

С.А. Балашова, И.В. Лазанюк

ЭКОНОМЕТРИКА В ЗАДАЧАХ И РЕШЕНИЯХ

EViews 7

ESTIMATION • FORECASTING • STATISTICAL ANALYSIS
GRAPHICS • DATA MANAGEMENT • SIMULATION



Москва
2017

С.А. Балашова, И.В. Лазанюк

ЭКОНОМЕТРИКА В ЗАДАЧАХ И РЕШЕНИЯХ

Учебное пособие для магистров

**Второе издание,
исправленное и дополненное**

**Москва
Российский университет дружбы народов
2017**

331.57
Б20

УДК 330.43:004.4(075.8)
ББК 65в6
Б20

Утверждено
РИС Ученого совета
Российского университета
дружбы народов

Рецензенты:

доктор экономических наук, профессор *Н.В. Бондарчук*;
ведущий научный сотрудник ЦЭМИ РАН
доктор экономических наук, профессор *Е.Ю. Хрусталева*

Балашова, С.А.

Б20 Эконометрика в задачах и решениях : учебное пособие для магистров / С. А. Балашова, И. В. Лазанюк. – 2-е изд., переп. и доп. – Москва : РУДН, 2017. – 188 с. : ил.

ISBN 978-5-209-07893-7

В учебном пособии подробно разбираются примеры эконометрических исследований, основанные на фактическом эмпирическом материале, которые иллюстрируют применение тех или иных эконометрических методов к решению экономических задач. Приводятся теоретические предпосылки рассматриваемых моделей, вытекающие из экономической теории, и обоснование применяемых эконометрических методов. За более подробными разъяснениями читатели отсылаются к рекомендованным себе учебникам по эконометрике и экономической теории.

Весьма подробно проиллюстрировано применение специализированного программного обеспечения Eviews 7 для проведения эконометрического анализа, наиболее трудные этапы решения снабжены пошаговыми инструкциями.

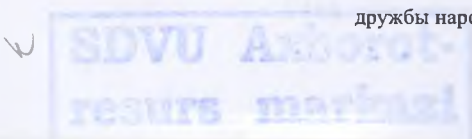
Последовательность изложения соответствует базовому учебнику по курсу М.Вербик «Путеводитель по современной эконометрике», данное пособие является хорошим дополнением к нему.

Для магистров первого года обучения, изучающих курс «Эконометрика. Продвинутый уровень», а также всех желающих повысить свой уровень владения эконометрическими методами.

УДК 330.43:004.4(075.8)
ББК 65в6

ISBN 978-5-209-07893-7

© Балашова С.А., Лазанюк И.В., 2017
© Российский университет
дружбы народов, 2017



Оглавление

| | |
|--|-----|
| ВВЕДЕНИЕ | 5 |
| Глава 1. Моделирование почасовой заработной платы. Линейная модель регрессии | 9 |
| Глава 2. Модель ценообразования финансовых активов (CAPM) как модель регрессии..... | 27 |
| Глава 3. Моделирование почасовой заработной платы (альтернативные модели). Выбор спецификации..... | 40 |
| Глава 4. Моделирование зависимости спроса от цены методом инструментальных переменных | 52 |
| Глава 5. Анализ временного ряда индекса цен на непродовольственные товары. Одномерные модели временных рядов | 66 |
| Глава 6. Моделирование волатильности обменных курсов. ARCH и GARCH модели..... | 84 |
| Глава 7. Моделирование совокупного потребления на основе динамических моделей временных рядов | 100 |
| Глава 8. Коинтеграция и модель коррекции ошибок для оценки потребления..... | 109 |
| Использованная литература..... | 121 |
| ПРИЛОЖЕНИЯ | 122 |
| Список используемых команд Eviews 7.0 | 122 |
| Перевод некоторых английских эконометрических терминов | 124 |
| Список некоторых обозначений в программе Eviews..... | 147 |
| Алгоритм используемых в работах тестов..... | 148 |
| Тест на проверку значимости коэффициентов регрессии | 149 |
| Тест на проверку значимости уравнения в целом | 150 |
| Тест на нормальность распределения..... | 153 |
| Тест на проверку ограничений на коэффициенты регрессии..... | 156 |

| | |
|--|-----|
| Тесты на гетероскедастичность..... | 157 |
| Тест Уайта..... | 157 |
| Тест Голдфелда–Квандта | 159 |
| Тесты на автокорреляцию..... | 159 |
| Тест Дарбина–Уотсона на автокорреляцию | 159 |
| Тест Льюнга–Бокса | 160 |
| Тесты на спецификацию | 160 |
| RESET-тест Рамсея | 160 |
| PE-тест..... | 161 |
| Тест Дики–Фуллера (ADF-тест)..... | 162 |
| Описание и программа курса..... | 163 |
| Вопросы для самопроверки..... | 172 |
| Творческая работа по дисциплине | 174 |
| Перечень тем для презентаций | 176 |
| Перечень вопросов в тестовой форме..... | 178 |
| Перечень вопросов итоговой аттестации..... | 181 |

ВВЕДЕНИЕ

Данное учебное пособие ориентировано на специфику преподавания продвинутого курса эконометрики на экономических специальностях. Оно разработано на основе опыта преподавания этой дисциплины в Российском университете дружбы народов студентам магистратуры экономических специальностей. Пособие охватывает основные темы курса и может быть полезно как магистрантам, изучающим этот предмет, так и аспирантам, применяющим в своих исследованиях эконометрические методы.

Эконометрика как наука получила свое развитие с середины XX-го века прошлого столетия. Наиболее точным представляется определение эконометрики как научной дисциплины, данное С.А. Айвазяном в учебнике «Методы эконометрики» (Айвазян, 2014) [1]: «Эконометрика - это самостоятельная научная дисциплина, объединяющая совокупность теоретических результатов, приемов, методов и моделей, предназначенных для того, чтобы на базе экономической теории, экономической статистики, математико-статистического инструментария придавать конкретное количественное выражение общим (качественным) закономерностям экономической теории». Применение современных методов эконометрических исследований представляется сегодня стандартной практикой в различных областях экономики. Такое стремительное развитие этой науки обусловлено появлением на рынке современных технологий большого количества программных средств для проведения эконометрических расчетов (STATA, SPSS, Eviews и др.).

Свидетельством всемирного признания эконометрики является присуждение Нобелевских премий по экономике ученым, внесшим значительный вклад в развитие этой дисциплины: премия 1969 г. была присуждена Р. Фришу и Я. Тинбергену, основоположникам эконометрики, за разработку математических методов анализа экономических процессов; премия 1980 г. — Л. Клейну за создание эконометрических моделей и их применение к анализу экономических колебаний и экономической политике; премия 1981 г. - Джеймсу Тобину, создателю регрессионных моделей с цензурированной зависимой переменной; премия 1989 г. — Т. Хаавелмо за прояснение вероятностных основ эконометрики и анализ одновременных экономических структур; премия 1995 г. - Р.Лукасу, который подверг конструктивной критике традиционные макроэкономические модели и дал толчок развитию новой области эконометрики, в основе которой лежит теория рациональных ожиданий; премия 2000 г. - Дж. Хекману за развитие теории и методов анализа селективных выборок и Д. Мак-Фаддену за развитие теории и методов анализа моделей дискретного выбора; премия 2003 г. — Р. Инглу и К. Грэнджеру. Оба ученых награждены за разработку методов макроэкономического анализа: Р. Ингл — за создание моделей с меняющейся волатильностью ARCH-моделей, а К. Грэнджер - метода коинтеграции. В 2011 г, Нобелевскую премию по экономике получили Томас Сарджент и Кристофер Симс за исследование причинно-следственных связей в макроэкономике, создание неструктурных методов макроэкономического прогнозирования, позволяющих рассчитать связь между экономической политикой и такими макроэкономическими параметрами, как ВВП, инфляция, безработица, инвестиции, что является одной из задач эконометрики как научной дисциплины.

Учебная дисциплина «Эконометрика (продвинутый уровень)» в соответствии с Федеральным государственным

образовательным стандартом высшего профессионального образования относится к базовой части профессионального цикла основной образовательной программы магистратуры по направлению экономика и читается во всех ВУЗах, реализующих это направление подготовки. В настоящее время на российском рынке учебной литературы присутствует большое разнообразие учебников по эконометрике, но ощущается острая нехватка учебных пособий с решением типовых практических задач, методическим описанием применения прикладного программного обеспечения и анализом экономического содержания полученных результатов. Настоящее пособие является попыткой внести некоторый вклад в создание такой учебно-методической литературы, которая является не только необходимым дополнением к аудиторным занятиям, но и полезна для самостоятельного освоения студентами этого курса.

Данное учебное пособие является продолжением учебного пособия Матюшка В.М., Балашовой С.А., Лазанюк И.В. «Основы эконометрического моделирования», которое ориентировано на изучение базового курса эконометрики.

Учебное пособие состоит из 8 глав, посвященных решению прикладных задач. Последовательность глав соответствует программе дисциплины, читаемой на экономическом факультете РУДН, и базовому учебнику по курсу М.Вербик «Путеводитель по современной эконометрике». Изложение материала начинается с повторения тем вводного курса (парная и множественная классическая линейная регрессия, главы 1-2), и постепенно усложняется, переходя к построению моделей множественной регрессии с фиктивными переменными взаимодействия, анализу отбора факторов в модель и выбору спецификации (глава 3), использованию метода инструментальных переменных (глава 4) при нарушении предпосылок классической регрессии. Затем следует изучение методов анализа одномерных (главы 5 - 6) и многомерных

временных рядов (главы 7 - 8). Модели авторегрессии - скользящего среднего рассматриваются в главе 5, глава 6 посвящена моделям с условной гетероскедастичностью. Построение и оценка динамических моделей, таких как моделирование совокупного потребления на основании гипотез Фридмана и Брауна, рассматривается в главе 7. Последняя глава посвящена моделированию нестационарных временных рядов, коинтеграции и модели коррекции ошибок.

Каждая глава данного пособия сопровождается необходимым теоретическим материалом и методическими рекомендациями для получения студентами практических навыков использования специализированного программного обеспечения Eviews 7.0, в конце каждой главы представлен перечень ключевых понятий и вопросов для повторения. Файлы с исходными данными для выполнения заданий расположены на учебном портале *Economist.rudn.ru* в разделе дисциплины «Эконометрика» (для магистров). В приложении приведены основные команды, необходимые при работе с пакетом Eviews 7, а также перевод основных английских терминов, встречающихся при работе с программой. Также в приложении изложены алгоритмы проведения тестов, которые используются в работах. Второе издание дополнено заданиями в тестовой форме.

Глава 1. Моделирование почасовой заработной платы. Линейная модель регрессии

Цель работы: применение метода наименьших квадратов для оценки простой (парной) и множественной моделей линейной регрессии. Интерпретация и сравнение моделей с разным количеством факторов. Оценка гендерных эффектов.

Исходные данные (wages1.wfl):

SCHOOL = количество лет обучения, лет

EXPER = опыт работы после обучения, лет

MALE = фиктивная переменная, равна 1 для мужчин и 0 для женщин, безразмерная

WAGE = почасовая заработная плата, долларов/час

Данные американского Национального периодического обследования (NLS) за 1987г., выборка из 3296 молодых работников, цитируется по (Вербик, 2008).

Задания

1. Определите дескриптивные (описательные) статистики и корреляционную матрицу для исходных переменных (отдельно для мужчин и для женщин). Проинтерпретируйте полученные результаты.
2. Проведите тест на равенство средних значений почасовой заработной платы для мужчин и женщин. Сделайте выводы.
3. Оцените модель вида

$$WAGE = \beta_1 + \beta_2 MALE + \varepsilon \quad (1.1)$$

Проинтерпретируйте коэффициенты уравнения.

4. Проинтерпретируйте коэффициент детерминации, проверьте на значимость уравнение и коэффициенты уравнения.
5. Постройте 95% доверительный интервал для коэффициента регрессии. Дайте интерпретацию полученным значениям.
6. Оцените модель влияния человеческого капитала на заработную плату (human capital earnings function) в линейной форме

$$WAGE = \beta_1 + \beta_2 MALE + \beta_3 SCHOOL + \beta_4 EXPER + \varepsilon \quad (1.2)$$

- Проинтерпретируйте коэффициенты регрессии.
7. Проведите тест Уальда на значимость улучшения оценки за счет введения новых переменных. Сделайте вывод.
8. По оцененному уравнению (1.2) получите расчетные значения, постройте графики расчетных и фактических значений заработной платы, проведите визуальное сравнение. Охарактеризуйте точность расчетных значений по относительной ошибке аппроксимации $MARE$.

Решения

Загрузите с портала *Economist.rudn.ru* файл *wages1.wfl*.

Откройте программу Eviews и в меню **File** → **Open** → **Eviews Workfile...** выберите файл *wages1.wfl*.

1. Под дескриптивными (или описательными) статистиками понимают группу статистических показателей, которые характеризуют наблюдаемую совокупность и дают количественную оценку ее свойств. К ним относятся меры среднего уровня (арифметическое среднее (*mean*), медиана (*median*), минимум (*min*), максимум (*max*)) и меры рассеяния (стандартное отклонение (*standard deviation, std.dev.*), дисперсия (*variance, var.*)). Для характеристики распределения случайной величины определяют коэффициент асимметрии (*skewness*) и эксцесса (или кurtosis, *kurtosis*). Эти коэффициенты дают возможность определить, в какой степени наблюдаемая совокупность может рассматриваться как выборка нормально распределенной случайной величины. Алгоритм теста на нормальность распределения подробно изложен в Приложении.

Для предварительного анализа данных, представленных для моделирования, надо получить описательные статистики, как для всей совокупности наблюдений, так и для двух групп данных.

Для определения дескриптивных (описательных) статистик отдельно для мужчин ($MALE=1$) и для женщин ($MALE=0$) выполните следующую последовательность действий.

Выделите переменные *EXPER*, *SCHOOL* и *WAGE* (при нажатой клавише **Ctrl**), нажмите правой клавишей мыши и выберите **Open** → **as Group**. Для последующих действий группу надо со-

хранить, выбрав в меню группы команду **Name**. Задайте имя по умолчанию (*Group01*) (рис.1.1).

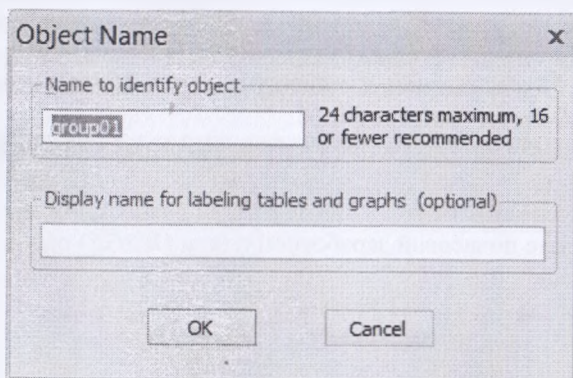


Рис. 1.1. Задание имени для объекта Группа

В верхней части рабочего листа дважды щелкните по надписи **Sample** и в открывшемся диалоговом окне в поле IF задайте *MALE=1* (рис.1.2). В результате выборка ограничится 1727 наблюдениями (только для мужчин).

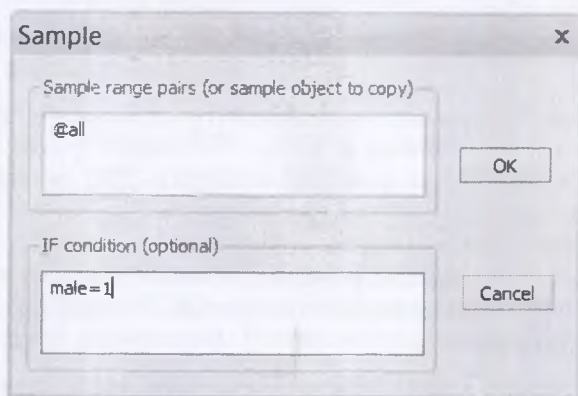


Рис. 1.2. Ограничение выборки только респондентами-мужчинами

Для получения результатов вычисления описательной статистики в меню группы **View** выберите команду **Descriptive Stats** → **Common Sample**. Получится таблица со значениями описательных статистик для каждой переменной. Для сохранения результатов выберите команду **Freeze**, будет создана таблица с результатами расчетов. Чтобы сохранить таблицу нажмите **Name** и задайте имя таблицы *Descript_male* (рис. 1.3).

Как видно из полученной таблицы, для мужчин среднее значение опыта работы (*EXPER*) составляет 8.32 года, среднее значение количества лет обучения (*SCHOOL*) равно 11.44 года, а среднее значение почасовой заработной платы (*WAGE*) равно 6.42 долларов в час.

| | A | B | C | D |
|----|--------------------------|-------------|-----------|----------|
| 1 | Date: 07/02/14 | Time: 13:25 | | |
| 2 | Sample: 1 3296 IF MALE=1 | | | |
| 3 | | | | |
| 4 | | EXPER | SCHOOL | WAGE |
| 6 | Mean | 8.323104 | 11.44181 | 6.424610 |
| 7 | Median | 8.000000 | 12.00000 | 5.681713 |
| 8 | Maximum | 16.00000 | 16.00000 | 112.7919 |
| 9 | Minimum | 2.000000 | 3.000000 | 0.153480 |
| 10 | Std Dev | 2.359826 | 1.736377 | 4.805478 |
| 11 | Skewness | 0.262091 | -0.474387 | 10.32415 |
| 12 | Kurtosis | 3.789847 | 4.244557 | 202.2089 |
| 14 | Jarque-Bera | 64.86350 | 176.2326 | 2886289 |
| 15 | Probability | 0.000000 | 0.000000 | 0.000000 |
| 17 | Sum | 14374.00 | 19760.00 | 11095.30 |
| 18 | Sum Sq Dev | 9611.708 | 5203.902 | 39857.85 |
| 19 | | | | |
| 20 | Observations | 1727 | 1727 | 1727 |

Рис. 1.3. Результаты расчета компонент описательной статистики для респондентов-мужчин, сохраненные в виде таблицы

Для расчета коэффициентов корреляции между переменными в меню группы **View** выберите команду **Covariance Analysis**. Снимите флажок напротив значения **Covariance** и поставьте флажок напротив значения **Correlation**. Включите также расчет *t*-

статистик (**t-statistic**) и **P-значения (Probability | t | = 0)** для проверки значимости рассчитываемых коэффициентов корреляции. Остальные условия оставьте по умолчанию (рис. 1.4). В результате будет получена корреляционная матрица. Результаты сохраните, выполнив последовательно команды **Freeze** → **Name**, задайте имя таблицы *Correl_male*.

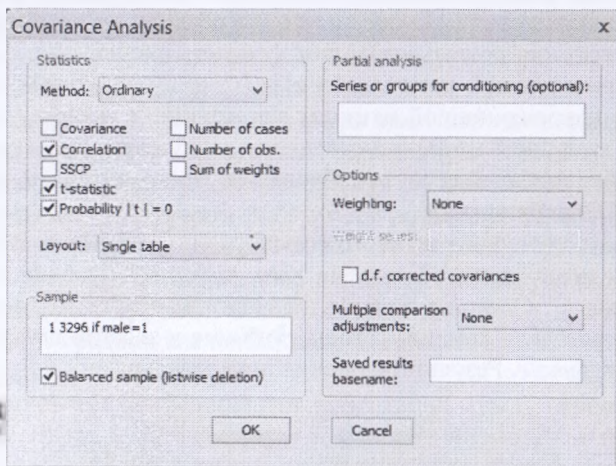


Рис. 1.4. Задание условий для расчета компонент матрицы парных корреляций

Полученные значения коэффициентов корреляции говорят о том, что между заработной платой и количеством лет обучения слабая прямая связь (коэффициент корреляции между *WAGE* и *SCHOOL* равен 0.29^1 , статистически значим на любом уровне значимости). Слабую отрицательную корреляционную связь имеют опыт работы и количество лет обучения (коэффициент корреляции между *EXPER* и *SCHOOL* равен -0.27 , статистически значим на любом уровне значимости). Незначимо отличается от нуля коэф-

¹ В тексте используется точка в качестве разделителя целой и десятичной части для согласованности обозначений со скриншотом программы.

коэффициент корреляции между опытом работы и заработной платой для выборки респондентов-мужчин.

Замечание. Коэффициент парной корреляции характеризует только тесноту линейной связи между двумя переменными (в случае совместной нормальной распределенности двух переменных (Айвазян, 2014)). При наличии «множественных» корреляционных связей или нелинейного характера зависимости между переменными знание только коэффициента корреляции недостаточно для характеристики тесноты связи между переменными.

Для получения результатов вычисления описательной статистики и корреляционной матрицы для женщин в меню группы выберите команду **Sample** и в открывшемся диалоговом окне в поле IF задайте $MALE=0$. Будут отображены наблюдения только для женщин (1569 наблюдений).

Для получения результатов вычисления описательной статистики в меню группы *Group01* View выберите команду **Descriptive Stats** → **Common Sample**. Результаты сохраните, выполнив последовательно команды **Freeze** → **Name** и задайте имя таблицы *Descript_female*. Результаты оценки приведены на рис. 1.5.

| View | Proc | Object | Print | Name | Edit+/- | CellFmt | Grid+/- | Title | Comments- |
|------|------|--------|-------|------|---------|---------|---------|-------|-----------|
| 1 | | | | A | | | | | |
| 2 | | | | B | | | | | |
| 3 | | | | | | | | | |
| 4 | | | | | | C | | | |
| 5 | | | | | | | | D | |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |
| 10 | | | | | | | | | |
| 11 | | | | | | | | | |
| 12 | | | | | | | | | |
| 13 | | | | | | | | | |
| 14 | | | | | | | | | |
| 15 | | | | | | | | | |
| 16 | | | | | | | | | |
| 17 | | | | | | | | | |
| 18 | | | | | | | | | |
| 19 | | | | | | | | | |
| 20 | | | | | | | | | |
| 21 | | | | | | | | | |

Рис. 1.5. Результаты расчета компонент описательной статистики для респондентов-женщин, сохраненные в виде таблицы

Как видно из полученной таблицы, для женщин среднее значение опыта работы (*EXPER*) составляет 7.73 года, среднее значение количества лет обучения (*SCHOOL*) равно 11.84 года, а среднее значение почасовой заработной платы (*WAGE*) равно 5.15 долларов в час. Таким образом, в среднем мужчины и женщины имеют практически равное значение количества лет обучения, близкие значения опыта работы (мужчины имеют в среднем на 0.6 года больше опыт работы). Однако мужчины существенно больше зарабатывают: среднее значение и медиана у мужчин значительно выше, чем у женщин. Это может объясняться наличием в выборке для мужчин двух наблюдений с очень высокими почасовыми заработными платами (если открыть переменную *WAGE* и выбрать сортировку по убыванию **Sort → Wage → Descending**, то видно, что респондент с номером 3026 имеет почасовой заработок 112.8 долл., а респондент с номером 2932 - 92.5 долл.). Но может присутствовать и дискриминация по половому признаку (так называемый гендерный эффект). Однако прежде чем искать причину различия двух средних значений, надо убедиться в том, что это различие статистически значимо.

Для расчета коэффициентов корреляции между переменными в меню группы **View** выберите команду **Covariance Analysis**, установите параметры, как показано на рис.1.4. Результаты сохраните, выполнив последовательно команды **Freeze → Name**, задайте имя таблицы *Correl_female* (рис. 1.6).

Заметим, что для женщин практически отсутствует линейная связь между опытом работы и количеством лет обучения (коэффициент корреляции равен -0.06, $tstat = -2.24$, $Prob=0.025$, т.е. коэффициент незначим на 1%-м уровне, хотя значим на 5% -м уровне) и присутствует статистически значимая, хотя и очень слабая прямая связь между опытом работы и заработной платой (коэффициент корреляции равен 0.09, $tstat = 3.58$, $Prob=0.0004$, т.е. коэффициент значим даже на уровне 1%).

Таким образом, парные коэффициенты корреляции между заработной платой и объясняющими переменными различны для женщин и мужчин. Однако, как отмечалось ранее, парные коэффициенты корреляции указывают только на характер линейной связи между двумя переменными, и их знания недостаточно, что-

бы сделать окончательный вывод о степени взаимосвязи факторов при наличии множественных или нелинейных зависимостей.

| View | Proc | Object | Print | Name | Edit +/- | CellFmt | Grid +/- | Title | Comments +/- |
|------|---|--------|-------|-----------|----------|----------|----------|----------|--------------|
| A | | | B | | C | | D | | |
| 1 | Covariance Analysis: Ordinary | | | | | | | | |
| 2 | Date: 14/02/14 Time: 22:02 | | | | | | | | |
| 3 | Sample (adjusted): 1 1569 | | | | | | | | |
| 4 | Included observations: 1559 after adjustments | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | Correlation | | | | | | | | |
| 7 | t-Statistic | | | | | | | | |
| 8 | Probability | | | | | | | | |
| 9 | | EXPER | | 1.000000 | | SCHOOL | | WAGE | |
| 10 | | | | ---- | | | | | |
| 11 | | | | ---- | | | | | |
| 12 | | SCHOOL | | -0.056524 | | 1.000000 | | | |
| 13 | | | | -2.241111 | | | | | |
| 14 | | | | 0.0252 | | | | | |
| 15 | | | | | | | | | |
| 16 | | WAGE | | 0.090110 | | 0.294314 | | 1.000000 | |
| 17 | | | | 3.581615 | | 12.19046 | | | |
| 18 | | | | | | | | | |
| 19 | | | | 0.0004 | | 0.0000 | | | |
| 20 | | | | | | | | | |

Рис. 1.6. Результаты расчета компонент матрицы парных корреляций для респондентов-женщин

Для дальнейшей работы закройте группу и все открытые объекты и вернитесь к полной выборке. Для этого в командной строке введите команду

`smpl @all` ←

2. Как показывают результаты расчета компонент описательной статистики, средние значения фактора *WAGE* различны для подвыборок, состоящих только из респондентов-мужчин и только респондентов-женщин. Однако полученные значения являются выборочными средними и рассчитаны с ошибками, величина которых зависит от дисперсии оценки среднего. Поэтому, чтобы сделать вывод о равенстве/неравенстве средних, надо провести тест с использованием критерия Стьюдента². Идея теста состоит в том, чтобы взять разность двух выборочных средних и нормировать их подходящим образом (поделив на стандартное отклонение). Нормирован-

² Стьюдент — псевдоним английского статистика В.Госсета

ную таким образом разность можно сравнивать с табличным значением.

Для проведения теста на равенство средних значений заработной платы у мужчин и женщин откройте двойным щелчком переменную *WAGE*, и в меню **View** выберите последовательность команд **Descriptive Statistics & Tests** → **Equality Tests by Classification**. Откроется диалоговое окно, в котором надо ввести переменную *MALE* в поле **SERIES/GROUP FOR CLASSIFY**. Так как тестируется равенство средних значений, то переключатель должен стоять в позиции **Mean**. Результаты теста отображаются в виде таблицы, которую надо сохранить, выполнив последовательность команд **Freeze** → **Name**, задайте имя таблицы *Wage_test*.

Как видно из результатов, приведенных на рис. 1.7, нулевая гипотеза о равенстве теоретических средних значений заработной платы для мужчин и женщин отвергается на любом уровне значимости по t-критерию Стьюдента, а также и по результатам дисперсионного анализа (ANOVA³ F-test и его модификации). Таким образом, можно говорить о статистически значимом различии в средних значениях заработной платы у мужчин и женщин.

| View | Proc | Object | Print | Name | Edit +/- | CellFmt | Grid +/- | Title | Comments +/- |
|------|---|--------------|-----------|-------------|----------|---------|----------|-------|--------------|
| | A | B | C | D | E | | | | |
| 1 | Test for Equality of Means of WAGE | | | | | | | | |
| 2 | Categorized by values of MALE | | | | | | | | |
| 3 | Date: 07/02/14 Time: 13:28 | | | | | | | | |
| 4 | Sample: 1 3296 | | | | | | | | |
| 5 | Included observations: 3296 | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | Method | df | Value | Probability | | | | | |
| 8 | | | | | | | | | |
| 9 | t-test | 3294 | -9.147693 | 0.0000 | | | | | |
| 10 | Satterthwaite-Welch t-test* | 2865.168 | -9.357350 | 0.0000 | | | | | |
| 11 | Anova F-test | (1, 3294) | 83.68028 | 0.0000 | | | | | |
| 12 | Welch F-test* | (1, 2865.17) | 87.55999 | 0.0000 | | | | | |
| 13 | | | | | | | | | |
| 14 | *Test allows for unequal cell variances | | | | | | | | |

DB = none WF = wages1_done

Рис. 1.7. Результаты теста на равенство средних значений заработной платы для респондентов-мужчин и респондентов-женщин, сохраненные в виде таблицы (фрагмент)

³ ANOVA - Analysis of Variance

3. Другой подход к проверке статистической значимости различия средних значений заработной платы у мужчин и женщин заключается в оценке линейного уравнения парной регрессии, где зависимой переменной является заработная плата *WAGE*, а единственным фактором – фиктивная переменная *MALE* (уравнение (1.1)). Для оценки уравнения в Главном меню выберите команду **Quick → Estimate Equation**, и в открывшемся диалоговом окне наберите через пробел *WAGE C MALE*

(зависимая переменная, свободный коэффициент и объясняющая переменная). В поле **Method** оставьте значение по умолчанию (LS-Least Squares - метод наименьших квадратов) и в поле **Sample** – значения 1 3296. Для дальнейшей работы с уравнением его нужно сохранить, выбрав команду **Name** (имя по умолчанию *Eq01* можно при желании изменить). В результате в окне рабочего файла появится новый объект с соответствующим именем (рис. 1.8).

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------------|-------------|-----------------------|-------------|--------|
| C | 5.146924 | 0.101103 | 50.90761 | 0.0000 |
| MALE | 1.277686 | 0.139673 | 9.147693 | 0.0000 |
| R-squared | 0.024774 | Mean dependent var | 5.816391 | |
| Adjusted R-squared | 0.024478 | SD. dependent var | 4.054694 | |
| S.E. of regression | 4.004760 | Akaike info criterion | 5.613451 | |
| Sum squared resid | 52829.51 | Schwarz criterion | 5.617153 | |
| Log likelihood | -9248.967 | Hannan-Quinn criter. | 5.614776 | |
| F-statistic | 83.68028 | Durbin-Watson stat | 1.929290 | |
| Prob(F-statistic) | 0.000000 | | | |

Рис. 1.8. Результаты оценки уравнения парной регрессии (1.1)

Оцененное уравнение, таким образом, имеет вид

$$WAGE^{\wedge} = 5.15 + 1.28MALE$$

Свободный коэффициент показывает, что без учета множества факторов, включенных в случайную составляющую в данной модели, среднее значение заработной платы для женщин ($MALE=0$) составляет 5.15, а для мужчин ($MALE=1$) $5.15 + 1.28 = 6.42$ долларов в час (что, конечно, совпадает с результатом расчета средних по выборке значений). Различие в средних зарплатах мужчин и женщин статистически значимо, так как коэффициент при факторе $MALE$ значим на любом уровне ($tstat=9.15$, $Prob.=0.0000$).

4. Коэффициент детерминации, значимость уравнения в целом и значимость коэффициентов определяются по таблице результатов оценки уравнения (рис. 1.8).

Для данного уравнения коэффициент детерминации очень мал $R^2 = 0.025$, что говорит о том, это фактором $MALE$ объясняется только 2.5% разброса в заработной плате. Однако, достаточно высокое значение статистики Фишера ($Fstat = 83.7$, $Prob. = 0.0000$) свидетельствует о статистической значимости уравнения в целом. И свободный коэффициент, и коэффициент при факторе $MALE$ статистически значимы на любом уровне.

5. Для получения интервальных оценок значений коэффициента регрессии при факторе $MALE$ создайте новый объект – вектор $b2$ размерности 3 и запишите в него расчетные значения коэффициента (в обозначениях Eviews это коэффициент $c(2)$), его нижней и верхней границы. Для этого в командной строке введите последовательно

```
vector (3) b2 ←
```

```
b2(1)=c(2) ←
```

```
b2(2)=c(2)-@stderrs(2)*1.96 ←
```

```
b2(3)=c(2)+@stderrs(2)*1.96 ←
```

Замечание. $@stderrs(2)$ – стандартное отклонение коэффициента при факторе $MALE$. 1.96 – критическое значение статистики Стьюдента при числе степеней свободы больше 1000. Если число степеней свободы мало, надо использовать соответствующее табличное значение. Для уровня $\gamma\%$ двухсторонний критерий равен одностороннему на уровне $\gamma/2\%$. В Eviews критическое значение для двухстороннего теста на 5% уровне обозначается $@qtdist(0.975,df)$, где df – число степеней свободы.

Командная строка и результат вычислений компонент вектора $b2$ представлены на рис. 1.9. В результате расчета имеем: нижняя

граница доверительного интервала равна 1, а верхняя – 1.55, т.е. с вероятностью 95% среднее значение заработной платы у мужчин превосходит зарплату женщин на величину в интервале от 1 до 1.55 долларов в час без учета других факторов.

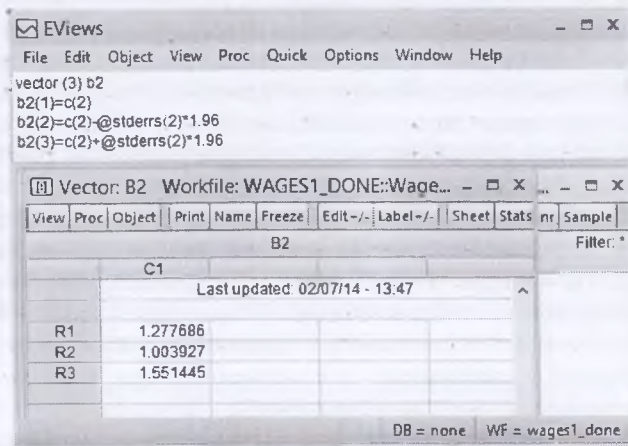


Рис. 1.9. Результат выполнения команд по расчету доверительного интервала коэффициента регрессии

6. Чтобы учесть такие факторы, как уровень образования и опыт работы, наряду с включением гендерного эффекта, оцените уравнение (1.2) с тремя факторами. Для этого в Главном меню выберите команду **Quick** → **Estimate Equation**, и в открывшемся диалоговом окне наберите через пробел

WAGE C MALE SCHOOL EXPER

(зависимая переменная, свободный коэффициент и объясняющие переменные). В поле **Method** оставьте значение по умолчанию (LS- Least Squares - метод наименьших квадратов) и в поле **Sample** – значения 1 3296 (или **@all**).

Для дальнейшей работы с уравнением его нужно сохранить, выбрав команду **Name** (*Eq02*). В результате в окне рабочего файла появится новый объект с соответствующим именем. Результаты оценки приведены на рис 1.10.

Оцененное уравнение, таким образом, имеет вид

$$WAGE^{\wedge} = -2.9 + 1.47MALE + 0.62SCHOOL + 0.09EXPER$$

Свободный коэффициент является отрицательным и в данной модели не имеет смысла, так как формально равен оцененному значению зависимой переменной ($WAGE^{\wedge}$) при равенстве нулю всех факторов (что, конечно, для переменной SCHOOL является чисто гипотетическим значением).

Коэффициент при факторе MALE несколько выше, чем в парной модели, и интерпретируется следующим образом: при одинаковых значениях факторов SCHOOL и EXPER мужчины в среднем (без учета других факторов) зарабатывают на 1.47 долларов в час больше, чем женщины. Это значение находится внутри доверительного интервала для соответствующего коэффициента парной регрессии (см. раздел Решения п. 5), т.е. оценки парной и множественной регрессии в отношении гендерного эффекта согласованы.

Коэффициент при факторе SCHOOL, равный 0.62, говорит о том, что при одинаковом опыте работы и одинаковом значении фактора MALE респонденты, которые учились на год дольше, получают в среднем на 0.62 доллара в час больше.

| Equation: EQ02 Workfile: WAGES1_DONE::Wages1\ | | | | | | | | | |
|---|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
| Dependent Variable: WAGE | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 07/02/14 Time: 14:27 | | | | | | | | | |
| Sample: 1 3296 | | | | | | | | | |
| Included observations: 3296 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| C | -2.890072 | 0.591789 | -4.883618 | 0.0000 | | | | | |
| MALE | 1.470217 | 0.137033 | 10.72890 | 0.0000 | | | | | |
| SCHOOL | 0.620407 | 0.041748 | 14.86063 | 0.0000 | | | | | |
| EXPER | 0.089617 | 0.030235 | 2.963956 | 0.0031 | | | | | |
| R-squared | 0.086104 | Mean dependent var | 5.816391 | | | | | | |
| Adjusted R-squared | 0.085271 | S.D. dependent var | 4.054694 | | | | | | |
| S.E. of regression | 3.877968 | Akaike info criterion | 5.549712 | | | | | | |
| Sum squared resid | 49507.19 | Schwarz criterion | 5.557116 | | | | | | |
| Log likelihood | -9141.925 | Hannan-Quinn criter | 5.552363 | | | | | | |
| F-statistic | 103.3869 | Durbin-Watson stat | 1.961224 | | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | |

Рис. 1.10. Результаты оценки уравнения трехфакторной регрессии (1.2)

Коэффициент при факторе *EXPER*, равный 0.09, говорит о том, что респонденты, которые имеют на один год больше опыта работы, получают в среднем на 0.09 доллара в час больше при одинаковом количестве завершенных лет обучения и одинаковом значении фактора *MALE*. При осмыслении полученного результата следует учесть, что в выборку включены респонденты в возрасте до 35 лет.

Уравнение *Eq02* имеет более высокое качество оценки, чем *Eq01*, если судить по коэффициенту детерминации. Однако чтобы ответить на вопрос, значимо ли это улучшение, надо провести тест на улучшение качества оценивания за счет включения в модель дополнительных факторов. Таким тестом может служить тест Уальда, который позволяет тестировать правомерность линейных ограничений на коэффициенты (в частности, равенство группы коэффициентов нулю)⁴.

7. Уравнение (1.1) можно рассматривать как частный случай уравнения (1.2), при котором коэффициенты при факторах *SCHOOL* и *EXPER* равны нулю одновременно. Математически такое условие является линейным ограничением на коэффициенты при этих факторах и может быть проверено по результатам оценки коэффициентов уравнения (1.2).

Для проведения теста Уальда откройте в меню уравнения *Eq02* команду **View** → **Coefficient Diagnostics** → **Wald Test – Coefficient Restrictions**. В окне записи ограничений введите $C(3)=C(4)=0$, что означает равенство нулю коэффициентов при факторах *SCHOOL* и *EXPER* (рис. 1.11). Используя команды **Freeze** → **Name**, сохраните результаты теста в виде таблицы (задайте имя *Wald_Eq02*).

⁴ Алгоритм проведения теста Уальда см в Приложении

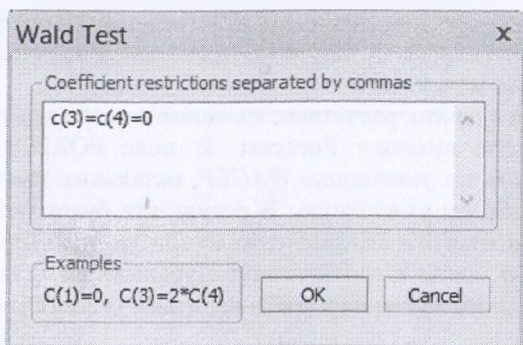


Рис. 1.11. Задание ограничений на коэффициенты при проведении теста Уальда

Как показывают результаты теста (рис. 1.12), нулевая гипотеза о правомерности ограничения $C(3)=C(4)=0$ отвергается ($F\text{stat} = 110.5$, $\text{Prob.}=0.0000$). Т.е. качество оценки переменной *WAGE* статистически значимо повысилось за счет введения двух факторов *SCHOOL* и *EXPER* (несмотря на низкие значения коэффициентов парной корреляции между этими факторами и зависимой переменной).

| View | Proc | Object | Print | Name | Edit +/- | CellFmt | Grid +/- | Title | Comm |
|------|------|--------|-------|------|----------|---------|----------|-------|------|
| 1 | | | | A | | | | B | |
| 2 | | | | | | | | C | |
| 3 | | | | | | | | D | |
| 4 | | | | | | | | | |
| 5 | | | | | | | | | |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |
| 8 | | | | | | | | | |
| 9 | | | | | | | | | |

Рис. 1.12. Результаты теста Уальда (фрагмент таблицы)

Таким образом, уравнение *Eq02* лучше описывает исходную выборку и его целесообразно использовать для получения оцененных значений зарплаты.

8. Для получения расчетных значений в меню уравнения *Eq02* выберите команду **Forecast**. В поле **FORECAST NAME** оставьте по умолчанию *WAGEF*, остальные значения также оставьте по умолчанию. В результате будет создана новая переменная *WAGEF*, содержащая расчетные значения заработной платы в соответствии с уравнением *Eq02*, и график расчетного значения и его доверительного интервала.

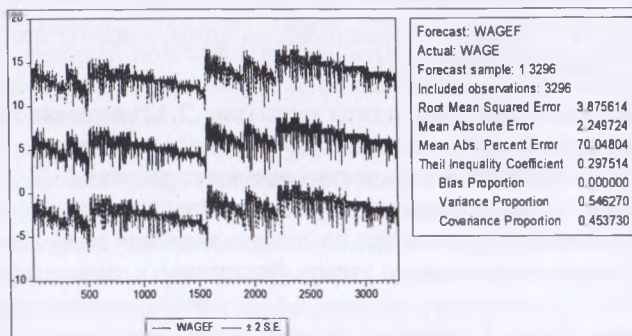


Рис. 1.13. Расчетное значение (средняя линия) и доверительный интервал зарплаты, полученный по уравнению с тремя факторами

Используя команды **Freeze** → **Name**, сохраните полученный график (*Graph01*).

Как видно из рис 1.13, доверительный интервал расчетного значения переменной *WAGE* достаточно широк, что является следствием низкого качества оценки уравнения *Eq02*. Видимый разрыв в графике связан с переходом от респондентов женщин к респондентам-мужчинам (начиная с 1570 наблюдения).

Для построения графика фактического и расчетного значений выделите при нажатой клавише **Ctrl** переменные *WAGE* и *WAGEF* и откройте их как группу (**Name** → *Group02*). В меню группы выберите **View** → **Graph...** Все параметры оставьте по умолчанию

(Basic type; Line & Symbol). График сохраните (Freeze → OK (параметры по умолчанию) → Name (Graph02).

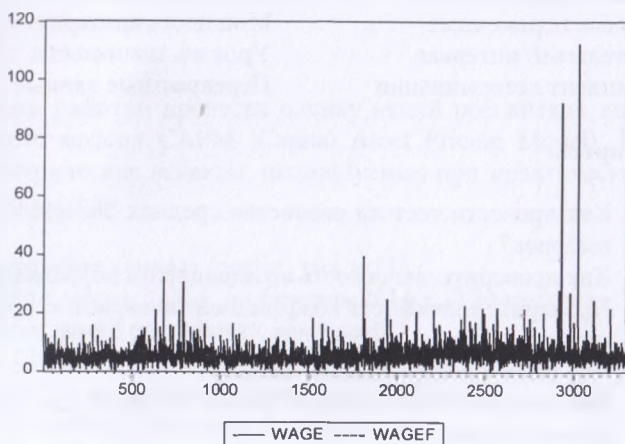


Рис. 1.14. Фактическое и расчетное значение зарплаты по уравнению с тремя факторами

Как видно из приведенного графика, anomalно большие значения заработной платы у двух респондентов-мужчин, на которые мы уже обращали внимание, никак не связаны с факторами, исключенными в уравнение, и в определенной степени отвечают и за низкое качество оценки, и за плохую точность расчетных значений (относительная ошибка аппроксимации $MAPE^5=70\%$, как видно из рис. 1.13). Но и для тех респондентов, зарплата которых только более чем на одно стандартное отклонение отличается от среднего, разница между фактическими и расчетными значениями достаточно велика. Расчетные значения близки к фактическим только для зарплат, мало отличающихся от средних, что, безусловно, говорит о низком качестве модели.

⁵ MAPE - mean absolute percent error

Ключевые понятия

Описательные статистики
Корреляционная матрица
Фиктивные переменные
Доверительный интервал
Коэффициент детерминации

Гипотеза нулевая, гипотеза
альтернативная
Мощность критерия
Уровень значимости
Перекрестные данные

Вопросы

1. Как провести тест на равенство средних значений у двух выборок?
2. Как проверить значимость коэффициента корреляции? Проверьте значимость коэффициентов парной корреляции, полученных в задании 1.
3. Что такое дисперсионный анализ?
4. Как построить 90% доверительный интервал для коэффициента регрессии?
5. В каких случаях используется тест Уальда?
6. Что такое относительная ошибка аппроксимации, как она рассчитывается? Каковы допустимые значения?

Глава 2. Модель ценообразования финансовых активов (CAPM) как модель регрессии

Цель работы: провести оценку акций российских компаний на основе модели CAPM (Capital Asset Pricing Model). Изучить эконометрические аспекты, используемые при практическом применении модели.

Исходные данные (capm_rus.wf1) :
за период с марта 2004 г. по февраль 2014 г. (120 месяцев)
котировки акций российских компаний:
LKOH (ЛУКОЙЛ – нефтяная компания)
MTSS (МТС – телекоммуникационная компания)
SBER (Сбербанк – банковский сектор)
и российского фондового индекса
MICEX
Безрисковая ставка
RF – доходность по трехмесячным казначейским векселям США, в процентах, в годовом выражении на 360-дневной базе
JAN – фиктивная переменная, принимающая значения 1 для января и 0 для остальных месяцев.

Источник данных: для котировок акций и индекса MICEX - www.finam.ru, для казначейских векселей - <http://www.federalreserve.gov/releases/h15/data.htm>

Задания

1. Постройте ряды простых годовых доходностей акций и индекса по правилу

$$r_j = \frac{p_{jt} - p_{jt-1}}{p_{jt-1}} * 12 * 100$$

где p_j - цена акций компании, $j=1$ (для LKOH), 2 (для MTSS), 3 (для SBER) или значение индекса MICEX (при $j=m$).

2. Оцените модель вида (без свободного члена)

$$r_j - r_f = \beta_j(r_m - r_f) + \varepsilon_j$$

(2.1)

где r_j - доходность акций компании, $j=1,2,3$, r_m - доходность рынка, оцененная по индексу MICEX, r_f - безрисковая доходность.

3. Протестируйте на 5%-м уровне гипотезу о равенстве коэффициента $\beta_j=1$ для каждой из компаний. В случае отклонения гипотезы H_0 протестируйте гипотезы о «защитном»/«агрессивном» характере рассматриваемых акций (проведите односторонний тест Стьюдента на уровне 1%). Сделайте выводы.
4. Включите в модель (1) свободный член:
$$r_j - r_f = \alpha + \beta_j(r_m - r_f) + \varepsilon_j \quad (2.2)$$
Оцените соответствующие регрессии. Протестируйте гипотезы о равенстве нулю свободного члена.
5. Дайте экономическую интерпретацию коэффициентов детерминации в рассмотренных регрессиях.
6. Проверьте наличие автокорреляции первого порядка по тесту Дарбина-Уотсона.
7. Проверьте гипотезу о существовании так называемого «эффекта января», т.е. предположение о том, что при прочих равных (*ceteris paribus*) доходность в январе отличается от доходности в другие месяцы.
8. Сделайте прогноз относительно ожидаемой доходности рассматриваемых акций на три месяца, если ожидается снижение индекса MICEX на 5 пунктов в месяц, а безрисковая доходность сохранится на уровне 0.04% годовых.

Решения

Загрузите с портала *Economist.rudn.ru* файл *capm_rus.wfl*.

Откройте программу Eviews и в меню **File** → **Open** → **Eviews**

Workfile... выберите файл *capm_rus.wfl*.

1. Для создания рядов простых годовых доходностей введите в командной строке

$$SERIES RM = (MICEX - MICEX(-1)) / MICEX(-1) * 12 * 100$$

В результате переменная RM появится в списке объектов вашего рабочего файла. Аналогично поступите для создания доходностей трех рассматриваемых акций.

2. Оцените уравнение (2.1) для акций трех различных компаний.

В Главном меню выберите команду **Quick → Estimate Equation**, и в открывшемся диалоговом окне наберите через пробел

$R1-RF$ $RM-RF$

(зависимая переменная $R1-RF$, объясняющая переменная $RM-RF$. *Без константы!*). В полях **Method** и **Sample** оставьте значение по умолчанию. Для дальнейшей работы с уравнением результат оценки нужно сохранить, выбрав команду **Name** (имя Eq01). В результате в окне рабочего файла появится новый объект с соответствующим именем.

Скопируйте полученное уравнение, используя в окне уравнения команду **Object → Copy object**. Появится точная копия результатов оценки уравнения (Equation: UNTITLED – Уравнение: без названия). В открывшемся окне выберите команду **Estimate** и измените зависимую переменную на $R2-RF$, т.е. в окне оценки уравнения должно быть записано

$R2-RF$ $RM-RF$

Нажмите ОК и дайте полученному уравнению имя Eq02.

Сделайте еще одну копию исходного уравнения, используя **Object → Copy object**. В открывшемся окне выберите команду **Estimate** и измените зависимую переменную на $R3-RF$, т.е. в окне оценки уравнения должно быть записано

$R3-RF$ $RM-RF$

Нажмите ОК и дайте полученному уравнению имя Eq03.

Таблица 2.1

Результаты оценки уравнения (2.1) для трех компаний

| Компания | Оценка бета-коэффициента | Стандартная ошибка оценки бета-коэффициента |
|----------|--------------------------|---|
| ЛУКОЙЛ | 0.86 | 0.055 |
| МТС | 0.77 | 0.066 |
| Сбербанк | 1.3 | 0.092 |

Полученные оценки бета-коэффициентов отличаются от 1 для всех компаний, однако, чтобы говорить о том, что это отличие статистически значимо, надо провести тестирование.

- Для тестирования гипотезы $\beta_j=1$, финансово-экономический смысл которой заключается в том, что акции компании ведут себя в соответствии с рынком, их ожидаемая доходность равна ожидаемой доходности рынка, а систематический риск равен рыночному, используйте тест Уальда.

Для тестирования характера поведения акций ЛУКОЙЛа откройте уравнение *Eq01*. Выполните последовательность команд **View** → **Coefficient Diagnostics** → **Wald Test-Coefficient Restrictions...** и в открывшемся диалоговом окне введите проверяемое ограничение

$$C(1) = 1.$$

Результат теста показывает, что для компании ЛУКОЙЛ гипотеза $H_0: \beta_1=1$ отвергается на уровне значимости более 1.5%, но не отвергается на уровне значимости 1%. Это следует из полученных значений статистик, которые рассчитываются для проверки ограничения (рис. 2.1).

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Res |
|--|-----------|-----------|-------------|------|--------|----------|----------|-------|-----|
| Wald Test | | | | | | | | | |
| Equation: EQ01 | | | | | | | | | |
| Test Statistic | Value | df | Probability | | | | | | |
| t-statistic | -2.568833 | 117 | 0.0115 | | | | | | |
| F-statistic | 6.598904 | (1, 117) | 0.0115 | | | | | | |
| Chi-square | 6.598904 | 1 | 0.0102 | | | | | | |
| Null Hypothesis: C(1)=1 | | | | | | | | | |
| Null Hypothesis Summary: | | | | | | | | | |
| Normalized Restriction (= 0) | Value | Std. Err. | | | | | | | |
| -1 + C(1) | -0.140877 | 0.054841 | | | | | | | |
| Restrictions are linear in coefficients. | | | | | | | | | |

Рис. 2.1. Результат теста Уальда для компании ЛУКОЙЛ

И для t-статистики, и для F-статистики и для χ^2 -статистики расчетные значения превосходят (по модулю в случае t-статистики) табличное значение на уровне 1.5%⁶, но меньше соответствующих значений на уровне 1%. Это следует также из приводимых P-значений (Prob.), которые в данном случае лежат в интервале от 0.01 до 0.015 ($0.01 < \text{Prob} < 0.015$).

Для компании ЛУКОЙЛ оцененное значение бета-коэффициента равно 0.86, поэтому имеет смысл проверить одностороннюю гипотезу о том, что истинный бета-коэффициент меньше 1 и акции этой компании можно отнести к «защитным». Для проверки этой гипотезы проведем односторонний (левосторонний) тест Стьюдента на уровне значимости 1%:

$$H_0: \beta_1 = 1$$

$$H_1: \beta_1 < 1$$

$$\text{Тстовая статистика } t_{stat} = (0.86 - 1) / 0.055 = -2.55$$

Табличное значение одностороннего критерия Стьюдента для числа степеней свободы $df=117$ (118 наблюдений, по которым оценен 1 коэффициент) и 1% уровня значимости равен $t_{kp}^{одн} = 2.36$

Для левостороннего теста гипотеза H_0 отвергается в том случае, когда $t_{stat} < -t_{kp}^{одн}$ (без модуля с учетом знака). Т.е. для компании ЛУКОЙЛ гипотеза H_0 отвергается, и акции компании можно отнести к менее рискованным (и менее доходным), чем рынок в целом на рассматриваемом временном интервале.

Аналогично проводятся тесты Уальда и Стьюдента для других компаний. Заметим, что для акций Сбербанка надо проводить правосторонний тест Стьюдента, так как оцененное значение бета-коэффициента больше 1.

Результаты тестов приведены в таблице 2.2.

⁶ Табличные значения можно найти в учебнике Большев Л.Н., Смирнов Н.В. «Таблицы математической статистики». М.: Наука. Главная редакция физико-математическом литературы, 1983 в полном объеме. Фрагменты таблиц приводятся во многих учебнике по эконометрике.

Таблица 2.2

Результаты тестов на значение бета-коэффициентов

| Компания | Тест Уальда $H_0: \beta_j=1$ $H_1: \beta_j \neq 1$ | Левосторонний тест Стьюдента $H_0: \beta_j=1$ $H_1: \beta_j < 1$ | Правосторонний тест Стьюдента $H_0: \beta_j=1$ $H_1: \beta_j > 1$ | Характер акций |
|----------|--|--|---|----------------|
| ЛУКОЙЛ | H_0 отвергается на уровне 1.5% | H_0 отвергается на уровне 1% | - | защитный |
| МТС | H_0 отвергается на уровне 1% | H_0 отвергается на уровне 1% | - | защитный |
| Сбербанк | H_0 отвергается на уровне 1% | - | H_0 отвергается на уровне 1% | агрессивный |

4. Оцените уравнение (2.2) для акций трех различных компаний, введя в каждый из них свободный коэффициент (константу). Согласно модели САМР оценка свободного коэффициента должна быть незначима (т.е. истинное значение коэффициента равно нулю). Оценка уравнения (2.2) со свободным коэффициентом позволяет, во-первых, убедиться в том, что эмпирические данные не противоречат модели, а, во-вторых, использовать стандартные эконометрические техники.

Для введения константы можно сначала сделать копию уже оцененного уравнения, а затем внести изменения в окне **Estimation**. Чтобы сделать копию уравнения, нажмите на его пиктограмму правой клавишей мыши и выберите команду **Object Copy**. Имя нового уравнения *Eq0101* можно оставить по умолчанию (рис. 2.2).

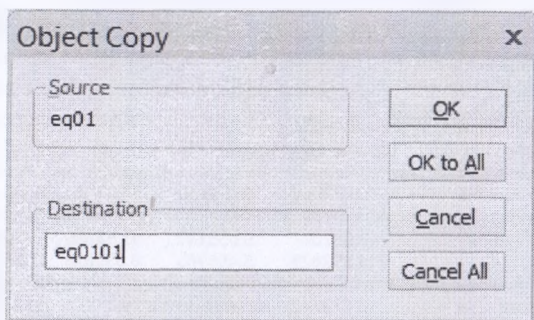


Рис. 2.2. Диалоговое окно при копировании объекта Eq01

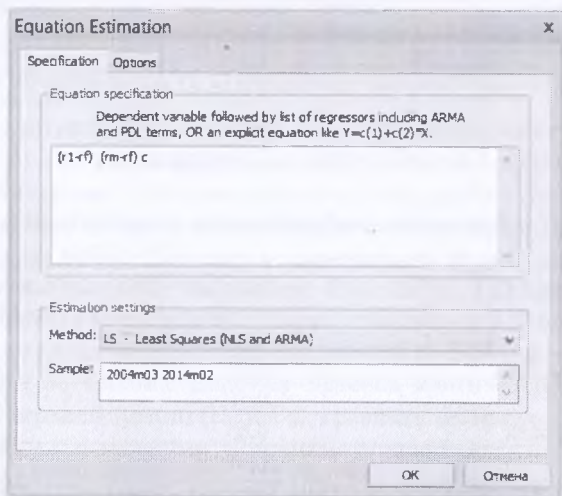


Рис. 2.3. Окно для ввода спецификации уравнения

Полученная оценка, как видно из рис. 2.4, практически не отличается от результата оценки уравнения без константы. Оценка свободного коэффициента статистически незначима ($t_{stat} = 0.33$, $Prob. = 0.74$), следовательно, по критерию Стьюдента гипотеза о том, что истинное значение свободного коэффициента равно 0, не отвергается.

Аналогичные результаты имеем для двух других компаний.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|--|-------------|-----------------------|-------------|----------|--------|----------|----------|-------|--------|
| Dependent Variable: R1-RF | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 24/03/14 Time: 19:14 | | | | | | | | | |
| Sample (adjusted): 2004M04 2014M02 | | | | | | | | | |
| Included observations: 119 after adjustments | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| RM-RF | 0.857087 | 0.055140 | 15.54385 | 0.0000 | | | | | |
| C | 1.832629 | 5.456963 | 0.335833 | 0.7376 | | | | | |
| R-squared | 0.673742 | Mean dependent var | | 11.39328 | | | | | |
| Adjusted R-squared | 0.670953 | S.D. dependent var | | 103.1144 | | | | | |
| S.E. of regression | 59.14909 | Akaike info criterion | | 11.01466 | | | | | |
| Sum squared resid | 409338.0 | Schwarz criterion | | 11.06137 | | | | | |
| Log likelihood | -653.3725 | Hannan-Quinn criter. | | 11.03363 | | | | | |
| F-statistic | 241.6114 | Durbin-Watson stat | | 2.015468 | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | |

Рис. 2.4. Результат оценки уравнения (2.2) для компании ЛУКОЙЛ

5. Значения коэффициентов детерминации для трех рассмотренных компаний приведены в таблице 2.3.

Таблица 2.3

Значения коэффициентов детерминации

| Компания | Коэффициент детерминации (в уравнении со свободным членом) |
|----------|---|
| ЛУКОЙЛ | 0.67 |
| МТС | 0.50 |
| Сбербанк | 0.63 |

Статистический смысл коэффициента детерминации заключается в том, что он показывает, какая доля общей дисперсии зависимой переменной объясняется уравнением регрессии. Так как дисперсия доходности является оценкой риска, то полный риск можно представить как сумму систематического (объясненного уравнением регрессии) и несистематического (собственного, необъясненного уравнением) риска. Таким образом, коэффициент детерминации, равный для компании ЛУКОЙЛ, 0.67 показывает, что доля систематического риска составляет для этой компании 67%.

Важно отметить, что бо́льшие значения коэффициента детерминации не обязательно соответствуют бо́льшим значениям оценок бета-коэффициентов. Это следует из связи между коэффициентом детерминации и оценкой коэффициента (подробнее см. Бердт, гл.2 [3])

$$R^2 = \beta_j^2 \frac{\sigma_m^2}{\sigma_j^2}$$

где σ_m^2 - дисперсия доходности рынка в целом

6. Для проверки гипотезы об отсутствии автокорреляции (АК) в остатках проведем тест Дарбина - Уотсона. Простейший вид АК

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t$$

$-1 < \rho < 1$ - коэффициент автокорреляции 1-го порядка

v_t - свободная от АК случайная составляющая

Гипотеза:

$H_0: \rho = 0$ - нет АК

$H_1: \rho > 0$ - есть положительная АК

Тест Дарбина - Уотсона входит в стандартный набор тестов, статистика для которых автоматически вычисляется при оценке уравнения по МНК. Поэтому в окне оценки уравнения (см. рис. 2.4) находим тестовую статистику $DW=2.015$. Так как тестовое значение практически равно 2, H_0 не отвергается, АК отсутствует. В случае, если тестовая статистика отлична от значения 2, необходимо найти по таблице Дарбина-Уотсона критические значения нижней и верхней границ (D_L и D_U), сравнить тестовую статистику с табличными значениями и принимать решение о наличии или отсутствии АК⁷.

7. Для проверки гипотезы о существовании так называемого «эффекта января», оцените уравнение для каждой ценной бумаги с фиктивной переменной, равной 1 для этого месяца. С этой целью для ЛУКОЙЛ скопируйте уравнение *Eq0101* и добавьте фиктивную переменную *JAN* (уравнение сохраните

⁷ Подробнее алгоритм проведения теста Дарбина-Уотсона см. в приложении

как Eq0102). Прделайте аналогичную процедуру для других ценных бумаг. Результат оценки представлен в таблице 2.4.

Таблица 2.4

Результат оценки CAPM (со свободным членом и фиктивной переменной)

| Компания | Оценка бета-коэффициента (P-значение) | Оценка коэф-фициента при фиктивной переменной (P-значение) | Коэффициент детерминации |
|----------|---------------------------------------|--|--------------------------|
| ЛУКОЙЛ | 0.86 (0.00) | 21.34 (0.28) | 0.68 |
| МТС | 0.72 (0.00) | -12.2 (0.61) | 0.50 |
| Сбербанк | 1.29 (0.00) | -45.66 (0.16) | 0.63 |

Результаты оценок показывают, что ни для одной из компаний мы не можем отклонить гипотезу об отсутствии эффекта января. Так как «эффект января» наблюдался для малых компаний на американском рынке, то отсутствие этого эффекта для крупных российских компаний не является удивительным.

8. Для расчета прогнозных значений выполните следующую последовательность действий.

- а. Расширьте диапазон данных рабочего файла **Proc→Structure/Resize Current Page...** добавьте 3 наблюдения, как показано на рис.2.5.

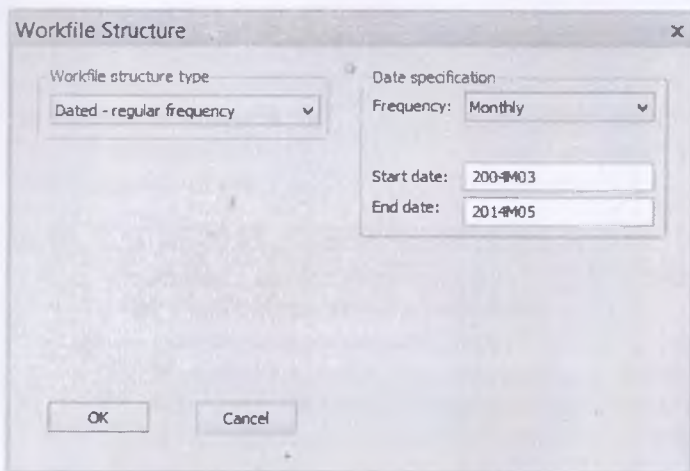


Рис. 2.5. Окно реструктурирования данных

- b. Откройте переменные *MICEX* и *RF* как группу и заполните данными по заданию (активировав команду *EDIT +/-*), как показано на рис. 2.6.

| obs | MICEX | RF |
|---------|---------|------|
| 2013M07 | 1377.60 | 0.04 |
| 2013M08 | 1364.54 | 0.04 |
| 2013M09 | 1463.13 | 0.02 |
| 2013M10 | 1509.62 | 0.05 |
| 2013M11 | 1479.35 | 0.07 |
| 2013M12 | 1503.39 | 0.07 |
| 2014M01 | 1454.05 | 0.04 |
| 2014M02 | 1494.32 | 0.04 |
| 2014M03 | 1489.32 | 0.04 |
| 2014M04 | 1484.32 | 0.04 |
| 2014M05 | 1479.32 | 0.04 |

Рис. 2.6. Окно заполнения данных для прогноза

- c. Откройте переменную *RM*, в окне переменной выберите **Proc** → **Generate by Equation**, в открывшемся диалоговом окне нажмите **OK** для перерасчета формулы (рис.2.7).

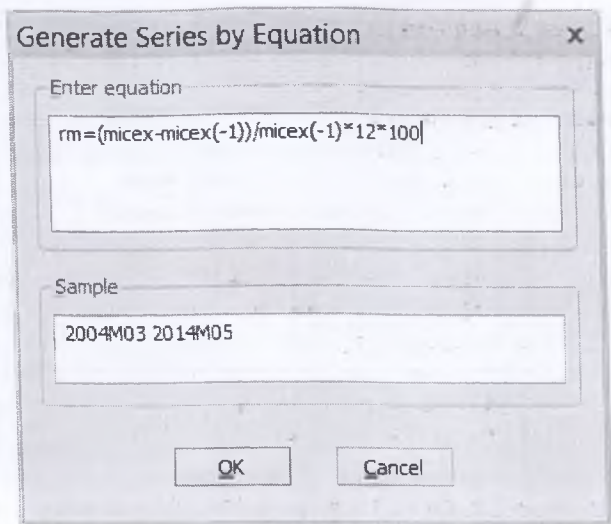


Рис. 2.7. Окно перерасчета формулы для доходности рынка

- d. Откройте уравнение *Eq0101* и нажмите команду **Forecast**. Все параметры диалогового окна оставьте по умолчанию (включая имя прогнозной переменной R1F). Аналогичную процедуру проделайте для *Eq0201* и *Eq0301*.

Сравните полученные результаты с данными, представленными в табл. 2.5.

Таблица 2.5

Результаты прогнозных значений доходностей акций по модели CAPM (со свободным членом)

| Дата/Компания | ЛУКОЙЛ | МТС | Сбербанк |
|---------------|-----------|-----------|----------|
| 2014M03 | -1.603035 | -2.809391 | 9.958957 |
| 2014M04 | -1.614588 | -2.819063 | 9.941644 |
| 2014M05 | -1.626220 | -2.828802 | 9.924215 |

Ключевые понятия

Модель CAPM

Избыточная доходность

Ожидаемая доходность

Систематический риск

Специфический риск

Полный риск

Рыночный риск

Диверсификация

«Эффект января»

Автокорреляция

Вопросы

1. Что такое систематический и несистематический риск? Как оцениваются эти величины?
2. Как выполнить прогноз доходности акций по модели CAPM, если ожидается рост рыночного индекса на 1 пп ежемесячно (безрисковая доходность не меняется)?
3. Какой инструмент используется для проверки «эффекта января»?
4. Какие возможности есть для минимизации специфического риска? Можно ли снизить систематический риск?
5. Какова экономическая интерпретация коэффициента детерминации в модели CAPM?
6. Как можно объяснить результаты прогноза для компании Сбербанк?

Глава 3. Моделирование почасовой заработной платы (альтернативные модели). Выбор спецификации

Цель работы: применение нелинейных моделей, использование фиктивных переменных взаимодействия и метода пошагового включения-исключения переменных, изучение методов тестирования спецификаций уравнения регрессии для вложенных и невложенных моделей

Исходные данные (eaf22.wf1):

EARNINGS – текущая почасовая заработная плата в долларах

MALE - фиктивная переменная, равная 1 для респондентов – мужчин

EXPER - опыт работы в годах

S - число завершенных лет обучения

EDUCBA – фиктивная переменная, равная 1 для респондента со степенью бакалавра

EDUCMAST – фиктивная переменная, равная 1 для респондента со степенью магистра

EDUCPHD – фиктивная переменная, равная 1 для респондента с докторской степенью

Данные американского Национального периодического обследования (NLS) за 1979г., выборка из 540 молодых работников, цитируется по Доутерги, 2009 [6].

Задания

1. Оцените полулогарифмическую модель (модель А)
$$\text{LOG}(EARNINGS) = \beta_1 + \beta_2 MALE + \beta_3 EXPER + \beta_4 S + \varepsilon^8 \quad (3.1)$$

Проинтерпретируйте коэффициенты уравнения.
2. Оцените модель, позволяющую ответить на вопрос, одинаково ли влияют опыт работы и образование на заработную плату мужчин и женщин, добавив к уравнению (3.1) члены, отвечаю-

⁸ Здесь и далее в пособии LOG означает натуральный логарифм

щие за взаимодействие объясняющих факторов с фиктивной переменной, т.е.

$MALE * EXPER$ и $MALE * S$ (уравнение 3.2).

$$\text{LOG}(EARNINGS) = \beta_1 + \beta_2 MALE + \beta_3 EXPER + \beta_4 S + \beta_4 MALE * EXPER + \beta_5 MALE * S + \varepsilon \quad (3.2)$$

Проинтерпретируйте коэффициенты при переменных взаимодействия, проверьте их на значимость.

- Используя метод пошагового включения-исключения переменных, оцените модель только со значимыми факторами. Выберите наилучшую модель по формальным критериям. Проведите тест Уальда на целесообразность введения ограничений, приводящих от модели с полным набором факторов (уравнение 3.2) к вложенной модели только со значимыми факторами.
- Найдите расчетное значение заработной платы по модели, полученной в п.3. для женщины и мужчины с опытом работы 10 лет и числом завершённых лет обучения равным 12.
- Учитывая результат п.3, проведите более подробный анализ влияния уровня образования на заработную плату. Ограничьтесь выявлением влияния достигнутой респондентом степени (бакалавр, магистр, PhD) на зарплату, учитывая опыт работы и пол респондента.

Оцените линейную модель

$$EARNINGS = \beta_1 + \beta_2 MALE + \beta_3 EXPER + \beta_4 EDUCBA + \beta_5 EDUCMAST + \beta_5 EDUCPHD + \varepsilon \quad (3.4)$$

и полулогарифмическую модель

$$\text{LOG}(EARNINGS) = \beta_1 + \beta_2 MALE + \beta_3 EXPER + \beta_4 EDUCBA + \beta_5 EDUCMAST + \beta_5 EDUCPHD + \varepsilon \quad (3.5)$$

Сравните результаты.

- Выполните PE-тестирование модели (3.4) против модели (3.5). Выполните PE -тестирование модели (3.5) против модели (3.4). Каково ваше заключение?

Решения

1. В главе 1 рассматривалась модель влияния человеческого капитала на заработную плату. На основании проведенного анализа было показано, что величина заработной платы в сильной степени зависит от уровня образования. Существенным оказался и гендерный эффект, различия в заработной плате мужчин и женщин статистически значимы. Эти оценки были получены на основании данных американского периодического обследования за 1987 г.

В данной главе продолжим исследование функции заработной платы, используя нелинейные модели, основываясь на более ранних данных этого же исследования.

Для оценки модели А в окне оценки уравнения **Quick → Estimate Equation...** наберите

LOG(EARNINGS) MALE EXPER S C

Результат оценки Eq01 (рис. 3.1) показывает, что при прочих равных условиях (одинаковый опыт работы и количество лет обучения) мужчины в среднем зарабатывают на 30% больше, чем женщины.

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| MALE | 0.304850 | 0.044831 | 6.800038 | 0.0000 |
| EXPER | 0.034822 | 0.004893 | 7.117383 | 0.0000 |
| S | 0.131725 | 0.009086 | 14.49779 | 0.0000 |
| C | 0.244945 | 0.164544 | 1.488627 | 0.1372 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.352318 | Mean dependent var | 2.782847 |
| Adjusted R-squared | 0.348693 | S.D. dependent var | 0.630120 |
| S.E. of regression | 0.508529 | Akaike info criterion | 1.492792 |
| Sum squared resid | 138.5107 | Schwarz criterion | 1.524581 |
| Log likelihood | -399.0538 | Hannan-Quinn criter. | 1.505225 |
| F-statistic | 97.18888 | Durbin-Watson stat | 1.967076 |
| Prob(F-statistic) | 0.000000 | | |

Рис. 3.1. Оценка модели А Eq01

Заработная плата респондентов, имеющих опыт работы с разницей в 1 год, отличается в среднем на 3.4% при прочих равных условиях (чем больше опыт работы, тем выше заработная плата).

Заработная плата респондентов, имеющих количество лет обучения с разницей в 1 год, отличается в среднем на 13.2% при прочих равных условиях (каждый дополнительный год обучения увеличивает заработную плату на 13.2%).

2. Для оценки влияния опыта работы и образования отдельно для мужчин и для женщин введем члены взаимодействия переменных *EXPER*, *S* и *MALE* (Eq02). Результат оценки представлен на рис. 3.2.

| File Edit Object View Proc Quick Options Window Help | | | | |
|---|-------------|-----------------------|-------------|--------|
| View Proc Object Print Name Freeze Estimate Forecast Stats Resids | | | | |
| Dependent Variable: LOG(EARNINGS) | | | | |
| Method: Least Squares | | | | |
| Date: 26/03/14 Time: 17:19 | | | | |
| Sample: 1 540 | | | | |
| Included observations: 540 | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| MALE | -0.712049 | 0.335732 | -2.120883 | 0.0344 |
| EXPER | 0.029277 | 0.006587 | 4.444385 | 0.0000 |
| S | 0.107368 | 0.012307 | 8.724481 | 0.0000 |
| MALE*EXPER | 0.015424 | 0.009817 | 1.571129 | 0.1167 |
| MALE*S | 0.055155 | 0.018217 | 3.027731 | 0.0026 |
| C | 0.667008 | 0.212581 | 3.137664 | 0.0018 |
| R-squared | 0.364053 | Mean dependent var | 2.782847 | |
| Adjusted R-squared | 0.358098 | S.D. dependent var | 0.630120 | |
| S.E. of regression | 0.504844 | Akaike info criterion | 1.481915 | |
| Sum squared resid | 136.0994 | Schwarz criterion | 1.529599 | |
| Log likelihood | -394.1171 | Hannan-Quinn criter. | 1.500564 | |
| F-statistic | 61.13841 | Durbin-Watson stat | 1.975380 | |
| Prob(F-statistic) | 0.000000 | | | |

Рис. 3.2. Оценка модели со взаимодействием Eq02

В такой спецификации влияние гендерного эффекта проявляется через три слагаемых: через переменную *MALE* (влияние пола на заработную плату в среднем независимо от других факторов), через переменную *MALE*EXPER* (различие во влиянии опыта работы для мужчин и женщин) и через переменную *MALE*S* (различие во влиянии образования для мужчин и женщин). Действительно, согласно полученным оценкам зависимость заработной платы от опыта работы и образования для мужчин имеет вид

$$\text{LOG}(\text{EARNINGS})^{\wedge} = (0.667 - 0.71) + (0.029 + 0.015) * \text{EXPER} + (0.107 + 0.055) * S$$

в то время как для женщин оцененное уравнение имеет вид

$$\text{LOG}(\text{EARNINGS})^{\wedge} = 0.667 + 0.029 * \text{EXPER} + 0.107 * S$$

Таким образом, коэффициенты при перекрестных членах показывают, на сколько процентов различается влияние того или иного фактора на моделируемую переменную в зависимости от гендерного признака.

Коэффициент при перекрестном члене $\text{MALE} * \text{EXPER}$ статистически незначим даже на 10% уровне значимости ($\text{Prob} = 0.12 > 0.1$), следовательно, данная оценка показывает, что влияние опыта работа на заработную плату незначимо различается для мужчин и для женщин. Однако коэффициент при $\text{MALE} * S$, равный 0.055, значим, и его величина говорит о том, что при одинаковом опыте работы каждый дополнительный год образования у мужчин увеличивает заработную плату на $(10.7 + 5.5 = 16.3\%)$, а у женщин только на 10.7%.

3. Используя метод пошагового исключения незначимых переменных, исключим на первом шаге перекрестный член $\text{MALE} * \text{EXPER}$ (рис.3.3).

| File Edit Object View Proc Quick Options Window Help | | | | | | | | | | |
|--|-------------|-----------------------|-------------|--------|------|--------|----------|----------|-------|--------|
| View | | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
| Dependent Variable: LOG(EARNINGS) | | | | | | | | | | |
| Method: Least Squares | | | | | | | | | | |
| Date: 26/03/14 Time: 17:19 | | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | | |
| Included observations: 540 | | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | | |
| MALE | -0.353812 | 0.246766 | -1.433800 | 0.1522 | | | | | | |
| EXPER | 0.036221 | 0.004891 | 7.405653 | 0.0000 | | | | | | |
| S | 0.109194 | 0.012268 | 8.900426 | 0.0000 | | | | | | |
| MALE*S | 0.047877 | 0.017642 | 2.713807 | 0.0069 | | | | | | |
| C | 0.533132 | 0.195023 | 2.733685 | 0.0065 | | | | | | |
| R-squared | 0.361113 | Mean dependent var | 2.782847 | | | | | | | |
| Adjusted R-squared | 0.356336 | S.D. dependent var | 0.630120 | | | | | | | |
| S.E. of regression | 0.505537 | Akaike info criterion | 1.482824 | | | | | | | |
| Sum squared resid | 136.7285 | Schwarz criterion | 1.522560 | | | | | | | |
| Log likelihood | -395.3624 | Hannan-Quinn criter. | 1.498364 | | | | | | | |
| F-statistic | 75.59840 | Durbin-Watson stat | 1.960610 | | | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | | |

Рис. 3.3. Оценка промежуточного уравнения

Полученная оценка содержит теперь незначимый коэффициент при факторе *MALE*. Исключение этого фактора дает следующий результат (рис. 3.4)

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|-----------------------------------|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LOG(EARNINGS) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 26/03/14 Time: 17:20 | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | |
| Included observations: 540 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| EXPER | 0.035223 | 0.004846 | 7.268508 | 0.0000 | | | | | |
| S | 0.120835 | 0.009206 | 13.12546 | 0.0000 | | | | | |
| MALE*S | 0.022998 | 0.003189 | 7.210835 | 0.0000 | | | | | |
| C | 0.383036 | 0.164708 | 2.325550 | 0.0204 | | | | | |
| R-squared | 0.358658 | Mean dependent var | 2.782847 | | | | | | |
| Adjusted R-squared | 0.355068 | S.D. dependent var | 0.630120 | | | | | | |
| S.E. of regression | 0.506034 | Akaike info criterion | 1.482955 | | | | | | |
| Sum squared resid | 137.2539 | Schwarz criterion | 1.514744 | | | | | | |
| Log likelihood | -396.3979 | Hannan-Quinn criter. | 1.495388 | | | | | | |
| F-statistic | 99.91580 | Durbin-Watson stat | 1.963459 | | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | |

Рис. 3.4. Оценка модели А только со значимыми факторами *Eq03*

По формальным критериям исключение факторов *MALE* и *MALE*EXPER* не ухудшило качества оценки, так как значения скорректированного коэффициента детерминации увеличилось, а значения информационных критериев Акайке и Шварца уменьшились.

Для проведения теста Уальда на целесообразность введения ограничений в меню уравнения с полным набором факторов выберите **View**→**Coefficient Diagnostics**→**Wald Test-Coefficient Restrictions**, в открывшемся окне запишите ограничения $C(1)=C(4)=0$. Здесь $C(1)$ - коэффициент при факторе *MALE*, $C(4)$ - коэффициент при факторе *MALE*EXPER* в уравнении *Eq02*.

Как показывает результат теста, различия в качестве оценки статистически незначимы ($Fstat=2.26$, $Prob=0.10$), поэтому для дальнейшего анализа выбираем более экономную модель *Eq03*.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|--|-----------|-----------|-------------|------|--------|----------|----------|-------|--------|
| Wald Test Equation: EQ02 | | | | | | | | | |
| Test Statistic | Value | df | Probability | | | | | | |
| F-statistic | 2.264934 | (2, 534) | 0.1048 | | | | | | |
| Chi-square | 4.529869 | 2 | 0.1038 | | | | | | |
| Null Hypothesis: C(1)=C(4)=0 Null Hypothesis Summary: | | | | | | | | | |
| Normalized Restriction (= 0) | Value | Std. Err. | | | | | | | |
| C(1) | -0.712049 | 0.335732 | | | | | | | |
| C(4) | 0.015424 | 0.009817 | | | | | | | |

Рис. 3.5. Результат теста Уальда на ограничения по уравнению $Eq02$

4. Расчетное значение заработной платы для женщины согласно $Eq03$ определяется из соотношения

$$\begin{aligned} \text{LOG}(\text{EARNINGS})^{\wedge} &= 0.383 + 0.035\text{EXPER} + 0.12S = \\ &= 0.383 + 0.035 \cdot 10 + 0.12 \cdot 12 = 2.17 \end{aligned}$$

$$\text{EARNINGS}^{\wedge} = \exp(2.17) \approx 8.76$$

Для мужчины имеем следующее значение;

$$\begin{aligned} \text{LOG}(\text{EARNINGS})^{\wedge} &= 0.383 + 0.035\text{EXPER} + (0.12 + 0.023)S = \\ &= 0.383 + 0.035 \cdot 10 + 0.143 \cdot 12 = 2.45 \end{aligned}$$

$$\text{EARNINGS}^{\wedge} = \exp(2.45) = 11.58$$

Таким образом, согласно полученным оценка почасовая заработная плата мужчин в среднем на 32% выше почасовой заработной платы женщин при одинаковом опыте работы (10 лет) и уровне образования (12 лет обучения), что согласуется с оценками модели без взаимодействия (3.1). Однако различия в заработной плате между мужчинами с разным уровнем образования больше, чем между женщинами с такими же характеристиками.

5. Составим модель, оценивающую влияние уровня образования респондента на заработную плату. Оценим уравнение

Eq04 в линейной форме и *Eq05* в полулогарифмической форме (рис. 3.6 и 3.7).

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|------------------------------|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: EARNINGS | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 17/06/14 Time: 22:20 | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | |
| Included observations: 540 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| C | 6.086387 | 2.266815 | 2.684995 | 0.0075 | | | | | |
| MALE | 6.843455 | 1.179279 | 5.803082 | 0.0000 | | | | | |
| EXPER | 0.426167 | 0.129183 | 3.298941 | 0.0010 | | | | | |
| EDUCBA | 12.37510 | 1.509624 | 8.197471 | 0.0000 | | | | | |
| EDUCMAST | 15.88195 | 2.563920 | 6.194402 | 0.0000 | | | | | |
| EDUCPHD | 14.04302 | 7.794170 | 1.801734 | 0.0722 | | | | | |
| R-squared | 0.207142 | Mean dependent var | 19.84939 | | | | | | |
| Adjusted R-squared | 0.199718 | S.D. dependent var | 14.95314 | | | | | | |
| S.E. of regression | 13.37685 | Akaike info criterion | 8.035977 | | | | | | |
| Sum squared resid | 95554.05 | Schwarz criterion | 8.083662 | | | | | | |
| Log likelihood | -2163.714 | Hannan-Quinn crite | 8.054627 | | | | | | |
| F-statistic | 27.90249 | Durbin-Watson stat | 1.935619 | | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | |

Рис. 3.6. Оценка линейной модели для выявления влияния уровня образования на заработную плату *Eq04*

Как показывает результат оценки модели в линейной форме, при прочих равных условиях респонденты, имеющие степень бакалавра, зарабатывают больше, чем респонденты без степени, в среднем на 12.4 доллара в час; респонденты, имеющие степень магистра, зарабатывают больше, чем респонденты без степени, в среднем на 15.9 доллара в час; респонденты, имеющие степень доктора (PhD), зарабатывают больше, чем респонденты без степени, в среднем на 14 долларов в час. Заметим, что последний коэффициент статистически значим только на 10% уровне.

Согласно полулогарифмической модели, респонденты, имеющие степень бакалавра, зарабатывают больше, чем респонденты без степени, в среднем на 56%. Эта оценка согласуется с оценкой, полученной по линейной модели. Как видно из рис. 3.6, среднее значение почасовой заработной платы (**Mean dependent var**) для

данной выборки составляет 19.85, т.е. увеличение на 12 долларов соответствует увеличению на 60%.

Заметим, что коэффициенты в нелинейной модели все статистически значимы. Однако сравнивать линейную и нелинейную модели по такому критерию, как коэффициент детерминации или информационные критерии, нельзя, так как при расчете этих критериев используются остатки (разность между фактическим и расчетным значением), которые имеют разный масштаб для самой переменной и ее логарифма (Вербик, гл. 3 [4]).

Поэтому для выявления, какая из моделей более адекватна по статистическим критериям, надо провести дополнительное исследование.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|-----------------------------------|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LOG(EARNINGS) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 17/06/14 Time: 22:22 | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | |
| Included observations: 540 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| C | 1.986542 | 0.091798 | 21.64048 | 0.0000 | | | | | |
| MALE | 0.310826 | 0.047756 | 6.508573 | 0.0000 | | | | | |
| EXPER | 0.029512 | 0.005231 | 5.641371 | 0.0000 | | | | | |
| EDUCBA | 0.561126 | 0.061134 | 9.178604 | 0.0000 | | | | | |
| EDUCMAST | 0.736568 | 0.103829 | 7.094037 | 0.0000 | | | | | |
| EDUCPHD | 0.846051 | 0.315635 | 2.680476 | 0.0076 | | | | | |
| R-squared | 0.267776 | Mean dependent var | 2.782847 | | | | | | |
| Adjusted R-squared | 0.260920 | S.D. dependent var | 0.630120 | | | | | | |
| S.E. of regression | 0.541712 | Akaike info criterion | 1.622886 | | | | | | |
| Sum squared resid | 156.7036 | Schwarz criterion | 1.670570 | | | | | | |
| Log likelihood | -432.1793 | Hannan-Quinn criter. | 1.641535 | | | | | | |
| F-statistic | 39.05698 | Durbin-Watson stat | 1.927794 | | | | | | |
| Prob(F-statistic) | 0.000000 | | | | | | | | |

Рис. 3.7. Оценка полулогарифмической модели для выявления влияния уровня образования на заработную плату Eq05

6. Нулевая гипотеза: линейная модель достаточна для объяснения зависимости, расчетные значения, полученные из нелинейной модели, не должны помогать в объяснении зависи-

мой переменной. Альтернативная гипотеза: линейная модель недостаточна.

Для тестирования нулевой гипотезы, т.е. модели (3.4) против модели (3.5) надо выполнить следующие действия.

- Получить расчетные значения из модели (3.4) EQ04 → Forecast → EARNINGSF_4 (EARNINGSF_4 - имя расчетной переменной)
- Получить логарифм расчетных значения из модели (3.5) EQ05 → Forecast → LEARNINGSF_5 (LEARNINGSF_5 - имя расчетной переменной в логарифмах)
- Построить вспомогательную регрессию, добавив в уравнение Eq04 разность $(\text{LOG}(\text{EARNINGSF}_4) - \text{LEARNINGSF}_5)$ ⁹. Т.е. в окне оценки уравнения записать

EARNINGS C MALE EXPER EDUCBA EDUCMAST EDUCPHD
(LOG(EARNINGSF_4)-LEARNINGSF_5)

- Провести тест Стьюдента для коэффициента при $(\text{LOG}(\text{EARNINGSF}_4) - \text{LEARNINGSF}_5)$.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|-------------------------------|-------------|------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: EARNINGS | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 17/06/14 Time: 22:46 | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | |
| Included observations: 540 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| C | 11.14505 | 2.519754 | 4.423069 | 0.0000 | | | | | |
| MALE | 9.481677 | 1.311572 | 7.229247 | 0.0000 | | | | | |
| EXPER | 0.395308 | 0.127305 | 3.105206 | 0.0020 | | | | | |
| EDUCBA | 13.78087 | 1.520645 | 9.062514 | 0.0000 | | | | | |
| EDUCMAST | 17.55432 | 2.552261 | 6.877950 | 0.0000 | | | | | |
| EDUCPHD | 9.888942 | 7.728930 | 1.279471 | 0.2013 | | | | | |
| LOG(EARNINGSF_4)-LEARNINGSF_5 | -42.93989 | 9.952761 | -4.314370 | 0.0000 | | | | | |

Рис. 3.8. Оценка вспомогательной регрессии для проведения PE теста

Как показывает результат оценки (рис.3.8), коэффициент при последнем слагаемом статистически значим, следовательно, нуле-

⁹ Расчетные значения линейной модели логарифмируются, чтобы иметь один масштаб с расчетными значениями нелинейной модели

вая гипотеза отвергается, и линейная модель недостаточная для объяснения влияния выбранных факторов на почасовую заработную плату.

Аналогично, для тестирования модели (3.5) против модели (3.4) надо:

- a. Получить расчетные значения из модели 4 EQ04 → Forecast → EARNINGSF_4 (EARNINGSF_4 - имя расчетной переменной, было получено ранее)
- b. Получить расчетные значения из модели 5 EQ05 → Forecast → EARNINGSF_5 (EARNINGSF_5 - имя расчетной переменной)
- c. Построить вспомогательную регрессию, добавив в уравнение Eq05 разность (EARNINGSF_4 - EARNINGSF_5). Т.е. в окне оценки уравнения записать LOG(EARNINGS) C MALE EXPER EDUCBA EDUCMAST EDUCPHD (EARNINGSF_4-EARNINGSF_5)

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|-----------------------------------|-------------|------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LOG(EARNINGS) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 17/06/14 Time: 23:00 | | | | | | | | | |
| Sample: 1 540 | | | | | | | | | |
| Included observations: 540 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| C | 2.065552 | 0.097405 | 21.20574 | 0.0000 | | | | | |
| MALE | 0.398613 | 0.060486 | 6.590122 | 0.0000 | | | | | |
| EXPER | 0.029820 | 0.005211 | 5.722386 | 0.0000 | | | | | |
| EDUCBA | 0.670521 | 0.076678 | 8.745973 | 0.0000 | | | | | |
| EDUCMAST | 0.871774 | 0.118339 | 7.366766 | 0.0000 | | | | | |
| EDUCPHD | 0.773447 | 0.315825 | 2.448974 | 0.0146 | | | | | |
| EARNINGSF_4-EARNINGSF_5 | -0.055648 | 0.023693 | -2.348668 | 0.0192 | | | | | |

Рис. 3.9. Оценка вспомогательной регрессии для проведения PE – теста

Как показывает результат оценки (рис.3.9), коэффициент при последнем слагаемом (EARNINGSF_4-EARNINGSF_5) статистически незначим на уровне 1%, следовательно, нулевая гипотеза не отвергается, и полугарифмическая модель достаточна для объяснения влияния выбранных факторов на почасовую заработную плату.

плату. Полученный вывод в определенной степени зависит от выбранного уровня значимости, однако нелинейная модель выглядит предпочтительнее линейной, если учитывать также значения коэффициентов при факторах, отражающих уровень образования.

Ключевые понятия

| | |
|--------------------------------|-------------------------------------|
| Переменные взаимодействия | Пошаговый метод исключения факторов |
| Функциональная форма | Тестирование функциональной формы |
| Вложенные и невложенные модели | Информационные критерии |
| Помогательная регрессия | |
| PE-тест | |

Вопросы

1. Приведите виды различных функциональных форм и примеры их использования в экономических моделях.
2. Какие есть способы тестирования функциональной формы, кроме приведенных в данной главе?
3. Дайте определение вложенных и невложенных моделей.
4. Проведите тестирование не вложенных моделей E_{q03} и E_{q05} , используя J-тест. Сделайте вывод.
5. Объясните основную идею теста Рамсея, как можно использовать тест Рамсея для данной задачи.
6. Объясните, почему тест установки Рамсея может указывать на проблему невключенных переменных.
7. Какие есть методы отбора факторов в моделях множественной регрессии?
8. Объясните полезность информационных критериев при сравнении двух вложенных моделей.
9. Объясните, почему не следует исключать из модели две переменные одновременно на основании только значений t-статистик.

Глава 4. Моделирование зависимости спроса от цены методом инструментальных переменных

Цель работы: изучение метода инструментальных переменных (ИП), применение двухшагового метода наименьших квадратов для оценки уравнения, изучение методов тестирования спецификации уравнения регрессии с использованием теста Дарбина-Ву-Хаусмана.

Исходные данные (*PACKPC.wf1*):

Данные по 48 штатам США за 1995 г. в ценах базового года
PACKPC - количество проданных упаковок сигарет на душу населения в год в штате

AVGPRS - средняя цена

PERINC - доход на душу населения

RTAXSO - средний налог с продаж

RTAXS - налог на табачные изделия

Источник данных: Stock and Watson 2010 (гл. 10), цитируются по MANUEL Eviews 7.0.

Задания

1. Оцените эластичность спроса на сигареты от цены методом наименьших квадратов

$$\text{LOG}(\text{PACKPC}) = \beta_1 + \beta_2 \text{LOG}(\text{AVGPRS}) + \varepsilon \quad (4.1)$$

Какие предпосылки МНК могут быть нарушены в такой модели и каковы последствия этих нарушений?

2. Предполагая, что цена является эндогенной переменной, проверьте возможность использования переменных *RTAXSO* и *RTAXS* в качестве инструментов.
3. Используя метод инструментальных переменных, получите состоятельную оценку эластичности спроса по цене. В качестве инструмента используйте *RTAXSO*. Оценку проведите двухшаговым методом наименьших квадратов.

4. Оцените эластичность спроса на сигареты от цены и дохода методом наименьших квадратов

$$\text{LOG}(\text{PACKPC}) = \beta_1 + \beta_2 \text{LOG}(\text{RAVGPRS}) + \beta_3 \text{LOG}(\text{PERINC}) + \varepsilon \quad (4.2)$$

5. Оцените нелинейную модель (4.2) двухшаговым методом наименьших квадратов (ДМНК, TSLS - two stage least square), используя в качестве инструментальных переменных RTAXSO и RTAXS.
6. Повторно оцените модель (4.2) методом ДМНК (TSLS), используя встроенный алгоритм программы Eviews. Сравните полученные результаты.
7. Проведите тест Дарбина-Ву-Хаусмана для сравнения оценок коэффициентов по методу МНК и ИП. Сделайте выводы.

Решения

1. Оценку модели (4.1) проведем обычным методом наименьших квадратов, используя командную строку, запишем

LS LOG(PACKPC) LOG(RAVGPRS) C

Результат оценки приведен на рис. 4.1

EViews - [Equation: EQ01_LS - Workfile: PACKPC:CI... - □ X

File Edit Object View Proc Quick Options Window Help

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: LOG(PACKPC)
 Method: Least Squares
 Date: 19/06/14 Time: 23:32
 Sample: 1 48
 Included observations: 48

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------|-------------|------------|-------------|--------|
| LOG(RAVGPRS) | -1.213057 | 0.216450 | -5.604337 | 0.0000 |
| C | 10.33892 | 1.035290 | 9.986500 | 0.0000 |

| | | | |
|--------------------|----------|-----------------------|-----------|
| R-squared | 0.405751 | Mean dependent var | 4.538837 |
| Adjusted R-squared | 0.392832 | S.D. dependent var | 0.243346 |
| S.E. of regression | 0.189618 | Akaike info criterion | -0.446838 |
| Sum squared resid | 1.653928 | Schwarz criterion | -0.368871 |
| Log likelihood | 12.72410 | Hannan-Quinn criter. | -0.417374 |
| F-statistic | 31.40860 | Durbin-Watson stat | 1.984850 |
| Prob(F-statistic) | 0.000001 | | |

DB = none WF = packpc

Рис. 4.1. Оценка эластичности спроса по цене методом наименьших квадратов в однофакторной модели *Eq01_ls*

Полученная оценка эластичности может быть смещенной, если нарушены условия применения МНК, а именно:

- пропущена существенная переменная (в данном случае, пропущена такая детерминанта спроса, как доход), смещение тем больше, чем сильнее пропущенная переменная связана со спросом;
- фактор цена измерен с ошибками;
- фактор цена является эндогенной переменной для модели спроса и коррелирует со случайной составляющей.

В последнем случае оценка по МНК является не только смещенной, но и несостоятельной.

Кроме того, поскольку мы имеем дело с данными перекрестной выборки, то стандартные ошибки оценок коэффициентов могут также иметь смещение из-за гетероскедастичности.

2. Предположим, что цена пачки сигарет является эндогенной переменной в уравнении для спроса. Действительно, цена формируется под воздействием спроса и предложения и, строго говоря, не является экзогенной по отношению к спросу. Следствием эндогенности цены является корреляция этого фактора со случайным членом уравнения (4.1), что в свою очередь является нарушением предпосылок МНК. В случае нарушения этой предпосылки МНК-оценки являются смещенными и несостоятельными.

Для того чтобы преодолеть эту проблему, используется метод инструментальных переменных. В качестве инструментов используются переменные, которые должны удовлетворять следующим требованиям:

- инструменты должны коррелировать с эндогенной переменной, чем выше эта корреляция, тем лучше;
- инструменты не должны коррелировать со случайным членом исходного уравнения;
- инструментальные переменные не должны быть объясняющими для исходной зависимой переменной.

В качестве инструментальных переменных для моделирования спроса на сигареты в учебнике Stock and Watson "Introduction to econometrics" предложено использовать налог с продаж (средний по рассматриваемым штатам) и специфический налог на табачные изделия.

Для проверки первого требования на инструментальные переменные построим корреляционную матрицу для *RAVGPRS*, *RTAXSO* и *RTAXS* (рис. 4.2)

| | A | B | C | D |
|----|-------------------------------|----------|----------|----------|
| 1 | Covariance Analysis: Ordinary | | | |
| 2 | Date: 22/06/14 Time: 19:48 | | | |
| 3 | Sample: 1 48 | | | |
| 4 | Included observations: 48 | | | |
| 5 | ----- | | | |
| 6 | Correlation | | | |
| 7 | t-Statistic | | | |
| 8 | Probability | RAVGPRS | RTAXS | RTAXSO |
| 9 | RAVGPRS | 1.000000 | | |
| 10 | | --- | | |
| 11 | | --- | | |
| 12 | | | | |
| 13 | RTAXS | 0.964311 | 1.000000 | |
| 14 | | 24.70154 | --- | |
| 15 | | 0.0000 | --- | |
| 16 | | | | |
| 17 | RTAXSO | 0.685614 | 0.711384 | 1.000000 |
| 18 | | 6.387747 | 6.865141 | --- |
| 19 | | 0.0000 | 0.0000 | --- |

Рис. 4.2. Корреляционная матрица переменных цена (*RAVGPRS*) и налоги (*RTAXS*, *RTAXSO*)

Таким образом, оба фактора *RTAXS* и *RTAXSO* удовлетворяют требованию высокой корреляции с эндогенной переменной.

Для того чтобы проверить отсутствие корреляции инструментов со случайной составляющей, можно использовать остатки регрессии по уравнению *Eq01_ls*. Откройте уравнение и выполните команду **Proc**→**Make Residual Series...** Дайте ряду остатков имя *RESID01_LS*. Рассчитав коэффициенты корреляции остатков с переменными, которые предполагается использовать в качестве инструментов, получаем, что коэффициент корреляции между *RESID01_LS* и *RTAXS* равен 0.05, а между *RESID01_LS* и *RTAXSO* равен 0.06. Оба коэффициента статистически незначимы. Таким образом, второе требование на инструментальные переменные также выполняется.

Третье требование также можно считать выполненным, так как непосредственного влияния налоги на спрос не оказывают, только опосредованно через фактор цена.

Д. Проведем оценку уравнения (4.1) двухшаговым методом наименьших квадратов, используя в качестве инструмента средний налог с продаж.

Алгоритм применения двухшагового метода состоит в следующем. На первом шаге оценивается регрессия эндогенной переменной на инструмент. Получаются расчетные значения эндогенной переменной. На втором шаге оценивается исходное уравнение, но вместо эндогенной переменной используются ее расчетные значения, полученные на первом шаге.

Таким образом, для данной задачи выполним следующую последовательность действий.

Первый шаг.

- а. МНК оценка уравнения, в котором цена является зависимой переменной, а налог с продаж - фактором. Для этого в командной строке запишите (уравнение сохраните как *Eq01_1*)

LS LOG(RAVGPRS) RTAXSO C

- б. Получение расчетных значений *RAVGPRSF_1* (через команду **Forecast** в окне уравнения *Eq01_1*)

Второй шаг.

МНК оценка уравнения (4.1) с заменой *LOG(RAVGPRS)* на *LOG(RAVGPRSF_1)*

LS LOG(PACKPC) LOG(RAVGPRSF_1) C

(сохраните уравнение как *Eq01_2*).

Результат второго шага приведен на рис. 4.3.

Если сравнить результаты МНК (LS) и ДМНК (TSL) оценки (рис. 4.1 и 4.3), то видим, что по МНК эластичность спроса по цене оценивается выше (по модулю). Если инструментальные переменные подобраны верно, то ДМНК оценка устраняет смещение, вызванное эндогенностью цены, и является состоятельной.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|---------------------------------|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LOG(PACKPC) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 22/06/14 Time: 20:46 | | | | | | | | | |
| Sample: 1 48 | | | | | | | | | |
| Included observations: 48 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| LOG(RAVGPRSF_1) | -1.083587 | 0.376649 | -2.876917 | 0.0061 | | | | | |
| C | 9.719877 | 1.801197 | 5.396344 | 0.0000 | | | | | |
| R-squared | 0.152490 | Mean dependent var | 4.538837 | | | | | | |
| Adjusted R-squared | 0.134066 | S.D. dependent var | 0.243346 | | | | | | |
| S.E. of regression | 0.226447 | Akaike info criterion | -0.091834 | | | | | | |
| Sum squared resid | 2.358809 | Schwarz criterion | -0.013867 | | | | | | |
| Log likelihood | 4.204012 | Hannan-Quinn criter. | -0.062370 | | | | | | |
| F-statistic | 8.276649 | Durbin-Watson stat | 2.099150 | | | | | | |
| Prob(F-statistic) | 0.006069 | | | | | | | | |

Рис. 4.3. Результат применения двухшагового метода наименьших квадратов (ДМНК) к оценке уравнения (4.1)

Заметим, однако, что рассчитанные по ДМНК стандартные ошибки коэффициентов регрессии не верны, так как при их расчете применяются обычные формулы и игнорируется то, что на втором шаге оценивается регрессия, в которой используются расчетные значения фактора.

Для устранения этого дефекта надо использовать специальные формулы для расчета ошибок коэффициентов регрессии. Проиллюстрируем, как это делается в программе EViews при выполнении следующих заданий.

4. Включение фактора доход и оценка двухфакторной модели по МНК устраняет возможное смещение в оценке эластичности спроса по цене, вызванное пропуском важной переменной.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|---------------------------------|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LCG(PACKPC) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 18/06/14 Time: 20:58 | | | | | | | | | |
| Sample: 1 48 | | | | | | | | | |
| Included observations: 48 | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| LOG(RAVGPRS) | -1.406500 | 0.251375 | -5.595218 | 0.0000 | | | | | |
| LOG(PERINC) | 0.343850 | 0.234967 | 1.463397 | 0.1503 | | | | | |
| C | 10.34203 | 1.022681 | 10.11267 | 0.0000 | | | | | |
| R-squared | 0.432746 | Mean dependent var | 4.538837 | | | | | | |
| Adjusted R-squared | 0.407535 | S.D. dependent var | 0.243346 | | | | | | |
| S.E. of regression | 0.187308 | Akaike info criterion | -0.451663 | | | | | | |
| Sum squared resid | 1.578794 | Schwarz criterion | -0.334713 | | | | | | |
| Log likelihood | 13.83991 | Hannan-Quinn criter. | -0.407467 | | | | | | |
| F-statistic | 17.16478 | Durbin-Watson stat | 1.933142 | | | | | | |
| Prob(F-statistic) | 0.000003 | | | | | | | | |

Рис. 4.4. Оценка двухфакторной модели спроса на сигареты методом наименьших квадратов $Eq02_Is$

Коэффициенты при факторах имеют ожидаемые знаки (коэффициент при цене - отрицательный, коэффициент при доходе - положительный), кроме того, так как сигареты являются специфическим товаром, спрос на который не эластичен по доходу, незначительность коэффициента при факторе доход не противоречит ожиданиям. Однако, чтобы доверять полученным оценкам, надо убедиться в применимости МНК для данной модели или воспользоваться более подходящими методами оценки.

5. Метод инструментальных переменных может быть обобщен на случай множественной регрессии и нескольких инструментов. В данном примере имеем двухфакторную регрессию, в которой один из факторов - экзогенный (доход на душу населения никак не зависит от спроса на сигареты и не может коррелировать со случайным членом уравнения (4.2)). Второй фактор является эндогенным, и для получения несмещенной и состоятельной оценки этого уравнения надо использовать инструменты. Инструменты можно расценивать как невключенные экзогенные переменные.

Если для одной эндогенной переменной используется один инструмент, то коэффициенты исходного уравнения являются в точности идентифицируемыми. Если для одной эндогенной переменной используются два и более инструментов, то уравнение является сверхидентифицируемым.

Двухшаговый метод наименьших квадратов выполняется в следующей последовательности.

Шаг первый.

- a. МНК оценка эндогенной переменной на все экзогенные (как включенные изначально в уравнение, так и инструменты).

Для нашего примера запишем в командной строке
`LS LOG(RAVGPRS) LOG(PERINC) RTAXSO RTAXS C`
Уравнение сохраним как `Eq02_1`.

- b. Получение расчетных значений `RAVGPRSF_2` (через команду **Forecast** в окне уравнения `Eq02_1`)

Второй шаг.

МНК оценка уравнения (4.2) с заменой `LOG(RAVGPRS)` на `LOG(RAVGPRSF_2)`

`LS LOG(PACKPC) LOG(RAVGPRSF_2) LOG(PERINC) C`
(сохраните уравнение как `Eq02_2`).

Результат второго шага приведен на рис. 4.4.

Полученные оценки имеют «хорошие» статистические свойства - несмещенность и состоятельность в том случае, если выполняются предпосылки использования метода инструментальных переменных (подробнее об этом см., например, Вербик, гл.5). Однако стандартные ошибки коэффициентов рассчитаны по некорректным формулам, поэтому говорить о значимости или незначимости полученных оценок преждевременно.

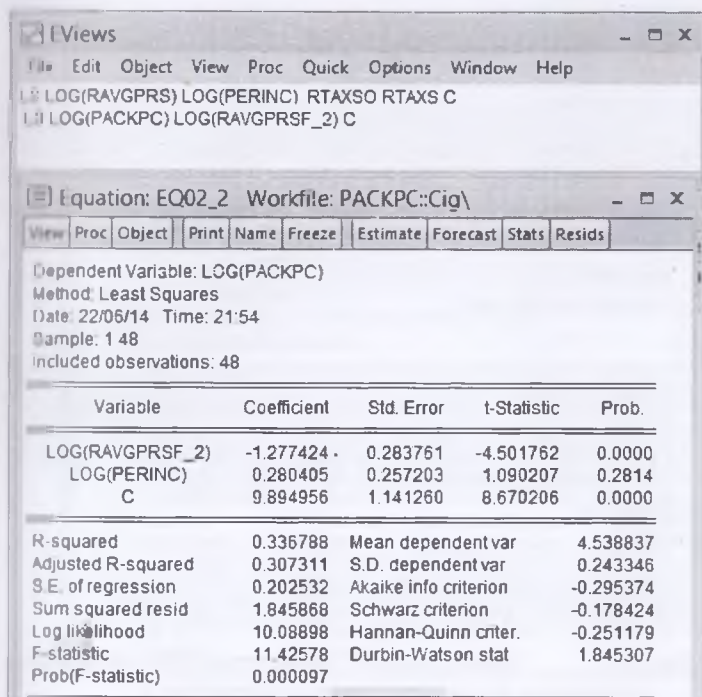


Рис. 4.5. Результат применения ДМНК к уравнению (4.2)

6. Для того чтобы получить корректные оценки и стандартные ошибки при использовании метода инструментальных переменных, воспользуемся встроенной процедурой в программе Eviews.

Сделайте копию уравнения *Eq02_ls*, откройте его и перейдите в окно **Estimate**. Выберите метод оценки - **TSLS**, при этом появится окно задания списка инструментов. В этом окне введите *LOG(PERINC) RTAXSO RTAXS*, не снимайте флажок у параметра *Include a constant* (рис. 4.6).

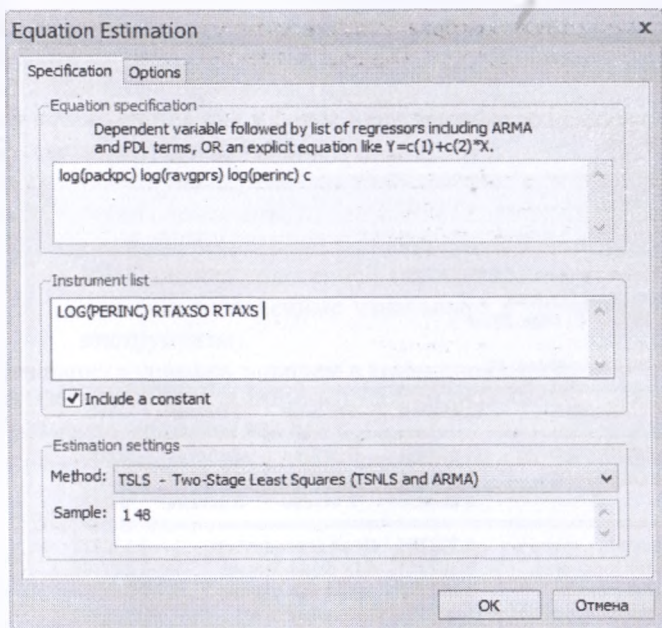


Рис. 4.6. Окно задания параметров при использовании метода инструментальных переменных

Результат оценки приведен на рис. 4.7.

Как видно из сравнения оценок, полученных двухшаговым методом «вручную» и с использованием встроенной процедуры (рис. 4.5 и 4.7), оценки коэффициентов совпадают, однако различаются стандартные ошибки коэффициентов и статистические характеристики, свидетельствующие о качестве оценки уравнения.

Приведенные на рис. 4.7 р-значения для t- и F-статистик верны в асимптотике. При использовании метода инструментальных переменных коэффициент детерминации теряет свою обычную интерпретацию и в некоторых случаях может быть даже отрицательным.

В результаты оценки включены значения ранга матрицы инструментальных переменных (Instrument rank). В данном примере ранг равен 4 по количеству инструментов (свободный коэффициент тоже учитывается). Кроме того, рассчитывается так называемый

мод J-статистика, которая связана с применением обобщенного метода моментов (ОМНК, Generalized Method of Moment - GMM).

Equation: EQ02_TSL5 Workfile: PACKPC::Cig\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: LOG(PACKPC)
 Method: Two-Stage Least Squares
 Date: 22/06/14 Time: 22:11
 Sample: 1 48
 Included observations: 48
 Instrument specification: LOG(PERINC) RTAXSO RTAXS
 Constant added to instrument list

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|--------------|-------------|------------|-------------|--------|
| LOG(RAVGPRS) | -1.277424 | 0.263199 | -4.853462 | 0.0000 |
| LOG(PERINC) | 0.280405 | 0.238565 | 1.175379 | 0.2460 |
| C | 9.894956 | 1.058560 | 9.347564 | 0.0000 |

| | | | |
|--------------------|----------|--------------------|----------|
| R-squared | 0.429422 | Mean dependent var | 4.538837 |
| Adjusted R-squared | 0.404063 | S.D. dependent var | 0.243346 |
| S.E. of regression | 0.187856 | Sum squared resid | 1.588044 |
| F-statistic | 13.28079 | Durbin-Watson stat | 1.946351 |
| Prob(F-statistic) | 0.000029 | Second-Stage SSR | 1.845868 |
| J-statistic | 0.311833 | Instrument rank | 4 |
| Prob(J-statistic) | 0.576557 | | |

Рис. 4.7. Результат оценки уравнения (4.2) с использованием двух инструментов с состоятельными стандартными ошибками

Таким образом, можно сделать вывод, что коэффициент при факторе цена статистически значим, а при факторе доход - нет, в то время как уравнение в целом статистически значимо. Эластичность по цене получилась меньше (по модулю), чем при использовании обычного МНК, но эта оценка все еще кажется слишком высокой.

7. Протестируем целесообразность применения метода инструментальных переменных для данной задачи. Воспользуемся тестом Дарбина-Ву-Хаусмана, который позволяет судить о том, значимо ли отличаются оценки, полученные по МНК или по методу ИП (подробнее о тесте см., например, Доугерти, гл.8).

Нулевая гипотеза состоит в том, что причины применения инструментальных переменных отсутствуют (нет эндогенности), и оценки, полученные по МНК и по методу ИП, не будут систематически отличаться. Тестовая статистика основана на разностях оценок коэффициентов, полученных этими двумя методами. При выполнении нулевой гипотезы об отсутствии значимых различий она имеет распределение χ^2 с числом степеней свободы, равным числу эндогенных переменных.

Чтобы выполнить тест Дарбина-Ву-Хаусмана в окне уравнения *Eq02_TSLs* выполните команду **View → IV Diagnostics & Tests → Regressor Endogeneity Test...** В открывшемся окне введите единственный подозреваемый в эндогенности регрессор уравнения (4.2), а именно *LOG(RAVGPRS)*.

Результат теста, приведенный на рис. 4.8, показывает, что нулевая гипотеза не отвергается на уровне 5%, но отвергается на уровне 10%. Такая неопределенность в выводе может являться следствием малости выборки.

| Endogeneity Test | | | |
|--|----------|----|-------------|
| Equation: EQ02_TSLs | | | |
| Specification: LOG(PACKPC) LOG(RAVGPRS) LOG(PERINC) C | | | |
| Instrument specification: C LOG(PERINC) RTAXSO RTAXS | | | |
| Endogenous variables to treat as exogenous: LOG(RAVGPRS) | | | |
| | Value | df | Probability |
| Difference in J-stats | 3.214428 | 1 | 0.0730 |

Рис. 4.8. Фрагмент таблицы с результатами теста на эндогенность регрессора *LOG(RAVGPRS)*

При выполнении нулевой гипотезы методы МНК и ИП дают состоятельные оценки, но МНК-оценки более эффективны. При выполнении альтернативной гипотезы МНК-оценки несостоятельны и смещены. В данном примере мы не можем сказать, что МНК-оценки имеют меньшие стандартные ошибки (и потому более эффективны) и более высокие *t*-статистики. На основании теста Дарбина-Ву-Хаусмана и полученных оценок следует отдать предпочтение методу инструментальных переменных для оценки эластичности спроса по цене.

Ключевые понятия

Двухшаговый метод наименьших квадратов (ДМНК)

Тест Дарбина-Ву-Хаусмана

Инструментальные переменные

Экзогенная переменная

Эндогенная переменная

Метод инструментальных переменных (ИП)

Смещенность

Состоятельность

Замещающие переменные

Вопросы

1. Что такое инструментальная переменная?
2. Для чего используется двухшаговый метод наименьших квадратов?
3. В чем суть метода инструментальных переменных? В каком случае его необходимо использовать?
4. Какие проблемы могут возникнуть из-за ошибок измерения данных?
5. В каком случае возникает проблема несостоятельности оценки?
6. Опишите алгоритм теста Дарбина-Ву-Хаусмана.

Глава 5. Анализ временного ряда индекса цен на непродовольственные товары. Одномерные модели временных рядов

Цель работы: изучение методов анализа и моделирования поведения временного ряда исходя только из его прошлых значений.

Исходные данные (PINOFOOD.wf1):

PINF– индекс цен на непродовольственные товары в период с января 2000 г. по июль 2009 г., ежемесячные данные

Источник данных: Федеральная служба государственной статистики

Задания

1. Постройте график временного ряда. Проанализируйте график, сделайте предварительные заключения о стационарности
2. Проверьте стационарность ряда с помощью подходящей процедуры. Определите порядок интегрируемости ряда
3. Оцените автокорреляционную и частную автокорреляционную функции процесса. Дайте им описание и интерпретацию, проверьте коэффициенты на значимость.
4. Какой тип базовой модели мог бы описывать ряд *PINF*?
5. Оцените параметры модели, являющейся ответом на вопрос п.4
6. Оцените альтернативные модели и сравните их с помощью соответствующих критериев (информационные критерии, единичные корни, коррелограмма теоретическая и расчетная).
7. Найдите расчетные значения индекса цен по выбранной модели, используя динамический и статический методы прогнозирования.

Решения

1. Для построения графика откройте временной ряд *PINF*, в меню **View** выберите **Graph** → **Basic type** → **Line & Symbol**. Полученный график сохраните под именем *Graph01*, оставив все параметры сохранения по умолчанию. Как видно из графика (рис.5.1), ряд претерпел значительные изменения, имея тенденцию к убыванию до середины 2000г., а затем к слабому росту. Выделяются резкие колебания значений *PINF* в начале и конце периода наблюдений. Явных признаков стационарности (постоянные средние значения и дисперсия) или нестационарности ряда (тренд или случайное блуждание) не наблюдается.

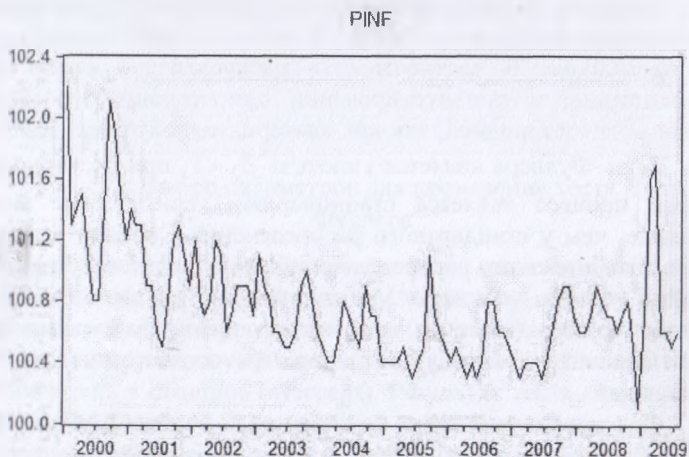


Рис. 5.1. График временного ряда *PINF*

Для проверки ряда на стационарность проведем тест Дики-Фуллера. Основная идея этого теста заключается в проверке коэффициента при переменной Y_{t-1} в предположении, что исходный ряд является процессом $AR(1)$:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \varepsilon_t \quad (5.1)$$

Уравнение процесса переписывается в виде

$$\Delta Y_t = \beta_1 + (\beta_2 - 1) Y_{t-1} + \varepsilon_t \quad (5.2)$$

и тестируется гипотеза о значимости коэффициента при Y_{t-1} .

Для тестирования нулевой гипотезы $H_0: (\beta_2 - 1) = 0$ нельзя, однако, использовать обычную статистику Стьюдента, так как при $\beta_2 = 1$ (или большем 1) стандартное t -отношение (коэффициент деленный на стандартную ошибку, t -статистика) не имеет t -распределения (распределения Стьюдента) даже асимптотически, что было показано Дики и Фуллером. Причиной этого является нестационарность процесса, при которой нарушаются предпосылки корректного применения стандартных процедур МНК-оценки. В частности, дисперсия такого процесса неограниченно возрастает со временем.

Выходом из положения является применение стандартного t -отношения, которое при нулевой гипотезе имеет нестандартное распределение. В частности, это распределение имеет правую асимметрию, и при тестировании односторонней гипотезы (а именно левосторонней, так как альтернативной гипотезой для теста Дики–Фуллера является гипотеза $\beta_2 < 1$, при истинности которой процесс является стационарным) критические значения меньше, чем у стандартного распределения Стьюдента. Критические значения этого распределения были определены Дики и Фуллером, а затем уточнены Маккинном. При проведении теста Дики–Фуллера (теста на единичный корень) Eviews использует критические значения Маккиннона (и соответствующие им p -значения).

Для проведения теста на единичный корень (**Unit Root Test**) откройте переменную *PINF* и в меню **View** выберите **Unit Root Test...**, установите переключатель, определяющий длину лага, в позицию **User specified** (установлено пользователем) и введите значение 0 - длину ряда для ряда в разностях (рис. 5.2). Остальные переключатели оставьте в положении, предложенном по умолчанию.

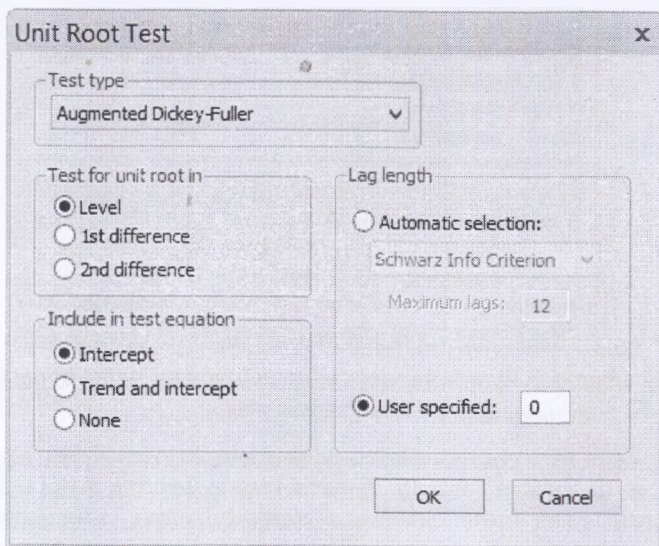


Рис. 5.2. Выбор параметров для проведения теста Дики-Фуллера

Это означает, что будет проведен тест, относящийся к типу Расширенные тесты Дики-Фуллера (**Augmented Dickey-Fuller**), тестироваться будут сами уровни ряда (переключатель в позиции **Level**) и в тестирующее уравнение (5.2) будет включена константа (переключатель в позиции **Intercept**). Результат теста сохраните в виде таблицы (**Freeze** → **Name** → **TABLE_UNITROOT1**).

Результат расширенного теста Дики-Фуллера представленный на (рис. 5. 3) показывает, что нулевая гипотеза о наличии единичного корня отвергается даже на 1% уровне значимости (значение t-статистики равно -5.135, что меньше 1% критического значения - 3.49).

Таким образом, гипотеза о нестационарности ряда *PINF* отвергается на всех уровнях значимости.

| | |
|---|---------------------|
| Null Hypothesis: $PINF$ has a unit root | |
| Exogenous: Constant | |
| Lag Length: 0 (Fixed) | |
| | t-Statistic Prob.* |
| Augmented Dickey-Fuller test statistic -5.135085 0.0000 | |
| Test critical values: | 1% level -3.488585 |
| | 5% level -2.886959 |
| | 10% level -2.580402 |
| *MacKinnon (1996) one-sided p-values. | |

Рис. 5.3. Результат теста Дики–Фуллера на наличие единичного корня у ряда $PINF$.

Этот вывод справедлив, если выполняются предпосылки теста, в частности, ряд $PINF$ является процессом $AR(1)$. Если же уровень ряда имеет значимые коэффициенты автокорреляции с более высокими лагами, то он может быть описан процессом $AR(p)$. В этом случае надо применять расширенный тест Дики–Фуллера, включающий в уравнение (5.2) лаги более высоких порядков.

Повторим тест на единичный корень, применяя расширенный тест Дики–Фуллера. Поскольку предварительного исследования о том, какой тип процесса наиболее точно описывает ряд $PINF$ (т.е. каково значение p), мы не проводили, то используем возможность автоматического выбора длины лага (т.е. переключатель установим в положение **Automatic selection**, см. рис.5.2)). Критерием выбора порядка процесса $AR(p)$ (а значит, длины лага) по умолчанию определен информационный критерий Шварца (**Schwarz Info Criteria**). Полученный результат теста сохраните в виде таблицы (**Freeze** → **Name** → **TABLE_UNITROOT2**). Результат расширенного теста Дики–Фуллера для тестирования ряда $PINF$ на стационарность представлен на рис. 5.4.

| | | |
|---|--------------------|-----------|
| Null Hypothesis: PINF has a unit root | | |
| Exogenous: Constant | | |
| Lag Length: 1 (Automatic - based on SIC, maxlag=12) | | |
| | t-Statistic | Prob.* |
| Augmented Dickey-Fuller test statistic | -4.753852 | 0.0001 |
| Test critical values: | 1% level | -3.489117 |
| | 5% level | -2.887190 |
| | 10% level | -2.580525 |
| *MacKinnon (1996) one-sided p-values. | | |

Рис. 5.4. Результат расширенного тест Дики-Фуллера для тестирования ряда *PINF* на стационарность.

В данном случае на основе критерия Шварца процесс был определен как AR(2) и длина лага установлена 1¹⁰. Это значит, что исходный ряд задается уравнением

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 Y_{t-2} + \varepsilon_t, \quad (5.3)$$

которое можно переписать в виде

$$Y_t - Y_{t-1} = \beta_1 + \beta_2 Y_{t-1} + \beta_3 Y_{t-2} - Y_{t-1} + \beta_3 Y_{t-1} - \beta_3 Y_{t-1} + \varepsilon_t \Rightarrow \Delta Y_t = \beta_1 + (\beta_2 + \beta_3 - 1)Y_{t-1} - \beta_3 \Delta Y_{t-1} + \varepsilon_t \quad (5.4)$$

Гипотеза о наличии единичного корня снова сводится к проверке на значимость коэффициента при Y_{t-1} , т.е. в данном случае

$$H_0: (\beta_2 + \beta_3 - 1) = 0 \text{ (нестационарный ряд)}$$

$$H_1: (\beta_2 + \beta_3 - 1) < 0 \text{ (стационарный ряд)}$$

Как видно из рис.5.4, нулевая гипотеза о наличии единичного корня снова отвергается даже на 1% уровне значимости (значение t-статистики равно -4,75, что меньше 1% критического значения -3,49).

Таким образом, изучаемый ряд является интегрируемым порядком ноль $I(0)$.

Построение коррелограммы ряда является вспомогательным инструментом для определения типа процесса. Под коррелограммой понимают график значений коэффициентов авто-

¹⁰ длина лага для ряда в разностях

корреляции разных порядков. Для построения коррелограммы в меню **View** выберите **Correlogram...** Так как мы определили ряд как интегрируемый нулевого порядка, то строить коррелограмму будем для самих уровней ряда (переключатель в позиции **Level**). Длину лага установите 12, основываясь на характере данных (месячные данные об индексе цен, нет оснований ожидать тесную связь между уровнями ряда при лагах более 12). Результаты сохраните в виде таблицы (**Freeze** → **Name** → **table_correlogram**).

Как видно из рис. 5.5, процесс не сводится к типу AR(1), так как в случае процесса AR(1) коэффициенты автокорреляции (AC) убывали бы по степенному закону, а среди коэффициентов частной автокорреляции (PAC) статистически значимым был бы только коэффициент первого порядка.

Значимость коэффициентов автокорреляции можно определить по Q-статистике Льюнга–Бокса (Q-stat).

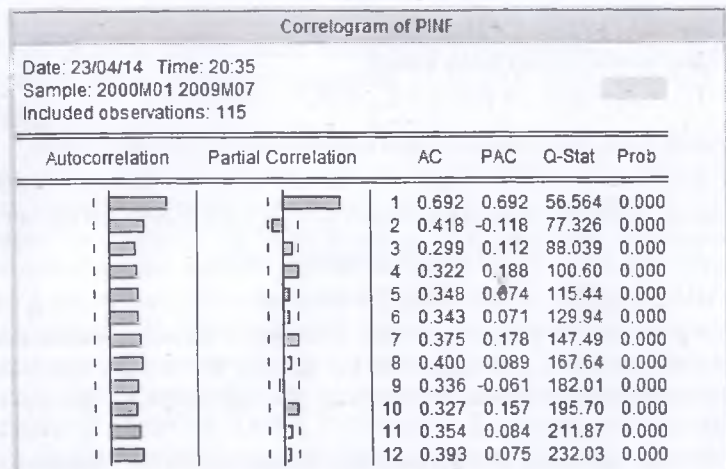


Рис. 5.5. Коррелограмма и таблица значений коэффициентов автокорреляции и частной автокорреляции уровней ряда PINF

При выполнении нулевой гипотезы (H_0 : коэффициент автокорреляции соответствующего порядка статистически незначим)

Q-стат имеет асимптотическое χ^2 распределение с числом степеней свободы, равным порядку автокорреляции. Q-статистика вычисляется по формуле

$$Q_k = T(T+2) \sum_{k=1}^K \frac{1}{T-k} r_k^2 \quad (5.5)$$

Здесь, T – число наблюдений, r_k – выборочные коэффициенты автокорреляции.

То есть для проверки на значимость коэффициента автокорреляции первого порядка надо сравнить

$$Q_1 = 115(115+2) \frac{1}{115-1} 0.692^2 = 56.56$$

с табличным значением статистики χ^2 . Для 5% уровня значимости и 1 степени свободы критическое значение равно 3.84. Следовательно, гипотеза об отсутствии автокорреляции первого порядка отвергается. Об этом же говорит и соответствующее Р-значение. Заметим, что чем больше число степеней свободы (выше лаг), тем больше критическое значение статистики χ^2 .

Для оценки значимости коэффициентов частной автокорреляции можно воспользоваться графическим представлением коррелограммы в Eviews. Пунктирной линией (см. рис.5.5) обозначена граница (приблизительно) 95% доверительного интервала для соответствующих коэффициентов. Как видно из рисунка, коэффициенты частной автокорреляции 4 и 7 порядков являются явно статистически значимыми, на значимость также претендуют коэффициенты 2, 3 и 10 порядков.

4. Исходя из анализа коррелограммы, рассматриваемый ряд может быть описан процессом AR(7), где должны быть включены, по крайней мере, первый, четвертый и седьмой лаги. Альтернативой является процесс AR(2) с включением константы, первого и второго лагов, как это следует из проведения расширенного теста Дики–Фуллера. Другой возможной альтернативой является ряд AR(12), если основываться на характере данных.

5. Оцените модель AR(7). На первом шаге включите все лаги, вплоть до 7-го, и константу. Для этого в окне спецификации уравнения Quick → Estimate Equation запишете
 PINF C AR(1) AR(2) AR(3) AR(4) AR(5) AR(6) AR(7)

Уравнение сохраните как Eq0101.

Заметим, что обозначение AR(P) в окне записи уравнения указывает на то, какой конкретно лаг включить в оценку, а не на AR(p) – процесс, который в общем случае включает все лаги вплоть до p. То есть оценивается уравнение

$$\text{PINF}_t = c + u_t \quad (5.6a)$$

где u_t – случайная составляющая, подверженная процессу авторегрессии порядка p:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t \quad (5.6b)$$

где ε_t – белый шум (если процесс подобран правильно)

Полученная оценка представлена на рис. 5.6 .

| Dependent Variable: PINF | | | | |
|--|-------------|-----------------------|-------------|--------|
| Method: Least Squares | | | | |
| Date: 24/04/14 Time: 21:27 | | | | |
| Sample (adjusted): 2000M08 2009M07 | | | | |
| Included observations: 108 after adjustments | | | | |
| Convergence achieved after 3 iterations | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 100.6795 | 0.133123 | 756.2905 | 0.0000 |
| AR(1) | 0.910657 | 0.096661 | 9.421290 | 0.0000 |
| AR(2) | -0.334515 | 0.132298 | -2.528506 | 0.0130 |
| AR(3) | -0.165945 | 0.138479 | -1.198347 | 0.2336 |
| AR(4) | 0.378087 | 0.142381 | 2.655464 | 0.0092 |
| AR(5) | -0.092381 | 0.152330 | -0.606453 | 0.5456 |
| AR(6) | -0.123430 | 0.145802 | -0.846560 | 0.3993 |
| AR(7) | 0.266019 | 0.097308 | 2.733780 | 0.0074 |
| R-squared | 0.653734 | Mean dependent var | 100.7370 | |
| Adjusted R-squared | 0.629495 | S.D. dependent var | 0.354304 | |
| S.E. of regression | 0.215662 | Akaike info criterion | -0.159025 | |
| Sum squared resid | 4.650994 | Schwarz criterion | 0.039651 | |
| Log likelihood | 16.58735 | Hannan-Quinn criter. | -0.078469 | |
| F-statistic | 26.97075 | Durbin-Watson stat | 2.053669 | |
| Prob(F-statistic) | 0.000000 | | | |

Рис. 5.6. Оценка AR(7) со всеми лагами

Таким образом, оценка имеет незначимые 3, 5 и 6 лаги. Исключение этих членов улучшает качество оценки (если основываться на критерии Шварца), все коэффициенты становятся статистически значимыми (по крайней мере, на уровне 2%).

Полученное оцененное уравнение (сохраните его как Eq0102) имеет вид (рис. 5.7):

$$PINF_t = 100.68 + u_t \quad (5.7a)$$

$$u_t = 0.90u_{t-1} - 0.41u_{t-2} + 0.20u_{t-4} + 0.16u_{t-7}$$

Это уравнение можно записать через лагированные значения урванной ряда, если использовать следующую цепочку замен:

$$PINF_{t-1} = 100.68 + u_{t-1} \Rightarrow$$

$$u_{t-1} = PINF_{t-1} - 100.68$$

Аналогично для 4-го и 7-го лага

$$u_{t-4} = PINF_{t-4} - 100.68$$

$$u_{t-7} = PINF_{t-7} - 100.68$$

Тогда оцененное уравнение имеет вид

$$PINF_t - 100.68 = 0.90(PINF_{t-1} - 100.68) - 0.41(PINF_{t-2} - 100.68) + 0.20(PINF_{t-4} - 100.68) + 0.16(PINF_{t-7} - 100.68)$$

(5.7b)

Таким образом, 100.68 – оценка среднего значения рассматриваемого ряда PINF, а данная форма записи уравнения является записью в отклонениях от среднего.

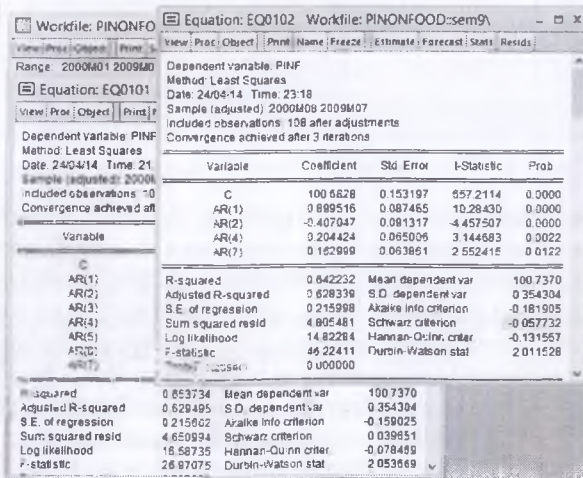


Рис. 5.7. Оценка модели AR(7) только со значимыми членами

б. Оценка альтернативной спецификации AR(2) (сохраните как *Eq02*) приводит к увеличению информационных критериев, и, следовательно, менее предпочтительна.

Оценка модели с 12 лагами неудовлетворительна, так как в силу высокой корреляционной связи между лагами имеет незначимые оценки всех коэффициентов (кроме коэффициента при первом лаге).

Для диагностики полученных уравнений можно использовать инструмент **ARMA Structure**. Чтобы диагностировать ARMA структуру, откройте в окне уравнения **View** → **ARMA Structure...**

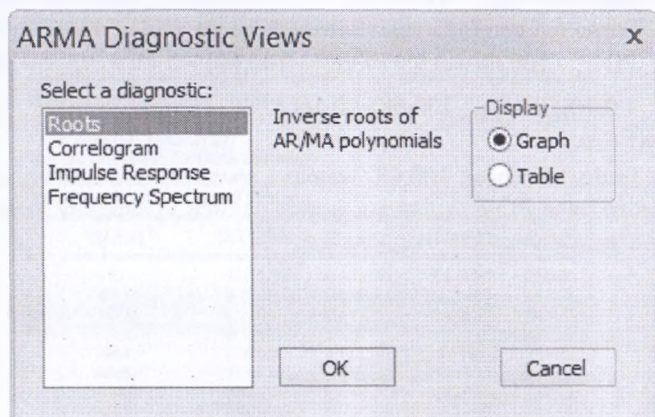


Рис. 5.8. Окно выбора инструментов для диагностики

Первым инструментом диагностики является расчет обратных корней характеристического уравнения полинома от оператора сдвига (более подробно, см. Вербик, гл.8)[4], который служит для проверки на стационарность полученной AR модели (рис.5.8). Если в этом окне выбрать позиции переключателей **Roots** и **Graph** (предложенные по умолчанию), то получим график обратных корней. В случае, когда все обратные корни лежат внутри единичного круга, рассматриваемый процесс является стационарным.

Если переключатель **Display** установить в позицию **Table**, то получим значения обратных корней в виде таблицы.

Рассмотрим инструмент **Correlogram** (см. рис. 5.8). Он позволяет сравнить теоретическую коррелограмму, построенную с учетом типа процесса, с его фактической коррелограммой. Чем ближе теоретическая и фактическая коррелограммы (для коэффициентов АС и РАС), тем лучше подобран тип процесса.

Для уравнения $Eq0102$ имеем диаграмму (рис.5.9) , где корни лежат внутри единичного круга, т.е. процесс является стационарным.

Для уравнения $Eq02$ условия стационарности также выполняются (проверьте самостоятельно, что все обратные корни лежат внутри единичного круга).

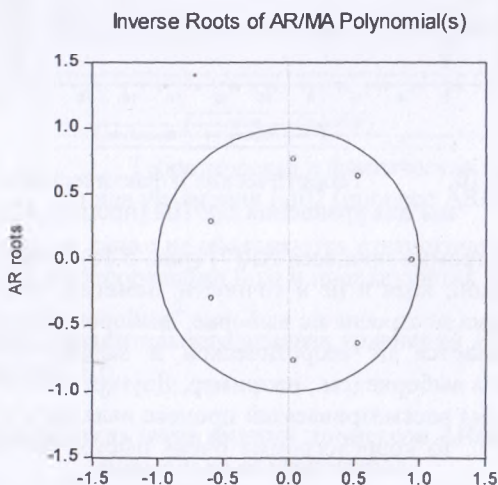


Рис. 5.9. Графическое изображение единичных корней

Для уравнений $Eq0102$ и $Eq02$ имеем следующие теоретическую и фактическую коррелограммы

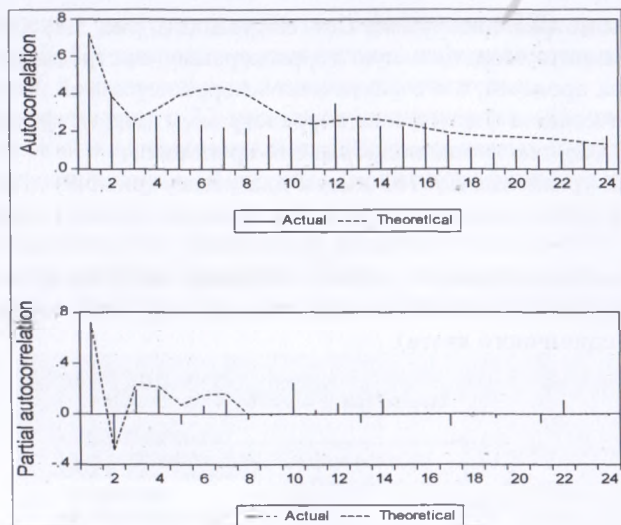


Рис. 5.10. Теоретическая и фактическая коррелограммы для уравнения $Eq0102$ (процесс $AR(7)$)

Коррелограмма процесса $AR(7)$ (рис. 5.10) более соответствует фактической, хотя и не в точности. Заметим, что фактическая коррелограмма построена по выборке, выборочная коррелограмма всегда отличается от теоретической и зависит от количества наблюдений в выборке (см., например, Доугерти, гл. 13)[6].

А если бы рассматриваемый процесс являлся в точности процессом $AR(2)$, то коррелограмма очень быстро бы убывала (рис. 5.11).

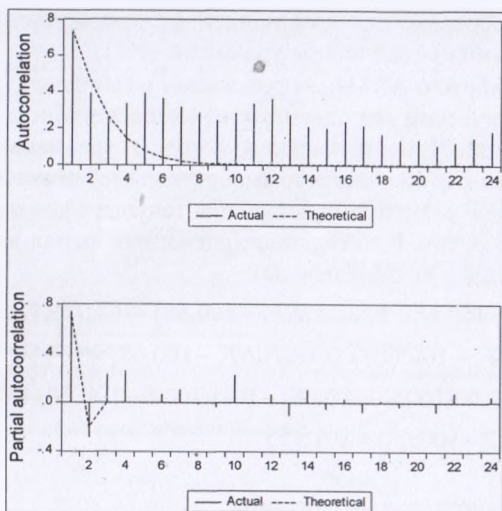


Рис. 5.11. Теоретическая и фактическая коррелограммы для уравнения $Eq02$ (процесс $AR(2)$)

В этом случае также не объясняются статистически значимые коэффициенты автокорреляции 4-го и последующих порядков (см. рис.5.5).

Результаты сравнительного анализа уравнений $Eq0102$ и $Eq02$ приведены в табл. 5.1.

Таблица 5.1
Сравнение результатов оценки процессов $AR(2)$ и $AR(7)$ по формальным критериям

| Уравнение/процесс | Информационный критерий Шварца | Обратные корни | Коррелограмма |
|-------------------|--------------------------------|-------------------------|-----------------------------------|
| $Eq0102/AR(7)$ | -0.0577 | Внутри единичного круга | Более соответствует теоретической |
| $Eq02/AR(2)$ | 0.0596 | Внутри единичного круга | Менее соответствует теоретической |

Напомним, что по информационным критериям выбирается та спецификация, у которой значение информационного критерия меньше (с учетом знака).

Таким образом, для дальнейшей работы останавливаемся на уравнении Eq0102 (оцененное уравнение (5.7)).

7. При работе с ARMA—процессами различают статический и динамический методы получения расчетных и прогнозных значений. При статическом методе в уравнение (5.7b) надо подставить фактические (наблюдаемые) значения переменной PINF с лагом для получения текущего значения.

Например, для 8-го периода (расчетные значения для предыдущих периодов не определены):

$$\begin{aligned} \text{PINF}_8^{\text{Static}} &= 100.68 + 0.90(\text{PINF}_7 - 100.68) - 0.41(\text{PINF}_6 - 100.68) + \\ &+ 0.20(\text{PINF}_4 - 100.68) + 0.16(\text{PINF}_1 - 100.68) = \\ &= 100.68 + 0.90(100.8 - 100.68) - 0.41(100.8 - 100.68) + 0.20(101.5 - 100.68) \\ &+ 0.16(102.2 - 100.68) = 101.155 \end{aligned}$$

Соответственно, для 9-го периода

$$\begin{aligned} \text{PINF}_9^{\text{Static}} &= 100.68 + 0.90(\text{PINF}_8 - 100.68) - 0.41(\text{PINF}_7 - 100.68) + \\ &+ 0.20(\text{PINF}_5 - 100.68) + 0.16(\text{PINF}_2 - 100.68) = 101.466 \end{aligned}$$

При динамическом методе в расчетном значении текущего периода используются расчетные (а не фактические) значения для предыдущих периодов. Т.е. для 8-го периода получим то же расчетное значение, что и статическим методом, а начиная с 9-го периода, в расчете участвуют оцененные значения уровней ряда, т.е.

$$\text{PINF}_8^{\text{Dynamic}} = \text{PINF}_8^{\text{Static}}$$

$$\begin{aligned} \text{PINF}_9^{\text{Dynamic}} &= 100.68 + 0.90(\text{PINF}_8^{\text{Dynamic}} - 100.68) - 0.41(\text{PINF}_7 - 100.68) + \\ &+ 0.20(\text{PINF}_5 - 100.68) + 0.16(\text{PINF}_2 - 100.68) = 101.246 \end{aligned}$$

Что касается прогноза, то статическим методом можно сделать прогноз только на 1 период вперед (так как используются только фактические значения), а динамическим – на любое число периодов.

Для получения расчетных значений в меню уравнения Eq0102 выберите команду **Forecast**. Установите переключатель **Method** в положение **Dymanic forecast**. Имя оцененного ряда задайте PINFFDYNAMIC (рис. 5.12). В результате будет создан ряд расчетных значений с заданным именем.

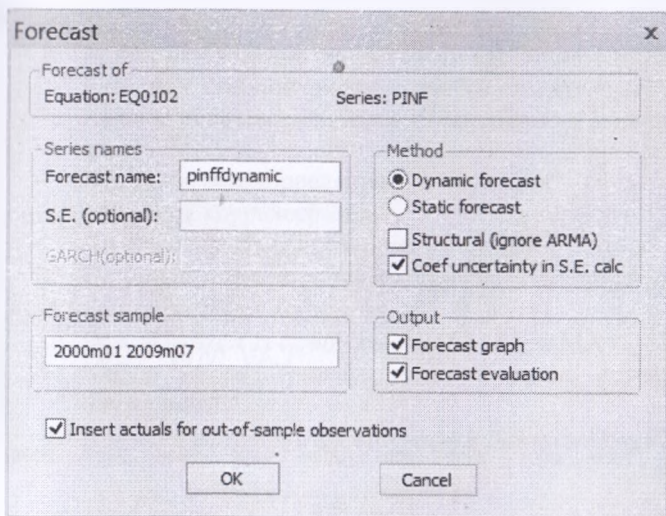


Рис. 5.12. Диалоговое окно для задания параметров прогноза

Повторно воспользуйтесь командой **Forecast** и получите ряд расчетных значений статическим методом (имя ряда PINFFSTATIC). Откройте полученные ряды и сверьте значения 8-го и 9-го периодов с приведенными выше.

Для наглядного сравнения фактических и расчетных значений, полученных разными методами, откройте ряды PINF, PINFFDYNAMIC, PINFFSTATIC как группу и постройте график (рис. 5.13).

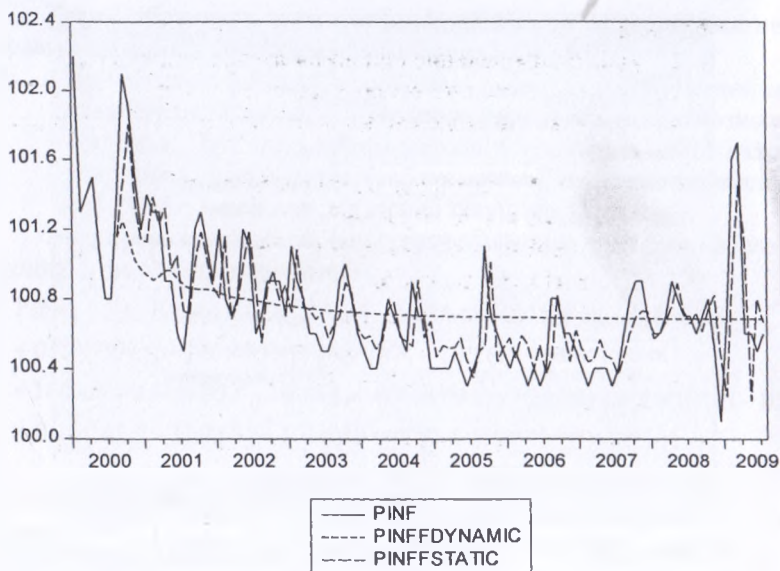


Рис. 5.13. График фактических и расчетных значений, полученных статическим и динамическим методами

Получение расчетных значений статическим методом дает лучшее приближение фактических и расчетных значений, однако динамический метод позволяет выявить тенденцию и пригоден для прогноза.

Ключевые понятия

ARMA процессы
 Коррелограмма
 Информационные критерии
 Стационарность
 Нестационарность
 Единичные корни

Тестирование единичных корней
 Расширенный тест Дики-Фуллера (ADF-тест)
 Статическое и динамическое прогнозирование

Вопросы

1. Дайте определение одномерного временного ряда
2. Что такое стационарность и слабая стационарность?
3. Каковы основные свойства стационарного временного ряда?
4. Что такое автокорреляционная функция?
5. Что такое коррелограмма?
6. Что такое частная автокорреляционная функция?
7. Опишите последовательность действий при проведении теста Дики–Фуллера.
8. Чем отличаются процессы ARMA, AR и MA?
9. Какую коррелограмму имеет процесс случайного блуждания?
10. Для чего используется статистика Льюнга–Бокса?

Глава 6. Моделирование волатильности обменных курсов. ARCH и GARCH модели

Цель работы: знакомство с понятием авторегрессионной условной гетероскедастичности, применение ARCH и GARCH моделей для оценки волатильности.

Исходные данные (EURUSD.wf1):

EURUSD – ежедневные данные об обменном курсе евро к доллару за период с 29/04/2002 по 28/04/2010. Всего 2072 наблюдения (исключая праздничные и выходные дни).

Источник данных: www.finam.ru

Задания

1. Проанализируйте исходный ряд и определите подходящий тип процесса и порядок интегрируемости ряда.
2. Постройте ряд приращений логарифма обменного курса (имя ряда *R_EURUSD*). Постройте график полученного ряда, проанализируйте его, сделайте предварительные заключения о поведении волатильности ряда.
3. Оцените тривиальную модель для ряда лог-доходности
 $R_EURUSD_t = \text{Const} + u_t$

(6.1)

и протестируйте наличие условной гетероскедастичности в ряде остатков u_t .

4. Оцените модель ARCH(1), проверьте, сохранились ли ARCH эффекты в остатках. Оцените альтернативные модели: ARCH(6), GARCH(1,1), EGARCH(1,1). Сравните полученные результаты.
5. Рассчитайте условные стандартные отклонения, полученные из разных моделей, постройте график, ограничив выборку 100 последними наблюдениями.

Решения

1. Для определения типа процесса построим график ряда, его коррелограмму и проведем тест на единичный корень. Как

показывает коррелограмма уровней ряда (рис. 6.1), коэффициент корреляции первого порядка практически равен 1 (0.997), коэффициенты последующих порядков убывают очень медленно, коэффициенты частной автокорреляции, начиная со второго порядка, близки к 0. Это свидетельствует о том, что ряд может быть описан процессом случайного блуждания.

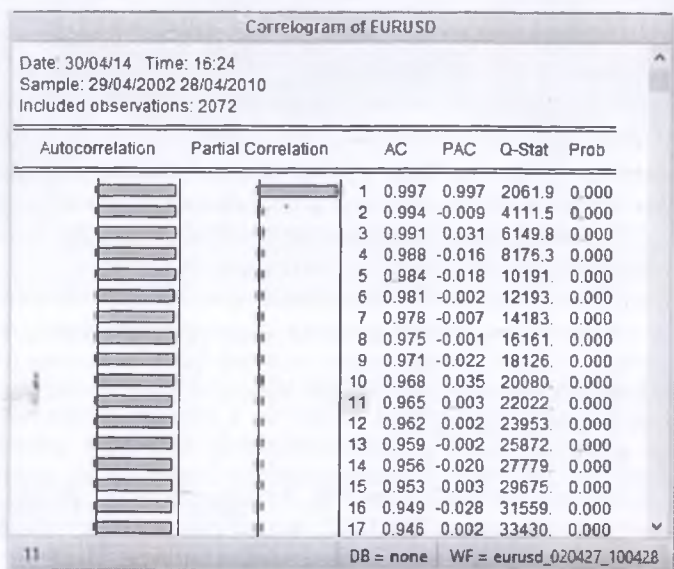


Рис. 6.1. Коррелограмма ряда котировок валют

Случайное блуждание является нестационарным процессом. Для тестирования нашего ряда на стационарность проведем тест Дики–Фуллера. Гипотеза о нестационарности исходного ряда не отвергается по результатам теста (рис. 6.2).

| | |
|---|--------------------|
| Null Hypothesis: EURUSD has a unit root | |
| Exogenous: Constant | |
| Lag Length: 0 (Automatic - based on SIC, maxlag=25) | |
| | t-Statistic Prob.* |

| | | |
|--|-----------|-----------|
| Augmented Dickey-Fuller test statistic | -2.543036 | 0.1054 |
| Test critical values: | 1% level | -3.433309 |
| | 5% level | -2.862733 |
| | 10% level | -2.567451 |
| *MacKinnon (1996) one-sided p-values. | | |

Рис. 6.2. Результат теста Дики–Фуллера для ряда котировок валют

Построение коррелограммы и результаты теста Дики–Фуллера для первой разности исходного ряда показывают, что первая разность ряда $\Delta EURUSD = EURUSD_t - EURUSD_{t-1}$ хорошо описывается процессом белого шума (все коэффициенты автокорреляции близки к нулю, гипотеза о нестационарности ряда отвергается). Таким образом, порядок интегрируемости ряда равен 1, т.е. исходный ряд котировок стационарен в разностях.

Доходность, которая определяется как отношение разности ряда к его значению, также является стационарным рядом в данном случае. Однако в финансовом анализе чаще используется логарифмическая доходность, которая имеет лучшие статистические свойства и определяется как

$$R_Y_t = LOG(Y_t / Y_{t-1})$$

2. Построим логарифмическую доходность R_EURUSD ряда котировок валют $EURUSD$, используя команду **Quick** → **Generate Series...** В диалоговом окне введем $R_EURUSD=LOG(EURUSD/EURUSD(-1))$

График временного ряда R_EURUSD (рис. 6.3) показывает существование периодов с низкой и высокой волатильностью¹¹. То есть большие возмущения (остатки) имеют тенденцию к последующим большим возмущениям и, соответственно, малые возмущения имеют тенденцию к малым возмущениям.

¹¹ Обычно волатильность моделируется дисперсией

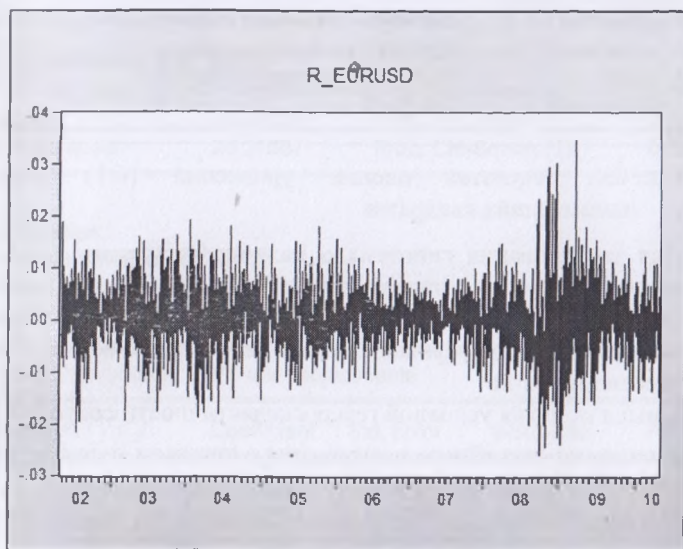


Рис. 6.3. График ряда лог-доходности

Одним из способов моделирования такой структуры является использование модели с условной гетероскедастичностью (Auto-Regressive Contional Heteroscedastisity, ARCH): модели, которая основана на введении предположения о зависимости дисперсии остатков от предыстории.

3. Результат оценки уравнения (6.1), приведен на рис 6.4, сохраните под именем (Eq01).

| Dependent Variable: R_EURUSD | | | | |
|---|-------------|--------------------|-------------|----------|
| Method: Least Squares | | | | |
| Date: 30/04/14 Time: 17:14 | | | | |
| Sample (adjusted): 30/04/2002 28/04/2010 | | | | |
| Included observations: 2071 after adjustments | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 0.000151 | 0.000140 | 1.078473 | 0.2809 |
| R-squared | 0.000000 | Mean dependent var | | 0.000151 |
| Adjusted R-squared | 0.000000 | S.D. dependent var | | 0.006366 |

| | | | |
|--------------------|----------|-----------------------|-----------|
| S.E. of regression | 0.006366 | Akaike info criterion | -7.275162 |
| Sum squared resid | 0.083893 | Schwarz criterion | -7.272440 |
| Log likelihood | 7534.430 | Hannan-Quinn criter. | -7.274164 |
| Durbin-Watson stat | 1.998271 | | |

Рис. 6.4. Результат оценки уравнения (6.1) методом наименьших квадратов

Для тестирования гипотезы о наличии/отсутствии условной гетероскедастичности остатков модели в окне уравнения *Eq01* выберите команды **View** → **Residual Diagnostics** → **Heteroskedasticity Tests...** В окне спецификации теста **Test Type** выберите **ARCH** (количество лагов 1).

Смысл понятия условной гетероскедастичности состоит в том, что дисперсия случайного возмущения в текущем периоде u_t^2 зависит от дисперсии предыдущего периода (т.е. квадрат остатков модели временного ряда, например, модели (6.1), является функцией квадрата остатков предыдущего периода)¹²:

$$u_t^2 = (\lambda_0 + \lambda_1 u_{t-1}^2) + \varepsilon_t \quad (6.2)$$

$$\lambda_1 < 1$$

¹² Такая модель называется моделью ARCH(1)

| Heteroskedasticity Test: ARCH | | | | |
|---|-------------|------------|---------------------|--------|
| F-statistic | 26.62816 | | Prob. F(1,2068) | 0.0000 |
| Obs*R-squared | 26.31507 | | Prob. Chi-Square(1) | 0.0000 |
| Test Equation: | | | | |
| Dependent Variable: RESID^2 | | | | |
| Method: Least Squares | | | | |
| Date: 30/04/14 Time: 17:43 | | | | |
| Sample (adjusted): 6/05/2002 28/04/2010 | | | | |
| Included observations: 2070 after adjustments | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
| C | 3.60E-05 | 1.94E-06 | 18.53512 | 0.0000 |
| RESID^2(-1) | 0.112753 | 0.021850 | 5.160248 | 0.0000 |

Рис. 6.5. Тестирование условной гетероскедастичности в модели (6.1)

где $\varepsilon_t \sim iid N(0, \sigma^2)$ - независимая нормально распределенная случайная величина с нулевым средним и постоянной дисперсией

Тогда условная дисперсия ошибки h_t^2 может быть выражена как

$$h_t^2 = \text{var}(u_t | u_{t-1}) = E(u_t^2 | u_{t-1}) = \lambda_0 + \lambda_1 u_{t-1}^2 \quad (6.3)$$

Условная дисперсия, как видно из (6.3), зависит от последнего значения u и является величиной переменной, в то время как безусловная дисперсия (6.2) постоянна (вывод см., например, Елисеева, гл 13)[7].

При проведении ARCH-теста проверяется гипотеза $H_0: \lambda_1 = 0$, против альтернативной $H_1: \lambda_1 \neq 0$ на основе ряда остатков из оцененного уравнения (6.1).

Результаты теста (рис.6.5) показывают, что нулевая гипотеза об отсутствии условной гетероскедастичности отвергается на всех уровнях значимости (по F-статистике и по χ^2 - статистике).

Проведенный тест является частным случаем теста Бреуша–Пагана. В качестве более общего варианта теста можно рассмотреть в уравнении для дисперсии (6.2) большее количество лагов¹³. В этом случае тестируется гипотеза об одновременном равенстве нулю коэффициентов при всех лагированных переменных.

Условная гетероскедастичность не нарушает предпосылок МНК, но игнорирование ARCH эффектов ведет к потере эффективности МНК-оценки.

4. Оценим модель ARCH(1). Для этого в окне записи уравнения выберем метод оценки ARCH (рис. 6.6)

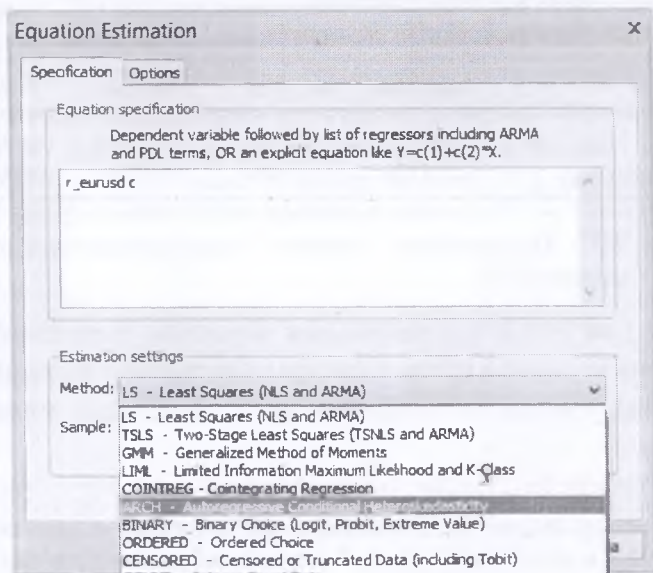


Рис. 6.6. Выбор метода оценки уравнения при наличии ARCH—эффектов

В диалоговом окне выбора спецификации установим: ARCH - 1 и GARCH - 0 (рис.6.7), результат сохраним под именем *Eq02*.

¹³ Модель ARCH(p), где p- количество лагов в уравнении для текущей дисперсии

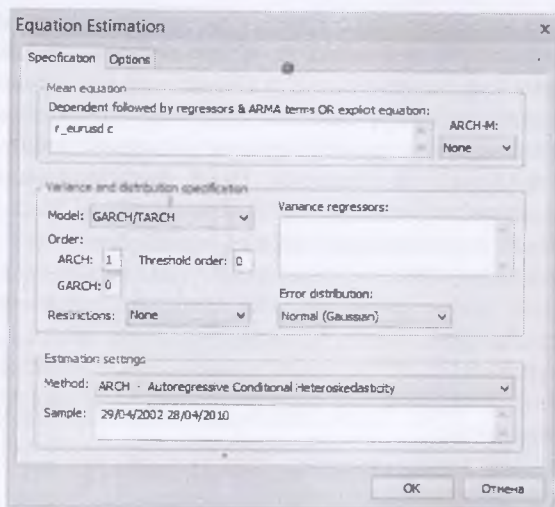


Рис. 6.7. Установка параметров для оценки модели ARCH(1)

Заметим, что при оценке моделей с условной гетероскедастичностью используется метод максимального правдоподобия (ML – Maximum Likelihood) (рис.6.8).

| Dependent Variable: R_EURUSD | | | | |
|---|-------------|------------|-------------|--------|
| Method: ML - ARCH | | | | |
| Sample (adjusted): 30/04/2002 28/04/2010 | | | | |
| Included observations: 2071 after adjustments | | | | |
| GARCH = C(2) + C(3)*RESID(-1)^2 | | | | |
| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
| C | 0.000198 | 0.000140 | 1.416350 | 0.1567 |
| Variance Equation | | | | |
| C | 3.70E-05 | 1.08E-06 | 34.25105 | 0.0000 |
| RESID(-1)^2 | 0.084992 | 0.019149 | 4.438416 | 0.0000 |

Рис. 6.8. Результат оценки модели ARCH(1)

В протоколе отчета о результатах оценки приводятся: оценка коэффициента C (константы из уравнения(6.1)), его стандартная ошибка, значение z -статистики¹⁴ и P -значение для проверки гипотезы на значимость коэффициента. В нижней части данного отчета приводится оценка уравнения для условной дисперсии, т.е. оценка коэффициентов уравнения (6.3). Eviews для записи ряда условной дисперсии использует обозначение GARCH.

В нижней половине окна оценки приводится стандартная регрессионная статистика, основанная на остатках уравнения (6.1). Заметим, что такие величины, как коэффициент детерминации, могут иметь бессмысленные значения в данных моделях, так как в основном уравнении нет ни одного регрессора. В данном примере коэффициент детерминации отрицательный и не имеет стандартной интерпретации.

Таким образом, согласно полученной оценке ряд лог-доходностей описывается моделью

$$R_EURUSD = 0.000198 + u_t$$

(0.00014)

$$h_t^2 = 3.7 \cdot 10^{-5} + 0.085 u_{t-1}^2$$

(1.08 \cdot 10^{-6}) \quad (0.019)

Проверьте полученную модель на адекватность. Для этого повторите процедуру тестирования остатков на условную гетероскедастичность (ARCH—эффектов) для уравнения Eq02, установите количество лагов, равное 6.

| Heteroskedasticity Test: ARCH | | | |
|-------------------------------|----------|---------------------|--------|
| F-statistic | 19.54287 | Prob. F(6,2058) | 0.0000 |
| Obs*R-squared | 111.3138 | Prob. Chi-Square(6) | 0.0000 |

Рис. 6.9. Результат тестирования дополнительных ARCH—эффектов

¹⁴ z -статистика есть оценка коэффициента, полученная методом максимального правдоподобия к его стандартной ошибке, при нулевой гипотезе имеет асимптотически нормальное распределение

Результат теста (рис. 6.9) показывает наличие дополнительных ARCH—эффектов, которые необходимо учесть в спецификации.

Оценим модель ARCH(6), так как тест показал, что, по крайней мере до 6 лага включительно, ARCH—эффекты присутствуют. Сохраним как уравнение *Eq03*.

Как видно из рис. рис. 6.10, все 6 лагов (за исключением первого) имеют значимый положительный эффект и не имеют тенденции к убыванию, т.е. влияние лагированных возмущений снижается очень медленно.

| Dependent Variable: R_EURUSD | | | | |
|--|-------------|------------|-------------|--------|
| Method: ML - ARCH | | | | |
| Sample (adjusted): 30/04/2002 28/04/2010 | | | | |
| Included observations: 2071 after adjustments | | | | |
| GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*RESID(-2)^2 + C(5)*RESID(-3)^2 + C(6)*RESID(-4)^2 + C(7)*RESID(-5)^2 + C(8)*RESID(-6)^2 | | | | |
| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
| C | 0.000260 | 0.000124 | 2.097675 | 0.0359 |
| Variance Equation | | | | |
| C | 2.04E-05 | 1.30E-06 | 15.75058 | 0.0000 |
| RESID(-1)^2 | 0.021251 | 0.019241 | 1.104451 | 0.2694 |
| RESID(-2)^2 | 0.056149 | 0.020684 | 2.714599 | 0.0066 |
| RESID(-3)^2 | 0.127277 | 0.024449 | 5.205882 | 0.0000 |
| RESID(-4)^2 | 0.073415 | 0.020365 | 3.605035 | 0.0003 |
| RESID(-5)^2 | 0.117044 | 0.020921 | 5.594449 | 0.0000 |
| RESID(-6)^2 | 0.098598 | 0.022465 | 4.388925 | 0.0000 |

Рис. 6.10. Результат оценки модели ARCH(6)

Для более полного учета воздействия всей предыдущей информации на текущую волатильность перейдем к обобщенной авторегрессионной модели с условной гетероскедастичностью GARCH.

Для этой модели условная дисперсия h_t^2 задается формулой

$$h_t^2 = \text{var}(u_t | \Omega_{t-1}) = \lambda_0 + \lambda_1 u_{t-1}^2 + \dots + \lambda_p u_{t-p}^2 + \omega_1 h_{t-1}^2 + \dots + \omega_q h_{t-q}^2 \quad (6.4),$$

где Ω_{t-1} - вся информация, доступная в момент времени $t-1$, а λ и ω - параметры модели, подлежащие определению.

На практике хорошо работает GARCH(1,1) с тремя неизвестными параметрами:

$$h_t^2 = \lambda_0 + \lambda u_{t-1}^2 + \omega h_{t-1}^2 \quad (6.5)$$

Таким образом, уравнение для условной дисперсии содержит константу (λ_0), информацию о волатильности предыдущего периода, измеренную как лагированное значение квадрата остатков уравнения (1) u_{t-1}^2 (ARCH-составляющая) и прогнозное значение условной дисперсии на предыдущий период h_{t-1}^2 (GARCH-составляющая). Обозначение GARCH(1,1) означает, что мы имеем авторегрессию первого порядка для GARCH-составляющей и MA-процесс первого порядка для ARCH-составляющей. ARCH(1) есть частный случай GARCH-модели, который может быть записан как GARCH(0,1).

Модель GARCH(1,1) (6.5) интерпретируется в финансовых задачах как прогнозное значение волатильности, которое является взвешенной суммой долгосрочного среднего (константа), прогноза волатильности на прошлый период (GARCH-составляющая) и наблюдаемой волатильности предыдущего периода (ARCH-составляющая). Если доходность актива имеет неожиданно большие отклонения в ту или другую сторону, то оценка волатильности на следующий период вырастет. Таким образом, эта модель подходит для описания кластеров волатильности, наблюдаемых на практике.

Для стационарности исходного процесса требуем, чтобы $(\lambda + \omega) < 1$. Если $(\lambda + \omega) \geq 1$, процесс является интегрированным первого порядка.

Спецификация GARCH является более экономным вариантом спецификации ARCH высокого порядка.

Для оценки GARCH(1,1) в диалоговом окне выбора спецификации (см. рис.6.7) установите: ARCH - 1 и GARCH - 1, результат сохраните под именем *Eq04*.

| Dependent Variable: R_EURUSD | | | | |
|---|-------------|------------|-------------|--------|
| Method: ML - ARCH (Marquardt) - Normal distribution | | | | |
| Included observations: 2071 after adjustments | | | | |
| GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1) | | | | |
| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
| C | 0.000241 | 0.000118 | 2.035117 | 0.0418 |
| Variance Equation | | | | |
| C | 1.44E-07 | 6.49E-08 | 2.214397 | 0.0268 |
| RESID(-1)^2 | 0.035683 | 0.005696 | 6.265016 | 0.0000 |
| GARCH(-1) | 0.960646 | 0.005667 | 169.5078 | 0.0000 |

Рис. 6.11. Результат оценки модели GARCH(1,1)

Полученная модель (рис.6.11) является достаточно простой: доходность R_EURUSD равна константе 0.000241 (т.е. 0.024% в среднем в день) с условной дисперсией по всей доступной информации к моменту времени t , задающейся формулой

$$h_t^2 = 1.44 \cdot 10^{-7} + 0.035u_{t-1}^2 + 0.96h_{t-1}^2$$

Сумма коэффициентов $(\lambda + \omega) \cong 1$, т.е. ряд условных дисперсий является нестационарным. Это говорит о том, что шоки волатильности очень медленно затухают. Это характерно для финансовых временных рядов, имеющих высокую частоту (например, ежедневных).

Недостатком ARCH и GARCH моделей является их симметрия: возмущения входят в уравнения (6.2) и (6.3) в квадрате, следовательно, нет различий между положительными и отрицательными возмущениями. Большое отрицательное возмущение имеет то же воздействие на будущую волатильность, что и большое положительное возмущение. Это допустимо для валютных рынков, но для фондовых рынков неожиданное снижение цены («плохие новости») имеет большее воздействие на будущую волатильность, чем неожиданное увеличение цены («хорошие новости»).

Устранение этого недостатка предпринято в асимметричных GARCH-моделях, в которых положительное и отрицательное возмущения имеют разные воздействия на будущую волатильность.

Основополагающий подход заложен в экспоненциальной GARCH-модели, предложенной Нельсоном (EGARCH). Условная дисперсия в EGARCH(1,1) задается формулой

$$\text{LOG}(h_t^2) = \lambda_0 + \beta \text{LOG}(h_{t-1}^2) + \gamma \frac{u_{t-1}}{h_{t-1}} + \alpha \frac{|u_{t-1}|}{h_{t-1}}$$

(6.5)

где α, β и γ – постоянные параметры. Так как модель включает $\frac{u_{t-1}}{h_{t-1}}$, то она асимметрична, если $\gamma \neq 0$. Действительно, при $\gamma < 0$ отрицательное воздействие имеет большее влияние, чем положительное.

Для оценки EGARCH(1,1) в диалоговом окне выбора спецификации (рис.6.12) выберите EGARCH и установите: ARCH - 1, GARCH - 1, Asymmetric order – 1, результат сохраните под именем *Eq05*.

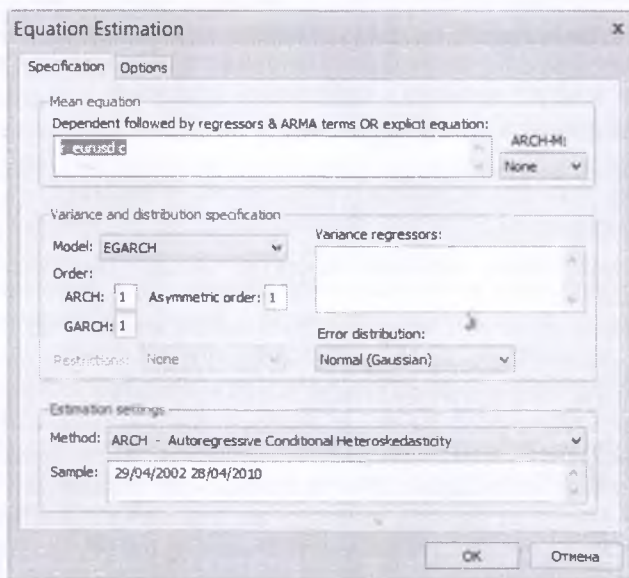


Рис. 6.12. Окно установки параметров оценки EGARCH(1,1)

Как видно из рис. 6.13, оценка коэффициента γ статистически незначима (в записи параметров в программе Eviews коэффициент C(4) = -0.0089, Prob=0.2). То есть асимметрия в модели отсутствует, что характерно для рядов обменных курсов. Также оценка коэффициента β близка к 1, что отражает высокую степень постоянства в волатильности обменных курсов.

| Dependent Variable: R_EURUSD | | | | |
|--|-------------|------------|-------------|--------|
| Method: ML - ARCH (Marquardt) - Normal distribution | | | | |
| Sample (adjusted): 30/04/2002 28/04/2010 | | | | |
| Included observations: 2071 after adjustments | | | | |
| LOG(GARCH) = C(2) + C(3)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(4)*RESID(-1)/@SQRT(GARCH(-1)) + C(5)*LOG(GARCH(-1)) | | | | |
| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
| C | 0.000257 | 0.000120 | 2.142908 | 0.0321 |
| Variance Equation | | | | |
| C(2) | -0.117328 | 0.024049 | -4.878650 | 0.0000 |
| C(3) | 0.077551 | 0.012168 | 6.373410 | 0.0000 |
| C(4) | -0.008882 | 0.007049 | -1.259951 | 0.2077 |
| C(5) | 0.994466 | 0.001972 | 504.2209 | 0.0000 |

Рис. 6.13. Оценка модели EGARCH(1,1)

- Для создания рядов условных дисперсий в окне уравнений *Eq03* выберите команды **Proc** → **Make GARCH Variance Series**. Переменную назовите GARCH03. Аналогично, создайте ряды по уравнениям *Eq04*, *Eq05*. Откройте созданные ряды как группу и постройте график, ограничив выборку 100 последними наблюдениями (в окне **Sample** задайте **@last-100 @last**).

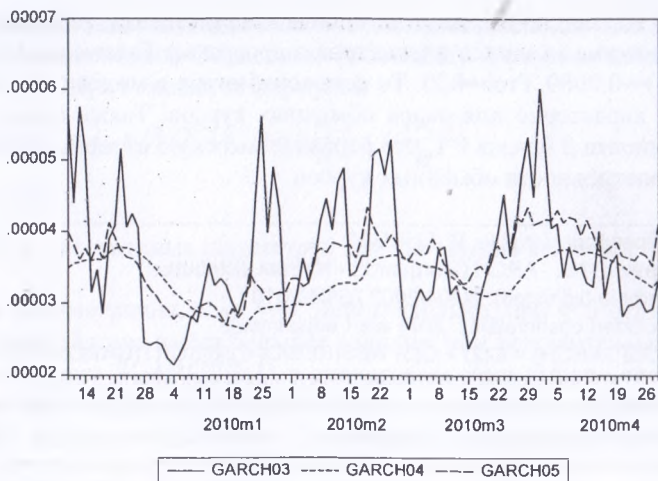


Рис. 6.14. График условной волатильности по моделям ARCH(6), GARCH(1,1) и EGARCH(1,1)

Как видно из графика, приведенного на рис. 6.14, волатильность, предполагаемая моделью ARCH(6), менее гладкая, чем для более общих моделей на основе учета всей предшествующей информации.

Ключевые понятия

Условная гетероскедастичность
 ARCH-модель
 GARCH-модель
 EGARCH-модель
 Волатильность

Порядок интегрируемости ряда
 Тестирование ARCH-эффектов
 Логарифмическая доходность

Вопросы

1. Что такое условная гетероскедастичность?
2. Как определить порядок интегрируемости ряда?
3. Как протестировать наличие ARCH-эффектов?
4. Что такое модель Бокса–Дженкинса?

5. Какие достоинства и недостатки есть у ARCH— и GARCH— моделей?
6. Приведите примеры обобщенных GARCH—моделей

Глава 7. Моделирование совокупного потребления на основе динамических моделей временных рядов

Цель работы: изучить подходы к моделированию динамики совокупного потребления на основе механизма адаптивных ожиданий (модель Фридмана) и механизма частичной корректировки (модель Брауна).

Исходные данные (Hamilton_coint.wfl):

LY – натуральный логарифм агрегированного располагаемого дохода (увеличенный в 100 раз)

LC – натуральный логарифм расходов на личное потребление (увеличенный в 100 раз).

Данные по США, квартальные, период с 1 квартала 1947 г. по 3 квартал 1989 г.

Источник данных: Hamilton, J. (1994) Time series analysis, Princeton: Princeton University Press, цитируется по Eviews 7 User Guide.

Задания

1. Оцените краткосрочную и долгосрочную эластичность к потреблению на основе модели адаптивных ожиданий, используя линейный и нелинейный МНК.
2. Определите 90% интервал продолжительности воздействия дохода на потребление на основании модели адаптивных ожиданий.
3. Объясните, как результаты оценки регрессионного уравнения, полученные в задании 1, могут быть интерпретированы в модели частичной корректировки. Поясните механизм процесса корректировки.
4. Проверьте остатки регрессионного уравнения на автокорреляцию. Какой вывод можно сделать в отношении двух предложенных моделей.

Решения

1. Рассмотрим гипотезу Фридмана о постоянном доходе. Суть заключается том, что фактическое потребление раскладывается на постоянную C_t^P и переменную C_t^T части (аналогично, фактический доход)

$$C_t = C_t^P + C_t^T \quad (7.1)$$

$$Y_t = Y_t^P + Y_t^T$$

Между постоянными составляющими устанавливается соотношение

$$C_t^P = \beta_2 Y_t^P v_t \quad (7.2)$$

где C_t^P - постоянное потребление; Y_t^P - постоянный доход; v_t - мультипликативный случайный член; β_2 - параметр, подлежащий оценке, а переменные C_t^T и Y_t^T являются случайными величинами с нулевым математическим ожиданием. Для разрешения проблемы не наблюдаемости постоянного дохода Фридман предложил механизм адаптации: приращение постоянного дохода пропорционально разности фактического и постоянного дохода предыдущего периода

$$Y_t^P - Y_{t-1}^P = \lambda(Y_t - Y_{t-1}^P) \quad (7.3)$$

Следовательно, постоянный доход в момент t (его ожидаемые значения) есть средневзвешенное значение фактического дохода в момент t и постоянного дохода в момент $t-1$:

$$Y_t^P = \lambda Y_t + (1 - \lambda) Y_{t-1}^P \quad (7.4)$$

В случае $\lambda = 1$ постоянный и фактический доход совпадают. В другом предельном случае $\lambda = 0$ ожидаемые значения постоянного дохода никак не корректируются значением фактического дохода.

Если равенство (7.4) выполняется для текущего момента времени, то оно выполняется и для предыдущего момента времени, т.е.

$$Y_{t-1}^P = \lambda Y_{t-1} + (1 - \lambda) Y_{t-2}^P$$

Подставляя его в формулу (7.4), получаем

$$Y_t^P = \lambda Y_t + (1 - \lambda) Y_{t-1} + (1 - \lambda)^2 Y_{t-2}^P$$

Проделив аналогичную процедуру с заменой Y_{t-2}^P на Y_{t-2} и Y_{t-3}^P и продолжив этот процесс до бесконечности, мы можем переписать выражение (7.1) в виде

$$C_t = \beta_2 \lambda Y_t + \beta_2 \lambda (1 - \lambda) Y_{t-1} + \beta_2 \lambda (1 - \lambda)^2 Y_{t-2} + \dots \\ + \beta_2 \lambda (1 - \lambda)^{s-1} Y_{t-s+1} + \beta_2 (1 - \lambda)^s Y_{t-s+1}^P + \dots + C_t^T$$

Таким образом, воздействие текущего дохода на текущее потребление задается коэффициентом $\beta_2 \lambda$, а с каждым предыдущим периодом воздействие изменяется в $(1 - \lambda)$ раз.

Введя разумное предположение о том, что λ должна быть в интервале от 0 до 1, из (7.1) - (7.4) с помощью преобразования Койка можно получить (подробнее см. Доугерти, гл.11)[6].

$$C_t = \lambda \beta_2 Y_t + (1 - \lambda) C_{t-1} + u_t \quad (7.5)$$

где случайный член

$$u_t = C_t^T - (1 - \lambda) C_{t-1}^T \quad (7.6)$$

В соотношении (7.5) входят только наблюдаемые величины, и его параметры могут быть оценены.

Если усреднить соотношение (7.5) по времени (т.е. рассмотреть долгосрочное соотношение), то

$$\bar{C} = \beta_1 + \beta_2 \bar{Y} \quad (7.7)$$

Таким образом, краткосрочная предельная склонность к потреблению равна в этой модели произведению $\lambda \beta_2$, а долгосрочная - β_2 . Поскольку $\lambda < 1$, то краткосрочная предельная склонность к потреблению меньше долгосрочной, что соответствует эмпирическим наблюдениям в период после Второй мировой войны.

Для определения краткосрочной и долгосрочной динамики оценим уравнение (7.5) для логарифмов дохода и потребления.

Заметим, что в этом случае величина $\lambda\beta_2$ будет иметь смысл краткосрочной эластичности, а β_2 - долгосрочной эластичности потребления по доходу:

$$LC_t = \beta_1 + \lambda\beta_2LY_t + (1-\lambda)LC_{t-1} + u_t$$

(7.8)

В командной строке рабочего файла запишем

$$LS LC=C(1)+C(2)*C(3)*LY+(1-C(2))*LC(-1)$$

Здесь LS - команда для проведения оценки уравнения, $C(1)$ - оценка свободного коэффициента, который в общем случае может присутствовать в уравнении, $C(2)$ - оценка коэффициента λ , $C(3)$ - оценка коэффициента β_2 . Результат оценки приведен на рис. 7.1.

Как видно из полученной оценки, свободный коэффициент оказался статистически незначимым, скорость адаптации λ оценивается как 0.107, а долгосрочная эластичность потребления по доходу равна 0.98.

Заметим, что уравнение (7.8) нелинейно относительно параметров, и для его оценки был использован нелинейный метод наименьших квадратов.

| Equation: EQ03 Workfile: HAMILTON_COINT:Untitled\ - □ × | | | | | | | | | |
|---|------|--------|-------|-------------|-----------------------|-------------|----------|----------|--------|
| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
| Dependent Variable: LC | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 21/05/14 Time: 15:37 | | | | | | | | | |
| Sample (adjusted): 1947Q2 1989Q3 | | | | | | | | | |
| Included observations: 170 after adjustments | | | | | | | | | |
| Convergence achieved after 1 iteration | | | | | | | | | |
| LC=C(1)+C(2)*C(3)*LY+(1-C(2))*LC(-1) | | | | | | | | | |
| | | | | Coefficient | Std. Error | t-Statistic | Prob. | | |
| | | | | C(1) | 0.992834 | 1.032533 | 0.961552 | 0.3377 | |
| | | | | C(2) | 0.107015 | 0.033304 | 3.213273 | 0.0016 | |
| | | | | C(3) | 0.984149 | 0.013234 | 74.36705 | 0.0000 | |
| R-squared | | | | 0.999659 | Mean dependent var | | | 720.5078 | |
| Adjusted R-squared | | | | 0.999655 | S.D. dependent var | | | 41.74069 | |
| S.E. of regression | | | | 0.775665 | Akaike info criterion | | | 2.347298 | |
| Sum squared resid | | | | 100.4767 | Schwarz criterion | | | 2.402636 | |
| Log likelihood | | | | -196.5204 | Hannan-Quinn criter. | | | 2.369754 | |
| F-statistic | | | | 244612.7 | Durbin-Watson stat | | | 1.834643 | |
| Prob(F-statistic) | | | | 0.000000 | | | | | |

Рис. 7.1. Оценка модели адаптивных ожиданий

Если воспользоваться другой формой записи уравнения $LS\ LC\ C\ LY\ LC(-1)$

то будет применен обычный МНК для определения коэффициентов при факторах (рис. 7.2).

Коэффициент при факторе LY равен 0.105, имеет смысл оценки краткосрочной эластичности $\lambda\beta_2$, коэффициент при лагированном значении зависимой переменной равен 0.893 и дает оценку величины $(1 - \lambda)$. Из полученных значений можно найти оценку β_2 :

$$\hat{\beta}_2 = \frac{0.105}{1 - 0.893} = 0.98$$

Такой способ расчета не позволяет, однако, проверить статистическую значимость оценки коэффициента β_2 , так как в явном виде не рассчитывается стандартная ошибка коэффициента.

Заметим, что применять МНК для оценки параметров некорректно из-за эндогенности регрессоров (нарушается четвертое условие Гаусса-Маркова).

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|----------|-------------|------------|-------------|--------|
| C | 0.992834 | 1.032533 | 0.961552 | 0.3377 |
| LY | 0.105319 | 0.032913 | 3.199890 | 0.0016 |
| LC(-1) | 0.892985 | 0.033304 | 26.81309 | 0.0000 |

| | | | |
|--------------------|-----------|-----------------------|----------|
| R-squared | 0.999659 | Mean dependent var | 720.5078 |
| Adjusted R-squared | 0.999655 | S.D. dependent var | 41.74069 |
| S.E. of regression | 0.775665 | Akaike info criterion | 2.347298 |
| Sum squared resid | 100.4767 | Schwarz criterion | 2.402636 |
| Log likelihood | -196.5204 | Hannan-Quinn criter. | 2.369754 |
| F-statistic | 244612.7 | Durbin-Watson stat | 1.834643 |
| Prob(F-statistic) | 0.000000 | | |

DB = none WF = hamilton_coint

Рис. 7.2. Альтернативная оценка модели адаптивных ожиданий

2. На основании модели адаптивных ожиданий получены оценки текущего воздействия дохода на потребление $\lambda\beta_2^{\wedge} = 0.105$. Воздействие предыдущего периода оценивается в $\lambda\beta_2(1 - \lambda)^{\wedge} = 0.105 \cdot 0.893 = 0.094$

За m периодов эффект воздействия будет равен (Берндт, гл.8)

[3].

$$\lambda\beta_2 \frac{1 - (1 - \lambda)^m}{\lambda} = \beta_2(1 - (1 - \lambda)^m)$$

При стремлении m к бесконечности общее кумулятивное воздействие будет равно β_2 .

Доля общего кумулятивного воздействия после m тактов времени, обозначенная как d , равна $(1 - (1 - \lambda)^m)$. Таким образом, для вычисления длительности периода времени, требующегося чтобы получить долю d в общей реакции потребления на доход, необходимо решить уравнение

$$d = (1 - (1 - \lambda)^m) \text{ относительно } m, \text{ т.е.}$$

$$m = \ln(1 - d) / \ln(1 - \lambda)$$

Таким образом, 90% интервал продолжительности воздействия дохода на потребление вычисляется по формуле

$$m = \ln(1 - 0.9) / \ln(0.893) = 20.3 \text{ кварталов.}$$

3. Предположим, что желаемое потребление C^* описывается моделью

$$C_t^* = \gamma_1 + \gamma_2 Y_t + \varepsilon_t, \quad (7.9)$$

т.е. доход определяет желаемое, а не фактическое потребление (модель Брауна). Приращение фактического потребления пропорционально разнице между желаемым и предыдущим значением фактического потребления:

$$C_t - C_{t-1} = \mu(C_t^* - C_{t-1}) \quad (7.10)$$

Здесь $0 < \mu < 1$ - скорость корректировки. Выражая C^* из (7.10) и подставляя его в (7.9), после преобразований получаем

$$\begin{aligned} C_t &= \mu(\gamma_1 + \gamma_2 Y_t + \varepsilon_t) + (1 - \mu)C_{t-1} \\ &= \gamma_1 \mu + \gamma_2 \mu Y_t + (1 - \mu)C_{t-1} + \mu \varepsilon_t \\ &= \beta_1 + \beta_2 Y_t + \beta_3 C_{t-1} + \mu \varepsilon_t \end{aligned} \quad (7.11)$$

Таким образом, модель, которая получила название модели частичной корректировки, связывает текущее значение фактического потребления с текущим значением дохода и лагированным значением потребления и имеет ту же структуру, что и модель адаптивных ожиданий после преобразования Койка.

4. Оцененные модели (и адаптивных ожиданий и частичной корректировки) содержат в правой части лаговую зависимую переменную. В этом случае для проверки автокорреляции в остатках нельзя использовать тест Дарбина-Уотсона. Значение тестовой статистики смещено в сторону 2, увеличивая тем самым вероятность ошибки II рода (принять неверную

гипотезу). В этом случае применяется асимптотический тест Дарбина, основанный на статистике:

$$h = \hat{\rho} \sqrt{\frac{n}{1 - ns_{by(t-1)}^2}} \quad (7.12)$$

Здесь $\hat{\rho}$ - оценка коэффициента автокорреляции для процесса $AR(1)$; n - число наблюдений при оценке уравнения регрессии, $s_{by(t-1)}^2$ - оценка дисперсии коэффициента при лаговой зависимой переменной. Для оценки коэффициента автокорреляции для больших выборок можно применить статистику Дарбина-Уотсона DW , принимая во внимание что

$$DW \rightarrow 2 - 2\rho$$

Вспользуемся результатом оценки уравнения, приведенном на рис. 7.2. Статистика $DW=1.83$, число наблюдений, включенных в оценку уравнения $n=170$, оценка дисперсии коэффициента $s_{by(t-1)}^2 = 0.033^2 = 0.001$. Подставляя полученные значения в соотношение (7.12), имеем

$$h = (2 - 1.83) / 2 \sqrt{\frac{170}{1 - 170 \cdot 0.001}} = 1.216$$

На больших выборках при выполнении нулевой гипотезы об отсутствии автокорреляции h -статистика распределена нормально, т.е. при $h < 1.96$ нулевая гипотеза не отвергается на 5% уровне (для 1% уровня критическое значение равно 2.58). Таким образом, в нашем примере гипотеза об отсутствии автокорреляции в остатках не отвергается.

Заметим, что в случае модели адаптивных ожиданий случайный член подвержен автокорреляции (как это следует из (7.6)) в результате преобразования Койка. Однако если в изначальном уравнении модели адаптивных ожиданий случайный член подвержен автокорреляции, то после преобразования Койка она может быть нейтрализована.

В случае модели частичной корректировки автокорреляция отсутствует, если она отсутствовала в изначальном уравнении (7.9). Однако если случайный член уравнения для ожидаемых зна-

чений потребления подвержен автокорреляции, то она сохранится и после применения механизма частичной корректировки.

Таким образом, на основании проверки остатков на автокорреляцию нельзя сделать однозначный вывод о том, какая из двух моделей более соответствует эмпирическим данным.

Ключевые понятия

Динамические модели

Модель адаптивных ожиданий

Модель частичной корректировки

Скорость корректировки

Долгосрочная и краткосрочная эластичность

Краткосрочная и долгосрочная предельная склонность к потреблению

Преобразование Койка

Лаговая переменная

Тест Дарбина

Вопросы

1. Запишите оцененное уравнение модели адаптивных ожиданий, используя результаты оценки, приведенные на рис. 7.1 и 7.2.
2. Как можно проинтерпретировать коэффициент при лаговой зависимой переменной с позиций модели адаптивных ожиданий и с позиций модели частичной корректировки? За какой временной период проводится корректировка в данной задаче?
3. Какие ограничения есть на применение тестовой статистики Дарбина?
4. В чем суть преобразования Койка?
5. Какие свойства имеют МНК оценки модели адаптивных ожиданий после преобразования Койка?

Глава 8. Коинтеграция и модель коррекции ошибок для оценки потребления

Цель работы: изучить подходы к моделированию динамики взаимосвязи нестационарных временных рядов, использовать освоенные методы для получения оценок долгосрочной и краткосрочной динамики потребления

Исходные данные (Hamilton_coint.wf1):

LY – натуральный логарифм агрегированного располагаемого дохода (увеличенный в 100 раз)

LC – натуральный логарифм расходов на личное потребление (увеличенный в 100 раз).

Данные по США, квартальные, период с 1 квартала 1947 г. по 3 квартал 1989 г.

Источник данных: Hamilton, J. (1994) Time series analysis, Princeton: Princeton University Press, цитируется по Eviews 7 User Guide.

Задания

1. Изучите динамику рядов *LY* и *LC*, определите порядок интегрируемости.
2. Опираясь на теорию Фридмана о постоянном доходе, проверьте гипотезу о том, что в долгосрочном периоде потребление и доход будут расти одинаковым темпом, т.е. ряды коинтегрированы.
3. Оцените коинтегрирующее соотношение подходящим методом. Проведите тест Уальда для проверки гипотезы о пропорциональности постоянного потребления постоянному доходу.
4. Оцените параметры модели коррекции ошибок, опираясь на найденное коинтегрирующее соотношение.

Решения

1. Постройте графики рядов LY и LC и их разности ($LY-LC$). Для этого выделите ряды LY и LC , откройте их как группу и в меню **View** → **Group Members** допишите $LY-LC$ как новый член группы. Сохраните изменения (**UpdateGroup**, сохраните как **GROUP01**). Затем постройте график **View** → **Graph**. На вкладке **Graph options** откройте список **Axes & Scaling** (оси и масштабирование) и установите для третьего ряда расположение по правой оси (рис. 8.1).

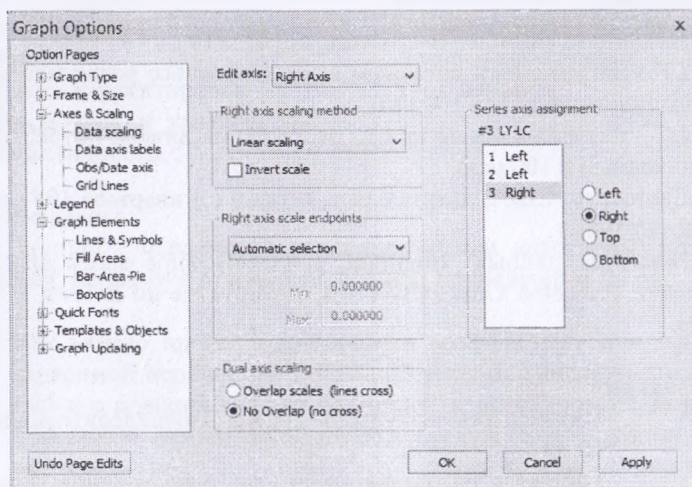


Рис. 8.1. Установка параметров графика

Из графиков видно (рис. 8.2), что ряды LY и LC нестационарны (имеют восходящий тренд).

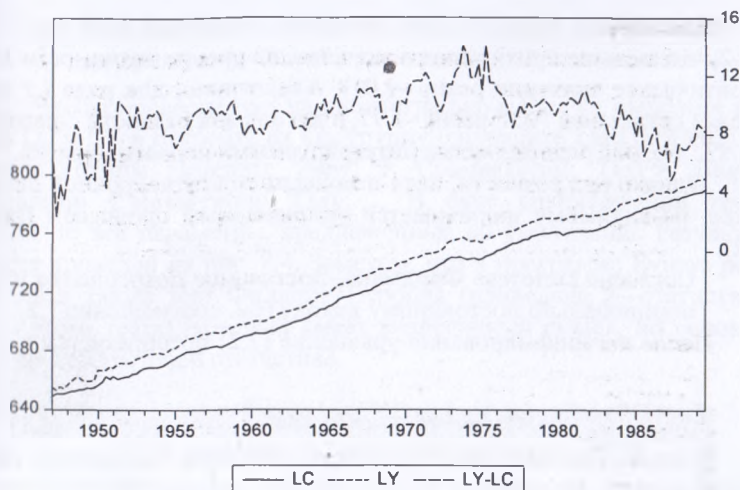


Рис. 8.2. Графики логарифмов расходов, доходов и их разности

Действительно, расширенный тест Дики–Фуллера с включением тренда показывает, что для ряда *LC* тестовая статистика ADF равна -2.39 , т.е. гипотеза о наличии единичного корня не отвергается даже на 10% уровне значимости (критическое значение равно -3.14 , см. рис. 8.3).

| Null Hypothesis: LC has a unit root | | |
|---|-------------|--------|
| Exogenous: Constant, Linear Trend | | |
| Lag Length: 2 (Automatic - based on SIC, maxlag=13) | | |
| | t-Statistic | Prob.* |
| Augmented Dickey-Fuller test statistic | -2.391784 | 0.3825 |
| Test critical values: 1% level | -4.013608 | |
| 5% level | -3.436795 | |
| 10% level | -3.142546 | |
| *MacKinnon (1996) one-sided p-values. | | |

Рис. 8.3. Результаты ADF—теста для ряда *LC*

Для первой разности ряда тестовая статистика ADF равна -7.12, что меньше критического значения на уровне значимости 1% (критическое значение равно -4.01). Аналогично, для ряда LY тестовая статистика ADF равна -1.77, а для первой разности равна -13.97, т.е. оба ряда являются интегрируемыми первого порядка.

Однако ряд разности, хотя и подвержен краткосрочной динамике, по-видимому, порождается стационарным процессом (рис. 8.2).

2. Согласно гипотезе Фридмана, постоянное потребление пропорционально постоянному доходу (см. соотношения 7.2).

После логарифмирования уравнение (7.2) преобразуется к виду

$$\text{Log}(C_t^p) = \text{Log}(\beta_2) + \text{Log}(Y_t^p) + \text{Log}(v_t) \quad (8.1)$$

Если не учитывать краткосрочную динамику и не делать различий между постоянным и фактическим доходом (постоянным/фактическим потреблением), то в долгосрочной перспективе доход и потребление будут расти с одинаковым темпом, а математическое ожидание разности двух рядов будет равно $\text{Log}(\beta_2)$ (Доугерти, гл.13)[6]. Если два или более нестационарных временных ряда связаны таким образом, то говорят, что они коинтегрированы. Соотношение (8.1) называют в этом случае коинтегрирующим соотношением.

Для проверки коинтегрированности рядов $LC = \text{Log}(C_t^p)$ и $LY = \text{Log}(Y_t^p)$ проведем тест Энгла-Грэнжера. Суть теста заключается в проверке на стационарность ряда остатков коинтегрирующего соотношения (8.1), используя расширенный тест Дики-Фуллера. Если ряд остатков стационарен, то исходные ряды коинтегрированы. В простейшем случае проверяется, порождаются ли остатки $\hat{u}_t = \text{Log}(\hat{v}_t)$ процессом случайного блуждания:

$$\Delta \hat{u}_t = (\rho - 1)\hat{u}_{t-1} + \varepsilon_t. \quad (8.2)$$

Нулевая гипотеза теста на коинтеграцию заключается в отсутствии коинтегрирующего соотношения (8.1), т.е. наличия единичного корня в ряде остатков ($\rho = 1$ в 8.2). Для проверки используется τ -статистика, основанная на обычной статистике Стьюдента (se - стандартная ошибка коэффициента)

$$\hat{\tau} = \frac{\hat{\rho} - 1}{se(\hat{\rho})}$$

Чтобы провести тест на коинтеграцию в Eviews, выделите ряды LC и LY и откройте их как группу (сохраните как GROUP02). В меню группы выберите View → Cointegration Test → Single Equation Cointegration Test ... В окне спецификации теста оставьте все параметры, предложенные по умолчанию. Результат теста приведен на рис. 8.4. Заметим, что в программе Eviews реализован также тест Филиппа–Перрона (приведена z-статистика для этого теста), который имеет аналогичный смысл, но основан на другой тестовой статистике.

| Engle-Granger Cointegration Test | | | | |
|---|---------------|-----------|-------------|--------|
| Date: 19/05/14 Time: 17:37 | | | | |
| Series: LC LY | | | | |
| Sample: 1947Q1 1989Q3 | | | | |
| Included observations: 171 | | | | |
| Null hypothesis: Series are not cointegrated | | | | |
| Cointegrating equation deterministics: C | | | | |
| Automatic lags specification based on Schwarz criterion (maxlag=13) | | | | |
| Dependent | tau-statistic | Prob * | z-statistic | Prob * |
| LC | -3.793029 | 0.0164 | -24.12538 | 0.0194 |
| LY | -3.849724 | 0.0139 | -24.47999 | 0.0179 |
| *Mackinnon (1996) p-values. | | | | |
| Intermediate Results: | | | | |
| | LC | LY | | |
| Rho - 1 | -0.179136 | -0.181619 | | |
| Rho S.E. | 0.047228 | 0.047177 | | |
| Residual variance | 1.053179 | 1.078710 | | |
| Long-run residual variance | 0.668823 | 0.686167 | | |
| Number of lags | 1 | 1 | | |
| Number of observations | 169 | 169 | | |
| Number of stochastic trends** | 2 | 2 | | |
| **Number of stochastic trends in asymptotic distribution | | | | |
| DB = none WF = hamilton_coint | | | | |

Рис. 8.4. Результат теста Энгла–Грэнжера на коинтеграцию

Как видно из результатов теста, гипотеза об отсутствии коинтеграции отвергается, по крайней мере, на уровне 2% (p-значение

равно 0.016, что меньше 0.02). Во второй строке оценивается обратное коинтегрирующее соотношение, аналогичное (8.1), где LY - зависимая переменная, а LC - объясняющая.

3. Оценка параметров коинтегрирующего соотношения (8.1) обычным методом наименьших квадратов является суперсостоятельной (см. Вербик, гл. 9)[4], так как она сходится к теоретическому значению с намного более высокой скоростью, чем в обычной асимптотике. Это порождает определенные проблемы при использовании обычного МНК, в частности, делает неприменимыми обычные статистические процедуры.

Для оценки коинтегрированных рядов разработано несколько специальных методов. В частности, полностью модифицированный метод наименьших квадратов (ПММК, Fully Modified OLS, FMOLS), который дает асимптотически несмещенные оценки и «хорошие» статистические свойства в асимптотике.

Для оценки коинтегрирующего соотношения методом FMOLS, в окне оценки уравнения **Quick** → **Estimate Equation** выберите метод оценки **Cointegrating Regression** В открывшемся окне спецификации уравнения в поле **Equation specification** введите через пробел зависимую переменную LC и коинтегрирующий регрессор LY . Во второй части определения параметров спецификации (**Cointegrating regressors specification**) ничего не добавляем, так как в нашем уравнении нет детерминированных регрессоров. Третья часть определения спецификации конкретизирует метод оценки уравнения (по умолчанию FMOLS) и оценки ковариационной матрицы (в долгосрочном периоде). Нажмите на выбор опций (**Options**), выберите метод **Andrews Automatic** (рис.8.5), который при оценке ковариационной матрицы не применяет коррекцию числа степеней свободы в отличие от метода **Newey-West**, который используется программой по умолчанию.

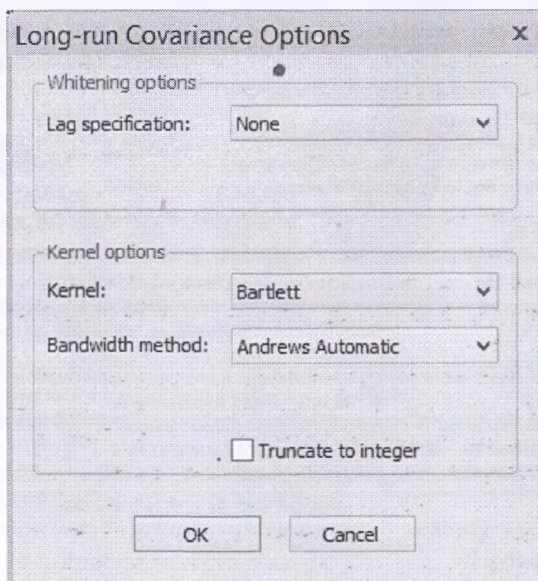


Рис. 8.5. Выбор параметров оценки ковариационной матрицы

Все остальные параметры спецификации оставьте по умолчанию. Результат оценки сравните с приведенным на рис. 8.6.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|---|-------------|--------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: LC | | | | | | | | | |
| Method: Fully Modified Least Squares (FMOLS) | | | | | | | | | |
| Date: 18/05/14 Time: 20:06 | | | | | | | | | |
| Sample (adjusted): 1947Q2 1989Q3 | | | | | | | | | |
| Included observations: 170 after adjustments | | | | | | | | | |
| Cointegrating equation deterministics: C | | | | | | | | | |
| Long-run covariance estimate (Bartlett kernel, Andrews bandwidth = 14.9878) | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| LY | 0.987548 | 0.009242 | 106.8539 | 0.0000 | | | | | |
| C | -0.035023 | 6.755216 | -0.005185 | 0.9959 | | | | | |
| R-squared | 0.998186 | Mean dependent var | 720.5078 | | | | | | |
| Adjusted R-squared | 0.998175 | S.D. dependent var | 41.74069 | | | | | | |
| S.E. of regression | 1.783125 | Sum squared resid | 534.1618 | | | | | | |
| Durbin-Watson stat | 0.409630 | Long-run variance | 25.76970 | | | | | | |

Рис. 8.6. Оценка коинтегрирующего соотношения по ПММНК

Таким образом, ряды LC и LY имеют коинтегрирующий вектор $(1, -0.99)^T$.

Согласно модели Фридмана, коэффициент при факторе LY равен единице (см. 8.1). Проверка этой гипотезы проводится с помощью обычного теста Уальда на линейное ограничение. В окне уравнения выберите **View** → **Coefficient Diagnostics** → **Wald Test - Coefficient Restrictions...** В окне задания ограничения введите $C(1)=1$. Результат теста (p -значения для t -статистики, F -статистики и χ^2 -статистики, рис. 8.7) говорит о том, что мы не можем отвергнуть нулевую гипотезу о равенстве единице коэффициента при факторе доход.

| | | | | |
|----|--|-----------|-----------|-------------|
| 1 | Wald Test: | | | |
| 2 | Equation: EQ_FMOLS | | | |
| 3 | | | | |
| 4 | Test Statistic | Value | df | Probability |
| 5 | | | | |
| 6 | t-statistic | -1.347366 | 168 | 0.1797 |
| 7 | F-statistic | 1.815394 | (1, 168) | 0.1797 |
| 8 | Chi-square | 1.815394 | 1 | 0.1779 |
| 9 | | | | |
| 10 | | | | |
| 11 | Null Hypothesis: C(1)=1 | | | |
| 12 | Null Hypothesis Summary: | | | |
| 13 | | | | |
| 14 | Normalized Restriction (= 0) | Value | Std. Err. | |
| 15 | | | | |
| 16 | -1 + C(1) | | -0.012452 | 0.009242 |
| 17 | | | | |
| 18 | Restrictions are linear in coefficients. | | | |

Рис. 8.7. Результат теста Уальда

Кointегрирующее соотношение описывает долговременные зависимости между двумя переменными. Оно не описывает краткосрочную динамику, но предполагает, что есть механизм, отвечающий за поддержание этой долгосрочной зависимости.

4. Рассмотрим модель коррекции ошибок, которая соединяет краткосрочную и долгосрочную динамику. Модель коррекции ошибок является разновидностью модели частичной корректировки. Как и в модели частичной корректировки, предполагается долгосрочная связь между X и Y :

$$Y_t^* = \gamma_1 + \gamma_2 X_t + u_t \quad (8.3)$$

где Y^* - уровень Y , который будет соответствовать уровню фактора X в долгосрочном соотношении.

В краткосрочном периоде изменения зависимой переменной, т.е. $\Delta Y = Y_t - Y_{t-1}$ определяются двумя компонентами:

частичным устранением различия между прошлым целевым и фактическим значениями $Y_{t-1}^* - Y_{t-1}$ и немедленной реакцией на скорость изменения фактора X

$$\Delta Y_t = \lambda(y_{t-1}^* - Y_{t-1}) + \delta \Delta X_t + u_t$$

(8.4)

Подставляя (8.3) в (8.4) и перегруппируя члены, получим

$$\begin{aligned} \Delta Y_t &= \lambda(\gamma_1 + \gamma_2 X_{t-1} - Y_{t-1}) + \delta(X_t - X_{t-1}) + u_t \\ &= \lambda\gamma_1 + \delta X_t + (\lambda\gamma_2 - \delta)X_{t-1} - \lambda Y_{t-1} + u_t \end{aligned}$$

В итоге уравнение имеет вид модели ADL(1,1) (модели с распределенными лагами, где числа указывают на порядок лага зависимой и независимой переменной в уравнении):

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 Y_{t-1} + \beta_4 X_{t-1} + u_t$$

(8.5)

Где

$$\begin{aligned} \beta_1 &= \lambda\gamma_1 & \beta_2 &= \delta \\ \beta_3 &= 1 - \lambda & \beta_4 &= \lambda\gamma_2 - \delta \end{aligned}$$

Тогда $\delta = \beta_2$ - краткосрочный эффект воздействия X на Y , $\lambda = 1 - \beta_3$ - скорость корректировки несоответствия между целевым и фактическим значениями, $\gamma_2 = (\beta_4 + \delta) / \lambda$ - долгосрочный эффект воздействия X на Y .

Оценка уравнения (8.5) непосредственно затрудняет получение экономически интерпретируемых коэффициентов. Кроме того, если исходные ряды нестационарны, то применение обычного МНК дает несостоятельные оценки. Однако если ряды имеют порядок интегрируемости 1, то их разности - стационарны.

Принимая во внимание, что в состоянии равновесия

$$\bar{Y} = \beta_1 + \beta_2 \bar{X} + \beta_3 \bar{Y} + \beta_4 \bar{X},$$

а, следовательно,

$$\bar{Y} = \frac{\beta_1}{1 - \beta_3} + \frac{\beta_2 + \beta_4}{1 - \beta_3} \bar{X}$$

и коинтегрирующее соотношение (т.е. соотношение в долгосрочном периоде) можно записать в виде

$$Y_t = \frac{\beta_1}{1 - \beta_3} + \frac{\beta_2 + \beta_4}{1 - \beta_3} X_t \quad (8.6)$$

Как видно из введенных обозначений, первый коэффициент в точности равняется γ_1 , а второй - γ_2 .

Соотношение (8.5) преобразуют, вычитая из обеих частей Y_{t-1} и перегруппировывая члены в правой части уравнения. В итоге получаем

$$\Delta Y_t = \lambda(Y_{t-1} - \gamma_1 - \gamma_2 X_{t-1}) + \delta \Delta X_t + u_t \quad (8.7)$$

Уравнение (8.7) называется *моделью коррекции ошибок*. Первое слагаемое является случайным членом коинтегрирующего соотношения и отвечает за коррекцию отклонения Y от долгосрочной динамики, второе отражает краткосрочное воздействие изменения X на изменение Y .

Для оценки модели коррекции ошибок для нашей задачи примените двухшаговую процедуру Энга–Гренжера.

На первом шаге получите остатки коинтегрирующего соотношения (8.1). Для этого в меню уравнения EQ01 выберите **Proc** → **Make Residual Series...** Сохраните название по умолчанию RESID01.

На втором шаге оцените по МНК уравнение (8.7), введя в командной строке

```
LS D(lc) resid01(-1) D(ly)
```

Заметим, что в уравнение (8.7) отсутствует константа, поэтому ее нет в записи спецификации, символом D обозначена разность, т.е.

$$D(LC) = LC_t - LC_{t-1}$$

Результат оценки модели коррекции ошибок приведен на рис. 8.8. Коэффициент при остатках коинтегрирующего соотношения статистически значим и показывает, что за один временной период (в данном случае - квартал) корректируется 15.7% отклонений фактических от долгосрочных значений.

| View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids |
|--|-------------|-----------------------|-------------|--------|--------|----------|----------|-------|--------|
| Dependent Variable: D(LC) | | | | | | | | | |
| Method: Least Squares | | | | | | | | | |
| Date: 28/05/14 Time: 17:32 | | | | | | | | | |
| Sample (adjusted): 1947Q3 1989Q3 | | | | | | | | | |
| Included observations: 169 after adjustments | | | | | | | | | |
| Variable | Coefficient | Std. Error | t-Statistic | Prob. | | | | | |
| RESID01(-1) | -0.157084 | 0.038960 | -4.031898 | 0.0001 | | | | | |
| D(LY) | 0.541022 | 0.049790 | 10.86605 | 0.0000 | | | | | |
| R-squared | -0.207205 | Mean dependent var | 0.823048 | | | | | | |
| Adjusted R-squared | -0.214434 | S.D. dependent var | 0.794158 | | | | | | |
| S.E. of regression | 0.875173 | Akaike info criterion | 2.582974 | | | | | | |
| Sum squared resid | 127.9100 | Schwarz criterion | 2.620014 | | | | | | |
| Log likelihood | -216.2613 | Hannan-Quinn criter. | 2.598006 | | | | | | |
| Durbin-Watson stat | 1.950434 | | | | | | | | |

Рис. 8.8. Результат оценки модели коррекции ошибок.

Ключевые понятия

Коинтеграция

Тест Энгла-Гренжера

Модель коррекции ошибок

Коинтегрирующее соотношение

Суперсостоятельность

Модель распределенных лагов

Долгосрочная динамика

Нестационарность

Вопросы

1. Поясните, что такое коинтегрирующее соотношение.
2. Поясните, что означает термин «суперсостоятельность».
3. Запишите оцененную модель ADL(1,1) (уравнение (8.5)) исходя из полученных оценок модели коррекции ошибок.
4. Покажите, при каком ограничении модель частичной коррекции является частным случаем модели ADL(1,1).
5. В чем заключается двухшаговая процедура Энгла-Гренжера?

Использованная литература

1. Айвазян С.А. (2014). *Методы эконометрики*. М.: Инфра-М.
2. Бабешко Л. (2007). *Основы эконометрического моделирования*. М.: КомКнига.
3. Берндт Эрнст Р. (2005). *Практика эконометрики: классика и современность*. М. : Юнити.
4. Вербик М. (2008). *Путеводитель по современной эконометрике*. М.: Научная книга.
5. Доугерти К. (2009). *Введение в эконометрику*. М.: ИНФРА-М.
6. Елисеева И.И. (2007). *Практикум по эконометрике*. М.: Финансы и статистика.
7. Елисеева И.И. (2011) *Эконометрика*. М.: Проспект.
8. Магнус Я.Р., Катышев П.К., Пересецкий А.А. (2005). *Эконометрика. Начальный курс*. М.: Дело.
9. Матюшок В.М., Балашова С.А., Лазанюк И.В. (2015). *Основы эконометрического моделирования с использованием Eviews*. М.: изд-во РУДН.
10. Hamilton J.D. (1994) *Time series analysis*. Princeton University Press, Princeton.
11. Green W.H. (2008). *Econometric Analysis*, 6th Edition. Prentice-Hall, Upper Saddle River, New Jersey.
12. Stock J., Watson M.W. (2010). *Introduction to Econometrics*, 3rd Edition. Addison-Wesley Series in Economics, Addison-Wesley.

Базы данных

www.gks.ru

www.cbr.ru

<http://censtats.census.gov/>

<http://data.worldbank.org/>

<http://pwt.econ.upenn.edu/>

ПРИЛОЖЕНИЯ

Список используемых команд Eviews 7.0

Основные команды в Eviews (приводится один из возможных вариантов выполнения команды)

1. Построение новой переменной **Quick → Generate series** (задать уравнение для переменной)
2. Построение графика переменной (выделить переменную) **View → Graph → Line → Name**
3. Создание группы переменных (выделить переменные) **Open → as Group → Name**
4. Построение диаграммы рассеяния (выделить сначала переменную по оси X, затем по оси Y, открыть как группу) **View → Graph → Scatter → Freeze → Name**
5. Получение описательной статистики (выделить переменные, открыть как группу) **View → Descriptive Statistics → Freeze → Name**
6. Расчет коэффициентов корреляции (выделить переменные, открыть как группу) **View → Covariance Analysis → Correlation → Freeze → Name**
7. Оценка уравнения **Quick → Estimate Equation** (записать через пробел зависимую переменную, независимые переменные и константу) **→ Name**
8. Получение расчетных значений и графика доверительного интервала для расчетных значений (открыть уравнение) **Forecast → Freeze → Name**
9. Получение графика фактических, расчетных значений и остатка (открыть уравнение) **Resids → Freeze → Name**
10. Расширение диапазона рабочего файла **Range →** (ввести нужное количество наблюдений)
11. Изменение выборки **Sample →** (ввести через пробел нужный диапазон наблюдений)

12. Задание прогнозных значений факторов (выделить факторы, открыть как группу) **Edit+/-** → ввести прогнозные значения
13. Расчет прогноза и стандартных ошибок прогноза по уравнению (открыть уравнение) **Forecast** (задать имя для прогнозной переменной и для стандартной ошибки и диапазон для прогноза) → **Freeze** → **Name**
14. Проведение специальных тестов на коэффициенты и остатки уравнения (открыть уравнение)
 - a. Тест Уальда на ограничения для коэффициентов **View** → **Coefficient Diagnostics** → **Wald Test -Coefficient Restriction** (ввести ограничения)
 - b. Тест на пропущенную переменную **View** → **Coefficient Diagnostics** → **Omitted Variable** (ввести переменную)
 - c. Тест на гетероскедастичность по Уайту **View** → **Residual Tests** → **White Heteroskedasticity Test**
15. Коррекция стандартных ошибок для коэффициентов уравнения (открыть уравнение, команда **Estimate**, вкладка **Options**)
 - a. В случае гетероскедастичности **Heteroskedasticity Consistent Covariance** → **White**
 - b. В случае автокорреляции **Heteroskedasticity Consistent Covariance** → **Newey-West**
16. Использование обобщенного МНК (в окне уравнения команда **Estimate**)
 - a. В случае гетероскедастичности - вкладка **Options** → **Wieghted LS/TLS** ввести вес
 - b. В случае автокорреляции в уравнение добавить член **AR(1)**

Перевод некоторых английских эконометрических терминов¹⁵

analysis анализ

~ **of covariance** ковариационный анализ

~ **of variance** дисперсионный анализ

autocorrelation автокорреляция, сериальная корреляция

partial ~ частная автокорреляция

residual ~ автокорреляция остатков

autocorrelogram автокоррелограмма

autoregression авторегрессия ~ **of n -th order** авторегрессия

n -го порядка

order of ~ порядок авторегрессии

vector ~ векторная авторегрессия

average среднее || средний (См. тж. **mean**)

autoregressive moving ~ авторегрессионное скользящее

среднее

moving ~ скользящее среднее

simple ~ простое среднее (\approx среднее арифметическое, невзвешенное)

weighted ~ взвешенное среднее

bias смещение

downward ~ смещение вниз

upward ~ смещение вверх

biased смещенный

causal причинный

causality причинность

Granger ~ причинность по Грейнджеру

chi-square хи-квадрат {**distribution, test**}

coefficient коэффициент

~ **of determination** коэффициент детерминации

¹⁵ <http://www.nsu.ru>

- ~ of determination adjusted for degrees of freedom коэффициент детерминации, скорректированный на количество степеней свободы
 ~ of variation коэффициент вариации
 adjusted ~ of determination скорректированный коэффициент детерминации
 correlation ~ коэффициент корреляции
 multiple correlation ~ коэффициент множественной корреляции
 regression ~ коэффициент регрессии
 sample correlation ~ выборочный коэффициент корреляции
 skewness ~ коэффициент асимметрии
 slope ~ угловой коэффициент, коэффициент наклона
 Spearman's rank correlation ~ коэффициент ранговой корреляции Спирмена
 cointegrated коинтегрированный
 cointegrating коинтегрирующий
 cointegration коинтеграция
 collection
 data ~ сбор данных
 collinearity коллинеарность
 combination комбинация, сочетание
 cointegrating ~ коинтегрирующая комбинация
 conditional условный
 confidence доверие || доверительный
 ~ interval - доверительный интервал
 consistency состоятельность
 consistent состоятельный
 constant константа (См. *тж.* **intercept**), постоянная || постоянный
 correction поправка, исправление, корректирование
 autocorrelation ~ поправка на автокорреляцию
 bias ~ поправка на смещение

- degrees of freedom** ~ поправка на количество степеней свободы
- error** ~ исправление ошибки
- correlated** коррелированный
- serially** ~ сериально коррелированный
- correlation** корреляция
- multiple** ~ множественная корреляция, множественный коэффициент корреляции
- rank** ~ ранговая корреляция
- simple** ~ простая корреляция
- serial** ~ сериальная корреляция
- spurious** ~ ложная корреляция
- correlogram** коррелограмма
- partial** ~ частная коррелограмма, график частной автокорреляционной функции
- residual** ~ коррелограмма остатков
- covariance** ковариация
- criterion** критерий
- Akaike information** ~ информационный критерий Акаике
- choice** ~ критерий выбора
- information** ~ информационный критерий
- fitting** ~ критерий подгонки
- Schwartz information** ~ информационный критерий Шварца
- cross** ~ перекрестный, взаимный
- cross-correlation** кросс-корреляция, перекрестная корреляция
- data** данные
- count** ~ счетные, дискретные данные
- differenced** ~ разности исходных данных
- grouped** ~ сгруппированные данные
- input** ~ входные данные, исходные данные
- integrated** ~ интегрированные данные
- missing** ~ пропущенные, отсутствующие данные

- noisy** ~ зашумленные данные (измеренные с большими ошибками)
ordered ~ упорядоченные данные
original ~ исходные данные
panel ~ панельные
raw ~ необработанные данные, «сырые» данные
ranked ~ упорядоченные данные
sample ~ выборочные данные
data-based основывающийся на данных
dating датировка
degree степень, порядок, градус
 ~ of accuracy степень точности
 ~ of belief степень уверенности, степень доверия
 ~ of dependence степень зависимости
 ~s of freedom степени свободы, количество степеней свободы
dependence зависимость
linear ~ линейная зависимость
deviation отклонение
absolute ~ абсолютное отклонение
standard ~ стандартное отклонение, среднеквадратическое отклонение
distribution распределение
asymptotic ~ асимптотическое распределение
chi-square(d) ~ распределение хи-квадрат
conditional ~ условное распределение
continuous ~ непрерывное распределение
empiric(al) ~ эмпирическое распределение
exponential ~ экспоненциальное распределение
F ~ F-распределение
Fisher ~ распределение Фишера
lognormal ~ логнормальное распределение
multivariate ~ многомерное распределение
non-central ~ нецентральное распределение

normal ~ нормальное распределение
Pearson ~ распределение Пирсона
Poisson ~ распределение Пуассона
sample ~ выборочное распределение
small-sample ~ распределение в малых выборках
Student ~ распределение Стьюдента
symmetric ~ симметричное распределение
t ~ t-распределение
unconditional ~ безусловное распределение
uniform ~ равномерное распределение
disturbance возмущение (\approx **error**)
random ~ случайное возмущение
efficiency
asymptotic ~ асимптотическая эффективность
endogeneity эндогенность
endogenous эндогенный, внутрисистемный
equation уравнение
recurrence ~s рекуррентные уравнения
regression ~ уравнение регрессии
simultaneous ~s система одновременных уравнений
error ошибка ~ **due to ...** ошибка, обусловленная ...
~s in variables ошибки в переменных
autoregressive ~s авторегрессионные ошибки
asymptotic standard ~ асимптотическая стандартная
ошибка
coefficient standard ~ стандартная ошибка коэффициента
forecast ~ ошибка прогноза
random ~ случайная ошибка
rounding ~ ошибка округления
sample ~ выборочная ошибка
specification ~ ошибка спецификации
systematic ~ систематическая ошибка
type I ~ ошибка первого рода
type II ~ ошибка второго рода

estimate оценка || оценивать **to estimate recursively** оценивать рекуррентно

à priori ~ априорная оценка

approximate ~ приближенная оценка

biased ~ смещенная оценка

maximum likelihood ~ оценка максимального правдоподобия

non-parametric ~ непараметрическая оценка

point ~ точечная оценка

regression ~ регрессионная оценка

restricted ~ ограниченная оценка, оценка с ограничениями

sample ~ выборочная оценка

unbiased ~ несмещенная оценка

estimator способ оценивания, метод оценивания, оценка

Aitken ~ метод оценивания Эйткена (\equiv **generalized least squares** ~)

best ~ наилучшая оценка

best linear unbiased ~ (BLUE) наилучшая линейная несмещенная оценка

consistent ~ состоятельная оценка

efficient ~ эффективная оценка

generalized least squares ~ оценка обобщенного метода наименьших квадратов

inconsistent ~ несостоятельная оценка

instrumental variable ~ оценка метода инструментальных переменных

linear ~ линейная оценка

minimum-variance ~ оценка с минимальной дисперсией

minimum-variance unbiased ~ несмещенная оценка с минимальной дисперсией

Newey-West [autocorrelation consistent] ~ [состоятельная при автокорреляции] оценка Ньюи-Уэста

- single equation** ~ метод оценивания одного уравнения (из системы одновременных уравнений)
- system** ~ системный метод оценивания (для всей системы одновременных уравнений в целом)
- unbiased** ~ несмещенная оценка
- exogeneity** экзогенность
- expectation** ожидание, математическое ожидание, первый момент ~s
- adaptive** ~s адаптивные ожидания
- conditional** ~ условное [математическое] ожидание
- data-based** ~s ожидания, основанные на данных
- backward-looking** ~s ретроспективные (построенные на прошлом опыте) ожидания
- rational** ~s рациональные ожидания
- static** ~s статические ожидания
- unconditional** ~ безусловное [математическое] ожидание
- unbiased** ~ несмещенные ожидания
- fit** подбор, соответствие, подобранная функция || подбирать, подгонять, соответствовать
- good** ~ точный подбор (предсказания модели хорошо соответствуют данным)
- poor** ~ плохой подбор (предсказания модели плохо соответствуют данным)
- fixed** постоянный, неизменный, фиксированный (\approx deterministic)
- ~ **in repeated samples** неизменный в повторных выборках
- forecast** прогноз, предсказание
- back** ~ обратный прогноз
- conditional** ~ условный прогноз
- dynamic** ~ динамический прогноз
- multi-step** ~ многошаговый прогноз
- one-step** ~ одношаговый прогноз
- static** ~ статический прогноз
- forecasting** прогнозирование

- form** вид, форма
- functional** ~ функциональная форма
 - quadratic** ~ квадратичная форма
 - reduced** ~ приведенная форма
 - restricted** ~ ограниченная форма
 - restricted reduced** ~ ограниченная приведенная форма
 - structural** ~ структурная форма
- function** функция
- autocorrelation** ~ автокорреляционная функция
 - beta** ~ бета-функция
 - characteristic** ~ характеристическая функция
 - cumulative distribution** ~ [кумулятивная, интегральная] функция распределения
 - demand** ~ функция спроса
 - density** ~ функция плотности распределения
 - distribution** ~ функция распределения
 - likelihood** ~ функция правдоподобия
 - loss** ~ функция потерь
 - partial autocorrelation** ~ частная автокорреляционная функция
 - production** ~ производственная функция
 - spline** ~ сплайн-функция
 - utility** ~ функция полезности
- general-to-specific** от общего к частному
- generator**
- random numbers** ~ датчик случайных чисел
- goodness**
- ~ **of fit** точность подбора, соответствие предсказаний модели данным, согласие
- goodness-of-fit** (\equiv **goodness of fit**)
- grid** сетка (на графике), решетка
- group** группа
- heteroscedasticity** гетероскедастичность

- autoregressive conditional** ~ авторегрессионная условная гетероскедастичность
- conditional** ~ условная гетероскедастичность
- multiplicative** ~ мультипликативная гетероскедастичность
- homoscedasticity** гомоскедастичность
- hypothesis** гипотеза, **to accept a** ~ принять гипотезу, **to reject a** ~ отвергнуть, отклонить гипотезу, **to test a** ~ проверить гипотезу
- alternative** ~ альтернативная гипотеза
- false** ~ неверная гипотеза, ложная гипотеза
- implicit** ~ неявная гипотеза, подразумеваемая гипотеза
- nested** ~ вложенная гипотеза
- non-nested** ~ невложенная гипотеза
- non-parametric** ~ непараметрическая гипотеза
- null** ~ нулевая гипотеза
- parametric** ~ параметрическая гипотеза
- simple** ~ простая гипотеза
- true** ~ верная гипотеза, истинная гипотеза
- identification** идентификация
- model** ~ идентификация модели
- identified** идентифицированный
- exactly** ~ точно идентифицированный
- poorly** ~ плохо идентифицированный
- identity** тождество
- ill-conditioned** плохо обусловленный
- inconsistent** несостоятельный
- information** информация
- a priori** ~ априорная информация
- extraneous** ~ внешняя информация
- prior** ~ априорная, предварительная информация
- sample** ~ выборочная информация
- insignificant** незначимый, **insignificant at ...% level** незначимый при уровне ...%
- instability** нестабильность, неустойчивость

- instrumental** инструментальный
- intercept** отрезок, отсекаемый на координатной оси; константа, свободный член (в уравнении регрессии)
- interdependence** взаимная зависимость
- internal** внутренний
- interpolation** интерполяция, интерполирование
- interval** интервал, отрезок, промежуток
- confidence** ~ доверительный интервал
- prediction** ~ интервал предсказаний
- probability** ~ интервал вероятности
- tolerance** ~ допустимый интервал, толерантный интервал
- J-test** J-критерий
- lag** лаг, запаздывание, задержка
- lagged** запаздывающий, с лагом (о переменной)
- lagging** запаздывание, отставание || запаздывающий
- large-sample** относящийся к большой выборке
- latent** скрытый, латентный
- law** закон
- ~ **of large numbers** закон больших чисел
- empirical** ~ эмпирический закон
- least-square(s)** наименьших квадратов, относящийся к методу наименьших квадратов
- least squares [method]** метод наименьших квадратов
- autoregressive** ~ авторегрессионный наименьших квадратов
- generalized** ~ обобщенный метод наименьших квадратов
- indirect** ~ косвенный метод наименьших квадратов
- non-linear** ~ нелинейный метод наименьших квадратов
- ordinary** ~ обычный метод наименьших квадратов
- recursive** ~ рекуррентный метод наименьших квадратов
- restricted** ~ метод наименьших квадратов с ограничениями
- level** уровень
- accuracy** ~ уровень точности

- confidence** ~ уровень доверия
- conventional significance** ~ общепринятый, традиционный
уровень значимости (обычно 5%)
- probability** ~ уровень вероятности
- risk** ~ уровень риска (\equiv **significance level**)
- significance** ~ уровень значимости
- likelihood** правдоподобие
- maximum** ~ максимальное правдоподобие, метод максимального правдоподобия
- loglinear** логлинейный (линейный в логарифмах)
- lognormal** логнормальный
- log-transform** логарифм переменной, логарифмическое преобразование
- long-run** долгосрочный
- loss** потеря, потери
~ **of efficiency** потеря эффективности
~ **of information** потеря информации
- minimum expected** ~ минимальные ожидаемые потери
- marginal** маргинальный (в статистике), предельный (в экономической теории)
- mean** среднее, среднее значение [математическое] ожидание
- population** ~ генеральное среднее (\equiv **expectation**)
- sample** ~ выборочное среднее
- unconditional** ~ \equiv **unconditional expectation**
- measure** мера {**unit of** ~ }
- of ~ fit** показатель точности подбора
Точность подбора обычно измеряют коэффициентом детерминации
- rank correlation** ~ мера ранговой корреляции
- method** метод, способ, прием
~ **of least squares** метод наименьших квадратов
~ **of maximum likelihood** метод максимального правдоподобия
~ **of moments** метод моментов

- Cochran–Orcutt** ~ метод Кохрейна–Оркатта
- estimation** ~ метод оценивания
- generalized ~ of moments** обобщенный метод моментов
- generalized instrumental variables** ~ обобщенный метод инструментальных переменных
- instrumental variables** ~ метод инструментальных переменных
- least squares** ~ метод наименьших квадратов
- maximum likelihood** ~ метод максимального правдоподобия
- principal components** ~ метод главных компонент
- misspecification** неправильная спецификация, ошибка спецификации
- functional form** ~ неправильная спецификация функциональной формы
- model** модель
- autoregression (autoregressive)** ~ авторегрессионная модель
- autoregressive (autoregression) moving average** ~ смешанная модель авторегрессии и скользящего среднего
- autoregressive distributed lag** ~ авторегрессионная модель с распределенным лагом
- binary choice** ~ модель бинарного выбора
- Box–Jenkins** ~ модель Бокса–Дженкинса
- discrete choice** ~ модель дискретного выбора
- dynamic** ~ динамическая модель
- fixed-effects** ~ модель с фиксированными эффектами
- general linear** ~ общая линейная модель
- limited dependent variable** ~ модель с ограниченной зависимой переменной
- linear** ~ линейная модель
- loglinear** ~ логлинейная (линейная в логарифмах) модель
- multiplicative** ~ мультипликативная модель
- moving average** ~ модель скользящего среднего

nested ~ вложенная модель
non-nested ~ невложенная модель
non-parametric ~ непараметрическая модель
partial adjustment ~ модель частичного приспособления
qualitative dependent variable ~ модель с качественной зависимой переменной
structural ~ структурная модель
Monte-Carlo относящийся к методу Монте-Карло
Метод Монте-Карло — любой метод, в котором используются случайные (псевдослучайные) числа
non-causality отсутствие причинной связи
non-linear нелинейный
 ~ **in variables** нелинейный в переменных
 intrinsically ~ внутренне нелинейный
normal нормальный, относящийся к нормальному распределению
 asymptotically normal ~ асимптотически нормальный
observation наблюдение
 artificial ~ искусственное наблюдение
 outlying ~ резко отклоняющееся наблюдение, выброс
observed наблюдаемый
omitted пропущенный
one-sided односторонний
one-step одношаговый
P-value ≡ **probability value**
period период, временной интервал, срок
 base ~ базисный период
 estimation ~ период оценивания
 sample ~ период выборки
plot график
 normal ~ нормальный график
point точка
 data ~ (отдельное) наблюдение
 percentage ~ процентная точка

population генеральная совокупность || относящийся к генеральной совокупности, генеральный, теоретический

power мощность (критерия), сила, степень

asymptotic ~ асимптотическая мощность

explanatory ~ объяснительная сила

forecasting ~ предсказательная сила, точность предсказаний

test ~ мощность критерия

prediction предсказание,

biased ~ смещенный прогноз

probability вероятность

conditional ~ условная вероятность

procedure процедура, метод действия

Hildreth-Lu ~ метод Хилдрета-Лу

proxy [variable] заместитель, эрзац-переменная (не очень качественный заместитель переменной, которая должна стоять в уравнении)

Q-statistic Q-статистика

qualitative качественный

quantile квантиль

quantitative количественный

rank ранг

~ **of a matrix** ранг матрицы

cointegration ~ ранг коинтеграции

full ~ полный ранг

ratio отношение, пропорция, коэффициент

likelihood ~ отношение правдоподобия

variance ~ дисперсионное отношение

recursive рекуррентный, рекурсивный

reduce сокращать, уменьшать, превращать, сводить, преобразовывать

reduced приведенный, сокращенный

redundant избыточный, излишний

regression регрессия, ~ **of Y on X** регрессия Y по/на X

- artificial** ~ искусственная регрессия
- auxiliary** ~ вспомогательная регрессия
- censored** ~ цензурированная регрессия
- dynamic** ~ динамическая регрессия
- logistic** ~ логистическая регрессия
- multiple** ~ множественная регрессия
- non-linear** ~ нелинейная регрессия
- non-parametric** ~ непараметрическая регрессия
- polynomial** ~ полиномиальная регрессия
- recursive** ~ рекуррентная регрессия
- simple** ~ простая регрессия
- spurious** ~ ложная регрессия
- static** ~ статическая регрессия
- regressor** регрессор, независимая переменная в регрессии
- reject** отвергать, отбрасывать {to ~ a hypothesis}
- relation** отношение, связь, зависимость
 - negative** ~ отрицательная связь
 - positive** ~ положительная связь
 - strong** ~ сильная связь
- relationship** отношение, соотношение, зависимость, связь
 - cointegrating** ~ коинтегрирующее отношение
 - functional** ~ функциональная связь
 - lagged** ~ зависимость с лагом
 - spurious** ~ ложная зависимость
- residual** остаток || остаточный
 - autocorrelated** ~s автокоррелированные остатки
 - generalized** ~s обобщенные остатки
 - squared** ~ квадрат остатка
 - standardized** ~s нормированные остатки
 - studentized** ~s студентизированные остатки
- residual-based** на основе остатков, основанный на остатках
- restriction** ограничение, функция ограничения to impose a
 - ~ наложить ограничение {set of ~s}
 - ~ on parameters ограничение на параметры

- a priori** ~ априорное ограничение
cointegration ~s коинтеграционные ограничения
equality ~ ограничение в виде равенства
identifying ~s идентифицирующие ограничения
inequality ~ ограничение в виде неравенства
linear ~ линейное ограничение
non-linear ~ нелинейное ограничение
zero ~ ограничение, состоящее в том, что параметр равен нулю
- robust** робастный, устойчивый к отклонениям от принятых предположений
robustness робастность, устойчивость к отклонениям от принятых предположений
sample выборка || выборочный
- artificial** ~ искусственная выборка
balanced ~ уравновешенная выборка
biased ~ смещенная выборка, необъективная выборка
censored ~ цензурированная выборка
finite ~ конечная выборка
infinite ~ бесконечная выборка
large ~ большая выборка
moderate ~ выборка среднего размера
ordered ~ упорядоченная выборка
random ~ случайная выборка
representative ~ представительная выборка, репрезентативная выборка
small ~ малая выборка
truncated ~ усеченная выборка
- seasonal** сезонный
semi-logarithmic полулогарифмический
semi-parametric полупараметрический
series ряд
- integrated time** ~ интегрированный временной ряд
power ~ степенной ряд

stationary ~ стационарный ряд
time ~ временной ряд
wandering ~ блуждающий ряд
set множество, набор, система, серия
 ~ **of equations** система уравнений
 ~ **of restrictions** набор ограничений
data ~ набор данных
information ~ информационное множество
short-run краткосрочный
significance значимость, уровень значимости
 regression ~ значимость регрессии
 statistical ~ статистическая значимость
 test ~ значимость критерия, уровень значимости критерия
significant значимый **significant at ...% level** значимый на уровне ...%
simulation моделирование, имитация, имитационное моделирование, имитационный эксперимент
 Monte-Carlo ~ метод Монте-Карло
simultaneous одновременный, совместный
skew асимметричный, скошенный
skewness асимметрия
slope угловой коэффициент, наклон (о линейной функции)
small-sample относящийся к малой выборке
specification спецификация
 dynamic ~ динамическая спецификация
 model ~ спецификация модели, описание модели
 system ~ спецификация системы
 temporal ~ временная спецификация (\equiv **dynamic** ~)
 weighted least ~s [method] взвешенный метод наименьших квадратов
 mean ~ средний квадрат
stability стабильность, устойчивость
 parameter ~ стабильность параметров

- structural** ~ стабильность структуры, структурная стабильность
- stationarity** стационарность
- strict** ~ строгая стационарность
- weak** ~ слабая стационарность
- stationary** стационарный, установившийся
- difference** ~ стационарный в первых разностях, стационарный в приращениях {process}
- trend** ~ стационарный относительно тренда
- weakly** ~ слабо стационарный
- statistic** статистика, статистический показатель
- descriptive** ~ описательная статистика
- Durbin-Watson** ~ статистика Дарбина-Уотсона
- likelihood ratio** ~ статистика отношения правдоподобия
- non-parametric** ~ непараметрическая статистика
- order** ~ порядковая статистика
- test** ~ статистика критерия
- trace** ~ статистика следа
- system** система, система уравнений
- ~ of demand equations система уравнений спроса
- dynamic** ~ [of equations] динамическая система [уравнений]
- singular equation** ~ вырожденная система уравнений
- recursive** ~ рекурсивная система
- simultaneous** ~ система одновременных уравнений
- static** ~ статическая система
- stochastic** ~ стохастическая система
- t-distribution** t-распределение, распределение Стьюдента
- t-statistic** t-статистика, статистика Стьюдента
- tail** хвост (распределения)
- technique** метод, способ, техника
- estimation** ~ метод оценивания
- temporal** временной, временный

test критерий, тест || проверять **to ~ a hypothesis** проверить гипотезу

asymptotic ~ асимптотический критерий

augmented Dickey-Fuller ~ дополненный критерий Дики-Фуллера

autocorrelation ~ критерий автокорреляции [остатков]

Box-Pierce ~ критерий Бокса-Пирса

break-point ~ критерий перелома в тенденции

chi-square ~ критерий хи-квадрат

Chow ~ критерий Чоу

cointegration ~ коинтеграционный критерий, критерий коинтеграции

common factor ~ критерий общего множитель

conditional moment ~ критерий условного момента

consistent ~ состоятельный критерий

diagnostic ~ диагностический критерий

Dickey-Fuller ~ критерий Дики-Фуллера

differencing ~ критерий приращений

Durbin's h ~ h-критерий Дарбина

Durbin-Watson ~ критерий Дарбина-Уотсона

efficient ~ эффективный критерий

exact ~ точный критерий

Glejser ~ критерий Глейзера

Godfrey ~ критерий Годфрея

Goldfeld-Quandt ~ критерий Голдфельда-Квандта

goodness-of-fit ~ критерий согласия

Hausman ~ критерий Хаусмана

homogeneity ~ критерий однородности

information matrix ~ критерий информационной матрицы

invariant ~ инвариантный критерий

Jarque-Bera ~ критерий Жака-Бера

likelihood ratio ~ критерий отношения правдоподобия

misspecification ~ критерий ошибки спецификации

most powerful ~ наиболее мощный критерий

- non-linearity** ~ критерий нелинейности
non-nested ~ невложенный критерий
non-parametric ~ непараметрический критерий
normality ~ критерий нормальности
omitted variables ~ критерий пропущенных переменных
one-sided ~ односторонний критерий
one-tail ~ \equiv **one-sided** ~
outer product of gradient ~ критерий внешнего произведения градиента
overidentifying restrictions ~ критерий сверхидентифицирующих ограничений
parameter constancy ~ критерий постоянства параметров
parameter-free ~ свободный от параметра критерий
predictive failure ~ критерий [не]адекватности предсказаний
preliminary ~ предварительный критерий
rank ~ ранговый критерий
specification ~ критерий правильности спецификации
two-sided ~ двусторонний критерий
two-tail ~ \equiv **two-sided** ~
unbiased ~ несмещенный критерий
unit root ~ критерий единичных корней
variable addition ~ критерий добавления переменных
variable deletion ~ критерий удаления переменных
Wald ~ критерий Вальда
White's ~ критерий Уайта
testable поддающийся проверке
testing проверка, проверка гипотез, тестирование
hypothesis ~ проверка гипотез
sequential ~ последовательная проверка гипотез
time-dependent зависящий от времени {parameter}
time-ordered упорядоченный по времени
time-series временной ряд || относящийся к временным рядам

- time-varying** меняющийся во времени (\approx **time-dependent**)
- trade-off** обратная зависимость, альтернатива, выбор
- bias-variance** ~ выбор между смещением и дисперсией (оценки)
- power-size** ~ выбор между мощностью и размером (критерия)
- transformation** преобразование
- Box-Cox** ~ преобразование Бокса-Кокса
- cointegrating** ~ коинтегрирующее преобразование
- Koyck** ~ преобразование Койка
- linear** ~ линейное преобразование
- logarithmic** ~ логарифмическое преобразование
- normalizing** ~ нормализующее преобразование
- orthogonal** ~ ортогональное преобразование
- Prais-Winsten** ~ преобразование Прейса-Уинстена
- square root** ~ преобразование квадратного корня
- variance-stabilizing** ~ преобразование, стабилизирующее (уравнивающее) дисперсию
- trend** тренд
- common** ~ общий тренд (\approx **cointegration**)
- deterministic** ~ детерминированный тренд
- downward** ~ убывающий тренд, тенденция к понижению
- exponential** ~ экспоненциальный тренд
- linear** ~ линейный тренд
- multiplicative** ~ мультипликативный тренд
- polynomial** ~ полиномиальный тренд
- quadratic** ~ квадратичный тренд
- stochastic** ~ стохастический тренд
- upward** ~ возрастающий тренд, тенденция к росту
- variable** переменная, величина **to drop a** ~ опустить переменную, не включить переменную
- artificial** ~ искусственная переменная
- auxiliary** ~ вспомогательная переменная

binary ~ бинарная переменная, двоичная переменная
 (принимающая только два значения, \approx **dummy**)
cointegrated ~s коинтегрированные переменные
controllable ~ управляемая переменная
dependent ~ зависимая переменная
deterministic ~ детерминированная переменная
discrete ~ дискретная переменная
dummy ~ фиктивная переменная
endogenous ~ эндогенная переменная
excluded ~ исключенная переменная
exogenous ~ экзогенная переменная
explanatory ~ объясняющая переменная (\approx **independent** ~
)
independent ~ независимая переменная
indicator ~ индикаторная переменная (\approx **dummy**)
instrumental ~ инструментальная переменная
instrumental ~s [**method**] метод инструментальных переменных
lagged ~ запаздывающая переменная, переменная с лагом
latent ~ скрытая переменная (\equiv **unobservable** ~)
leading ~ опережающая переменная
limited dependent ~ ограниченная зависимая переменная
omitted ~ пропущенная переменная
predetermined ~ predetermined переменная
proxy ~ эрзац-переменная
qualitative dependent ~ качественная зависимая переменная
random ~ случайная переменная, случайная величина
redundant ~ лишняя переменная
response ~ переменная-отклик
seasonal ~ сезонная переменная
stationary ~ стационарная переменная
stochastic ~ случайная переменная, случайная величина
unrestricted ~ свободная переменная

- unobservable** ~ ненаблюдаемая переменная
- vector random** ~ векторная случайная величина, случайный вектор
- variance** дисперсия, второй центральный момент ~ **about mean** дисперсия относительно среднего, ~ **about regression** дисперсия относительно регрессии
- asymptotic(al)** ~ асимптотическая дисперсия
- bounded** ~ ограниченная дисперсия
- conditional** ~ условная дисперсия
- error** ~ дисперсия ошибки
- generalized** ~ обобщенная дисперсия
- limiting** ~ предельная дисперсия
- relative** ~ относительная дисперсия
- residual** ~ остаточная дисперсия
- sample** ~ выборочная дисперсия
- unequal** ~ неодинаковая дисперсия (\approx **heteroscedasticity**)
- unit** ~ единичная дисперсия
- variation** вариация, разброс, варьирование, коэффициент вариации, изменение, отклонение
- explained** ~ объясненная вариация
- excess** ~ избыточная вариация
- random** ~ случайная вариация, случайное изменение
- seasonal** ~s сезонные колебания
- total** ~ общая вариация
- walk** блуждание
- random** ~ случайное блуждание
- random** ~ **with drift** случайное блуждание с дрейфом
- weight** вес
- lag** ~s веса лага
- weighted** взвешенный

Список некоторых обозначений в программе Eviews

| Рус./англ. | | Eviews, принятые по умолчанию |
|---|---|-------------------------------------|
| Оценки коэффициентов регрессии/ Estimated coefficients | a, b - для парной регрессии b_0, b_1, b_2, \dots - для множественной регрессии | c(1), c(2), ... |
| Стандартные ошибки коэффициентов регрессии / Standard errors | c.o., se | <i>Std. Error,</i> <i>se</i> |
| Расчетное значение зависимой переменной (y)/ Fitted value | $y_p,$ \hat{y} | <i>yf, yhat</i> |
| Статистика Фишера / F-statistics | $F_{\text{стат}}$ | <i>F-statistic</i> |
| p-ЗНАЧЕНИЕ - величина, применяемая при проверке гипотез | P-значение, P-value | <i>Prob</i> |
| Коэффициент детерминации / Coefficient of determination | R^2 | R-squared |
| Скорректированный коэффициент детерминации/ Adjusted coefficient of determination | R^2_{adj} | Adjusted R-squared |
| Сумма квадратов остатков/ Residual Sum of Squares | RSS | Sum squared resid |
| Остатки/Residuals | e | Resid |
| Метод наименьших квадратов / Least squares | MНК | LS |

Алгоритм используемых в работах тестов

Важное правило современного эконометрического моделирования заключается во всесторонней проверке оцениваемой модели на предмет нарушения тех или иных предположений, лежащих в ее основе. Перечислим основные критерии (тесты), которые применяются в настоящее время для того, чтобы проверить, насколько правильно специфицирована оцененная регрессионная модель.

1. Критерий пропущенных переменных.
2. Критерии функциональной формы. Критерий RESET.
3. Критерии структурных изменений (критерий Чоу, CUSUM и CUSUMSQ) и критерии выбросов.
4. Критерии автокорреляции остатков (критерий Дарбина–Уотсона, критерий Годфрея (альтернативный критерий Дарбина), точечно-оптимальные критерии Кинга).
5. Критерии экзогенности регрессоров (критерий Дарбина–Ву–Хаусмана).
6. Критерий первых разностей и другие критерии преобразования данных.
7. Невложенные критерии (nonnested tests).
8. Критерии стационарности переменных.
9. Критерии гетероскедастичности ошибки.
10. Критерии нормальности (критерий Жака–Бера).

Каждому из критериев соответствует статистика, которая является функцией данных.

В предположении, что использованная вероятностная модель верна, можно теоретически вывести распределение данной статистики (зачастую распределение выводится из асимптотической теории и известно только приближенно).

Процедура проверки нулевой гипотезы состоит в том, что если полученная на основе имеющегося набора данных статистика выходит за некоторый установленный заранее доверительный интервал, то нулевая гипотеза отклоняется.

В случае отклонения нулевой гипотезы делается вывод, что принятые допущения неверны. Доверительный интервал задается, как правило, указанием критической границы. Вероятность того, что статистика выйдет за пределы доверительного интервала, за-

данного данной критической границей, и, тем самым, будет отклонена верная нулевая гипотеза, называют уровнем значимости.

Понятно, что полученной на основе имеющегося набора данных статистике соответствует некоторый уровень значимости. Этот уровень значимости сам по себе можно рассматривать как статистику и использовать для проверки нулевой гипотезы. Чаще всего на практике используют 5%-ю границу. Если получен уровень значимости менее 5%, то нулевая гипотеза отклоняется в пользу альтернативной.

К показателям качества параметров регрессии относят:

1. Стандартные ошибки оценок (анализ точности определения оценок).
2. Значения t -статистик (проверка гипотез относительно коэффициентов регрессии).
3. Интервальные оценки коэффициентов линейного уравнения регрессии.
4. Доверительные области для зависимой переменной.

Тест на проверку значимости коэффициентов регрессии

Значимость коэффициентов множественной регрессии проверяется по t -критерию Стьюдента:

t -тесты обеспечивают проверку значимости предельного вклада каждой переменной при допущении, что все остальные переменные уже включены в модель.

Незначимость коэффициента регрессии не всегда может служить основанием для исключения соответствующей переменной из модели.

Для проверки нулевой гипотезы H_0 о равенстве нулю коэффициента регрессионного уравнения ($H_0: \beta_i=0$) необходимо сравнить фактическое значение статистики с критическим значением $t_{(\alpha, n-k)}$.

Порядок работы при проверке значимости коэффициента по t -статистике:

1. Формулируется гипотеза $H_0: \beta_i=0$

2. Рассчитывается величина $t_{\text{стат}} = \frac{b_i - 0}{Se(b)}$
3. Выбирается уровень значимости γ (обычно 1% , 5% или 10%).
4. Вычисляется число степеней свободы $(n-k)$.
5. По таблицам распределения Стьюдента определяется критическое значение $t_{(\gamma; n-k)}$ (двухсторонний критерий)
6. Если модуль t -статистики больше критического значения, то коэффициент является значимым на уровне значимости γ .
7. В противном случае коэффициент не значим (на данном уровне γ).

Правило оценки значимости коэффициентов регрессии без использования таблиц:

1. Если $|t_{b_i}| \leq 1$, то коэффициент β_i не значим, т.к. доверительная вероятность менее 0,7.
2. Если $1 < |t_{b_i}| \leq 2$, то найденная оценка может рассматриваться как относительно (слабо) значимая. При этом доверительная вероятность лежит между 0,7 и 0,95 при числе степеней свободы, больших 60.
3. Если $2 < |t_{b_i}| \leq 3$, то коэффициент значим. при числе степеней свободы, больших 60. Доверительная вероятность лежит между значениями 0,95 и 0,99.
4. Если $|t_{b_i}| > 3$, то это почти полная гарантия значимости коэффициента.

Тест на проверку значимости уравнения в целом

Суть проверки общего качества уравнения регрессии – оценить, насколько хорошо эмпирическое уравнение регрессии согла-

соединяется со статистическими данными. Основные показатели качества:

- коэффициент детерминации R^2
- скорректированный коэффициент детерминации
- значение F -статистики
- сумма квадратов остатков (RSS)
- оцененная сумма квадратов остатков (ESS)
- общая сумма квадратов остатков (TSS)
- стандартная ошибка регрессии S_e
- прочие показатели: средняя ошибка аппроксимации, индекс множественной корреляции и т.д.

Коэффициент детерминации R^2 достигает максимума из возможных значений, когда RSS принимает наименьшее из достижимых значений (критерий максимума R^2 эквивалентен принципу МНК)

Коэффициент R^2 показывает долю объясненной вариации зависимой переменной:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Используется для предварительной оценки качества модели (по степени корреляции между объясняемой переменной и ее предсказанным значением).

Величину $1-R^2$ можно трактовать как вероятность необъяснения разброса значений зависимой переменной выбранными регрессорами.

При правильном включении факторов в модель индекс множественной корреляции будет существенно превосходить наибольшее из значений коэффициента парной корреляции.

F -статистика для проверки качества уравнения регрессии

F -статистика представляет собой отношение объясненной суммы квадратов (в расчете на одну независимую переменную) к остаточной сумме квадратов (в расчете на одну степень свободы):

$$F(k-1, n-k) = \frac{ESS/(k-1)}{RSS/(n-k)}$$

Поскольку и ESS , и RSS , по сути – дисперсии, умноженные на соответствующие числа степеней свободы, то их распределение будет связано со статистикой Пирсона, а потому эта дробь окажется распределенной по закону Фишера.

$$F(k-1, n-k) = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{\frac{ESS}{TSS}/(k-1)}{\frac{RSS}{TSS}/(n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

F-статистики в разных моделях с разным числом наблюдений и (или) переменных несравнимы.

Порядок действий при проверке значимости уравнения по F -статистике:

1. Формулируем гипотезы
2. Выбираем уровень значимости (1%, 5%, 10% и т.п.).
3. Вычисляем число степеней свободы k и $(n-k)$.
4. По таблицам F -распределения определяем критическое значение $F_{\gamma; k; n-k}$.
5. Если F -статистика больше $F_{\gamma; k; n-k}$, то уравнение в целом является значимым на уровне значимости γ .
6. В противном случае уравнение в целом не значимо (на данном уровне γ).

В парной регрессии F -статистика равна квадрату t -статистики; то же верно и для их критических уровней (двухсторонний для t -статистики).

В парной регрессии значимость коэффициента регрессии и значимость уравнения в целом эквивалентны.

Сумма квадратов остатков RSS является оценкой необъясненной части вариации зависимой переменной и используется как основная минимизируемая величина в МНК, а также для расчета других показателей.

Значения RSS в разных моделях с разным числом наблюдений и (или) переменных несравнимы.

Стандартная ошибка регрессии S_e является оценкой величины квадрата ошибки, приходящейся на одну степень свободы модели, и используется как основная величина для измерения качества модели (чем она меньше, тем лучше).

Значения S_e в однотипных моделях с разным числом наблюдений и (или) переменных сравнимы.

Типичные ошибки в использовании показателей качества регрессии:

- величина коэффициентов регрессии не указывает на силу связи или силу влияния на зависимую переменную;
- значимость коэффициентов по t -тестам не позволяет сделать вывод о справедливости тех или иных теорий;
- t -статистики не указывают на относительную важность коэффициентов регрессии;
- t -статистики предназначены для использования исключительно для выборки и бесполезны для анализа всей совокупности;
- нельзя сравнивать t -статистики, F -статистики, коэффициенты детерминации в разных уравнениях.

Тест на нормальность распределения

Критерий Жака-Бера (Jarque-Bera) используется для проверки гипотезы о том, что исследуемая выборка является выборкой нормально распределенной случайной величины с нулевым математическим ожиданием и дисперсией.

Этот критерий используется в ряде пакетов статистического анализа данных (например, в Eviews) для проверки гипотезы H_0 нормальности ошибок в модели наблюдений, точнее,

$$H_0 : \varepsilon_1, \dots, \varepsilon_n \sim i.i.d.N(0, \sigma^2)$$

(значение σ^2 не конкретизируется). Если эта гипотеза верна, то при большом количестве наблюдений JB статистика

$$JB = n \left[\frac{(\text{sample skewness})^2}{6} + \frac{(\text{sample skewness} - 3)^2}{24} \right]$$

имеет распределение, близкое к распределению хи-квадрат с двумя степенями свободы $\chi^2(2)$, функция плотности которого имеет вид

$$p(x) = \frac{1}{2} e^{-x/2}, x > 0$$

Здесь «sample skewness» — *выборочный коэффициент асимметрии*,

$$\text{sample skewness} = \frac{m_3}{(m_2)^{3/2}},$$

«sample kurtosis» — *выборочный коэффициент эксцесса*,

$$\text{sample kurtosis} = \frac{m_4}{m_2^2},$$

где

$$m_k = \frac{1}{n} \sum_{i=1}^n e_i^k$$

и e_1, \dots, e_n — остатки, полученные при оценивании модели.

Если распределение ошибок действительно является нормальным, то значения выборочного коэффициента асимметрии близки к нулю, а значения выборочного коэффициента эксцесса близки к 3.

Существенное отличие выборочного коэффициента асимметрии от нуля указывает на несимметричность (относительно нуля) графика функции плотности распределения ошибок («скошенность» распределения). Существенное отличие от 3 выборочного коэффициента эксцесса указывает на не характерные для нормаль-

ного распределения «островершинность» (при значении этого коэффициента, большем трех) или излишнюю «сглаженность» (при значении этого коэффициента, меньшем трех) графика функции плотности распределения ошибок.

При нарушении условия нормальности распределения ошибок значения статистики JB имеют тенденцию к возрастанию. Поэтому гипотеза нормальности ошибок *отвергается, если* значения этой статистики «слишком велики», а именно, если

$$JB > \chi^2_{1-\alpha}(2),$$

где $\chi^2_{1-\alpha}(2)$ — квантиль распределения $\chi^2(2)$, соответствующая уровню $1-\gamma$.

Замечание. Критерии Дарбина–Уотсона и Голдфелда–Квандта являются точными, в том смысле, что они непосредственно учитывают количество наблюдений n . В противоположность этому, критерий Жака–Бера является *асимптотическим критерием*: распределение статистики JB хорошо приближается распределением $\chi^2(2)$, только при большом количестве наблюдений. Поэтому вполне полагаться на результаты применения критерия Жака–Бера можно только в таких ситуациях. Следует отметить, что при росте n статистика Жака–Бера сходится к распределению хи-квадрат с двумя степенями свободы, поэтому в практике иногда используют таблицу квантилей распределения хи-квадрат. Однако это является ошибкой - сходимость слишком медленная и неравномерная.

Например, даже при $n = 70$ (что уже немало) и значении статистики Жака–Бера $JB = 5$ таблица квантилей распределения хи-квадрат дает нам p -значение, равное 0,08, в то время как на самом деле соответствующее p -значение равно 0.045 - то есть мы можем ошибочно принять гипотезу, которую скорее следовало бы отвергнуть. Поэтому более правильным является использование специально составленной таблицы квантилей распределения Жака–Бера. Помимо критерия Жака–Бера в специализированные пакеты программ статистического анализа данных часто встраиваются и другие асимптотические критерии, например, критерии Уайта, Бройша–Годфри и др.

Тест на проверку ограничений на коэффициенты регрессии

Для проверки ограничений на коэффициенты используется тест Уальда.

Рассмотрим модель регрессии в общем виде:

$$Y = F(X, \beta) + \varepsilon$$

где $F(X, \beta)$ – функция регрессоров X и параметров β .

Суть теста состоит в проверке ограничений на коэффициенты регрессии. Если гипотеза о правомерности ограничений верна, то величина W , называемая статистикой Уальда, имеет в асимптотике распределение χ^2 с числом степеней свободы q , равным количеству ограничений на параметры. В том случае, когда расчетное значение статистики Уальда превосходит критическое значение статистики χ^2 , то гипотеза о правомерности ограничений отклоняется.

В том случае, когда исходная модель является линейной многофакторной моделью, уравнение для которой можно записать в матричной форме

$$Y = X\beta + \varepsilon,$$

а ограничения на параметры регрессии – линейны, статистика Уальда преобразуется в статистику, имеющую распределение Фишера.

Порядок проведения теста:

1. Формулируется нулевая гипотеза, которая содержит линейные ограничения на коэффициенты в виде некой математической функции.

$$H_0: R\beta - r = 0,$$

где R – матрица размерности $q \times k$ (q – число ограничений, k – число параметров в уравнении регрессии, r – вектор-столбец размерности q).

Конкурирующая гипотеза H_1 отвергает сформулированные ограничения.

2. Оценивается исходная регрессия без ограничений, для которой рассчитывается сумма квадратов остатков RSS_{UR} (UR - unrestricted – без ограничений)
3. Оценивается регрессия с введенными ограничениями, для нее также рассчитывается сумма квадратов остатков RSS_R (R - restricted – с ограничениями).
4. Рассчитывается статистика $F_{стат}$, имеющая распределение Фишера $F(q, n-k)$.

$$F_{стат} = \frac{W}{q} = \frac{(RSS_R - RSS_{UR}) / q}{RSS_{UR} / (n - k)}$$

5. Выбирается уровень значимости γ (обычно 5%).
6. Сравнивается полученное значение $F_{стат}$ с табличным (критическим) значением $F_{кр}(\gamma, q, n-k)$ или рассчитывается P -значение, соответствующее полученной $F_{стат}$.
7. Если расчетное значение $F_{стат} > F_{кр}(\gamma, q, n-k)$ (или P -значение $< \gamma$), то нулевая гипотеза H_0 отклоняется в пользу H_1 , и ограничение в модель не вводится.
8. Если расчетное значение $F_{стат} < F_{кр}(\gamma, q, n-k)$ (или P -значение $> \gamma$), то нулевая гипотеза H_0 не отклоняется, и ограничение в модель вводится.

Тесты на гетероскедастичность

Для проверки наличия или отсутствия гетероскедастичности на практике используются несколько тестов (Голдфелда–Квандта, Спирмена и др.).

Тест Уайта

Рассмотрим сначала тест Уайта. Содержательный смысл теста в том, что часто гетероскедастичность модели вызвана зависимостью дисперсий ошибок от признаков. Уайт предложил метод тестирования гипотезы H_0 без каких-либо предположений о структуре гетероскедастичности. Сначала к исходной модели применяется

обычный метод наименьших квадратов, и находятся остатки регрессии e_i .

Затем осуществляется регрессия квадратов этих остатков на все признаки, их квадраты, попарные произведения и константу. Тогда при гипотезе H_0 величина nR^2 асимптотически имеет распределение $\chi^2(N-1)$, где R^2 — коэффициент детерминации, а — N число регрессоров второй регрессии.

Предполагается, что дисперсии $\sigma_{e_i}^2$ связаны с объясняющими переменными $X_j, j = \overline{1, m}$ в виде

$$\sigma_{e_i}^2 = f(X_{i1}, X_{i2}, \dots, X_{im}) + \eta_i, i = \overline{1, n}$$

где $f(\cdot)$ — квадратичная функция от аргументов.

Т.к. дисперсии $\sigma_{e_i}^2$ неизвестны, то их заменяют оценками квадратов отклонений e_i^2 .

1. Формулируется нулевая гипотеза о гомоскедастичности остатков
2. Оценивается уравнение регрессии (приведен пример с тремя регрессорами):

$$\hat{y}_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i3}$$

3. Вычисляются остатки

$$e_i = y_i - \hat{y}_i, i = \overline{1, n}$$

4. Оценивают вспомогательное уравнение регрессии:

$$e_i^2 = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i3} + \alpha_4 X_{i1}^2 + \alpha_5 X_{i2}^2 + \alpha_6 X_{i3}^2 + \alpha_7 X_{i1} X_{i2} + \alpha_8 X_{i1} X_{i3} + \alpha_9 X_{i2} X_{i3} + \eta_i$$

5. Определяют из вспомогательного уравнения тестовую статистику $U = nR^2$.

6. Проверяют общую значимость вспомогательного уравнения с помощью критерия χ^2 . Если

$$U > \chi_{r, k}^2$$

то гипотеза гомоскедастичности отвергается. Число степеней свободы k равно числу объясняющих переменных вспомогательного уравнения. В частности, для рассматриваемого случая $k = 9$.

Плюс данного теста — его универсальность. Минусы : 1) если гипотеза H_0 отвергается, то никаких указаний на функциональную форму гетероскедастичности мы не получаем; 2) несомненным минусом является поиск вслепую вида регрессии (как один из вариантов приближение простыми полиномами второй степени без каких бы то ни было объяснений этому).

Тест Голдфелда–Квандта

Алгоритм применения теста Голдфелда–Квандта.

1. Выделяется фактор пропорциональности Z . Данные упорядочиваются в порядке возрастания величины Z .

2. Выбрасывается средняя треть упорядоченных наблюдений. Для первой и последней трети строятся две отдельные регрессии (каждая по n' наблюдений), используя спецификацию первоначального уравнения.

3. Берутся суммы квадратов остатков для регрессий по первой трети RSS_1 и последней трети RSS_3 и рассчитывается их отношение

$$GQ = \frac{RSS_3}{RSS_1}$$

(количество наблюдений одинаково).

Используется F -тест для проверки гипотезы о гомоскедастичности: если величина GQ превышает критическое значение F с $n'-k-1$ степенями свободы (и для числителя и для знаменателя, n' — объем подвыборки), то гипотеза о гомоскедастичности отвергается.

Тесты на автокорреляцию

Тест Дарбина–Уотсона на автокорреляцию

$H_0: \rho=0$ — нет АК

$H_1: \rho>0$ — есть положительная АК

Проверка гипотезы:

1. Рассчитывается тестовая статистика

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{RSS}$$

$$DW \rightarrow 2 - 2\rho$$

- По таблице Дарбина-Уотсона находятся значения d_L , d_U (нижняя и верхняя границы для критических значений, определяются для выбранного уровня значимости по числу наблюдений)
- Сравнивается расчетное значение с табличным. Если

$$DW \in [0, d_L] \Rightarrow \rho > 0$$

$$DW \in (d_U, 2] \Rightarrow \rho = 0$$

$$DW \in [d_L, d_U] \Rightarrow ???$$

Тест Льюнга-Бокса

H_0 : нет автокорреляции порядка K

Тестовая статистика

$$Q_k = T(T+2) \sum_{k=1}^K \frac{1}{T-k} r_k^2$$

T – число наблюдений

r_k – выборочные коэффициенты автокорреляции остатков

Порядок K выбирается исследователем

$K > p+q$

Если H_0 отклоняется, подбор модели надо продолжать

Тесты на спецификацию

RESET-тест Рамсея

RESET-тест Рамсея - это обобщенный тест на наличие следующих ошибок спецификации модели линейной регрессии:

- наличие пропущенных переменных. Регрессия содержит не все объясняющие переменные;
- неверная функциональная форма. Некоторые или все переменные должны быть преобразованы с помощью логарифмической, степенной, обратной или какой-либо другой функции;
- корреляция между фактором X и случайной составляющей модели, которая может быть вызвана ошибками измерения факторов, рассмотрением систем уравнений или другими причинами.

Ошибки такого рода приводят к смещению среднего остатков регрессионной модели.

Тестируется спецификация: $Y = X\beta + \varepsilon$

H_0 : Функциональная форма приемлема $H_0: \varepsilon \sim N(0, \sigma^2 I)$

H_1 : Функциональная форма не приемлема $H_1: \varepsilon \sim N(\mu, \sigma^2 I)$

Алгоритм теста

- оценивается уравнение: $Y = X\beta + \varepsilon$
- определяются расчетные значения
- оценивается вспомогательная регрессия $Y = X\beta + \alpha_1 \bar{Y}^2 + \alpha_2 \bar{Y}^3 + \alpha_4 \bar{Y}^4 + \dots + \varepsilon$
- проверяется гипотеза, что все коэффициенты при степенях расчетных равны нулю одновременно

Для принятия решения используется:

- статистика Фишера (F_{stat});
- статистика Хи-квадрат (χ^2);

Если $|T_{stat}| < T_{cr}(5\%, df)$, т.е. $Prob(T_{stat}) > 0.05$, то H_0 не отвергается

Или:

$Prob > 0,1$ гипотеза о приемлемости функциональной формы принимается

$Prob < 0,01$ гипотеза о приемлемости функциональной формы отклоняется

РЕ-тест

Тестируется линейная модель А против лог-линейной модели В

$$Y = X\beta + \varepsilon \quad (A)$$

$$\text{LOG}(Y) = \text{LOG}(X)\beta + \varepsilon \quad (B)$$

Идея: расчетные значения, полученные по модели B, не должны помогать в объяснении Y, и наоборот.

Алгоритм теста

- Оценивается модель (A), получают расчетные значения
- Оценивается модель (B), получают расчетные значения $\text{LOG}(\hat{Y})_B$
- Оценивается вспомогательная регрессия

$$Y = X\beta + \delta_{lin}(\text{LOG}(\hat{Y}_A) - \text{LOG}(Y^{\wedge}_B)) + \varepsilon \quad (C)$$
- Если верна нулевая гипотеза (модель (A)), то t-статистики для коэффициента δ_{lin} имеет асимптотическое нормальное распределение
 $|t \text{ stat}| < tcr(5\%, df)$, т.е. $\text{Prob}(t \text{ stat}) > 0.05$, то H_0 не отвергается.

Аналогично тестируется модель (B) против модели (A)

Тест Дики-Фуллера (ADF-тест)

Тест Дики-Фуллера (Dickey and Fuller) основан на модели

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \delta t + \varepsilon_t$$

Альтернативы $-1 < \beta_2 < 1$ $\beta_2 = 1$ $\delta = 0$ $\delta \neq 0$

Гипотеза H_0 : ряд нестационарен

либо (стационарен в разностях)

либо (трендово стационарен)

Гипотеза H_1 : ряд стационарен

Если $|t \text{ stat}| < tcr(5\%, df)$, т.е. $\text{Prob}(t \text{ stat}) > 0.05$, то H_0 не отвергается

Заметим:

- если мы не можем отклонить гипотезу о наличии единичного корня, то это не обязательно означает, что гипотеза верна;
- не все временные ряды, для которых H_0 не отклоняется, являются интегрируемыми процессами первого порядка;
- тест ADF имеет малую мощность.

Описание и программа курса «ЭКОНОМЕТРИКА» (ПРОДВИНУТЫЙ УРОВЕНЬ)

Рекомендуется для направления подготовки

38.04.01 (ЭКОНОМИКА)

Программы в рамках направления:

- Финансовое управление в секторах экономики
- Экономика фирмы и отраслевых рынков
- Международная торговля
- Страхование и управление страховой деятельностью
- Бухучет, внутренний контроль и аудит
- Экономика устойчивого развития

Квалификация (степень) выпускника Магистр

Цели и задачи дисциплины: Целью курса «Эконометрика. Продвинутый уровень» является ознакомление студентов с широким кругом тем в современной эконометрике. Изучаются методы анализа пространственных данных, временных рядов и панельных данных, развивающие базовые методы, освоенные студентами в базовом курсе эконометрики. Акцент делается на современных методах исследования, иллюстрируется приложениями из микро- и макроэкономики, а также из области финансов. Особенностью курса является применение специализированного программного обеспечения, позволяющего выполнять эконометрическое моделирование на профессиональном уровне.

Место дисциплины в структуре ООП: Курс относится к базовой части, рассчитан на 18 лекционных часов, 36 часа практических занятий и 54 часа самостоятельной работы студентов.

Для освоения дисциплины студент должен

знать:

- основы математического анализа, линейной алгебры, теории вероятности и математической статистики,
- закономерности функционирования современной экономики на макро- и микроуровне;
- основные методы построения эконометрических моделей объектов, явлений и процессов;

уметь:

- применять методы математического анализа и моделирования для решения экономических задач;
- анализировать во взаимосвязи экономические явления и процессы на микро- и макроуровне;
- использовать источники экономической, социальной, управленческой информации;
- анализировать и интерпретировать данные отечественной и зарубежной статистики о социально-экономических процессах и явлениях, выявлять тенденции изменения социально-экономических показателей;
- строить на основе описания ситуаций стандартные эконометрические модели, анализировать и содержательно интерпретировать полученные результаты;

- прогнозировать на основе стандартных эконометрических моделей поведение экономических агентов, развитие экономических процессов и явлений на микро- и макроуровне;

а также **владеть MS Excel** и современными методами сбора, обработки и анализа экономических данных.

Требования к результатам освоения дисциплины:

Процесс изучения дисциплины направлен на формирование следующих компетенций:

- способность к абстрактному мышлению, анализу, синтезу (ОК-1);

- готовность к коммуникации в устной и письменной формах на русском и иностранном языках для решения задач профессиональной деятельности (ОПК-1);

- способность проводить самостоятельные исследования в ответствии с разработанной программой (ПК-3);

способность представлять результаты проведенного исследования научному сообществу в виде статьи или доклада (ПК-4);

- способность анализировать и использовать различные источники информации для проведения экономических расчетов (ПК-9);

- способность составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом (ПК-10);

В результате изучения дисциплины студент должен:

Знать:

- основные результаты новейших исследований, опубликованные в ведущих профессиональных журналах, по проблемам эконометрики;

- современные методы эконометрического анализа;

- современные программные продукты для проведения эконометрического исследования;

Уметь:

- применять различные методы эконометрического анализа для решения задач в области микро- и макроэкономики, и в области финансов;

- использовать программу Eviews для решения эконометрических задач;
- строить прогнозы развития конкретных экономических процессов на основе эконометрической модели.

Владеть:

- современной методикой построения эконометрических моделей;
- навыками самостоятельной исследовательской работы.

Сведения об авторах

Балашова Светлана Алексеевна, кандидат физико-математических наук, доцент кафедры экономико-математического моделирования экономического факультета Российского университета дружбы народов.

Лазанюк Инна Васильевна, кандидат экономических наук, доцент кафедры экономико-математического моделирования экономического факультета Российского университета дружбы народов.

Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 3 зачетных единицы.

| Вид учебной работы | Всего часов | Семестры | | | |
|--|-------------|----------|----|---|---|
| | | 1 | 2 | 3 | 4 |
| Аудиторные занятия (всего) | 54 | | 54 | | |
| В том числе: | - | - | - | - | - |
| Лекции | 18 | | 18 | | |
| Практические занятия (ПЗ) | 36 | | 36 | | |
| Семинары (С) | | | | | |
| Лабораторные работы (ЛР) | | | | | |
| Из них в интерактивной форме (ИФ) | 36 | | 36 | | |

| | | | | | |
|------------------------------------|-----|---|---|---|---|
| Самостоятельная работа (всего) | 54 | | | | |
| В том числе: | - | - | - | - | - |
| Курсовой проект (работа) | | | | | |
| Расчетно-графические работы | 30 | | | | |
| Реферат | | | | | |
| Другие виды самостоятельной работы | 24 | | | | |
| Общая трудоемкость час | 108 | | | | |
| зач. ед. | 3 | | | | |

Содержание разделов дисциплины

| № п/п | Наименование раздела дисциплины | Содержание раздела |
|-------|---|--|
| 1. | Модель множественной линейной регрессии | <p>Природа эконометрики и ее место в современной экономической науке. Метод наименьших квадратов (МНК). Парная модель линейной регрессии. Многофакторная модель линейной регрессии. Матричные обозначения. Свойства МНК-оценок при выполнении условий Гаусса-Маркова. Асимптотические свойства МНК-оценок. Линеаризация</p> <p>Отбор множества объясняющих переменных. Сравнение вложенных и не вложенных моделей. Тестирование функциональной формы.</p> <p>Тест Чоу. Фиктивные переменные в моделях регрессии.</p> |

| | | |
|----|--|--|
| 2. | Методы оценки при нарушении условий классической регрессии | <p>Проблема гетероскедастичности, автокорреляция, влияние на МНК-оценки, обобщенный метод наименьших квадратов (ОМНК), тесты на гетероскедастичность, стандартные ошибки в форме Уайта.</p> <p>Эндогенность, инструментальные переменные. Оценивание методом инструментальных переменных. Обобщенный метод инструментальных переменных. Двухшаговый метод наименьших квадратов.</p> <p>Метод максимального правдоподобия (ММП). Формулировка метода, оценка по ММП, свойства оценок. Тесты на спецификацию модели.</p> |
| 3. | Одномерные модели временных рядов | <p>Стационарность. Автокорреляционная функция. Модели авторегрессии и скользящего среднего. ARMA-модели. Единичные корни. Тесты на единичный корень. Условная гетероскедастичность. ARCH и GARCH модели.</p> |
| 4. | Многомерные модели временных рядов | <p>Динамические модели со стационарными переменными. Модели с нестационарными переменными. Коинтеграция. Тестирование на коинтеграцию.</p> |

Разделы дисциплин и виды занятий

| № п/п | Наименование раздела дисциплины | Лекция. | Практ. занятие | СРС | Всего час. |
|-------|---------------------------------|---------|----------------|-----|------------|
|-------|---------------------------------|---------|----------------|-----|------------|

| | | | | | |
|----|---|---|----|----|----|
| 1. | Модель множественной линейной регрессии | 4 | 8 | 12 | 24 |
| 2. | Методы оценки при нарушении условий МНК | 4 | 8 | 12 | 24 |
| 3. | Одномерные модели временных рядов | 4 | 8 | 12 | 24 |
| 4. | Многомерные модели временных рядов | 6 | 12 | 18 | 36 |

Практические занятия (семинары)

| № п/п | № раздела дисциплины | Тематика практических занятий (семинаров) | Трудоемкость (час.) |
|-------|----------------------|--|---------------------|
| 1. | 1 | Оценивание доходности акций по модели CAPM | 2 |
| 2. | 1 | Оценивание доходности акций по модели арбитражного ценообразования | 2 |
| 3. | 1 | Оценивание параметров кривой обучения | 2 |
| 4. | 1 | Оценивание альтернативных спецификаций отдачи от масштаба | 2 |
| 5. | 2 | Автокорреляция в модели кривой обучения | 2 |
| 6. | 2 | Гетероскедастичность в модели спроса на труд | 2 |
| 7. | 2 | Оценивание отдачи от рекламы | 2 |
| 8. | 2 | Оценивание нормальной регрессионной модели методом максимального правдоподобия | 2 |

| | | | |
|-----|-----|---|---|
| 9. | 3 | Оценивание ARMA модели | 2 |
| 10. | 3 | Тестирование единичных корней | 2 |
| 11. | 3 | Оценивание ARCH и GARCH моделей | 4 |
| 12. | 4 | Оценивание модели адаптивных ожиданий | 6 |
| 13. | 4 | Оценивание модели частичной корректировки | 6 |
| 14. | 1-4 | Итоговая контрольная работа | 2 |

Учебно-методическое и информационное обеспечение дисциплины:

а) основная литература

1. Балашова С.А., Лазанюк И.В. Эконометрика в задачах и решениях: учебное пособие для магистров. М.: РУДН, 2017.
2. Балашова С.А., Лазанюк И.В. Эконометрика в задачах и решениях: учебное пособие для магистров. М.: РУДН, 2014.
3. Берндт Эрнст Р. Практика эконометрики: классика и современность: Учебник. М. : Юнити, 2005.
4. Verbeek Marno A Guide to Modern Econometrics. 4th edition. Wiley, 2012.
5. Вербик М. Путеводитель по современной эконометрике. М.: Научная книга, 2008.
6. Елисева И.И. Эконометрика: учебник для магистров/ Под ред. И.И.Елисеевой.- М.: Юрайт, 2012.

б) дополнительная литература

1. Бабешко Л. Основы эконометрического моделирования : Учебное пособие . М. : КомКнига, 2007.
2. Доугерти Кристофер. Введение в эконометрику: Учебник для вузов: Пер. с англ.- 2-е изд.- М.: ИНФРА-М, 2009.

3. Матюшок В.М., Балашова С.А., Лазанюк И.В. «Основы эконометрического моделирования с использованием Eviews». – М.: изд-во РУДН, 2015.
4. Матюшок В.М., Балашова С.А., Лазанюк И.В. «Основы эконометрического моделирования с использованием Eviews». – М.: изд-во РУДН, 2011.
5. Dougherty: Introduction to Econometrics 4e. Oxford University Press, 2011
6. Green W.H. Econometric Analysis. 7th edition , 2011.

в) программное обеспечение Eviews 5.0, Eviews 7.0

г) базы данных, информационно-справочные и поисковые системы

www.gks.ru – портал Федеральной службы государственной статистики РФ

www.cbr.ru – портал Центрального банка РФ

<http://censtats.census.gov/> - социально-экономическая статистика по штатам США

<http://www.bls.gov/> - Бюро статистики труда США

<http://data.worldbank.org/> - база данных Мирового банка

<http://pwt.econ.upenn.edu/> - база данных для проведения меж-
страновых сравнений

Методические рекомендации по организации изучения дисциплины:

После каждого раздела студентами выполняется и защищается самостоятельная работа на материале реальных статистических данных. Работа выполняется в группе с использованием программного обеспечения.

Вопросы для самопроверки

1. Какие предположения надо сделать, чтобы МНК-оценка была несмещенной оценкой вектора неизвестных параметров?
2. Объясните, как можно построить доверительный интервал для коэффициента регрессии? Какие дополнительные предположения необходимы?
3. Какие предположения надо сделать, чтобы МНК-оценка была состоятельной оценкой вектора неизвестных параметров?
4. Объясните полезность критериев R^2 , AIC, SIC
5. Как можно протестировать модель А против модели В, если они объясняют одну и ту же переменную?
6. Как влияет включение лагированной зависимой переменной в спецификацию на свойства МНК-оценок?
7. Поясните алгоритм теста Голдфелда-Квандта
8. Каковы последствия ошибок измерения регрессора?
9. Каковы последствия ошибок измерения зависимой переменной?
10. Какие требования предъявляются в инструментальным переменным?
11. Приведите примеры инструментальных переменных
12. Что такое функция правдоподобия?
13. Какие вы знаете спецификационные тесты, основанные на методе максимального правдоподобия?
14. Что такое AR(1) процесс
15. Что такое MA процесс?
16. Что такое ARMA процесс?
17. Опишите алгоритм теста Дики-Фулера
18. Что такое условная гетероскедастичность?
19. Каковы эффекты влияния автокорреляции на свойства МНК-оценок? Какие есть подходы для устранения этих последствий?
20. Объясните, что подразумевается под понятием «ложная регрессия»

21. Какие тесты используются для проверки гипотезы о некоинтегрированности двух рядов динамики?
22. Что такое механизм коррекции остатков?

Творческая работа по дисциплине

В работе необходимо:

1. Собрать данные и провести эконометрический анализ по выбранной теме (изучение и использование теоретического материала по теме **ОБЯЗАТЕЛЬНО!**).
2. Рассчитать не менее 3 спецификаций (линейная, нелинейная, парная, множественная.).

При выборе спецификации использовать следующие приемы:

Выбор доминирующего фактора на основе матрицы парных корреляций

Тест на пропущенные переменные

Тест на ограничения

Тест на целесообразность введения фиктивных переменных

РЕ-тест

Тест Рамсея

3. Провести тест на значимость параметров (Стьюдента) и уравнения (Фишера).
4. Проверить все условия (тесты на гетероскедастичность, автокорреляцию, эндогенность) по теореме Гаусса-Маркова. В случае обнаружения проблем, провести процедуру корректировки. Необходимо подтвердить, что проблема устранена.
5. Построить доверительные интервалы для всех коэффициентов окончательной модели.
6. Построить прогноз на 2-3 периода (прогнозные значения факторов задать по своему усмотрению).
7. Построить доверительный интервал прогноза. Прокомментировать полученный результат
8. Сделать окончательный вывод с использованием всей проделанной работы (не менее 0,5 стр. А4).
9. Подготовить презентацию (10-12) слайдов.
10. Работу оформить в виде тезисов статьи. Требования к тезисам:

- A. Введение (актуальность, цель, задачи)
- B. Основная часть: что делаем и с какой целью. Изложение материала должно быть логично, последовательно (одно вытекает из другого). Выводы по ходу изложения приветствуются.
- C. Заключение (основные выводы) .

Работа выполняется самостоятельно или в группе не более 3 человек.

Этапы сдачи работы:

1. Выбор темы, обоснование актуальности, формулировка целей и задач, подбор источников данных и литературы – до 22 февраля.
 Выслать на почту преподавателя 1 стр. текста и ссылки на источники теоретического материала и исходные данные.
 Результаты рассмотрения 1-го этапа – до 29 февраля
 Ко второму этапу допускаются работы, своевременно прошедшие 1-й этап.
2. Расчетная часть работы с описанием этапов эконометрического моделирования – до 14 марта.
 Загрузить на портал файл с расчетами и 4-5 стр. Текст в распечатанном виде сдать преподавателю
 Результаты рассмотрения 2-го этапа - до 21 марта.
3. Формулировка выводов, окончательное оформление работы, подготовка к защите – до 28 марта
 Загрузить на портал готовую работу, файл с окончательными расчетами и презентацию

Критерии оценки (всего 20 баллов):

- Полнота информационного массива, качество данных - 4 бал.
- Обоснованность априорных предположений - 4 бал.
- Правильность выбора модели и методов оценки - 4 бал.
- Тестирование результатов по статистическим критериям - 4 бал.
- Обоснованность выводов - 4 бал.

Перечень тем для презентаций

Требования к подготовке презентации

(презентацию готовит 1-2 человека (кроме биографий), 10-12 слайдов на 10 баллов. Обязательно наличие слайда с источниками информации (не менее 7-10 источников)

1. Краткая биография, основные работы, награды
 - 1) Рагнар Фриш
 - 2) Ян Тинберген
 - 3) Рональд Фишер
 - 4) Джеймс Дарбин
 - 5) Джордж Бокс
 - 6) Гвилим Дженкинс
 - 7) К. Грэнджер
 - 8) Лоуренс Клейн
 - 9) Дж. Тобин
 - 10) Д. Мак-Фадден
 - 11) Дж. Хекман
 - 12) Р. Энгл
2. Нобелевские премии по эконометрике (лауреат, дата вручения премии, вклад в науку)
3. Модели (описание, математическая форма, основные следствия)
 - 1) Макроэкономическая модель Клейна (США)
 - 2) Макроэкономическая модель Клейна—Гольдбергера (США)
 - 3) Макроэкономическая модель Макарова-Айвазяна (Россия)
 - 4) Модель Бокса-Дженкинса
 - 5) Модель Лизера
 - 6) Модель Фамы-Френча
 - 7) Модель арбитражного ценообразования
 - 8) Модели с распределенными лагами
4. Методы оценки (основная идея, условия применения, свойства оценок)
 - 1) Метод максимального правдоподобия

- 2) Метод инструментальных переменных
- 3) Двухшаговый МНК
- 4) Трехшаговый МНК
5. Модели с ограниченными зависимыми переменными
 - 1) Логит-регрессии
 - 2) Пробит-регрессии
 - 3) Тобит-модели
6. Модели, основанные на панельных данных (описание, математическая форма, основные следствия)
 - 1) Модель с фиксированными эффектами
 - 2) Модель со случайными эффектами

Перечень вопросов в тестовой форме

(на остаточные знания по базовому курсу)

1. Установить правильную последовательность.

Этапы построения эконометрической модели:

- 1) оценка параметров модели (параметризация);
- 2) спецификация модели;
- 3) проверка адекватности модели;
- 4) сбор статистической информации об объекте исследования.

2. Выбрать правильный ответ.

Принцип спецификации модели, лежащий в основании классификации: экономические модели; эконометрические модели

- 1) формализация экономических закономерностей;
- 2) равенство числа уравнений модели числу эндогенных переменных;
- 3) датирование переменных;
- 4) включение случайных возмущений.

3. Пространственные данные фиксируют значения:

- a) по одному объекту за период времени;
- b) по нескольким объектам за один тот же момент времени;
- c) по нескольким объектам за период времени;
- d) по одному объекту за один тот же момент времени.

4. Выбрать правильный ответ.

Часть зависимой переменной в регрессионной модели, которая полностью объясняется значением регрессора:

- 1) случайное возмущение;
- 2) фактор;
- 3) уравнение регрессии.

5. Выбрать правильный ответ.

В качестве критерия отбора уравнения регрессии $f(X, P)$ в МНК используется:

$$1) F = \sum_{i=1}^n (Y_i - f(X_i, \beta));$$

$$2) F = \sum_{i=1}^n (Y_i - f(X_i, \beta))^2;$$

$$3) F = \max(Y_i - f(X_i, \beta)).$$

$Y_t, X_t, t = 1, \dots, n$ — выборочные данные зависимой переменной и регрессора.

6. Выбрать правильный ответ.

Несмещенной оценкой дисперсии возмущений в парной регрессии со свободным членом является оценка:

$$1) \frac{1}{n-2} \sum_{t=1}^n e_t^2; \quad 2) \frac{1}{n-1} \sum_{t=1}^n e_t; \quad 3) \frac{1}{n} \sum_{t=1}^n e_t^2; \quad 4) \frac{1}{n-2} \sum_{t=1}^n e_t,$$

где e_t — остаток регрессии, $t = 1, \dots, n$, n — объем выборки.

7. Выбрать правильные ответы.

Спецификация парной регрессионной модели может иметь вид:

1) $Y = a + bX + \varepsilon$;

2) $Y = bX + \varepsilon$;

3) $Y = a + bX$;

4) $Y = bX$,

где Y — зависимая переменная, X — регрессор, ε — случайное возмущение, a, b — параметры модели.

8. Уровнем значимости называется:

- совокупность значений критерия проверки, при которых нулевую гипотезу не отклоняют;
- совокупность значений критерия проверки, при которых нулевую гипотезу отклоняют;
- вероятность отвергнуть правильную нулевую гипотезу.

9. Выбрать правильный ответ.

Регрессор X спецификации $Y = a + bX + \varepsilon$ значим, если:

$$1) |t| = \left| \frac{\hat{b}}{s_b} \right| > t_{\text{кр}}; \quad 2) |t| = \left| \frac{\hat{b}}{s_b} \right| < t_{\text{кр}};$$

где t — статистика Стьюдента, $t_{\text{кр}}$ — критическое значение статисти-

стики, \hat{b} — оценка параметра b , s_b — стандартная ошибка оценки b

10. Установить соответствие:

| Структура спецификации парной регрессии | Выполнение равенства $TSS = RSS + ESS$ |
|--|--|
| 1) уравнение регрессии включает константу | а) не выполняется; |
| 2) уравнение регрессии включает только константу | б) выполняется. |
| 3) уравнение регрессии не включает константу | |

11. Остаточная сумма квадратов в модели парной регрессии равна нулю в том случае, когда:

- значения Y , рассчитанные по уравнению регрессии, равны среднему значению Y ;
- Y связан с X функционально;
- общая дисперсия Y обусловлена влиянием прочих факторов
- дисперсия X равна нулю.

12. Выбрать правильный ответ.

Для проверки значимости уравнения в целом используется статистика с распределением:

- 1) Стьюдента;
- 2) Гаусса;
- 3) Фишера;
- 4) хи-квадрат.

13. Добавление новой объясняющей переменной:

- иногда уменьшает значение коэффициента детерминации;
- не оказывает влияния на значение коэффициента детерминации;
- никогда не уменьшает значение коэффициента детерминации

14. Если коэффициент детерминации равен нулю, то F-критерий Фишера равен:

- единице;
- больше или равен единице;
- нулю.

15. С увеличением объема выборки ширина доверительного интервала индивидуального значения эндогенной переменной:

- увеличивается
- уменьшается
- не меняется
- зависит от спецификации

Перечень вопросов итоговой аттестации

1. Условия Гаусса-Маркова
2. Свойства МНК- оценок
3. Асимптотические свойства МНК оценок
4. Тестирование ограничений
5. Отбор множества объясняющих переменных
6. Тестирование невложенных моделей
7. Тестирование функциональной формы модели
8. Понятие гетероскедастичности. Последствия для МНК –оценки
9. Методы выявления и устранения гетероскедастичности
10. Понятие автокорреляции. Последствия для МНК –оценки
11. Методы выявления и устранения автокорреляции
12. Автокорреляция высоких порядков
13. Проблема эндогенности в моделях регрессии
14. Метод инструментальных переменных
15. Оценивание систем одновременных уравнений
16. Структурная и приведенная форма системы одновременных уравнений
17. Двухшаговый метод наименьших квадратов
18. Метод максимального правдоподобия
19. Процессы авторегрессии и скользящего среднего
20. Стационарность и автокорреляционная функция
21. Стационарность и единичные корни
22. Тестирование единичных корней
23. Тестирование единичных корней в моделях авторегрессии высоких порядков
24. ARCH и GARCH модели
25. Модель частичной корректировки
26. Модель адаптивных ожиданий
27. Векторные модели авторегрессии
28. Коинтеграция

Учебное издание

**Светлана Алексеевна Балашова
Инна Васильевна Лазанюк**

ЭКОНОМЕТРИКА В ЗАДАЧАХ И РЕШЕНИЯХ

Редактор *Ж.В. Медведева*
Технический редактор *Е.Н. Собанина*
Компьютерная верстка *М.Н. Заикина*
Дизайн обложки *Ю.Н. Ефремова*

Подписано в печать 21.03.17 г. Формат 60×84/16. Печать офсетная.
Усл. печ. л. 10,7. Тираж 500 экз. Заказ 259

Российский университет дружбы народов
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3

Типография РУДН
115419, ГСП-1, г. Москва, ул. Орджоникидзе, д. 3, тел. (495) 952-04-41

Для заметок

50,000 сум

Для заметок

Директор филиала
И. И. Иванов

Иванов И. И.
Иванов И. И.
Иванов И. И.

50 000 руб



Балашова Светлана Алексеевна, кандидат физико-математических наук, доцент кафедры экономико-математического моделирования Российского университета дружбы народов. Автор более 70 научных работ, ряда учебников и учебных пособий, в том числе по эконометрике и финансовой математике.



Лазанюк Инна Васильевна, кандидат экономических наук, доцент кафедры экономико-математического моделирования Российского университета дружбы народов. Автор более 30 публикаций. Сфера научных интересов - анализ развития сектора информационно-коммуникационных технологий с использованием эконометрических методов.



9 785209 07893 7