

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение
высшего профессионального образования
«Оренбургский государственный университет»

Кафедра математических методов и моделей в экономике

О.С. БРАВИЧЕВА, О.И. СТЕБУНОВА

ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ПАКЕТЕ EViews

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ЛАБОРАТОРНОМУ ПРАКТИКУМУ И
САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ

Рекомендовано к изданию Редакционно-издательским советом
государственного образовательного учреждения высшего профессионального
образования «Оренбургский государственный университет»

Оренбург 2005

УДК 519.86 (07)

ББК 22.172 я7

Б 87

Рецензент

кандидат технических наук, доцент А.Г. Реннер

Б 87 Бравичева О.С., Стебунова О.И.
Эконометрическое моделирование в пакете EViews: Методические указания к лабораторному практикуму и самостоятельной работе студентов. – Оренбург: ГОУ ОГУ, 2005. – 33 с.

Методические указания содержат описание работы по оцениванию параметров эконометрической модели и ее исследованию в пакете EViews.

Методические указания могут использоваться для выполнения лабораторных работ и самостоятельной работы студентов специальностей 061800, 061700 и других специальностей по дисциплинам «Эконометрика» и «Эконометрическое моделирование».

ББК 22.172 я7

© Стебунова О.И., 2005

© Бравичева О.С., 2005

© ГОУ ОГУ, 2005

Содержание

Введение	4
1 Постановка задачи	5
2 Ввод и первичный анализ данных	5
3 Построение линейной модели множественной регрессии	19
4 Исследование модели на гетероскедастичность и автокорреляцию	24
5 Двухшаговая процедура устранения гетероскедастичности	30
Список использованных источников	33

Введение

Получившие широкое распространение в учебном процессе диалоговая система Stadia и статистический пакет Statistica хотя и позволяют решать задачи эконометрического моделирования, но в недостаточной степени приспособлены к исследованию, например, регрессионной модели на гетероскедастичность и автокорреляцию. Эконометрический пакет EViews позволяет исследовать более широкий круг задач эконометрического моделирования.

Целью данных методических указаний является освоение студентами практики построения и статистического исследования эконометрической модели в пакете EViews.

1 Постановка задачи

На основании данных файла Flat.xls, содержащего информацию о стоимости однокомнатных квартир в Москве в 1998, оценить параметры линейной модели множественной регрессии, провести статистическое исследование модели, в том числе на гетероскедастичность и автокорреляцию.

2 Ввод и первичный анализ данных

Запуск эконометрического пакета **EViews** осуществляется либо через кнопку «Пуск» на панели инструментов и последующего выбора пунктов меню «Программы»/«EViews 3»/«EViews 3.1», либо через ярлык к программе EViews на рабочем столе. После запуска на экране появится стартовое окно пакета, представленное на рисунке 1.

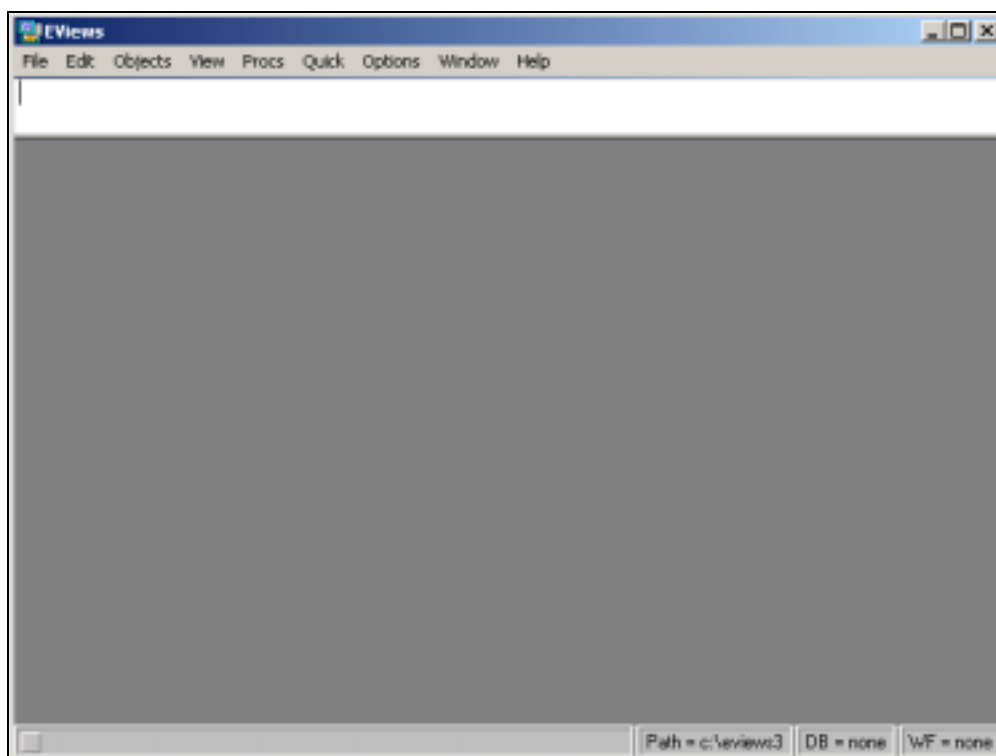


Рисунок 1 – Стартовое окно пакета EViews

Рассмотрим структуру стартового окна. Если окно, содержащее пакет, является активным, то первая строка экрана (Title Bar) будет темнее остальных. Она содержит стандартные кнопки «свернуть», «распахнуть», «закрыть». При переключении в другое окно цветовая окраска данной строки изменит цвет на более приглушенный (серый). Ниже следует строка основного меню. Принцип работы меню прост – при выборе пункта меню появляется вложенное подменю. Доступные для выбора пункты являются активными, те пункты, с которыми в настоящий момент работа невозможна, являются неактивными.

Ниже основного меню располагается командная строка (окно). В нем происходит непосредственный набор команд, которые выполняются после нажатия клавиши **Enter**. Многие команды можно выполнить через соответствующие пункты основного меню.

Большая часть окна пакета отведена под рабочую область. В ней размещаются рабочие объекты. Переключение между ними осуществляется нажатием клавиши **F6**. Последняя область экрана показывает текущее состояние пакета (рабочий каталог, текущий файл и др.).

Завершение работы с пакетом осуществляется путем выбора пункта **File** основного меню, подпункта **Exit** выпадающего роруп-меню. Система предложит сохранить имеющиеся данные. Если имя файла не было задано ранее, автоматически будет предложено имя **UNTITLED**. Его можно изменить на любое другое.

Пакет имеет обширную справочную систему (пункт основного меню **Help**).

Рассмотрим работу пакета на примере. Создадим новый рабочий файл (workfile). Для этого выберем пункты меню **File/New/Workfile**. Вид экрана при этом представлен на рисунке 2.

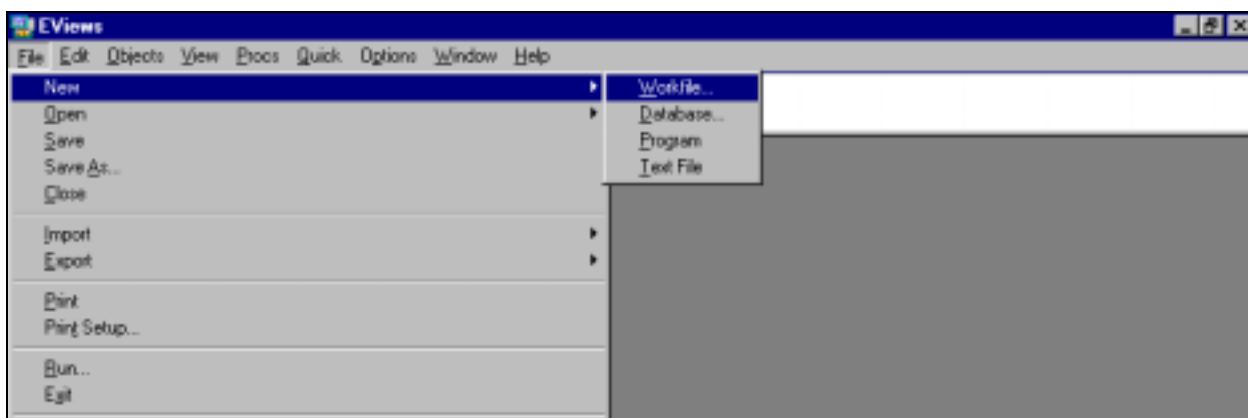


Рисунок 2 – Создание рабочего файла

После выбора указанных пунктов на экране появится диалоговое окно, представленное на рисунке 3. В данном окне задается тип вводимых данных.

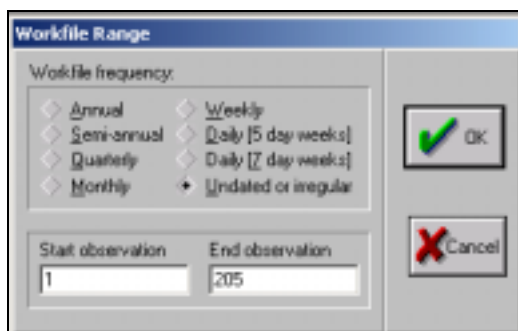


Рисунок 3 – Задание типа данных

В пакете допускается восемь типов данных:

- годовые (Annual) – годы 20 века идентифицируются по последним двум цифрам (97 эквивалентно 1997), для данных, относящихся к 21 веку необходима полная идентификация (например, 2020);
- полугодовые (Semi-annual) – 1999:1, 2001:2 (формат – год и номер полугодия);
- квартальные (Quarterly) – 1992:1, 65:4, 2005:3 (формат – год и номер квартала);
- ежемесячные (Monthly) – 1956:1, 1990:11 (формат – год и номер месяца);
- недельные (Weekly) и дневные (5/7 day weeks) – допускаются форматы Месяц/День/Год (по умолчанию) и День/Месяц/Год. Настроить эту опцию можно с помощью пунктов меню **Options/Frequency Conversion & Date Display**. Так, введенные числа 8:10:97 будут интерпретированы как Август, 10, 1997. Для установки, принятой в Европе, начальная дата будет выглядеть как Октябрь, 8, 1997;
- недатированные или нерегулярные (Undated or irregular) – допускают работу с данными, строго не привязанными к определенным временным периодам.

Важным является указание начальной (start) и конечной (end) даты/наблюдения (date/observation).

В рассматриваемом примере начальным наблюдением является первый объект, конечным – объект с номером 205.

После нажатия на кнопку **OK** на форме **Workfile Range** пакет создаст рабочий файл без имени, и на экране в рабочей области появится окно, изображенное на рисунке 4. Все рабочие файлы пакета всегда содержат вектор коэффициентов **C** и переменную (серию) **RESID**.

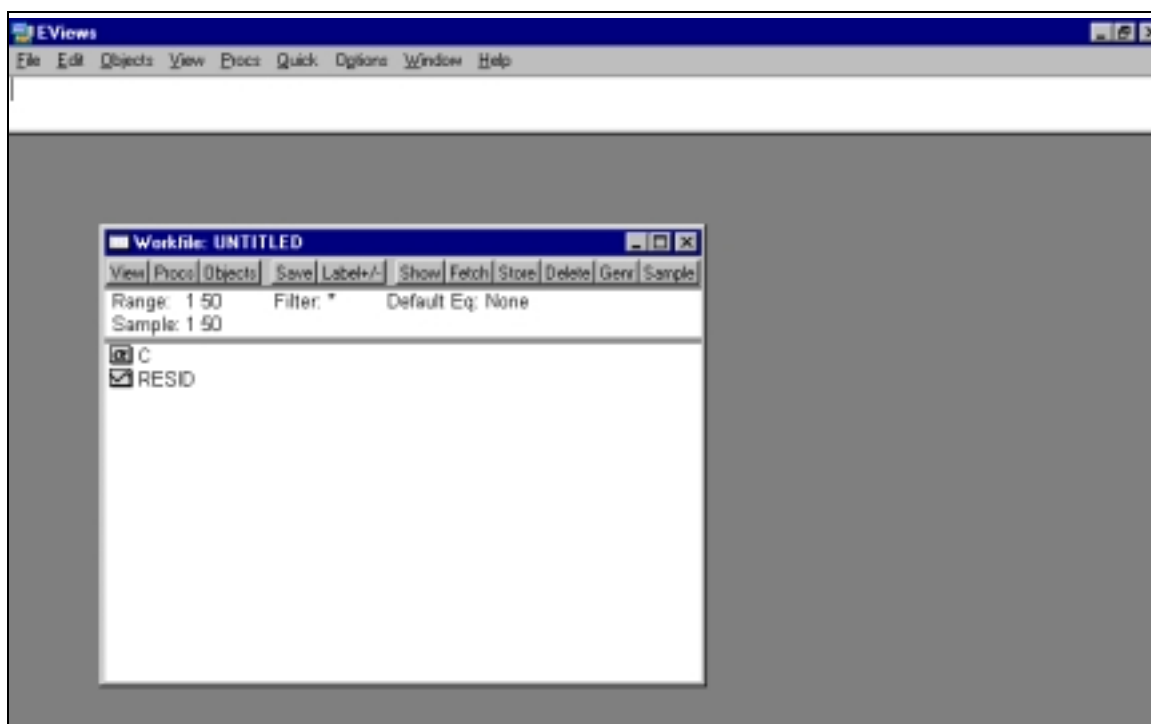


Рисунок 4 – Окно рабочего файла

Исходные данные для анализа находятся в файле Flat.xls. В пакете EViews предусмотрена возможность импорта данных. Для визуализации данных необходимо запустить табличный процессор Excel и открыть файл Flat.xls. Вид экрана представлен на рисунке 5. Ознакомившись с данными, файл, подлежащий экспортированию, необходимо закрыть.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	N	BAL	BRICK	DIST	FLOOR	KITSP	LIVSP	METRDIST	PRICE	TOTSP	WALK				
2	1	1	1	4,9	1	13	19,4	7	53	43,6	1				
3	2	1	0	15,7	1	8,8	20,5	5	36	39	0				
4	3	0	0	4,9	1	10,1	16,9	2	36	35,2	1				
5	4	1	0	15,7	1	8,5	17,5	5	44	35	0				
6	5	1	0	11,6	1	10,2	20	10	29,35	40	0				
7	6	1	0	15,7	1	8,5	19,4	10	30	38,9	0				
8	7	1	0	8,8	1	9	21	5	38	42	0				
9	8	0	0	8,8	0	5,7	21,5	5	24	33	1				
10	9	1	0	8,8	1	8,5	19	15	32	38	1				
11	10	0	1	4,9	1	4,5	15,6	7	30	28	1				
12	11	1	0	17,2	1	8,5	19	15	32	39	0				
13	12	0	0	14,6	1	7,5	17,9	5	26,5	33,3	0				
14	13	0	0	14,5	1	7,3	18	15	31	34	1				
15	14	1	0	14,5	1	8,7	19,3	5	38	39,1	0				
16	15	1	0	11,6	1	8,5	21	15	28	39	0				
17	16	1	1	8	1	6	18,2	10	28	30	0				
18	17	0	0	8	0	6	19,5	5	30	33,3	1				
19	18	1	0	12,4	1	9	18	9	44	37,5	1				
20	19	1	0	14,6	0	6	16	5	27,5	30	0				
21	20	1	0	14,5	1	8,5	19,8	10	32	39	0				
22	21	1	0	12,4	1	5,5	21,3	15	30	31,4	1				
23	22	1	0	14,6	1	8,4	19	5	36,5	38	1				
24	23	1	0	11,6	1	9	20,8	10	33	38,7	0				

Рисунок 5 – Исходные данные для анализа в пакете Excel

Файл Flat.xls содержит информацию о стоимости однокомнатных квартир в Москве в 1998:

- N – номер объекта наблюдения;
- BAL – наличие балкона (1 – есть, 0 – нет);
- BRICK – дом кирпичный? (1 – да, 0 – нет);
- DIST – расстояние от центра города до дома, км;
- FLOOR – первый или последний этаж (1 – нет, 0 – да);
- KITSP – площадь кухни, кв.м;
- LIVSP – жилая площадь, кв.м;
- METRDIST – расстояние до метро, минут;
- PRICE – цена квартиры, тыс.дол;
- TOTSP – общая площадь, кв.м;
- WALK – до метро идти пешком? (1 – да, 0 – нет).

Для чтения данных, созданных в других программах, необходимо выбрать в рабочем файле EViews опцию **Procs/Import/Read Text-Lotus-Excel...** Вид экрана представлен на рисунке 6.

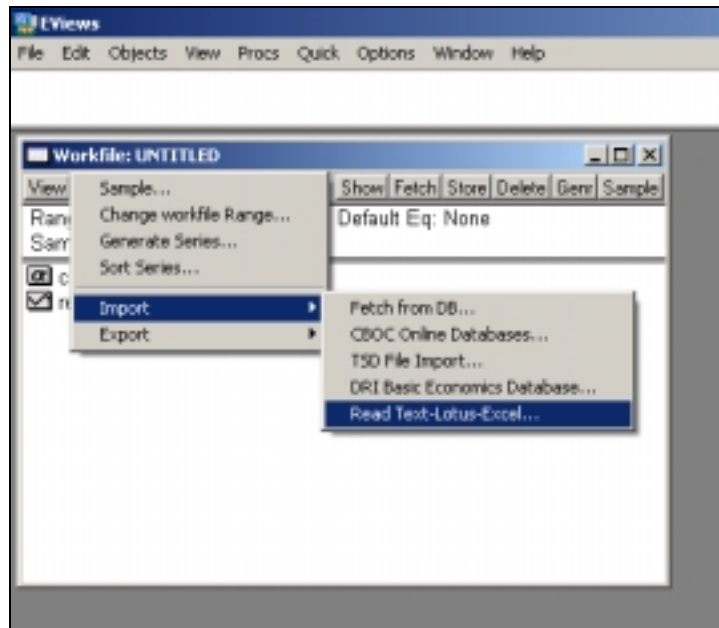


Рисунок 6 – Импорт данных из пакета Excel

После выбора указанных пунктов на экране появится диалог открытия данных, представленный на рисунке 7. Необходимо выбрать папку, в которой находится искомый файл. Для упрощения поиска в опции **Тип файлов (Files of type)** можно выбрать **Excel.xls**. Для того чтобы пакет «помнил» перемещения по папкам, можно поставить флажок в опции **Update default directory**. Вид окна представлен на рисунке 7.

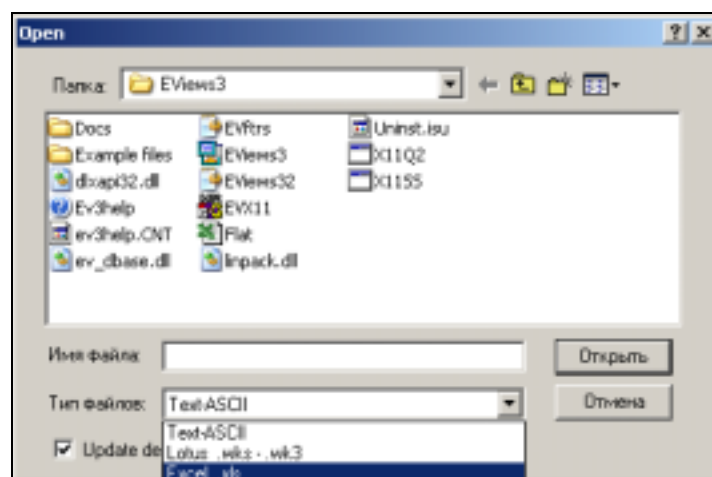


Рисунок 7 – Открытие файла

В окне необходимо выбрать файл **Flat.xls** и нажать кнопку **Открыть**. На экране появится диалог открытия электронных таблиц формата Excel, представленный на рисунке 8.

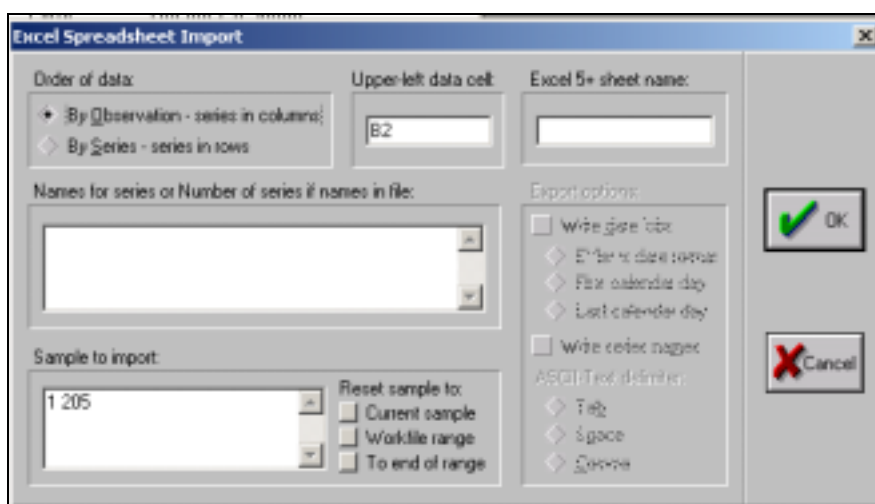


Рисунок 8 – Открытие электронной таблицы формата Excel

По умолчанию в окне, представленном на рисунке 8, предполагается, что данные находятся в столбцах (**by observation - series in columns**). Если данные представлены в виде переменных (признаков, серий) в строках, то надо отметить другую опцию (**By series - series in rows**).

В поле **Upper-left data cell** (левая верхняя ячейка данных) автоматически отображена клетка **B2**. Это означает, что данные будут импортироваться из исходной таблицы с клетки, указанной в этом окне (тем самым первая строка и первый столбец будут пропущены). Это не соответствует структуре файла Flat.xls, поэтому в данном поле следует записать **A2**.

В поле **Names for series or Number of series if names in file** (имена для переменных или число переменных, если имена содержатся в файле) указываем цифру 11. Это связано с тем, что файл данных содержит 11 переменные, находящиеся в столбцах. Имена для этих переменных будут взяты из первой строки электронной таблицы (клетки A1:K1). В том случае, когда необходимо импортировать часть данных (например, только первые две переменные), надо ввести их количество. Если имена переменных, по каким либо причинам, в исходном файле не заданы, можно вместо цифр ввести их имена (латинскими буквами). Если количество переменных, введенных в рассматриваемом окне, превышает количество реально существующих, то в рабочий файл будет введен столбец с заданным именем без данных (обозначаются такие клетки как NA).

После того, как сделаны все установки можно нажать кнопку **ОК**. Появится окно, изображенное на рисунке 9.

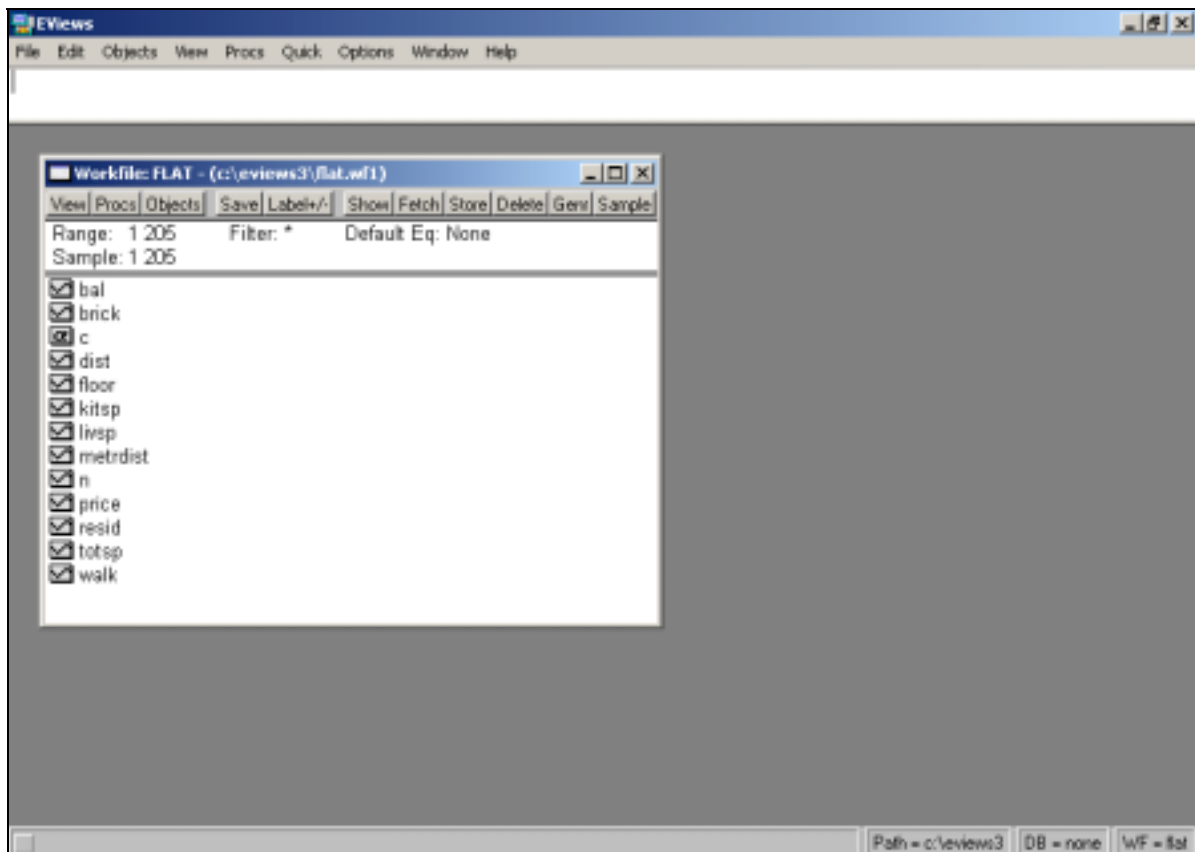


Рисунок 9 – Рабочий файл с данными

После того, как исходные данные перенесены в рабочую область пакета (появились имена переменных), необходимо провести их верификацию (проверку правильности). Для просмотра значений каждой переменной требуется два раза «кликнуть» мышью по имени переменной (например, PRICE). В окне откроется столбец со значениями выбранной переменной, представленный на рисунке 10.

The screenshot shows the EViews software window titled 'Series: PRICE Workfile: UNTITLED'. The menu bar includes 'View', 'Proc', 'Objects', 'Print', 'Name', 'Freeze', 'Edit+/-', 'Simpl+/-', 'Label+/-', 'Wide+', 'InsDel', 'Title', and 'Sa'. The main area displays a table titled 'PRICE' with the following data:

PRICE		
Last updated: 10/17/04 - 20:10		
1	1	53.00000
2	2	36.00000
3	3	36.00000
4	4	44.00000
5	5	29.35000
6	6	30.00000
7	7	38.00000
8	8	24.00000
9	9	32.00000
10	10	30.00000
11	11	32.00000
12	12	26.50000
13	13	31.00000
14	14	30.00000

Рисунок 10 – Наблюдаемые значения признака PRICE

Если необходимо просмотреть данные в виде матрицы типа «объект-свойство», то для этого потребуется создать новую группу, содержащую все импортированные переменные. Это осуществляется следующим образом: выделяются переменные, входящие в группу, используя клавишу **CTRL**; подводится курсор мыши на выделенную область и нажимается правая кнопка мыши. Далее необходимо выбрать опцию **Open**. Вид экрана при этом представлен на рисунке 11.

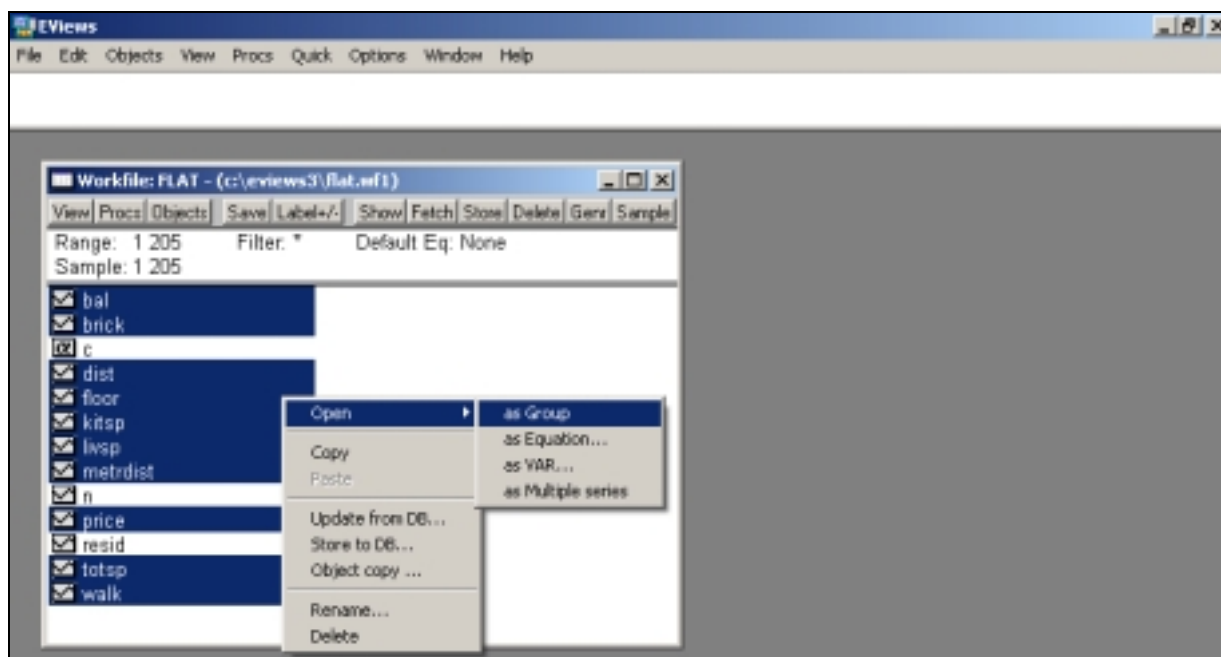


Рисунок 11 – Создание группы признаков

После выбора **Open/as Group** (открыть в одной группе) пакет создаст группу с именем **UNTITLED**, в которую войдут все выделенные переменные. По умолчанию, данные будут представлены в виде электронной таблицы (возможны другие варианты представления), изображенной на рисунке 12.

obs	BAL	BRICK	DIST	FLOOR	KITSP
1	1.000000	1.000000	4.900000	1.000000	13.00
2	1.000000	0.000000	15.70000	1.000000	8.800
3	0.000000	0.000000	4.900000	1.000000	10.10
4	1.000000	0.000000	15.70000	1.000000	8.500
5	1.000000	0.000000	11.60000	1.000000	10.20
6	1.000000	0.000000	15.70000	1.000000	8.500
7	1.000000	0.000000	8.800000	1.000000	9.000
8	0.000000	0.000000	8.800000	0.000000	5.700
9	1.000000	0.000000	8.800000	1.000000	8.500
10	0.000000	1.000000	4.900000	1.000000	4.500
11	1.000000	0.000000	17.20000	1.000000	8.500
12	0.000000	0.000000	14.60000	1.000000	7.500
13	0.000000	0.000000	14.50000	1.000000	7.300
14	1.000000	0.000000	14.50000	1.000000	8.700
15	1.000000	0.000000	11.60000	1.000000	8.500
16	1.000000	1.000000	8.000000	1.000000	6.000
17	0.000000	0.000000	8.000000	0.000000	6.000

Рисунок 12 – Просмотр группы признаков

Группе данных можно дать имя. Для этого необходимо нажать кнопку **Name** в текущем окне, изображенном на рисунке 12. Появится диалоговое окно, представленное на рисунке 13. Автоматически предложено имя **GROUP01**. Его можно принять, нажав кнопку **OK**. В рабочем файле добавится одна переменная с введенным именем. Теперь к ней всегда можно перейти простым нажатием клавиши мыши.

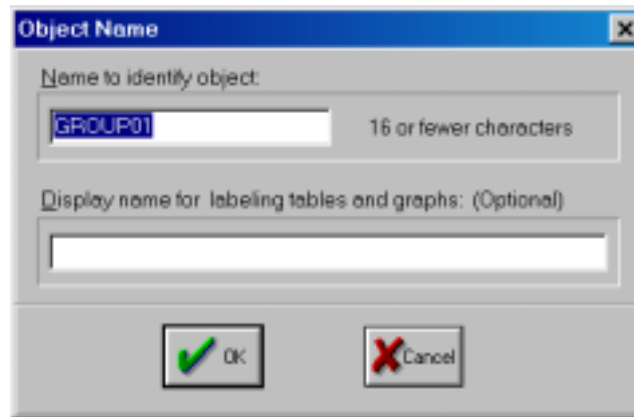


Рисунок 13 – Присвоение имени группе признаков

Образованную группу можно просматривать не только в виде электронной таблицы. Если в окне со значениями признаков в группе **GROUP01** выбрать **View/Multiple Graphs/Line** (рисунок 14), то данные предстанут не в виде таблицы, а как линейные графики по каждой переменной. Вид экрана представлен на рисунке 15.

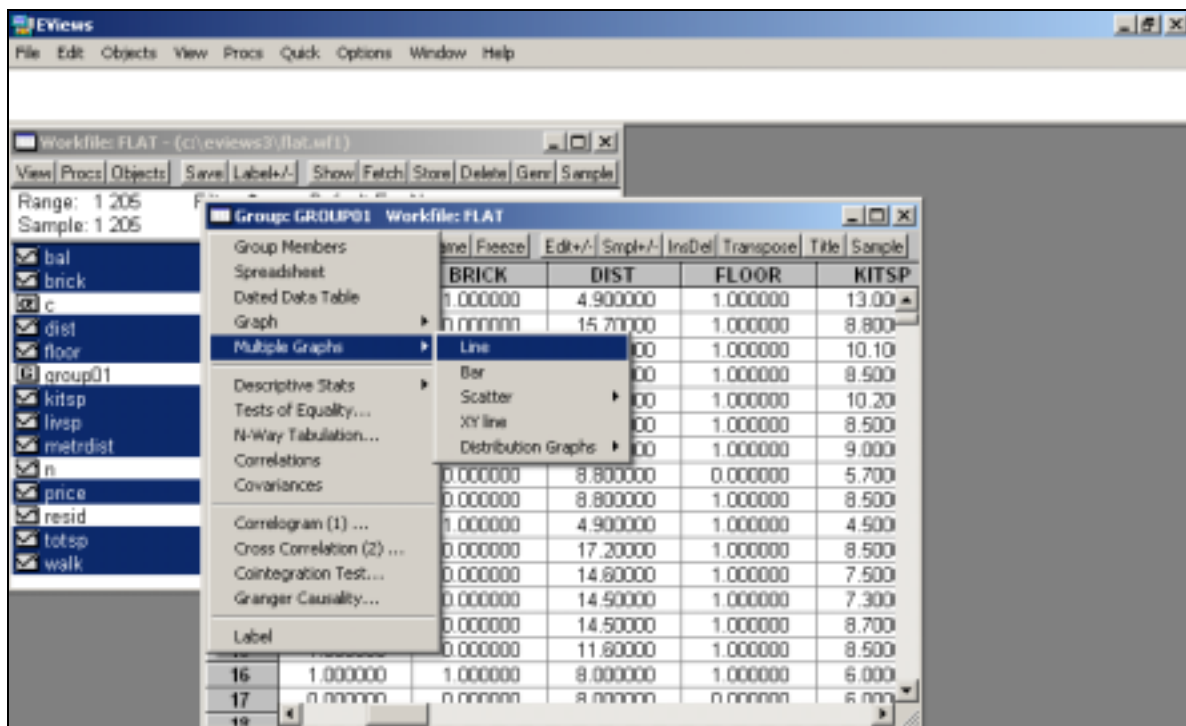


Рисунок 14 – Пункты меню для просмотра значений признаков в виде линейных графиков

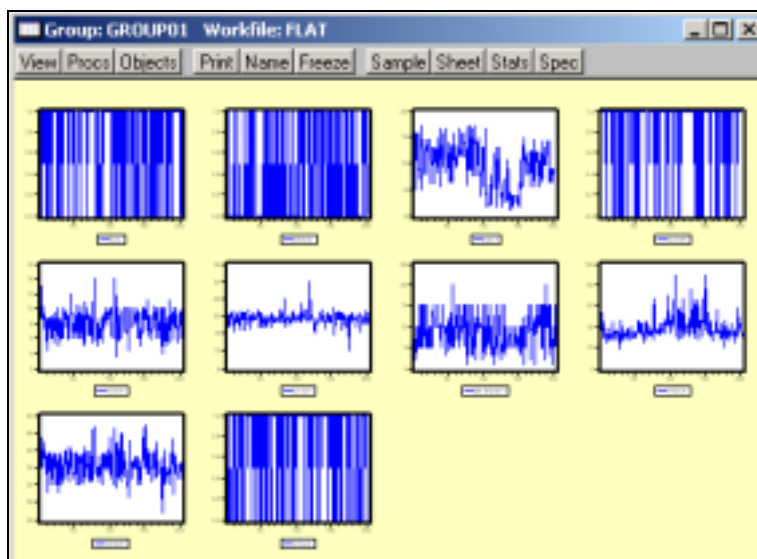


Рисунок 15 – Значения признаков в виде линейных графиков

Для того чтобы вернуться к прежней форме представления данных (например, электронной таблице), необходимо выбрать **View/Spreadsheet**.

Для просмотра оценок основных параметров распределения (описательных статистик) выделенных переменных необходимо выбрать **View/Descriptive Stats/Individual Samples** (рисунок 16).

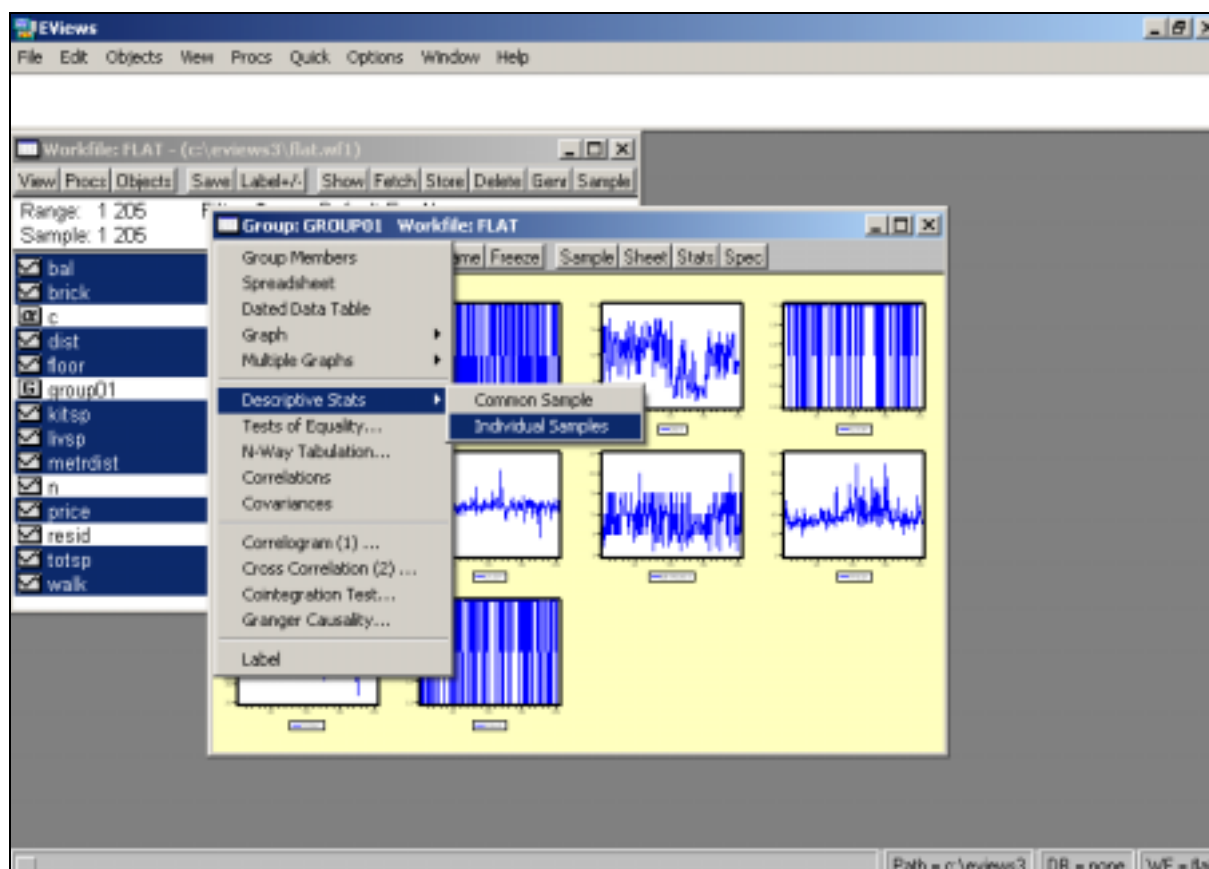


Рисунок 16 – Выбор пунктов меню для просмотра оценок основных параметров распределения признаков, входящих в группу **GROUP01**

В результате появится окно, представленное на рисунке 17. В нем содержатся:

- Mean** – среднее арифметическое значение;
- Median** – оценка медиана;
- Maximum** – максимальное значение;
- Minimum** – минимальное значение;
- Std. Dev.** – оценка среднего квадратического отклонения;
- Skewness** – оценка коэффициента асимметрии;
- Kurtosis** – оценка коэффициента эксцесса;
- Jarque-Bera** – статистика Jarque-Bera;
- Probability** – Вероятность;
- Observations** – Количество наблюдений.

	BAL	BRICK	DIST	FLOOR	KITSP	LIVSP	METRDIST	PRICE	T
Mean	0.687805	0.356098	9.360976	0.800000	8.053659	19.16634	8.712195	38.63244	
Median	1.000000	0.000000	9.700000	1.000000	8.500000	19.00000	10.00000	35.50000	
Maximum	1.000000	1.000000	17.20000	1.000000	14.00000	30.00000	20.00000	89.00000	
Minimum	0.000000	0.000000	1.100000	0.000000	3.500000	10.00000	1.000000	18.00000	
Std. Dev.	0.464523	0.480017	4.266644	0.400979	1.835128	1.831823	4.258797	11.06564	
Skewness	-0.810571	0.601040	-0.179861	-1.500000	0.065455	0.346913	0.214677	1.790523	
Kurtosis	1.657026	1.361249	2.078517	3.250000	2.964392	11.26945	2.230532	7.371428	
Jarque-Bera	37.85396	35.28136	8.358285	77.40885	0.157211	588.2237	6.631975	272.7633	
Probability	0.000000	0.000000	0.015312	0.000000	0.924405	0.000000	0.036298	0.000000	
Observations	205	205	205	205	205	205	205	205	

Рисунок 17 – Значения описательных статистик

Для оценки матрицы парных коэффициентов корреляции необходимо выбрать **View/Correlations**. Результат представлен на рисунке 18.

Correlation Matrix									
	BAL	BRICK	DIST	FLOOR	KITSP	LIVSP	METRDIST	PRICE	TOTSP
BAL	1.000000	-0.070563	0.148404	0.163167	0.266439	0.007177	0.132765	0.119672	0.143697
BRICK	-0.070563	1.000000	-0.446982	-0.086591	-0.198757	-0.160794	-0.280528	0.348263	-0.071726
DIST	0.148404	-0.446982	1.000000	0.110885	-0.006668	-0.049284	0.215763	-0.465964	0.022443
FLOOR	0.163167	-0.086591	0.110885	1.000000	0.251810	0.070207	0.063726	0.173970	-0.058684
KITSP	0.266439	-0.198757	-0.006668	0.251810	1.000000	0.157983	0.062763	0.405642	-0.075343
LIVSP	0.007177	-0.160794	-0.049284	0.070207	0.157983	1.000000	-0.045044	0.205703	-0.044686
METRDIST	0.132765	-0.280528	0.215763	0.063726	0.062763	-0.045044	1.000000	-0.277628	-0.154642
PRICE	0.119672	0.348263	-0.465964	0.173970	0.405642	0.205703	-0.277628	1.000000	0.260951
TOTSP	0.143697	-0.071726	0.022443	0.200142	0.678079	0.465946	-0.059760	0.577812	1.000000
WALK	-0.116138	0.305756	-0.230832	-0.058684	-0.075343	-0.044686	-0.154642	0.260951	-0.059760

Рисунок 18 – Оценка матрицы парных коэффициентов корреляции

С помощью пунктов **View/Covariances** рассчитывается оценка ковариационной матрицы. Пункт меню **View/Test of Equality...** предназначен для проверки гипотез об однородности ряда генеральных средних, медиан и

дисперсий. Для анализа таблиц сопряженностей используется пункт меню **View/N-Way Tabulation...**

Можно исследовать характеристики отдельных признаков, например, признака цена квартиры. Для этого в рабочем файле необходимо дважды «кликнуть» по признаку **PRICE** (рисунок 10). С помощью пунктов меню **View/Descriptive Stats/Histogram and Stats** (рисунок 19) рассчитываются оценки основных параметров распределения и выводится гистограмма частот. Результаты представлены на рисунке 20.

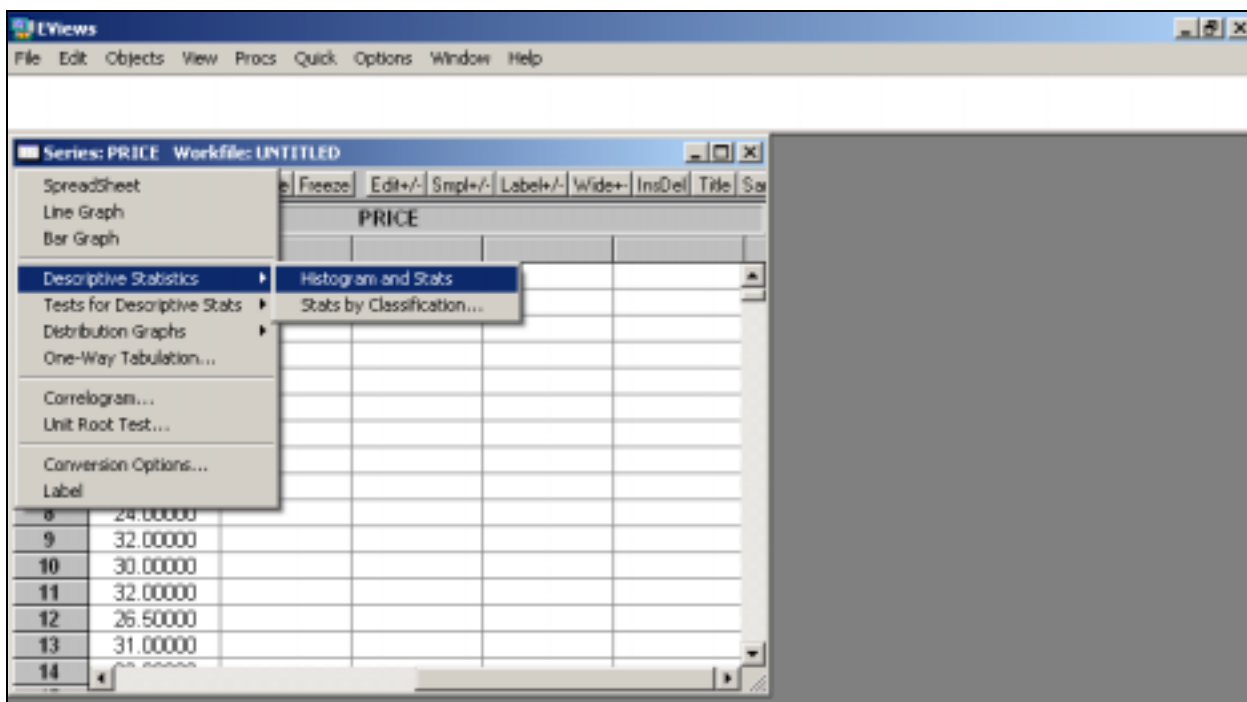


Рисунок 19 – Выбор пунктов меню для расчета основных статистик для признака PRICE

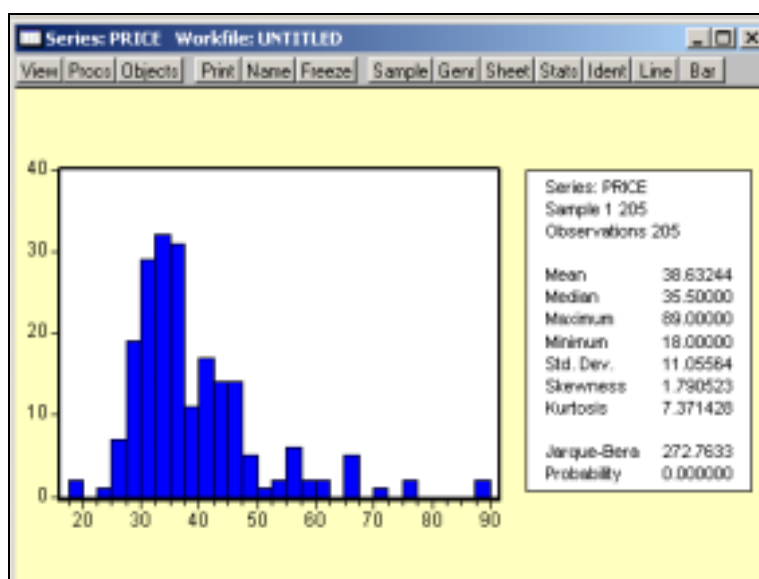


Рисунок 20 – Гистограмма частот и оценки основных параметров распределения признака PRICE

При выборе пунктов меню **View/Tests for Descriptive Stats/Simple Hypothesis Tests** на экране появится форма, представленная на рисунке 21. Она предназначена для проверки гипотез о значении математического ожидания, медианы, дисперсии. Так для проверки гипотезы о равенстве математического ожидания **30** введем данное значение в поле рядом с надписью **Mean**.

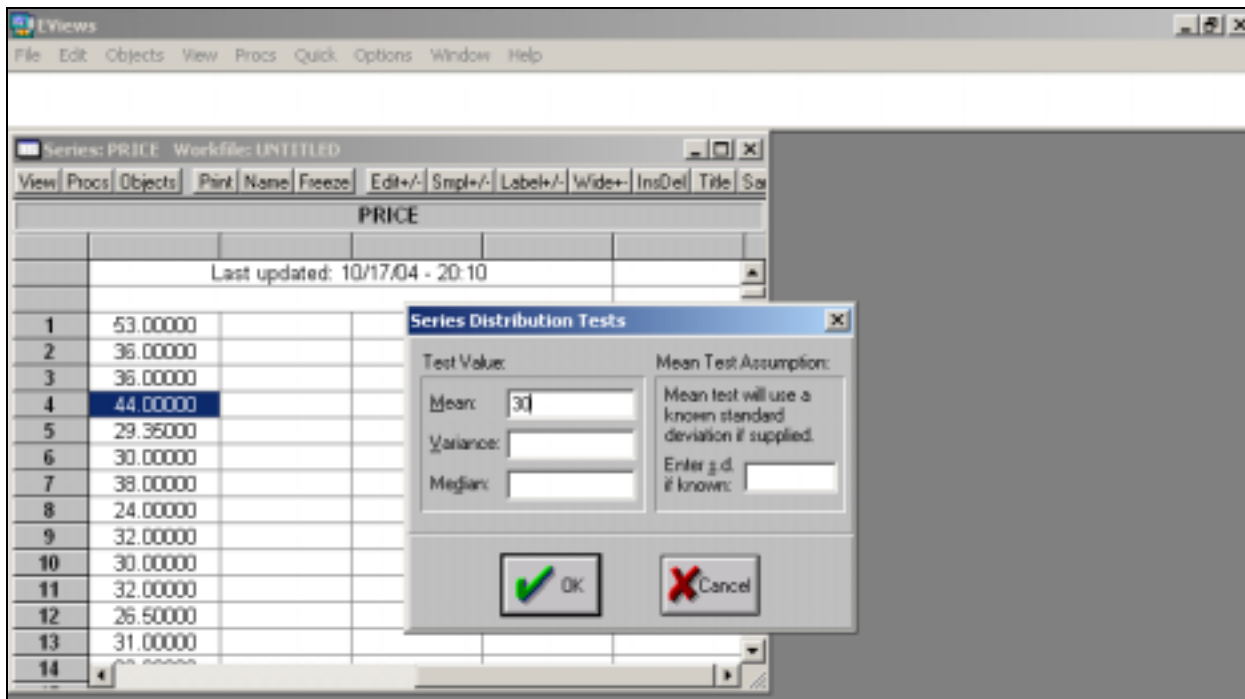


Рисунок 21 – Проверка гипотезы о значении параметров распределения

После нажатия кнопки **ОК** на экране появляются результаты проверки гипотезы, представленные на рисунке 22.

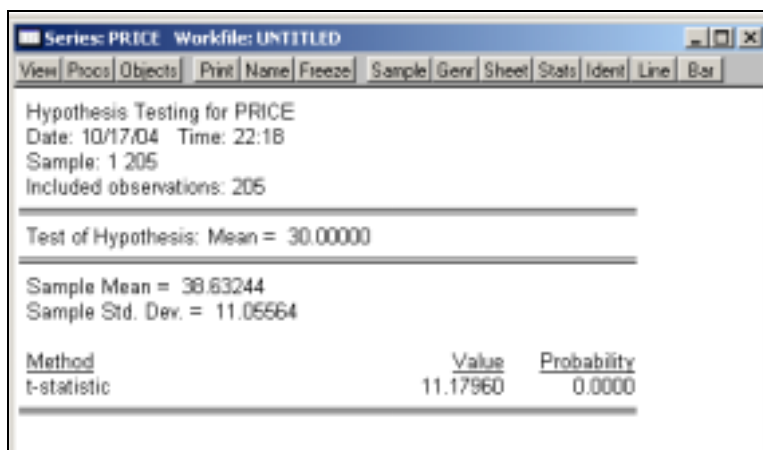


Рисунок 22 – Результаты проверки гипотезы о значении математического ожидания

Так как наблюдаемый уровень значимости равен $0,0000 < 0,05$, то гипотеза о значении математического ожидания равного 30 отвергается.

Оценку функции распределения можно получить с помощью пунктов меню **View/Distribution Graphs/CDF-Survivor-Quantile...** На экране появится форма, представленная на рисунке 23.

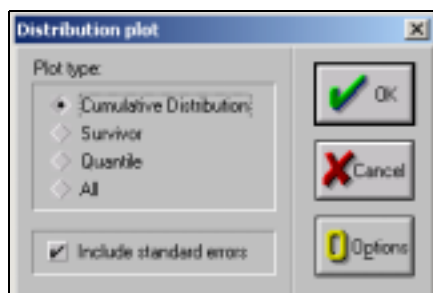


Рисунок 23 – Форма для оценки функции распределения

После нажатия кнопки **ОК** на экране появится график кумулятивной функции распределения, изображенный на рисунке 24.

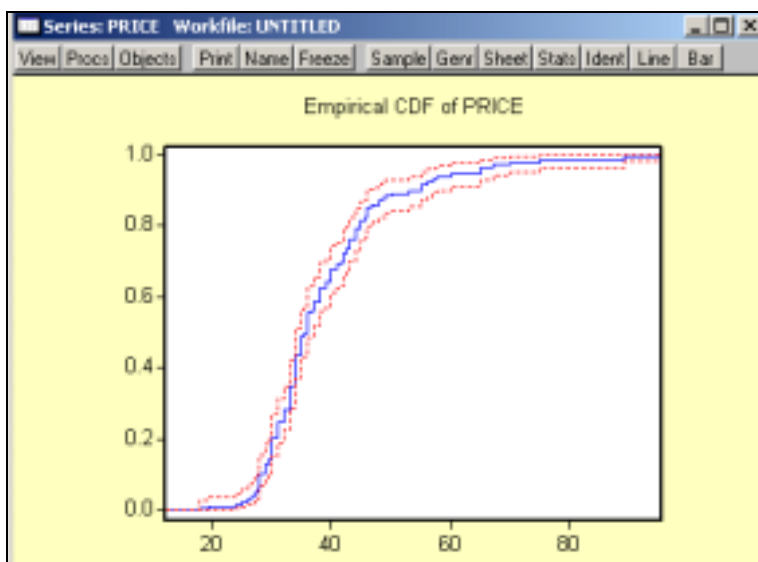


Рисунок 24 – График кумулятивной функции распределения

При выборе пунктов меню **View/One-Way Tabulation...** на экране появится форма, представленная на рисунке 25.



Рисунок 25 – Построение вариационного ряда

После нажатия на кнопку **ОК** на экране появятся вариационный ряд, изображенный на рисунке 26.

Value	Count	Percent	Cumulative Count	Cumulative Percent
18.00000	1	0.49	1	0.49
19.50000	1	0.49	2	0.98
24.00000	1	0.49	3	1.46
25.00000	2	0.98	5	2.44
26.00000	2	0.98	7	3.41
26.50000	1	0.49	8	3.90
27.00000	2	0.98	10	4.88
27.50000	2	0.98	12	5.85
28.00000	9	4.39	21	10.24
29.00000	6	2.93	27	13.17
29.35000	1	0.49	28	13.66

Рисунок 26 – Дискретный вариационный ряд для признака PRICE

3 Построение линейной модели множественной регрессии

Для построения линейной модели множественной регрессии методом наименьших квадратов используется команда следующего формата: LS «эндогенная переменная», «константа», «список экзогенных переменных».

Команда для построения линейной модели зависимости стоимости квартиры от всех остальных признаков представлена на рисунке 27. Уравнение регрессии можно также построить с помощью пунктов меню **QUICK/ESTIMATE EQUATION** (рисунок 28).

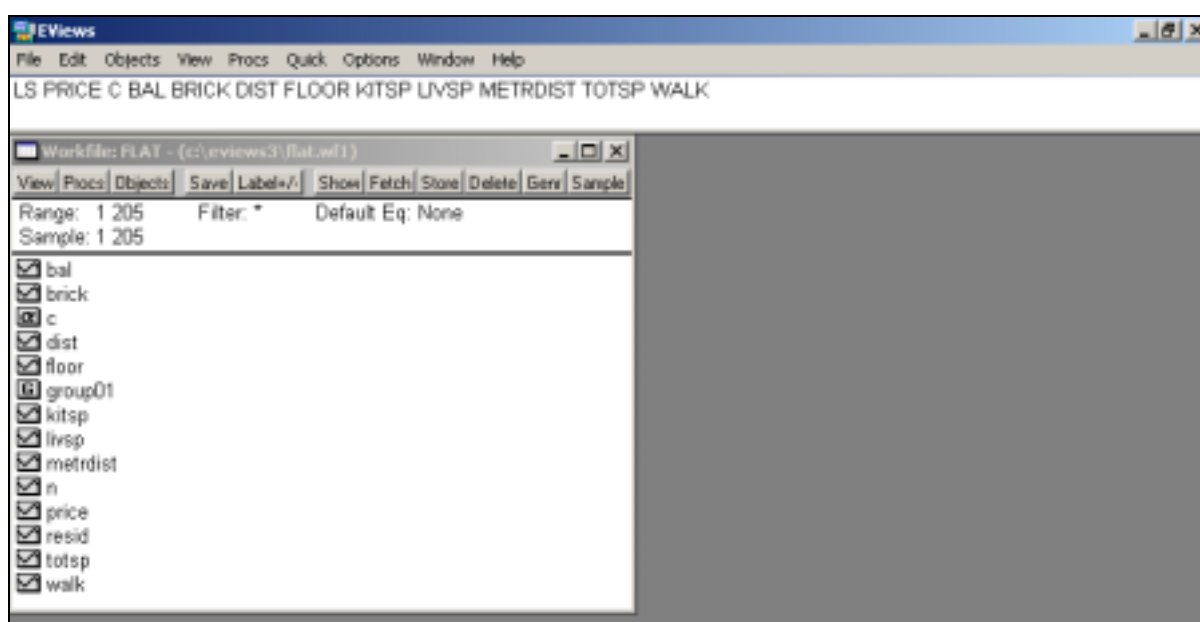


Рисунок 27 – Команда для построения линейной модели множественной регрессии

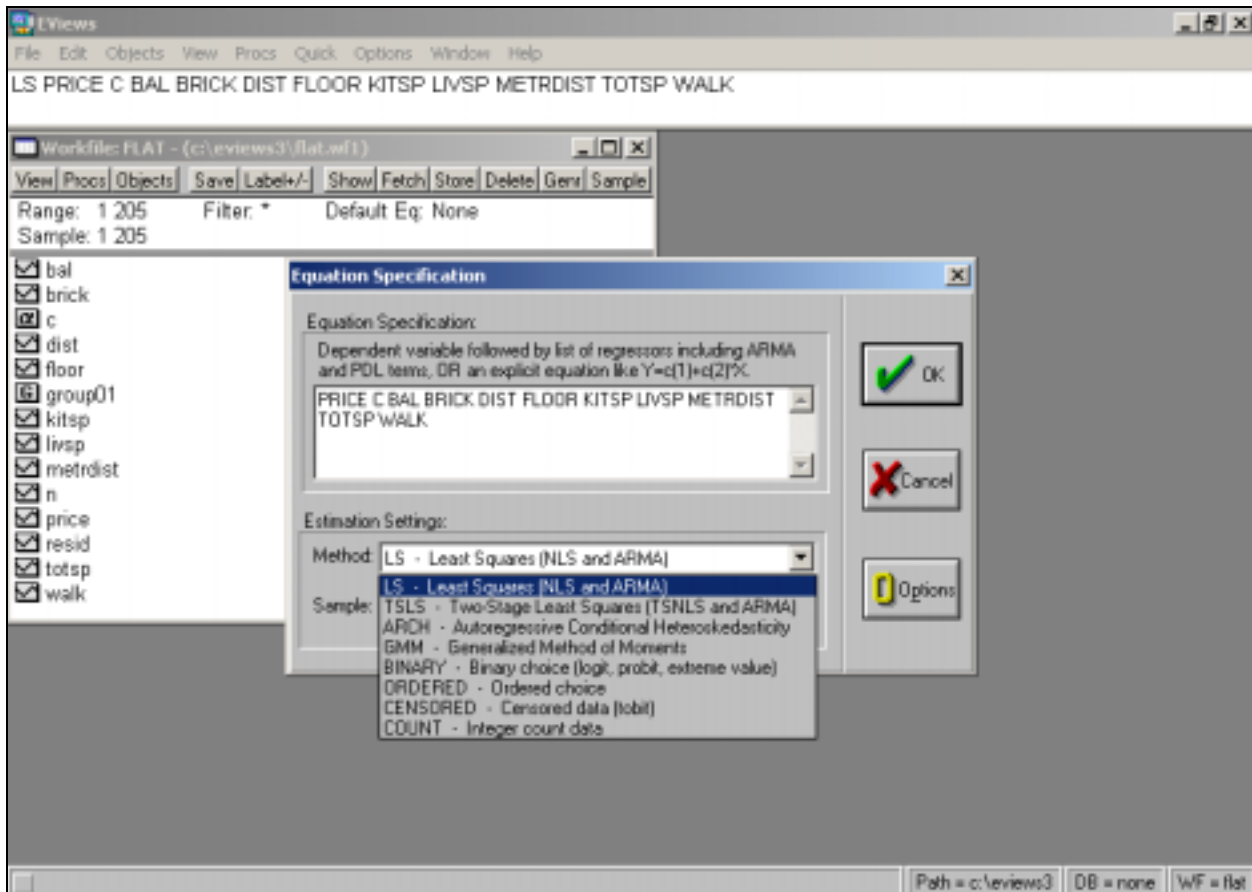


Рисунок 28 – Построение линейной модели множественной регрессии с помощью меню

Результаты оценивания представлены на рисунке 29. В открывшемся окне отражены: зависимая переменная, применяемый метод, число наблюдений, набор экзогенных переменных, оценки параметров уравнения регрессии, стандартные ошибки, значения t-статистик и соответствующие им вероятности, оценка коэффициента детерминации R^2 и ряд других показателей /1/-/3/.

Как видно из рисунка 29 незначимыми на уровне $\alpha = 0,05$ являются коэффициенты при признаках **KITSP** и **LIVSP**. Однако следует отметить, что на стоимость квартиры помимо общей площади, площади кухни и жилой площади, возможно, влияет оставшаяся площадь квартиры (площадь коридора и др.). Проверим эту гипотезу. Сгенерируем новую переменную **OSTSP**, учитывающую оставшуюся площадь квартиры. Для этого введем команду, вид которой представлен на рисунке 30.

Построим уравнение регрессии, учитывающее помимо прежних признаков влияние признака **OSTSP**. При этом для избежания полной мультиколлинеарности в уравнение регрессии не следует включать признак **TOTSP**. Команда для построения уравнения регрессии представлена на рисунке 31.

EViews - [Equations: UNTITLED Workfile: FLAT]

File Edit Objects View Procs Quick Options Window Help

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: PRICE
 Method: Least Squares
 Date: 10/25/04 Time: 00:27
 Sample: 1 205
 Included observations: 205

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7.210555	5.781264	-1.247228	0.2138
BAL	2.950858	1.024219	2.881080	0.0044
BRICK	2.800640	1.171889	2.389650	0.0178
DIST	-1.005091	0.123405	-8.144623	0.0000
FLOOR	3.117493	1.162693	2.681269	0.0080
KITSP	-0.180979	0.376289	-0.480957	0.6311
LIVSP	-0.560457	0.296526	-1.856356	0.0649
METRDIST	-0.291332	0.111608	-2.610316	0.0097
TOTSP	1.707652	0.177854	9.601428	0.0000
WALK	5.171294	1.048920	4.930115	0.0000

R-squared	0.683031	Mean dependent var	38.63244
Adjusted R-squared	0.668402	S.D. dependent var	11.05564
S.E. of regression	6.366345	Akaike info criterion	6.587479
Sum squared resid	7903.418	Schwarz criterion	6.749577
Log likelihood	-665.2166	F-statistic	46.68913
Durbin-Watson stat	1.856846	Prob(F-statistic)	0.000000

Рисунок 29 – Результаты оценивания параметров линейной модели множественной регрессии

EViews

File Edit Objects View Procs Quick Options Window Help

GENR OSTSP=(TOTSP-(KITSP+LIVSP))

Workfile: FLAT - (c:\eviews3\flat.wf1)

View Procs Objects Save Label+/- Show Fetch Store Delete Genr Sample

Range: 1 205 Filter: * Default Eq: None
 Sample: 1 205

- bal
- brick
- c
- dist
- floor
- group01
- kitsp
- livsp
- metrdist
- n
- price
- resid
- totsp
- walk

Рисунок 30 – Создание переменной OSTSP

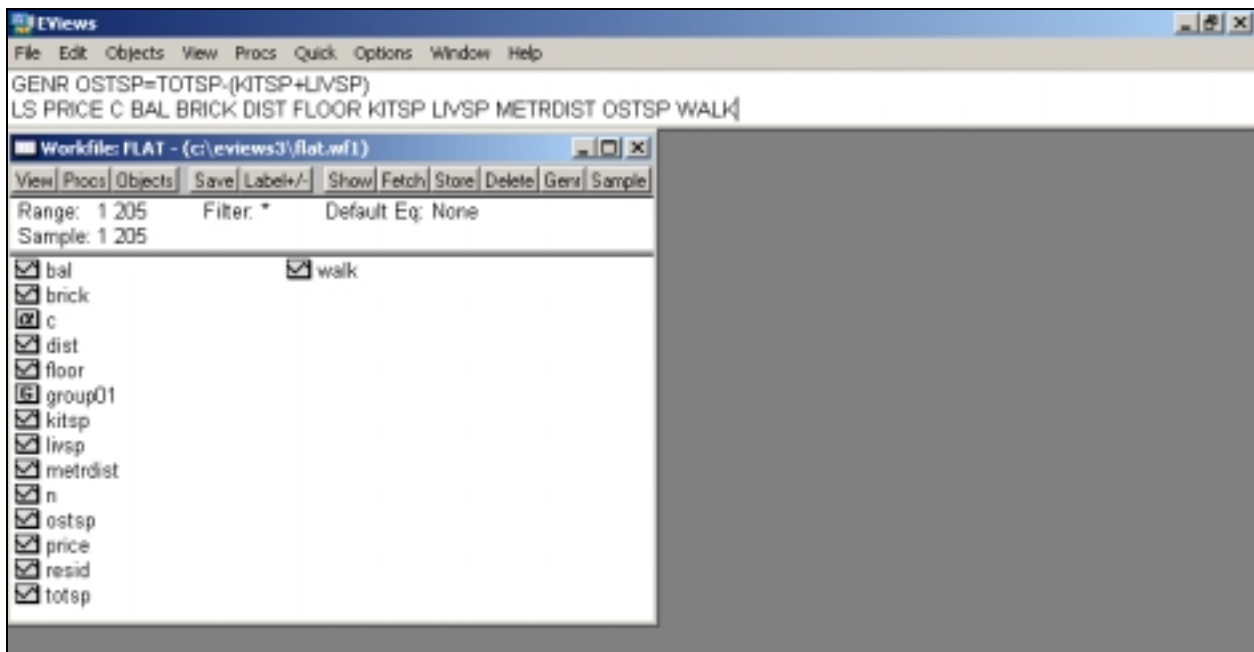


Рисунок 31 – Построение линейной модели множественной регрессии

Результаты оценивания параметров линейной модели множественной регрессии представлены на рисунке 32.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7.210555	5.781264	-1.247228	0.2138
BAL	2.950858	1.024219	2.881080	0.0044
BRICK	2.600640	1.171889	2.389850	0.0178
DIST	-1.005091	0.123405	-8.144623	0.0000
FLOOR	3.117493	1.162693	2.681269	0.0080
KITSP	1.526673	0.281242	5.428334	0.0000
LVSP	1.157195	0.252713	4.579062	0.0000
METRDIST	-0.291332	0.111608	-2.610316	0.0097
OSTSP	1.707652	0.177854	9.601428	0.0000
WALK	5.171294	1.048920	4.930115	0.0000

R-squared	0.683031	Mean dependent var	38.63244
Adjusted R-squared	0.668402	S.D. dependent var	11.05564
S. E. of regression	6.366345	Akaike info criterion	6.587479
Sum squared resid	7903.418	Schwarz criterion	6.749577
Log likelihood	-665.2166	F-statistic	46.68913
Durbin-Watson stat	1.856846	Prob(F-statistic)	0.000000

Рисунок 32 – Результаты оценивания параметров линейной модели множественной регрессии

Как видно из рисунка 32 построенная модель значима, значимыми являются все коэффициенты модели.

Для ответа на вопрос «Что стоит дороже: квадратный метр кухни или квадратный метр жилой площади (комнаты)?», проверим гипотезу

$H_0 : \beta_6 = \beta_7$. Для этого в меню окна **Equation** выберем **View/Coefficient Tests/Wald – Coefficient Restrictions**. Вид экрана представлен на рисунке 33.

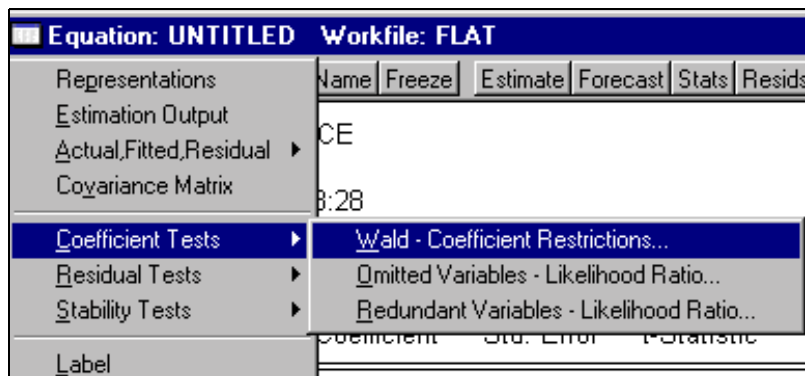


Рисунок 33 – Проверка гипотезы о равенстве коэффициентов регрессии

На экране появится окно, в котором необходимо сформулировать нулевую гипотезу. Вид окна представлен на рисунке 34.

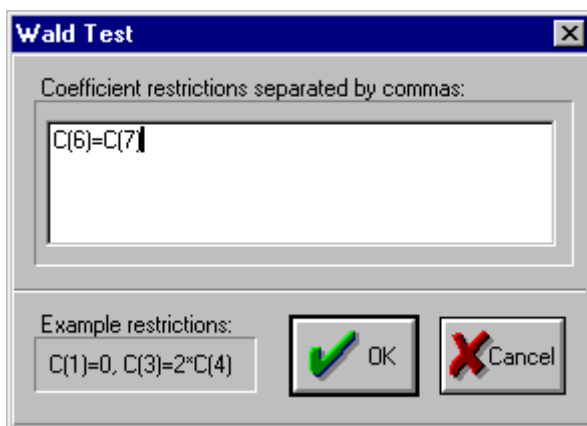


Рисунок 34 – Формулировка гипотезы о равенстве коэффициентов регрессии β_6 и β_7

Результаты проверки гипотезы представлены на рисунке 35. Как видно из рисунка нулевая гипотеза о равенстве коэффициентов не отвергается, т.е. можно сделать вывод о том, что увеличение стоимости одного квадратного метра кухни или одного квадратного метра комнаты одинаково влияет на изменение стоимости квартиры. Этот факт подтверждает и полученная оценка уравнения регрессии: коэффициенты при переменных **KITSP** и **LIVSP** имеют практически одинаковые значения.

Equation: UNTITLED Workfile: FLAT			
View	Proc	Objects	Print Name Freeze Estimate Forecast Stats Resids
Wald Test: Equation: Untitled			
Null Hypothesis: C(5)=C(7)			
F-statistic	0.849103	Probability	0.357945
Chi-square	0.849103	Probability	0.356806

Рисунок 35 – Результаты проверки гипотезы о равенстве коэффициентов регрессионной модели

4 Исследование модели на гетероскедастичность и автокорреляцию

Согласно результатам регрессионного анализа, представленным на рисунке 32, оценка линейной модели множественной регрессии имеет вид:

$$\hat{y} = -7.21 + 2.95BAL + 2.80BRICK - 1.005DIST + 3.12FLOOR + 1.53KITS + 1.16LIVSP - 0.29METRDIST + 1.71OSTSP + 5.17WALK, \quad (1)$$

где \hat{y} – модельное значение стоимости квартиры.

Для проверки гипотезы об отсутствии гетероскедастичности воспользуемся тестом **White** /1/, /3/. Для этого в меню окна **Equation** выберем пункты **View/Residual Test -- White heteroscedasticity(no cross terms)**. Вид экрана представлен на рисунке 36.

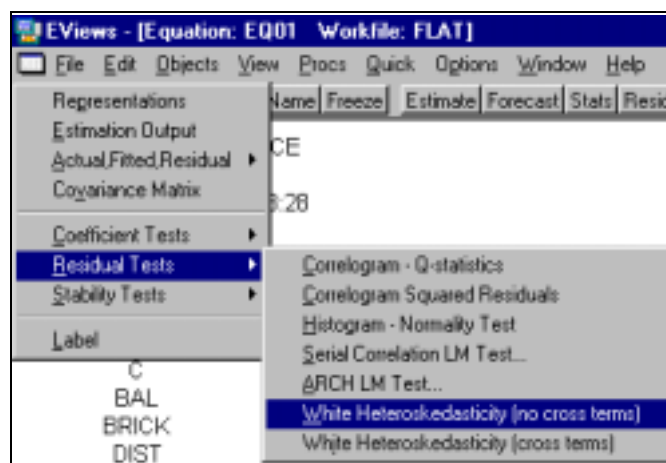


Рисунок 36 – Тест White

Результаты проверки гипотезы об отсутствии гетероскедастичности представлены на рисунке 37.

White Heteroskedasticity Test:			
F-statistic	4.843393	Probability	0.000000
Obs*R-squared	53.91829	Probability	0.000001

Рисунок 37- Результаты проверки гипотезы об отсутствии гетероскедастичности с помощью теста White

На рисунке 37 видно, что вероятность ошибки первого рода равна 0,0000, меньше, чем уровень значимости 0,05. Следовательно, нулевая гипотеза об отсутствии гетероскедастичности отвергается.

На рисунке 32 наряду с основными характеристиками модели представлена статистика Дарбина-Уотсона, с помощью которой можно проверить гипотезу об отсутствии автокорреляции /1/-/3/. Так как $DW=1,86 < 2$, то проверим гипотезу об отсутствии положительной автокорреляции. По таблице определим нижнюю и верхнюю границы критической области: $d_H=1,57$ и $d_B=1,78$. Поскольку $DW > d_B$, то нулевая гипотеза не отвергается.

В нашем случае, когда гетероскедастичность присутствует, проблему можно решить, сделав коррекцию стандартных ошибок в форме Уайта. Для этого выберем в текущем окне пункты **Proc/Specify/Estimate...** Вид экрана представлен на рисунке 38. Появится форма оценки регрессии, где необходимо нажать кнопку **Options**. В появившемся окне отметить **Heteroskedasticity** и указать вид коррекции – **White**. Форма представлена на рисунке 39. Пакет Eviews позволяет также производить коррекцию стандартных ошибок в форме Невье–Веста.

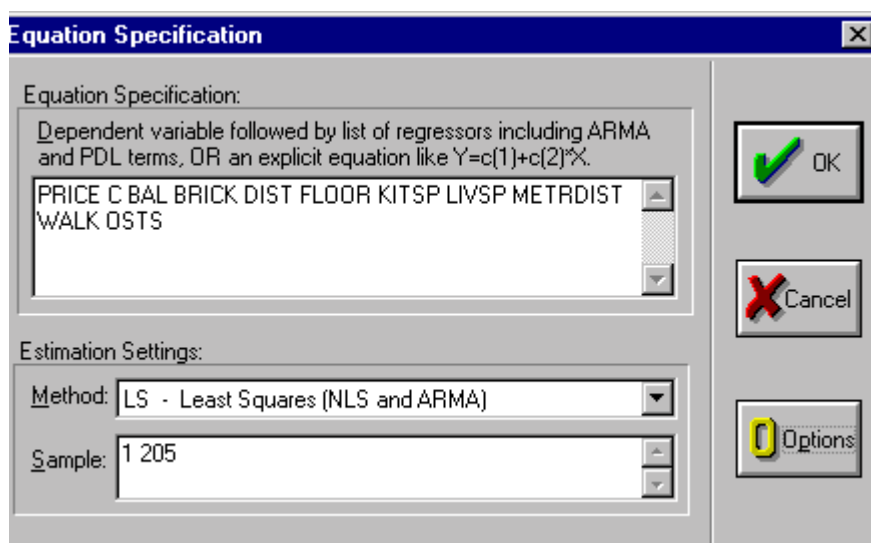


Рисунок 38 – Процедура коррекции регрессионной модели

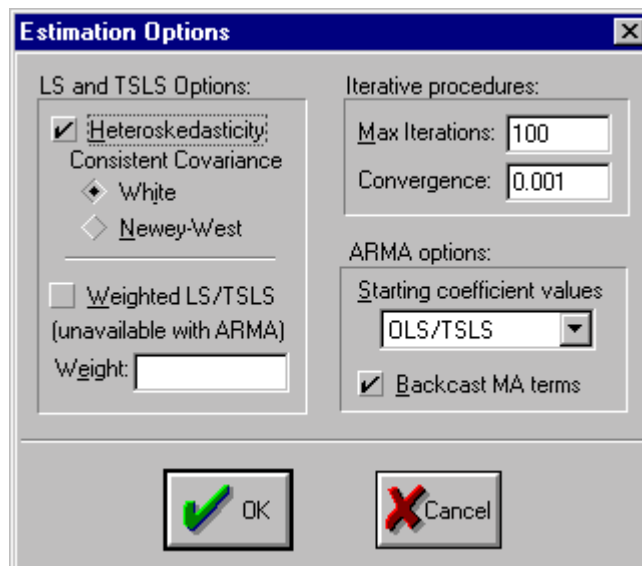


Рисунок 39 – Выбор вида коррекции стандартных ошибок

После нажатия на кнопку **ОК** на экране появится окно, содержащее результаты коррекции (рисунок 40). Если сравнить результаты оценки модели, представленные на рисунках 32 и 40, можно заметить, что оценки параметров модели остались прежними, а изменились лишь стандартные ошибки коэффициентов.

EViews - [Equation: EQ01 Workfile: FLAT]

File Edit Objects View Proc Quick Options Window Help

View Proc Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: PRICE
 Method: Least Squares
 Date: 11/11/04 Time: 19:05
 Sample: 1 205
 Included observations: 205
 White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7.210555	6.728172	-1.071696	0.2852
BAL	2.950858	1.015814	2.904919	0.0041
BRICK	2.800640	1.186815	2.359795	0.0193
DIST	-1.005091	0.136665	-7.354431	0.0000
FLOOR	3.117493	1.080704	2.884688	0.0044
KITSP	1.526673	0.300654	5.077847	0.0000
LIVSP	1.157195	0.293261	3.945953	0.0001
METRDIST	-0.291332	0.103958	-2.802404	0.0056
WALK	5.171294	0.926937	5.578903	0.0000
OSTS	1.707652	0.226001	7.555946	0.0000

R-squared	0.683031	Mean dependent var	38.63244
Adjusted R-squared	0.668402	S.D. dependent var	11.05564
S.E. of regression	6.366345	Akaike info criterion	6.587479
Sum squared resid	7903.418	Schwarz criterion	6.749577
Log likelihood	-665.2166	F-statistic	46.68913
Durbin-Watson stat	1.856846	Prob(F-statistic)	0.000000

Рисунок 40 – Результаты коррекции стандартных ошибок коэффициентов регрессионной модели в форме Уайта

Для проверки гипотезы об отсутствии гетероскедастичности можно также воспользоваться и другими тестами, например, тестом Бреуш-Пагана /2/. Реализуем данный тест в Eviews.

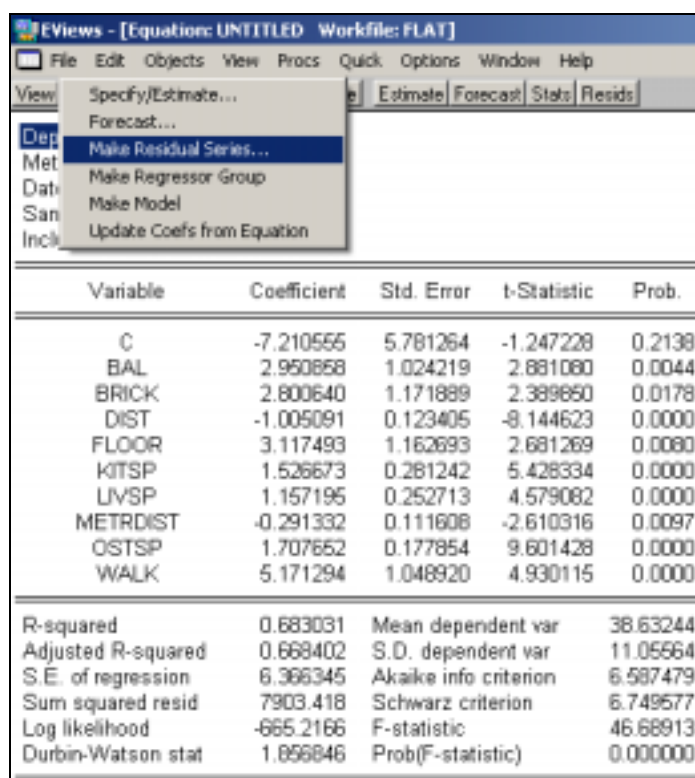
Предположим следующую форму гетероскедастичности:

$$\sigma^2 = C1 + C2 \cdot KITSP + C3 \cdot LIVSP + C4 \cdot OSTSP + C5 \cdot BRICK + C6 \cdot TEL,$$

где *TEL* – признак, указывающий на наличие в квартире телефона (1 – есть, 0 – нет).

Ниже приведена последовательность шагов для реализации этого теста.

1) Оценим снова исходное уравнение и построим ряд остатков регрессии **Res**. Для этого необходимо выбрать пункты меню **Procs/Make residual series**. Вид экрана представлен на рисунке 41.



Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-7.210555	5.781264	-1.247228	0.2138
BAL	2.950858	1.024219	2.881080	0.0044
BRICK	2.800640	1.171889	2.389650	0.0178
DIST	-1.005091	0.123405	-8.144623	0.0000
FLOOR	3.117493	1.162893	2.681269	0.0080
KITSP	1.526673	0.281242	5.428334	0.0000
LIVSP	1.157195	0.252713	4.579082	0.0000
METRDIST	-0.291332	0.111608	-2.610316	0.0097
OSTSP	1.707652	0.177854	9.601428	0.0000
WALK	5.171294	1.048920	4.930115	0.0000

R-squared	0.683031	Mean dependent var	38.63244
Adjusted R-squared	0.668402	S. D. dependent var	11.05564
S. E. of regression	6.366345	Akaike info criterion	6.587479
Sum squared resid	7903.418	Schwarz criterion	6.749577
Log likelihood	-665.2166	F-statistic	46.68913
Durbin-Watson stat	1.856846	Prob(F-statistic)	0.000000

Рисунок 41 – Формирования вектора регрессионных остатков

На экране появится форма ввода имени переменной, содержащей регрессионные остатки. Вид формы представлен на рисунке 42.

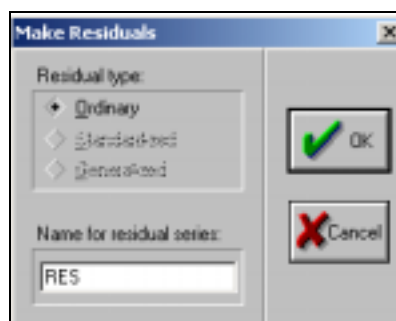


Рисунок 42 – Ввод имени переменной, содержащей регрессионные остатки

После нажатия кнопки ОК в рабочем файле будет создана переменная с указанным именем, в которой будут содержаться регрессионные остатки для модели (1).

2) Сгенерируем ряд квадратов остатков. Для этого в командной строке введем: **genr res2 = res^2**

3) Сгенерируем ряд нормированных остатков. Для этого в командной строке необходимо ввести:

scalar cc=@sum(res2)/205

genr res2n=res2/cc

4) Оценим вспомогательную регрессию:

$$REZ2N = C1 + C2 \cdot KITSP + C3 \cdot LIVSP + C4 \cdot OSTSP + C5 \cdot BRICK + C6 \cdot TEL. \quad (2)$$

Для этого необходимо ввести команду, представленную на рисунке 43. Результаты оценки параметров модели (2) представлены на рисунке 44.

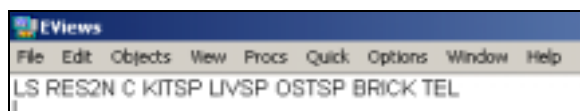


Рисунок 43 – Команда для оценки модели (2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.683727	1.274817	-2.889611	0.0043
KITSP	0.174268	0.064728	2.692618	0.0077
LIVSP	0.133040	0.062048	2.144161	0.0332
OSTSP	0.107433	0.042289	2.540412	0.0118
BRICK	0.963920	0.242823	3.969641	0.0001
TEL	-0.597250	0.308802	-1.934085	0.0545

R-squared	0.162093	Mean dependent var	1.000000
Adjusted R-squared	0.141040	S.D. dependent var	1.693816
S.E. of regression	1.569831	Akaike info criterion	3.768644
Sum squared resid	490.4083	Schwarz criterion	3.965903
Log likelihood	-380.2860	F-statistic	7.699309
Durbin-Watson stat	1.789155	Prob(F-statistic)	0.000001

Рисунок 44 – Результаты оценивания вспомогательной модели (2)

На основании данных рисунка 44 можно сделать вывод, что модель (2) значима, на уровне $\alpha = 0,1$ значимы все коэффициенты модели.

5) Вычислим и сохраним ряд прогнозных (модельных) значений результативной переменной. Для этого выберем в текущем окне пункты меню **Procs/Forecast**. На экране появится форма, изображенное на рисунке 45. В поле **Forecast name** необходимо ввести имя переменной (**RES2NF**), которая будет содержать модельные значения.



Рисунок 45 – Формирование прогнозных значений переменной **RES2N**

После нажатия на кнопку **OK** в рабочем файле появится новая переменная **RES2NF**. Щелкнув два раза по этой переменной, получим вектор прогнозных значений результирующей переменной **RES2N**. Результаты представлены на рисунке 46.

Obs	RES2NF
1	2.732902
2	1.022168
3	1.206900
4	0.495560
5	1.210394
6	1.560451
7	1.370641
8	0.793180
9	1.453518
10	0.391377
11	1.560951
12	0.853962
13	0.320704
14	0.995497
15	1.014915
16	0.773104
17	0.794251
18	0.810372
19	-0.247152
20	0.984187
21	

Рисунок 46 – Прогнозные значения переменной **RES2N**

б) На последнем этапе проверим нулевую гипотезу об отсутствии гетероскедастичности с помощью статистики:

$$F = \frac{RSS}{2}, \quad (3)$$

где RSS – оценка дисперсии признака RES2N.

Статистика (3) при справедливости нулевой гипотезы распределена по закону Хи-квадрат с числом степеней свободы k (число экзогенных переменных). Для проверки гипотезы в командной строке введем следующие команды:

```
scalar mm=@mean(res2n)
genr xx=(res2nf-mm)^2
=0.5*@sum(xx)
```

После этого в левой нижней части рабочего окна появится наблюдаемое значение статистики. В нашем случае этого значение составляет 47,43 (рисунок 47).



Рисунок 47 – Наблюденное значение статистики для проверки нулевой гипотезы об отсутствии гетероскедастичности

По таблице распределения Хи-квадрат определим критическое значение статистики: $\chi^2(0,05;5) = 11,07$. Так как наблюдаемое значение статистики больше, чем критическое, то нулевую гипотезу отвергаем и делаем вывод о наличии гетероскедастичности.

5 Двухшаговая процедура устранения гетероскедастичности

При выявлении гетероскедастичности с помощью теста Бреуша-Пагана можно попытаться осуществить коррекцию с помощью взвешенного метода наименьших квадратов, выбирая в качестве весов $\hat{\sigma}_i = (RES2NF)^{-1/2} / 2$. При этом может оказаться, что некоторые прогнозные значения (RES2NF) являются отрицательными. Если число таких наблюдений невелико, то их можно просто отбросить. В противном случае можно попытаться использовать

мультипликативную форму гетероскедастичности: $\sigma_i^2 = e^{\gamma_0 + \bar{z}_i^T \bar{\gamma}}$, где γ_j являются МНК-оценками модели регрессии $\ln(e_i^2) = \gamma_0 + \bar{z}_i^T \gamma + \delta_i$.

Поскольку в нашем случае среди прогнозных значений переменной RES2N шесть отрицательных, то исключим соответствующие им наблюдения из дальнейшего рассмотрения. Это можно сделать с помощью двойного щелчка мыши на слове Sample в верхней части окна, содержащего рабочий файл. Появится форма, где необходимо задать условие отбора наблюдений. Эту форму следует заполнить так, как показано на рисунке 48.

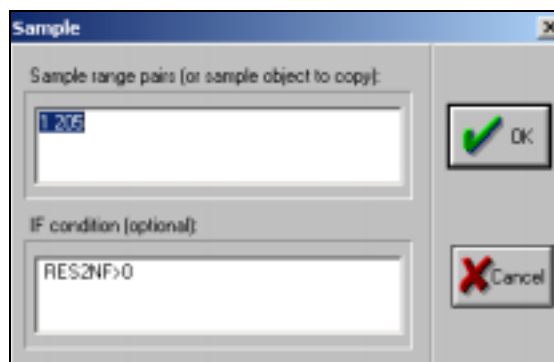


Рисунок 48 – Задание условия отбора наблюдений

Реализуем взвешенный метод наименьших квадратов, выбирая в качестве весов $\hat{\sigma}_i = (RES2NF)^{-1/2}$. Для этого сгенерируем ряд оценок стандартных отклонений регрессионных остатков исходной регрессии с помощью команды: **genr sgm = sqr(RES2NF)**.

После этого перейдем к оценке исходного уравнения взвешенным методом наименьших квадратов. Для этого в окне исходного уравнения (рисунок 32) выберем пункты меню **Estimate/ Options /Weighted LS/TSLS(отметить квадрат)**. В окне **Weight** (вес) запишем имя переменной, значения которой выступают в качестве весов, в нашем случае – 1/SGM. Вид формы представлен на рисунке 49.

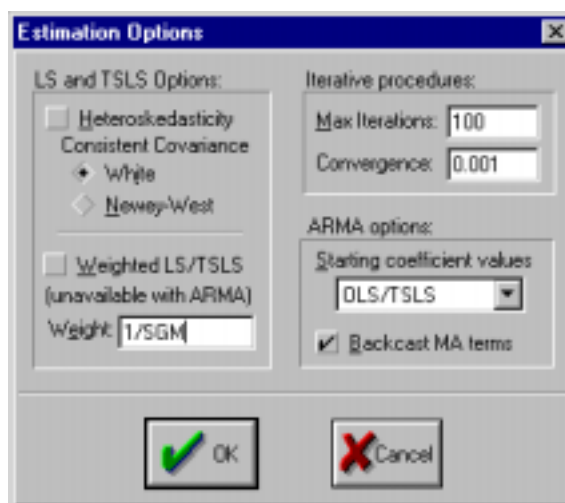


Рисунок 49 – Оценивания параметров регрессионной модели взвешенным методом наименьших квадратов

После нажатия на кнопку **ОК** программа выдает окно с результатами оценивания. Результаты представлены на рисунке 50.

EViews - [Equation: EQ04 Workfile: FLAT]				
File Edit Objects View Proc Quick Options Window Help				
View Proc Objects Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: PRICE				
Method: Least Squares				
Date: 11/11/04 Time: 19:59				
Sample: 1 205				
Included observations: 199				
Excluded observations: 6				
Weighting series: 1/SGM				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.341018	4.476058	0.523009	0.6016
BAL	1.576642	0.723848	2.178141	0.0306
BRICK	1.378808	1.087666	1.267676	0.0206
DIST	-0.757253	0.082550	-9.173281	0.0000
FLOOR	3.711416	0.858049	4.325410	0.0000
KITSP	1.283772	0.209694	6.122124	0.0000
LIVSP	0.797807	0.195531	4.080206	0.0001
METRDIST	-0.247200	0.080801	-3.059366	0.0025
WALK	4.449723	0.777066	5.726315	0.0000
OSTS	1.484688	0.185936	7.984929	0.0000
Weighted Statistics				
R-squared	0.963928	Mean dependent var	36.26212	
Adjusted R-squared	0.962210	S.D. dependent var	26.36512	
S.E. of regression	5.125298	Akaike info criterion	6.155199	
Sum squared resid	4964.781	Schwarz criterion	6.320692	
Log likelihood	-602.4423	F-statistic	51.29824	
Durbin-Watson stat	1.762927	Prob(F-statistic)	0.000000	

Рисунок 50 – Результаты оценивания параметров регрессионной модели взвешенным методом наименьших квадратов.

Оценка уравнения регрессии имеет вид:

$$\hat{y} = 2,34 + 1,57BAL + 1,37BRICK - 0,757DIST + 3,711FLOOR + 1,28KITS + 0,798LIVSP - 0,247METRDIST + 1,48OSTSP + 4,45WALK. \quad (4)$$

Модель является значимой, значимы все коэффициенты модели. Кроме того, улучшилось качество модели, так как повысилось значение коэффициента детерминации с 68,3% до 96,4%.

Проведем снова тест White, чтобы убедиться в отсутствии гетероскедастичности. На рисунке 51 представлены результаты оценивания, которые подтверждают отсутствие гетероскедастичности в модели (4).

White Heteroskedasticity Test:			
F-statistic	0.958235	Probability	0.865731
Obs*R-squared	9.927352	Probability	0.762567

Рисунок 51 – Результаты проверки гипотезы об отсутствии гетероскедастичности с помощью теста White

Список использованных источников

1 Айвазян, С.А. Прикладная статистика и основы эконометрики [Текст]: учебник для вузов / С.А. Айвазян, В.С. Мхитарян. – М.: ЮНИТИ, 1998. – 1022с.

2 Магнус, Я.Р. Эконометрика. Начальный курс [Текст]: учебник / Я.Р. Магнус, П.К. Катышев, А.А. Пересецкий. – 6-е изд., перераб. и доп. – М.: Дело, 2004. – 576 с.

3 Тихомиров, Н.П. Эконометрика [Текст]: учебник / Н.П. Тихомиров, Е.Ю. Дорохина. – М.: Издательство «Экзамен», 2003. – 512 с.