СВЯЗЬ ИНТЕГРАЛЬНОГО ИНДИКАТОРА УСЛОВИЙ ЖИЗНИ С ЭКЗОГЕННЫМИ ПЕРЕМЕННЫМИ РЕГИОНАЛЬНОЙ ПОЛИТИКИ

В. С. Степанов

Центральный экономико-математический институт РАН, Москва, Россия
E-mail: _Stepanov@cemi.rssi.ru

В работе выполняется эконометрический анализ региональных данных. Сначала по набору показателей 2017 года на основе компонентного анализа рассчитывается интегральный индикатор условий жизни в регионах (субъектах РФ). Затем по другой таблице данных этого же года оценивается классическая линейная регрессионная модель, в которой зависимой (эндогенной) переменной является индикатор условий жизни населения региона, а экзогенными – пять показателей, зависящих прямо или косвенно от решений территориальных органов власти. Модель оказалась достаточно точной почти в половине регионов РФ (как на материале «обучения», так и на независимой выборке за 2016 год). Она может быть полезной при имитации сценариев их развития и выявлении того, как повлияют изменения экзогенных переменных в регионе на поведение эндогенной.

THE ASSOSIATION BETWEEN INTERGRAL INDICATOR OF LIVING CONDITIONS AND EXOGENOUS VARIABLES OF REGIONAL POLICY

V. S. Stepanov

This paper performs an econometric analysis of regional data. First, an integral indicator of living conditions in the “regions” (constituent entities of the Russian Federation) is calculated. The regions are described by eleven variables of 2017 and the integral indicator is calculated through the principal component analysis approach. Then, according to another data of the same year, a classical linear regression model is estimated, in which the endogenous variable is the integral indicator and exogenous ones are five variables that depend on decisions of territorial authorities. The model turned out to be quite accurate in almost half of the Russian regions: both on the basis of the “learning data” and on another data for 2016. It can be useful in simulating scenarios of region development and identifying how some changes in the exogenous variables will affect endogenous behavior.

Вопрос оценки качества жизни населения возникает при социально-экономическом планировании, разработке стратегии развития территорий. В [1, 2] была развита методология построения интегральных индикаторов (ИИ) качества жизни. Она применялась к данным по странам, регионам и др. [1, 3]. Ниже она прикладывается к оценке уровня благосостояния (условий жизни) населения в 76 регионах РФ по базе данных из [4]; шесть исключённых регионов перечислены в [5]. В результате строится измеритель Y условий жизни, который может принимать вещественные значения от 0 до 10. Этим измерителем является ИИ, который формируется как нелинейная свёртка набора показателей из [1, 5], что описывают условия жизни.

Целью работы является построение регрессионной модели, связываю-

щей эндогенную переменную Y с пятью экзогенными, которые прямо или косвенно зависят от решений федеральных и региональных органов власти.

Ниже сначала описывается, как строилась переменная Y в заданный год t ; во второй части приводятся результаты по модели. Перечень показателей, измеряющих условия жизни, приводится в [1,Таб.П2.6] и в чуть расширенном списке $x_1, \ldots, x_{11}$ – в [5, Таб.5]. Содержательно, а также с учётом корреляций, переменные (показатели) сгруппированы в [1,Таб.2.6], [5] в три блока: “доходы-расходы” (душевые показатели ВРП и товарооборота, автомобилизация и др.), “жилищные условия, инфраструктура”, “дифференциация по доходам” (коэффициент фондов *Кф*). В настоящей работе два показателя из [5] скорректированы с учётом паритета покупательной способности доходов [7]: ВРП и розничный товарооборот в сумме с объемом платных услуг.

Первоначально выполнялась унификация шкалы показателя x согласно [1]; формулы (2.2) и (2.3), кроме случая x = *Кф*. Они описывают связь переменной x с условиями жизни: "чем больше значение x, тем лучше” или, соответственно – наоборот. Переменная *Кф* преобразуется по формуле $u = 10 \cdot \left[1 - \left| x - x_0 \right| \Big/ \max\left\{ \left(x^* - x_0 \right), \left(x_0 - x_{\&} \right) \right\} \right]$ с предварительным выбором оптимального значения $x_0 =$ 9.5 [1,(2.4)], где $x -$ значение *Кф* для рассматриваемого региона, $x_* -$ минимальное значение x среди всех в год t, x^* – аналогичное максимальное значение. Формулы (2.2), (2.3) порождаются из (2.4), когда $x_0 = x_*$, либо $x_0 = x^*$. После унификации самый лучший по x регион имеет значение $u = 10$, а слабейший – 0 . Затем, после объединения переменных в блоки, исходя из коэффициентов корреляций, на такие же блоки разбивается и ковариационная матрица $\mathbf{S}$ всех 11-ти переменных. Иначе говоря, после перестановки в списке переменных она имеет структуру блочно-диагонального типа. Затем на уровне каждой из блочных матриц $\mathbf{S}^{(k)}, \mathrm{k} = 1, 2, 3$ рассчитываются главные компоненты [6]; в результате в качестве ИИ для k-го блока берётся чуть изменённая первая главная компонента $y^{(k)}$ [1] (индекс региона i=1,…, 76 для простоты здесь опущен). Расчёт делается через решение задачи на собственные значения $\mathbf{S}^{(k)}$ и дальнейшим использованием лишь одного собственного вектора, что соответствует максимальному собственному числу $\lambda_1^{(\mathrm{k})}$ этой матрицы. В итоге удаётся снизить размерность и вместо всех переменных, входящих в состав k-го блока получить единственную: $y^{(k)}$ – первую главную компоненту, что является ”линейной свёрткой” переменных блока. Для рассматриваемого случая, унифицированные данные из 11- мерного пространства проектируются в 3-D. При этом, за счёт предложенной в [1] модификации при расчёте $y^{(k)}$, они преобразуются в точки, лежащие внутри и на поверхности куба с рёбрами длины 10; это расписано с иллюстрацией в [5]. Кроме того, вводится “эталонный регион”, у которого все x равны 10 и поэтому эталонная точка лежит в

вершине ***10***=(10,10,10)^{T}. Затем к точке $\boldsymbol{y_i}=(y_i^{(1)}, y_i^{(2)}, y_i^{(3)})^{T}$ *i*-го региона из 3-D применяется нелинейное преобразование; оно задается через расчёт квадрата расстояния от $\boldsymbol{y_i}$ до эталонной точки ***10*** через вычисление суммы $\rho_i^2 = \sum_{k=1}^{3} q_k \cdot \left(10 - y_i^{(k)}\right)^2$, взятой с "весами" q_k из (0,1) ; последние задаются согласно [1] пропорционально дисперсии $y^{(k)}$. Величина Y искомого ИИ условий жизни в *i*-том регионе определяется как разность $Y_i = 10 - \rho_i$. Для вышеперечисленных блоков веса в 2017 году оказались равны $q_1 = 0.464$, $q_2 =$ 0.362, $q_3 = 0.174$, так что унифицированная величина *Кф* из третьего блока существенно влияет на *Y*. После нахождения значений *Y* для регионов и их ранжирования, последние можно упорядочить по убыванию *Y* . Если всё аналогично повторить за несколько лет, то можно получить эволюцию блочных ИИ по годам, а также – для *Y* и её рангов [4, 5], а затем найти лидеров и аутсайдеров, провести дальнейший анализ [1,3]. В частности, в [8] было предложено использовать величину *Y* как критерий эффективности региональной социально-экономической политики.

Переходя к этапу построению модели, в развитие этой идеи нами были выбраны пять факторных переменных, от которых зависит *Y* (синхронно по *t*); на них так или иначе могут влиять органы власти, управляющие регионами. Это были (см. таблицу), [4, 9]: Z_1*–уровень безработицы* в регионе, Z_5*– доля его изношенных основных фондов* (ОФ), Z_4*–соотношение безвозмездных поступлений в бюджет региона с ВРП,* Z_3 *–инвестиции в основной капитал на душу* (за исключением бюджетных средств)*;* аналогичный ему показатель имеется в [10]. Наконец, вводился показатель *риска потери здоровья и жизни* Z_2 [4,5,11] :

$$Z_2 = 100\left[0.85\left(m + s_n\right) + 0.15\left(v - s_n + n_s\right)\right]/N \qquad (1)$$

где m –зарегистрированное в МВД за год t число умышленных убийств в регионе (с покушениями), v – число случаев умышленного нанесения тяжкого вреда здоровью (УНТВЗ), s_n – число случаев смерти по неосторожности из-за УНТВЗ и n_s – число изнасилований (с покушениями), N – средняя численность постоянных жителей региона; его индекс в (1) для краткости опущен. Перед включением в модель переменные Z_3 и Z_5 предварительно подвергались нелинейному преобразованию через [1,(2.4)] с заданием оптимальных значений: 100 тыс руб.(на чел.) и соответственно – 47% для износа ОФ [5].

Параметры β_j , j = 0, 1, ...5 и σ в модели

$$Y = \beta_0 + \sum_{j=1}^{5} \beta_j \, Z_j + \varepsilon \qquad (2)$$

были оценены посредством МНК по выборке с показателями Z_1 ,…, Z_5 для 67 регионов за 2017 год, где β_j – неизвестные коэффициенты; случайное возмущение ε по предпосылке имеет нулевое математическое ожидание и неизвестную дисперсию σ^2. Эти 67 регионов входили в список наших 76 регио-

нов; из перечня регионов страны исключались: Москва, Севастополь ; республики Бурятия, Ингушетия, Калмыкия, Марий Эл, Саха-Якутия, Тыва, Чеченская; Чукотский АО; Иркутская, Тамбовская, Тюменская области. Свободный член β_0 учитывает влияние на Y всех прочих факторов, не учтённых в (2).

В таблице приводятся результаты оценивания по МНК β_j-коэффициентов (1-я строка), модули t-статистик для проверки Ho: “$\beta_j=0$ для каждого j “ (2-я строка), достигнутые уровни значимости и средние частные коэффициенты эластичности (3- и 4-я строки), коэффициент детерминации R^2 и его подправленное значение, несмещённая оценка стандартного отклонения σ (5-я строка), результаты дисперсионного анализа остатков и статистика DW, статистики критериев согласия распределения остатков с нормальным: 6-, 7-я строка.

Результаты оценивания параметров в (2); статистики и тесты

	Y	Z_1	Z_2	Z_3	Z_4	Z_5
	Const	Безработица, %	Криминал. риск на 100 тыс. нас.	Инвест. в ОК, балл	Дотационность, %	Процент износа ОФ, балл
1.	7,06877	–0,16648	–0,12525	0,06626	–5,36373	0,08089
2.	(16,5)	(5)	(9,7)	(2,2)	(5,1)	(3,5)
3.	< 0.001	< 0,001	< 0,001	0,0295	< 0,001	< 0,001
4.	Коэффициент эластичности	–0,18	–0,27	0,09	–0,06	0,11
5.	R^2	81,3%	R^2_{adj}	79,8%	Оценка *б*	0,43071
6.	F-statistics	53	$v = 61$	< 0,001	DW	1,8133
7.	$\chi 2$	1,256	Dn	0,067	ω^{*2}	0,0475

Средние коэффициенты эластичности из таблицы показывают, на сколько процентов от своего среднего значения изменится Y при увеличении каждого фактора на 1% (и фиксировании всех прочих). Далее проверялись предпосылки МНК. Коэффициенты корреляции между Z_j малы; мультиколлинеарности нет [12,13]. Гипотеза об отсутствии в остатках автокорреляции 1-го порядка на 5% уровне не отвергается: статистика DW попадает в ($Du, 4-Du$); где порог Du =1.77 [12]. Также рассматривался график зависимости остатков для (2) от номера региона; они лежат вдоль оси абсцисс в ($-1, 1$). Другие графики свидетельствуют о независимости ε_i от Z_j ; остатки имеют случайный характер. Затем для каждого Z_j проверялись гипотезы о гомоскедастичности тестами: ранговой корреляции и Гольдфельда – Квандта [12, 13]; они не отвергались на 5% уровне значимости. В итоге можно считать, что предпосылки МНК ([13], формулы (5.34-38)) не противоречат выборочным данным; отсюда оценки параметров в (2) находятся по МНК в Excel. Кроме того, можно использовать критерии Стьюдента и Фишера, т.к. гипотеза о нормальности не отвергнута. Она проверялась через сравнение графиков двух функций: эмпирической функции распределения остатков и гауссовой, взятой при $E\varepsilon=0$ и оценке σ

=0,4307. Кроме графиков, применялись три критерия согласия [12-15]: χ^2 , Колмогорова-Смирнова, критерий ω^2. Значение из таблицы критической статистики χ^2 Фишера для проверки сложной гипотезы меньше критического значения $\chi^2(6)$ при 5% уровне значимости. Второй критерий основан на максимальном уклонении этих функций распределения (т.е. статистике *Dn* А.Н.Колмогорова) с учётом результатов моделирования по Монте-Карло [14] для проверки сложной гипотезы. Критерий ω^2 строится по статистике ω^{*2} [15,(10.8)], что была чуть модифицирована Стефенсом; для вычислений используется формула (10.5) из [15].

Оценка Y по модели (2) и факторам Z_j за 2016, 2017 гг. имеет модуль относительной ошибки меньше 5.5% для регионов: Ставропольский край; республики Башкортостан, Дагестан, Коми, Карачаево-Черкесская ,Чувашия; Белгородская, Брянская, Владимирская, Воронежская, Калужская, Мурманская, Новгородская, Омская, Оренбургская, Орловская, Самарская, Тверская, Тульская, Ульяновская области. Данные за 2016 г. никак не использовались при оценивании. Максимальная ошибка менее 9% в эти годы наблюдалась в регионах: Пермский, Приморский, Хабаровский края; республики Карелия и Татарстан, а также области: Архангельская, Астраханская, Ивановская, Курская, Ленинградская, Липецкая, Новосибирская, Ростовская, Рязанская, Саратовская.

СПИСОК ЛИТЕРАТУРЫ

1. *Айвазян С. А.* Анализ качества и образа жизни населения: эконометрический подход: монография, М. : Наука, 2012. 432 с.
2. *Aivazian Sergey A.* Quality of Life and Living Standard Analysis: an Econometric Approach. Germany : De Gruyter, 2014. 399 p.
3. *Айвазян С. А., Степанов В. С., Козлова М. И.* Измерение синтетических категории качества жизни населения региона и выявление ключевых направлений совершенствования социально-экономической политики // Прикладная эконометрика. 2006. № 2. С. 18–84. [Электронный ресурс]. URL http://www.appliedeconometrics.ru/general/upload/articles/pe0206-18-renamed.pdf (дата обращения: 27.08.2019).
4. Регионы России. Соц.-экономические показатели : стат. сб. / Росстат. М., 2018.
5. *Степанов В. С.* Индикатор уровня благосостояния населения Крыма и близких ему регионов: связь с факторными переменными // Вестник ЦЭМИ. Доступ для зарегистрированных пользователей. [Электронный ресурс]. URL: http://ras.jes.su/cemi/s265838870004976-6-1 (дата обращения: 13.08.2019).
6. *Андерсон Т.* Введение в многомерный статистический анализ. М. : Мир,1963. 500 с.
7. *Гранберг А. Г., Зайцева Ю. С.* Производство и использование ВРП : Межрегиональные сопоставления. Статья 2 // Российский экономический журнал. 2002. № 11-12.
8. *Айвазян С. А., Исакин М. А.* Интегральные индикаторы качества жизни населения как критерии эффективности социально-экономической политики, проводимой органами региональной власти // Прикладная эконометрика, 2005. № 4. С. 16–22.
9. Финансы России : Приложение в разрезе субъектов РФ : стат. сб. / Росстат. М. : 2018.
10. *Путин В. В.* Указ Президента РФ от 25.04.2019. № 193.
11. Портал правовой статистики Генеральной Прокуратуры РФ. [Электронный ресурс]. URL: http://crimestat.ru/regions_chart_total (дата обращения: 02.06.2019).

12. *Кремер Н. Ш., Путко Б. А.*Эконометрика : учебник. М. : ЮНИТИ-ДАНА, 2002. 311 с.

13. *Гладилин А. В.* Эконометрика : учеб. пособие для вузов. М. : Кнорус, 2006. 232 с.

14. *Lilliefors H. W.* On the Kolmogorov-Smirnov Test for Normality with Mean and Variance are Unknown // J. American Statist. Assoc. 1967. 62. № 318. P. 399-402.

15. *Тюрин Ю. Н., Макаров А. А.* Статистический анализ данных на компьютере. М. : Инфра-М, 1998. 528 с.