3.4. ОЦЕНКА СТОИМОСТИ КОТТЕДЖА В НОВОЙ МОСКВЕ

Галочкин В.Т., к.ф.-м.н., доцент кафедры «Моделирование экономических и информационных систем»

Финансовый университет при Правительстве РФ

Рассмотрена линейная множественная регрессионная модель применительно к оценке стоимости нового коттеджа в Новой Москве. Получено удовлетворительное согласие рыночной стоимости квартиры, заявленной риелтером, с оценкой, полученной по уравнению регрессии, и выявлены основные факторы, определяющие выбор коттеджа потенциальным покупателем.

Эконометрика – научная дисциплина, объединяющая совокупность теоретических результатов, приемов, методов и моделей, предназначенных для того, чтобы на базе экономической теории и математического инструментария придавать конкретные количественные выражения общим закономерностям, обусловленным экономической теорией.

Одной из основных задач эконометрики является выявление взаимосвязи между экономическими переменными. Как правило, результирующий показатель зависит от нескольких факторов. В то же время, в модель необходимо ввести остаточную случайную составляющую, отражающую влияние на результирующий показатель всех неучтенных факторов. Присутствие случайной составляющей отражает стохастический характер зависимости. Например, наблюдая спрос рынка в разные моменты времени, мы увидим случайное варьирование спроса около некоторого определенного уровня даже при фиксировании всех объясняющих переменных.

Таблица 1

ИСХОДНЫЕ ДАННЫЕ, ПОЛУЧЕННЫЕ ИЗ ЭЛЕКТРОННЫХ БАЗ РИЕЛТЕРОВ

№	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	У
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	25	15	430	0	0	0	0	0	1	0	1	1	0	17
2	25	14	363	1	1	1	0	0	1	1	1	1	0	62,5
3	25	14	510	1	1	1	0	0	1	1	1	1	0	70
4	25	14	428	1	1	1	0	0	1	1	1	1	0	78,5
5	19	36	745	1	1	0	1	1	1	1	1	0	0	61
6	19	94	810	1	1	0	1	1	1	1	1	0	0	77
7	35	8	150	1	1	1	0	0	1	0	1	1	1	15,5
8	35	17	250	1	1	0	0	0	1	0	1	1	1	32,5
9	35	23	350	1	1	0	0	0	1	0	1	1	1	57,5
10	15	19	370	1	1	1	0	0	1	1	1	1	0	39
11	15	120	950	1	1	1	0	0	1	1	1	1	0	127
12	10	9	400	1	1	1	1	0	1	0	1	1	0	39,6
13	10	12	500	1	1	1	1	0	1	0	1	1	0	50,2
14	10	15	520	1	1	1	1	0	1	0	1	1	0	60,4
15	10	17	570	1	1	1	1	0	1	0	1	1	0	70,4
16	12	18	300	1	1	1	1	0	1	1	1	1	0	35
17	12	28	400	1	1	1	1	0	1	1	1	1	0	61
18	12	39	500	1	1	1	1	0	1	1	1	1	0	100
19	6	14	368	1	1	1	0	0	1	1	1	0	0	31,5
20	6	14	373	1	1	1	0	0	1	1	1	0	0	29,7
21	33	12	343	1	1	1	0	0	1	0	1	1	1	25
22	33	18	406	1	1	1	0	0	1	0	1	1	1	33,5
23	33	30	665	1	1	1	0	0	1	0	1	1	1	41,3
24	31	7	91	0	0	0	0	0	0	0	1	1	1	5,2
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
25	31	10	150	0	0	0	0	0	0	0	1	1	1	12,3
26	31	12	240	0	0	0	0	0	0	0	1	1	1	17,6
27	7	16	335	1	1	1	1	0	1	1	1	1	0	20,2
28	7	16	610	1	1	1	1	0	1	1	1	1	0	37,5
29	30	4,4	230	1	1	1	1	0	1	0	1	1	0	5,3
30	30	23	250	1	1	1	1	0	1	0	1	1	0	21,5
31	23	7	481	1	1	1	0	0	1	1	1	0	0	13
32	23	15	524	1	1	1	0	0	1	1	1	0	0	56,3
33	23	28	667	1	1	1	0	0	1	1	1	0	0	93,3
34	12	12	300	1	1	1	0	0	1	0	1	1	0	18
35	12	58	750	1	1	1	0	0	1	0	1	1	0	90,3
36	20	21	420	1	1	1	0	0	1	0	1	1	0	55
37	20	45	670	1	1	1	0	0	1	0	1	1	0	75
38	20	60	1250	1	1	1	0	0	1	1	1	1	0	100
39	25	20	435	1	1	0	1	1	1	1	1	1	0	15
40	25	20	550	1	1	0	1	1	1	1	1	1	0	35
41	25	20	572	1	1	0	1	1	1	1	1	1	0	55
42	25	20	580	1	1	1	1	1	1	1	1	1	0	85
43	25	20	750	1	1	1	1	1	1	1	1	1	0	100
44	16	7	220	0	0	1	0	0	0	0	0	0	0	6,95
45	20	5	220	0	0	0	1	1	0	0	0	0	0	6,9
46	20	5	230	0	0	0	1	1	0	0	0	0	0	6,9
47	40	14,6	237	1	1	0	1	1	1	0	0	0	0	7,5
48	0	8,5	200	1	1	1	1	1	1	0	0	0	0	17
49	10	15	342	1	1	0	1	1	1	0	0	0	0	75
50	20	20	150	1	1	0	0	0	1	0	0	0	0	10,5
51	25	12	230	1	1	0	0	0	1	0	0	0	0	21
52	28	8	180	1	1	0	0	0	1	0	0	0	0	7,5
53	35	13	250	1	1	0	0	0	1	0	0	0	0	19
54	35	13	250	1	1	0	0	0	1	0	0	0	0	17,3
55	40	16	320	0	0	1	0	0	1	0	0	0	0	9
56	20	10	130	1	1	0	0	0	1	0	0	0	0	7,5
57	35	12,5	247	1	1	0	0	0	1	0	0	0	0	17,3
58	45	15	340	1	1	0	1	1	1	0	0	0	0	18,8
59	90	7	125	0	0	1	1	0	0	0	0	0	0	2,2
60	26	6	116	1	1	1	1	1	1	0	1	0	0	7
61	26	7,5	185	1	1	1	1	1	1	0	1	0	0	15
62	26	9	231	1	1	1	1	1	1	0	1	0	0	21
63	26	4	115	1	1	1	1	1	1	0	1	0	0	5,9
64	21	12	150	1	1	1	0	0	1	0	1	1	0	20
65	21	24	300	1	1	1	0	0	1	0	1	1	0	40
66	21	36	450	1	1	1	0	0	1	0	1	1	0	60
67	21	48	600	1	1	1	0	0	1	0	1	1	0	80
68	24	9	264	1	0	0	0	0	1	0	1	1	0	5
69	24	12	375	1	0	0	0	0	1	0	1	1	0	9
70	24	15	498	1	0	0	0	0	1	0	1	1	0	12
71	87	10	86	1	1	1	1	0	1	0	1	1	1	2,1
72	87	15	120	1	1	1	1	0	1	0	1	1	1	4,5
73	87	20	186	1	1	1	1	0	1	0	1	1	1	11
74	87	25	212	1	1	1	1	0	1	0	1	1	1	15,5
75	87	30	252	1	1	1	1	0	1	0	1	1	1	21
76	20	7	150	1	1	1	1	0	1	0	1	1	0	11
77	20	14	220	1	1	1	1	0	1	0	1	1	0	22
78	20	20	300	1	1	1	1	0	1	0	1	1	0	28
79	25	12,3	290	1	1	0	0	0	1	0	1	1	0	10,9
80	25	12	250	1	1	0	0	0	1	0	1	1	0	9,9

В работе [1, с. 111] рассмотрены примеры линейной множественной регрессии применительно к описанию простейших производственных процессов. Показано, что использование эконометрических моделей дает преимущества:

- во-первых, основываясь на количественных параметрах эконометрической модели, выявлять характер и направление связей между отдельными структурными элементами и факторами, формирующими объект исследования

и прогнозирования, а также влияние каждого их них на его состояние и развитие;

- во-вторых, позволяет исследователю изучать альтернативные варианты принимаемых решений, т.е. в конечном итоге создает основу для выбора наилучшего экономического решения, оптимальной тактики и стратегии.

В настоящей работе методами эконометрики оценена стоимость нового коттеджа в Новой Москве и проведен полный анализ компонентов, составляющих его стоимость. Исходные данные приведены в табл. 1.

Приведенные в табл. 1 данные взяты из:

- поз. 1-35 (http://guberny.ru/);
- поз. 36-38 (http://www.idyll.ru/);
- поз. 39-43 (http://www.imperialdom.ru/);
- поз. 44-59 (http://msk.ners.ru/);
- поз. 60-63 (http://www.ostrov-erin.ru/);
- поз. 64-67 (http://www.discontru.ru/);
- поз. 68-70 (http://www.discontru.ru/);
- поз. 71-75 (http://www.discontru.ru/);
- поз. 76-78 (http://www.discontru.ru/);
- поз.79-80 (http://www.2220444.ru/).

В таблице обозначены:

- X_1 – расстояние от коттеджного поселка до Московской кольцевой автомобильной дороги (МКАД);
- ***X2*** – площадь земельного участка;
- ***X3*** – общая площадь коттеджа;
- ***X4*** – наличие центрального водоснабжения в поселке;
- ***X5*** – наличие общих очистных сооружений;
- ***X6*** – наличие в шаговой доступности реки, пруда, бассейна;
- ***X7*** – наличие в поселке магазина;
- ***X8*** – наличие детского сада, школы;
- ***У*** – цена коттеджа в млн. руб.

В табл. 2 и 3 приведены пояснения определения качества коттеджа и общего количества коттеджей в поселке. Значение «1 дом» в табл. 3 означает, что коттедж построен в поселении, расположенном в черте Новой Москвы.

Таблица 2

ОПРЕДЕЛЕНИЕ КАЧЕСТВА КОТТЕДЖА

Качество коттеджа	X_9	X_{10}
Деревянный	0	0
Кирпичный	1	0
Элитный	1	1

Таблица 3

ОПРЕДЕЛЕНИЕ КОЛИЧЕСТВА ДОМОВ В ПОСЕЛКЕ

Число домов	X_{11}	X_{12}	X_{13}
1 дом	0	0	0
< 10	1	0	0
< 100	1	1	0
> 100	1	1	1

Таким образом, общее число параметров, принятых в обсуждаемой ниже модели, составляло 13. Отметим, что параметры $X_4 \div X_{13}$ качественные (фиктивные), принимающие всего два значения – ноль и единица. Более подробно о таких параметрах см., например, в [5, с. 74]. Обратим внимание на поз. 48 в табл. 1: расстояние до МКАД равно нулю. Это означает, что новый коттедж построен в шаговой доступности от станции метро (в Бутове). Выбор исходных данных (табл. 1) обусловлен желанием, по возможности, охватить всю территорию Новой Москвы, приближаясь к границам Калужской области.

Требуется построить модель множественной линейной регрессии:

$$y = b_0 + \Sigma b_i x_i,$$

где $i = 1, 2, \ldots, 13$.

На основании полученной модели надо определить рыночную стоимость коттеджа в Новой Москве; определить статистическую значимость модели в целом и ее коэффициентов, интерпретировать коэффициенты и сделать вывод о возможности применения полученной модели для предварительной оценки стоимости коттеджа. Для решения задачи построения модели множественной линейной регрессии воспользуемся пакетом «анализ данных» Microsoft Excel. Задачу разделим на этапы с целью получения уравнения множественной линейной регрессии, наиболее адекватно отвечающего поставленной цели.

Этап 1

С помощью пакета «анализ данных» Microsoft Excel для исходных данных (табл. 1) по методу наименьших квадратов получили уравнение множественной линейной регрессии:

$$y = -5{,}73 - 0{,}167x_1 + 0{,}37x_2 + 0{,}073x_3 - 0{,}61x_4 + 20{,}8x_5 + 10{,}29x_6 - 5{,}23x_7 + 6{,}91x_8 - 15{,}09x_9 + 6{,}64x_{10} - 10{,}73x_{11} + 12{,}74x_{12} + 0{,}24x_{13}. \quad (1)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,4,5,6,7,8,9,10,11,.12,13} = 0{,}885$, коэффициента детерминации $R^2 = 0{,}784$ и скорректированного коэффициента детерминации (зависит от числа объясняющих переменных) $\check{R}^2 = 0{,}741$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. Чтобы полученная модель множественной линейной регрессии адекватно описывала действительные процессы, необходимо, чтобы коэффициенты регрессии имели правильное экономическое толкование и были значимы.

В данной модели не соответствуют экономическому смыслу коэффициенты при x_4, x_7, x_9, x_{11} (наличие центрального водоснабжения и наличие магазина в поселке, качество коттеджа и общее количество коттеджей в поселке). Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}23;$

$|t_{b2}| = 2{,}60;$

$|t_{b3}| = 4{,}96;$

$|t_{b4}| = 0{,}04;$

$|t_{b5}| = 1{,}97;$

$|t_{b6}| = 2{,}14;$

$|t_{b7}| = 1{,}02;$

$|t_{b8}| = 1{,}03;$

$|t_{b9}| = 1{,}17;$

$|t_{b10}| = 1{,}29;$

$|t_{b11}| = 1{,}41;$

$|t_{b12}| = 2{,}11;$

$|t_{b13}| = 0{,}03.$

Критическое значение статистики Стьюдента для доверительной вероятности $\gamma = 0{,}95$, числа наблюдений $n = 80$, $t_{кр} = 1{,}99$. Видим, что большинство коэффициентов, за исключением b_2, b_3, b_6 и b_{12}, статистически незначимы.

Следовательно, полученная модель внутренне противоречива. Значимость модели в целом высокая

($F_{набл}$ = 18,46), качество модели высокое (R^2 = 0,784), однако коэффициенты регрессии интерпретируются не все.

Таблица 4

КОРРЕЛЯЦИОННАЯ МАТРИЦА ОБЪЯСНЯЮЩИХ ПЕРЕМЕННЫХ ДЛЯ ПЕРВОГО ЭТАПА ВЫЧИСЛЕНИЙ

Показатель	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}
X_1	1	-0,102	-0,352	-0,119	-0,087	0,013	0,162	-0,101	-0,109	-0,303	-0,067	0,050	0,598
X_2	-0,102	1	0,712	0,203	0,224	0,102	-0,082	-0,022	0,207	0,334	0,229	0,165	-0,067
X_3	-0,352	0,712	1	0,242	0,209	0,153	-0,089	0,053	0,273	0,567	0,330	0,211	- 0,259
X_4	-0,119	0,203	0,242	1	0,847	0,236	0,092	-0,008	0,869	0,226	0,316	0,164	-0,148
X_5	-0,087	0,224	0,209	0,847	1	0,352	0,179	0,047	0,737	0,266	0,227	0,070	-0,083
X_6	0,013	0,102	0,153	0,236	0,352	1	0,155	-0,259	0,236	0,176	0,419	0,252	-0,007
X_7	0,162	-0,082	-0,089	0,092	0,179	0,155	1	0,56	0,021	0,075	0,025	-0,027	-0,239
X_8	-0,101	-0,022	0,053	-0,008	0,047	-0,259	0,56	1	-0,055	0,142	-0,199	-0,405	-0,239
X_9	-0,109	0,207	0,273	0,869	0,737	0,236	0,021	-0,055	1	0,196	0,287	0,153	-0,207
X_{10}	-0,303	0,334	0,567	0,226	0,266	0,176	0,075	0,142	0,196	1	0,3176	0,044	-0,293
X_{11}	-0,067	0,229	0,330	0,316	0,227	0,419	0,025	-0,199	0,287	0,3176	1	0,7	0,23
X_{12}	0,050	0,165	0,211	0,164	0,070	0,252	-0,027	-0,405	0,153	0,044	0,7	1	0,328
X_{13}	0,598	-0,067	-0,259	-0,148	-0,083	-0,007	-0,097	-0,239	-0,207	-0,293	0,23	0,328	1

Предполагается, что в полученной модели присутствует мультиколлинеарность. Следовательно, необходимо уточнение модели. Исследуем модель на наличие мультиколлинеарности. Вновь используя пакет «анализ данных» Microsoft Excel, получим корреляционную матрицу (табл. 4). Видно, что наиболее сильная корреляция наблюдается между переменными x_2 и x_3, x_4 и x_5, x_4 и x_9. Эти данные свидетельствуют о присутствии мультиколлинеарности в модели.

Отметим наличие некоторых из обнаруженных корреляционных связей. Например, наиболее коррелируют переменные x_2 (площадь земельного участка) и x_3 (общая площадь коттеджа), x_4 и x_5 (центральное водоснабжение и очистные сооружения), x_4 и x_9 (центральное водоснабжение и качество коттеджа). Действительно, это согласуется со здравым смыслом: чем больше земельный участок, тем обычно больше коттедж, причем эта зависимость, как правило, прямо пропорциональна. Также обычно взаимосвязаны центральное водоснабжение и очистные сооружения, то же относится и к качеству коттеджа.

К другим методам выявления мультиколлинеарности можно отнести исследование матрицы X^tX, использование «ридж-регрессии» [3, с. 108]. (Матрица X имеет размерность 80 * 13, ее элементы берутся из табл. 1). Следует обратить внимание, что если по какому-либо из методов в модели не обнаружена мультиколлинеарность, это еще не значит, что ее в модели нет. В этом случае исследование модели желательно провести несколькими различными методами. Таким образом, в исследуемой модели есть внутренние противоречия, несмотря на хорошие показатели качества и значимости в целом. Не все коэффициенты регрессии значимы и экономически интерпретируемы.

Вывод

В модели присутствует мультиколлинеарность. В полученном виде модель линейной множественной регрессии не подходит для практического использования и нуждается в уточнении.

Существуют разные способы устранения мультиколлинеарности: способ перехода от исходных переменных, связанным между собой достаточно тесной корреляционной зависимостью, к новым переменным, представляющим линейные комбинации исходных [см., например, 4]. Можно попытаться увеличить число объясняющих переменных, в нашем случае это невозможно. Другим способом устранения (уменьшения) мультиколлинеарности является использование пошаговых процедур отбора наиболее информативных переменных. Этот способ обычно применятся, когда невозможно получить дополнительные данные.

Воспользуемся последним способом: будем оставлять в модели те факторы, которые наиболее сильно влияют на результат.

Этап 2

Анализируя статистическую значимость коэффициентов уравнения регрессии, корреляционную матрицу и знаки коэффициентов у объясняющих переменных, считаем, что переменную x_{13}, имеющую наименьшее значение статистики Стьюдента (количество коттеджей в поселке более ста), можно удалить. Для корректности процедуры удаления объясняющих переменных необходимо пошагово каждый раз удалять только одну переменную. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии:

$$y = -5{,}74 - 0{,}165x_1 + 0{,}37x_2 + 0{,}073x_3 - 0{,}658x_4 + 20{,}91x_5 + 10{,}26x_6 - 5{,}27x_7 + 6{,}916x_8 - 15{,}2x_9 + 6{,}607x_{10} - 10{,}65x_{11} + 12{,}77x_{12}. \quad (2)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,4,5,6,7,8,9,10,11,12}$ = 0,885, коэффициента детерминации R^2 = 0,784 и скорректированного коэффициента детерминации $\check{R}^2$ = 0,745. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. В данной модели не соответствуют экономическому смыслу коэффициенты при x_4, x_7, x_9, x_{11}. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}64;$

$|t_{b2}| = 2{,}62;$

$|t_{b3}| = 5{,}00;$

$|t_{b4}| = 0{,}04;$

$|t_{b5}| = 2{,}09;$

$|t_{b6}| = 2{,}18;$

$|t_{b7}| = 1{,}07;$

$|t_{b8}| = 1{,}04;$

$|t_{b9}| = 1{,}23;$

$|t_{b10}| = 1,34;$

$|t_{b11}| = 1,49;$

$|t_{b12}| = 2,16.$

Видим, что большинство коэффициентов, за исключением b_2, b_3, b_5, b_6 и b_{12}, статистически незначимы. Полученная модель не отражает реальность: коэффициенты b_4, b_7, b_9 и b_{11} имеют противоположные экономическому смыслу знаки. Также наблюдается сильная корреляция[1] объясняющих переменных x_4, x_5 и x_9. Модель нуждается в дальнейшем улучшении.

Этап 3

Удаляем переменную x_8 (наличие в коттеджном поселке детского сада или школы). Большинство коттеджных поселков не имеет детских учреждений. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии:

$$y = -4,11 – 0,18x_1 + 0,36x_2 + 0,076x_3 – 1,275x_4 + 21,74x_5 + 8,12x_6 – 1,76x_7 – 15,65x_9 + 6,416x_{10} – 9,08x_{11} + 9,945x_{12}. \quad (3)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,4,5,6,7,9,10,11,12} = 0,883$, коэффициента детерминации $R^2 = 0,78$ и скорректированного коэффициента детерминации $\bar{R}^2 = 0,745$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии.

В данной модели не соответствуют экономическому смыслу коэффициенты при x_4, x_7, x_9, x_{11}. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1,80;$

$|t_{b2}| = 2,54;$

$|t_{b3}| = 5,33;$

$|t_{b4}| = 0,09;$

$|t_{b5}| = 2,19;$

$|t_{b6}| = 1,92;$

$|t_{b7}| = 0,49;$

$|t_{b9}| = 1,27;$

$|t_{b10}| = 1,29;$

$|t_{b11}| = 1,30;$

$|t_{b12}| = 1,89.$

Видим, что большинство коэффициентов, за исключением b_2, b_3 и b_5, статистически незначимы. Полученная модель не отражает реальность: знаки коэффициентов b_4, b_7, b_9 и b_{11} не имеют экономического смысла. Также наблюдается сильная корреляция объясняющих переменных x_4, x_5 и x_9. Модель нуждается в дальнейшем улучшении.

Этап 4

Удаляем переменную x_7 (наличие в коттеджном поселке магазина). С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии

$$y = -4,62 – 0,18x_1 + 0,36x_2 + 0,076x_3 – 1,45x_4 + 21,07x_5 + 7,94x_6 – 14,95x_9 + 6,16x_{10} – 9,08x_{11} + 10,01x_{12}. \quad (4)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,4,5,6,9,10,11,12} = 0,883$, коэффициента детерминации $R^2 = 0,78$ и скорректированного коэффициента детерминации $\bar{R}^2 = 0,748$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии.

В данной модели не соответствуют экономическому смыслу коэффициенты при x_4, x_9, x_{11}. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1,92;$

$|t_{b2}| = 2,60;$

$|t_{b3}| = 5,37;$

$|t_{b4}| = 0,10;$

$|t_{b5}| = 2,15;$

$|t_{b6}| = 1,89;$

$|t_{b9}| = 1,23;$

$|t_{b10}| = 1,25;$

$|t_{b11}| = 1,31;$

$|t_{b12}| = 1,91.$

Видим, что большинство коэффициентов, за исключением b_2, b_3 и b_5, статистически незначимы. Полученная модель не отражает экономическую реальность. Также наблюдается сильная корреляция объясняющих переменных x_4, x_5 и x_9. Модель нуждается в дальнейшем улучшении.

Этап 5

Удаляем переменную x_4 (наличие в коттеджном поселке центрального водоснабжения). Центральное водоснабжение заменяется индивидуальной скважиной. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии:

$$y = -4,62 – 0,18x_1 + 0,36x_2 + 0,076x_3 + 20,44x_5 + 8,04x_6 – 15,77x_9 + 6,18x_{10} – 9,19x_{11} + 10,00x_{12}. \quad (5)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,5,6,9,10,11,12} = 0,883$, коэффициента детерминации $R^2 = 0,78$ и скорректированного коэффициента детерминации $\bar{R}^2 = 0,751$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. В данной модели не соответствуют экономическому смыслу коэффициенты при x_9, x_{11}. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1,93;$

$|t_{b2}| = 2,63;$

$|t_{b3}| = 5,40;$

$|t_{b5}| = 2,73;$

$|t_{b6}| = 1,98;$

$|t_{b9}| = 1,72;$

$|t_{b10}| = 1,26;$

$|t_{b11}| = 1,35;$

$|t_{b12}| = 1,92.$

Видим, что большинство коэффициентов, за исключением b_2, b_3 и b_5, статистически незначимы. Полученная модель все еще не отражает экономическую реальность, но корреляция между объясняющими перемен-

[1] Автором проведены вычисления корреляционной матрицы для всех этапов вычислений; в работе приведены корреляционные матрицы только для первого и последнего этапов.

ными существенно уменьшилась. Модель нуждается в дальнейшем улучшении.

Этап 6

Удаляем переменную x_9. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии

$$y = -9{,}47 - 0{,}18x_1 + 0{,}39x_2 + 0{,}071x_3 + 11{,}15x_5 + 8{,}86x_6 + 7{,}46x_{10} - 10{,}86x_{11} + 10{,}06x_{12}. \quad (6)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,5,6,10,11,12} = 0{,}877$, коэффициента детерминации $R^2 = 0{,}77$ и скорректированного коэффициента детерминации $\check{R}^2 = 0{,}745$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. В данной модели не соответствуют экономическому смыслу коэффициенты при x_{11}. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}90;$

$|t_{b2}| = 2{,}80;$

$|t_{b3}| = 5{,}10;$

$|t_{b5}| = 2{,}12;$

$|t_{b6}| = 2{,}17;$

$|t_{b10}| = 1{,}52;$

$|t_{b11}| = 1{,}59;$

$|t_{b12}| = 1{,}91.$

Полученная модель все еще не в полной мере отражает экономическую реальность: коэффициент b_{11} имеет противоположный знак, не все коэффициенты статистически значимы. Модель нуждается в дальнейшем улучшении.

Этап 7

Удаляем переменную x_{11}. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии:

$$y = -12{,}01 - 0{,}18x_1 + 0{,}39x_2 + 0{,}071x_3 + 10{,}59x_5 + 7{,}03x_6 + 5{,}32x_{10} + 4{,}25x_{12}. \quad (7)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,5,6,10,12} = 0{,}873$, коэффициента детерминации $R^2 = 0{,}762$ и скорректированного коэффициента детерминации $\check{R}^2 = 0{,}739$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости. Посмотрим значимость каждого коэффициента регрессии.

Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}83;$

$|t_{b2}| = 2{,}76;$

$|t_{b3}| = 5{,}01;$

$|t_{b5}| = 1{,}99;$

$|t_{b6}| = 1{,}77;$

$|t_{b10}| = 1{,}11;$

$|t_{b12}| = 1{,}11.$

В полученной модели не все коэффициенты статистически значимы. Модель нуждается в дальнейшем улучшении.

Этап 8

Удаляем переменную x_{12}. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии:

$$y = -10{,}87 - 0{,}17x_1 + 0{,}39x_2 + 0{,}07x_3 + 10{,}41x_5 + 8{,}03x_6 + 4{,}79x_{10}. \quad (8)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,5,6,10} = 0{,}870$, коэффициента детерминации $R^2 = 0{,}758$ и скорректированного коэффициента детерминации $\check{R}^2 = 0{,}738$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости. Посмотрим значимость каждого коэффициента регрессии.

Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}72;$

$|t_{b2}| = 2{,}74;$

$|t_{b3}| = 5{,}30;$

$|t_{b5}| = 1{,}96;$

$|t_{b6}| = 2{,}07;$

$|t_{b10}| = 1{,}01.$

В полученной модели не все коэффициенты статистически значимы. Модель нуждается в дальнейшем улучшении.

Этап 9

Удаляем переменную x_{10}. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии

$$y = -12{,}02 - 0{,}18x_1 + 0{,}37x_2 + 0{,}08x_3 + 11{,}30x_5 + 8{,}25x_6. \quad (9)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/1,2,3,5,6} = 0{,}868$, коэффициента детерминации $R^2 = 0{,}755$ и скорректированного коэффициента детерминации $\check{R}^2 = 0{,}738$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости.

Посмотрим значимость каждого коэффициента регрессии. Получена следующая статистическая значимость коэффициентов:

$|t_{b1}| = 1{,}84;$

$|t_{b2}| = 2{,}65;$

$|t_{b3}| = 6{,}34;$

$|t_{b5}| = 2{,}16;$

$|t_{b6}| = 2{,}13.$

В полученной модели не все коэффициенты статистически значимы. Модель нуждается в дальнейшем улучшении.

Этап 10

Удаляем переменную x_1. С помощью пакета «анализ данных» Microsoft Excel получили уравнение множественной линейной регрессии

$$y = -19{,}37 + 0{,}31x_2 + 0{,}09x_3 + 11{,}98x_5 + 7{,}57x_6. \quad (10)$$

Полученная модель имеет высокие величины множественного коэффициента корреляции $r_{y/2,3,5,6} = 0{,}862$, коэффициента детерминации $R^2 = 0{,}743$ и скорректированного коэффициента детерминации $\check{R}^2 = 0{,}730$. Полученная модель в целом имеет хорошие показатели коэффициентов качества и значимости $F = 54{,}432$. Посмотрим значимость каждого коэффициента регрессии.

Получена следующая статистическая значимость коэффициентов:

$|t_{b2}| = 2,24;$

$|t_{b3}| = 7,62;$

$|t_{b5}| = 2,26;$

$|t_{b6}| = 2,00.$

В полученной модели все коэффициенты статистически значимы. Используя пакет «анализ данных» Microsoft Excel получим корреляционную матрицу для последнего этапа (табл. 5).

Таблица 5

КОРРЕЛЯЦИОННАЯ МАТРИЦА ОБЪЯСНЯЮЩИХ ПЕРЕМЕННЫХ И У ДЛЯ ЭТАПА 10

Показатель	X_2	X_3	X_5	X_6	y
X_2	1	0,712	0,224	0,102	0,688
X_3	0,712	1	0,209	0,153	0,823
X_5	0,224	0,209	1	0,352	0,364
X_6	0,102	0,153	0,352	1	0,290
y	0,688	0,823	0,364	0,290	1

Представленная корреляционная матрица адекватно отражает результаты, полученные по уравнению (10). Наблюдается сильная связь объясняемой переменной (цена) с общей площадью коттеджа и площадью земельного участка. Несколько менее связь цены с факторами:

- X_5 – наличие очистных сооружений;
- X_6 – наличие в шаговой доступности реки, пруда, бассейна.

Экономическое объяснение полученного результата.

1. Из всех исходных факторов наиболее важными оказались площадь коттеджа (X_3), площадь земельного участка (X_2), наличие очистных сооружений (X_5) и наличие в шаговой доступности реки, пруда или бассейна (X_6). Наличие последних двух объяснимо: современные способы очистки стоков, конечно, хороши для индивидуального пользователя, но в масштабах коттеджного поселка это нонсенс. Также понятно желание потенциального покупателя иметь в летнее время место отдыха на воде, особенно это важно для семей с детьми.
   Современные коттеджи строятся в основном для покупателей с достатком выше среднего, для них не так важно наличие магазина, больших детских учреждений. Коттеджи строятся, как правило, из качественных материалов, поэтому и эти параметры выпали из рассмотрения. Отсутствие такого параметра, как «расстояние до МКАД», можно объяснить тем, что шаговая доступность от МКАД сопоставима с ценой за удаленность от городской агломерации в плане комфорта (отсутствие шума, лучшая экология и так далее). Кроме того, средства связи, включая Интернет, позволяют частично исполнять свои обязанности, не выезжая к месту работы.
2. Сравним полученные результаты стоимости коттеджа со стоимостью новой квартиры в Новой Москве [5, с. 103]. Из всех исследованных в этой работе факторов (общая площадь, количество комнат, площадь кухни, этаж, расстояние до МКАД и срок сдачи жилья), наиболее важными оказались площадь новой квартиры и расстояние до МКАД. Именно эти два показателя в новостройках являются определяющими (квартир с маленькими кухнями и без балконов сейчас не строят). Это легко объяснимо как логически, так и экономически. Для жителей Новой Москвы, работающих «внутри МКАД», время подъезда до места работы имеет чрезвычайно важное значение. Совпадение с настоящей работой наблюдается только по одному параметру – общая площадь жилья.
3. Выводы в работе сформулированы для конкретных исходных данных по различным законченным коттеджным поселкам Новой Москвы. Задача состояла в попытке оценить стоимость коттеджа и выявить наиболее значимые факторы его стоимости в Новой Москве.
4. Ошибка оценки стоимости коттеджа по уравнению (10) составляет около 10%. Возможной причиной этого является недостаточность, а иногда и неточность исходных данных, публикуемых риэлторами. Автором сделаны попытки уменьшения величины ошибки, но это возможно только за счет дальнейшего уменьшения числа объясняющих переменных.

Литература

1. Галочкин В.Т. Применения эконометрических методов в построении оптимальной экономической модели [Текст] / В.Т. Галочкин // Аудит и финансовый анализ. – 2013. – №6. – С. 111-112.
2. Галочкин В.Т. Прогнозирование производства метровагонов методами эконометрики [Текст] / В.Т. Галочкин // Аудит и Финансовый анализ. – 2014. – №3.
3. Галочкин В.Т. Оценка стоимости квартиры в Новой Москве [Текст] / В.Т. Галочкин // Сб. ст. Междунар. науч.-практ. конф. 31 янв. 2014 г. – Уфа, 2014. – С. 103-109.
4. Кремер Н.Ш. Эконометрика [Текст] / Н.Ш. Кремер, Б.А. Путко. – М., 2010.
5. Новиков А.И. Эконометрика [Текст] / А.И. Новиков. – М. : Дашков и К, 2013.

Ключевые слова

Эконометрика; линейная регрессия; метод наименьших квадратов; множественная линейная регрессия; факторы регрессии; корреляция; прогнозирование экономических явлений; коттедж.

Галочкин Валерий Тимофеевич

РЕЦЕНЗИЯ

Актуальность проблемы. Рецензируемая статья предлагает методами эконометрики оценку стоимости нового коттеджа в Новой Москве. За базу автор взял официальные данные риэлтеров по 80 коттеджам.

Актуальность поставленных автором задач обусловлена проблемой внедрения и практического использования в реальной экономике математических методов эконометрического прогнозирования. Теоретической основой методов эконометрического прогнозирования являются математические дисциплины (прежде всего теория вероятностей и математическая статистика, дискретная математика и исследование операций), а также экономическая теория, экономическая статистика и другие социально-экономические науки.

Среди различных моделей прогнозирования наиболее употребительными на практике являются регрессионные модели: линейные, нелинейные и системы одновременных уравнений. Однако последние трудно использовать применительно к задаче, поставленной автором. Опираясь на опубликованные данные, автор строит уравнение множественной линейной регрессии, исследует его качественные и количественные характеристики и дает прогноз оценки стоимости коттеджа в Новой Москве.

Научная новизна и практическая значимость. Построенное автором уравнение множественной линейной регрессии для предварительной оценки стоимости нового коттеджа встречается в экономической литературе нечасто. Это делает работу новой, актуальной и приближенной к реальной жизни. Безусловно, предложенная автором модель может обладать определенной схематичностью, поскольку в ней не учтены все характеристики, так или иначе влияющие на итоговый результат. Следует добавить также некоторую осторожность в использовании исходных данных, предлагаемых риэлтерами. Однако при всех недостатках таких моделей принятие решений на их основе в конечном итоге создает основу для выбора наилучшего экономического решения.

Заключение. Статья написана грамотным математическим языком, приведенный список литературы адекватно отражает содержание статьи.

Рассматриваемая статья может быть рекомендована для опубликования в журнале «Аудит и финансовый анализ».

Савинов Ю.А., д.э.н., профессор Всероссийской академии внешней торговли.