

Министерство науки и высшего образования Российской Федерации
Владивостокский государственный университет экономики и сервиса (ВГУЭС)

**ИНТЕЛЛЕКТУАЛЬНЫЙ ПОТЕНЦИАЛ ВУЗОВ –
НА РАЗВИТИЕ ДАЛЬНЕВОСТОЧНОГО РЕГИОНА
РОССИИ И СТРАН АТР**

Материалы XX международной научно-практической
конференции студентов, аспирантов и молодых ученых

18–20 апреля 2018 г.

В четырех томах

Том 2

Под общей редакцией д-ра экон. наук Т.В. Терентьевой

Владивосток
Издательство ВГУЭС
2018

4. Одишко, Н.Н. Бизнес-планирование как переход к новому качеству экономического развития в рыночных условиях / Н.Н. Одишко, Е.Г. Гусев, Н.Ю. Голодная // Экономика и предпринимательство. 2015. № 5-1 (58-1).
5. Одишко, Н.Н. Имитационное моделирование в анализе инвестиционного проекта / Н.Н. Одишко, С.М. Гриванова, А.Г. Гузенко // Экономика и предпринимательство. 2014. № 12-3 (53-3).
6. Гузенко, А.Г. Моделирование потребительского рынка торговой марки / А.Г. Гузенко, Н.Н. Одишко, К.И. Лавренко // Экономика и предпринимательство. 2013. № 5 (34). С. 384-389.
7. Одишко, Н.Н. Оптимизация деятельности предприятий малого бизнеса как основа повышения качества обслуживания населения / Н.Н. Одишко, Н.Ю. Голодная, А.Г. Гузенко, Е.Г. Гусев // Экономика и предпринимательство. 2016. № 6 (71).
8. Одишко, Н.Н. Оптимизация работы информационной системы предприятия малого бизнеса по оказанию консультационных услуг / Н.Н. Одишко, Н.Ю. Голодная, Р.С. Реуцкий.
9. Одишко, Н.Н. Применение аддитивной и мультипликативной моделей прогнозирования / Н.Н. Одишко, Н.Ю. Голодная // Экономика и предпринимательство. 2013. № 12-1 (411). С. 667-674.
10. Одишко, Н.Н. Создание интерфейса взаимодействия пользователя с базой данных предприятия / Н.Н. Одишко, И.В. Ермак // Интеллектуальный потенциал вузов - на развитие дальневосточного региона России и стран АТР: материалы XVIII Международной конференции студентов, аспирантов и молодых ученых. 2016. С. 147-149.

Рубрика: Эконометрия

УДК 330.43

ИЗУЧЕНИЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ФАКТОРОВ, ВЛИЯЮЩИХ НА ЗАБОЛЕВАЕМОСТЬ ВИЧ

Ю.Д. Чердонова

бакалавр, 1 курс

К.Л. Шаркова

бакалавр, 1 курс

Е.Д. Емцева

канд. физ.-мат. наук, доцент кафедры математики и моделирования

*Владивостокский государственный университет экономики и сервиса
Россия, Владивосток*

В работе представлены результаты оценки закономерности динамики показателя заболеваемости ВИЧ в регионах России с использованием методов эконометрического моделирования. В работе рассмотрен временной период с 2005-го по 2015 год. В качестве объекта взяты 78 регионов РФ. Все исследования и расчеты проведены с использованием программы Rstudio.

Ключевые слова и словосочетания: ВИЧ, регионы РФ, панельные данные, метод главных компонент, PCA.

STUDY OF SOCIO-ECONOMIC FACTORS INFLUENCING THE MORBIDITY OF HIV

The article describes the results of the evaluation of the pattern of the incidence rate of HIV in the regions of Russia using methods of econometric modeling. The period of our study from 2005 to 2015. The 78 regions of Russia are considered. For statistical processing of data and working with graphics we used Rstudio.

Keywords: HIV, regions of the Russian Federation, panel data, principal component analysis, PCA.

На сегодняшний день проблема заболеваемости ВИЧ инфекцией относится к глобальным проблемам. На начало 2017 года общее количество случаев заболеваемости ВИЧ инфекцией среди граждан России достигло 1 114 815 человек (в мире – 36,7 миллионов ВИЧ-инфицированных, в том числе 2,1 млн детей). На конец прошлого года Россия уже являлась третьей страной в мире по числу новых случаев ВИЧ-инфекции после ЮАР и Нигерии.

Целью нашего исследования является изучение влияния различных социально-экономических факторов на динамику показателя заболеваемость ВИЧ-инфекцией по регионам РФ. Наиболее существенный рост (скорость, темп роста появления новых случаев ВИЧ в единицу времени) заболеваемости в 2016 г. наблюдался в Республике Крым, Карачаево-Черкесской Республике, Чукотском АО, Камчатском крае, Белгородской, Ярославской, Архангельской областях, г. Севастополь, Чувашской, Кабардино-Балкарской Республиках, Ставропольском крае, Астраханской области, Ненецком АО, Самарской области и Еврейской АО. При проведении исследования мы использовали показатель y – количество больных ВИЧ на 1000 человек, и такие факторы, как: x_1 – доля занятого населения (отношение численности занятого населения определенной возрастной группы к общей численности населения соответствующей возрастной группы), x_2 – доля городского населения, x_3 – коэффициент дифференциации доходов (характеризует степень социального расслоения и определяется как соотношение между средними уровнями денежных доходов 10% населения с самыми высокими доходами и 10% населения с самыми низкими доходами), x_4 – число мигрантов на 1000 человек населения, x_5 – число безработных на 1000 человек населения, x_6 – денежные доходы в среднем на душу населения, x_7 – ВРП на душу населения в ценах 2001 года, x_8 – доля населения в процентах, с уровнем доходов ниже прожиточного минимума, x_9 – продажа пива литров на душу населения, x_{10} – продажа водки литров на душу населения, x_{11} – продажа вина, игристых вин литров на душу населения, x_{12} – нагрузка на одного врача (измеряется в численности населения на одного врача), x_{13} – мощность амбулаторно-политехнических учреждений (измеряется в количестве посещений в смену в тысячах), x_{14} – количество прерываний беременности (измеряется на 1000 женщин 15-49 лет), x_{15} – число зарегистрированных преступлений на 100 тысяч человек.

Все эти данные были собраны на сайтах государственной статистики [1,2] по 78 регионам с 2002 года по 2015 год, а моделирование производилось с использованием программы Rstudio [3,4]. В качестве эконометрического инструмента исследования взята модель с учетом панельных структурных данных [5]. Общая специфика, которой имеет вид:

$$y_{it} = \alpha + X_{it}^k \beta + v_{it} \cdot i = \overline{1, N}, t = \overline{1, T}, \quad (1)$$

где i – номер объекта, t – время, α – свободный член, β – вектор-столбец коэффициентов размерности $k \times 1$, $X_{it} = (X_{1,it}; X_{2,it}; \dots; X_{k,it})$ – вектор-строка матрицы k объясняющих переменных, $v_{it} = u_i + \varepsilon_{it}$ – случайная ошибка с ненаблюдаемыми индивидуальными эффектами u_i и остаточным возмущением ε_{it} .

Данные используемые, для построения данного типа модели подверглись предварительной стандартизации.

Методом Best Subsets построена модель панельных данных с детерминированными эффектами зависимости показателя заболеваемости от экологических факторов для 78 регионов, которая получилась значимой, со значимыми коэффициентами при переменных $x_2, x_7, x_8, x_{10}, x_{11}, x_{12}, x_{13}$, результаты моделирования представлены на рис. 1.

При увеличении, например, доли городского населения значение (фактор x_2) показатель заболеваемости ВИЧ увеличивается, возможно, в сельской местности меньше заболевших ВИЧ или существует проблема диагностики данного заболевания. Так же положительное воздействие на рост заболеваемости оказывают факторы $x_7, x_{11}, x_{12}, x_{13}$ и отрицательное x_8, x_{10}, x_{13} . Влияние некоторых факторов вполне поддаются интерпретации с точки зрения здравого смысла, иные требуют более вдумчивого и доказательного осмысления. Так, например, кажется вполне логичным прямая связь (положительный коэффициент) заболеваемости с ростом потребления вина, большой нагрузке на врачей и мощность амбулаторно-политехнических учреждений. Но положительная связь заболеваемости с показателем ВРП, а также отрицательная с показателем доли населения с уровнем доходов ниже прожиточного минимума и потреблением водки, хотя и кажется нам небезосновательным, но, тем не менее, требует более серьезного обоснования.

```

Call:
plm(formula = y ~ x2 + x7 + x8 + x10 + x11 + x12 + x13, data = pdatas,
     model = "within")

Balanced Panel: n = 78, T = 14, N = 1092

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-1.882312 -0.176742 -0.032923  0.157648  2.559214

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
x2    0.381188   0.150056   2.5403 0.0112249 *
x7    0.232172   0.039116   5.9354 4.028e-09 ***
x8   -0.100373   0.015215  -6.5970 6.767e-11 ***
x10  -0.311207   0.023007 -13.5266 < 2.2e-16 ***
x11   0.087054   0.028234   3.0833 0.0021029 **
x12   0.189032   0.062475   3.0257 0.0025433 **
x13   0.319780   0.082989   3.8533 0.0001239 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    314.32
Residual Sum of Squares: 176.86
R-Squared:                0.43732
Adj. R-Squared:          0.39039
F-statistic: 111.808 on 7 and 1007 DF, p-value: < 2.22e-16
> |

```

Рис.1. Модель с детерминированными эффектами

Следующий этап анализа данных, который мы выполнили - это использование «Метода главных компонент» (PCA-principal components analysis)[5].

Этот метод нередко позволяет построить более адекватные модели со всеми факторами, при этом преобразовать большое число скоррелированных переменных в гораздо меньший набор нескоррелированных переменных, называемых главными компонентами. Главные компоненты сохраняют максимально возможную часть информации, которая содержалась в исходном наборе переменных.

Согласно критерию Кайзера определено число главных компонент, равное трем и четырем - для компонент с вращением.

```

call:
plm(formula = y ~ PC1 + PC2 + PC3, data = pdata2, model = "within")

balanced panel: n = 78, T = 14, N = 1092

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-1.777153 -0.182191 -0.006708  0.166781  2.888923

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
PC1    0.160122   0.041500   3.8583 0.0001214 ***
PC2   -0.366332   0.023305 -15.7193 < 2.2e-16 ***
PC3    0.095703   0.021694   4.4115 1.137e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    314.32
Residual sum of squares: 170.51
R-Squared:                0.45753
Adj. R-Squared:          0.41461
F-statistic: 284.235 on 3 and 1011 DF, p-value: < 2.22e-16
> |

```

Рис. 2. Модель с детерминированными эффектами главным компонентам без вращения

```

call:
plm(formula = y ~ RC1 + RC2 + RC3 + RC4, data = pdata3, model = "within")

Balanced Panel: n = 78, T = 14, N = 1092

Residuals:
    Min.   1st Qu.   Median   3rd Qu.    Max.
-1.825731 -0.177373 -0.010846  0.161355  2.913604

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
RC1  0.105051  0.037087  2.8325  0.004710 **
RC2 -0.285905  0.018127 -15.7725 < 2.2e-16 ***
RC3  0.304967  0.013590  22.4409 < 2.2e-16 ***
RC4  0.132316  0.043526  3.0400  0.002427 **
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:  314.32
Residual Sum of Squares: 164.58
R-Squared:  0.47637
Adj. R-Squared: 0.43438
F-statistic: 229.714 on 4 and 1010 DF, p-value: < 2.22e-16

```

Рис.3. Модель с детерминированными эффектами главным компонентам с вращением

Построенные модели с детерминированными эффектами по главным компонентам получились статистически значимые, со значимыми коэффициентами при главных компонентах, с коэффициентом детерминации, близким к 0.5.

Но главные компоненты и нагрузки на них факторов зачастую нелегко интерпретировать, поэтому на следующем этапе исследования факторы предварительно разбили на группы с учетом межфакторной корреляции и аналитических соображений. В первую группу, названную нами «социальные факторы», вошли показатели: $x_1, x_2, x_4, x_5, x_{15}$, вторую группу («факторы материального состояния») составили показатели x_3, x_6, x_7, x_8 ; в третью группу («факторы, отражающие потребление алкоголя и качество медицинского обслуживания») вошли показатели $x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}$. Последнюю группу мы сформировали, объединив в одну предварительно исследованные две группы факторов – потребление алкоголя и факторов медицинского характера, руководствуясь результатами поэтапных построений моделей и поиска наиболее адекватных.

```

Call:
plm(formula = y ~ PC1 + PC1.1 + PC1.2, data = pdata4, model = "within")

Balanced Panel: n = 78, T = 14, N = 1092

Residuals:
    Min.   1st Qu.   Median     3rd Qu.    Max.
-1.692453 -0.187573 -0.029642  0.158281  3.039212

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)
PC1      0.065402    0.024452  2.6747  0.0076 **
PC1.1    0.310812    0.018259 17.0222 < 2.2e-16 ***
PC1.2   -0.150648    0.022231 -6.7766 2.087e-11 ***
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:   314.32
Residual Sum of Squares: 180.21
R-Squared:              0.42666
Adj. R-Squared:         0.38129
F-statistic: 250.782 on 3 and 1011 DF, p-value: < 2.22e-16

```

Рис.4. Модель с детерминированными эффектами по групповым главным компонентам

Таким образом, результаты моделирования зависимости заболеваемости ВИЧ от групповых главных компонент представлены на рисунке 4. Модель статистически значима, со всеми значимыми коэффициентами при компонентах, с коэффициентом детерминации 0.4.

Предполагаются дальнейшие исследования в данном направлении, с надеждой, что использование методов статистического анализа помогут разрешить, прояснить и, возможно, регулировать заболеваемость населения ВИЧ-инфекцией.

1. Официальный сайт Министерства Здравоохранения Российской Федерации. [Электронный ресурс]: URL: <http://www.gosminzdrav.ru>.
2. Официальный сайт Федеральной службы государственной статистики. [Электронный ресурс]: URL: <http://www.gks.ru>.
3. Официальный сайт проекта R. [Электронный ресурс]: URL: <https://cran.r-project.org/>.
4. Kabacoff R.I. R in Action: Data Analysis and Graphics with R. Manning Publications, 2011. 447 p.
5. Елисеева, И.И. Эконометрика: учебник / И.И. Елисеева, Курышева С.В., Костеева Т.В., Пангина И.В., Михайлов Б.А., Нерадовская Ю.В., Штрое Г.Г., Бартелс К., Рыбкина Л.Р.. М.: Финансы и статистика, 2007. 576 с.

