

ИСПОЛЬЗОВАНИЕ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ ПРИ ВЫЯВЛЕНИИ ПРИОРИТЕТОВ РЕГИОНАЛЬНОЙ ИНВЕСТИЦИОННОЙ ПОЛИТИКИ В ОТНОШЕНИИ ИНОСТРАННЫХ ИНВЕСТОРОВ В РЕГИОНЫ РФ

УДК 339.727.22

Лилия Валериевна Матраева,
к.э.н., доцент кафедры «Финансы и Банковское дело», Российский государственный торгово-экономический университет, Тульский филиал
Тел.: 8 (903) 841-89-49
Эл. почта: matraeva@rambler.ru

Данная статья посвящена проблемам исследования инвестиционного климата страны и регионов. Поскольку дискуссионной является сама возможность статистической оценки инвестиционного климата в силу субъективности инструментов, используемых для его измерения, то в статье предпринята попытка разработки модели, отражающей влияние инструментов региональной инвестиционной политики на потоки прямых иностранных инвестиций, т.е. решить не общий вопрос измерения инвестиционного климата региона, а частный – какие стимулирующие инструменты являются наиболее востребованными иностранными инвесторами в рамках конкретной территории.

Ключевые слова: прямые иностранные инвестиции, логистическая регрессия, региональная инвестиционная политика.

Lilia V. Matraeva,
PhD in economics, associate professor, the Department of «the Finance and Banking», Russian State University of Trade and Economics, Tula branch
Tel.: 8 (903) 841-89-49
E-mail: matraeva@rambler.ru

USE OF LOGISTIC REGRESSION IN IDENTIFYING PRIORITIES FOR REGIONAL INVESTMENT POLICY IN RESPECT OF FOREIGN INVESTORS IN THE REGIONS OF THE RUSSIAN FEDERATION

This paper focuses on the problems of the study of the investment climate in the country and the regions. Since the possibility of statistical assessment of the investment climate is itself under discussion due to the subjectivity of instruments used to measure it, the article attempts to develop a model that reflects the impact of the investment instruments of regional policy on flows of foreign direct investments, i.e. not to solve the general problem of measurement of the investment climate in the region, but to solve the particular one – what incentive tools are the most in demand by foreign investors in the particular area or region.

Keywords: foreign direct investments, logistic regression, regional investment policy.

1. Введение

Для успешной разработки и реализации стратегических и тактических инвестиционных планов развития территории необходимо всестороннее изучение ситуации в регионе – его инвестиционного климата. Поскольку при создании подобных планов наиболее остро встает вопрос о том, какие стимулирующие инструменты являются наиболее востребованными иностранными инвесторами в рамках конкретной территории, то можно сместить фокус исследования с возможности комплексной оценки инвестиционного климата и ее субъективизма на углубленное изучение эффективности (с точки зрения достижения поставленной цели), осуществляемых в регионе стимулирующих мероприятий. При этом необходимо использовать дифференцированный подход, на основании которого могут быть найдены собственные детерминанты инвестиционной политики в отношении иностранных инвесторов в зависимости от их интересов и мотивации. Таким образом, на данном этапе была поставлена задача разработки модели, отражающей влияние инструментов региональной инвестиционной политики на потоки прямых иностранных инвестиций. Это позволит:

Во-первых, выделить наиболее эффективные направления стимулирования для различных стадий инвестиционного процесса;

Во вторых, определить вероятность увеличения потоков ПИИ или их качества в зависимости от инвестиционной политики, осуществляемой конкретным регионом.

2. Выбор метода и построение моделей

Поскольку в нашем случае отклик может принимать только два значения: будет или не будет достигнута поставленная цель при использовании определенного набора инструментов инвестиционной политики в отношении иностранных инвесторов, то для решения поставленной задачи можно использовать инструментарий моделей бинарного выбора, наиболее распространенными среди которых являются логистическая регрессия и пробит-регрессия. Логистическая регрессия позволяет исследовать зависимость дихотомических переменных от независимых переменных, имеющих любой вид шкалы. Обычно, со значениями дихотомических переменных связывают вероятность наступления какого-то события в зависимости от исследуемых независимых переменных. В отличие от логит-модели, основанной на логистическом законе распределения вероятностей, пробит-модель основана на законе нормального распределения. Эти модели, тем не менее, имеют достаточно схожие графики распределения вероятностей, и, очень часто исследователи получают схожие результаты, особенно если предварительно использовалась соответствующая нормировка. Однако, более детальный анализ, позволяет увидеть, что схожие результаты получаются, только если они не слишком близки к нулю или единице, так как пробит-модель быстрее стремится к этим значениям. Таким образом, если ожидаемые значения близки к краям функции распределения, то более «чувствительной» в данном случае является логит-модель.

Альтернативным инструментом статистического исследования в данном случае является линейный дискриминантный анализ, который также позволяет оценивать принадлежность к выделенным группам, однако использование этого инструмента предполагает нормальность распределения независимых переменных, а для применения логистической регрессии это условие не является обязательным, в тоже время результаты сравнимы с линейным дискриминантным анализом.

Поэтому, учитывая специфику исходных данных, в качестве инструмента исследования был выбран аппарат бинарной логистической регрессии.

Модель бинарной логистической регрессии, в случае если имеется более одной независимой переменной, имеет вид (форм.1) [1]:

$$P = \frac{1}{1 + e^{-y}}, \quad (1)$$

где P – вероятность того, что произойдет интересующее событие;
 e – основание натуральных логарифмов 2,71...;
 y – стандартное уравнение регрессии.

Зависимая переменная в случае бинарной логистической регрессии может принимать только два значения (наступление или ненаступление события). На данном этапе для оценки эффективности инвестиционной политики были исследованы мероприятия, по стимулированию иностранных инвесторов, осуществляемые в регионах РФ на 2012 год. С помощью инструментария логистической регрессии решается задача построения двух моделей, позволяющей предсказывать вероятность наступления следующих событий:

А) Для регионов, находящихся на стадии допуска инвесторов: произойдет ли увеличение притоков ПИИ?

$$\hat{y}_1 = [0; 1] \quad (2)$$

1 – если объем ПИИ выше инвестиционного потенциала региона
 0 – в противном случае

Б) Для регионов, находящихся на стадии функционирования инвесторов: произойдет ли устойчивый рост ПИИ?

$$\hat{y}_2 = [0; 1] \quad (3)$$

1 – если объем ПИИ выше инвестиционного потенциала региона
 0 – в противном случае

В качестве исходной статистической базы использовались данные о наличии мероприятий по стимулированию иностранных инвесторов в конкретном регионе, сформированные автором с использованием данных для проведения подобной оценки представлены в Базе данных «Инвестиционные характеристики субъектов Российской Федерации» (официальный сайт <http://polpred.com>) и в социальных атласах российских регионов официальный сайт <http://www.socpol.ru>). На этих сайтах отражены наличие основных инструментов в инвестиционной политике, используемые в практике конкретного региона, а также приводятся ссылки на нормативные акты конкретных регионов в данной сфере. На основании этих данных по каждому региону был составлен «Паспорт стимулирующих инвестиционных мероприятий».

Отбор регрессоров в модель осуществлялся в ходе процедуры корреляционного анализа, где были отобраны характеристики, наиболее сильно влияющие на результирующую переменную.

На стадии допуска инвестора с вероятностью больше чем 95% (χ^2 факт > 3,84 с учетом $P < 0,05$) существует статистически значимая связь между ожиданиями инвестора и следующим набором показателей: Наличие инвестиционных площадок с готовой инфраструктурой ($x_{1.1}$); Преступления, совершенные в сфере экономики на 100000 населения ($x_{1.2}$); Участие в создании организаций ($x_{1.3}$); Преступления на 100000 населения ($x_{1.4}$); Поручительства администрации региона в обеспечении возврата привлекаемых инвестором средств ($x_{1.5}$); Наличие свободных экономических зон ($x_{1.6}$); Инвестиционные налоговые кредиты ($x_{1.7}$).

На стадии функционирования инвесторов с вероятностью больше чем 95% (χ^2 факт > 3,84 с учетом $P < 0,05$) существует статистически значимая связь с показателями: Наличие государственных программ кластеризации ($x_{2.1}$); Поручительства администрации региона в обеспечении возврата при-

влекаемых инвестором средств ($x_{2.2}$); Создание организационных структур по поддержке инвестиционной деятельности ($x_{2.3}$); Льготные условия пользования недвижимостью ($x_{2.4}$).

Для решения поставленной задачи использовался ППП SPSS 20.0 и метод пошагового (прямого) включения переменных. Порог отсекания был принят равным 0,5. Если в результате расчетов, полученное значение вероятности будет меньше 0,5, то предполагается, что событие не наступит. Если это значение равно или превышает 0,5, то предполагается, что событие наступит. В результате были получены два уравнения логистической регрессии (форм. 4 и 5), где в качестве предикторов остались все указанные выше переменные. Итоги оценки моделей приведены в табл. 1.

$$\begin{cases} \hat{y}_1 = (1 + e^{-z})^{-1} \\ z' = -2,72 - 2,045x_{1.1} - 2,188x_{1.2} + \\ + 1,540x_{1.3} + 3,831x_{1.4} - 3,522x_{1.5} + \\ + 3,967x_{1.6} + 2,957x_{1.7} \end{cases} \quad (4)$$

$$\begin{cases} \hat{y}_2 = (1 + e^{-z})^{-1} \\ z' = -4,107 - 0,936x_{2.1} + 2,509x_{2.2} + \\ + 1,961x_{2.3} - 1,767x_{2.4} \end{cases} \quad (5)$$

Таблица 1

Оценка переменных в уравнениях логистической регрессии

Переменные	В	Стд. ошибка	Вальд	Знч.	Exp(B)
Стадия допуска инвестора (\hat{y}_1)					
Наличие инвестиционных площадок с готовой инфраструктурой	-2,045	1,097	3,477	0,062	0,129
Преступления совершенные в сфере экономики	-2,188	1,041	4,415	0,036	0,112
Участие в создании организаций	1,540	1,252	1,513	0,219	4,666
Преступления на 100000 населения	3,831	1,513	6,410	0,011	46,092
Поручительства администрации в обеспечении возврата средств	-3,522	1,507	5,465	0,019	0,030
Наличие СЭЗ	3,967	1,783	4,946	0,026	52,803
Инвестиционные скидки	2,957	1,670	3,136	0,077	19,240
Константа	-2,720	1,435	3,594	0,058	0,066
Стадия функционирования инвестора (\hat{y}_2)					
Наличие государственных программ кластеризации	0,936	0,732	1,637	0,201	2,551
Поручительства администрации в обеспечении возврата средств	2,509	1,157	4,704	0,030	12,298
Создание организационных структур по поддержке инвестиционной деятельности	1,961	0,791	6,146	0,013	7,107
Льготные условия пользования недвижимостью	-1,767	0,847	4,353	0,037	0,171
Константа	-4,107	1,311	9,812	0,002	0,016

3. Интерпретация и оценка качества моделей

Дальнейший анализ полученных моделей уравнения множественной логистической регрессии традиционно проводится по следующим направлениям:

1. Интерпретация уравнения по оцениваемым параметрам

Смысловое значение логистических коэффициентов интерпретируют как изменение логарифмических шансов, вызванных единичным изменением независимой переменной. Экспонированное значение шансов ($\text{Exp}(B)$) показывает во сколько раз изменится шанс возникновения изучаемого события, если значения одного из предикторов изменится на единицу при фиксированных значениях прочих предикторов.

На основании данных, приведенных в табл. 1, можно отметить следующее:

На стадии допуска инвесторов наибольшее влияние оказывают следующие факторы:

- наличие СЭЗ увеличивает шанс притока ПИИ в 52 раза;

- наличие в регионе инвестиционных скидок, представляемых иностранным инвесторам, также существенно увеличивает шанс притока ПИИ – в 19 раз;

- необходимо отметить, что несмотря на невысокую статистическую значимость показателя «Участие в создании организаций» расчеты позволяют отметить его существенную роль в увеличении шансов притока ПИИ (в 4,6 раза).

- факторы «Наличие инвестиционных площадок с готовой инфраструктурой», «Преступления совершенные в сфере экономики», «Поручительства администрации в обеспечении возврата средств» показали статистически значимую связь с результирующим признаком, однако их наличие среди мер стимулирования иностранных инвесторов на данном этапе увеличивает шансы притока ПИИ в регион менее чем на 1%.

На стадии функционирования инвесторов ситуация выглядит несколько иначе:

- на первое место выходит фактор «Поручительства администрации в обеспечении возврата средств»: его наличие увеличивает шансы устойчивого роста ПИИ в 12 раз;

- на второе место по увеличению шансов (в 7 раз) занимает фактор «Создание организационных структур по поддержке инвестиционной деятельности»;

- далее следует «Наличие государственных программ кластеризации» (в 2,5 раза);

- фактор «Льготные условия пользования недвижимостью» показал статистически значимое, но не существенное влияние на увеличение шансов.

2. Оценка качества регрессионной модели оценивается по нескольким критериям.

Во-первых, для оценки качества логистической регрессии используют меры определенности Кокса&Шнела и Нэйджелкерка. Они характеризуют часть дисперсии, объясняемой с помощью построенной модели логистической регрессии [2]. Для первой модели R квадрат Нэйджелкерка (псевдокоэффициент детерминации) составил 64,9%. Для второй модели – 50,9%.

Во-вторых, для оценки доли правильно определенных (переклассифицированных) наблюдений используют проценты конкордации. Обе полученные модели обладают достаточно высокими прогностическими свойствами. Первая модель в целом правильно описывает 85,3% наблюдений, в том числе было предсказано верно 56,3% из общего числа наблюдений, в которых результат был положительным; 94,2% – отрицательным. Вторая модель имеет результаты чуть хуже: в общей выборке ею предсказано верно 82,4% наблюдений. Она правильно описывает 50% наблюдений, результат которых был положительным и 92,3% наблюдений, результат которых был отрицательным.

Таким образом, учитывая небольшой объем статистической выборки и высокую неопределенность социально-экономических процессов, а также непараметрический характер исследуемых переменных, прогностические свойства построенных моделей можно признать удовлетворительными.

Самым наглядным методом, позволяющим оценить качество модели, является построение ROC-кривой. Однако, как справедливо отмечается в работе Илышева А.М. и Шубата О.М., посвященной исследованию практике применения многомерных методов классификации [3], «... этот метод еще не достаточно применяется

в отечественной науке и крайне редко встречается в работах российских статистиков» представляется совершенно обоснованным утверждение авторов статьи о том, что наиболее полно техника применения логистической регрессии представлена в работах И.И. Елисейевой и В.С. Мхитаряна. Среди последних публикаций в этой области можно выделить работы Н.Б. Паклина, В.И. Орешкова, Цыпиной Н.Ю [4]. В связи с вышесказанным рассмотрим методику построения ROC-кривой для решения поставленной в исследовании задачи более подробно.

ROC-кривая показывает зависимость количества верно классифицированных положительных примеров (истинно положительное множество) от количества неверно классифицированных отрицательных примеров (ложно отрицательное множество) [4].

Положительные примеры, классифицированные как отрицательные в подобных терминах трактуются как ошибка I рода. То есть в этом случае построенная модель не выявила интересующее нас событие. В свою очередь отрицательные примеры, классифицированные как положительные представляют собой ошибку II рода. [4]

Для определения качества бинарного классификатора рассчитывают два показателя: чувствительность (S_e) и специфичность (S_p) модели [1].

Само по себе исследование таких характеристик модели как чувствительность и специфичность, дают возможность определить две очень важных показателя модели:

1. оптимальный порог отсечения, в котором ошибки ошибок I и II рода минимальны ($S_p + S_e \Rightarrow \max$), то есть это точка, в которой чувствительность и специфичность модели максимальна.

2. точку баланса, в которой чувствительность и специфичность примерно совпадают ($|S_e - S_p| \Rightarrow \min$).

Значения специфичности и чувствительности для исследуемых моделей представлены в табл. 2.

Как показывают данные, приведенные в табл. 2, для обеих моделей точка баланса и оптимальный порог отсечения совпадают. Для первой модели это значение равно порядка 0,2; для второй – 0,26.

На основании полученных можно построить график ROC-кривой и для большей точности вычислить площадь под ROC-кривой (AUC – Area

Таблица 2

Значения чувствительности и специфичности моделей

Порог классификации	$S_e, \%$	$S_p, \%$	$S_p + S_e, \%$	$ S_e - S_p , \%$
Модель стадии допуска инвестора				
...				
0,1	100,00	69,23	169,23	30,77
0,2	87,50	84,48	171,98	3,02
0,3	81,25	86,54	167,79	5,29
0,4	56,25	90,38	146,63	34,13
0,5	56,25	94,23	150,48	37,98
0,6	56,25	94,23	150,48	37,98
0,7	37,50	98,08	135,58	60,58
0,8	37,50	100,00	137,50	62,50
0,9	31,25	100,00	131,25	68,75
...				
Модель стадии функционирования инвестора				
...				
0,1	100,00	46,15	146,15	53,85
0,2	81,25	67,31	148,56	13,94
0,26	75,00	78,85	153,85	3,85
0,3	68,75	80,77	149,52	12,02
0,4	50,00	92,31	142,31	42,31
0,5	50,00	92,31	142,31	42,31
0,6	50,00	92,31	142,31	42,31
0,7	0,00	100,00	100,00	100,00
...				

Таблица 3

Площади под ROC-кривой

Модель	Показатель AUC (площадь под кривой)	Стандартная ошибка	Асимптотическая значимость
Модель стадии допуска инвестора	0,813	0,078	0,001
Модель стадии функционирования инвестора	0,762	0,085	0,005

Таблица 4

Регионы, инвестиционный климат в которых, с вероятностью более 50% обеспечивает увеличение потоков ПИИ

Субъект РФ	Вероятность увеличения потоков ПИИ
Амурская область	0,99993
Архангельская область	0,99993
Томская область	0,99968
Новгородская область	0,99634
Еврейская АО	0,95669
Сахалинская область	0,82561
Липецкая область	0,77665
Республика Карелия	0,75219
Республика Коми	0,75219
Республика Хакасия	0,75219
Курганская область	0,63304
Кемеровская область	0,61357

Таблица 5

Регионы, инвестиционный климат в которых, с вероятностью более 50% обеспечивает устойчивый рост ПИИ

Субъект РФ	Вероятность устойчивого роста ПИИ
Калининградская область	0,78588
Алтайский край	0,78588
г.Санкт-Петербург	0,78588
Нижегородская область	0,78588
Самарская область	0,78588
Республика Татарстан	0,78588
Саратовская область	0,78588
Чувашская республика	0,78588
Сахалинская область	0,58999
Псковская область	0,58999
Удмуртская Республика	0,58999
Тюменская область	0,58999

Under Curve). Однако используют этот показатель в основном для сравнения нескольких моделей. Что касается трактовки значений данного показателя, то устоявшегося мнения не существует. Наиболее распространенная [7]: 0.9–1.0 – Отличное, 0.8–0.9 – Очень хорошее, 0.7–0.8 – Хорошее, 0.6–0.7 – Среднее, 0.5–0.6 – Неудовлетворительное (табл. 3).

Результаты оценки вероятности увеличения потоков ПИИ для регионов РФ и результаты оценки вероятности устойчивого роста ПИИ для регионов РФ приведены соответственно в табл. 4 и 5.

В наибольшей степени **соответствуют ожиданиям инвесторов**, находящихся на стадии допуска инвестиционная политика регионов, перечень которых приведен в табл. 4. Вероятность увеличения потоков ПИИ для этих регионов составляет более 50%. *Здесь следует особенно подчеркнуть, что в данном случае и далее речь идет об увеличении потоков ПИИ, обусловленном исключительно эффективностью (в нашем случае это соответствие ожиданиям иностранных инвесторов) проводимой в регионе инвестиционной политикой, хотя она, безусловно, не является единственно определяющим фактором, обуславливающим потоки иностранных инвестиций в регион.*

В табл. 5 соответственно приведен перечень регионов, для которых вероятность устойчивого роста инвестиционных потоков составляет более 50%, то есть инструменты, используемые в рамках инвестиционной политики данных регионов, в основном соответствуют ожиданиям инвесторов, находящихся на стадии функционирования.

Данные табл. 5 свидетельствуют, что инвестиционная политика Алтайского края, Чувашской республики, Удмуртской Республики, Тюменской области соответствует ожиданиям инвесторов, находящихся на стадии функционирования, тем не менее, фактически устойчивый рост ПИИ в данных регионах отсутствует. Одной из причин является то, что очень часто в российских регионах политика в отношении иностранных инвесторов формируется либо интуитивно, либо «по образцу», когда за основу берется инвестиционная политика, проводимая в «успешном» регионе, в надежде

повторить историю его успеха. Тем не менее, не всегда история успеха может быть адаптирована к конкретному региону.

4. Заключение

Таким образом, используемый инструментарий бинарной логистической регрессии позволил не только выявить, но и количественно описать наиболее существенные индикаторы, детерминирующие процесс притока ПИИ на различных стадиях инвестиционного цикла.

На стадии допуска иностранных инвесторов основой для принятия решения об инвестировании является ресурсный фактор. Тем не менее, толчком для ускорения притока ПИИ в российских регионах являются в основном методы фискального стимулирования. Кроме того, устойчивое влияние на этот процесс имеет наличие гарантий администрации региона, представляемый иностранным инвесторам и качество правовой среды.

На этапе функционирования инвестор смещает свои приоритеты, уделяя более внимание совершенности бизнес-среды региона, качеству административного процесса и заинтересованности администрации региона

в ПИИ. Влияние финансового и налогового стимулирования на данном этапе перестает быть приоритетным для инвестора.

При детальном рассмотрении полученные результаты могут быть использованы в качестве базы для разработки комплексной дифференцированной по отношению каждому региону страны политики в отношении ПИИ. Так, используя инструментарий логистической регрессии, можно оценить вероятность увеличения потоков ПИИ или устойчивости их роста в зависимости от инвестиционной политики, осуществляемой конкретным регионом, что косвенно отражает эффективность применяемых инструментов стимулирования иностранных инвесторов в регионе.

Литература

1. Илышев А.М., Шубат О.М. Возможности углубления статистического изучения малого предпринимательства // Вопросы статистики. – 2009. – №2. – С. 8–17.
2. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. – М. : Изд. дом ГУ ВШЭ, 2006. – 281 с.
3. Илышев А.М., Шубат О.М. Многомерная классификация дан-

ных: особенности методики, анализ практики и перспектив применения // Вопросы статистики. – 2010. – №10. – С. 34–40.

4. Паклин Н.Б. Логистическая регрессия и ROC-анализ – математический аппарат // Официальный сайт компании BaseGroup Labs URL: <http://www.basegroup.ru/library/analysis/regression/logistic/> (вход свободный, дата обращения 12.05.2012).

References

1. Ilyshev A.M., Shubat O.M. Opportunities to deepen the statistical study of small business // Voprosy statistiki. – 2009. – № 2. – P.8–17.
2. Kryshchanovskiy A.O. The analysis of social science data using the package SPSS. – M. Izdat. dom HSE, 2006. – 281 p.
3. Ilyshev A.M., Shubat O.M. Multidimensional data classification : features techniques, analysis of practices and perspectives of application // Questions of Statistics. – 2010. – № 10. – P. 34–40.
4. Paklin N.B. Logistic regression and ROC- analysis – a mathematical tool// Official site of BaseGroup Labs URL: <http://www.basegroup.ru/library/analysis/regression/logistic/> (free admission, date of access 12.05.2012).