

ЭКОНОМЕТРИЧЕСКИЙ ПРОЕКТ И ЕГО РЕАЛИЗАЦИИ С ПОМОЩЬЮ СТАТИСТИЧЕСКОГО ПАКЕТА ECONOMETRIC VIEWS

В этой статье построено простейший эконометрический модели регрессию цены квартиры

ЭКОНОМЕТРИКАЛЫК ДОЛБООР ЖАНА АНЫ ECONOMETRIC VIEWS СТАТИСТИКАЛЫК ПАКЕТИНИН ЖАРДАМЫ МЕНЕН ИШКЕ АШЫРУУ

Бул макалада батирлердин регрессиялык баасынын эконометриялык модели берилет

THE ECONOMETRIC PROJECT AND ITS REALIZATION WITH THE HELP STATISTICAL ECONOMETRIC VIEWS PACKAGE

In this article, build a simple econometric regression model prices were flat

Эконометрический пакет Eviews обеспечивает особо сложный и тонкий инструментарий обработки данных, позволяет выполнять регрессионный анализ, строить прогнозы в Windows-ориентированной компьютерной среде. С помощью этого программного средства можно очень быстро выявить наличие статистической зависимости в анализируемых данных и затем, используя полученные взаимосвязи, сделать прогноз изучаемых показателей.

Целесообразно выделить следующие сферы применения Eviews:

- анализ научной информации и оценивание;
- финансовый анализ;
- макроэкономическое прогнозирование;
- моделирование;
- прогнозирование состояния рынков.

Особо широкие возможности открывает Eviews при анализе данных, представленных в виде временных рядов.

Данные по однокомнатным квартирам, анализируемые в проекте, были взяты с Интернет-сайта: <http://www.kont.ru> (агентство «Контакт»). При отборе данных уделялось внимание тому, чтобы были представлены различные районы Москвы, для чего отбор выборки производился по принципу ближайшего метро (то есть были взяты квартиры, расположенные у станций метро, названия которых начинаются с букв А-К; такие станции сравнительно равномерно разбросаны по Москве).

Первоначально было собрано порядка 500 записей, из которых часть пришлось отбросить либо из-за отсутствия некоторых данных (наличие телефона, балкона и т.д.), либо из-за содержания явно ошибочной информации (нулевая или несоизмеримо высокая цена и т.д.). В результате осталось около 400 наблюдений (см. файл data.xls). Каждая запись, описывающая отдельную квартиру, содержит следующие поля:

Price – цена в 1\$

Totsp – общая площадь (м²)

Kitsp – площадь кухни (м²)

Livsp – площадь жилых помещений (м²)

Dist – расстояние до центра (км)

Metrdist – расстояние до метро (мин.)

Walk – 1: до метро пешком, 0: в противном случае

Brick – 1: дом кирпичный, 0: в противном случае

Tel – 1: телефон есть, 0: в противном случае

Bal – 1: балкон есть, 0: в противном случае

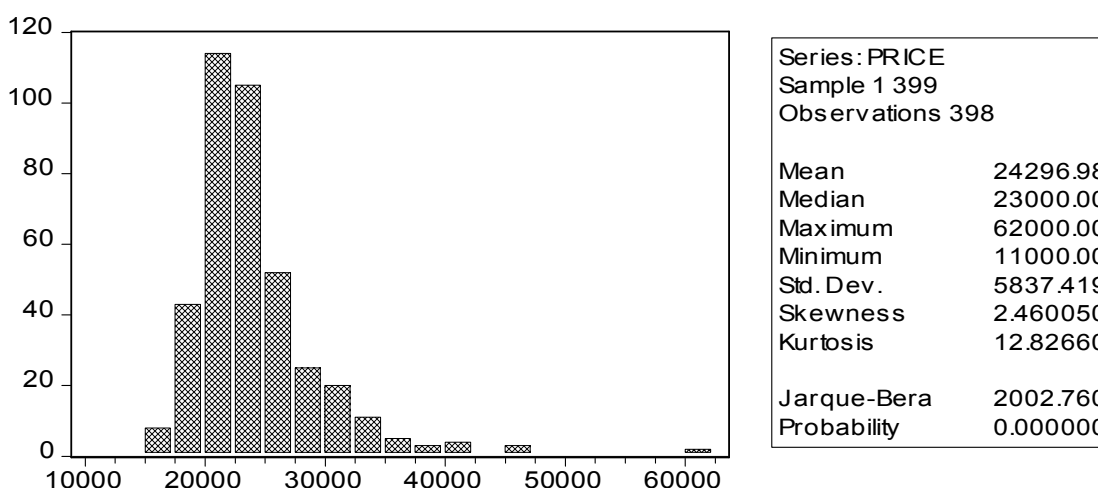
Floor – 1: если не первый или последний этаж, 0: в противном случае.

Обработка полученных данных была проведена с помощью статистического пакета Econometric Views.

Сначала были найдены стандартные статистики, по которым можно составить первоначальное представление об исследуемой выборке (см Таблицу 1 и гистограммы). Как видно из таблицы, средняя цена однокомнатной квартиры составляет 24300\$, причем разброс цен достаточно велик – от 11000\$ до 62000\$. Средняя жилая площадь в рассматриваемой выборке составляет 19 кв.м., разброс также велик – от 10 кв.м. до 28 кв.м. Рассматриваемая выборка содержит данные о квартирах как вблизи, так и вдали от центра (от 2.65 до 17.34 км). Около 70% квартир имеют балкон. В большинстве квартир (81%) имеется телефон. Чуть меньше чем из половины квартир до метро надо добираться на транспорте.

	PRICE	TOTSP	LIVSP	KITSP	DIST	METRDIST
Mean	24296.98	35.87940	18.95477	8.185176	9.932736	8.582915
Median	23000.00	36.00000	19.00000	8.500000	10.47500	10.00000
Maximum	62000.00	56.00000	28.00000	17.00000	17.34800	20.00000
Minimum	11000.00	24.00000	10.00000	4.600000	2.654000	1.000000
Std. Dev.	5837.419	4.408091	1.680520	1.963099	3.745408	3.686986
Skewness	2.460050	0.930911	-0.502625	0.632886	-0.014358	0.436759
Kurtosis	12.82660	6.077857	7.325501	4.171747	2.046812	2.755857
	BAL	BRICK	TEL	WALK	FLOOR	
Mean	0.703518	0.314070	0.816583	0.582915	0.824121	
Median	1.000000	0.000000	1.000000	1.000000	1.000000	
Maximum	1.000000	1.000000	1.000000	1.000000	1.000000	
Minimum	0.000000	0.000000	0.000000	0.000000	0.000000	
Std. Dev.	0.457281	0.464729	0.387495	0.493698	0.381197	
Skewness	-0.891241	0.801171	-1.636053	-0.336315	-1.702683	
Kurtosis	1.794310	1.641875	3.676670	1.113108	3.899129	

Таблица 1.



На основании такого беглого осмотра можно сделать вывод, что исследуемая выборка имеет довольно неплохое качество, и можно надеяться, что выводы, полученные в результате ее дальнейшей, более детальной обработки, будут реально отражать особенности рынка однокомнатных квартир в Москве.

В качестве первого шага рассмотрим корреляционную матрицу для исследуемых переменных:

	PRICE	TOTSP	LIVSP	KITSP	DIST	METRDIST
PRICE	1.000000	0.489103	0.197391	0.476101	-0.466783	-0.114823
TOTSP	0.489103	1.000000	0.475302	0.758995	0.041884	0.113910
LIVSP	0.197391	0.475302	1.000000	0.271764	0.069713	0.000607
KITSP	0.476101	0.758995	0.271764	1.000000	-0.056649	0.150216
DIST	-0.466783	0.041884	0.069713	-0.056649	1.000000	0.040311
METRDIST	-0.114823	0.113910	0.000607	0.150216	0.040311	1.000000
BAL	0.179427	0.074688	0.057897	-0.044473	-0.135166	0.029557
BRICK	0.133499	-0.238447	-0.313970	-0.241994	-0.333848	-0.132107
FLOOR	0.133335	0.032316	0.038668	-0.038163	-0.053142	-0.143728
TEL	0.053875	-0.181094	0.114877	-0.082724	-0.031654	0.000975
WALK	0.225675	-0.116925	-0.110838	-0.062794	-0.184280	-0.135940
	BAL	BRICK	FLOOR	TEL	WALK	
PRICE	0.179427	0.133499	0.133335	0.053875	0.225675	
TOTSP	0.074688	-0.238447	0.032316	-0.181094	-0.116925	
LIVSP	0.057897	-0.313970	0.038668	0.114877	-0.110838	
KITSP	-0.044473	-0.241994	-0.038163	-0.082724	-0.062794	
DIST	-0.135166	-0.333848	-0.053142	-0.031654	-0.184280	
METRDIST	0.029557	-0.132107	-0.143728	0.000975	-0.135940	
BAL	1.000000	0.048127	0.234763	0.104580	0.008747	
BRICK	0.048127	1.000000	0.028223	0.040944	0.177148	
FLOOR	0.234763	0.028223	1.000000	-0.031363	0.024146	
TEL	0.104580	0.040944	-0.031363	1.000000	0.033612	
WALK	0.008747	0.177148	0.024146	0.033612	1.000000	

Из неё видно, что цена квартиры достаточно сильно связана с величиной общей площади квартиры, с площадью кухни и расстоянием до центра, причем, как и следовало ожидать, корреляция с расстоянием до центра отрицательная. Также следует отметить высокие выборочные коэффициенты корреляции между Totsp, Livsp и Kitsp. Это и не удивительно, т.к. естественно ожидать, что между этими переменными существует приблизительная зависимость: в больших по площади квартирах делают как более просторные кухни, так и увеличивают жилую площадь. Также не вызывает удивления и

тот факт, что цена отрицательно коррелирована со временем, которое требуется для того, чтобы добраться до метро. Интуитивно очевидно, что квартиры, расположенные вблизи метро должны стоить дороже; это также подтверждается положительной корреляцией переменных Price и Walk.

Таким образом, корреляционная матрица еще раз подтверждает неплохое качество рассматриваемой выборки и говорит в пользу того, что можно перейти к более детальному исследованию имеющихся наблюдений.

Простейшая модель

В качестве простейшей эконометрической модели рассмотрим регрессию цены квартиры на все имеющиеся переменные:

$$PRICE = C(1) + C(2)*TOTSP + C(3)*LIVSP + C(4)*KITSP + C(5)*DIST + C(6)*METRDIST + C(7)*TEL + C(8)*BAL + C(9)*BRICK + C(10)*WALK + C(11)*FLOOR$$

Оценки, полученные методом наименьших квадратов, имеют следующий вид

Dependent Variable: PRICE

Method: Least Squares

Date: 02/22/01 Time: 20:04

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

Variable	Coefficien	Std. Error	t-Statistic	Prob.
	t			
C	2658.624	2690.136	0.988286	0.3236
TOTSP	566.8810	79.47475	7.132844	0.0000
LIVSP	-18.66373	141.7399	-0.131676	0.8953
KITSP	596.6298	161.4919	3.694486	0.0003
DIST	-602.2474	56.67157	-10.62698	0.0000
METRDIST	-213.0281	54.18210	-3.931706	0.0001
TEL	1824.812	530.2565	3.441376	0.0006
BAL	944.1807	448.8724	2.103450	0.0361
BRICK	1183.404	478.5794	2.472743	0.0138
WALK	2071.598	403.7721	5.130612	0.0000
FLOOR	1026.588	526.7818	1.948791	0.0520
R-squared	0.580513	Mean dependent var	24296.98	
Adjusted R-squared	0.569674	S.D. dependent var	5837.419	
S.E. of regression	3829.302	Akaike info criterion	19.36600	
Sum squared resid	5.67E+09	Schwarz criterion	19.47618	
Log likelihood	-3842.834	F-statistic	53.55565	
Durbin-Watson stat	1.276260	Prob(F-statistic)	0.000000	

Из таблицы видно, что переменная livsp является незначимой, более того, оценка коэффициента перед ней отрицательна. На первый взгляд этот результат кажется довольно странным, так как из эвристических соображений было бы естественно ожидать, что цена существенным образом зависит от жилой площади квартиры и эта зависимость положительная, в то время как проведенная регрессия свидетельствует об обратном. Однако, все становится на свои места, если принять во внимание тесную линейную зависимость между переменными totsp, livsp и kitsp. В результате этого в нашей модели появляется проблема мультиколлинеарности, которая и приводит к незначимости коэффициента перед livsp. Подтверждением наличия мультиколлинеарности можно также

считать и высокое значение F-статистики (т.е. модель в целом является значимой) при малой значимости и больших стандартных ошибках некоторых коэффициентов.

Чтобы преодолеть эту проблему, введем новую переменную $dopsp = totsp - livsp - kitsp$ и рассмотрим модель

$$PRICE = C(1) + C(2)*LIVSP + C(3)*KITSP + C(4)*DOPSP + C(5)*DIST + C(6)*METRDIST + C(7)*TEL + C(8)*BAL + C(9)*BRICK + C(10)*WALK + C(11)*FLOOR.$$

Результаты регрессии имеют вид

Dependent Variable: PRICE

Method: Least Squares

Date: 02/22/01 Time: 20:24

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

Variable	Coefficien t	Std. Error	t-Statistic	Prob.
C	2658.624	2690.136	0.988286	0.3236
LIVSP	548.2172	126.6886	4.327283	0.0000
KITSP	1163.511	114.3304	10.17674	0.0000
DOPSP	566.8810	79.47475	7.132844	0.0000
DIST	-602.2474	56.67157	-10.62698	0.0000
METRDIST	-213.0281	54.18210	-3.931706	0.0001
TEL	1824.812	530.2565	3.441376	0.0006
BAL	944.1807	448.8724	2.103450	0.0361
BRICK	1183.404	478.5794	2.472743	0.0138
WALK	2071.598	403.7721	5.130612	0.0000
FLOOR	1026.588	526.7818	1.948791	0.0520
R-squared	0.580513	Mean dependent var	24296.98	
Adjusted R-squared	0.569674	S.D. dependent var	5837.419	
S.E. of regression	3829.302	Akaike info criterion	19.36600	
Sum squared resid	5.67E+09	Schwarz criterion	19.47618	
Log likelihood	-3842.834	F-statistic	53.55565	
Durbin-Watson stat	1.276260	Prob(F-statistic)	0.000000	

Модель 1.

В данной регрессионной модели почти все переменные значимы на 5% уровне, причём значение F-статистики осталось на прежнем высоком уровне. Поэтому можно с большой долей уверенности говорить о том, что проблема мультиколлинеарности решена.

Проверка на гетероскедастичность

Чтобы проверить качество построенной модели, протестируем ее на гетероскедастичность.

White Heteroskedasticity Test:

F-statistic	5.641630	Probability	0.000000
Obs*R-squared	72.17910	Probability	0.000000

Из приведенной таблицы видно, что в рассматриваемой модели гетероскедастичность есть. Это означает, что оценки, полученные выше для стандартных ошибок коэффициентов регрессии, посчитаны по неверным формулам, а, следовательно, и

все тесты, в которых эти ошибки использовались, дают неверные результаты. Однако оценки метода наименьших квадратов, как и в случае гомоскедастичной модели, являются несмещенными. Для оценок же дисперсий коэффициентов воспользуемся оценками в форме Уайта, что частично снимает проблему гетероскедастичности.

Dependent Variable: PRICE

Method: Least Squares

Date: 02/22/01 Time: 20:37

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficien t	Std. Error	t-Statistic	Prob.
C	2658.624	3494.077	0.760895	0.4472
DOPSP	566.8810	110.9675	5.108530	0.0000
LIVSP	548.2172	163.4696	3.353634	0.0009
KITSP	1163.511	155.1793	7.497846	0.0000
DIST	-602.2474	67.34220	-8.943091	0.0000
METRDIST	-213.0281	50.14988	-4.247829	0.0000
TEL	1824.812	557.6384	3.272393	0.0012
BAL	944.1807	485.5144	1.944702	0.0525
BRICK	1183.404	549.4526	2.153787	0.0319
FLOOR	1026.588	465.3781	2.205922	0.0280
WALK	2071.598	338.0945	6.127274	0.0000
R-squared	0.580513	Mean dependent var	24296.98	
Adjusted R-squared	0.569674	S.D. dependent var	5837.419	
S.E. of regression	3829.302	Akaike info criterion	19.36600	
Sum squared resid	5.67E+09	Schwarz criterion	19.47618	
Log likelihood	-3842.834	F-statistic	53.55565	
Durbin-Watson stat	1.276260	Prob(F-statistic)	0.000000	

Из таблицы видно, что произошло изменение стандартных ошибок, а следовательно, и t-статистик. Теперь незначимым на 5% уровне является коэффициент перед bal (а также перед константой).

Еще одним способом борьбы с гетероскедастичностью является двухшаговая процедура оценки. Однако для ее использования нужно иметь априорное знание о том, от чего зависят дисперсии ошибок. Мы рассмотрели регрессии квадратов остатков исходной модели на различные комбинации регрессоров в надежде, что найдется один или несколько факторов, от которых значимо зависят дисперсии ошибок. Однако R^2 в таких регрессиях оказался крайне низким (порядка нескольких сотых), а также очень низкие F-статистики (порядка нескольких единиц). Более того, применение взвешенного метода наименьших квадратов, где в качестве весов были взяты найденные оценки для $\hat{\sigma}_i$, гетероскедастичность, как показал тест Уайта, не устранило. Это видимо означает, что дисперсии ошибок зависят от каких-то факторов, которые мы не рассматриваем в качестве регрессоров в исходной модели. В связи с этим мы ограничимся использованием МНК-оценок с исправленными по методу Уайта оценками дисперсий.

Выбор модели

Прежде всего заметим, что с помощью имеющихся данных можно попытаться получить более богатую информацию по сравнению с тем, что даёт предыдущая модель.

В наших данных квартиры на первом и последнем этажах объединены в одну категорию. При этом, однако, квартиры на первом этаже часто не имеют балкона.

Попытаемся учесть это обстоятельство, построив регрессию, в которой переменная floor заменена на две переменные: $(1-bal)*floor$ и $bal*floor$. В результате получим:

Dependent Variable: PRICE

Method: Least Squares

Date: 02/23/01 Time: 14:00

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2313.579	3480.108	0.664801	0.5066
DOPSP	566.7114	110.1797	5.143518	0.0000
LIVSP	532.9353	161.4909	3.300095	0.0011
KITSP	1170.097	154.6508	7.566060	0.0000
DIST	-602.1330	67.27358	-8.950512	0.0000
METRDIST	-218.1824	49.81084	-4.380219	0.0000
TEL	1756.691	559.7023	3.138617	0.0018
BRICK	1217.729	551.1492	2.209436	0.0277
BAL	2382.707	782.3125	3.045723	0.0025
(1-BAL)*FLOOR	1996.267	741.8912	2.690781	0.0074
BAL*FLOOR	162.1983	581.8945	0.278742	0.7806
WALK	2066.963	336.0001	6.151675	0.0000
R-squared	0.583821	Mean dependent var	24296.98	
Adjusted R-squared	0.571961	S.D. dependent var	5837.419	
S.E. of regression	3819.113	Akaike info criterion	19.36311	
Sum squared resid	5.63E+09	Schwarz criterion	19.48331	
Log likelihood	-3841.259	F-statistic	49.22596	
Durbin-Watson stat	1.294102	Prob(F-statistic)	0.000000	

Видно, что коэффициент перед $(1-bal)*floor$ значим, тогда как гипотеза о том, что коэффициент перед $bal*floor$ равен нулю, не отвергается на уровне доверия 78%. В содержательных терминах это может означать следующее. Пусть балкона нет ($bal=0$). Тогда рост цены квартиры при изменении переменной floor довольно велик, значим и в среднем равен коэффициенту перед $(1-bal)*floor$. С другой стороны, если балкон имеется (в этом случае квартира вряд ли может находиться на первом этаже), то изменение floor практически не сказывается на цене квартиры (коэффициент перед $bal*floor$ незначим). Другими словами, стоимость квартиры, имеющей балкон, мало зависит от того, расположена ли она на последнем или на одном из промежуточных этажей. Это наводит на мысль о том, что квартиры на первом этаже стоят дешевле квартир на всех остальных этажах (включая последний). Таким образом, объединение квартир, находящихся на первом и последнем этажах, в одну ценовую группу не всегда может оказаться целесообразным.

До сих пор мы использовали одну из самых простых моделей. В надежде получить оценки более высокого качества, был осуществлён перебор целого ряда моделей, включающих помимо исходных переменных их логарифмы и различные кросс-члены. В результате была найдена еще одна неплохая регрессия:

$$\text{LOG}(\text{PRICE}) = \text{C}(1) + \text{C}(2)*\text{LOG}(\text{DOPSP}) + \text{C}(3)*\text{LOG}(\text{LIVSP}) + \text{C}(4)*\text{LOG}(\text{KITSP}) + \text{C}(5)*(\text{BRICK}*\text{LOG}(\text{DIST})) + \text{C}(6)*((1-\text{BRICK})*\text{LOG}(\text{DIST})) + \text{C}(7)*\text{METRDIST} + \text{C}(8)*\text{TEL} + \text{C}(9)*\text{WALK} + \text{C}(10)*\text{BAL} + \text{C}(11)*\text{FLOOR} + \text{C}(12)*\text{BRICK}$$

Dependent Variable: LOG(PRICE)

Method: Least Squares

Date: 02/23/01 Time: 12:05

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.471434	0.237958	35.60061	0.0000
LOG(DOPSP)	0.134868	0.021040	6.410064	0.0000
LOG(LIVSP)	0.286217	0.077746	3.681447	0.0003
LOG(KITSP)	0.356533	0.031501	11.31801	0.0000
BRICK*LOG(DIST)	-0.281484	0.026884	-10.47051	0.0000
(1-BRICK)*LOG(DIST)	0.165158	0.020708	-7.975658	0.0000
METRDIST	-0.007277	0.001862	-3.907592	0.0001
TEL	0.051839	0.017721	2.925342	0.0036
WALK	0.070531	0.013923	5.065874	0.0000
BAL	0.049158	0.015295	3.213962	0.0014
FLOOR	0.043373	0.017967	2.413999	0.0162
BRICK	0.270324	0.074087	3.648739	0.0003
R-squared	0.610978	Mean dependent var	10.07498	
Adjusted R-squared	0.599892	S.D. dependent var	0.206460	
S.E. of regression	0.130594	Akaike info criterion	-1.203757	
Sum squared resid	6.583170	Schwarz criterion	-1.083563	
Log likelihood	251.5477	F-statistic	55.11192	
Durbin-Watson stat	1.337755	Prob(F-statistic)	0.000000	

Поясним содержательный смысл коэффициентов перед кросс-членами. Коэффициент перед $\text{brick}*\log(\text{dist})$ показывает, на сколько процентов в среднем уменьшается стоимость квартиры в кирпичном доме при увеличении расстояния до центра на 1%. Коэффициент перед $(1-\text{brick})*\log(\text{dist})$ даёт аналогичную характеристику для панельных домов.

К сожалению, в этой модели также присутствует сильная гетероскедастичность, о чём свидетельствует тест Уайта:

White Heteroskedasticity Test:

F-statistic	4.821314	Probability	0.000000
Obs*R-squared	70.61400	Probability	0.000000

Поэтому, как и в случае первой модели, используем оценки дисперсий оценок в форме Уайта

Dependent Variable: LOG(PRICE)

Method: Least Squares

Date: 02/23/01 Time: 11:59

Sample(adjusted): 1 213 215 399

Included observations: 398 after adjusting endpoints

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.471434	0.301886	28.06166	0.0000
LOG(DOPSP)	0.134868	0.028111	4.797619	0.0000
LOG(LIVSP)	0.286217	0.096425	2.968279	0.0032
LOG(KITSP)	0.356533	0.036436	9.785244	0.0000
BRICK*LOG(DIST)	-0.281484	0.038941	-7.228513	0.0000
(1- BRICK)*LOG(DIS T)	0.165158	0.020685	-7.984579	0.0000
METRDIST	-0.007277	0.001872	-3.886980	0.0001
TEL	0.051839	0.020831	2.488491	0.0132
WALK	0.070531	0.012747	5.532954	0.0000
BAL	0.049158	0.016778	2.929925	0.0036
FLOOR	0.043373	0.018340	2.364971	0.0185
BRICK	0.270324	0.095843	2.820474	0.0050
R-squared	0.610978	Mean dependent var	10.07498	
Adjusted R-squared	0.599892	S.D. dependent var	0.206460	
S.E. of regression	0.130594	Akaike info criterion	-	1.203757
Sum squared resid	6.583170	Schwarz criterion	-	1.083563
Log likelihood	251.5477	F-statistic	55.11192	
Durbin-Watson stat	1.337755	Prob(F-statistic)	0.000000	

Модель 2.

Тестирование гипотез

С помощью рассмотренных моделей 1 и 2 можно протестировать ряд гипотез.

Гипотеза 1. Стоимости жилой и дополнительной площади одинаковы. Применение стандартного теста в модели 1 дает

Wald Test:

Equation: GEN REGR W

Null C(2)=C(3)

Hypothesis:

F-statistic	0.016511	Probability	0.897825
Chi-square	0.016511	Probability	0.897759

Очевидным образом гипотеза не отвергается.

Гипотеза 2. Стоимость 1 кв.м. кухни в два раза больше стоимости 1 кв.м. жилой площади.

Тест, проведенный в модели 1, дает

Wald Test:

Equation: GEN REGR W

Null $2 * C(2) = C(4)$

Hypothesis:

F-statistic	0.010829	Probability	0.917174
-------------	----------	-------------	----------

Chi-square	0.010829	Probability	0.917120
------------	----------	-------------	----------

Эта гипотеза также не отвергается.

Гипотеза 3. Влияние на цену расстояния до центра одинаково для кирпичных и некирпичных домов. Для тестирования этой гипотезы хорошо приспособлена модель 2:

Wald Test:

Equation: LOG REGR W

Null $C(5) = C(6)$

Hypothesis:

F-statistic	7.013650	Probability	0.008421
-------------	----------	-------------	----------

Chi-square	7.013650	Probability	0.008089
------------	----------	-------------	----------

Из результатов тестирования видно, что гипотеза отвергается на 1% уровне.

Анализ результатов

Проанализируем полученные результаты. Коэффициент при $livsp$ в регрессии 1 показывает, что в среднем цена за 1 квадратный метр жилой площади составляет при прочих равных примерно 550\$, что также равно и стоимости одного метра дополнительной площади (следует из справедливости гипотезы 1). В то же время, цена 1 кв. метра площади кухни примерно вдвое выше (гипотеза 2). В логарифмической регрессии коэффициенты перед $\log(livsp)$ и $\log(dopsp)$ равны соответственно 0.28 и 0.13. Это означает, что при увеличении жилой площади на 1% цена квартиры увеличивается в среднем на 0.28%. Учитывая то, что средняя жилая площадь в рассматриваемой выборке составляет 19 кв.м., а средняя цена квартиры 24300\$, получаем, что цена одного квадратного метра в среднем составляет 360\$. Видно, что согласие между регрессиями не самое хорошее. Это может быть связано с тем, что, с одной стороны, разброс квартир по величине жилой площади весьма велик, и, с другой стороны, стандартные ошибки оценок коэффициентов достаточно большие. Из результатов регрессии 1 видно, что каждый километр от центра уменьшает стоимость квартиры в среднем на 600\$. Логарифмическая регрессия позволяет более детально проанализировать этот вопрос. Коэффициент -0.28 в регрессии (2) перед комбинацией переменных $brick * \log(dist)$ говорит об уменьшении цены квартир в кирпичных домах на 850\$ при увеличении расстояния на 1 км. Аналогичный показатель для некирпичных домов составляет 365\$. Заметим, что по крайней мере качественно результаты регрессий неплохо согласуются. Из логарифмической регрессии следует, что наличие телефона, возможность добраться до метро пешком, а также то, что квартира расположена не на первом или последнем этажах, увеличивает стоимость квартиры в среднем примерно на 5% (естественно, каждый из факторов по отдельности). Это составляет 1200\$, что, в целом, неплохо согласуется с результатами регрессии (1). Логарифмическая модель показывает, что при увеличении времени, требуемого для того, чтобы добраться до метро, на 1 минуту, стоимость квартиры уменьшается в среднем на 0.007%. Отсюда получаем, что каждая дополнительная минута пути до метро при прочих равных в среднем уменьшает цену квартиры на 170\$, что не противоречит результатам первой модели (200\$).

Литература:

1. Кристофер Доугерти Введение в эконометрику. Москва, 2011
2. Бердимуратов А.М. Эконометрика. Методическое пособие для ст.очного отделения,Бишкек,МАУПФиБ,2012
3. Бердимуратов А.М. Эконометрика. Методическое пособие для ст.заочного
4. отделения,Бишкек, МАУПФиБ ,2012
5. ЭНДРЮ Ф.СИГЕЛ Практическая бизнес-СТАТИСТИКА,четвертое издание,Москва,2002