

ШАМАЕВ Б. Д.
ЭКОНОМЕТРИЧЕСКИЙ АНАЛИЗ И МОДЕЛИРОВАНИЕ
СТОИМОСТИ ЖИЛОЙ НЕДВИЖИМОСТИ

Аннотация. В статье представлены результаты эконометрического анализа рынка жилья города Саранска. Построена модель зависимости средней стоимости квартиры от общей площади, площади кухни, наличия/отсутствия балкона и этажа расположения. Представлены расчеты стоимости нескольких квартир с заданными характеристиками.

Ключевые слова: рынок жилья, эконометрический анализ, уравнение регрессии, стоимость квартиры.

SHAMAEV B. D.
ECONOMETRIC ANALYSIS AND MODELING
OF THE COST OF RESIDENTIAL REAL ESTATE

Abstract. The article presents the results of an econometric analysis of the housing market in the city of Saransk. The model of dependence of the average cost of an apartment on the total area, kitchen area, presence/absence of a balcony and floor location is constructed. The cost of several apartments with specified characteristics was calculated.

Keywords: housing market, econometric analysis, regression equation, apartment cost.

В настоящее время особо остро стоит вопрос приобретения собственного жилья, т. к. цены на квадратный метр жилой недвижимости неуклонно растут с каждым годом, а реальные доходы населения падают. Для того чтобы понимать, от каких составляющих зависит стоимость квартиры как на первичном, так и на вторичном рынке жилья в настоящее время, целесообразно использовать соответствующие статистические методы. Для изучения связи между двумя и более переменными используется корреляционно-регрессионный анализ, который позволяет определить, насколько сильна связь между изучаемыми переменными и построить эконометрическую модель в форме уравнения регрессии, определяющего зависимость результативного показателя от обуславливающих его факторных показателей [2].

В качестве источника информации стоимости квартир на первичном и вторичном рынках жилья был выбран сайт «Avito» [1]. Было изучено 100 квартир, которые составили выборку исследования. Были использованы факторы, характеризующиеся как количественными, так и фиктивными переменными. Под последними понимаются атрибутивные или качественные факторы, представленные посредством определенного цифрового кода [3]. Перечень факторов представлен на рисунке 1.

X_1 – общая площадь, м2;						
X_2 – число комнат в квартире;						
X_3 – площадь кухни, м2;						
D_4 – района города (Ленинский, Пролетарский и Октябрьский);						
D_5 – тип строения (кирпичный, панельный и монолитный);						
X_6 – год ввода здания в эксплуатацию;						
D_7 – наличие балкона (лоджии);						
D_8 – этаж в доме;						
D_9 – санузел (совмещенный/раздельный);						
X_{10} – высота потолков, м.						

В данной регрессионной модели предполагается введение следующих двух бинарных переменных D_{41} и D_{42} для признака «район города» и двух бинарных переменных D_{51} и D_{52} для признака «тип строения»:

$$D_{41} = \begin{cases} 1, \text{ Ленинский район} \\ 0, \text{ Октябрьский район} \\ 0, \text{ Пролетарский район} \end{cases}$$

$$D_{51} = \begin{cases} 1, \text{ Кирпичный} \\ 0, \text{ Монолитный} \\ 0, \text{ Панельный} \end{cases}$$

$$D_{42} = \begin{cases} 0, \text{ Ленинский район} \\ 1, \text{ Октябрьский район} \\ 0, \text{ Пролетарский район} \end{cases}$$

$$D_{52} = \begin{cases} 0, \text{ Кирпичный} \\ 1, \text{ Монолитный} \\ 0, \text{ Панельный} \end{cases}$$

Также предполагается введение фиктивной бинарной переменной D_7 для признака «наличие балкона», D_8 для признака «этаж в доме», D_9 для признака «санузел (совм./разд.)».

$$D_7 = \begin{cases} 1, \text{ присутствует} \\ 0, \text{ отсутствует} \end{cases}$$

$$D_8 = \begin{cases} 0, \text{ первый/последний} \\ 1, \text{ средний} \end{cases}$$

$$D_9 = \begin{cases} 1, \text{ совмещённый} \\ 0, \text{ раздельный} \end{cases}$$

Рис. 1. Факторы, используемые для статистического анализа рынка недвижимости.

Для того чтобы выявить тесноту связи между исследуемыми факторами целесообразно составить корреляционную матрицу коэффициентов корреляции с помощью пакета прикладных программ Excel (табл. 1). Из таблицы 1 видно, что наиболее влияющими на стоимость квартиры факторами являются её общая площадь, количество комнат и площадь кухни. Это выражается в высоких значениях коэффициентов корреляции r : $r_{x1} = 0,82$; $r_{x2} = 0,45$; $r_{x3} = 0,71$. Поскольку данные значения превышают 0,7; можно сделать вывод о высоком влиянии данных факторов на результативный признак. Также можно заметить, что между некоторыми признаками X и D присутствует мультиколлинеарность – тесная корреляционная связь между отбираемыми для анализа факторами [4]. Если значение коэффициента корреляции между ними превышает значение $|0,7|$, можно делать вывод о высокой зависимости между исследуемыми признаками.

Таблица 1

Матрица коэффициентов корреляции исследуемых признаков

	Y	X ₁	X ₂	X ₃	D ₄₁	D ₄₂	D ₅₁	D ₅₂	X ₆	D ₇	D ₈	D ₉	X ₁₀
Y	1												
X ₁	0,82	1											
X ₂	0,75	0,81	1										
X ₃	0,71	0,55	0,08	1									
D ₄₁	0,44	0,26	0,01	0,49	1								
D ₄₂	-0,32	-0,19	-0,10	-0,21	-0,49	1							
D ₅₁	0,27	0,05	-0,03	0,24	0,43	-0,05	1						
D ₅₂	-0,06	-0,06	-0,21	0,10	-0,13	0,33	-0,36	1					
X ₆	0,39	0,24	-0,21	0,71	0,35	0,01	0,05	0,31	1				
D ₇	0,18	0,16	0,23	-0,10	-0,05	-0,05	-0,14	0,05	-0,09	1			
D ₈	0,12	-0,01	0,00	0,11	-0,16	0,12	0,03	0,16	0,06	0,01	1		
D ₉	-0,28	-0,41	-0,48	-0,05	0,05	-0,04	0,15	0,18	0,12	-0,10	0,02	1	
X ₁₀	0,47	0,42	0,03	0,54	0,42	-0,17	0,26	0,03	0,49	-0,13	-0,02	0,02	1

Так как $r_{x_1x_2} = 0,81 > 0,7$, то между общей площадью квартиры и числом комнат наблюдается мультиколлинеарность. Также сильная зависимость характерна для площади кухни и годом ввода дома в эксплуатацию ($r_{x_3x_6} = 0,711 > 0,7$). Это связано с тем, что в домах более старого периода площади кухни были меньше (4 – 6,5 м²). Для современного человека площадь кухни является одним из важнейших факторов выбора квартиры.

Дальнейшее исследование следует проводить без одного из факторов из мультиколлинеарной пары. Принять решение, какой фактор следует исключить, можно исходя из сравнения парных коэффициентов корреляции факторных и результативного признака. Так $r_{yx1} = 0,824$, а $r_{yx2} = 0,555$, значит, из дальнейшего исследования исключаем X₂, т.к. $0,555 < 0,824$. Таким же образом сравним коэффициенты корреляции следующей мультиколлинеарной пары: $r_{yx3} = 0,668$, а $r_{yx6} = 0,387$, значит, из дальнейшего исследования исключаем X₆, т.к. $0,387 < 0,668$. Следовательно, дальнейшее исследование будет включать в себя оставшиеся после проверки на мультиколлинеарность факторы.

На следующем этапе статистического анализа с помощью пакета прикладных программ Excel был проведён корреляционный анализ. Полученные результаты представлены на рисунке 2.

<i>Регрессионная статистика</i>					
Множественный R	0,908				
R-квадрат	0,824				
Нормированный R	0,804				
Стандартная ошибка	725,426				
Наблюдения	100,000				
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	10,000	219118316,434	21911831,643	41,638	0,000
Остаток	89,000	46835659,096	526243,361		
Итого	99,000	265953975,530			
<i>Коэффициент стандартная ошибка статистика P-Значение Нижние 95%</i>					
Y-пересечение	-1219,204	1575,636	-0,774	0,441	-4349,958
X1	59,101	6,302	9,378	0,000	46,579
X3	61,508	19,603	3,138	0,002	22,558
D41	127,239	229,944	0,553	0,581	-329,656
D42	-561,747	204,012	-2,753	0,007	-967,114
D51	714,705	198,045	3,609	0,001	321,195
D52	352,574	257,453	1,369	0,174	-158,981
D7	656,227	257,294	2,550	0,012	144,989
D8	385,456	169,810	2,270	0,026	48,047
D9	-247,767	181,889	-1,362	0,177	-609,177
X10	223,945	628,786	0,356	0,723	-1025,439

Рис. 2. Результаты регрессионного анализа.

Можно заметить, что значение коэффициента, нормированного R, равно 0,804. Это говорит о том, что на стоимость квартиры в Саранске по состоянию на декабрь 2023 г. на 80,4% оказывают влияние факторы, представленные в полученной модели, и на 19,6% оказывают влияние другие факторы, не учтённые в исследовании.

Для итоговой модели будем использовать лишь те факторы, которые, по нашему мнению, наиболее сильно влияют на стоимость квартиры. Таким образом, в дальнейшее исследование войдут факторы, не зачёркнутые на рисунке 3.

X₂	число комнат в квартире;						
X₆	год ввода здания в эксплуатацию;						
D₄	района города (Ленинский, Пролетарский и Октябрьский);						
D₅	тип строения (кирпичный, панельный и монолитный);						
D₉	санузел (совмещенный/раздельный);						
X ₁	общая площадь, м2;						
X ₃	площадь кухни, м2;						
D ₇	наличие балкона (лоджии);						
D ₈	этаж в доме;						
X ₁₀	высота потолков, м.						

Рис.

Рис. 3. Факторы, входящие в итоговую модель.

Исходя из первоначальных данных, можно заметить, что в выборку входят выбросы, значения которых сильно отклоняются от среднего уровня ряда в большую или меньшую сторону. Для получения адекватной модели, пригодной для прогнозирования, целесообразно убрать из выборки 20 наблюдений. Таким образом, в дальнейшее исследование войдёт выборка из 80 квартир. На заключительном этапе регрессионного анализа ещё раз оценим значимость выбранных факторов с помощью регрессионного анализа (рис. 4).

<i>Регрессионная статистика</i>					
Множест	0,901				
R-квадра	0,812				
Нормиро	0,802				
Стандарт	603,372				
Наблюде	80,000				
<i>Дисперсионный анализ</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регресси	4,000	117595415,633	29398853,908	80,753	0,000
Остаток	75,000	27304300,755	364057,343		
Итого	79,000	144899716,388			
	<i>Коэффициент</i>	<i>стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>
Y-пересе	-390,379	316,560	-1,233	0,221	-1020,999
X1	55,093	5,512	9,994	0,000	44,112
X3	81,991	16,006	5,123	0,000	50,106
D7	466,172	226,469	2,058	0,043	15,022
D8	294,462	151,735	1,941	0,056	-7,810

Рис. 4. Параметры окончательной модели регрессии.

По представленным параметрам можно составить итоговое уравнение искомой регрессии. Оно имеет следующий вид:

$$Y = -390,38 + 55,01 \cdot X_1 + 81,10 \cdot X_3 + 466,17 \cdot D_7 + 294,46 \cdot D_8 + \varepsilon$$

Дадим экономическую интерпретацию полученным коэффициентам регрессии. Коэффициент a_1 показывает, что при увеличении площади квартиры на 1 м² стоимость квартиры в Саранске в среднем увеличится на 55,093 тыс. руб. при закреплении других факторов на неизменном среднем уровне. Коэффициент a_3 показывает, что при увеличении площади кухни в квартире на 1 м² стоимость квартиры в Саранске в среднем увеличится на 81,991 тыс. руб. при закреплении других факторов на неизменном среднем уровне. Коэффициент a_7 показывает, что при наличии балкона стоимость квартиры в Саранске в среднем будет выше на 466,172 тыс. руб., чем у квартиры без балкона. Коэффициент a_8 показывает, что при расположении квартиры на промежуточном этаже в среднестатистическом доме Саранска её стоимость в среднем будет выше на 294,462 тыс. руб., чем у квартиры, находящейся на первом или последнем этажах.

Вычисленный коэффициент детерминации, равный 0,8116 позволяет сделать следующий вывод. Вариация стоимости квартиры на первичном и вторичном рынках жилья в Саранске на 81,16 % обусловлена изменением общей площади квартиры, площадью кухни в квартире, наличием/отсутствием балкона и расположением квартиры на промежуточном/крайнем этаже дома. На 18,84 % результативный признак зависит от других факторов, не входящих в данное исследование.

Составим частные уравнения стоимости некоторых разновидностей исследованных квартир и проранжируем по итоговой сумме (табл. 2).

Таблица 2

Частные уравнения стоимости квартир с учётом рангов

Вид квартиры	Уравнение регрессии	Ранг
Квартира с балконом на крайнем этаже	$Y=75,79+55,01 \cdot X_1+81,10 \cdot X_3$	2
Квартира с балконом на среднем этаже	$Y=370,26+55,01 \cdot X_1+81,10 \cdot X_3$	1
Квартира без балкона на крайнем этаже	$Y=-390,38+55,01 \cdot X_1+81,10 \cdot X_3$	4
Квартира без балкона на среднем этаже	$Y=-95,92+55,01 \cdot X_1+81,10 \cdot X_3$	3

Далее осуществим расчёты стоимости квартир с заданными условиями (табл. 3).

Таблица 3

Расчёт стоимости квартир с заданными условиями

Заданные характеристики квартир	Итоговая стоимость, тыс. руб.
3-х комнатная квартира на 1-м этаже с балконом	4123,58
2-х комнатная квартира на среднем этаже с балконом	3945,59
1-х комнатная квартира на среднем этаже без балкона	2193,85

На основании таблицы 3 можно сделать вывод: квартиры с балконами на средних этажах стоят дороже, чем квартиры на первом/последнем этаже без балкона при условии соблюдения неизменности прочих факторов.

Исходя из проведённого исследования, подведём итоги. С помощью статистического метода корреляционно-регрессионного анализа мы выяснили, какие факторы и в какой степени влияют на стоимость квартиры на первичном и вторичном рынках жилья в настоящее время. По составленной регрессионной модели можно рассчитать среднерыночную стоимость квартиры с различными заданными параметрами. На основании этого можно сделать вывод, что такой эконометрический метод, как корреляционно-регрессионный анализ, может быть применимым в различных отраслях экономики, в том

числе – в исследовании рынка недвижимости. Это говорит о практической возможности применения статистических методов для изучения различных экономических явлений и сфер жизни населения.

СПИСОК ЛИТЕРАТУРЫ

1. Интернет-сервис для размещения объявлений «Avito» [Электронный ресурс]. – Режим доступа: <https://www.avito.ru> (дата обращения: 28.12.2024).
2. Марченко В. М., Можей Н. П., Шинкевич Е. А. Эконометрика: учебное пособие. – Минск: БГТУ, 2011. – 157 с.
3. Сажин Ю. В., Иванова И. А. Эконометрика: учебник. – Саранск: Мордовский гос. ун-т, 2014. – 316 с.
4. Хайруллина О. И., Баянова О. В. Эконометрика: базовый курс: учебник. – Пермь, 2019. – 288 с.