0.91	0.24	-0.07	0.11	-0.04	-0.11
Обрабатывающие производства	-0.35	0.81	-0.31	-0.04	-0.07	-0.16
Строительство	0.00	-0.17	-0.16	-0.48	0.75	-0.01
Оптовая и розничная торговля	-0.16	-0.05	0.63	0.33	0.13	-0.57
Финансовая деятельность	0.00	0.00	0.01	0.00	-0.02	0.02
Операции с недвижимым имуществом	-0.07	0.11	0.39	0.13	-0.04	0.78
Государственное управление и обеспечение военной безопасности	-0.06	-0.36	-0.17	-0.31	-0.55	-0.16
Образование	-0.03	-0.12	-0.05	-0.08	-0.21	-0.06
Здравоохранение и предоставление социальных услуг	-0.04	-0.09	-0.02	-0.10	-0.23	0.05
Предоставление прочих коммунальных, социальных и персональных услуг	-0.01	-0.01	0.02	-0.02	-0.03	0.01
Объясненная дисперсия (%)	64.98	20.14	6.29	3.46	2.81	1.57

Из первого столбца таблицы видно, что максимальные нагрузки на первую главную компоненту (индекс промышленной специализации) приходятся на два показателя — долю добычи полезных и долю обрабатывающих производств в ВРП. При этом объясненная дисперсия, приходящаяся на эту компоненту, равна 64.98%. Нагрузка доли добычи полезных ископаемых в ВРП на первую главную компоненту имеет положительный знак, а нагрузка доли обрабатывающих производств в ВРП — отрицательный. Для второй главной компоненты (индекс индустриализации) максимальные нагрузки приходятся на показатели *доля обрабатывающих и добывающих производств в ВРП*, которой соответствует положительная нагрузка, а также *доля в ВРП сельского хозяйства, доля государственного управления и обеспечение военной безопасности*, которым отвечает отрицательная нагрузка. Объясненная дисперсия для второй главной компоненты равна 20.14%.

На рисунке 3 точками изображены субъекты РФ на плоскости (горизонтальная ось — индекс промышленной специализации, вертикальная ось — индекс индустриализации).

Из рисунка 3 видно, что точки, которым соответствуют минимальные значения индекса промышленной специализации и максимальные значения индекса индустриализации, формируют сгусток в виде полосы, а вся совокупность точек образует треугольник. Регионы, которым соответствуют наибольшие значения индекса промышленной специализации, специализируются на добыче полезных ископаемых, главным образом нефти и газа. Среди таких регионов следует выделить Ханты-Мансийский АО, Ненецкий АО, Сахалинскую область, Тюменскую область, Ямало-Ненецкий АО. Регионам с минимальными значениями первой главной компоненты отвечает специализация в обрабатывающей промышленности. В их число входят Калужская область, Омская область, Республика Башкортостан.



С. А. Айвазян, М. Ю. Афанасьев, А. В. Кудров

**Рис. 3.** Индексы индустриализации и промышленной специализации субъектов РФ по данным за 2012 г.

Регионы, имеющие наибольшие значения индекса индустриализации, включают в себя перечисленные выше регионы, для которых абсолютное значение индекса промышленной специализации принимает сравнительно большие значения. Тогда как регионы с минимальными значениями индекса индустриализации являются сельскохозяйственными и (или) дотационными, в которых уровень развития промышленности крайне низок. Среди таких регионов следует выделить Калмыкию, Алтай, Ингушетию, Чечню, Тыву.

Отдельно следует выделить группу регионов, которым соответствует высокий уровень индекса индустриализации и локализация индекса промышленной специализации вокруг нуля. Такие регионы характеризуются развитостью как обрабатывающей, так и добывающей промышленности. В число таких регионов входит Пермский край, Красноярский край и Самарская область.

### 5. Эмпирический анализ: модель граничного производственного потенциала с локально-постоянными коэффициентами для ВРП регионов РФ

В модели граничного производственного потенциала с локально-постоянными коэффициентами для ВРП регионов РФ, см. раздел 1, используется переменная  $q_i(t)$  — вектор-характеристика структуры экономики для  $i$ -го региона в год  $t$ . В качестве значений этой переменной будет использоваться индикаторы отраслевой структуры ВРП:

- индекс промышленной специализации;
- индекс индустриализации для  $i$ -го региона в год  $t$  (см. раздел 4).

Вектор-переменная  $q_i(t)$  при этом будет иметь размерность  $1 \times 2$ .

Как видно из результатов раздела 4 (см. табл. 1), данным по отраслевой структуре ВРП регионов РФ за 2008 г. соответствует другая информативность, чем данным за период с 2009 по 2012 г. С одной стороны, это указывает на то, что в 2008 г. были другие условия функционирования экономики, а с другой стороны, говорит о том, что данные о структуре ВРП за период с 2009–2012 гг. можно считать однородными. Поэтому из совокупной выборки был исключен 2008 г. и брались данные только с 2009 по 2012 г.

**Регионы со специализацией в добывающей промышленности.** Выберем в качестве граничной точки (см. описание алгоритма в разделе 2) вектор  $q_i(t)$ , которому соответствует максимальное значение индекса промышленной специализации. Эта точка соответствует региону с максимальной долей добывающей промышленности в структуре ВРП — Тюменской области. Применяв алгоритм, описанный в разделе 2, получим график средних значений логарифма функции правдоподобия, упорядоченных по числу добавленных объектов, ближайших относительно значений  $q_i(\cdot)$ -вектора (рис. 4). На этом графике по оси  $X$  стоит число объектов, ближайших к региону с максимальным значением индекса промышленной специализации и минимальным значением индекса индустриализации; а по оси  $Y$  — отношение значения функции правдоподобия к числу объектов, по которым она оценивалась.



Рис. 4. Средние значения функции правдоподобия (добывающая промышленность)

Формируя «базовую» выборку (см. раздел 2), оцениваем параметры модели (1) при  $m = 1$ . Затем строим доверительный интервал для средних значений логарифма функции правдоподобия, и с его помощью находим итоговую группу регионов со специализацией в добывающей промышленности, см. табл. 9.

Оценки параметров модели (1) для «базовой» выборки, отвечающей специализации в добывающей промышленности, представлены в табл. 10.

Следует отметить, что эластичности по труду и по капиталу ( $\beta_1$  и  $\beta_2$  соответственно) являются статистически значимыми, а значение коэффициента  $R^2$  весьма близко к единице.

**Регионы со специализацией в обрабатывающей промышленности.** Возьмем в качестве граничной точки (см. описание алгоритма в разделе 2) вектор  $q_i(t)$ , которому со-

**Таблица 9.** Список регионов со специализацией в добывающей промышленности

Архангельская область	Республика Коми
Тюменская область	Кемеровская область
Сахалинская область	Томская область
Республика Саха (Якутия)	Удмуртская Республика
Чукотский автономный округ	Республика Татарстан
Оренбургская область	

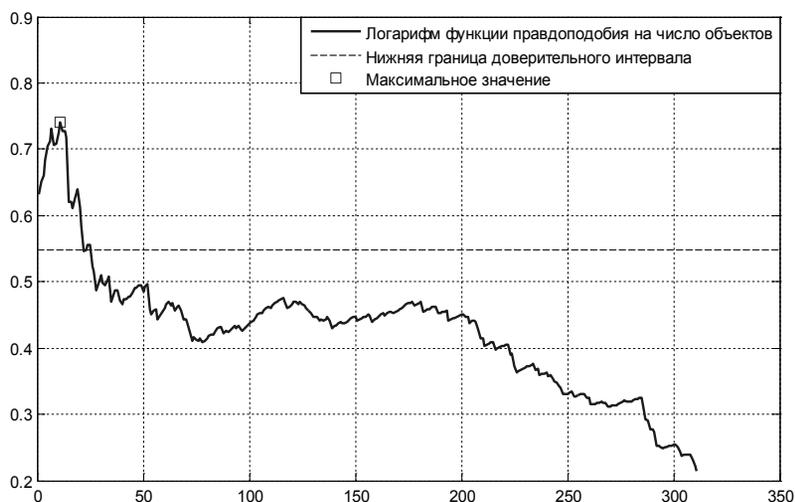
**Таблица 10.** Оценки модели для группы регионов со специализацией в добывающей промышленности

	Параметр				
	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma_v^2$	$\sigma_u^2$
Оценка	1.84*** (0.33)	0.11*** (0.03)	0.75*** (0.03)	0.04	$4.19 \cdot 10^{-8}$
$R^2$	0.97				

*Примечание.* В скобках — стандартные ошибки, \*\*\* — значимость на 1%-ном уровне.

ответствует минимальное значение индекса индустриализации. Эта граничная точка соответствует региону с максимальной долей обрабатывающей промышленности в структуре ВРП — Калужской области.

На рисунке 5 по оси  $X$  — число ближайших объектов к региону с максимальным значением индекса промышленной специализации и минимальным значением индекса индустриализации; по оси  $Y$  — значение функции правдоподобия, поделенное на число объектов, по которым она оценивалась.

**Рис. 5.** Средние значения функции правдоподобия (обрабатывающая промышленность)

В результате формирования «базовой» выборки (см. раздел 2), оценки по ней параметров модели (1) при  $m = 1$  и построения доверительного интервала для средних значений логарифма функции правдоподобия образуем итоговую группу регионов со специализацией в обрабатывающей промышленности, см. табл. 11.

**Таблица 11.** Список регионов со специализацией в обрабатывающей промышленности

Калужская область	Республика Башкортостан
Омская область	Новгородская область
Липецкая область	Нижегородская область
Челябинская область	Владимирская область
Вологодская область	Свердловская область
Тульская область	г. Санкт-Петербург

Отметим, что только в 2011–2012 гг. Республика Башкортостан попадает в список регионов со специализацией в обрабатывающей промышленности, что подтверждается завершением в эти годы проектов по запуску новых обрабатывающих производств в регионе.

Оценки параметров «базовой» выборки, отвечающей регионам со специализацией в обрабатывающей промышленности, представлены в табл. 12.

**Таблица 12.** Оценки модели для группы регионов со специализацией в обрабатывающей промышленности

	Параметр				
	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma_v^2$	$\sigma_u^2$
Оценка	0.93 (0.63)	0.45*** (0.10)	0.65*** (0.09)	0.02	$5.78 \cdot 10^{-8}$
$R^2$	0.96				

*Примечание.* В скобках — стандартные ошибки, \*\*\* — значимость на 1%-ном уровне.

Как видно из таблицы 12, эластичности по труду и по капиталу ( $\beta_1$  и  $\beta_2$  соответственно) являются статистически значимыми, значение коэффициента  $R^2$  близко к единице.

Отметим, что значения параметров модели для групп регионов со специализацией в добывающей промышленности и в обрабатывающей промышленности отличаются. Для последней характерна большая эластичность по труду и меньшая эластичность по капиталу.

**Регионы с формирующейся экономикой (подгруппа 1).** Рассмотрим в качестве граничной точки (см. описание алгоритма в разделе 2) вектор  $q_i(t)$ , которому соответствует минимальное значение индекса промышленной специализации. Эта граничная точка соответствует Республике Калмыкия.

Согласно описанному ранее алгоритму, сформируем «базовую» выборку и построим доверительный интервал для средних значений логарифма функции правдоподобия. На рисунке 6 по оси  $X$  представлено число объектов, ближайших к региону с минимальным значением индекса индустриализации, а по оси  $Y$  — значение функции правдоподобия, поделенное на число объектов, по которым она оценивалась.

В результате получим итоговую группу регионов с формирующейся экономикой, см. табл. 13.

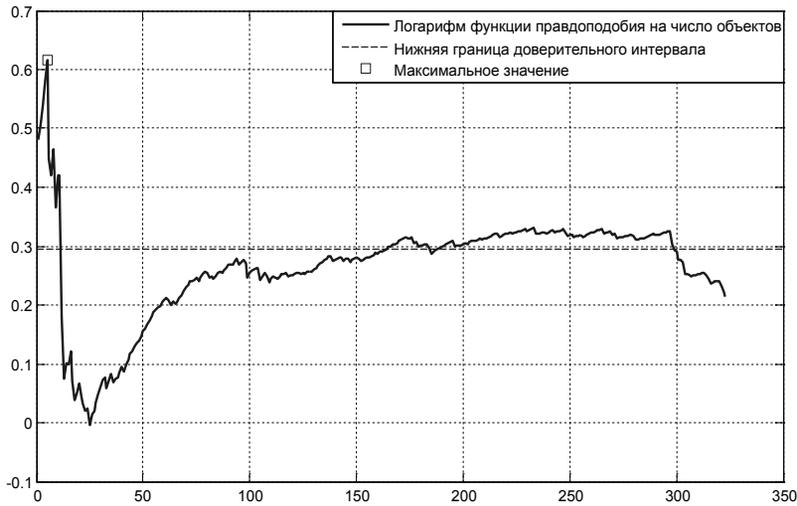


Рис. 6. Средние значения функции правдоподобия (формирующаяся экономика, подгруппа 1)

Таблица 13. Список регионов с формирующейся экономикой (подгруппа 1)

Республика Калмыкия	Еврейская автономная область
Республика Алтай	Республика Ингушетия
Чеченская Республика	Республика Дагестан

Оценки параметров «базовой» выборки, отвечающей регионам с формирующейся экономикой, представлены ниже в табл. 14.

Таблица 14. Оценки модели для группы регионов с формирующейся экономикой (подгруппа 1)

	$\beta_0$	$\beta_1$	Параметр $\beta_2$	$\sigma_v^2$	$\sigma_u^2$
Оценка	5.44*** (0.99)	0.62*** (0.14)	0.16 (0.13)	0.03	$2.67 \cdot 10^{-8}$
$R^2$	0.90				

Примечание. В скобках — стандартные ошибки, \*\*\* — значимость на 1%-ном уровне.

Как видно из таблицы 14, эластичность по капиталу ( $\beta_2$ ) является статистически незначимой.

**Регионы с формирующейся экономикой (подгруппа 2).** Аналогично удастся выделить еще одну группу регионов с формирующейся экономикой, среди которых, главным образом, регионы Дальнего Востока и Сибири, см. рис. 7 и табл. 15. На этом рисунке по оси  $X$  откладывается число ближайших объектов к региону с минимальным значением индекса индустриализации (в выборке с исключенными регионами из подгруппы 1); по оси  $Y$  — значение функции правдоподобия, поделенное на число объектов, по которым она оценивалась.

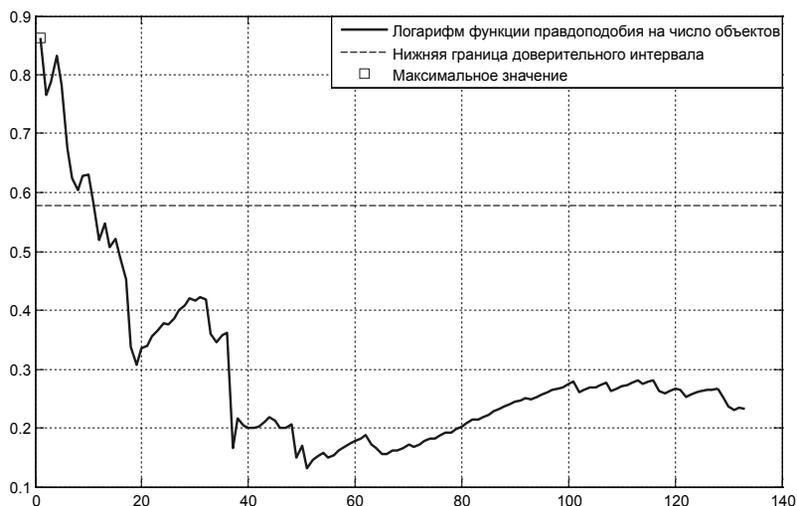


Рис. 7. Средние значения функции правдоподобия (формирующаяся экономика, подгруппа 2)

Таблица 15. Список регионов с формирующейся экономикой (подгруппа 2)

Республика Тыва	Камчатский край
Магаданская область	Амурская область
Забайкальский край	Республика Хакасия
Мурманская область	

В таблице 16 представлены оценки параметров «базовой» выборки, отвечающей регионам с формирующейся экономикой.

Таблица 16. Оценки модели для группы регионов с формирующейся экономикой (подгруппа 2)

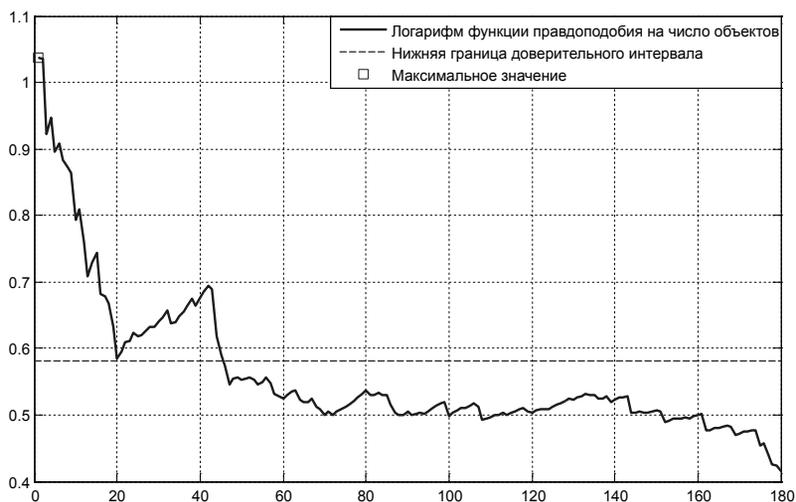
	Параметр				
	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma_v^2$	$\sigma_u^2$
Оценка	3.17*** (0.31)	0.06 (0.08)	0.65*** (0.05)	0.01	$1.2 \cdot 10^{-8}$
$R^2$	0.97				

Примечание. В скобках — стандартные ошибки, \*\*\* — значимость на 1%-ном уровне.

Следует отметить, что эластичность на труд является незначимой. Это является отражением того факта, что регионы из этой группы являются слабо заселенными.

**Регионы со специализацией в сельском хозяйстве и рыболовстве.** Из общей совокупности исключим наблюдения, отвечающие регионам с формирующейся экономикой, и возьмем наблюдение, которому соответствует минимальное значение индекса промышленной специализации. Эта граничная точка соответствует Кабардино-Балкарской Республике (см. рис. 8).

На рисунке 8 представлено: по оси  $X$  — число ближайших объектов к региону с минимальным значением индекса индустриализации в выборке с исключенными регионами с формирующейся экономикой, а по оси  $Y$  — значение функции правдоподобия, поделенное на число объектов, по которым она оценивалась.



**Рис. 8.** Средние значения функции правдоподобия (сельское хозяйство и рыболовство)

Список регионов со специализацией в сельском хозяйстве и рыболовстве представлен в табл. 17. Оценки параметров модели (1) для группы регионов со специализацией в сельском хозяйстве и рыболовстве приведены в табл. 18.

**Таблица 17.** Список регионов со специализацией в сельском хозяйстве и рыболовстве

Карачаево-Черкесская Республика	Ростовская область
Приморский край	Орловская область
Кабардино-Балкарская Республика	Воронежская область
Краснодарский край	Псковская область
Тамбовская область	Брянская область
Ставропольский край	Алтайский край
Республика Адыгея	Курганская область
Республика Северная Осетия–Алания	

**Таблица 18.** Оценки параметров модели (1) при  $m = 1$  для группы регионов со специализацией в сельском хозяйстве и рыболовстве

	Параметр				
	$\beta_0$	$\beta_1$	$\beta_2$	$\sigma_v^2$	$\sigma_u^2$
Оценка	1.06** (0.40)	0.51*** (0.08)	0.60*** (0.07)	0.01	$7.9 \cdot 10^{-11}$
$R^2$	0.98				

*Примечание.* В скобках — стандартные ошибки; \*\*, \*\*\* — значимость на 5 и 1%-ном уровне соответственно.

**Регионы со специализацией смешанного типа.** Для всех нераспределенных по группам регионов расчеты, аналогичные прежним, приводят к выводу, что дополнительных групп выделить не удастся (см. рис. 9 и табл. 19, 20). На рисунке 9 по горизонтали откладывается число ближайших объектов к региону с минимальным значением индекса индустриализации в выборке с исключенными регионами, уже распределенными по группам; а по вертикальной оси — значения функции правдоподобия, поделенное на число объектов, по которым она оценивалась.



**Рис. 9.** Средние значения функции правдоподобия (специализация смешанного типа)

**Таблица 19.** Список регионов со специализацией смешанного типа

Астраханская область	Белгородская область
Волгоградская область	г. Москва
Ивановская область	Иркутская область
Калининградская область	Кировская область
Костромская область	Красноярский край
Курская область	Ленинградская область
Московская область	Новосибирская область
Пензенская область	Пермский край
Республика Бурятия	Республика Карелия
Республика Марий Эл	Республика Мордовия
Рязанская область	Самарская область
Саратовская область	Смоленская область
Тверская область	Ульяновская область
Хабаровский край	Чувашская Республика
Ярославская область	

**Таблица 20.** Оценки параметров модели (1) при  $m = 1$  для группы регионов со специализацией смешанного типа

	Параметр			$\sigma_v^2$	$\sigma_u^2$
	$\beta_0$	$\beta_1$	$\beta_2$		
Оценка	-1.49*** (0.41)	0.22** (0.10)	0.93*** (0.07)	0.03	$3.2 \cdot 10^{-8}$
$R^2$	0.96				

Примечание. В скобках — стандартные ошибки; \*\*, \*\*\* — значимость на 5 и 1%-ном уровне соответственно.

Заметим, что эластичность по труду является незначимой. А значение коэффициента  $R^2$  мало отличается от единицы.

## 6. Заключение

Основным результатом данного исследования является разбиение совокупности регионов на группы, для каждой из которых характерна своя производственная функция и специализация экономики.

Проведенные расчеты подтвердили наличие большей чувствительности ВРП к изменению стоимости основных фондов для групп регионов со специализацией в добывающей промышленности, специализацией смешанного типа и регионов с формирующейся экономикой (подгруппа 2), чем для групп регионов со специализацией в обрабатывающей промышленности, в сельском хозяйстве и регионов с формирующейся экономикой (подгруппа 1). С другой стороны, последние имеют большую чувствительность к изменению численности занятых.

Представляется, что разработанная методология может быть одним из условий выработки процедур прогнозирования и планирования вариантов устойчивого социально-экономического развития.

Предложенная методика анализа годится и для более широкого исследования (при наличии соответствующих данных) в задаче идентификации типов производственных функций в зависимости от структуры занятости, географии и т. д.

## Список литературы

Айвазян С. А., Афанасьев М. Ю., Руденко В. А. (2014). Оценка эффективности регионов РФ на основе модели производственного потенциала с характеристиками готовности к инновациям. *Экономика и математические методы*, 50 (4), 34–70.

Айвазян С., Бухштабер В., Енюков И., Мешалкин Л. (1989). *Прикладная статистика. Классификация и снижение размерности*. М.: Финансы и статистика.

Зубаревич Н. В., Горина Е. (2015). Социальные расходы в России: федеральный и региональные бюджеты. М.: НИУ ВШЭ.

Aigner D. J., Lovell C. A. K., Schmidt P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21–37.

Bera A., Sharma S. (1999). Estimating production uncertainty in stochastic frontier production function models. *Journal of Productivity Analysis*, 12, 187–210.

Farrell M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society (A, general)*, 120, 253–281.

Kumbhakar S. C., Lovell K. C. A. (2000). *Stochastic frontier analysis*. Cambridge: Cambridge University Press.

Meeusen W., van den Broeck J. (1977). Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review*, 8, 435–444.

Tsionas E. G. (2006). Inference in dynamic stochastic frontier models. *Journal of Applied Econometrics*, 21, 669–676.

Tsionas E. G., Kumbhakar S. C. (2004). Markov switching stochastic frontier model. *Econometrics Journal*, 7 (2), 398–425.

Van den Broeck J., Koop J., Osiewalski J., Steel M. F. J. (1994). Stochastic frontier models: A Bayesian perspective. *Journal of Econometrics*, 61 (2), 273–303.

Поступила в редакцию 07.03.2016;  
принята в печать 17.03.2016.

## Приложение

**1. Логарифм функции правдоподобия** модели (1) граничного производственного потенциала с постоянными коэффициентами при наличии только одной группы ( $m = 1$ ) для выборки размера  $N$  имеет вид:

$$l(\varepsilon(\alpha_0, \alpha_1, \alpha_2) | \lambda, \sigma, \alpha_0, \alpha_1, \alpha_2, N) = N \ln \left( \frac{\sqrt{2}}{\sqrt{\pi}} \right) + N \ln \left( \frac{1}{\sigma} \right) + \sum_{i=1}^N \ln (1 - \Phi(\varepsilon_i \lambda \sigma^{-1})) - \frac{1}{2\sigma^2} \sum_{i=1}^N \varepsilon_i^2,$$

где  $\sigma^2 = \sigma_u^2 + \sigma_v^2$ ,  $\lambda = \sigma_u / \sigma_v$ ;  $\varepsilon(\alpha_0, \alpha_1, \alpha_2) = \ln P - \alpha_0 - \alpha_1 \ln L - \alpha_2 \ln K$ .

Таким образом, выражение для среднего значения логарифма функции правдоподобия равно:

$$El(\varepsilon(\alpha_0, \alpha_1, \alpha_2) | \lambda, \sigma, \alpha_0, \alpha_1, \alpha_2, N) = \ln \left( \frac{\sqrt{2}}{\sqrt{\pi}} \right) + \ln \left( \frac{1}{\sigma} \right) + \frac{1}{N} \sum_{i=1}^N \ln (1 - \Phi(\varepsilon_i \lambda \sigma^{-1})) - \frac{1}{2\sigma^2 N} \sum_{i=1}^N \varepsilon_i^2.$$

## 2. Алгоритм формирования однородной группы и оценки параметров.

*Шаг 1.* Формирование базы из  $M$  объектов ( $M = 15$ ):

*Шаг 1.1.* Выбираем граничную точку области возможных значений для  $q_i(t_0)$ , одна из координат которой соответствует экстремальному из возможных для нее значений (максимуму или минимуму). Пусть эта точка имеет номер  $i_0$ .

*Шаг 1.2.* Находим точку, ближайшую к найденной на шаге 1.1:

$$(i_1, t_1) = \operatorname{argmin}_{D \setminus \{(i_0, t_0)\}} d^2(q_{i_0}(t_0), q_i(t)), \text{ где } D = \{(i, t) : 1 \leq i \leq I, 1 \leq t \leq T\}.$$

*Шаг 1.3.* Находим ближайшую точку к классу, составленному из двух объектов

$$\{q_{i_0}(t_0), q_{i_1}(t_1)\}; (i_2, t_2) = \operatorname{argmin}_{D \setminus \{(i_0, t_0), (i_1, t_1)\}} \rho_{\tau}^{(K)}(\{q_{i_0}(t_0), q_{i_1}(t_1)\}, q_i(t)).$$

*Шаг 1.4.* Повторяем процедуру добавления ближайшего к классу объекта  $M$  три раза. В результате получаем класс  $G_{M_3}$ , составленный из  $M$  объектов.

*Шаг 2.* Используя метод максимума правдоподобия, оцениваем параметры модели граничного производственного потенциала с постоянными коэффициентами для базовых  $M$  объектов, найденных на шаге 1:

$$\operatorname{argmax} l(\ln P - \gamma_d(0) - \gamma_d(1)\ln L - \gamma_d(2)\ln K \mid \gamma_d(0), \gamma_d(1), \gamma_d(2), \lambda, \sigma, M),$$

где  $\lambda = \delta_d(2) / \delta_d(1)$ ,  $\sigma^2 = \delta_d^2(1) + \delta_d^2(2)$ ;  $d$  — номер однородной группы ( $d = 1$ );  $l(\cdot)$  — логарифм функции правдоподобия для рассматриваемых  $M$  объектов (см. п. 1 Приложения).

В качестве начальных значений параметров в процедуре максимизации логарифма функции правдоподобия использовались оценки, полученные методом моментов. Обозначим максимальное значение логарифма функции правдоподобия для базовых  $M$  объектов через  $l^M(1, M)$ .

*Шаг 3.* Находим ближайший объект к классу  $G_M$ :

$$(i_{M+1}, t_{M+1}) = \operatorname{argmin}_{D \setminus \{(i,t): q_i(t) \in G_M\}} \rho_\tau^{(K)}(G_M, q_i(t)).$$

Оцениваем параметры модели граничного производственного потенциала с постоянными коэффициентами для найденных  $M+1$  объектов. Значение логарифма функции правдоподобия обозначим через  $l^{M+j}(1, M+j)$ ,  $j=1$ .

*Шаг 4.* Итеративно повторяем шаг 3 до тех пор, пока закончатся объекты в совокупной выборке.

*Шаг 5.* Находим шаг итерации  $j_0$ , для которого  $j_0 = \operatorname{argmax}_{j=1, \dots, N-M} \frac{l^{M+j}(1, M+j)}{M+j}$ .

*Шаг 6.* Из модели, полученной на шаге итерации  $j_0$  (см. шаг 5), смоделируем  $H = 10000$  раз последовательность из  $(M + j_0)$  ошибок и построим 90%-ный доверительный интервал

$$[r^d, r^u] \text{ для логарифма функции правдоподобия } \frac{l^{M+j_0}(1, M+j_0)}{M+j_0}.$$

*Шаг 7.* К множеству  $\{1, \dots, j_0\}$  добавим  $\{j_0, \dots, j_0 + k_0\}$ ,

где  $k_0 = \min \{k : k \geq 0, l^{M+k}(1, M+k) < (M+k)r^d\}$ .

*Шаг 8.* Исключаем из имеющейся совокупности множество объектов из группы, определенной на шаге 7, и для оставшихся объектов повторяем процедуру, начиная с шага 1.

Параметр  $m$ , т.е. число однородных групп, определяется, когда все объекты будут распределены по соответствующим группам однородности.

Aivazian S., Afanasiev M., Kudrov A. Clustering methodology of the Russian Federation regions with account of sectoral structure of GRP. *Applied Econometrics*, 2016, 41, pp. 24–46.

### Sergei Aivazian

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

### Mikhail Afanasiev

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

### Aleksander Kudrov

Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow, Russia; kovlal@inbox.ru

## Clustering methodology of the Russian Federation regions with account of sectoral structure of GRP

The methodology to identify similar groups of Russian regions is presented, each of which has its own model of productive capacity, which determines the dependence of the GRP from the fixed assets and number of employees. It is carried out to test the hypothesis that the parameters of the functions describing the production potential of the regions included in the different groups, different by reason of the sectoral structure of GRP. The procedure for constructing the auxiliary integral indicators based on component analysis is offered, it is reflected the specialization of regions. Developed and tested on data for the period 2009–2013 algorithm of forming a similar group of regions. It is revealed the presence of a greater sensitivity to changes in GRP value of fixed assets for the group of regions with a specialization in mining and mixed type specialty than for groups of regions specializing in manufacturing and agriculture. At the same time, last two groups of regions and the group of regions with emerging economies are more sensitive to changes in the number of employees. The developed methodology is an element of forecasting procedures and planning options for sustainable socio-economic development of regions of the Russian Federation.

**Keywords:** regional economy; production potential; econometric modeling; hypothesis testing.

**JEL classification:** C12; C51; R15.

### References

- Aivazian S. A., Afanasiev M. Yu., Rudenko V. A. (2014). Ocenka jeffektivnosti regionov RF na osnove modeli proizvodstvennogo potenciala s karakteristikami gotovnosti k innovacijam. *Ekonomika i Matematicheskie Metody*, 50 (4), 34–70 (in Russian).
- Aivazian S., Buhshtaber V., Enjukov I., Meshalkin L. (1989). *Prikladnaja statistika. Klassifikacija i snizhenie razmernosti*. M.: Finansy i statistika (in Russian).
- Zubarevich N. V., Gorina E. (2015). *Social'nye rashody v Rossii: federal'nyj i regional'nye bjudzhet*. M.: NRU HSE (in Russian).
- Aigner D. J., Lovell C. A. K., Schmidt P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21–37.
- Bera A., Sharma S. (1999). Estimating production uncertainty in stochastic frontier production function models. *Journal of Productivity Analysis*, 12, 187–210.
- Farrell M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society (A, general)*, 120, 253–281.
- Kumbhakar S. C., Lovell K. C. A. (2000). *Stochastic frontier analysis*. Cambridge: Cambridge University Press.
- Meeusen W., van den Broeck J. (1977). Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review*, 8, 435–444.
- Tsionas E. G. (2006). Inference in dynamic stochastic frontier models. *Journal of Applied Econometrics*, 21, 669–676.
- Tsionas E. G., Kumbhakar S. C. (2004). Markov switching stochastic frontier model. *Econometrics Journal*, 7 (2), 398–425.
- Van den Broeck J., Koop J., Osiewalski J., Steel M. F. J. (1994). Stochastic frontier models: A Bayesian perspective. *Journal of Econometrics*, 61 (2), 273–303.

Received 07.03.2016; accepted 17.03.2016.