

РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ

СЕРИЯ «АКАДЕМИЧЕСКИЙ УЧЕБНИК»

В.П. Носко

Эконометрика

Книга вторая

Часть 3

*Системы одновременных уравнений, панельные данные, модели
с дискретными и ограниченными объясняемыми переменными*

Часть 4

*Временные ряды: дополнительные главы.
Модель стохастической границы*

Рекомендовано

*Федеральным государственным бюджетным учреждением
высшего профессионального образования
«Российская академия народного хозяйства и государственной службы
при Президенте Российской Федерации»
в качестве учебника для студентов высших учебных заведений,
обучающихся по экономическим специальностям*

МОСКВА

ИЗДАТЕЛЬСКИЙ ДОМ «ДЕЛО»

2011

УДК 330.43(075.8)
ББК 65в6я73
Н84

Р е ц е н з е н т:

И.И. Елисеева, доктор экономических наук,
профессор, член-корреспондент РАН, заслуженный деятель науки РФ,
зав. кафедрой статистики и эконометрики Санкт-Петербургского
государственного университета экономики и финансов

А В Т О Р:

Носко Владимир Петрович,
кандидат физико-математических наук,
старший научный сотрудник механико-математического факультета
МГУ им. М.В. Ломоносова, зав. кафедрой эконометрики
и математической экономики РАНХиГС.
Преподает эконометрику с 1994 г., читает курсы лекций
в МГУ им. М.В. Ломоносова, в Российской академии
народного хозяйства и государственной службы при Президенте РФ
и в Институте экономической политики им. Е.Т. Гайдара.
Автор более 60 научных работ

Носко В.П.

Н84 Эконометрика. Кн. 2. Ч. 3, 4: учебник / В.П. Носко. — М.: Изда-
тельский дом «Дело» РАНХиГС, 2011. — 576 с. (Сер. «Академиче-
ский учебник».)

ISBN 978-5-7749-0655-3

В учебнике излагаются методы эконометрического анализа — от самых простых до весьма продвинутых. В основе учебника — курсы лекций, прочитанные автором в Институте экономической политики им. Е.Т. Гайдара, на механико-математическом факультете Московского государственного университета им. М.В. Ломоносова и на экономическом факультете РАНХиГС.

Учебник состоит из двух книг (четырёх частей): в кн. 1 рассматриваются линейные модели регрессии; модели стационарных и нестационарных временных рядов, особенности регрессионного анализа для стационарных и нестационарных переменных; в кн. 2 — модели одновременных уравнений, модели с дискретными и цензурированными объясняемыми переменными, модели для анализа панельных данных; модель стохастической границы производственных возможностей, а также содержится дополнительный материал по анализу временных рядов (прогнозирование, методология векторных авторегрессий и др.). В каждой части учебника имеется словарь употребляемых в ней терминов.

Для студентов, аспирантов, преподавателей, а также для специалистов по прикладной экономике.

УДК 330.43(075.8)
ББК 65в6я73

ISBN 978-5-7749-0655-3

© ФГБОУ ВПО «Российская академия народного
хозяйства и государственной службы
при Президенте Российской Федерации», 2011

Содержание

Предисловие	6
Предисловие ко второй книге	8

Часть 3

СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ, ПАНЕЛЬНЫЕ ДАННЫЕ, МОДЕЛИ С ДИСКРЕТНЫМИ И ОГРАНИЧЕННЫМИ ОБЪЯСНЯЕМЫМИ ПЕРЕМЕННЫМИ

Р а з д е л 1. СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ	11
Тема 1.1. Идентифицируемость структурной формы системы одновременных уравнений	11
Тема 1.2. Оценивание систем одновременных уравнений	42
Р а з д е л 2. СТРУКТУРНЫЕ И ПРИВЕДЕННЫЕ ФОРМЫ МОДЕЛЕЙ КОРРЕКЦИИ ОШИБОК	86
Тема 2.1. Структурные и приведенные формы векторных авторегрессий и моделей коррекции ошибок	86
Р а з д е л 3. ПАНЕЛЬНЫЕ ДАННЫЕ	105
Тема 3.1. Панельные данные: модель пула, модель ковариационного анализа, модель кажущихся несвязанными регрессий	105
Тема 3.2. Модели с фиксированными и случайными эффектами	129
Тема 3.3. Двухнаправленные модели	156
Тема 3.4. Несбалансированные панели, эндогенные объясняющие переменные, модели с индивидуально-специфическими переменными	163
Тема 3.5. Динамические модели	173

Р а з д е л 4. МОДЕЛИ С ДИСКРЕТНЫМИ И ОГРАНИЧЕННЫМИ ОБЪЯСНЯЕМЫМИ ПЕРЕМЕННЫМИ.....	185
Тема 4.1. Модели, в которых объясняемая переменная принимает только два различных значения	185
Тема 4.2. Модели, в которых объясняемая переменная принимает несколько различных значений.....	212
Тема 4.3. Цензурированные модели регрессии (тобит-модели).....	228
Тема 4.4. Модели бинарного выбора для панельных данных	249
Тема 4.5. Тобит-модели для панельных данных.....	261
Задания для семинарских занятий, работы в компьютерном классе и для самостоятельной работы.....	271
<i>Приложение. Статистические данные к заданиям</i>	<i>306</i>
Литература	311
Глоссарий	313

Часть 4

ВРЕМЕННЫЕ РЯДЫ: ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ. МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ

Р а з д е л 5. СГЛАЖИВАНИЕ И ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ.....	327
Тема 5.1. Адаптивные методы, метод наименьших квадратов.....	327
Тема 5.2. Прогнозирование по моделям AR, MA, ARMA, ARIMA	360
Р а з д е л 6. МЕТОДОЛОГИЯ ВЕКТОРНЫХ АВТОРЕГРЕССИЙ	392
Тема 6.1. Прогнозирование по модели векторной авторегрессии, проверка наличия причинности по Грейнджеру для двух и более рядов.....	392
Тема 6.2. Методология VAR.....	414
Тема 6.3. Эмпирические исследования.....	440
Тема 6.4. Нестабильные VAR.....	467
Р а з д е л 7. ТЕСТЫ НА ЕДИНИЧНЫЕ КОРНИ И НЕЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ. ДИНАМИЧЕСКИЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ	489
Тема 7.1. Тесты на единичные корни и нелинейные преобразования	489
Тема 7.2. Динамический метод наименьших квадратов для оценивания коинтегрирующего вектора системы интегрированных рядов	504

Р а з д е л 8. МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ	515
Тема 8.1. Модель стохастической границы для перекрестной выборки	515
Тема 8.2. Модели стохастической границы для панельных данных	531
Задания для семинарских занятий, работы в компьютерном классе и для самостоятельной работы	537
<i>Приложение.</i> Таблицы статистических данных к заданиям.....	558
Литература	563
Глоссарий	567
Предметный указатель	572

Предисловие

Учебник содержит изложение основ эконометрики и написан на базе курсов лекций, прочитанных автором в Институте экономической политики им. Е.Т. Гайдара, на механико-математическом факультете Московского государственного университета им. М.В. Ломоносова и на отделении экономики экономического факультета Российской академии народного хозяйства и государственной службы при Президенте РФ.

Учебник состоит из четырех частей, объединенных в две книги. В первой части изучаются линейные модели регрессии, методы статистического анализа таких моделей, методы выявления нарушений стандартных предположений, лежащих в основе статистического анализа линейных моделей, и методы коррекции статистических выводов при выявлении таких нарушений. Во второй части рассматриваются модели стационарных и нестационарных временных рядов, особенности регрессионного анализа для стационарных и нестационарных переменных, в третьей — модели одновременных уравнений, модели, объясняющие наличие или отсутствие у субъекта некоторого признака значениями тех или иных характеристик субъекта, модели с цензурированными данными, модели, служащие для описания панельных данных. Четвертая часть содержит дополнительный материал по анализу временных рядов (прогнозирование, методология векторных авторегрессий и др.), в ней также рассматривается модель стохастической границы производственных возможностей. Материал каждой части рассчитан на изучение его в течение одного семестра (2 часа лекций и 2 часа практических занятий в неделю).

Каждая часть учебника состоит из разделов, объединяющих несколько тем. В конце темы приводятся контрольные вопросы, позволяющие закрепить усвоенный материал. В каждой части имеется набор заданий для самостоятельной работы и работы в компьютерном классе под руководством преподавателя. Методические указания по выполнению практических заданий на компьютере ориентированы в основном на использование пакета эконометрического анализа *Econometric Views*, а для некоторых разделов курса — на использование пакета *Stata*. В конце каждой части приведен словарь употребляемых в ней терминов.

Для удобства читателя при первом упоминании в тексте основные термины выделяются жирным шрифтом, а в скобках приводятся их англоязычные

эквиваленты. Некоторые слова или целые предложения, требующие привлечения внимания читателя, выделены светлым курсивом.

Автор считает своим приятным долгом выразить признательность академику РАН Револьду Михайловичу Энтову и доктору экономических наук Сергею Германовичу Синельникову-Мурылеву, которые инициировали работу по написанию данного учебника и поддерживали автора на всех этапах этой продолжительной работы. В значительной мере на изложение материала повлияли заинтересованные обсуждения лекций автора по различным аспектам эконометрических исследований в коллективе Института экономики переходного периода (в настоящее время – Институт экономической политики им. Е.Т. Гайдара). Автор благодарен Марине Юрьевне Турунцевой и Илье Борисовичу Воскобойникову, которые внимательно прочитали материал, вошедший во вторую часть учебника, и сделали ряд замечаний, способствовавших улучшению изложения. Автор весьма признателен Ирине Михайловне Промахиной, апробировавшей все задания, содержащиеся в учебнике, на занятиях со студентами отделения экономики экономического факультета Академии народного хозяйства при Правительстве РФ, что позволило устранить имевшиеся неточности в формулировках заданий и в методических указаниях по их выполнению. Автор благодарен Надежде Викторовне Андриановой за тщательную правку текста при подготовке учебника к изданию.

Предисловие ко второй книге

Вторая книга учебника состоит из двух частей (части 3 и 4) и предполагает свободное владение материалом, содержащимся в первой книге. Как и в первой книге, основные акценты в изложении смещены в сторону разъяснения базовых понятий и основных процедур статистического анализа с привлечением смоделированных и реальных экономических данных.

В третьей части учебника рассматриваются методы статистического анализа моделей с дискретными объясняющими переменными, систем одновременных уравнений, панельных данных, а также структурные и приведенные формы векторных авторегрессий и моделей коррекции ошибок. Включенный в третью часть лекционный материал в основном соответствует материалу, содержащемуся в ранее изданной книге автора¹.

В разд. 1 третьей части рассматривается возможность получения подходящих оценок параметров в ситуациях, когда объясняющие переменные, входящие в уравнение регрессии, коррелированы с ошибкой в этом уравнении. Именно такое положение наблюдается в имеющих широкое применение моделях, известных как «системы одновременных уравнений». Это модели, состоящие из нескольких уравнений регрессии и такие, что переменные, выступающие объясняемыми переменными в одних уравнениях, являются объясняющими переменными в других уравнениях. Здесь основным методом оценивания параметров является метод инструментальных переменных, состоящий в «очистке» объясняющей переменной, коррелированной с ошибкой, от этой коррелированности, и в подстановке в правую часть уравнения вместо этой объясняющей переменной ее очищенного варианта. Рассматриваются и другие методы оценивания систем одновременных уравнений, обсуждаются связи между различными методами, недостатки и преимущества каждого метода.

¹ Носко В.П. Эконометрика для начинающих: дополнительные главы. М.: ИЭПП, 2005.

Раздел 2 дополняет содержащийся в разд. 11 первой книги учебника материал, касающийся моделей векторной авторегрессии и моделей коррекции ошибок для системы коинтегрированных временных рядов. Это дополнение связано с рассмотрением возможности построения и оценивания структурной формы модели коррекции ошибок.

Раздел 3 посвящен методам статистического анализа панельных данных, т.е. данных, содержащих наблюдения за некоторым достаточно большим количеством субъектов в течение некоторого достаточно небольшого количества периодов времени. Особенностью многих моделей, используемых для статистического анализа таких данных, является предположение о наличии различий между субъектами исследования, которые постоянны во времени, но которые не удается реально измерить в виде значений некоей объясняющей переменной. Такие различия специфицируются в этих моделях как фиксированные или случайные эффекты. В зависимости от пригодности той или иной интерпретации этих эффектов используются различные методы оценивания параметров модели (обычный или обобщенный метод наименьших квадратов). Метод инструментальных переменных находит новое применение в динамических моделях панельных данных (в которых в качестве объясняющих переменных в правых частях уравнения могут выступать и запаздывающие значения объясняемой переменной) и реализуется в виде обобщенного метода моментов, ставшего весьма популярным в последние годы.

В разд. 4 обсуждаются особенности статистического анализа моделей, в которых объясняемая переменная имеет лишь конечное количество возможных значений или только частично наблюдаема. При оценивании этих моделей на первый план выступает метод максимального правдоподобия. Сначала рассматриваются модели бинарного выбора с двумя значениями объясняющей переменной (пробит, логит, гомпит) и модели с несколькими значениями объясняющей переменной (порядковая пробит-модель, мультиномиальная модель), а затем — модели с частично наблюдаемой (цензурированной) объясняемой переменной. При этом цензурирование может определяться значениями как самой объясняемой переменной (модель тобит I), так и некоторой дополнительной функции полезности (модель тобит II). В заключительной части этого раздела рассматриваемые модели (пробит, логит, тобит) распространяются на случай панельных данных.

В четвертой части учебника рассматриваются некоторые методы прогнозирования временных рядов, методология век-

торных авторегрессий (*VAR*) и ее применение к анализу систем макроэкономических рядов (построение структурных *VAR*, построение функций импульсного отклика и декомпозиций дисперсий ошибок прогнозов, проверка на отсутствие/наличие причинности по Грейнджеру и блочную экзогенность группы переменных в модели *VAR*), влияние нелинейных преобразований ряда на результаты проверки гипотезы единичного корня, применение динамического метода наименьших квадратов (*DOLS*) для оценивания коинтеграционного соотношения в системе рядов, имеющих порядки интегрированности 1 и 2. Заключительный раздел этой части посвящен модели стохастической границы производственных возможностей, которая не только применима к анализу эффективности производственных предприятий, но и используется, например, для анализа эффективности банков и эффективности сбора налогов.

Материал каждой части разбит на разделы, объединяющие несколько тем. По каждой теме кроме лекционного материала приведены контрольные вопросы, позволяющие провести контроль усвоения лекционного материала, а также задания для самостоятельной работы и практических занятий, которые выполняются студентами в компьютерном классе под руководством преподавателя. Методические указания к выполнению заданий на компьютере ориентированы на использование пакетов эконометрического анализа *Econometric Views* (*EViews*) и *Stata*.

Для удобства читателя при первом упоминании в тексте те или иные термины выделены жирным шрифтом, а в скобках приведены их англоязычные эквиваленты. В конце каждой части есть словарь употребляемых в ней терминов.

ЧАСТЬ 3
СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ,
ПАНЕЛЬНЫЕ ДАННЫЕ,
МОДЕЛИ С ДИСКРЕТНЫМИ И ОГРАНИЧЕННЫМИ
ОБЪЯСНЯЕМЫМИ ПЕРЕМЕННЫМИ

Р а з д е л 1

СИСТЕМЫ
ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

_____ Тема 1.1 _____
ИДЕНТИФИЦИРУЕМОСТЬ СТРУКТУРНОЙ ФОРМЫ
СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

В разд. 6 первой части учебника уже рассматривалась система одновременных уравнений, соответствующая модели замкнутой экономики без правительства. В этой модели уравнение

$$C_t = \alpha + \beta Y_t + \varepsilon_t,$$

где C_t — реальное потребление на душу населения;

Y_t — реальный доход на душу населения;

ε_t — случайная ошибка,

дополняется соотношением

$$Y_t = C_t + I_t,$$

где I_t — реальные инвестиции на душу населения.

Это приводит к системе уравнений

$$\begin{cases} C_t = \alpha + \beta Y_t + \varepsilon_t \\ Y_t = C_t + I_t. \end{cases}$$

О такой системе уравнений говорят как о **структурной форме системы одновременных уравнений** (*structural form of simultaneous equations*), подразумевая под этим, что она представляет в явном виде взаимные связи между входящими в модель переменными, показывает, как эти переменные взаимо-

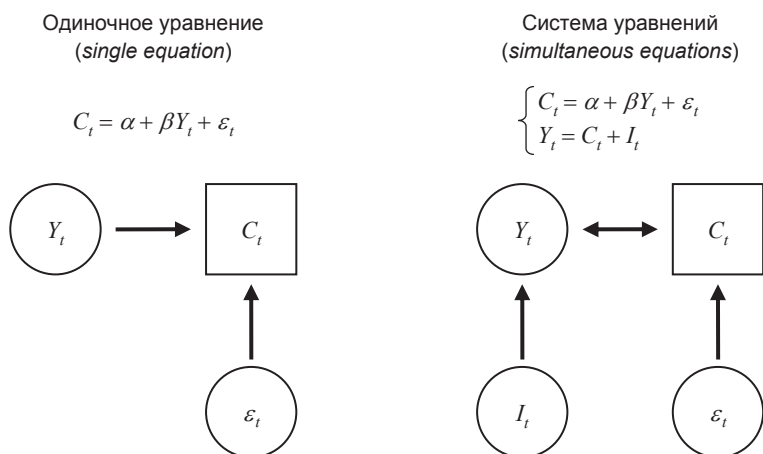


Рис. 1.1

действуют друг с другом (в данном случае Y_t воздействует на C_t , а C_t воздействует на Y_t — см. рис. 1.1). В структурной форме модели одновременных уравнений переменная, являющаяся объясняемой переменной в одном из уравнений, может входить в другое уравнение в качестве объясняющей переменной.

Выражая из этой системы C_t и Y_t через I_t , получим **приведенную форму системы одновременных уравнений** (*reduced form of simultaneous equations*) в виде:

$$\begin{cases} C_t = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I_t + \frac{1}{1-\beta} \varepsilon_t \\ Y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I_t + \frac{1}{1-\beta} \varepsilon_t, \end{cases}$$

или

$$\begin{cases} C_t = \tilde{\alpha} + \tilde{\beta} I_t + \tilde{\varepsilon}_t \\ Y_t = \tilde{\gamma} + \tilde{\delta} I_t + \tilde{\varepsilon}_t, \end{cases}$$

где

$$\tilde{\alpha} = \tilde{\gamma} = \frac{\alpha}{1-\beta}, \quad \tilde{\beta} = \frac{\beta}{1-\beta}, \quad \tilde{\delta} = \frac{1}{1-\beta},$$

$$\tilde{\varepsilon}_t = \frac{\varepsilon_t}{1-\beta}, \quad E(\tilde{\varepsilon}_t) = 0, \quad D(\tilde{\varepsilon}_t) = \sigma_{\varepsilon}^2 = \frac{\sigma_{\varepsilon}^2}{(1-\beta)^2}.$$

В такой форме взаимодействие между C_t и Y_t уже не представлено в явном виде, однако коэффициенты приведенной формы отражают итог взаимодействия этих переменных.

В разд. 6 части 1 предполагали, что $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$ и что для каждого t случайные величины I_t и ε_t независимы. Тогда из второго уравнения приведенной формы находим:

$$Cov(Y_t, \varepsilon_t) = \frac{1}{1-\beta} Cov(\varepsilon_t, \varepsilon_t) = \frac{\sigma^2}{1-\beta} \neq 0,$$

так что в исходном уравнении для C_t объясняющая переменная Y_t коррелирована с ошибкой. При этом для оценки $\hat{\beta}$ коэффициента β , получаемой по n наблюдениям применением метода наименьших квадратов к исходному уравнению, имеем:

$$p \lim_{n \rightarrow \infty} \hat{\beta} = \beta + \frac{Cov(Y, \varepsilon_t)}{D(Y_t)},$$

где

$$D(Y) = \frac{1}{(1-\beta)^2} (D(I_t) + \sigma^2),$$

и

$$p \lim_{n \rightarrow \infty} \hat{\beta} = \beta + (1-\beta) \frac{\sigma^2}{D(I_t) + \sigma^2}.$$

Поскольку $\sigma^2 > 0$ и в модели Кейнса $0 < \beta < 1$, то $\hat{\beta}$ переоценивает значение нормы потребления.

Учитывая это осложнение, в разд. 6 ч. 1 рассматривалось опосредованное получение оценок коэффициентов структурной формы через оценки коэффициентов уравнений приведенной формы. Применив метод наименьших квадратов к первому уравнению приведенной формы, найдем оценки коэффициентов $\tilde{\alpha}$ и $\tilde{\beta}$ и оценку дисперсии σ_{ε}^2 . После этого можно найти оценки для параметров исходного уравнения, используя соотношения

$$\beta = \frac{\tilde{\beta}}{1 + \tilde{\beta}}, \quad \alpha = \frac{\tilde{\alpha}}{1 + \tilde{\beta}}, \quad \sigma_{\varepsilon}^2 = \frac{\sigma_{\tilde{\varepsilon}}^2}{(1 + \tilde{\beta})^2}.$$

Таким образом, структурная форма восстанавливается по первому уравнению приведенной формы. Второе уравнение оказывается в этом плане избыточным. Но, используя одно это уравнение, также можно восстановить структурную форму. Применив метод наименьших квадратов к этому уравнению, найдем оценки коэффициентов $\tilde{\gamma}$ и $\tilde{\delta}$ и оценку дисперсии σ_{ε}^2 . После этого можно найти оценки для параметров исходного уравнения, используя соотношения

$$\beta = \frac{\tilde{\delta} - 1}{\tilde{\delta}}, \quad \alpha = \frac{\tilde{\gamma}}{\tilde{\delta}}, \quad \sigma_{\varepsilon}^2 = \frac{\sigma_{\tilde{\varepsilon}}^2}{\tilde{\delta}^2}.$$

Возникает вопрос: совпадают ли результаты восстановления параметров структурной формы, полученные по двум различным уравнениям приведенной формы?

Используя выражения для α , β и σ_ε^2 , найдем:

$$\frac{\tilde{\gamma}}{\tilde{\delta}} = \frac{\tilde{\alpha}}{1 + \tilde{\beta}}, \quad \frac{\tilde{\delta} - 1}{\tilde{\delta}} = \frac{\tilde{\beta}}{1 + \tilde{\beta}}, \quad \frac{\sigma_\varepsilon^2}{\tilde{\delta}^2} = \frac{\sigma_\varepsilon^2}{(1 + \tilde{\beta})^2}.$$

Таким образом, зная истинные значения параметров уравнений приведенной формы, однозначно восстанавливаем по ним значения параметров структурной формы. Однако это если знаем истинные значения параметров уравнений приведенной формы, но они нам не известны, их приходится оценивать по имеющимся статистическим данным. При этом оценки параметров структурной формы, полученные с помощью оценок коэффициентов для разных уравнений приведенной формы, могут в общем случае отличаться друг от друга. И это связано с тем, что количество параметров приведенной формы больше минимально необходимого их числа для восстановления значений параметров структурной формы.

Между тем в примере 6.2.2 разд. 6 ч. 1 оценки параметров структурной формы, полученные с помощью оценок коэффициентов для разных уравнений приведенной формы, оказались одинаковыми. Почему это произошло?

Использование оценок коэффициентов первого уравнения приведенной формы приводит к следующим оценкам для α и β :

$$\hat{\beta} = \frac{\hat{\tilde{\beta}}}{1 + \hat{\tilde{\beta}}}, \quad \hat{\alpha} = \frac{\hat{\tilde{\alpha}}}{1 + \hat{\tilde{\beta}}};$$

использование оценок коэффициентов второго уравнения приведенной формы приводит к оценкам:

$$\hat{\beta} = \frac{\hat{\tilde{\delta}} - 1}{\hat{\tilde{\delta}}}, \quad \hat{\alpha} = \frac{\hat{\tilde{\gamma}}}{\hat{\tilde{\delta}}}.$$

Полученные этими двумя способами оценки коэффициента β будут совпадать, если $\frac{\hat{\tilde{\beta}}}{1 + \hat{\tilde{\beta}}} = \frac{\hat{\tilde{\delta}} - 1}{\hat{\tilde{\delta}}}$, т.е. если $\hat{\tilde{\delta}} = \hat{\tilde{\beta}} + 1$. В то же время

$$C_t = \tilde{\alpha} + \tilde{\beta} I_t + \tilde{\varepsilon}_t \rightarrow \hat{\tilde{\beta}} = \frac{\text{Cov}(C_t, I_t)}{\text{Var}(I_t)},$$

$$Y_t = \tilde{\gamma} + \tilde{\delta} I_t + \tilde{\varepsilon}_t \rightarrow \hat{\tilde{\delta}} = \frac{\text{Cov}(Y_t, I_t)}{\text{Var}(I_t)} = \frac{\text{Cov}(C_t + I_t, I_t)}{\text{Var}(I_t)} = \frac{\text{Cov}(C_t, I_t)}{\text{Var}(I_t)} + 1 = \hat{\tilde{\beta}} + 1.$$

Таким образом, указанное соотношение выполняется, поэтому оценки коэффициента β , полученные с использованием оценок коэффициентов для разных уравнений приведенной формы, совпадают. Аналогичный результат

справедлив и для оценок коэффициента α , полученных с использованием оценок коэффициентов для разных уравнений приведенной формы.

Рассмотрим теперь простейшую модель рынка некоторого товара:

$$\begin{cases} Q^d = a_0 + a_1 P \\ Q^s = b_0 + b_1 P \\ Q^s = Q^d, \end{cases}$$

где Q^s — предложение товара (*supply*);

Q^d — спрос на товар (*demand*);

P — цена единицы товара (*price*);

$a_1 < 0$, $b_1 > 0$.

Если правила определения объемов предложения и спроса известны, т.е. известны коэффициенты a_0 , a_1 , b_0 и b_1 , то при отсутствии флуктуаций равновесное решение для цены P и спроса Q находится без труда. Здесь в наличии система двух линейных уравнений с двумя неизвестными P и Q :

$$\begin{cases} Q - a_1 P = a_0 \\ Q - b_1 P = b_0, \end{cases}$$

решением которой является пара значений

$$Q = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1}, \quad P = \frac{a_0 - b_0}{b_1 - a_1},$$

которые есть координаты точки пересечения прямых спроса и предложения (прямые d и s на рис. 1.2). При $a_0 > b_0$ оба эти значения положительны.

Предположим, что спрос подвержен случайным флуктуациям, изменяющим a_0 до значения $(a_0 + u_t)$ в t -м наблюдении, а предложение подвержено флуктуациям, изменяющим в t -м наблюдении b_0 до значения $(b_0 + v_t)$, так что,

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t. \end{cases}$$

Тогда каждому t соответствуют свои равновесные значения цены P_t и спроса Q_t , являющиеся решениями системы

$$\begin{cases} Q_t - a_1 P_t = a_0 + u_t \\ Q_t - b_1 P_t = b_0 + v_t. \end{cases}$$

Поскольку значения P_t и Q_t определяются внутри системы, о переменных P_t и Q_t говорят как об **эндогенных** (*endogenous*) переменных. Их значения в t -м наблюдении определяются коэффициентами a_0 , a_1 , b_0 , b_1 и внешними случайными воздействиями («шоками») u_t , v_t .

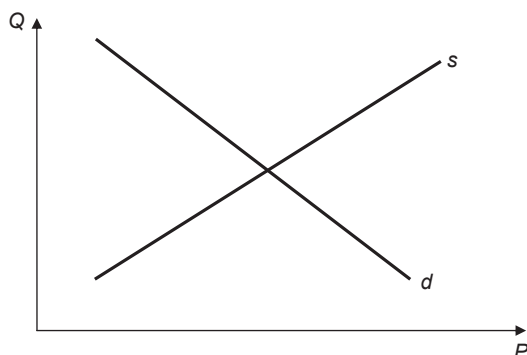


Рис. 1.2

Обращаясь к оцениванию уравнения $Q_t = \alpha + \beta P_t + \varepsilon_t$ по наблюдаемым значениям P_t и Q_t , $t = 1, \dots, n$, статистик даже не знает, что он оценивает: прямую спроса или прямую предложения. Так, оценивание методом наименьших квадратов линейной модели зависимости потребления свинины на душу населения США от оптовых цен на свинину по годовым данным за период с 1948 по 1961 г. дает результаты, приведенные в табл. 1.1.

Таблица 1.1

Оценивание зависимости потребления свинины на душу населения США от оптовых цен на нее по годовым данным с 1948 по 1961 г.

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
1	77.484	13.921	5.556	0.0001
Цена	-24.775	29.794	-0.832	0.4219

Хотя формально отрицательное значение оценки коэффициента при цене говорит о том, что мы имеем дело с уравнением спроса, эта оценка оказывается статистически незначимой при любом разумном выборе уровня значимости, так что в доверительный интервал для данного коэффициента попадают как отрицательные, так и положительные значения.

Получив выражение для P_t из второго уравнения системы и подставив это выражение вместо P_t в первое уравнение, найдем:

$$Q_t = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} + \frac{b_1 u_t - a_1 v_t}{b_1 - a_1} = \pi_1 + w_{t1}.$$

Выражение для Q_t из первого уравнения системы подставим во второе уравнение, в результате получим:

$$P_t = \frac{a_0 - b_0}{b_1 - a_1} + \frac{u_t - v_t}{b_1 - a_1} = \pi_2 + w_{t2}.$$

Шоки u_t, v_t сдвигают точку равновесия. Сдвиг вдоль оси ординат:

$$\Delta Q = \frac{b_1 u_t - a_1 v_t}{b_1 - a_1};$$

сдвиг вдоль оси абсцисс:

$$\Delta P = \frac{u_t - v_t}{b_1 - a_1}.$$

Направление сдвига:

$$\frac{\Delta Q}{\Delta P} = \frac{b_1 u_t - a_1 v_t}{u_t - v_t}.$$

Если имеется только сдвиг спроса ($v_t = 0$), то

$$\frac{\Delta Q}{\Delta P} = \frac{b_1 u_t}{u_t} = b_1 \text{ — сдвиг вдоль прямой предложения.}$$

Если имеется только сдвиг предложения ($u_t = 0$), то

$$\frac{\Delta Q}{\Delta P} = \frac{a_1 v_t}{v_t} = a_1 \text{ — сдвиг вдоль прямой спроса.}$$

ПРИМЕР 1.1.1

На рис. 1.3 показана реализация следующего процесса порождения данных:

$$DGP: Q_t = 20 - 0.5P_t + u_t, \text{ (спрос)}$$

$$Q_t = 10 + 0.5P_t + v_t, \text{ (предложение)}$$

$$u_t, v_t \sim i.i.d. N(0, 1), \quad t = 1, 2, \dots, 100.$$

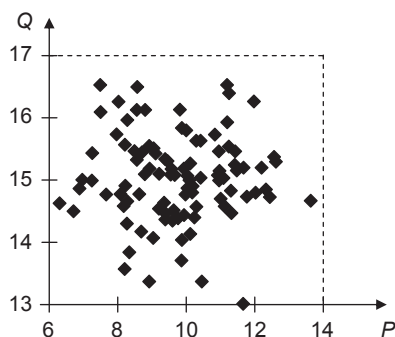


Рис. 1.3

Рассеяние происходит вокруг точки $P = 10, Q = 15$; направления прямых спроса и предложения на диаграмме не выявляются. ■

Пара уравнений

$$\begin{cases} Q_t = \pi_1 + w_{t1} \\ P_t = \pi_2 + w_{t2} \end{cases}$$

представляет приведенную форму системы. Заметим, что здесь существует корреляция между ошибками в разных уравнениях приведенной формы при одном и том же t , даже если ошибки в структурной системе не коррелированы: в последнем случае

$$\text{Cov}(w_{t1}, w_{t2}) = \text{Cov}\left(\frac{b_1 u_t - a_1 v_t}{b_1 - a_1}, \frac{u_t - v_t}{b_1 - a_1}\right) = \frac{1}{(b_1 - a_1)^2} (b_1 D(u_t) + a_1 D(v_t)).$$

Однако, поскольку в правых частях обоих уравнений приведенной формы одни и те же объясняющие переменные (в данном случае одна объясняющая переменная, тождественно равная 1), оценки коэффициентов приведенной формы, учитывающие наличие корреляции между ошибками в разных уравнениях (т.е. оценки, получаемые применением обобщенного метода наименьших квадратов — *GLS*), численно совпадают с оценками, получаемыми раздельным оцениванием обоих уравнений обычным методом наименьших квадратов — *OLS*¹. Получив таким образом оценки $\hat{\pi}_1$, $\hat{\pi}_2$, тем самым найдем оценки для дробей $\frac{a_0 b_1 - a_1 b_0}{b_1 - a_1}$ и $\frac{a_0 - b_0}{b_1 - a_1}$. Но этих двух оценок недостаточно для восстановления по ним значений 4 коэффициентов a_0 , a_1 , b_0 , b_1 структурных уравнений, так что здесь имеет место **недоидентифицированность** (*underidentification*) или **неидентифицируемость** (*non-identification*) структурной формы системы.

Включим в правую часть уравнения спроса доход (например, совокупный располагаемый доход) Y_t , система примет вид:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t. \end{cases}$$

Действуя аналогично предыдущему случаю, находим приведенную форму:

$$Q_t = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} + \frac{a_2 b_1}{b_1 - a_1} Y_t + \frac{b_1 u_t - a_1 v_t}{b_1 - a_1} = \pi_{11} + \pi_{21} Y_t + w_{t1},$$

$$P_t = \frac{a_0 - b_0}{b_1 - a_1} + \frac{a_2}{b_1 - a_1} Y_t + \frac{u_t - v_t}{b_1 - a_1} = \pi_{12} + \pi_{22} Y_t + w_{t2}.$$

В приведенной форме 4 коэффициента, тогда как в структурной форме — 5 коэффициентов. Поэтому и здесь невозможно восстановить все коэффици-

¹ Этот факт был установлен в работе (Dwivedi, Srivastava, 1978).

енты структурной формы по коэффициентам приведенной формы. Однако кое-что сделать все же можно.

Прежде всего заметим, что

$$\frac{\pi_{21}}{\pi_{22}} = b_1, \quad \pi_{11} - b_1 \pi_{12} = b_0,$$

так что коэффициенты уравнения предложения восстанавливаются по коэффициентам приведенной формы. В то же время для восстановления коэффициентов уравнения спроса остается только два выражения, так что восстановить однозначно их значения не представляется возможным. Таким образом, здесь уравнение предложения идентифицируемо, а уравнение спроса неидентифицируемо: система **частично идентифицируема** (*partially identified*).

ПРИМЕР 1.1.2

На рис. 1.4 показана реализация следующего процесса порождения данных:

$$DGP: Q_t = 20 - 0.5P_t + 0.1Y_t + u_t, \quad (\text{спрос})$$

$$Q_t = 10 + 0.5P_t + v_t, \quad (\text{предложение})$$

$$u_t, v_t \sim i.i.d. N(0, 1), \quad Y_t = 1000 + 100\eta_t, \quad t = 1, 2, \dots, 100,$$

где $\eta_1, \eta_2, \dots, \eta_{100}$ — случайная выборка из равномерного распределения в интервале $(0, 1)$.

В этом случае на диаграмме выявляется направление прямой предложения. ■

Пополним теперь уравнение предложения. Если рассматриваемый товар — продукт сельскохозяйственного производства, то в качестве объясняющей переменной в правую часть этого уравнения естественно включить какой-либо подходящий индекс климатических условий (например, среднее количество осадков в соответствующий период R_t). Тогда получим систему:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + v_t \end{cases}$$

с 6 коэффициентами. Найдем приведенную форму этой системы:

$$\begin{cases} Q_t = \pi_{11} + \pi_{21} Y_t + \pi_{31} R_t + w_{t1} \\ P_t = \pi_{12} + \pi_{22} Y_t + \pi_{32} R_t + w_{t2}, \end{cases}$$

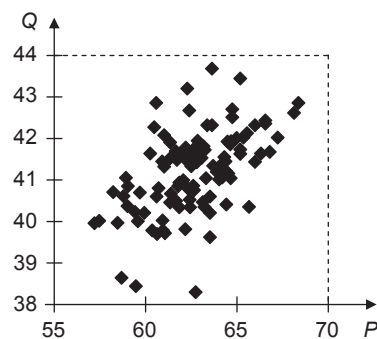


Рис. 1.4

применяя матричный подход, обычно используемый для анализа и оценивания систем одновременных уравнений. Заметим, что структурную форму системы можно записать в виде:

$$\begin{cases} Q_t - a_1 P_t = a_0 + a_2 Y_t + u_t \\ Q_t - b_1 P_t = b_0 + b_2 R_t + v_t, \end{cases}$$

или

$$(Q_t, P_t) \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix} = (1, Y_t, R_t) \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix} + (u_t, v_t),$$

а приведенную форму — в виде:

$$(Q_t, P_t) = (1, Y_t, R_t) \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} + (w_{t1}, w_{t2}).$$

Приведенную форму системы получаем из структурной формы, умножая обе части предпоследнего уравнения на матрицу, обратную матрице, стоящей в левой части:

$$\begin{aligned} (Q_t, P_t) &= \left\{ (1, Y_t, R_t) \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix} + (u_t, v_t) \right\} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1} = \\ &= (1, Y_t, R_t) \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1} + (u_t, v_t) \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1} = \\ &= (1, Y_t, R_t) \Pi + (u_t, v_t) \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1}, \end{aligned}$$

где Π — матрица коэффициентов приведенной формы,

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1}.$$

Но

$$\begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1} = \frac{1}{a_1 - b_1} \begin{pmatrix} -b_1 & -1 \\ a_1 & 1 \end{pmatrix},$$

так что

$$Q_t = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} + \frac{a_2 b_1}{b_1 - a_1} Y_t - \frac{a_1 b_2}{b_1 - a_1} R_t + \frac{b_1 u_t - a_1 v_t}{b_1 - a_1},$$

$$P_t = \frac{a_0 - b_0}{b_1 - a_1} + \frac{a_2}{b_1 - a_1} Y_t - \frac{b_2}{b_1 - a_1} R_t + \frac{u_t - v_t}{b_1 - a_1}.$$

Поскольку матрица коэффициентов приведенной формы получается как

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1},$$

то

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix}.$$

Это дает нам 6 уравнений для восстановления 6 коэффициентов структурной формы по коэффициентам приведенной формы:

$$\pi_{11} - \pi_{12}a_1 = a_0, \quad \pi_{11} - \pi_{12}b_1 = b_0,$$

$$\pi_{21} - \pi_{22}a_1 = a_2, \quad \pi_{21} - \pi_{22}b_1 = 0,$$

$$\pi_{31} - \pi_{32}a_1 = 0, \quad \pi_{31} - \pi_{32}b_1 = b_2.$$

Из этих уравнений находим:

$$a_1 = \frac{\pi_{31}}{\pi_{32}}, \quad b_1 = \frac{\pi_{21}}{\pi_{22}},$$

$$a_2 = \pi_{21} - \pi_{22} \frac{\pi_{31}}{\pi_{32}}, \quad b_2 = \pi_{31} - \pi_{32} \frac{\pi_{21}}{\pi_{22}},$$

$$a_0 = \pi_{11} - \frac{\pi_{12}\pi_{31}}{\pi_{32}}, \quad b_0 = \pi_{11} - \frac{\pi_{12}\pi_{21}}{\pi_{22}}.$$

Таким образом, здесь идентифицируемы и уравнение предложения, и уравнение спроса.

Рассмотрим теперь систему, в которой доход не включен в уравнение спроса, а уравнение предложения дополнено еще одной объясняющей переменной — S_t (пусть это будет, скажем, индекс стоимости горюче-смазочных материалов, используемых при производстве соответствующего продукта сельского хозяйства). Тогда система принимает вид:

$$\begin{cases} Q_t - a_1 P_t = a_0 + u_t \\ Q_t - b_1 P_t = b_0 + b_2 R_t + b_3 S_t + v_t, \end{cases}$$

или

$$(Q_t, P_t) \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix} = (1, R_t, S_t) \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix} + (u_t, v_t).$$

Матрица коэффициентов приведенной формы

$$\begin{cases} Q_t = \pi_{11} + \pi_{21}R_t + \pi_{31}S_t + w_{t1} \\ P_t = \pi_{12} + \pi_{22}R_t + \pi_{32}S_t + w_{t2} \end{cases}$$

получается как

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1},$$

так что

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix}.$$

Здесь опять получаем 6 уравнений для восстановления 6 коэффициентов структурной формы:

$$\pi_{11} - \pi_{12}a_1 = a_0, \quad \pi_{11} - \pi_{12}b_1 = b_0,$$

$$\pi_{21} - \pi_{22}a_1 = 0, \quad \pi_{21} - \pi_{22}b_1 = b_2,$$

$$\pi_{31} - \pi_{32}a_1 = 0, \quad \pi_{31} - \pi_{32}b_1 = b_3.$$

Однако ситуация с идентифицируемостью резко отличается от предыдущего случая.

Для коэффициентов первого структурного уравнения (уравнения спроса) находим:

$$a_0 = \pi_{11} - \pi_{12}a_1, \quad a_1 = \frac{\pi_{21}}{\pi_{22}}, \quad a_1 = \frac{\pi_{31}}{\pi_{32}},$$

так что для восстановления коэффициента a_1 имеем два соотношения. Поскольку

$$\begin{aligned} \Pi &= \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}^{-1} = \\ &= \frac{1}{a_1 - b_1} \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix} \begin{pmatrix} -b_1 & -1 \\ a_1 & 1 \end{pmatrix} = \frac{1}{a_1 - b_1} \begin{pmatrix} -a_0b_1 + a_1b_0 & -a_0 + b_0 \\ a_1b_2 & b_2 \\ a_1b_3 & b_3 \end{pmatrix}, \end{aligned}$$

то $\frac{\pi_{21}}{\pi_{22}} = \frac{\pi_{31}}{\pi_{32}}$, так что коэффициент a_1 восстанавливается однозначно, если

известны точно коэффициенты приведенной формы. Если же производим свободное оценивание уравнений приведенной формы по имеющимся статистическим данным, не принимая во внимание ограничения на их коэффици-

енты, накладываемые структурной формой (в данном случае ограничения $\frac{\pi_{21}}{\pi_{22}} - \frac{\pi_{31}}{\pi_{32}} = 0$), то на основании оценок $\hat{\pi}_{21}$, $\hat{\pi}_{22}$, $\hat{\pi}_{31}$, $\hat{\pi}_{32}$ получим, как пра-

вило, различные значения отношений $\frac{\hat{\pi}_{21}}{\hat{\pi}_{22}}$ и $\frac{\hat{\pi}_{31}}{\hat{\pi}_{32}}$, два варианта оценок для

коэффициента a_1 и соответственно два варианта для коэффициента a_0 . Таким образом, уравнение спроса оказывается **сверхидентифицируемым** (*overidentified*) — для восстановления его коэффициентов имеется количество соотношений, большее минимально необходимого.

Для коэффициентов второго структурного уравнения (уравнения предложения) также имеем три соотношения:

$$\pi_{11} - \pi_{12}b_1 = b_0, \quad \pi_{21} - \pi_{22}b_1 = b_2, \quad \pi_{31} - \pi_{32}b_1 = b_3.$$

Однако во втором структурном уравнении 4 неизвестных коэффициента — b_0 , b_1 , b_2 , b_3 , и этих трех соотношений недостаточно для их восстановления — этим соотношениям удовлетворяет бесконечно много наборов значений b_0 , b_1 , b_2 , b_3 .

Таким образом, для рассмотренной модели:

- уравнение спроса сверхидентифицировано;
- уравнение предложения недоидентифицировано.

Последний пример показывает, что имеет смысл говорить не только об идентифицируемости или неидентифицируемости системы в целом, но и об идентифицируемости или неидентифицируемости отдельных уравнений системы.

Проверка выполнения условий идентифицируемости структурных уравнений

При рассмотрении условий идентифицируемости отдельных структурных уравнений, входящих в систему одновременных уравнений¹, прежде всего предполагается, что переменные, задействованные в системе, подразделяются на три типа:

- **эндогенные переменные** (*endogenous variables*);
- **экзогенные переменные** (*exogenous variables*);
- **предопределенные переменные** (*predetermined variables*).

Значения эндогенных переменных определяются внутри рассматриваемой системы; эндогенная переменная, входящая в i -е уравнение системы, коррелирована с ошибкой в этом уравнении. Значения экзогенных переменных

¹ В смысле возможности восстановления коэффициентов структурных уравнений на основе коэффициентов уравнений приведенной формы.

определяются вне рассматриваемой системы; экзогенные переменные не коррелированы с ошибками во всех уравнениях системы для всех моментов времени. Понятие предопределенной переменной относится к системам, в которых наблюдения производятся в последовательные моменты времени. Значения предопределенных переменных, как и значения эндогенных переменных, определяются внутри системы. Однако значение в момент t предопределенной переменной, входящей в i -е уравнение, не должно быть коррелированным со значениями ошибки в этом уравнении, соответствующими моментам $t, t + 1, \dots$ Например, в системе

$$\begin{cases} Q_t = a_1 P_t + a_2 Q_{t-1} + u_t \\ P_t = b_1 Q_{t-1} + v_t \end{cases}$$

Q_t и P_t — эндогенные переменные;

Q_{t-1} — предопределенная переменная.

Предполагается, что:

- система состоит из g уравнений, в каждое из которых входит хотя бы одна эндогенная переменная;
- в систему входят g эндогенных переменных;
- в систему входят K экзогенных и предопределенных переменных;
- каждое из g уравнений **нормировано** (*normalized*), так что коэффициент при одной из эндогенных переменных, входящих в уравнение, равен 1.

(В последнем примере $g = 2, K = 1$, уравнения нормированы.)

При выводе условий идентифицируемости можно не различать предопределенные и экзогенные переменные, и для краткости будем называть их в контексте проблемы идентифицируемости предопределенными.

Если собрать все эндогенные переменные в левых частях структурных уравнений, то систему одновременных уравнений можно записать в виде:

$$\begin{cases} \gamma_{11}y_{t1} + \dots + \gamma_{g1}y_{tg} = \beta_{11}x_{t1} + \dots + \beta_{K1}x_{tK} + u_{t1} \\ \dots \\ \gamma_{1g}y_{t1} + \dots + \gamma_{gg}y_{tg} = \beta_{1g}x_{t1} + \dots + \beta_{Kg}x_{tK} + u_{tg}, \end{cases}$$

где $t = 1, \dots, n$;

y_{t1}, \dots, y_{tg} — эндогенные переменные;

x_{t1}, \dots, x_{tK} — предопределенные переменные;

u_{t1}, \dots, u_{tg} — случайные ошибки;

γ_{ji} — в этой записи коэффициент при j -й эндогенной переменной в i -м уравнении;

β_{ji} — в этой записи коэффициент при j -й предопределенной переменной в i -м уравнении.

(Разумеется, часть коэффициентов в конкретных системах равна нулю.) Заметим, что последнюю запись можно также представить как

$$\begin{cases} y_{t1}\gamma_{11} + \dots + y_{tg}\gamma_{g1} = x_{t1}\beta_{11} + \dots + x_{tK}\beta_{K1} + u_{t1} \\ \dots \\ y_{t1}\gamma_{1g} + \dots + y_{tg}\gamma_{gg} = x_{t1}\beta_{1g} + \dots + x_{tK}\beta_{Kg} + u_{tg}, \end{cases}$$

и обозначим:

$$\Gamma = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g1} & \dots & \gamma_{gg} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{11} & \dots & \beta_{1g} \\ \vdots & \ddots & \vdots \\ \beta_{g1} & \dots & \beta_{gg} \end{pmatrix},$$

$$y_t = (y_{t1}, \dots, y_{tg}), \quad x_t = (x_{t1}, \dots, x_{tK}), \quad u_t = (u_{t1}, \dots, u_{tg})$$

(последние три вектора здесь удобнее представлять как векторы-строки). Тогда система записывается в компактном виде:

$$y_t \Gamma = x_t B + u_t, \quad t = 1, \dots, n.$$

Предполагая невырожденность матрицы Γ , так что для этой матрицы существует обратная, умножим обе части последнего уравнения на Γ^{-1} . Получим приведенную форму системы:

$$y_t = x_t B \Gamma^{-1} + u_t \Gamma^{-1} = x_t \Pi + w_t.$$

Здесь

$$\Pi = B \Gamma^{-1} = \begin{pmatrix} \pi_{11} & \dots & \pi_{1g} \\ \vdots & \ddots & \vdots \\ \pi_{K1} & \dots & \pi_{Kg} \end{pmatrix}, \quad w_t = u_t \Gamma^{-1} = (w_{t1}, \dots, w_{tg}),$$

где w_{ti} — случайная ошибка в i -м уравнении приведенной формы в момент t .

Выше уже фактически использовали это представление для получения приведенных форм систем

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + v_t \end{cases}$$

и

$$\begin{cases} Q_t - a_1 P_t = a_0 + u_t \\ Q_t - b_1 P_t = b_0 + b_2 R_t + b_3 S_t + v_t. \end{cases}$$

Следует отметить, что, даже если векторы $u_i = (u_{i1}, \dots, u_{ig})$, $i = 1, \dots, n$, взаимно независимы и имеют одинаковое g -мерное нормальное распределение с нулевым средним и ковариационной матрицей $\Sigma = \sigma^2 I_g$, где I_g — единичная матрица, векторы $w_t = (w_{t1}, \dots, w_{tg})$ могут иметь коррелированные между

собой и неодинаково распределенные компоненты. Однако это не препятствует получению оценок элементов матрицы Π обычным методом наименьших квадратов: достаточно применить этот метод отдельно к каждому уравнению приведенной системы. Это связано с тем, что в правых частях всех уравнений приведенной формы присутствует один и тот же набор объясняющих переменных (см. сделанную выше ссылку на работу (Dwivedi, Srivastava, 1978)).

Поскольку

$$\Pi = B\Gamma^{-1} = \begin{pmatrix} \pi_{11} & \dots & \pi_{1g} \\ \vdots & \ddots & \vdots \\ \pi_{K1} & \dots & \pi_{Kg} \end{pmatrix},$$

то $\Pi\Gamma = B$, и мы использовали это соотношение для восстановления коэффициентов структурных уравнений двух последних систем. Вопрос об идентифицируемости структурной формы — это вопрос о возможности однозначного восстановления всех коэффициентов структурной формы, т.е. восстановления матриц Γ и B , на основании матрицы $\Pi = B\Gamma^{-1}$. Заметим, что в совокупности матрицы Γ и B состоят из $g^2 + Kg$ элементов, тогда как в матрице Π всего Kg элементов. Поэтому однозначное восстановление коэффициентов структурной формы по коэффициентам приведенной формы невозможно без использования дополнительной информации в виде невключения в отдельные уравнения тех или иных переменных, нормировки коэффициентов, линейных ограничений на параметры структуры¹.

Если нас интересует i -е структурное уравнение, то его идентифицируемость — это возможность однозначного восстановления на основании коэффициентов приведенной формы:

Γ_i — i -го столбца матрицы Γ , который содержит коэффициенты при эндогенных переменных, входящих в i -е структурное уравнение;

B_i — i -го столбца матрицы B , который содержит коэффициенты при предопределенных переменных, входящих в i -е структурное уравнение.

При этом, по существу, достаточно иметь возможность восстановления Γ_i и B_i с точностью до умножения их на один и тот же числовой множитель: единственность достигается в этом случае указанием правила нормировки, в соответствии с которым коэффициент при определенной эндогенной переменной в i -м структурном уравнении полагается равным 1.

¹ Как будет продемонстрировано ниже в данном разделе (см. Замечание 1.1.5), коэффициенты структурной формы могут не восстанавливаться однозначно по одним только коэффициентам приведенной формы и в то же время однозначно восстанавливаться при привлечении дополнительной информации в виде ограничений на элементы ковариационной матрицы ошибок в правых частях уравнений структурной формы и использовании элементов ковариационной матрицы ошибок в правых частях уравнений приведенной формы.

Пусть матрица A размера $(g + K) \times g$ составлена из матриц Γ и B таким образом, что матрица Γ располагается над матрицей B :

$$A = \begin{bmatrix} \Gamma \\ B \end{bmatrix}.$$

Коэффициенты при g эндогенных и K предопределенных переменных в i -м структурном уравнении составляют i -й столбец α_i матрицы A .

Существенным является то обстоятельство, что коэффициенты i -го структурного уравнения не могут быть восстановлены на основании коэффициентов приведенной формы, если в это уравнение входят все (g) эндогенные и все (K) предопределенные переменные системы.

Поэтому далее будем предполагать, что на элементы вектора α_i помимо нормировочного накладываются еще и некоторые дополнительные однородные линейные ограничения в виде уравнений

$$\Phi_i \alpha_i = 0,$$

где Φ_i — матрица размера $R_i \times (g + K)$;

R_i — количество линейных ограничений.

Неспецифицированные коэффициенты i -го уравнения определяются по матрице $\Pi = B\Gamma^{-1}$ после применения правила нормировки однозначным образом тогда и только тогда, когда выполнено следующее **ранговое условие идентифицируемости** (*rank condition for identification*):

$$\text{rank}(\Phi_i A) = g - 1$$

(матрица $\Phi_i A$ имеет R_i строк и g столбцов).

Пусть A_i — матрица, получаемая из матрицы A вычеркиванием ее i -го столбца α_i , так что $A = [\alpha_i : A_i]$. Тогда

$$\text{rank}(\Phi_i A) = \text{rank}(\Phi_i [\alpha_i : A_i]) = \text{rank}(\Phi_i \alpha_i : \Phi_i A_i),$$

и поскольку $\Phi_i \alpha_i = 0$, то

$$\text{rank}(\Phi_i A) = \text{rank}(0 : \Phi_i A_i) = \text{rank}(\Phi_i A_i).$$

Но матрица $\Phi_i A_i$ имеет размер $R_i \times (g - 1)$, и, чтобы ее ранг был равен $(g - 1)$, во всяком случае необходимо выполнение следующего **порядкового условия идентифицируемости** (*order condition for identification*) i -го структурного уравнения:

$$R_i \geq g - 1.$$

Предположим, что все линейные ограничения, накладываемые на элементы столбца α_i (помимо условия нормировки), являются **исключающими ограничениями** (*exclusion restrictions*), т.е. все они состоят в приравнивании определенных элементов столбца α_i нулю и исключении из i -го уравнения g_i^*

эндогенных и K_i^* предопределенных переменных. Тогда общее количество исключенных переменных равно $(g_i^* + K_i^*)$, и необходимое условие идентифицируемости i -го структурного уравнения принимает вид:

$$g_i^* + K_i^* \geq g - 1,$$

или

$$K_i^* \geq (g - g_i^*) - 1.$$

Иначе говоря, количество предопределенных переменных в системе, не включенных в i -е структурное уравнение, должно быть не меньше количества эндогенных переменных, включенных в i -е уравнение, уменьшенного на 1.

Если в левой части i -го структурного уравнения находится единственная эндогенная переменная, то $(g - g_i^*) - 1$ есть просто количество эндогенных переменных, включенных в правую часть этого уравнения. В этом случае порядковое условие идентифицируемости i -го структурного уравнения можно сформулировать следующим образом:

Количество предопределенных переменных в системе, не включенных в i -е структурное уравнение, должно быть не меньше количества эндогенных переменных, включенных в правую часть этого уравнения.

Теперь можно охарактеризовать три ситуации, возникающие при оценивании i -го структурного уравнения:

- 1) $\text{rank } \Phi_i A < g - 1 \rightarrow i$ -е уравнение **неидентифицируемо (недоопределено — *underidentified*)**;
- 2) $\text{rank } \Phi_i A = g - 1$ и $R_i = g - 1 \rightarrow i$ -е уравнение **идентифицируемо точно (*just identified*)**;
- 3) $\text{rank } \Phi_i A = g - 1$ и $R_i > g - 1 \rightarrow i$ -е уравнение **сверхидентифицируемо (переопределено — *overidentified*)**.

В ситуации 1 не выполнено необходимое условие идентифицируемости. В ситуациях 2 и 3 коэффициенты i -го структурного уравнения однозначно восстанавливаются на основании коэффициентов приведенной системы. Однако эти две ситуации различаются существенным образом, если рассматривать задачу восстановления коэффициентов i -го структурного уравнения на основании оценок коэффициентов приведенной формы, которые получены методом наименьших квадратов, примененным к каждому отдельному уравнению приведенной системы и не учитывающим ограничения на коэффициенты приведенной формы, накладываемые на них соотношением $\Pi = B\Gamma^{-1}$. Если $\hat{\Pi}$ — оценка матрицы Π , полученная таким свободным оцениванием, то в ситуации 2 коэффициенты i -го структурного уравнения восстанавливаются по матрице $\hat{\Pi}$ однозначным образом, тогда как в ситуации 3 существует несколько вариантов восстановления, приводящих к различным результатам.

Заметим, что разным уравнениям системы могут соответствовать разные ситуации из трех перечисленных.

Проанализируем рассмотренные в этом разделе системы одновременных уравнений.

Первой была система

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t. \end{cases}$$

Здесь список эндогенных переменных: (Q_t, P_t) , а список предопределенных переменных ограничивается переменной, тождественно равной 1, так что полный список переменных в системе: $(Q_t, P_t, 1)$. При этом $g = 2$, $K = 1$, матрицы Γ , B и A имеют вид:

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad B = (a_0, b_0), \quad A = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & b_0 \end{pmatrix}.$$

На столбцы матрицы A не накладывается никаких ограничений, кроме нормировочных, так что $g_1^* = g_2^* = 0$, $K_1^* = K_2^* = 0$, и ни для одного из двух уравнений не выполнено порядковое условие $g_i^* + K_i^* \geq g - 1$. Следовательно, система неидентифицируема.

Следующей была система:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t, \end{cases}$$

т.е.

$$\begin{cases} Q_t - a_1 P_t = a_0 + a_2 Y_t + u_t \\ Q_t - b_1 P_t = b_0 + v_t. \end{cases}$$

Здесь список эндогенных переменных тот же: (Q_t, P_t) . В список предопределенных переменных входят две переменные: $(1, Y_t)$. Полный список переменных в системе: $(Q_t, P_t, 1, Y_t)$. При этом $g = 2$, $K = 2$, матрицы Γ , B и A имеют вид:

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad B = \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \end{pmatrix},$$

$$A = (\alpha_1 \quad \alpha_2) = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \\ \alpha_{41} & \alpha_{42} \end{pmatrix} = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & b_0 \\ a_2 & 0 \end{pmatrix}.$$

На элементы первого столбца матрицы A накладывается только условие нормировки $\alpha_{11} = 1$. Поэтому первое уравнение системы неидентифицируемо. На элементы второго столбца, помимо нормировочного, накладывается одно исключающее ограничение $\alpha_{42} = 0$, так что для этого столбца $g_2^* = 0$, $K_2^* = 1$, и $g_2^* + K_2^* = g - 1 = 1$, т.е. порядковое условие идентифицируемости выполняется.

Заметим далее, что ограничение $\alpha_{42} = 0$ можно записать в виде $\Phi_2 \alpha_2 = 0$, где $\Phi_2 = (0 \ 0 \ 0 \ 1)$. Тогда

$$\Phi_2 A_2 = (0 \ 0 \ 0 \ 1)(1, -a_1, a_0, a_2)^T = (a_2),$$

$$\text{rank}(\Phi_2 A) = \text{rank}(\Phi_2 A) = \text{rank}(a_2) = 1,$$

так что $\text{rank}(\Phi_2 A) = g - 1$, и ранговое условие идентифицируемости выполнено. Наконец, поскольку $g_2^* + K_2^* = g - 1 = 1$, второе уравнение идентифицируемо точно.

Следующая система:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + v_t, \end{cases}$$

т.е.

$$\begin{cases} Q_t - a_1 P_t = a_0 + a_2 Y_t + u_t \\ Q_t - b_1 P_t = b_0 + b_2 R_t + v_t. \end{cases}$$

Список эндогенных переменных: (Q_t, P_t) . Список предопределенных переменных: $(1, Y_t, R_t)$. Полный список переменных в системе: $(Q_t, P_t, 1, Y_t, R_t)$. При этом $g = 2$, $K = 3$, матрицы Γ , B и A имеют вид:

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad B = \begin{pmatrix} a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix}, \quad A = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & b_0 \\ a_2 & 0 \\ 0 & b_2 \end{pmatrix}.$$

Соответственно здесь для каждого столбца матрицы A , помимо нормирующего ограничения, имеется по одному исключающему ограничению на экзогенные переменные, так что $g_1^* = g_2^* = 0$, $K_1^* = K_2^* = 1$, $g_1^* + K_1^* = g - 1$, и порядковое условие выполнено.

Ограничение $\alpha_{51} = 0$ в первом столбце можно записать в виде $\Phi_1 \alpha_1 = 0$, где $\Phi_1 = (0 \ 0 \ 0 \ 0 \ 1)$. Тогда

$$\Phi_1 A_1 = (0 \ 0 \ 0 \ 0 \ 1)(1, -b_1, b_0, 0, b_2)^T = (b_2),$$

$$\text{rank}(\Phi_1 A) = \text{rank}(\Phi_1 A_1) = \text{rank}(b_2) = 1,$$

так что $\text{rank}(\Phi_1 A) = g - 1$, и для первого уравнения выполнено ранговое условие идентифицируемости. Наконец, поскольку $g_1^* + K_1^* = g - 1$, первое уравнение идентифицируемо точно.

Ограничение $\alpha_{42} = 0$ во втором столбце можно записать в виде $\Phi_2 \alpha_2 = 0$, где $\Phi_2 = (0 \ 0 \ 0 \ 1 \ 0)$. Тогда

$$\Phi_2 A_2 = (0 \ 0 \ 0 \ 1 \ 0) (1, -a_1, a_0, a_2, 0)^T = (a_2),$$

$$\text{rank}(\Phi_2 A) = \text{rank}(\Phi_2 A_2) = \text{rank}(a_2) = 1,$$

так что $\text{rank}(\Phi_2 A) = g - 1$, и для второго уравнения также выполнено ранговое условие идентифицируемости. Наконец, поскольку $g_2^* + K_2^* = g - 1$, второе уравнение идентифицируемо точно.

Таким образом, в данной системе одновременных уравнений оба уравнения идентифицируемы, причем идентифицируемы точно.

Наконец, в системе

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + b_3 S_t + v_t, \end{cases}$$

т.е.

$$\begin{cases} Q_t - a_1 P_t = a_0 + u_t \\ Q_t - b_1 P_t = b_0 + b_2 R_t + b_3 S_t + v_t, \end{cases}$$

эндогенные переменные те же, а список предопределенных переменных: $(1, R_t, S_t)$. Полный список переменных в системе: $(Q_t, P_t, 1, R_t, S_t)$. При этом $g = 2$, $K = 3$, матрицы Γ , B и A имеют вид:

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad B = \begin{pmatrix} a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix}, \quad A = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix}.$$

На элементы второго столбца накладывается только условие нормировки. Поэтому второе уравнение системы неидентифицируемо. На элементы первого столбца, помимо условия нормировки, накладываются два исключаяющих ограничения: $\alpha_{41} = 0$, $\alpha_{51} = 0$. При этом $g_1^* = 0$, $K_1^* = 2$, $g_1^* + K_1^* = 2 > g - 1$, так что первое уравнение идентифицируемо. Исключающие ограничения можно записать в форме $\Phi_1 \alpha_1 = 0$, где

$$\Phi_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Тогда

$$\Phi_1 A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & b_0 \\ 0 & b_2 \\ 0 & b_3 \end{pmatrix} = \begin{pmatrix} 0 & b_2 \\ 0 & b_3 \end{pmatrix},$$

$$\Phi_1 A_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} (1, -b_1, b_0, b_2, b_3)^T = \begin{pmatrix} b_2 \\ b_3 \end{pmatrix},$$

$$\text{rank}(\Phi_1 A) = \text{rank}(\Phi_1 A_1) = \text{rank} \begin{pmatrix} b_2 \\ b_3 \end{pmatrix} = 1,$$

так что $\text{rank}(\Phi_1 A) = g - 1$, и для первого уравнения выполнено ранговое условие идентифицируемости. Поскольку $g_1^* + K_1^* = 2 > g - 1$, первое уравнение сверхидентифицируемо.

Приведем пример системы (упрощенный вариант модели мультипликатора-акселератора), в которой присутствуют линейные **ограничения неисключающего типа** (*non-excluding conditions*):

$$\begin{cases} C_t = a_0 + a_1 Y_t + a_2 C_{t-1} + u_{t1} \\ I_t = b_0 + b_1 (Y_t - Y_{t-1}) + u_{t2} \\ Y_t = C_t + I_t, \end{cases}$$

где C_t — потребление;

I_t — инвестиции;

Y_t — доход.

Подставляя выражение для I_t из последнего тождества во второе уравнение, запишем систему в виде:

$$\begin{cases} C_t - a_1 Y_t = a_0 + a_2 C_{t-1} + u_{t1} \\ -C_t + (1 - b_1) Y_t = b_0 - b_1 Y_{t-1} + u_{t2}. \end{cases}$$

Список эндогенных переменных: (C_t, Y_t) , список предопределенных переменных: $(1, C_{t-1}, Y_{t-1})$. Полный список: $(C_t, Y_t, 1, C_{t-1}, Y_{t-1})$. Матрица A :

$$A = \begin{pmatrix} 1 & -1 \\ -a_1 & 1 - b_1 \\ a_0 & b_0 \\ a_2 & 0 \\ 0 & -b_1 \end{pmatrix}.$$

В первом столбце одно исключаящее ограничение $\alpha_{51} = 0$, т.е. $\Phi_1 \alpha_1 = 0$, где $\Phi_1 = (0 \ 0 \ 0 \ 0 \ 1)$. При этом

$$\text{rank}(\Phi_1 A) = \text{rank}(\Phi_1 A_1) = \text{rank}(-b_1) = 1 = g - 1,$$

так что первое уравнение идентифицируемо. Поскольку $g_1^* + K_1^* = 1 = g - 1$, это уравнение идентифицируемо точно.

Во втором столбце одно исключаящее ограничение $\alpha_{42} = 0$ и одно неисключающее ограничение $\alpha_{12} + \alpha_{22} = \alpha_{52}$. Эту пару ограничений можно записать в виде $\Phi_2 \alpha_2 = 0$, где

$$\Phi_2 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & -1 \end{pmatrix}.$$

Тогда

$$\Phi_2 A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -a_1 & 1-b_1 \\ a_0 & b_0 \\ a_2 & 0 \\ 0 & -b_1 \end{pmatrix} = \begin{pmatrix} a_2 & 0 \\ 1-a_1 & 0 \end{pmatrix},$$

$\text{rank}(\Phi_2 A) = 1$, так что $\text{rank}(\Phi_2 A) = g - 1$, и для второго уравнения также выполнено ранговое условие идентифицируемости. Поскольку $R_2 = 2 > g - 1 = 1$, второе уравнение сверхидентифицируемо.

✓ **Замечание 1.1.1.** Константа играет в проблеме идентификации такую же роль, что и остальные предопределенные переменные.

Проиллюстрируем Замечание 1.1.1. Как уже выяснили ранее, в системе

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t \end{cases}$$

оба уравнения неидентифицируемы. Исключим константу из правой части второго уравнения:

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_1 P_t + v_t. \end{cases}$$

Для измененной системы имеем те же списки эндогенных и предопределенных переменных, полный список переменных в системе: $(Q_t, P_t, 1)$. При этом $g = 2$, $K = 1$, матрица Γ не изменяется, а матрицы B и A принимают вид:

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad B = (a_0, 0), \quad A = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \\ a_0 & 0 \end{pmatrix}.$$

На первый столбец матрицы A не накладывается никаких ограничений, кроме нормировочных, так что $g_1^* = 0$, $K_1^* = 0$, и для первого уравнения не выполнено порядковое условие $g_i^* + K_i^* \geq g - 1$. Следовательно, первое уравнение неидентифицируемо. Однако на второй столбец на этот раз накладывается исключающее ограничение $\alpha_{32} = 0$, т.е. $\Phi_2 \alpha_2 = 0$, где $\Phi_2 = (0 \ 0 \ 1)$. При этом

$$\text{rank}(\Phi_2 A) = \text{rank}(a_0 \ 0) = \text{rank}(\Phi_2 A_1) = (a_0) = 1 = g - 1,$$

так что второе уравнение идентифицируемо. Поскольку $g_2^* + K_2^* = 1 = g - 1$, это уравнение идентифицируемо точно.

✓ **Замечание 1.1.2.** Критерий идентифицируемости дает один и тот же результат в отношении i -го стохастического структурного уравнения (содержащего случайные ошибки в правой части) независимо от того, рассматривается полная система вместе с тождествами или система, в которой тождества учтены и исключены.

Проиллюстрируем Замечание 1.1.2. При исследовании вопроса об идентифицируемости модели

$$\begin{cases} Q_t^d = a_0 + a_1 P_t + u_t \\ Q_t^s = b_0 + b_1 P_t + v_t \\ Q_t^s = Q_t^d \end{cases}$$

и различных ее расширений мы, исключая (и учитывая) тождество, сводили эти модели к системам без тождеств, так что в правых частях всех уравнений преобразованных систем присутствовали случайные ошибки. Поступая, например, таким образом с системой трех уравнений

$$\begin{cases} Q_t^d = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t^s = b_0 + b_1 P_t + v_t \\ Q_t^d = Q_t^s, \end{cases}$$

проверяли условия идентифицируемости системы двух уравнений, полученных на основании этой системы:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t, \end{cases}$$

и обнаружили, что первое уравнение системы неидентифицируемо, а второе идентифицируемо точно.

Попробуем проверить условия идентифицируемости непосредственно в рамках исходной системы трех уравнений, так что $g = 3$. Для этой системы список эндогенных переменных полнее, чем у преобразованной системы: (Q_t^d, Q_t^s, P_t) , тогда как список предопределенных переменных $(1, Y_t)$ не изменяется. Полный список содержит теперь 5 переменных: $(Q_t^d, Q_t^s, P_t, 1, Y_t)$. Перенесем все эндогенные переменные в левые части уравнений:

$$\begin{cases} Q_t^d - a_1 P_t = a_0 + a_2 Y_t + u_t \\ Q_t^s - b_1 P_t = b_0 + v_t \\ Q_t^d - Q_t^s = 0. \end{cases}$$

Матрица A имеет вид:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ -a_1 & -b_1 & 0 \\ a_0 & b_0 & 0 \\ a_2 & 0 & 0 \end{pmatrix}.$$

На элементы первого столбца накладывается исключающее ограничение $\alpha_{21} = 0$, т.е. $\Phi_1 \alpha_1 = 0$, где $\Phi_1 = (0 \ 1 \ 0 \ 0 \ 0)$. При этом

$$\text{rank}(\Phi_1 A) = \text{rank} \begin{pmatrix} 0 & 1 & -1 \end{pmatrix} = 1 < g - 1 = 2,$$

так что первое уравнение неидентифицируемо. На элементы второго столбца накладывается $K_2^* = 2$ исключающих ограничения: $\alpha_{12} = 0$ и $\alpha_{52} = 0$, т.е. $\Phi_2 \alpha_2 = 0$, где

$$\Phi_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

При этом

$$\text{rank}(\Phi_2 A) = \text{rank} \begin{pmatrix} 1 & 0 & 1 \\ a_2 & 0 & 0 \end{pmatrix} = 2 = g - 1,$$

так что второе уравнение идентифицируемо, причем идентифицируемо точно, поскольку $g_2^* + K_2^* = 2 = g - 1$. Результаты в отношении каждого из двух стохастических уравнений оказались одинаковыми для систем из трех и из двух уравнений.

До сих пор мы рассматривали только возможность восстановления коэффициентов структурных уравнений по коэффициентам приведенной формы. Однако идентифицируемость i -го стохастического структурного уравнения, строго говоря, означает идентифицируемость не только коэффициентов этого уравнения, но и дисперсии случайной составляющей в этом уравнении.

Идентифицируемость системы структурных уравнений в целом (на основании приведенной формы системы) означает идентифицируемость не только всех коэффициентов системы, но и ковариационной матрицы случайных ошибок, входящих в правые части уравнений системы. При этом при восстановлении коэффициентов и ковариационной матрицы ошибок в структурной форме используются не только коэффициенты приведенной формы, но и ковариационная матрица ошибок в приведенной форме.

Обратимся опять к общей форме системы:

$$y_t \Gamma = x_t B + u_t, \quad t = 1, \dots, n,$$

где

$$\Gamma = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g1} & \dots & \gamma_{gg} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{11} & \dots & \beta_{1g} \\ \vdots & \ddots & \vdots \\ \beta_{K1} & \dots & \beta_{Kg} \end{pmatrix},$$

$$y_t = (y_{t1}, \dots, y_{tg}), \quad x_t = (x_{t1}, \dots, x_{tK}), \quad u_t = (u_{t1}, \dots, u_{tg}),$$

и предполагается невырожденность матрицы Γ . Приведенная форма системы:

$$y_t = x_t B \Gamma^{-1} + u_t \Gamma^{-1} = x_t \Pi + w_t,$$

где

$$\Pi = B \Gamma^{-1} = \begin{pmatrix} \pi_{11} & \dots & \pi_{1g} \\ \vdots & \ddots & \vdots \\ \pi_{K1} & \dots & \pi_{Kg} \end{pmatrix}, \quad w_t = u_t \Gamma^{-1} = (w_{t1}, \dots, w_{tg}).$$

Пусть

$$E(u_t) = 0, \quad (Cov(u_{ti}, u_{tj})) = \Sigma = (\sigma_{ij}),$$

$$(Cov(u_{ti}, u_{sj})) = 0 \quad \text{для } t \neq s,$$

так что ошибки не коррелированы по времени, но для одного и того же момента ошибки в разных уравнениях могут быть коррелированными между собой. Тогда $E(w_t) = 0$ и для ковариационной матрицы $\Omega = (\omega_{ij}) = Cov(w_t) = (Cov(w_{ti}, w_{tj}))$ вектора w_t ошибок в приведенном уравнении имеем:

$$\Omega = Cov(w_t) = Cov(u_t \Gamma^{-1}) = (\Gamma^{-1})^T \Sigma (\Gamma^{-1}),$$

так что

$$\Sigma = \Gamma^T \Omega \Gamma.$$

Следовательно, если структурная система идентифицируема (коэффициенты структурной системы однозначно восстанавливаются по коэффициентам приведенной формы), то тогда, восстановив по коэффициентам приведенной формы матрицу Γ , можно, используя эти восстановленные коэффициенты и матрицу Ω , восстановить ковариационную матрицу Σ .

Если структурная форма не восстанавливается целиком, а возможно лишь восстановление некоторых ее уравнений, то тогда для полной идентификации i -го стохастического структурного уравнения надо восстановить все его коэффициенты и дисперсию случайной составляющей этого уравнения. Пусть нас интересует, например, первое уравнение системы. Представим тогда матрицу Γ в виде

$$\Gamma = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g1} & \cdots & \gamma_{gg} \end{pmatrix} = [\gamma_1 : \Gamma_1], \quad \text{где } \gamma_1 = \begin{pmatrix} \gamma_{11} \\ \vdots \\ \gamma_{g1} \end{pmatrix}, \quad \Gamma_1 = \begin{pmatrix} \gamma_{12} & \cdots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g2} & \cdots & \gamma_{gg} \end{pmatrix}.$$

Дисперсия случайной составляющей в первом структурном уравнении равна $\sigma_{11} = \gamma_1^T \Omega \gamma_1$, так что для ее восстановления по приведенной форме достаточно предварительно восстановить только коэффициенты первого уравнения. Аналогично, если нас интересует i -е стохастическое структурное уравнение, то дисперсия случайной составляющей в этом структурном уравнении равна $\sigma_{ii} = \gamma_i^T \Omega \gamma_i$, где γ_i — i -й столбец матрицы Γ , и для восстановления σ_{ii} достаточно предварительно восстановить коэффициенты i -го уравнения.

В качестве примера рассмотрим опять систему

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + v_t. \end{cases}$$

Ранее установили, что в этой системе коэффициенты первого уравнения неидентифицируемы, а коэффициенты второго идентифицируемы точно. Для этой системы

$$\Gamma = \begin{pmatrix} 1 & 1 \\ -a_1 & -b_1 \end{pmatrix}, \quad \gamma_2 = \begin{pmatrix} 1 \\ -b_1 \end{pmatrix},$$

так что (поскольку ковариационные матрицы симметричны)

$$D(v_t) = \sigma_{22} = (1 - b_1) \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} \begin{pmatrix} 1 \\ -b_1 \end{pmatrix} = \omega_{11} - 2b_1\omega_{12} + b_1^2\omega_{22}.$$

Оценив наряду с коэффициентами приведенной формы элементы ковариационной матрицы ошибок приведенной формы, можно получить оценку для коэффициента b_1 , а через нее — и оценку для $D(v_t)$.



Замечание 1.1.3. Если проанализировать примеры, в которых на уравнения накладывались только исключаяющие ограничения, нетрудно заметить, что проверку рангового условия идентифицируемости i -го стохастического структурного уравнения, по существу, можно проводить следующим образом.

Составляется таблица, в «шапке» которой перечисляются эндогенные и предопределенные переменные, задействованные в системе, а в i -й строке находятся коэффициенты при этих переменных в левой и правой частях i -го уравнения (как они есть, без переносов в левую часть).

Например, для системы

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + v_t \end{cases}$$

это будет табл. 1.2.

Таблица 1.2

i	Q_t	P_t	1	Y_t	R_t
1	1	a_1	a_0	a_2	0
2	1	b_1	b_0	0	b_2

Для исследования i -го уравнения достаточно рассмотреть матрицу, образованную теми столбцами таблицы, элементы которых, стоящие в i -й строке, равны нулю, и всеми строками таблицы кроме i -й. В рассматриваемом примере при исследовании 1-го уравнения такая матрица состоит из единственного элемента b_2 , а при исследовании 2-го уравнения — из единственного элемента a_2 . В обоих случаях ранг выделенной матрицы равен 1, и поскольку $g - 1 = 1$, оба уравнения идентифицируемы.

Для системы

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + b_3 S_t + v_t \end{cases}$$

получаем табл. 1.3.

Таблица 1.3

i	Q_t	P_t	1	R_t	S_t
1	1	a_1	a_0	0	0
2	1	b_1	b_0	b_2	b_3

Для второго уравнения нет ни исключяющих, ни других линейных ограничений — только нормирующее ограничение, так что второе уравнение неидентифицируемо. На коэффициенты первого уравнения, помимо нормирующего, накладываются только исключяющие ограничения. Выделяемая матрица сводится к одной строке с двумя элементами: $(b_2 \ b_3)$. Ранг этой матрицы равен 1, так что $g - 1 = 1$, и первое уравнение идентифицируемо.

✓ **Замечание 1.1.4.** В реальных ситуациях, если порядковое условие выполнено, то, как правило, выполняется и ранговое условие. Приводимые в литературе контрпримеры носят явно искусственный характер. В качестве такого контрпримера выступает, например, система трех стохастических структурных уравнений

$$\begin{cases} a_{11}y_{t1} + a_{12}y_{t2} + a_{13}y_{t3} = a_{14}x_{t1} + a_{15}x_{t2} + u_{t1} \\ a_{21}y_{t1} + a_{22}y_{t2} + a_{23}y_{t3} = a_{24}x_{t1} + a_{25}x_{t2} + u_{t2} \\ a_{31}y_{t1} + a_{32}y_{t2} + a_{33}y_{t3} = a_{34}x_{t1} + a_{35}x_{t2} + u_{t3}, \end{cases}$$

в которой на коэффициенты первого уравнения накладываются линейные ограничения $a_{14} = 0$, $a_{12} = a_{13}$. Эти ограничения записываются в стандартной форме как $\Phi_1 \alpha_1 = 0$, где

$$\Phi_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{pmatrix},$$

так что

$$\Phi_1 A = \begin{pmatrix} a_{14} & a_{24} & a_{34} \\ a_{12} - a_{13} & a_{22} - a_{23} & a_{32} - a_{33} \end{pmatrix} = \begin{pmatrix} 0 & a_{24} & a_{34} \\ 0 & a_{22} - a_{23} & a_{32} - a_{33} \end{pmatrix}.$$

Ранговое условие не выполняется, если строки этой матрицы пропорциональны. Последнее может осуществляться, если:

- на уравнения системы накладываются одинаковые ограничения;
- переменная x_{t1} не входит в систему;
- коэффициенты при y_{t2} и y_{t3} равны во всех уравнениях.

✓ **Замечание 1.1.5.** До сих пор мы не предполагали никаких ограничений на ковариационную матрицу Σ вектора ошибок в структурной форме. Между тем введение ограничений на структуру этой матрицы в некоторых ситуациях может помочь идентификации уравнений, которые без таких ограничений неидентифицируемы.

Для иллюстрации Замечания 1.1.5 рассмотрим систему

$$\begin{cases} Q_t = a_1 P_t + a_2 Q_{t-1} + u_{t1} \\ P_t = b_1 Q_{t-1} + u_{t2}, \end{cases}$$

где P_t и Q_t — эндогенные переменные;

Q_{t-1} — единственная предопределенная переменная.

При этом

$$\Gamma = \begin{pmatrix} 1 & 0 \\ -a_1 & 1 \end{pmatrix}, \quad B = (a_2 \quad b_1), \quad \Pi = (\pi_{11} \quad \pi_{12}),$$

так что соотношение $\Pi\Gamma = B$ принимает вид:

$$(\pi_{11} \quad \pi_{12}) \begin{pmatrix} 1 & 0 \\ -a_1 & 1 \end{pmatrix} = (a_2 \quad b_1),$$

откуда $(\pi_{11} - a_1\pi_{12} \quad \pi_{12}) = (a_2 \quad b_1)$, т.е. $a_2 = \pi_{11} - a_1\pi_{12}$, $b_1 = \pi_{12}$. Таким образом, единственный коэффициент второго уравнения восстанавливается по матрице Π однозначно, а для восстановления двух коэффициентов первого уравнения имеется только одно уравнение, и первое уравнение оказывается неидентифицируемым. Вспомним, однако, соотношение между ковариационными матрицами ошибок в приведенной и структурной формах:

$$\Sigma = \Gamma^T \Omega \Gamma.$$

В нашем примере оно принимает вид:

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 1 & -a_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -a_1 & 1 \end{pmatrix} = \\ = \begin{pmatrix} \omega_{11} - 2a_1\omega_{12} + a_1^2\omega_{22} & \omega_{12} - a_1\omega_{22} \\ \omega_{12} - a_1\omega_{22} & \omega_{22} \end{pmatrix}.$$

Если предположить дополнительно, что $\sigma_{12} = \sigma_{21} = 0$, т.е. что ошибки в разных уравнениях не коррелированы между собой, то из последнего соотношения получаем:

$$\omega_{12} - a_1\omega_{22} = 0,$$

так что коэффициент a_1 первого структурного уравнения восстанавливается

по матрице Ω : $a_1 = \frac{\omega_{12}}{\omega_{22}}$. После этого восстанавливается и коэффициент a_2

первого структурного уравнения: $a_2 = \pi_{12} - a_1\pi_{22}$. Тем самым оказывается идентифицируемым все первое уравнение структурной формы.

✓ **Замечание 1.1.6.** Рассмотренная в Замечании 1.1.5 система

$$\begin{cases} Q_t = a_1 P_t + a_2 Q_{t-1} + u_{t1} \\ P_t = b_1 Q_{t-1} + u_{t2} \end{cases}$$

при выполнении условия $\sigma_{12} = \sigma_{21} = 0$ принадлежит классу **рекурсивных систем** (*recursive systems*). Благодаря последовательному определению эндогенных переменных в таких системах при переходе от уравнения к уравнению, в правых частях каждого из уравнений системы не оказываются переменных, значения которых коррелированы со значением ошибки в этом уравнении при одном и том же t . Во втором уравнении рассматриваемой системы $Cov(Q_{t-1}, u_{t2}) = 0$,

так как значение Q_{t-1} определяется ранее момента t . В правой части первого уравнения $Cov(Q_{t-1}, u_{t1}) = 0$ по той же причине и

$$\begin{aligned} Cov(P_t, u_{t1}) &= Cov(b_1 Q_{t-1} + u_{t2}, u_{t1}) = \\ &= b_1 Cov(Q_{t-1}, u_{t1}) + Cov(u_{t2}, u_{t1}) = 0, \end{aligned}$$

так что при выполнении условия $\sigma_{12} = \sigma_{21} = 0$ переменная P_t не является эндогенной. Если же $\sigma_{12} \neq 0$, то P_t становится эндогенной переменной, а система перестает быть рекурсивной.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какая эконометрическая модель называется системой одновременных уравнений?
2. Что понимается под структурной и приведенной формами системы одновременных уравнений?
3. Что понимается под идентифицируемостью отдельного структурного уравнения системы одновременных уравнений?
4. Что понимается под идентифицируемостью структурной формы системы одновременных уравнений?
5. Могут ли быть в неидентифицируемой системе одновременных уравнений отдельные структурные уравнения, являющиеся идентифицируемыми?
6. На какие три типа подразделяются переменные при рассмотрении вопроса об идентифицируемости структурных уравнений системы?
7. В чем состоит порядковое условие идентифицируемости отдельного структурного уравнения?
8. В чем состоит ранговое условие идентифицируемости отдельного структурного уравнения?
9. Охарактеризуйте три ситуации, возникающие при оценивании отдельного структурного уравнения, с привлечением порядкового и рангового условий идентифицируемости.
10. Какова роль константы в проблеме идентифицируемости структурных уравнений?
11. Как производится проверка рангового условия идентифицируемости в ситуациях, когда на структурные уравнения накладываются только исключающие ограничения?
12. Пусть критерий идентифицируемости применяется в отношении выделенного структурного уравнения, содержащего случайные ошибки в правой части. Зависит ли результат применения этого критерия от того, что рассматривается — полная система вместе с тождествами или система, в которой тождества учтены и исключены?
13. Может ли помочь в восстановлении коэффициентов структурной формы использование ковариационной матрицы ошибок в приведенной форме?
14. Может ли введение ограничений на структуру ковариационной матрицы ошибок в структурной форме системы обеспечить идентификацию уравнений, которые без таких ограничений неидентифицируемы?
15. В каком случае система одновременных уравнений называется рекурсивной?
16. Останется ли рекурсивная система рекурсивной, если снять условие некоррелированности ошибок в разных уравнениях?

_____ Тема 1.2 _____

ОЦЕНИВАНИЕ СИСТЕМ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

Косвенный метод наименьших квадратов

Если i -е стохастическое уравнение структурной формы идентифицируемо точно, то параметры этого уравнения (коэффициенты уравнения и дисперсия случайной ошибки) восстанавливаются по параметрам приведенной системы однозначно. Поэтому для оценивания параметров такого уравнения достаточно оценить коэффициенты каждого из уравнений приведенной формы методом наименьших квадратов (отдельно для каждого уравнения) и получить оценку ковариационной матрицы Ω ошибок в приведенной форме, после чего воспользоваться соотношениями $\Pi\Gamma = B$ и $\Sigma = \Gamma^T\Omega\Gamma$, подставляя в них вместо Π оцененную матрицу коэффициентов приведенной формы $\hat{\Pi}$ и оцененную ковариационную матрицу ошибок в приведенной форме $\hat{\Omega}$. Такая процедура называется **косвенным методом наименьших квадратов** (*ILS — indirect least squares*). Полученные в результате оценки коэффициентов i -го стохастического уравнения структурной формы наследуют свойство состоятельности оценок приведенной формы. Однако они не наследуют таких свойств оценок приведенной формы, как несмещенность и эффективность, из-за того что получаются в результате некоторых нелинейных преобразований. Соответственно при небольшом количестве наблюдений даже у этих естественных оценок может возникать заметное смещение. В связи с этим при рассмотрении различных методов оценивания коэффициентов структурных уравнений в первую очередь заботятся об обеспечении именно состоятельности получаемых оценок.

Двухшаговый метод наименьших квадратов

Фактически мы уже воспользовались этим методом при рассмотрении системы

$$\begin{cases} C_t = \alpha + \beta Y_t + \varepsilon_t \\ Y_t = C_t + I_t \end{cases}$$

в разд. 6 первой части учебника (тема 6.2). Там подменили переменную Y_t в первом структурном уравнении искусственной инструментальной переменной $\hat{Y}_t = \hat{\gamma} + \hat{\delta}I_t$, где $\hat{\gamma}$ и $\hat{\delta}$ — оценки наименьших квадратов, получаемые при оценивании модели $Y_t = \gamma + \delta I_t + \varepsilon_t$. После такой подмены уравнение $C_t = \alpha + \beta\hat{Y}_t + \varepsilon_t$ состоятельно оценивается обычным методом наименьших квадратов, поскольку «объясняющая переменная» \hat{Y}_t в этом уравнении не коррелирована с ε_t .

Рассмотрим систему g одновременных уравнений

$$y_t \Gamma = x_t B + u_t, \quad t = 1, \dots, n,$$

$$\text{где } \Gamma = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g1} & \dots & \gamma_{gg} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{11} & \dots & \beta_{1g} \\ \vdots & \ddots & \vdots \\ \beta_{K1} & \dots & \beta_{Kg} \end{pmatrix},$$

$y_t = (y_{t1}, \dots, y_{tg})$ — вектор значений эндогенных переменных в t -м наблюдении;

$x_t = (x_{t1}, \dots, x_{tK})$ — вектор значений предопределенных переменных в t -м наблюдении;

$u_t = (u_{t1}, \dots, u_{tg})$ — вектор значений случайных ошибок в t -м наблюдении,

при этом предполагается невырожденность матрицы Γ .

Пусть наибольший интерес представляет первое уравнение системы. (Это не уменьшает общности, поскольку уравнения всегда можно надлежащим образом перенумеровать.) Считая, что первое уравнение нормировано на коэффициент при переменной y_{t1} , преобразуем уравнение к виду:

$$y_{t1} = \alpha_{11} y_{t1}^* + \dots + \alpha_{g1,1} y_{tg1}^* + \theta_{11} x_{t1}^* + \dots + \theta_{K1,1} x_{tK1}^* + u_{t1},$$

или

$$y_{t1} = Y_{t1} \alpha_1 + X_{t1} \theta_1 + u_{t1},$$

где $Y_{t1} = (y_{t1}^*, \dots, y_{tg1}^*)$ — вектор значений g_1 эндогенных переменных, включенных в правую часть первого уравнения;

$X_{t1} = (x_{t1}^*, \dots, x_{tK1}^*)$ — вектор значений K_1 предопределенных переменных, включенных в правую часть первого уравнения;

α_1 и θ_1 — векторы коэффициентов при эндогенных и предопределенных переменных, включенных в первое уравнение.

Состоятельному оцениванию коэффициентов уравнения мешает эндогенность переменных $y_{t1}^*, \dots, y_{tg1}^*$. Это затруднение преодолевается за два шага — отсюда название метода: **двухшаговый метод наименьших квадратов** (2SLS — *two-stage least squares*).

1. Производится оценивание уравнений регрессии каждой из этих переменных на все предопределенные переменные, включенные в систему. В качестве очищенных вариантов переменных $y_{t1}^*, \dots, y_{tg1}^*$ берутся предсказанные значения $\hat{y}_{t1}^*, \dots, \hat{y}_{tg1}^*$ этих переменных (в этом контексте предопределенные переменные понимаются как *инструменты* для очистки эндогенных переменных).

2. Значения $y_{t1}^*, \dots, y_{tg_1}^*$ эндогенных переменных в первом уравнении заменяются значениями $\hat{y}_{t1}^*, \dots, \hat{y}_{tg_1}^*$. Полученное уравнение оценивается методом наименьших квадратов.

Как и в случае оценивания обычным методом наименьших квадратов (*OLS* — *ordinary least squares*) единственного уравнения в линейной множественной регрессии, процедуру *2SLS* можно представить в матричном виде. Для этого предположим, что в левой части i -го стохастического структурного уравнения находится единственная эндогенная переменная y_{ti} , и обозначим:

$y_i = (y_{1i}, \dots, y_{ni})^T$ — вектор-столбец значений i -й эндогенной переменной в n наблюдениях;

$Y_i = \begin{pmatrix} y_{11}^* & \dots & y_{1g_i}^* \\ \vdots & \ddots & \vdots \\ y_{n1}^* & \dots & y_{ng_i}^* \end{pmatrix}$ — матрица значений в n наблюдениях g_i эндогенных переменных, включенных в правую часть i -го уравнения;

$X_i = \begin{pmatrix} x_{11}^* & \dots & x_{1K_i}^* \\ \vdots & \ddots & \vdots \\ x_{n1}^* & \dots & x_{nK_i}^* \end{pmatrix}$ — матрица значений в n наблюдениях K_i предопределенных переменных, включенных в правую часть i -го структурного уравнения;

$u_i = (u_{1i}, \dots, u_{ni})^T$ — вектор-столбец значений ошибки в i -м структурном уравнении в n наблюдениях;

$X = \begin{pmatrix} x_{11} & \dots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nK} \end{pmatrix}$ — матрица значений в n наблюдениях всех K предопределенных переменных, включенных в систему;

$\Pi = \begin{pmatrix} \pi_{11} & \dots & \pi_{1g} \\ \vdots & \ddots & \vdots \\ \pi_{K1} & \dots & \pi_{Kg} \end{pmatrix}$ — матрица коэффициентов приведенной формы системы;

$W = \begin{pmatrix} w_{11} & \dots & w_{1g} \\ \vdots & \ddots & \vdots \\ w_{n1} & \dots & w_{ng} \end{pmatrix}$ — матрица значений в n наблюдениях ошибок в g уравнениях приведенной формы.

Заметим во избежание путаницы, что $y_i = (y_{1i}, \dots, y_{ni})^T$ и $u_i = (u_{1i}, \dots, u_{ni})^T$ — векторы-столбцы, содержащие значения объясняемой переменной и случайных ошибок в i -м структурном уравнении в моменты времени $t = 1, \dots, n$. Их не следует путать с ранее использовавшимися векторами-строками $y_t = (y_{t1}, \dots, y_{tg})$ и $u_t = (u_{t1}, \dots, u_{tg})$, которые содержат значения объясняемой

переменной и случайных ошибок в g уравнениях, относящиеся к одному и тому же моменту времени t .

Тогда **первый шаг процедуры 2SLS** оценивания i -го структурного уравнения состоит в оценивании методом наименьших квадратов отдельных уравнений системы

$$Y_i = X\Pi_i + W_i,$$

где Π_i — подматрица матрицы Π коэффициентов приведенной формы, образованная теми ее столбцами, которые соответствуют эндогенным переменным, включенным в правую часть i -го структурного уравнения;
 W_i — подматрица матрицы W , образованная столбцами матрицы W , соответствующими тем же эндогенным переменным.

Оценив отдельные уравнения методом наименьших квадратов, получим оценку $\hat{\Pi}_i$ матрицы Π_i и на ее основе — оценку матрицы Y_i в виде $\hat{Y}_i = X\hat{\Pi}_i$. Матрица \hat{Y}_i содержит значения эндогенных переменных, включенных в правую часть i -го структурного уравнения, «очищенных» от корреляции с ошибкой в этом уравнении.

Обозначив α_i и θ_i — векторы коэффициентов при эндогенных и предопределенных переменных, включенных в i -е структурное уравнение, запишем это уравнение в виде:

$$y_i = Y_i\alpha_i + X_i\theta_i + u_i = Z_i\delta_i + u_i,$$

где

$$\delta_i = \begin{pmatrix} \alpha_i \\ \theta_i \end{pmatrix}, \quad Z_i = [Y_i \ X_i].$$

Это уравнение можно записать также в следующей форме:

$$y_i = Y_i\alpha_i + X_i\theta_i + u_i = \hat{Y}_i\alpha_i + X_i\theta_i + ((Y_i - \hat{Y}_i)\alpha_i + u_i),$$

или

$$y_i = \hat{Z}_i\delta_i + \varepsilon_i,$$

где

$$\hat{Z}_i = [\hat{Y}_i \ X_i].$$

Второй шаг процедуры 2SLS состоит в вычислении оценки наименьших квадратов вектора δ_i в последнем уравнении. Эта оценка вычисляется по обычной формуле:

$$\hat{\delta}_i^{2SLS} = (\hat{Z}_i^T \hat{Z}_i)^{-1} \hat{Z}_i^T y_i.$$

При этом, естественно, требуется, чтобы матрица $\hat{Z}_i^T \hat{Z}_i$ была невырожденной. Заметим, что $\text{rank}(\hat{Z}_i^T \hat{Z}_i) = \text{rank} \hat{Z}_i^T \leq \text{rank} X$. Но матрица $\hat{Z}_i^T \hat{Z}_i$ имеет порядок, равный $(g_i + K_i)$, а $\text{rank} X = K$, так что необходимо, чтобы $K \geq g_i + K_i$, т.е.

$K - K_i \geq g_i = (g - g_i^*) - 1$. Это есть не что иное, как известное нам порядковое условие идентифицируемости. Для состоятельности $\hat{\delta}_i^{2SLS}$ требуется еще, чтобы предельная матрица

$$p \lim_{n \rightarrow \infty} \left(\frac{1}{n} \hat{Z}_i^T \hat{Z}_i \right) = Q_i$$

имела конечные элементы и была невырожденной, а для этого матрица Π_i должна иметь полный столбцовый ранг. Последнее же выполняется в случае идентифицируемости i -го уравнения (см. (Schmidt, 1976, p. 150—151)).

Несмотря на то что на втором шаге используется *OLS*, непосредственное использование для построения t -статистик и доверительных интервалов вычисленных (по формулам *OLS*) значений стандартных ошибок коэффициентов невозможно. Эти значения должны быть скорректированы. Например, если дело касается первого уравнения, то на втором шаге оценивается уравнение

$$y_{t1} = \alpha_{11} \hat{y}_{t1}^* + \dots + \alpha_{g_1,1} \hat{y}_{tg_1}^* + \theta_{11} x_{t1}^* + \dots + \theta_{K_1,1} x_{tK_1}^* + \tilde{u}_{t1},$$

где $\hat{y}_{t1}^*, \dots, \hat{y}_{tg_1}^*$ — «очищенные» на первом шаге значения эндогенных переменных.

При вычислении стандартных ошибок коэффициентов по обычным формулам *OLS* используется оценка дисперсии случайной ошибки в правой части уравнения в виде:

$$\tilde{S}^2 = \frac{\sum_{t=1}^n (y_{t1} - \hat{\alpha}_{11} \hat{y}_{t1}^* - \dots - \hat{\alpha}_{g_1,1} \hat{y}_{tg_1}^* - \hat{\theta}_{11} x_{t1}^* - \dots - \hat{\theta}_{K_1,1} x_{tK_1}^*)^2}{n - g_1 - K_1}.$$

Вместо этого следует использовать другую оценку:

$$S^2 = \frac{\sum_{t=1}^n (y_{t1} - \hat{\alpha}_{11} y_{t1}^* - \dots - \hat{\alpha}_{g_1,1} y_{tg_1}^* - \hat{\theta}_{11} x_{t1}^* - \dots - \hat{\theta}_{K_1,1} x_{tK_1}^*)^2}{n - g_1 - K_1},$$

в которой вместо «очищенных» значений $\hat{y}_{t1}^*, \dots, \hat{y}_{tg_1}^*$ используются «сырые» значения $y_{t1}^*, \dots, y_{tg_1}^*$ (в специализированных программах статистического анализа систем одновременных уравнений такая коррекция предусмотрена).

В обеих формулах оценки параметров получены методом *2SLS*, для сокращения записи в обозначениях этих оценок опущен верхний индекс, указывающий на *2SLS* (например, $\hat{\alpha}_{11}$ вместо $\hat{\alpha}_{11}^{2SLS}$).

✓ **Замечание 1.2.1.** На втором шаге *2SLS* не следует особенно полагаться на указываемые в распечатках значения t -статистик, поскольку если данных мало, то асимптотическая теория неприменима, и эти статистики не имеют ни нормального, ни t -распределения.

✓ **Замечание 1.2.2.** Вернемся к рассмотренной ранее системе

$$\begin{cases} Q_t = a_0 + a_1 P_t + u_t \\ Q_t = b_0 + b_1 P_t + b_2 R_t + b_3 S_t + v_t, \end{cases}$$

в которой, как мы установили, первое уравнение сверхидентифицируемо. Следуя методу инструментальных переменных, можно произвести «очистку» эндогенной переменной P_t в правой части первого уравнения, используя в качестве инструмента только одну из переменных — R_t или S_t . В $2SLS$ в качестве инструментов для «очистки» P_t используются сразу обе эти переменные. «Очистка» с использованием только одной из них также приводит к состоятельной оценке, но эта оценка менее эффективна (ее асимптотическая дисперсия больше, чем у оценки, полученной с применением пары инструментальных переменных). В этом смысле $2SLS$ оценка является **наилучшей инструментальной оценкой** среди этих трех альтернатив.

✓ **Замечание 1.2.3** (*проверка адекватности*). Применив двухшаговый метод наименьших квадратов к i -му структурному уравнению, получим оценку $\hat{\delta}_i^{2SLS}$, оцененное значение $\hat{y}_i^{2SLS} = Z_i \hat{\delta}_i^{2SLS}$ вектора y_i и вектор остатков $\hat{y}_i^{2SLS} = y_i - \hat{y}_i^{2SLS}$ (заметим, что эти остатки отличаются от остатков, непосредственно получаемых на втором шаге $2SLS$ и равных $y_i - \hat{Z}_i \hat{\delta}_i^{2SLS}$). Опираясь на эти остатки, можно обычным образом проверять адекватность этого уравнения, используя критерии:

- нормальности (Харке — Бера, по оцененным асимметрии и кurtозису последовательности остатков);
- линейности (добавляя степени и перекрестные произведения предопределенных переменных и проверяя гипотезу обнуления коэффициентов при «лишних» составляющих);
- гомоскедастичности (Уайта, Пагана — Холла);
- независимости — против автокоррелированности остатков (Бройша — Годфри).

При обнаружении нарушений стандартных предположений необходимо соответствующим образом изменить спецификацию модели или, не изменяя спецификации, скорректировать статистические выводы.

✓ **Замечание 1.2.4.** Критерий Пагана — Холла предназначен специально для выявления гетероскедастичности в отдельном уравнении системы. В отличие от других критериев гомоскедастичности, он не предполагает отсутствия гетероскедастичности в других уравнениях системы. Этот критерий реализован, например, в пакете *Stata*.

GLS-оценивание систем одновременных уравнений. Трехшаговый метод наименьших квадратов

Если нас интересует оценивание коэффициентов всех g структурных уравнений и каждое из уравнений идентифицируемо (идентифицируемо точно или сверхидентифицируемо), то можно сразу получить *OLS*-оценку $\hat{\Pi}$ матрицы коэффициентов приведенной формы системы и на ее основе сформировать подматрицы $\hat{\Pi}_1, \dots, \hat{\Pi}_g$, соответствующие эндогенным переменным, включенным в отдельные уравнения структурной формы

$$y_i = Y_i \alpha_i + X_i \theta_i + u_i = Z_i \delta_i + u_i, \quad i = 1, \dots, g.$$

После этого можно вычислить $\hat{y}_i = X_i \hat{\Pi}_i$ и получить $2SLS$ -оценку $\hat{\delta}_i^{2SLS}$ для δ_i , применив *OLS* к уравнению $y_i = \hat{Z}_i \delta_i + \varepsilon_i$.

Совокупность оценок $\hat{\delta}_i^{2SLS}$, $i = 1, \dots, g$, в форме вектора

$$\hat{\delta}^{2SLS} = \begin{pmatrix} \hat{\delta}_1^{2SLS} \\ \vdots \\ \hat{\delta}_g^{2SLS} \end{pmatrix}$$

фактически получается как *OLS*-оценка уравнения регрессии

$$\begin{pmatrix} y_1 \\ \vdots \\ y_g \end{pmatrix} = \begin{pmatrix} \hat{Z}_1 & 0 & \dots & 0 \\ 0 & \hat{Z}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{Z}_g \end{pmatrix} \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_g \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_g \end{pmatrix},$$

в сокращенной форме:

$$y = \hat{Z} \delta + \varepsilon,$$

где

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_g \end{pmatrix}, \quad \hat{Z} = \begin{pmatrix} \hat{Z}_1 & 0 & \dots & 0 \\ 0 & \hat{Z}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{Z}_g \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_g \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_g \end{pmatrix},$$

Соответственно вектор $\hat{\delta}^{2SLS}$ находится по обычной формуле:

$$\hat{\delta}^{2SLS} = (\hat{Z}^T \hat{Z})^{-1} \hat{Z}^T y.$$

Эта оценка неэффективна вследствие коррелированности ошибок $\varepsilon_{t1}, \dots, \varepsilon_{tg}$ (имеющей место даже при некоррелированности ошибок u_{t1}, \dots, u_{tg} в струк-

турных уравнениях) и различия матриц $\hat{Z}_1, \dots, \hat{Z}_g$. Эффективную оценку можно получить, если использовать здесь вместо *OLS* обобщенный метод наименьших квадратов (*GLS*), но для этого надо знать ковариационную матрицу вектора ε . Поскольку эта матрица неизвестна, можно довольствоваться только ее оценкой, и такая оценка должна быть состоятельной, если хотим получить в итоге асимптотически эффективную оценку вектора δ .

Заметим, что ковариационная матрица вектора ε при наших предположениях имеет весьма специфический вид:

$$\text{Cov}(\varepsilon) = \text{Cov} \begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{n1} \\ \vdots \\ \varepsilon_{1g} \\ \vdots \\ \varepsilon_{ng} \end{pmatrix} = \begin{pmatrix} \lambda_{11}I_g & \lambda_{12}I_g & \dots & \lambda_{1g}I_g \\ \lambda_{21}I_g & \lambda_{22}I_g & \dots & \lambda_{2g}I_g \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{g1}I_g & \lambda_{g2}I_g & \dots & \lambda_{gg}I_g \end{pmatrix},$$

или

$$\text{Cov}(\varepsilon) = \Lambda \otimes I_g,$$

где I_g — единичная матрица порядка g ;

$$\lambda_{ij} = \text{Cov}(\varepsilon_{i1}, \varepsilon_{ij});$$

$\Lambda = (\lambda_{ij})$ — ковариационная матрица вектора $(\varepsilon_{11}, \dots, \varepsilon_{1g})$;

$\Lambda \otimes I_g$ — формализованное обозначение матрицы, являющейся **кронекеровским произведением** (*Kronecker product*) матриц Λ и I_g .

Таким образом, для реализации доступного *GLS* необходимо состоятельно оценить ковариации λ_{ij} . Это можно сделать, используя для λ_{ij} естественную оценку

$$\hat{\lambda}_{ij} = \frac{1}{n} (\hat{u}_i^{2SLS})^T \hat{u}_j^{2SLS}.$$

Заменив в выражении для $\text{Cov}(\varepsilon)$ истинные значения λ_{ij} их оценками $\hat{\lambda}_{ij}$, получим состоятельную оценку ковариационной матрицы $\text{Cov}(\varepsilon)$ в виде $\hat{\Lambda} \otimes I_g$. В результате приходим к **доступной обобщенной оценке наименьших квадратов** (*FGLS* — *feasible generalized least squares*):

$$\hat{\delta}^{3SLS} = \left(\hat{Z}^T (\hat{\Lambda}^{-1} \otimes I_g) \hat{Z} \right)^{-1} \hat{Z}^T (\hat{\Lambda}^{-1} \otimes I_g) y,$$

которая известна как **трехшаговая оценка наименьших квадратов**, или **3SLS-оценка** (*three-stage least squares*).

Оценивание систем одновременных уравнений с использованием метода максимального правдоподобия

Запишем совокупность g одновременных уравнений для n наблюдений в виде:

$$Y\Gamma = XB + U,$$

где $Y = \begin{pmatrix} y_{11} & \dots & y_{1g} \\ \vdots & \ddots & \vdots \\ y_{n1} & \dots & y_{ng} \end{pmatrix}$ — матрица значений в n наблюдениях всех g эндогенных переменных, включенных в систему;

$X = \begin{pmatrix} x_{11} & \dots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nK} \end{pmatrix}$ — матрица значений в n наблюдениях всех K предопределенных переменных, включенных в систему;

$U = \begin{pmatrix} u_{11} & \dots & u_{1g} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{ng} \end{pmatrix}$ — матрица значений в n наблюдениях случайных ошибок в g структурных уравнениях системы;

$\Gamma = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{1g} \\ \vdots & \ddots & \vdots \\ \gamma_{g1} & \dots & \gamma_{gg} \end{pmatrix}$, $B = \begin{pmatrix} \beta_{11} & \dots & \beta_{1g} \\ \vdots & \ddots & \vdots \\ \beta_{K1} & \dots & \beta_{Kg} \end{pmatrix}$ — матрицы коэффициентов.

Относительно векторов $u_t = (u_{t1}, \dots, u_{tg})$ ошибок в g уравнениях в t -м наблюдении сделаем следующие предположения:

- векторы u_1, \dots, u_n имеют одинаковое g -мерное нормальное распределение $N_g(0, \Sigma)$ с нулевым вектором математических ожиданий и положительно определенной ковариационной матрицей Σ ;
- векторы u_1, \dots, u_n независимы между собой.

При таких предположениях совместная плотность распределения случайных векторов u_1, \dots, u_n имеет вид:

$$p_u(u_1, \dots, u_n) = \prod_{t=1}^n \left[\frac{1}{(\sqrt{2\pi})^g \sqrt{\det \Sigma}} \cdot \exp \left\{ -\frac{1}{2} u_t \Sigma^{-1} u_t^T \right\} \right].$$

Поскольку $u_t = y_t \Gamma - x_t B$, то, перейдя от переменных u_1, \dots, u_n к переменным y_1, \dots, y_n , получим выражение для совместной плотности значений векторов y_1, \dots, y_n в виде:

$$p_Y(y_1, \dots, y_n) = |J| \left((\sqrt{2\pi})^g \sqrt{\det \Sigma} \right)^{-n} \times \exp \left\{ -\frac{1}{2} (y_t \Gamma - x_t B) \Sigma^{-1} (y_t \Gamma - x_t B)^T \right\},$$

где $|J|$ — якобиан перехода от переменных u_1, \dots, u_n к переменным y_1, \dots, y_n ,
 $|J| = |\det \Gamma|^n$.

Для взаимной однозначности такого перехода требуется, чтобы $|\det \Gamma| \neq 0$. Рассматривая правую часть последнего выражения как функцию от неизвестных параметров, составляющих матрицы Γ , B , Σ , при известных значениях Y , X , получаем функцию правдоподобия

$$L(\Gamma, B, \Sigma | Y, X) = |\det \Gamma|^n \cdot \left((\sqrt{2\pi})^g \sqrt{\det \Sigma} \right)^{-n} \times \exp \left\{ -\frac{1}{2} \sum_{t=1}^n (y_t \Gamma - x_t B) \Sigma^{-1} (y_t \Gamma - x_t B)^T \right\},$$

логарифм которой равен:

$$\ln L(\Gamma, B, \Sigma | Y, X) = n \ln |\det \Gamma| - \frac{n}{2} \ln(\det \Sigma) - \frac{ng}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n (y_t \Gamma - x_t B) \Sigma^{-1} (y_t \Gamma - x_t B)^T.$$

Продифференцировав последнее выражение по элементам матрицы Σ^{-1} и приравняв найденные производные к нулю, получим:

$$\Sigma = n^{-1} (Y\Gamma - XB)^T (Y\Gamma - XB).$$

Подставив данное соотношение для Σ в правую часть выражения для $\ln L(\Gamma, B, \Sigma | Y, X)$ и отбросив слагаемые, не зависящие от неизвестных параметров, получим **концентрированную логарифмическую функцию правдоподобия** (*concentrated log-likelihood function*):

$$\ln L^*(\Gamma, B | Y, X) = n \ln |\det \Gamma| - \frac{n}{2} \ln |n^{-1} (Y\Gamma - XB)^T (Y\Gamma - XB)|.$$

Оценки матриц Γ и B находятся максимизацией $\ln L^*$ с учетом всех ограничений, накладываемых на коэффициенты структурных уравнений. В частности, если используются только исключающие и нормировочные ограничения, то максимизация проводится только по неспецифицированным элементам этих матриц. В процессе такой минимизации матрица $(Y\Gamma - XB)^T (Y\Gamma - XB)$ должна иметь полный ранг для всех допустимых значений Γ и B . Необходимым для этого является условие $n \geq g + K$, которое может быть ограничительным для систем с большим количеством переменных.

Матрицы $\hat{\Gamma}$ и \hat{B} , получаемые в результате максимизации $\ln L^*$, и матрица $\hat{\Sigma} = n^{-1} (Y\hat{\Gamma} - X\hat{B})^T (Y\hat{\Gamma} - X\hat{B})$ вместе образуют оценку максимального правдо-

подобия, учитывающую полную информацию о структуре модели одновременных уравнений. Поэтому такую оценку называют **оценкой максимального правдоподобия с полной информацией** (*FIML* — *full information maximum likelihood*).

Пусть:

- выполнены сделанные выше предположения;
- ранговое условие идентифицируемости выполняется для всех структурных уравнений системы;
- матрица X имеет полный ранг;
- предельная матрица $p \lim_{n \rightarrow \infty} \left(\frac{1}{n} X^T X \right) = Q$ имеет конечные элементы и положительно определена.

Тогда *FIML*-оценка состоятельна и асимптотически нормальна. При этом требование нормальности распределения векторов u_1, \dots, u_n необязательно (см., например, (Amemiya, 1985, ch. 7)).

- ✓ **Замечание 1.2.5.** Если при выводе *FIML* отправляться не от структурной, а от приведенной формы системы, учитывающей ограничения, накладываемые на коэффициенты структурной формы, то тогда дело сводится к максимизации концентрированной функции правдоподобия

$$\ln L^{**}(\Gamma, B) = -\frac{n}{2} \ln \left| n^{-1} (Y - XB\Gamma^{-1})^T (Y - XB\Gamma^{-1}) \right|,$$

т.е. с учетом соотношения $\Pi = B\Gamma^{-1}$ — к минимизации целевой функции

$$Q(\Gamma, B) = \tilde{Q}(\Pi) = (Y - X\Pi)^T (Y - X\Pi)$$

по элементам матрицы Π при ограничениях, накладываемых на коэффициенты этой матрицы выбранной спецификацией матриц Γ и B .

Если не учитывать эти ограничения при минимизации целевой функции $\tilde{Q}(\Pi)$, то при сверхидентифицируемости отдельных уравнений системы возникает неоднозначность восстановления Γ и B по полученной оценке $\hat{\Pi}$. Если же все уравнения системы идентифицируемы точно, то значения Γ и B восстанавливаются однозначно, и восстановленные значения $\hat{\Gamma}$ и \hat{B} совпадают с оценками, полученными при минимизации $Q(\Gamma, B)$ по Γ и B .

- ✓ **Замечание 1.2.6.** При практической реализации метода *FIML* приходится использовать итерационные процедуры. Для обеспечения состоятельности и асимптотической нормальности *FIML*-оценки в качестве начальных значений параметров необходимо использо-

вать их состоятельные оценки. Их можно получить двухшаговым методом наименьших квадратов. Если система неидентифицируема, то итерационный процесс может не сходиться.

- ✓ **Замечание 1.2.7.** Перед применением *FIML* обычно производят исключение из системы тождеств и недоидентифицируемых уравнений.
- ✓ **Замечание 1.2.8.** В рекурсивной системе с диагональной ковариационной матрицей Σ оценка *FIML* получается путем применения *OLS* отдельно к каждому уравнению.

Пусть первое стохастическое структурное уравнение

$$y_1 = Y_1\alpha_1 + X_1\theta_1 + u_1$$

идентифицируемо, а другие уравнения либо неидентифицируемы, либо имеются сомнения в правильности их спецификации. Пусть при этом известны список всех предопределенных переменных, включаемых в систему, и значения этих предопределенных переменных в n наблюдениях, так что известна матрица X этих значений и можно говорить о приведенной форме системы:

$$Y = X\Pi + W.$$

Удалим из приведенной формы часть, относящуюся к y_1 , оставшаяся часть принимает вид:

$$Y_1 = X\Pi_1 + W_1.$$

Вместо полной системы структурных уравнений, или полной приведенной системы, рассмотрим теперь смешанную систему:

$$\begin{cases} y_1 = Y_1\alpha_1 + X_1\theta_1 + u_1 \\ Y_1 = X\Pi_1 + W_1. \end{cases}$$

Эту систему можно записать в виде:

$$\begin{bmatrix} y_1 & Y_1 \end{bmatrix} \begin{pmatrix} 1 & O \\ -\alpha_1 & I_{g-1} \end{pmatrix} = \begin{bmatrix} X_1 & X_1^* \end{bmatrix} \begin{pmatrix} \theta_1 \\ 0 \end{pmatrix} \Big| \Pi_1 + \begin{bmatrix} u_1 \\ W_1 \end{bmatrix}$$

и применить к ней метод *FIML*. Такая процедура приводит к оценке параметров α_1 , θ_1 , Π , называемой **оценкой максимального правдоподобия с ограниченной информацией** (*LIML* — *limited information maximum likelihood*).

При практической реализации этой процедуры итерационными методами в качестве начальных значений следует использовать состоятельные оценки параметров α_1 , θ_1 , Π_1 . Состоятельную оценку матрицы Π_1 можно получить

путем непосредственного применения метода наименьших квадратов к уравнениям системы $Y_1 = X\Pi_1 + W_1$. Состоятельные оценки параметров α_1 и θ_1 можно получить, применив двухшаговый метод наименьших квадратов.

Связь между различными оценками систем одновременных уравнений

Пусть интерес представляет оценивание отдельного уравнения структурной формы системы. Тогда:

- если оно идентифицируемо без запаса (точно), то достаточно применить косвенный метод наименьших квадратов — *ILS*;
- если оно сверхидентифицируемо, то можно применить *2SLS*, *LIML*, *3SLS*, *FIML*.

Чтобы применить *3SLS* и *FIML*, необходимо знать структуру всех уравнений системы и убедиться в идентифицируемости всех этих уравнений. Чтобы применить *2SLS* и *LIML*, достаточно знать только структуру рассматриваемого уравнения и список (и значения) всех предопределенных переменных, включенных в систему. В *2SLS* и *LIML* ошибка спецификации одного структурного уравнения системы локализуется в пределах этого уравнения, в *3SLS* и *FIML* такая ошибка влияет на оценку всех уравнений.

Предположим теперь, что выполнены указанные выше условия состоятельности оценок. Тогда:

- если i -е структурное уравнение идентифицируемо точно, то оценки *2SLS*, *LIML* и *ILS* совпадают; если же оно сверхидентифицируемо, то оценки *2SLS* и *LIML* имеют одинаковое асимптотическое распределение, но оценка *LIML* предпочтительнее при малом количестве наблюдений;
- если i -е структурное уравнение идентифицируемо, то *2SLS* дает состоятельные оценки параметров и $\sqrt{n}(\hat{\delta}_i^{2SLS} - \delta_i) \xrightarrow{d} N(0, C_2)$;
- если все структурные уравнения идентифицируемы, то *3SLS* дает состоятельные оценки параметров и $\sqrt{n}(\hat{\delta}_i^{3SLS} - \delta_i) \xrightarrow{d} N(0, C_3)$, причем матрица $C_2 - C_3$ неотрицательно определена (положительно полуопределена), так что *3SLS* приводит к более эффективным оценкам;
- если $\Sigma = I_g$ и все структурные уравнения идентифицируемы точно, то $\hat{\delta}_i^{3SLS} = \hat{\delta}_i^{2SLS}$;
- оценки *FIML* и *3SLS* имеют одинаковое асимптотическое распределение, при конечных n предпочтительнее *FIML*.

✓ **Замечание 1.2.9.** Если в i -м структурном уравнении системы $Y\Gamma = XB + U$ ошибки автокоррелированы, то для учета этой автокоррелированности можно использовать комбинацию *2SLS* и *GLS*, не при-

бегая к 3SLS. Пусть, например, ошибки в i -м уравнении следуют процессу авторегрессии первого порядка:

$$u_{ti} = \rho u_{t-1,i} + \eta_t, \quad |\rho| < 1.$$

Тогда естественно применить к i -му уравнению авторегрессионное преобразование (Кохрейна — Оркатта). Состоятельную оценку для ρ можно получить, оценив обычным методом наименьших квадратов (OLS) уравнение

$$\hat{u}_{ti}^{IV} = \rho \hat{u}_{t-1,i}^{IV} + v_t,$$

где \hat{u}_{ti}^{IV} — остатки, полученные в результате применения к i -му уравнению метода инструментальных переменных.

При этом для повышения эффективности оценивания к используемым в качестве инструментов в 2SLS переменным $x_t = (x_{t1}, \dots, x_{tK})$ можно добавить $y_{t-1} = (y_{t-1,1}, \dots, y_{t-1,g})$ и $x_{t-1} = (x_{t-1,1}, \dots, x_{t-1,K})$.

Проверка правильности спецификации системы одновременных уравнений

Выше уже говорилось (см. Замечание 1.2.3) о возможности проверки адекватности i -го структурного уравнения, опираясь на остатки $\hat{u}_i^{2SLS} = y_i - Z_i \hat{\delta}_i^{2SLS}$, полученные в результате применения двухшагового метода наименьших квадратов к этому уравнению, в отношении таких стандартных предположений, как линейность уравнения, нормальность, гомоскедастичность и некоррелированность ошибок.

Между тем не менее важным является вопрос о правильности подразделения включенных в систему переменных на эндогенные и экзогенные, произведенного на основании соответствующих экономических (предлагаемых экономической теорией) и логических представлений о связях между переменными. Еще одна проблема спецификации структурных уравнений состоит в том, что сверхидентифицируемость i -го уравнения системы

$$y_i = Y_i \alpha_i + X_i \theta_i + u_i = Z_i \delta_i + u_i$$

может быть просто следствием того, что на коэффициенты этого уравнения наложены ограничения, которых в действительности нет. Например, из i -го уравнения могут быть ошибочно исключены некоторые предопределенные переменные, включенные в другие уравнения системы. Хотелось бы иметь какой-то статистический инструментарий, позволяющий ответить на такие вопросы. Ряд статистических критериев, служащих этой цели, используют следующую идею Хаусмана (Hausman, 1978).

Пусть для $(p \times 1)$ -вектора параметров θ имеются две различные оценки $\hat{\theta}$ и $\tilde{\theta}$, причем оценка $\tilde{\theta}$ состоятельна и при гипотезе H_0 , и при альтернатив-

ной гипотезе H_A , а оценка $\hat{\theta}$ состоятельна и асимптотически эффективна при гипотезе H_0 , но не является состоятельной при гипотезе H_A . Рассмотрим разность этих двух оценок $\hat{q} = \hat{\theta} - \tilde{\theta}$. Поскольку при гипотезе H_0 обе оценки состоятельны, т.е. сходятся по вероятности к истинному значению θ , то их разность \hat{q} сходится по вероятности к нулю. Следовательно, если гипотеза H_0 верна, то не ожидается больших отклонений значения \hat{q} от нуля, и наличие таковых может трактоваться как указание на невыполнение гипотезы H_0 .

Критерий Хаусмана для проверки правильности спецификации системы одновременных уравнений (Hausman, 1978) использует в качестве $\hat{\theta}$ трехшаговую оценку наименьших квадратов, а в качестве $\tilde{\theta}$ — двухшаговую оценку наименьших квадратов. Если все структурные уравнения специфицированы правильно, то 3SLS состоятельна и эффективна; если же хотя бы одно из уравнений специфицировано неправильно, то 3SLS перестает быть состоятельной оценкой.

Однако, как было отмечено в работе (Spencer, Berk, 1981), для применения этого критерия необходима спецификация всех структурных уравнений системы, тогда как на практике чаще представляет интерес правильность спецификации какого-то отдельного структурного уравнения. В таком случае речь идет о проверке правильности спецификации i -го структурного уравнения при ограниченной информации об остальной части системы (как при построении LIML-оценки).

Статистика критерия Хаусмана определяется следующим образом:

$$H = n \hat{q}^T [asCov(\hat{q})]^{-1} \hat{q},$$

где $asCov(\hat{q})$ — состоятельная оценка асимптотической ковариационной матрицы $asCov(\hat{q})$ разности $\hat{q} = \hat{\theta} - \tilde{\theta}$, при выполнении достаточно общих условий она имеет при гипотезе H_0 асимптотическое распределение хи-квадрат.

Некоторые трудности при использовании этой статистики вызывает то обстоятельство, что компоненты вектора $\hat{q} = \hat{\theta} - \tilde{\theta}$ в общем случае линейно зависимы, вследствие чего матрица $asCov(\hat{q})$ может быть вырожденной и не иметь обратной в обычном смысле. В связи с этим в формуле для статистики H следует использовать не обычную обратную, а так называемую обобщенную обратную матрицу.

Указанные трудности можно обойти, если использовать асимптотически эквивалентные версии критерия Хаусмана, основанные на оценивании тех или иных уравнений регрессии. В таких вариантах критерия дело сводится к проверке значимости оцененных коэффициентов соответствующих уравнений.

Версия критерия Хаусмана, приведенная в (Davidson, MacKinnon, 1993) и называемая там **критерием Дарбина — Ву — Хаусмана** (*Durbin-Wu-Hausman test*), состоит в следующем.

Пусть

$$y_i = Y_i \alpha_i + X_i \theta_i + u_i = Z_i \delta_i + u_i,$$

где X_i — матрица значений инструментальных переменных в i -м уравнении;

Y_i — матрица значений тех объясняющих переменных в i -м уравнении, которые не входят в состав инструментальных переменных и не являются линейными комбинациями последних.

Гипотеза H_0 : в i -м уравнении отсутствует проблема эндогенности, т.е. все объясняющие переменные в составе Y_i не коррелированы с u_i . Иначе говоря, это **гипотеза экзогенности** (предопределенности) переменных, входящих в состав Y_i . Если эта гипотеза выполнена, то оценивание i -го уравнения можно производить обычным методом наименьших квадратов (*OLS*). В противном случае надо применять метод инструментальных переменных.

Сначала производится *OLS*-оценивание уравнений регрессии объясняющих переменных, входящих в состав Y_i , на инструментальные переменные:

$$Y_i = X \Pi_i + W_i$$

и вычисляются прогнозные значения \hat{Y}_i этих переменных. Затем эти прогнозные значения добавляются в качестве дополнительных объясняющих переменных в правую часть i -го уравнения, что приводит к расширенному уравнению

$$y_i = Z_i \delta_i + \hat{Y}_i \gamma + \eta_i,$$

производится *OLS*-оценивание расширенного уравнения и проверяется гипотеза H_0 : $\gamma = 0$. Для проверки этой гипотезы используется обычный *F*-критерий, хотя, вообще говоря, он является в этой ситуации только приближенным критерием.

Вместо \hat{Y}_i в расширенном уравнении можно использовать остатки $\hat{W}_i = Y_i - \hat{Y}_i$, т.е. оценивать уравнение

$$y_i = Z_i \delta_i + \hat{W}_i \gamma + \eta_i$$

и проверять гипотезу H_0 : $\gamma = 0$ в рамках этого уравнения. В любом случае отклонение гипотезы H_0 трактуется как наличие проблемы эндогенности, вызывающей несостоятельность *OLS*-оценок параметров i -го структурного уравнения.

Еще один вариант критерия Хаусмана для проверки той же гипотезы состоит в следующем. Наряду с остатками $\hat{W}_i = Y_i - \hat{Y}_i$, определенными выше, рассмотрим остатки \hat{u}_i , которые получаются при оценивании i -го уравнения обычным методом наименьших квадратов (*OLS*). Пусть R^2 — коэффициент детер-

минации, определяемый при OLS -оценивании уравнения $\hat{u}_i = Z_i\delta_i + \hat{W}_i\gamma + \xi_i$. Тогда при выполнении гипотезы экзогенности статистика nR^2 имеет асимптотическое (при $n \rightarrow \infty$) распределение $\chi^2(g_i)$, где g_i — количество переменных в составе Y_i . Эта гипотеза отвергается при $nR^2 > \chi^2_{1-\alpha}(g_i)$, где α — выбранный уровень значимости критерия.

Можно указать и некоторые другие варианты реализации критерия Хаусмана для проверки гипотезы об отсутствии проблемы эндогенности в i -м уравнении. Но как бы там ни было, прежде чем производить проверку на эндогенность тех или иных переменных, включенных в структурное уравнение, рекомендуется предварительно *проверить пригодность самих выбранных инструментов*. Такую проверку можно осуществлять в том случае, когда количество имеющихся инструментов превышает минимально необходимое их количество, и сделать это можно, используя, например, **J -статистику** (*J-statistic for valid instruments*), предложенную в работе (Godfrey, Hutton, 1994).

Пусть для очистки эндогенных переменных, входящих в правую часть i -го уравнения системы

$$y_i = Y_i\alpha_i + X_i\theta_i + u_i = Z_i\delta_i + u_i,$$

используется уравнение

$$Y_i = X \Pi_i + W_i,$$

где X — матрица значений инструментальных переменных.

Применив к i -му уравнению двухшаговый метод наименьших квадратов, получим $2SLS$ -остатки в виде

$$\hat{u}_i^{2SLS} = y_i - Z_i\hat{\delta}_i^{2SLS}.$$

После этого оценим линейную модель регрессии \hat{u}_i^{2SLS} на переменные, входящие в состав X . Пусть R^2 — полученное при этом значение коэффициента детерминации. Указанная J -статистика равна: $J = nR^2$ и имеет асимптотическое (при $n \rightarrow \infty$) распределение хи-квадрат с числом степеней свободы, равным разности между количеством переменных в составе X и количеством объясняющих переменных в i -м уравнении.

Гипотеза пригодности выбранного множества инструментов отвергается при значениях J -статистики, превышающих критическое значение, рассчитанное по указанному распределению хи-квадрат (т.е. при значениях J -статистики, для которых P -значение оказывается меньше заданного уровня значимости). Если это происходит, нет смысла заниматься оцениванием коэффициентов рассматриваемого уравнения с выбранным множеством инструментов, поскольку в этом случае или сами эти инструменты непригодны, или уравнение неправильно специфицировано.

Если указанная гипотеза не отвергается J -критерием, переходят ко второму шагу, на котором используется критерий Хаусмана (в том или ином его

варианте) для проверки переменных в i -м уравнении системы на эндогенность/экзогенность.

В работе (Godfrey, Hutton, 1994) показано, что статистики, используемые в такой двухступенчатой процедуре, асимптотически независимы, так что вероятность ошибочного решения в этой процедуре приближенно равна: $1 - (1 - \alpha_J)(1 - \alpha_H) = \alpha_J + \alpha_H - \alpha_J\alpha_H$, где α_J — уровень значимости J -критерия; α_H — уровень значимости критерия Хаусмана, используемого на втором шаге.

✓ **Замечание 1.2.10.** Отклонение нулевой гипотезы при применении критериев экзогенности означает только то, что проблема эндогенности существует. Однако степень влияния обнаруженной эндогенности на смещение обычных оценок наименьших квадратов остается при этом неизвестной. В то же время мощность критериев типа критерия Хаусмана становится довольно низкой, если инструменты слабо коррелированы с эндогенными переменными. И это означает, что нулевая гипотеза экзогенности может быть не отвергнута, а смещение OLS -оценок велико. Поэтому во многих практических исследованиях авторы сообщают и результаты оценивания методом инструментальных переменных, и результаты OLS -оценивания.

Примеры оценивания систем одновременных уравнений

ПРИМЕР 1.2.1

Рассмотрим модель «спрос — предложение» в виде:

$$\begin{cases} P = \alpha_1 Q + \theta_{11} + \theta_{12} DPI + u_1 \\ Q = \alpha_2 P + \theta_{21} + \theta_{22} Weather + \theta_{23} Invest + u_2, \end{cases}$$

где P — розничная цена свежих фруктов, выраженная в постоянных ценах с использованием индекса розничных цен;

Q — объем потребления свежих фруктов на душу населения;

DPI — располагаемый доход на душу населения, дефлированный на индекс потребительских цен (CPI);

$Weather$ — климатическая характеристика, отражающая размер потенциальных потерь урожая из-за неблагоприятных погодных условий;

$Invest$ — дефлированный на CPI объем на душу населения чистых инвестиций производителей свежих фруктов, отражающий издержки производства.

Первое уравнение является уравнением спроса, а второе — уравнением предложения.

Всего имеется 30 наблюдений, все переменные выражены в индексной форме с одним и тем же базовым периодом (табл. 1.4).

Таблица 1.4

Данные 30 наблюдений в модели «спрос — предложение»

t	Цена (P)	Потребление (Q)	DPI	$Weather$	$Invest$
1	108.9	127.4	97.6	99.1	142.9
2	100.6	105.1	98.2	98.9	123.8
3	109.7	76.7	99.8	110.8	111.9
4	111.6	93.8	100.5	108.2	121.4
5	109.8	88.3	96.6	108.7	92.9
6	104.4	78.4	88.9	100.6	97.6
7	89.6	89.6	84.6	70.9	64.3
8	117.2	75.3	96.4	110.5	78.6
9	109.3	109.1	104.4	92.5	109.5
10	114.9	121.3	110.7	89.3	128.6
11	112.0	106.3	99.1	90.3	95.8
12	112.9	129.1	105.6	95.2	130.9
13	121.0	118.6	116.8	98.6	125.7
14	112.8	94.3	105.3	105.7	109.8
15	102.9	81.0	85.6	107.8	88.4
16	86.0	104.9	84.8	80.4	96.9
17	95.7	94.6	89.8	90.7	90.8
18	104.9	102.9	93.2	88.9	101.7
19	114.0	110.6	105.9	96.9	110.8
20	121.9	111.7	110.8	101.9	117.9
21	127.2	117.6	115.3	104.9	134.8
22	128.3	125.1	120.6	103.6	140.2
23	125.0	87.4	105.7	106.2	78.3
24	117.1	84.6	103.5	100.8	94.7
25	122.7	107.8	110.6	110.5	135.9
26	111.6	120.7	109.3	86.7	126.8
27	114.1	102.8	99.5	93.8	90.5
28	110.4	99.2	105.9	99.9	134.8
29	109.2	107.1	102.7	104.0	123.8
30	108.9	127.4	97.6	99.1	142.9

С помощью обозначений, использованных ранее при рассмотрении систем одновременных уравнений, запишем систему в виде:

$$\begin{cases} y_{t1} = \alpha_{11}y_{t2} + \theta_{11}x_{t1} + \theta_{21}x_{t2} + u_{t1} \\ y_{t2} = \alpha_{12}y_{t1} + \theta_{12}x_{t1} + \theta_{22}x_{t3} + \theta_{32}x_{t4} + u_{t2}, \end{cases}$$

где $y_{t1} = P_t$, $y_{t2} = Q_t$, $x_{t1} \equiv 1$, $x_{t2} = (DPI)_t$, $x_{t3} = (Weather)_t$, $x_{t4} = (Invest)_t$.

Список эндогенных переменных: (y_{t1}, y_{t2}) . Список экзогенных переменных: $(1, x_{t2}, x_{t3}, x_{t4})$. Полный список переменных, включенных в систему: $(y_{t1}, y_{t2}, 1, x_{t2}, x_{t3}, x_{t4})$. Соответственно $g = 2, K = 4$,

$$\Gamma = \begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{11} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & 0 \\ 0 & \theta_{22} \\ 0 & \theta_{32} \end{pmatrix}, \quad A = \begin{pmatrix} \Gamma \\ B \end{pmatrix} = \begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{11} & 1 \\ \theta_{11} & \theta_{12} \\ \theta_{21} & 0 \\ 0 & \theta_{22} \\ 0 & \theta_{32} \end{pmatrix}.$$

На элементы первого столбца матрицы A помимо нормировочного накладывается два исключаяющих ограничения $\alpha_{51} = 0, \alpha_{61} = 0$, так что для этого столбца $g_1^* = 0, K_1^* = 2$, и $g_1^* + K_1^* > g - 1$, т.е. порядковое условие идентифицируемости выполняется. На элементы второго столбца помимо нормировочного накладывается одно исключаяющее ограничение $\alpha_{42} = 0$, так что для этого столбца $g_2^* = 0, K_2^* = 1$, и $g_2^* + K_2^* = g - 1$, т.е. порядковое условие идентифицируемости выполняется.

Для проверки выполнения ранговых условий идентифицируемости воспользуемся Замечанием 1.1.3. В соответствии с ним построим таблицу коэффициентов (табл. 1.5).

Таблица 1.5

Коэффициенты для полного списка переменных

i	y_{t1}	y_{t2}	1	x_{t2}	x_{t3}	x_{t4}
1	1	α_{11}	θ_{11}	θ_{21}	0	0
0	α_{12}	1	θ_{12}	0	θ_{22}	θ_{32}

Для первого уравнения выделяемая матрица сводится к одной строке с двумя элементами: $(\theta_{22} \ \theta_{32})$. Ранг этой матрицы равен 1, что совпадает со значением $g - 1 = 1$, так что первое уравнение идентифицируемо. Для второго уравнения выделяемая матрица сводится к одному элементу: (θ_{21}) . Ранг этой матрицы также равен 1, так что и второе уравнение идентифицируемо. Разница только в том, что для первого уравнения $g_1^* + K_1^* > g - 1$, а для второго $g_2^* + K_2^* = g - 1$, т.е. первое уравнение сверхидентифицируемо, а второе — идентифицируемо точно. Соответственно для оценивания второго уравнения можно использовать косвенный метод наименьших квадратов, а для оценивания первого уравнения этот метод не годится.

Чтобы применить косвенный метод наименьших квадратов, сначала отдельно оценим методом наименьших квадратов уравнения приведенной формы:

$$y_{t1} = \pi_{11} + \pi_{21}x_{t2} + \pi_{31}x_{t3} + \pi_{41}x_{t4} + w_{t1},$$

$$y_{t2} = \pi_{12} + \pi_{22}x_{t2} + \pi_{32}x_{t3} + \pi_{42}x_{t4} + w_{t2}.$$

В пакете *EViews* это дает результаты, приведенные в табл. 1.6 (при расчетах использовались значения P и Q с 4 десятичными знаками).

Таблица 1.6

Оценка уравнений с помощью метода наименьших квадратов

Объясняемая переменная Y1				
Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	−12.409540	8.192675	−1.514712	0.1419
X2	1.030854	0.090988	11.32951	0.0000
X3	0.361564	0.066508	5.436388	0.0000
X4	−0.152442	0.040203	−3.791820	0.0008
R-squared	0.902361	Mean dependent var		111.1445
Adjusted R-squared	0.891094	S.D. dependent var		9.877858
S.E. of regression	3.259777	Akaike info criterion		5.324760
Sum squared resid	276.2797	Schwarz criterion		5.511587
Log likelihood	−75.87140	F-statistic		80.09524
Durbin-Watson stat	2.016289	Prob. (F-statistic)		0.000000
Объясняемая переменная Y2				
Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	81.78495	17.56752	4.655463	0.0001
X2	0.581396	0.195106	2.979896	0.0062
X3	−0.924096	0.142613	−6.479734	0.0000
X4	0.475229	0.086207	5.512656	0.0000
R-squared	0.824120	Mean dependent var		101.8111
Adjusted R-squared	0.803826	S.D. dependent var		15.78165
S.E. of regression	6.989927	Akaike info criterion		6.850383
Sum squared resid	1270.336	Schwarz criterion		7.037209
Log likelihood	−98.755750	F-statistic		40.60943
Durbin-Watson stat	2.084533	Prob. (F-statistic)		0.000000

Используем теперь соотношение $\Pi\Gamma = V$, которое в нашем примере принимает вид:

$$\begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \\ \pi_{31} & \pi_{32} \\ \pi_{41} & \pi_{42} \end{pmatrix} \begin{pmatrix} 1 & -\alpha_{12} \\ -\alpha_{11} & 1 \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & 0 \\ 0 & \theta_{22} \\ 0 & \theta_{32} \end{pmatrix}$$

и приводит к уравнениям:

$$\begin{aligned} \pi_{11} - \pi_{12}\alpha_{11} &= \theta_{11}, & \pi_{12} - \pi_{11}\alpha_{12} &= \theta_{12}, \\ \pi_{21} - \pi_{22}\alpha_{11} &= \theta_{21}, & \pi_{22} - \pi_{21}\alpha_{12} &= 0, \\ \pi_{31} - \pi_{32}\alpha_{11} &= 0, & \pi_{32} - \pi_{31}\alpha_{12} &= \theta_{22}, \\ \pi_{41} - \pi_{42}\alpha_{11} &= 0, & \pi_{42} - \pi_{41}\alpha_{12} &= \theta_{32}, \end{aligned}$$

Поскольку точно идентифицируемо только второе структурное уравнение системы, интерес для применения косвенного метода наименьших квадратов представляют лишь коэффициенты этого уравнения α_{12} , θ_{12} , θ_{22} и θ_{32} . Это означает, что из 8 приведенных уравнений достаточно рассмотреть только 4 уравнения, стоящих в правом столбце. Решив эти уравнения, имеем:

$$\alpha_{12} = \frac{\pi_{22}}{\pi_{21}}, \quad \theta_{12} = \pi_{11} - \pi_{12} \frac{\pi_{22}}{\pi_{21}}, \quad \theta_{22} = \pi_{31} - \pi_{32} \frac{\pi_{22}}{\pi_{21}}, \quad \theta_{32} = \pi_{41} - \pi_{42} \frac{\pi_{22}}{\pi_{21}}.$$

Подставив в правые части оцененные значения коэффициентов π_{ki} , найдем оценки для коэффициентов второго структурного уравнения. Например:

$$\hat{\alpha}_{12} = \frac{\hat{\pi}_{22}}{\hat{\pi}_{21}} = \frac{0.581396}{1.030854} = 0.5639945.$$

В отношении трех остальных коэффициентов получим:

$$\hat{\theta}_{12} = 88.78386, \quad \hat{\theta}_{22} = -1.128017, \quad \hat{\theta}_{32} = 0.561206. \blacksquare$$

✓ **Замечание 1.2.11.** В таблицах результатов применения косвенного метода наименьших квадратов обычно не приводятся значения стандартных ошибок коэффициентов, поскольку из-за нелинейности соотношений между коэффициентами структурной и приведенной форм вычисление стандартных ошибок оценок коэффициентов при конечных n затруднительно. В то же время при применении двухшагового метода наименьших квадратов для вычисления этих ошибок имеются соответствующие формулы. Поэтому можно было вычислить искомые стандартные ошибки оценок коэффициентов первого уравнения рассматриваемой системы, используя *2SLS* и имея в виду, что в случае точно идентифицируемого уравнения результаты оце-

нивания его коэффициентов методами *ILS* и *2SLS* совпадают. Проблема, однако, в том, что в этой ситуации у *2SLS*-оценки не существует конечных выборочных моментов. Соответственно по сравнению с нормальным распределением оценки более часто далеко отклоняются от истинных значений параметров, и это затрудняет интерпретацию полученных результатов.

Имея в виду сделанное замечание, применим все же двухшаговый метод наименьших квадратов для оценивания обоих структурных уравнений (результаты приведены в табл. 1.7).

Таблица 1.7

Оценка уравнений с помощью двухшагового метода наименьших квадратов
Estimation Method: Two-Stage Least Squares (2SLS)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.361831	0.081657	−4.431122	0.0000
C(2)	18.16639	9.760145	1.861283	0.0683
C(3)	1.279190	0.126018	10.15089	0.0000
C(4)	0.563994	0.175307	3.217171	0.0022
C(5)	88.78386	15.06578	5.893080	0.0000
C(6)	−1.128017	0.158217	−7.129565	0.0000
C(7)	0.561206	0.065442	8.575557	0.0000
Determinant residual covariance		479.5253		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.781184	Mean dependent var		111.1445
Adjusted R-squared	0.764975	S.D. dependent var		9.877858
S.E. of regression	4.788722	Sum squared resid		619.1603
Durbin-Watson stat	2.036078			
Equation: Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4; Observations: 30				
R-squared	0.849107	Mean dependent var		101.8111
Adjusted R-squared	0.831696	S.D. dependent var		15.78165
S.E. of regression	6.474401	Sum squared resid		1089.865
Durbin-Watson stat	2.152230			

Оценки всех коэффициентов, кроме постоянной составляющей в первом уравнении, имеют высокую статистическую значимость. Отрицательное значение оценки коэффициента при переменной y_{t2} в первом уравнении согласо-

ется с тем, что это уравнение является уравнением спроса. Положительное значение оценки при переменной y_{i1} во втором уравнении согласуется с тем, что второе уравнение является уравнением предложения. Априорным предположениям также соответствуют знаки оцененных коэффициентов при переменных x_{i1} , x_{i2} , x_{i3} и x_{i4} (увеличение спроса — при возрастании дохода, снижение предложения — при усилении неблагоприятных погодных факторов и увеличение предложения — при возрастании инвестиций).

Вместе с тем, если произвести оценивание корреляционной матрицы ошибок в структурных уравнениях, то оцененная на основании векторов остатков $\hat{u}_i^{2SLS} = y_i - Z_i \hat{\delta}_i^{2SLS}$ ковариационная матрица примет вид:

$$\begin{pmatrix} 20.638675 & 16.439387 \\ 16.439387 & 36.328821 \end{pmatrix}.$$

Ей соответствует оцененная корреляционная матрица

$$\begin{pmatrix} 1 & 0.600370 \\ 0.600370 & 1 \end{pmatrix},$$

указывающая на наличие заметной корреляции между ошибками в разных уравнениях. Это означает, что потенциально имеется возможность повысить эффективность оценивания, если учесть такую коррелированность и применить трехшаговый метод наименьших квадратов или метод максимального правдоподобия с полной информацией.

При применении 3SLS получим результаты, приведенные в табл. 1.8.

Таблица 1.8

Оценка уравнений с помощью трехшагового метода наименьших квадратов

*Estimation Method: Iterative Three-Stage Least Squares (3SLS);
Convergence achieved after: 2 weight matrices, 3 total coef iterations*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	-0.361831	0.077466	-4.670812	0.0000
C(2)	18.16639	9.259287	1.961964	0.0550
C(3)	1.279190	0.119551	10.69998	0.0000
C(4)	0.592909	0.150576	3.937596	0.0002
C(5)	89.91484	13.79846	6.516294	0.0000
C(6)	-1.159822	0.130071	-8.916811	0.0000
C(7)	0.550297	0.056107	9.808057	0.0000
<i>Determinant residual covariance</i>		486.0481		

Окончание табл. 1.8

Equation: $Y1=C(1)*Y2+C(2)+C(3)*X2$; Observations: 30			
<i>R-squared</i>	0.781184	<i>Mean dependent var</i>	111.1445
<i>Adjusted R-squared</i>	0.764975	<i>S.D. dependent var</i>	9.877858
<i>S.E. of regression</i>	4.788722	<i>Sum squared resid</i>	619.1603
<i>Durbin-Watson stat</i>	2.036078		
Equation: $Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4$; Observations: 30			
<i>R-squared</i>	0.849356	<i>Mean dependent var</i>	101.8111
<i>Adjusted R-squared</i>	0.831974	<i>S.D. dependent var</i>	15.78165
<i>S.E. of regression</i>	6.469047	<i>Sum squared resid</i>	1088.063
<i>Durbin-Watson stat</i>	2.139746		

Сравним оцененные коэффициенты и оцененные стандартные ошибки оценок коэффициентов, полученные двумя методами (табл. 1.9). Оцененные значения коэффициентов практически не изменились. При этом произошло некоторое уменьшение всех оцененных стандартных ошибок коэффициентов.

Таблица 1.9

Данные, полученные с помощью методов 2SLS и 3SLS

Переменная	Коэффициент		Стандартная ошибка	
	2SLS	3SLS	2SLS	3SLS
C(1)	-0.361831	-0.361831	0.081657	0.077466
C(2)	18.16639	18.16639	9.760145	9.259287
C(3)	1.279190	1.279190	0.126018	0.119551
C(4)	0.563994	0.592909	0.175307	0.150576
C(5)	88.78386	89.91484	15.06578	13.79846
C(6)	-1.128017	-1.159822	0.158217	0.130071
C(7)	0.561206	0.550297	0.065442	0.056107

Результаты применения метода *FIML* приведены в табл. 1.10. Оцененные значения коэффициентов незначительно отличаются от полученных ранее.

В связи с рассматриваемым примером обратим внимание на следующее обстоятельство. Оценим уравнения структурной формы обычным методом наименьших квадратов, игнорируя эндогенность некоторых переменных в правых частях этих уравнений (результаты приведены в табл. 1.11).

Таблица 1.10

Оценка уравнений с помощью метода *FIML*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.363540	0.097355	−3.734157	0.0005
C(2)	18.18156	13.65099	1.331886	0.1886
C(3)	1.280771	0.167147	7.662534	0.0000
C(4)	0.593625	0.221715	2.677418	0.0099
C(5)	88.82550	19.84547	4.526248	0.0000
C(6)	−1.159160	0.151139	−7.669539	0.0000
C(7)	0.549823	0.065905	8.342675	0.0000
Log Likelihood		−172.0962		
Determinant residual covariance		486.9397		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.780336	Mean dependent var		111.1445
Adjusted R-squared	0.764065	S.D. dependent var		9.877858
S.E. of regression	4.797990	Sum squared resid		621.5590
Durbin-Watson stat	2.030828			
Equation: Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4; Observations: 30				
R-squared	0.849367	Mean dependent var		101.8111
Adjusted R-squared	0.831986	S.D. dependent var		15.78165
S.E. of regression	6.468814	Sum squared resid		1087.984
Durbin-Watson stat	2.137757			

Таблица 1.11

Оценка уравнений с помощью метода *OLS* без учета эндогенности переменных
Estimation Method: Two-Stage Least Squares

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.213716	0.063693	−3.355431	0.0015
C(2)	18.159740	8.908678	2.038432	0.0465
C(3)	1.130662	0.108276	10.44241	0.0000
C(4)	0.614038	0.159561	3.848293	0.0003
C(5)	86.61460	14.71164	5.887487	0.0000
C(6)	−1.154140	0.153404	−7.523517	0.0000
C(7)	0.553841	0.064458	8.592225	0.0000
Determinant residual covariance		517.6508		

Окончание табл. 1.11

Equation: $Y1=C(1)*Y2+C(2)+C(3)*X2$; Observations: 30			
<i>R-squared</i>	0.817697	<i>Mean dependent var</i>	111.1445
<i>Adjusted R-squared</i>	0.804193	<i>S.D. dependent var</i>	9.877858
<i>S.E. of regression</i>	4.370958	<i>Sum squared resid</i>	515.8424
<i>Durbin-Watson stat</i>	2.442191		
Equation: $Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4$; Observations: 30			
<i>R-squared</i>	0.849675	<i>Mean dependent var</i>	101.8111
<i>Adjusted R-squared</i>	0.832330	<i>S.D. dependent var</i>	15.78165
<i>S.E. of regression</i>	6.462188	<i>Sum squared resid</i>	1085.757
<i>Durbin-Watson stat</i>	2.138293		

Сравним оценки коэффициентов, полученные при применении системного метода оценивания *FIML* и несистемного *OLS* (табл. 1.12).

Таблица 1.12

Коэффициенты, полученные с помощью методов *FIML* и *OLS*

Переменная	Коэффициент	
	<i>FIML</i>	<i>OLS</i>
$C(1)$	-0.363540	-0.213716
$C(2)$	18.18156	18.15974
$C(3)$	1.280771	1.130662
$C(4)$	0.593625	0.614038
$C(5)$	88.82550	86.61460
$C(6)$	-1.159165	-1.154140
$C(7)$	0.549823	0.553841

За исключением коэффициента при переменной y_{t2} в первом уравнении, оценки всех остальных коэффициентов структурных уравнений, полученные двумя методами, довольно близки друг к другу. И это характерно для ситуаций, когда при оценивании уравнений приведенной формы получаются высокие значения коэффициентов детерминации. В нашем примере это значения 0.902361 для первого уравнения и 0.824120 для второго уравнения приведенной формы.

Рассмотрим теперь, что получится, если во второе уравнение не включена одна из переменных — x_{t3} или x_{t4} (при этом система остается идентифицируемой). Будем ориентироваться на результаты применения *3SLS*. Сначала исключим переменную x_{t4} — издержки производства (результаты приведены в табл. 1.13).

Таблица 1.13

Оценка уравнений с исключенной переменной с помощью метода 3SLS*Estimation Method: Iterative Three-Stage Least Squares (3SLS);**Convergence achieved after: 1 weight matrix, 2 total coef iterations*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.393021	0.106679	−3.684137	0.0005
C(2)	18.16779	9.612186	1.890079	0.0641
C(3)	1.310468	0.142630	9.187882	0.0000
C(4)	1.603271	0.324620	4.938909	0.0000
C(5)	68.85896	29.34434	2.346584	0.0226
C(6)	−1.469719	0.308619	−4.762247	0.0000
Determinant residual covariance		3322.482		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.764186	Mean dependent var		111.1445
Adjusted R-squared	0.746719	S.D. dependent var		9.877858
S.E. of regression	4.971235	Sum squared resid		667.2557
Durbin-Watson stat	1.946538			
Equation: Y2=C(4)*Y1+C(5)+C(6)*X3; Observations: 30				
R-squared	0.360021	Mean dependent var		101.8111
Adjusted R-squared	0.312615	S.D. dependent var		15.78165
S.E. of regression	13.08436	Sum squared resid		4622.411
Durbin-Watson stat	1.199501			

Сравним коэффициенты, полученные при исключении x_{t4} из второго уравнения, с коэффициентами, полученными 3SLS без исключения этой переменной из второго уравнения (табл. 1.14).

Таблица 1.14

Коэффициенты уравнений с исключенной переменной и с ней (3SLS)

Переменная	Коэффициент	
	<i>X4 excluded</i>	<i>X4 included</i>
C(1)	−0.393021	−0.361831
C(2)	18.16779	18.16639
C(3)	1.310468	1.279190
C(4)	1.603271	0.592909
C(5)	68.85896	89.91484
C(6)	−1.469719	−1.159822

Произошло более чем двукратное возрастание коэффициента при переменной y_{t1} во втором уравнении. Проанализировав результаты в табл. 1.13, отметим довольно малое значение статистики Дарбина — Уотсона для второго уравнения, что может указывать на пропуск существенной объясняющей переменной в этом уравнении. Отметим также довольно большое значение оцененной стандартной ошибки случайной составляющей во втором уравнении — 13.08436 против 4.971235, что, конечно, вполне допустимо, но также может указывать на пропуск объясняющей переменной во втором уравнении.

Если вместо переменной x_{t4} исключить из второго уравнения переменную x_{t3} — погодный фактор, это приведет к результату, указанному в табл. 1.15.

Таблица 1.15

Оценка уравнений с исключенной переменной с помощью метода 3SLS*Estimation Method: Iterative Three-Stage Least Squares (3SLS)*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.318323	0.116740	−2.726758	0.0086
C(2)	18.16444	8.863637	2.049321	0.0453
C(3)	1.235561	0.145848	8.471570	0.0000
C(4)	1.274776	0.265732	1.034037	0.3057
C(5)	16.26065	26.12243	0.622478	0.5362
C(7)	0.499415	0.113699	4.392423	0.0001
Determinant residual covariance		2157.560		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.799484	Mean dependent var		111.1445
Adjusted R-squared	0.784631	S.D. dependent var		9.877858
S.E. of regression	4.584099	Sum squared resid		567.3771
Durbin-Watson stat	2.166955			
Equation: Y2=C(4)*Y1+C(5)+C(7)*X4; Observations: 30				
R-squared	0.490954	Mean dependent var		101.8111
Adjusted R-squared	0.453247	S.D. dependent var		15.78165
S.E. of regression	11.66938	Sum squared resid		3676.710
Durbin-Watson stat	1.624247			

Сравним коэффициенты, полученные при исключении x_{t3} из второго уравнения, с коэффициентами, полученными 3SLS без исключения этой переменной из второго уравнения (табл. 1.16).

Таблица 1.16

Коэффициенты уравнений без одной переменной и с ней (3SLS)

Переменная	Коэффициент	
	$X3\ excluded$	$X3\ included$
$C(1)$	−0.318323	−0.361831
$C(2)$	18.16444	18.16639
$C(3)$	1.235561	1.279190
$C(4)$	0.274776	0.592909
$C(5)$	16.26065	89.91484
$C(6)$	—	−1.159822
$C(7)$	0.499415	0.550297

На этот раз произошло более чем двукратное уменьшение коэффициента при переменной y_{t1} во втором уравнении. Опять наблюдается довольно большое значение оцененной стандартной ошибки случайной составляющей во втором уравнении — 11.66938 против 4.584099 в первом уравнении, что может указывать на пропуск объясняющей переменной во втором уравнении. Заметим, что при включении в правую часть второго уравнения обеих переменных — x_{t3} и x_{t4} такого большого различия оцененных стандартных ошибок случайных составляющих не наблюдается: в первом уравнении оцененное значение равно 4.788722, а во втором 6.469047. При этом вполне удовлетворительными выглядят и значения статистик Дарбина — Уотсона: 2.036078 в первом и 2.139746 во втором уравнениях.

ПРИМЕР 1.2.2

Чтобы составить представление о возможности более или менее точного восстановления параметров структурной формы, сгенерируем 30 новых «наблюдений», используя процесс порождения данных в виде:

$$\begin{cases} y_{t1} = -0.5y_{t2} + 80 + 0.7x_{t2} + u_{t1} \\ y_{t2} = 0.75y_{t1} + 10 - 1.5x_{t3} + 1.5x_{t4} + u_{t2}, \end{cases}$$

где переменные имеют ту же интерпретацию и значения переменных x_{t2} , x_{t3} и x_{t4} те же, что и в предыдущем примере, а случайные составляющие u_{t1} , u_{t2} не коррелированы во времени, имеют нулевые ожидания, одинаковые дисперсии $D(u_{t1}) = D(u_{t2}) = 36$ и $Corr(u_{t1}, u_{t2}) = 0.7$. Полученные «данные» приведены в табл. 1.17.

Таблица 1.17

Сгенерированные данные 30 наблюдений

t	Y_1	Y_2	X_2	X_3	X_4
1	88.1	152.4	97.6	99.1	142.9
2	87.1	114.6	98.2	98.9	123.8
3	101.1	74.1	99.8	110.8	111.9
4	98.6	99.2	100.5	108.2	121.4
5	109.4	74.0	96.6	108.7	92.9
6	105.2	72.4	88.9	100.6	97.6
7	100.6	69.2	84.6	70.9	64.3
8	121.7	51.3	96.4	110.5	78.6
9	96.0	109.7	104.4	92.5	109.5
10	90.3	135.5	110.7	89.3	128.6
11	105.7	100.2	99.1	90.3	95.8
12	91.3	144.5	105.6	95.2	130.9
13	95.5	126.1	116.8	98.6	125.7
14	101.0	90.6	105.3	105.7	109.8
15	110.2	66.7	85.6	107.8	88.4
16	87.1	104.0	84.8	80.4	96.9
17	97.4	85.3	89.8	90.7	90.8
18	99.8	102.6	93.2	88.9	101.7
19	99.9	110.8	105.9	96.9	110.8
20	102.8	114.8	110.8	101.9	117.9
21	99.8	131.2	115.3	104.9	134.8
22	95.8	141.2	120.6	103.6	140.2
23	123.4	62.6	105.7	106.2	78.3
24	110.4	73.2	103.5	100.8	94.7
25	98.7	121.1	110.6	110.5	135.9
26	88.2	134.5	109.3	86.7	126.8
27	110.0	91.8	99.5	93.8	90.5
28	88.3	115.1	105.9	99.9	134.8
29	93.6	114.5	102.7	104,0	123.8
30	103.5	76.9	96.8	108.4	104.5

Оценим по сгенерированным данным уравнения структурной формы обычным методом наименьших квадратов, игнорируя наличие в правых частях этих уравнений эндогенных переменных (результаты приведены в табл. 1.18).

Таблица 1.18

Оценка уравнений с помощью обычного метода наименьших квадратов*Estimation Method: Least Squares (OLS)*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.363218	0.041816	−8.686139	0.0000
C(2)	88.69833	10.49401	8.452281	0.0000
C(3)	0.476608	0.122697	3.884420	0.0003
C(4)	1.019773	0.297976	3.422330	0.0012
C(5)	−11.07961	28.30156	−0.391484	0.6970
C(6)	−1.752307	0.214654	−8.163409	0.0000
C(7)	1.672787	0.127403	13.12991	0.0000
Determinant residual covariance		576.6877		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.743523	Mean dependent var		100.0167
Adjusted R-squared	0.724524	S.D. dependent var		9.334551
S.E. of regression	4.899311	Sum squared resid		648.0877
Durbin-Watson stat	2.154319			
Equation: Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4; Observations: 30				
R-squared	0.950061	Mean dependent var		102.0033
Adjusted R-squared	0.944299	S.D. dependent var		27.23976
S.E. of regression	6.428860	Sum squared resid		1074.586
Durbin-Watson stat	2.151388			

Применим двухшаговый метод наименьших квадратов (результаты приведены в табл. 1.19).

Таблица 1.19

Оценка уравнений с помощью двухшагового метода наименьших квадратов*Estimation Method: Two-Stage Least Squares (2SLS)*

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.424633	0.045795	−9.272577	0.0000
C(2)	83.95884	10.96205	7.659048	0.0000
C(3)	0.585040	0.130033	4.499177	0.0000
C(4)	0.901844	0.420366	2.145379	0.0365
C(5)	−1.113488	37.813630	−0.029447	0.9766
C(6)	−1.684157	0.274838	−6.127827	0.0000
C(7)	1.628247	0.169687	9.595614	0.0000
<i>Determinant residual covariance</i>		522.6707		

Окончание табл. 1.19

Equation: $Y1=C(1)*Y2+C(2)+C(3)*X2$; Observations: 30			
<i>R-squared</i>	0.723032	<i>Mean dependent var</i>	100.0167
<i>Adjusted R-squared</i>	0.702516	<i>S.D. dependent var</i>	9.334551
<i>S.E. of regression</i>	5.091263	<i>Sum squared resid</i>	699.8659
<i>Durbin-Watson stat</i>	1.952744		
Equation: $Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4$; Observations: 30			
<i>R-squared</i>	0.949761	<i>Mean dependent var</i>	102.0033
<i>Adjusted R-squared</i>	0.943964	<i>S.D. dependent var</i>	27.23976
<i>S.E. of regression</i>	6.448196	<i>Sum squared resid</i>	1081.060
<i>Durbin-Watson stat</i>	2.162986		

Применив 3SLS, получим результаты, представленные в табл. 1.20.

Таблица 1.20

Оценка уравнений с помощью трехшагового метода наименьших квадратов

Estimation Method: Iterative Three-Stage Least Squares (3SLS)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	-0.424633	0.043445	-9.774155	0.0000
C(2)	83.95884	10.39951	8.073345	0.0000
C(3)	0.585040	0.123360	4.742549	0.0000
C(4)	0.889189	0.369342	2.407495	0.0196
C(5)	-0.653169	34.88671	-0.018723	0.9851
C(6)	-1.671481	0.220626	-7.576096	0.0000
C(7)	1.624186	0.152419	10.65606	0.0000
<i>Determinant residual covariance</i>		518.5160		
Equation: $Y1=C(1)*Y2+C(2)+C(3)*X2$; Observations: 30				
<i>R-squared</i>	0.723032	<i>Mean dependent var</i>	100.0167	
<i>Adjusted R-squared</i>	0.702516	<i>S.D. dependent var</i>	9.334551	
<i>S.E. of regression</i>	5.091263	<i>Sum squared resid</i>	699.8659	
<i>Durbin-Watson stat</i>	1.952744			
Equation: $Y2=C(4)*Y1+C(5)+C(6)*X3+C(7)*X4$; Observations: 30				
<i>R-squared</i>	0.949688	<i>Mean dependent var</i>	102.0033	
<i>Adjusted R-squared</i>	0.943883	<i>S.D. dependent var</i>	27.23976	
<i>S.E. of regression</i>	6.452837	<i>Sum squared resid</i>	1082.617	
<i>Durbin-Watson stat</i>	2.161325			

Таблица 1.21

Оценки коэффициентов уравнений с помощью методов OLS, 2SLS и 3SLS

Переменная	Коэффициент			
	OLS	2SLS	3SLS	True
<i>Первое уравнение</i>				
C(1)	−0.363218	−0.424633	−0.424633	−0.5
C(2)	88.69833	83.95884	83.95884	80.0
C(3)	0.476608	0.585040	0.585040	0.7
<i>Второе уравнение</i>				
C(4)	1.019773	0.901844	0.889189	0.75
C(5)	−11.07961	−1.113488	−0.653169	10.0
C(6)	−1.752307	−1.684157	−1.671481	−1.5
C(6)	1.672787	1.628247	1.624186	1.5

Сравним полученные этими методами оценки коэффициентов (табл. 1.21).

Для иллюстрации приведем результаты применения двухступенчатой процедуры Годфри — Хаттона к рассмотренной системе уравнений:

$$\begin{cases} y_{i1} = \alpha_{11}y_{i2} + \theta_{11}x_{i1} + \theta_{21}x_{i2} + u_{i1} \\ y_{i2} = \alpha_{12}y_{i1} + \theta_{12}x_{i1} + \theta_{22}x_{i3} + \theta_{32}x_{i4} + u_{i2}, \end{cases}$$

по данным табл. 1.7. В качестве инструментальных переменных используются $x_1 = 1, x_2, x_3, x_4$.

Выше было произведено оценивание обоих уравнений двухшаговым методом наименьших квадратов. Полученные при этом 2SLS-остатки обозначим соответственно $\hat{u}_{i1}^{2SLS}, \hat{u}_{i2}^{2SLS}$. Оценим уравнение регрессии \hat{u}_{i1}^{2SLS} на $x_1 = 1, x_2, x_3, x_4$. Полученное значение коэффициента детерминации равно 0.000319, так что $J = nR^2 = 0.00957$. Число степеней свободы равно: $4 - 3 = 1$. Поскольку P -значение равно 0.922, гипотеза пригодности использованных инструментов не отвергается. (Если оценить уравнение регрессии \hat{u}_{i2}^{2SLS} на $x_1 = 1, x_2, x_3, x_4$, то в этом случае $J = 0$, число степеней свободы равно: $4 - 4 = 0$, и J -критерий неприменим.)

Поскольку гипотеза пригодности инструментов не отвергнута, перейдем ко второму этапу и используем на этом этапе критерий Дарбина — Ву — Хаусмана.

Сначала оцениваем уравнение

$$y_{i2} = \pi_{12} + \pi_{22}x_{i2} + \pi_{32}x_{i3} + \pi_{42}x_{i4} + w_{i2}$$

и получаем ряд остатков $\hat{w}_{i2} = y_{i2} - \hat{y}_{i2}$. Затем оцениваем расширенное первое уравнение

$$y_{i1} = \alpha_{11}y_{i2} + \theta_{11}x_{i1} + \theta_{21}x_{i2} + \gamma\hat{w}_{i2} + \eta_{i1}$$

и проверяем гипотезу $H_0: \gamma = 0$. Поскольку здесь γ — скалярная величина, для проверки этой гипотезы можно использовать t -критерий. Результаты оценивания расширенной модели представлены в табл. 1.22.

Таблица 1.22

Объясняемая переменная Y1
Method: Least Squares

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
Y2	−0.424633	0.023668	−17.94140	0.0000
C	83.95884	5.665468	14.81940	0.0000
X2	0.585040	0.067204	8.705405	0.0000
W2	0.616608	0.074993	8.222165	0.0000
R-squared	0.928759	Mean dependent var		100.0167
Durbin-Watson stat	2.089751	Prob. (F-statistic)		0.000000

Оцененный коэффициент $\hat{\gamma} = 0.617$ имеет очень высокую статистическую значимость, что говорит о наличии проблемы эндогенности в первом уравнении.

Применив тот же критерий ко второму уравнению, получим для расширенного уравнения

$$y_{i2} = \alpha_{12}y_{i1} + \theta_{12}x_{i1} + \theta_{22}x_{i3} + \theta_{32}x_{i4} + \gamma\hat{w}_{i1} + \eta_{i2}$$

результаты, приведенные в табл. 1.23.

Таблица 1.23

Объясняемая переменная Y2
Method: Least Squares

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
Y1	−0.901844	0.426087	2.116571	0.0444
C	−1.113488	38.32829	−0.029051	0.9771
X3	−1.684157	0.278578	−6.045545	0.0000
X4	1.628247	0.171996	9.466768	0.0000
W1	0.238481	0.605917	0.393586	0.6972
R-squared	0.950369	Mean dependent var		102.0033
Durbin-Watson stat	2.187028	Prob. (F-statistic)		0.000000

В этом случае оцененный коэффициент при дополнительной переменной статистически незначим, так что эндогенность во втором уравнении не выявляется. Между тем и в первом, и во втором уравнениях *OLS*-оценки всех параметров имеют большее смещение по сравнению с другими методами оценивания, а во втором уравнении оценка постоянной имеет большое смещение при всех методах оценивания. ■

Заметим, что используемые на втором шаге процедур 2SLS и 3SLS очищенные значения \tilde{y}_{i1} , \tilde{y}_{i2} эндогенных переменных y_{i1} , y_{i2} являются линейными комбинациями экзогенных переменных x_{i2} , x_{i3} и x_{i4} , в число которых входят переменные, находящиеся в правых частях соответствующих структурных уравнений. В правой части первого уравнения оказались переменные \tilde{y}_{i1} и x_{i2} , а в правой части второго уравнения — переменные \tilde{y}_{i2} , x_{i3} и x_{i4} . Однако при этом опасной мультиколлинеарности переменных в правых частях каждого из уравнений не возникло. Это происходит благодаря тому что:

- в состав \tilde{y}_{i1} входит не только x_{i2} , но и переменные x_{i3} и x_{i4} , не слишком сильно коррелированные с x_{i2} : $Corr(x_3, x_2) = 0.308$, $Corr(x_4, x_2) = 0.680$;
- в состав \tilde{y}_{i2} входят не только x_{i3} и x_{i4} , но и переменная x_{i2} , не слишком сильно коррелированная с x_{i3} и x_{i4} .

Попробуем смоделировать ситуацию, в которой происходит критическая потеря точности оценивания. Для этого реализуем процесс порождения данных, соответствующий системе

$$\begin{cases} y_{i1} = -0.5y_{i2} + 50 + 1.1x_{i2} + u_{i1} \\ y_{i2} = 0.5y_{i1} - 45 + 1.1x_{i4} + u_{i2}, \end{cases}$$

с теми же значениями переменных x_{i2} , x_{i4} и ошибок u_{i1} , u_{i2} , что и в рассмотренном выше примере. Эта система идентифицируема точно, и ее оценивание методом 3SLS дает результаты, указанные в табл. 1.24.

Таблица 1.24

Оценка уравнений с помощью трехшагового метода наименьших квадратов

Estimation Method: Iterative Three-Stage Least Squares (3SLS)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−0.412458	0.062834	−6.564206	0.0000
C(2)	53.69121	10.66377	5.034918	0.0000
C(3)	0.955053	0.151915	6.286768	0.0000
C(4)	0.545426	0.213593	2.553572	0.0135
C(5)	−56.93625	25.48195	−2.234375	0.0296
C(7)	1.167493	0.064218	18.18025	0.0000
Determinant residual covariance		506.2606		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	0.489024	Mean dependent var		98.82331
Adjusted R-squared	0.451174	S.D. dependent var		6.911911
S.E. of regression	5.120537	Sum squared resid		707.9373
Durbin-Watson stat	1.741253			

Окончание табл. 1.24

Equation: $Y_2 = C(4) \cdot Y_1 + C(5) + C(7) \cdot X_4$; Observations: 30			
<i>R-squared</i>	0.929722	<i>Mean dependent var</i>	125.5639
<i>Adjusted R-squared</i>	0.924517	<i>S.D. dependent var</i>	23.42729
<i>S.E. of regression</i>	6.436470	<i>Sum squared resid</i>	1118.560
<i>Durbin-Watson stat</i>	2.049311		

Все оцененные коэффициенты имеют высокую статистическую значимость. Заменяем во втором уравнении переменную x_{t4} новой переменной x_{t5} , порожаемой соотношением

$$x_{t5} = x_{t2} + 2\eta_t, \quad \eta_t \sim i.i.d. N(0, 1), \quad t = 1, \dots, 30,$$

так что система примет вид:

$$\begin{cases} y_{t1} = -0.5y_{t2} + 50 + 1.1x_{t2} + u_{t1} \\ y_{t2} = 0.5y_{t1} - 45 + 1.1x_{t5} + u_{t2}. \end{cases}$$

С точки зрения критериев идентифицируемости формально ничего не изменилось: первое уравнение сверхидентифицируемо, а второе — идентифицируемо точно. Посмотрим, однако, на результаты оценивания (табл. 1.25).

Таблица 1.25

Оценка уравнений с помощью трехшагового метода наименьших квадратов

Estimation Method: Iterative Three-Stage Least Squares (3SLS)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C(1)	−2.109257	3.560708	−0.592370	0.5561
C(2)	9.837905	90.42874	0.108792	0.9138
C(3)	3.365924	4.983926	0.675356	0.5020
C(4)	0.940376	0.427263	2.200928	0.0320
C(5)	−82.48076	30.85206	−2.673428	0.0099
C(8)	1.023874	0.193137	5.301291	0.0000
Determinant residual covariance		3003.720		
Equation: Y1=C(1)*Y2+C(2)+C(3)*X2; Observations: 30				
R-squared	−7.828599	Mean dependent var		102.6282
Adjusted R-squared	−8.482569	S.D. dependent var		5.264654
S.E. of regression	16.21186	Sum squared resid		7096.260
Durbin-Watson stat	1.700910			

Окончание табл. 1.25

Equation: $Y_2 = C(4) \cdot Y_1 + C(5) + C(8) \cdot X_5$; Observations: 30			
<i>R-squared</i>	0.825653	<i>Mean dependent var</i>	117.9538
<i>Adjusted R-squared</i>	0.812738	<i>S.D. dependent var</i>	14.61378
<i>S.E. of regression</i>	6.323932	<i>Sum squared resid</i>	1079.7870
<i>Durbin-Watson stat</i>	2.085919		

Здесь оценки стандартных ошибок оцененных коэффициентов заметно выросли, особенно в первом уравнении, так что оказались статистически незначимыми все три оцененных коэффициента этого уравнения. Значительно изменились и оценки коэффициентов, особенно в первом уравнении (табл. 1.26).

Таблица 1.26

Уравнения, оцененные 3SLS

Переменная	Коэффициент		Стандартная ошибка	
	Уравнение 1	Уравнение 2	Уравнение 1	Уравнение 2
$C(1)$	-0.412458	-2.109257	0.062834	3.560708
$C(2)$	53.69121	9.837905	10.66377	90.42874
$C(3)$	0.955053	3.365924	0.151915	4.983926
$C(4)$	0.545426	0.940376	0.213593	0.427263
$C(5)$	-56.93625	-82.48076	25.48195	30.85206
$C(7)$	1.167493	—	0.064218	—
$C(8)$	—	1.023874	—	0.193137

Столь драматические изменения произошли по той причине, что переменная x_{15} имеет очень высокую корреляцию с x_{12} в выборке: $\text{Corr}(x_{15}, x_{12}) = 0.969$. И хотя формально эти переменные различны и оба уравнения идентифицируемы, со статистической точки зрения эти переменные «слишком близки». Это порождает проблему опасной мультиколлинеарности и приводит к практической неидентифицируемости коэффициентов структурных уравнений, поскольку теперь:

- в состав \tilde{y}_{11} помимо x_{12} и константы входит только переменная x_{15} , сильно коррелированная с x_{12} ;
- в состав \tilde{y}_{12} помимо x_{15} и константы входит только переменная x_{12} , сильно коррелированная с x_{15} .

Прогнозирование по оцененной системе одновременных уравнений

Если нас интересует только предсказание значений $y_{n+1,1}, \dots, y_{n+1,g}$ эндогенных переменных в новом наблюдении по заданным (планируемым) значениям экзогенных переменных $x_{n+1,1}, \dots, x_{n+1,K}$, это можно осуществить непосредственно с использованием оцененной матрицы $\hat{\Pi}$ приведенной формы:

$$(\hat{y}_{n+1,1}, \dots, \hat{y}_{n+1,g}) = (x_{n+1,1}, \dots, x_{n+1,K}) \hat{\Pi}.$$

Возьмем смоделированные данные из табл. 1.7 и на основании оцененной приведенной формы построим точечные прогнозы значений $(y_{30+t,1}, y_{30+t,2})$, $t = 1, \dots, 30$, для новой последовательности значений экзогенных переменных $(x_{30+t,2}, x_{30+t,3}, x_{30+t,4})$, $t = 1, \dots, 30$, воспроизводящей основные характеристики последовательности $(x_{t,2}, x_{t,3}, x_{t,4})$, $t = 1, \dots, 30$. Оценивание уравнений приведенной формы дает результаты, представленные в табл. 1.27.

Таблица 1.27

Оценка уравнений приведенной формы Method: Least Squares

Объясняемая переменная Y1				
Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	61.35073	7.478051	8.204107	0.0000
X2	0.428163	0.083052	5.155376	0.0000
X3	0.511247	0.060707	8.421559	0.0000
X4	-0.502120	0.036696	-13.68320	0.0000
Объясняемая переменная Y2				
C	54.21529	18.22556	2.974685	0.0063
X2	0.386136	0.202414	1.907652	0.0675
X3	-1.223092	0.147955	-8.266637	0.0000
X4	1.175413	0.089436	13.14250	0.0000
R-squared	0.936459	Mean dependent var		102.0033
Adjusted R-squared	0.929127	S.D. dependent var		27.23976
S.E. of regression	7.251752	Akaike info criterion		6.923929
Sum squared resid	1367.286	Schwarz criterion		7.110755
Log likelihood	-99.85894	F-statistic		127.7280
Durbin-Watson stat	2.066287	Prob. (F-statistic)		0.000000

Соответственно прогнозы для $y_{t,1}$, $y_{t,2}$, $t = 31, \dots, 60$ вычисляются по формулам:

$$\hat{y}_{t,1} = 61.351 + 0.428x_{t2} + 0.511x_{t3} - 0.502x_{t4},$$

$$\hat{y}_{t,2} = 54.215 + 0.386x_{t2} - 1.223x_{t3} + 1.175x_{t4}.$$

Параллельно вычислим для $t = 31, \dots, 60$ «теоретические» значения $y_{t,1_true}$, $y_{t,2_true}$, порожденные системой без случайных ошибок:

$$\begin{cases} y_{t1} = -0.5y_{t2} + 80 + 0.7x_{t2} \\ y_{t2} = 0.75y_{t1} + 10 - 1.5x_{t3} + 1.5x_{t4}. \end{cases}$$

В табл. 1.28 приведены полученные таким образом прогнозные и «теоретические» («истинные») значения (округленные).

Таблица 1.28

Прогнозные и «истинные» значения

t	$Y1_PREDICT$	$Y1_TRUE$	$Y2_PREDICT$	$Y2_TRUE$
1	118.0	119.2	49.4	53.3
2	98.6	98.4	101.2	101.2
3	95.2	94.6	106.7	106.4
4	116.3	117.9	70.3	73.0
5	92.1	91.1	107.2	106.6
6	95.6	94.6	91.5	91.9
7	77.9	75.9	145.6	141.9
8	87.4	85.8	114.1	113.1
9	98.3	98.6	121.6	120.8
10	102.0	103.0	116.0	114.3
11	97.0	96.6	105.2	105.2
12	122.0	124.2	64.4	67.7
13	108.9	109.4	73.2	75.6
14	98.6	99.1	117.3	115.7
15	92.6	92.1	119.4	117.9
16	88.2	87.4	132.9	130.3
17	91.8	91.2	124.1	122.9
18	107.1	108.3	93.1	92.7
19	115.2	116.5	68.0	70.9
20	98.8	98.0	79.7	80.5
21	100.1	100.1	100.6	100.9
22	92.8	91.8	107.2	106.8
23	99.4	99.1	94.1	94.5
24	86.5	85.4	133.1	130.4

Окончание табл. 1.28

t	$Y1_PREDICT$	$Y1_TRUE$	$Y2_PREDICT$	$Y2_TRUE$
25	110.2	111.4	89.0	90.5
26	106.2	106.0	61.9	64.5
27	83.8	82.6	140.6	137.4
28	93.0	92.3	117.7	116.9
29	89.7	89.9	155.3	151.4
30	88.3	87.8	141.7	138.5

Как видно, прогнозные значения оказались весьма близкими к «истинным». Если использовать при сравнении этих значений **среднюю абсолютную процентную ошибку прогноза** ($MAPE$ — *mean absolute percent error*), вычисляемую по формуле

$$MAPE(i) = \frac{1}{30} \sum_{t=1}^{30} 100 \left| \frac{y_{t,i} - predict - y_{t,i} - true}{y_{t,i} - true} \right|,$$

то получим следующие результаты:

$$MAPE(1) = 0.881\%, \quad MAPE(2) = 1.797\%.$$

Подойдем теперь к вопросу прогнозирования с другой стороны: попробуем получить прогнозные значения на основании оцененных структурных уравнений. При этом будем ориентироваться на оценки, полученные ранее (см. табл. 1.20) методом 3SLS (табл. 1.29).

Таблица 1.29

Коэффициенты, полученные 3SLS

Переменная	Коэффициент
$C(1)$	−0.424633
$C(2)$	83.95884
$C(3)$	0.585040
$C(4)$	0.889189
$C(5)$	−0.653169
$C(6)$	−1.671481
$C(7)$	1.624186

Иначе говоря, будем опираться на оцененные структурные уравнения:

$$\begin{cases} y_{t1} = -0.425y_{t2} + 83.959 + 0.585x_{t2} \\ y_{t2} = 0.889y_{t1} - 0.653 - 1.671x_{t3} + 1.624x_{t4}. \end{cases}$$

Получим прогнозные значения ($Y1F$, $Y2F$), приведенные в табл. 1.30.

Таблица 1.30

Прогнозные и «истинные» значения

t	$Y1F$	$Y1_PREDICT$	$Y2F$	$Y2_PREDICT$
1	117.9539	116.3	49.43060	49.4
2	98.62752	92.1	101.2295	101.2
3	95.21508	95.6	106.7986	106.7
4	116.2262	77.9	70.28356	70.3
5	92.14986	87.4	107.2833	107.2
6	95.65463	98.3	91.58455	91.5
7	77.93844	102.0	145.6529	145.6
8	87.49209	97.0	114.3175	114.1
9	98.28987	122.0	121.6912	121.6
10	101.7954	108.9	115.6271	116.0
11	97.07454	98.6	105.3218	105.2
12	121.9508	92.6	64.30229	64.4
13	108.9582	88.2	73.29769	73.2
14	98.49096	91.8	117.0801	117.3
15	92.60201	107.1	119.3591	119.4
16	88.14985	115.2	132.8158	132.9
17	91.87218	98.8	124.2472	124.1
18	106.9471	100.1	92.66480	93.1
19	115.1834	92.8	68.00794	68.0
20	98.80164	99.4	79.70102	79.7
21	100.0887	86.5	100.6958	100.6
22	92.81281	110.2	107.3434	107.2
23	99.37046	106.2	94.12591	94.1
24	86.43539	83.8	133.1007	133.1
25	110.1443	93.0	89.02851	89.0
26	106.2692	89.7	62.01926	61.9
27	83.80259	88.3	140.5781	140.6
28	93.01326	116.3	117.8567	117.7
29	89.61098	92.1	155.1329	155.3
30	88.19077	95.6	141.5809	141.7

Во второй и четвертой графе табл. 1.30 приведены для сравнения прогнозные значения, полученные ранее на основании оцененной приведенной формы.

Для прогнозов, полученных по структурным уравнениям:

$$MAPE(1) = 0.906\%, \quad MAPE(2) = 1.773\%,$$

по сравнению со значениями:

$$MAPE(1) = 0.881\%, \quad MAPE(2) = 1.797\%$$

для прогнозов на основании оцененных приведенных уравнений. В целом прогнозы обоих типов близки по указанной характеристике качества.

Графики на рис. 1.5 и 1.6 иллюстрируют качество прогнозов по сравнению с «теоретическими» значениями.

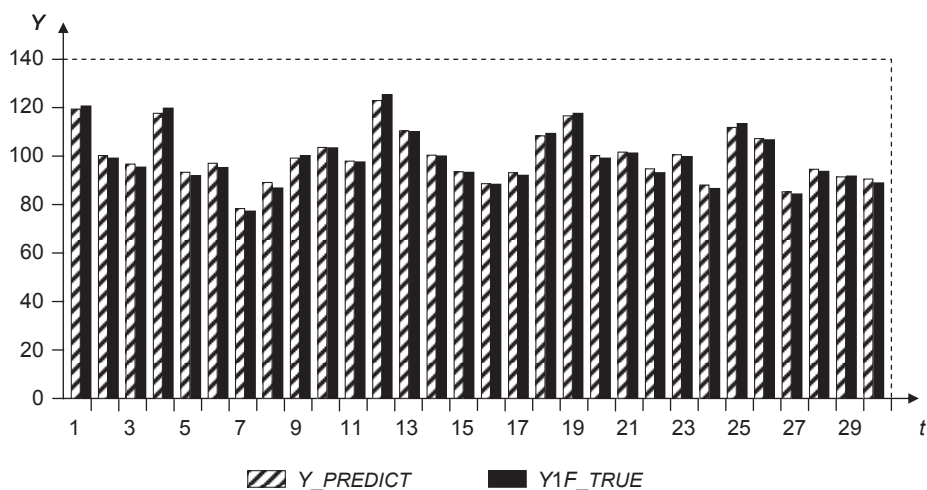


Рис. 1.5

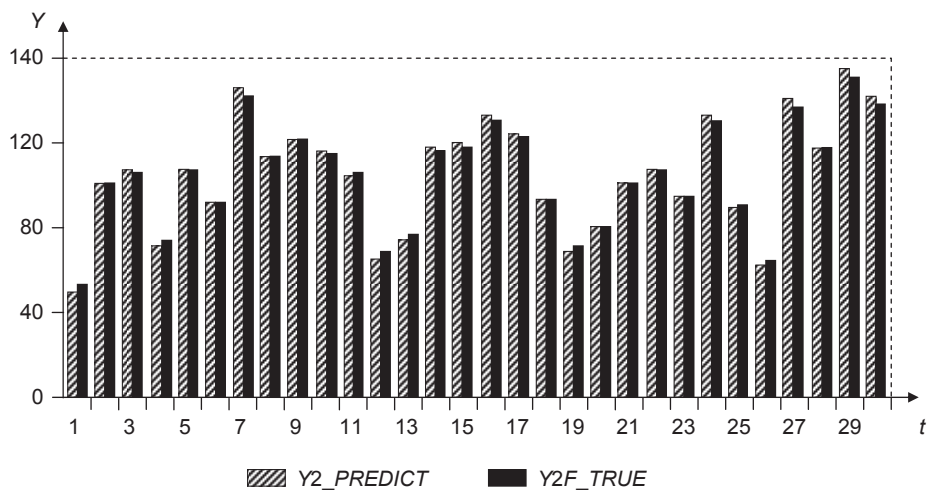


Рис. 1.6

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В чем состоит косвенный метод наименьших квадратов, применяемый для оценивания коэффициентов структурной формы системы одновременных уравнений? В каких случаях можно применять этот метод?
2. В чем состоит двухшаговый метод наименьших квадратов, применяемый для оценивания коэффициентов структурной формы системы одновременных уравнений? В каких случаях можно применять этот метод?
3. Как связан двухшаговый метод наименьших квадратов с методом инструментальных переменных?
4. Как можно проверить адекватность модели при использовании двухшагового метода наименьших квадратов? Какие остатки при этом следует использовать?
5. Можно ли для построения t -статистик и доверительных интервалов использовать вычисленные (по формулам *OLS*) на втором шаге двухшагового метода наименьших квадратов значения стандартных ошибок коэффициентов?
6. Какой критерий предназначен специально для выявления гетероскедастичности в отдельном уравнении системы?
7. Как получить трехшаговую оценку наименьших квадратов и почему возникает необходимость в построении такой оценки? В каких случаях трехшаговая оценка наименьших квадратов не является состоятельной оценкой коэффициентов?
8. Как используется метод максимального правдоподобия для оценивания систем одновременных уравнений?
9. Почему может не сходиться итерационный процесс при практической реализации метода максимального правдоподобия с полной информацией (*FIML*)?
10. Как можно получить *FIML*-оценку в рекурсивной системе с диагональной ковариационной матрицей?
11. Как получить оценку максимального правдоподобия с ограниченной информацией?
12. Опишите связь между различными оценками отдельного структурного уравнения системы.
13. Чем определяется выбор того или иного метода оценивания системы одновременных уравнений?
14. Как следует поступать при наличии автокоррелированности случайных ошибок в оцениваемом структурном уравнении системы?
15. На чем основан критерий Хаусмана для проверки правильности спецификации системы одновременных уравнений?
16. Как проверяется гипотеза экзогенности переменных с использованием критерия Дарбина — Ву — Хаусмана? В каких случаях и каким образом можно произвести проверку пригодности выбранных инструментов?
17. Почему решению об экзогенности переменной не всегда можно доверять?
18. Как производится прогнозирование по оцененной системе одновременных уравнений?

Р а з д е л 2

СТРУКТУРНЫЕ И ПРИВЕДЕННЫЕ ФОРМЫ МОДЕЛЕЙ КОРРЕКЦИИ ОШИБОК

Тема 2.1

СТРУКТУРНЫЕ И ПРИВЕДЕННЫЕ ФОРМЫ ВЕКТОРНЫХ АВТОРЕГРЕССИЙ И МОДЕЛЕЙ КОРРЕКЦИИ ОШИБОК

Кратко напомним необходимые для последующего изложения факты, относящиеся к построению и статистическому анализу моделей коррекции ошибок для коинтегрированных временных рядов, приведенные во второй части учебника. Пусть имеем N временных рядов y_{1t}, \dots, y_{Nt} ¹, каждый из которых является интегрированным порядка 1, так что в общепринятых обозначениях $y_{jt} \sim I(1), j = 1, \dots, N$. Если существует вектор $\beta = (\beta_1, \dots, \beta_N)^T$, отличный от нулевого, для которого

$$\beta_1 y_{1t} + \dots + \beta_N y_{Nt} \sim I(0) \text{ — стационарный ряд,}$$

то ряды коинтегрированы (в узком смысле)²; такой вектор β называется *коинтегрирующим*. Если при этом

$$c = E(\beta_1 y_{1t} + \dots + \beta_N y_{Nt}),$$

то тогда можно говорить о долговременном положении равновесия системы в виде

$$\beta_1 y_{1t} + \dots + \beta_N y_{Nt} = c.$$

В каждый конкретный момент времени t существует некоторое отклонение системы от этого положения равновесия, характеризующееся величиной

$$z_t = \beta_1 y_{1t} + \dots + \beta_N y_{Nt} - c.$$

¹ Здесь в подстрочных индексах номер (момент) наблюдения стоит на втором месте — *после* номера ряда. Это отличается от системы обозначений, использованной в разд. 1, где номер наблюдения стоял на первом месте, а за ним следовал номер уравнения.

² Такое положение называют еще *детерминистской* коинтеграцией.

При этом столбцы $\beta_{(1)}, \dots, \beta_{(r)}$ матрицы β являются линейно независимыми коинтегрирующими векторами, а элементы α_{ij} матрицы α являются коэффициентами при стационарных линейных комбинациях

$$z_{1,t-1} = \beta_{(1)}^T y_{t-1}, \dots, z_{r,t-1} = \beta_{(r)}^T y_{t-1}$$

(представляющих отклонения в момент $(t - 1)$ от r долговременных соотношений между рядами y_{1t}, \dots, y_{Nt}) в правых частях уравнений для $\Delta y_{1t}, \dots, \Delta y_{Nt}$. Будем говорить о такой модели коррекции ошибок как о **модели коррекции ошибок без ограничений** (*UECM — unrestricted error correction model*).

Представление коинтегрированной *VAR* в форме модели коррекции ошибок не единственно, поскольку в качестве набора $\beta_{(1)}, \dots, \beta_{(r)}$ можно взять любой базис коинтеграционного пространства. Соответственно неоднозначность имеется и в отношении матрицы α . В практических задачах на первый план (наряду с определением ранга коинтеграции) выходит идентификация коинтегрирующих векторов, выражающих осмысленные с экономической точки зрения (экономической теории) долговременные связи между рассматриваемыми переменными (например, паритет покупательной способности, спрос на деньги и т.п.). Это, в свою очередь, требует наложения на коинтегрирующие векторы соответствующих **идентифицирующих ограничений** (*identifying restrictions*), позволяющих различать эти векторы, выделяя их из всего множества линейных комбинаций базисных векторов.

Заметим, что в правую часть уравнений стандартной формы *VAR* включаются только запаздывающие значения переменных y_{1t}, \dots, y_{Nt} . Поэтому в правой части ЕСМ, соответствующей такой *VAR*, оказываются только запаздывающие значения приращений (y_{1t}, \dots, y_{Nt}). Между тем в практических исследованиях при рассмотрении оцененной корреляционной матрицы вектора приращений ($\Delta y_{1t}, \dots, \Delta y_{Nt}$) часто наблюдаются достаточно удаленные от нуля значения недиагональных элементов этой матрицы. Последнее указывает на возможную коррелированность приращений, соответствующих одному и тому же моменту времени. Непосредственно учесть такого рода коррелированность можно, если перейти к модели **структурной VAR** (*SVAR — structural vector autoregression*) и соответствующей ей модели **структурной ЕСМ** (*SECM — structural error correction model*). Уравнения структурной *VAR* содержат в правых частях уравнений для отдельных переменных не только запаздывающие, но и текущие значения других переменных. Уравнения структурной ЕСМ содержат в правых частях уравнений для приращений отдельных переменных не только запаздывающие, но и текущие значения приращений других переменных.

Рассмотрим пару рядов y_{1t}, y_{2t} , составляющих векторный случайный процесс $y_t = (y_{1t}, y_{2t})^T$, порождаемый структурной ЕСМ:

$$\begin{aligned}\Delta y_{1t} &= \varphi_{12} \Delta y_{2t} + \gamma_{11}^* \Delta y_{1,t-1} + \gamma_{12}^* \Delta y_{2,t-1} + \\ &\quad + a_{11}(\beta_{11} y_{1,t-1} + \beta_{21} y_{2,t-1}) + a_{12}(\beta_{12} y_{1,t-1} + \beta_{22} y_{2,t-1}) + \zeta_{1t}, \\ \Delta y_{2t} &= \varphi_{21} \Delta y_{1t} + \gamma_{21}^* \Delta y_{1,t-1} + \gamma_{22}^* \Delta y_{2,t-1} + \\ &\quad + a_{21}(\beta_{11} y_{1,t-1} + \beta_{21} y_{2,t-1}) + a_{22}(\beta_{12} y_{1,t-1} + \beta_{22} y_{2,t-1}) + \zeta_{2t}.\end{aligned}$$

Обозначив

$$\begin{aligned}\beta &= \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}, \quad \Gamma_1^* = \begin{pmatrix} \gamma_{11}^* & \gamma_{12}^* \\ \gamma_{21}^* & \gamma_{22}^* \end{pmatrix}, \quad a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \\ \Phi &= \begin{pmatrix} 1 & -\varphi_{12} \\ -\varphi_{21} & 1 \end{pmatrix}, \quad \zeta_t = \begin{pmatrix} \zeta_{1t} \\ \zeta_{2t} \end{pmatrix},\end{aligned}$$

получим для этой структурной ЕСМ компактное выражение:

$$\Phi \Delta y_t = \Gamma_1^* \Delta y_{t-1} + a \beta^T y_{t-1} + \zeta_t.$$

Умножив обе части последнего выражения слева на матрицу Φ^{-1} , найдем **приведенную форму ЕСМ** (*reduced form error correction model*):

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \alpha \beta^T y_{t-1} + \varepsilon_t,$$

где $\Gamma_1 = \Phi^{-1} \Gamma_1^*$, $\alpha = \Phi^{-1} a$, $\varepsilon_t = \Phi^{-1} \zeta_t$. Но

$$\Phi^{-1} = \frac{1}{(1 - \varphi_{12} \varphi_{21})} \begin{pmatrix} 1 & \varphi_{12} \\ \varphi_{21} & 1 \end{pmatrix},$$

так что, обозначив $\delta = 1 - \varphi_{12} \varphi_{21}$, получим:

$$\begin{aligned}\gamma_{11} &= \frac{\gamma_{11}^* + \varphi_{12} \gamma_{21}^*}{\delta}, \quad \gamma_{12} = \frac{\gamma_{12}^* + \varphi_{12} \gamma_{22}^*}{\delta}, \\ \gamma_{21} &= \frac{\gamma_{21}^* + \varphi_{21} \gamma_{11}^*}{\delta}, \quad \gamma_{22} = \frac{\gamma_{22}^* + \varphi_{21} \gamma_{12}^*}{\delta}, \\ \alpha &= \frac{1}{\delta} \begin{pmatrix} 1 & \varphi_{12} \\ \varphi_{21} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \\ \alpha_{11} &= \frac{a_{11} + \varphi_{12} a_{21}}{\delta}, \quad \alpha_{12} = \frac{a_{12} + \varphi_{12} a_{22}}{\delta}, \\ \alpha_{21} &= \frac{a_{21} + \varphi_{21} a_{11}}{\delta}, \quad \alpha_{22} = \frac{a_{22} + \varphi_{21} a_{12}}{\delta}.\end{aligned}$$

Рассматриваемая структурная ЕСМ не имеет ограничений в том смысле, что в ней не приравниваются нулю:

- никакие внедиагональные элементы матрицы Φ ;
- никакие элементы матрицы Γ_1^* ;
- никакие элементы вектора a ;
- никакие элементы вектора β .

Но в конкретных примерах некоторые из перечисленных элементов аннулируются. Например, если ряды $y_{1t}, y_{2t} \sim I(1)$ коинтегрированы, то коинтегрирующий вектор единствен. Если это вектор $(\beta_{11}, \beta_{21})^T$, тогда можно положить $\beta_{12} = \beta_{22} = 0, \alpha_{12} = \alpha_{22} = 0$, что уменьшает количество неизвестных коэффициентов.

В общем случае структурная ЕСМ имеет вид:

$$\Phi \Delta y_t = \Gamma_1^* \Delta y_{t-1} + \dots + \Gamma_p^* \Delta y_{t-p} + a \beta^T y_{t-1} + \zeta_t,$$

где Φ — недиагональная невырожденная квадратная матрица размера $(N \times N)$.

Умножив обе части уравнения на Φ^{-1} , получим приведенную форму ЕСМ:

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_p \Delta y_{t-p} + \alpha \beta^T y_{t-1} + \varepsilon_t,$$

где $\Gamma_j = \Phi^{-1} \Gamma_j^*, \alpha = \Phi^{-1} a, \varepsilon_t = \Phi^{-1} \zeta_t$.

В этих двух формах общим является только $\beta^T y_{t-1}$, т.е. долговременное соотношение, тогда как коэффициенты адаптации (в приведенной форме ЕСМ) являются функциями от коэффициентов структурной ЕСМ: $\alpha = \Phi^{-1} a$. Последнее приводит к тому, что отсутствие некоторой корректирующей составляющей в одном из уравнений SECM отнюдь не означает, что эта составляющая будет отсутствовать и в соответствующем уравнении приведенной ЕСМ. Соответственно коррекция ошибок в одном уравнении SECM может распространяться и на все остальные уравнения приведенной ЕСМ.

При рассмотрении вопроса об идентифицируемости параметров структурной ЕСМ естественно выделить отдельно идентификацию коинтегрирующих векторов и идентификацию коэффициентов, связанных с динамической адаптацией, т.е. элементов матриц $\Phi, \Gamma_1^*, \dots, \Gamma_p^*, a$. Поскольку коинтегрирующие соотношения в структурной и приведенной ЕСМ одни и те же, можно использовать результаты, касающиеся идентифицируемости коинтегрирующих векторов в приведенной ЕСМ.

Вопрос об идентифицируемости коинтегрирующих векторов естественно возникает при наличии двух и более коинтегрирующих векторов и связан с возможностью различения таких векторов. В процедуре Йохансена, реализованной в пакете *EViews*, такое различение производится исходя из того, что если ранг коинтеграции равен $1 < r < N$, то существует $(N - r)$ некоинтегрированных между собой (в совокупности) переменных — «общих трендов» («*common*

trends»), таких, что добавление к этой совокупности любой из оставшихся r переменных приводит к коинтегрированности пополненного множества из $(N - r + 1)$ переменных. Это означает, что в качестве линейно независимых коинтегрирующих векторов можно взять любой набор из r векторов вида

$$\begin{pmatrix} \beta_{11} \\ 0 \\ \vdots \\ 0 \\ \beta_{1,r+1} \\ \vdots \\ \beta_{1N} \end{pmatrix}, \begin{pmatrix} 0 \\ \beta_{22} \\ \vdots \\ 0 \\ \beta_{2,r+1} \\ \vdots \\ \beta_{2N} \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \beta_{rr} \\ \beta_{r,r+1} \\ \vdots \\ \beta_{rN} \end{pmatrix}.$$

Конечно, при этом предполагается, что переменные занумерованы так, что «общие тренды» соответствуют переменным с номерами $r + 1, \dots, N$. Выделение из этого множества возможных наборов единственного набора производится в пакете *EViews* нормализацией каждого из этих векторов, так что j -й вектор нормализуется делением всех его элементов на β_{jj} , вследствие чего получаем набор из r векторов:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \beta_{1,r+1}^* \\ \vdots \\ \beta_{1N}^* \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \beta_{2,r+1}^* \\ \vdots \\ \beta_{2N}^* \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \beta_{r,r+1}^* \\ \vdots \\ \beta_{r,N}^* \end{pmatrix}.$$

Такой набор образует базис r -мерного линейного пространства коинтегрирующих векторов при ранге коинтеграции r .

Проблема, однако, в том, что с точки зрения экономической теории нас могут интересовать коинтегрирующие векторы другого вида (являющиеся, конечно, линейными комбинациями векторов, принадлежащих базису). При этом на соответствующие коинтегрирующие векторы накладываются определенные ограничения, вытекающие из экономической теории: невхождение в коинтегрирующую линейную комбинацию тех или иных переменных, равенство некоторых компонент коинтегрирующего вектора или наличие у них противоположных знаков при одинаковой абсолютной величине и т.п. В такой ситуации возникает вопрос о необходимости и достаточности множества накладываемых ограничений для идентификации (т.е. для различения) коинтегрирующих векторов.

Обычно рассматривают лишь линейные ограничения, в том числе исключаяющие появление отдельных переменных в коинтегрирующей линейной комбинации. При этом ограничения могут быть представлены как в явной, так и в неявной форме. Если векторы уже нормализованы, то **необходимым условием идентифицируемости r коинтегрирующих векторов** является наложение на каждый из r векторов не менее $(r - 1)$ линейных ограничений. Об этом условии говорят как о **порядковом условии идентифицируемости**.

Порядковое условие, вообще говоря, не является достаточным для идентифицируемости, поскольку при его выполнении полученные r векторов могут все же оказаться линейно зависимыми, так что, скажем, вектор β_1 нельзя отличить от некоторой линейной комбинации векторов β_2, \dots, β_r . Поэтому, в принципе, следует производить еще и проверку линейной независимости полученных r векторов. Для этого можно воспользоваться достаточными условиями идентифицируемости, формулируемыми в терминах матриц, участвующих в формировании явной и неявной форм линейных ограничений.

Если на i -й коинтегрирующий вектор накладывается r_i линейных ограничений, то их можно записать в двух формах: явной и неявной. Под **неявной формой** (*indirect restrictions*) понимается представление этих ограничений в виде:

$$R_i \beta_i = 0,$$

где R_i — матрица размера $(r_i \times N)$ ранга r_i .

Ту же самую совокупность ограничений можно представить в **явной форме** (*direct restrictions*) в виде:

$$\beta_i = H_i \mathcal{G}_i,$$

где H_i — матрица размера $N \times (N - r_i)$ ранга $N - r_i$;

\mathcal{G}_i — вектор размера $(N - r_i) \times 1$.

При этом выполняется соотношение

$$R_i H_i = 0,$$

т.е. строки матрицы R_i ортогональны столбцам матрицы H_i .

Поясним это на примере модели IS/LM, связывающей следующие макроэкономические параметры:

$m_t = \ln M_t$, где M_t — номинальная денежная масса;

$\ln c_t = \ln GDP_t$, где GDP_t — реальный валовой внутренний продукт;

$p_t = \ln P_t$, где P_t — дефлятор GDP ;

r_t^s — краткосрочная процентная ставка;

r_t^b — долгосрочная процентная ставка.

Пусть все эти ряды — интегрированные порядка 1, ранг коинтеграции этих временных рядов равен $r = 3$ и в коинтеграционное соотношение приходится включать еще и временной тренд t . Тогда речь идет об идентификации трех коинтегрирующих векторов:

$$\beta_1 = \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \\ \beta_{41} \\ \beta_{51} \\ \beta_{61} \end{pmatrix}, \quad \beta_2 = \begin{pmatrix} \beta_{12} \\ \beta_{22} \\ \beta_{32} \\ \beta_{42} \\ \beta_{52} \\ \beta_{62} \end{pmatrix}, \quad \beta_3 = \begin{pmatrix} \beta_{13} \\ \beta_{23} \\ \beta_{33} \\ \beta_{43} \\ \beta_{53} \\ \beta_{63} \end{pmatrix},$$

обеспечивающих стационарность линейных комбинаций:

$$\begin{aligned} z_{1t} &= \beta_{11}m_t + \beta_{21}inc_t + \beta_{31}p_t + \beta_{41}r_t^s + \beta_{51}r_t^b + \beta_{61}t, \\ z_{2t} &= \beta_{12}m_t + \beta_{22}inc_t + \beta_{32}p_t + \beta_{42}r_t^s + \beta_{52}r_t^b + \beta_{62}t, \\ z_{3t} &= \beta_{13}m_t + \beta_{23}inc_t + \beta_{33}p_t + \beta_{43}r_t^s + \beta_{53}r_t^b + \beta_{63}t. \end{aligned}$$

Ограничения на коэффициенты этих векторов могут быть получены, например, из следующих соображений.

Если спрос на реальные деньги рассматривается как функция от реального дохода, краткосрочной ставки и тренда, т.е. $m_t - p_t = f_1(inc_t, r_t^s, t)$, то это подразумевает наличие долгосрочной связи между переменными $m_t - p_t$, inc_t , r_t^s , t без включения в нее переменной r_t^b , так что стационарной является линейная комбинация $z_{1t} = \beta_{11}m_t + \beta_{21}inc_t - \beta_{11}p_t + \beta_{41}r_t^s + \beta_{61}t$. Ограничения на вектор β_1 принимают вид: $\beta_{31} = -\beta_{11}$, $\beta_{51} = 0$. Эти ограничения можно записать в виде $R_1\beta_1 = 0$ (неявная форма), где

$$R_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

или в виде $\beta_1 = H_1\mathcal{G}_1$, где

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{G}_1 = \begin{pmatrix} \mathcal{G}_{11} \\ \mathcal{G}_{21} \\ \mathcal{G}_{31} \\ \mathcal{G}_{41} \end{pmatrix}, \quad \text{так что} \quad \beta_1 = \begin{pmatrix} \mathcal{G}_{11} \\ \mathcal{G}_{21} \\ -\mathcal{G}_{11} \\ \mathcal{G}_{31} \\ 0 \\ \mathcal{G}_{41} \end{pmatrix}.$$

Нетрудно проверить, что $R_1H_1 = 0$.

Если дифференциал процентных ставок $(r_t^s - r_t^b)$ определяется через остальные переменные без включения тренда, т.е. $r_t^s - r_t^b = f_1(m_t, inc_t, p_t)$, то это подразумевает наличие долговременной связи между переменными $r_t^s - r_t^b$, m_t , inc_t , p_t без включения в нее переменной t , так что стационарной является линейная комбинация $z_{2t} = \beta_{12}m_t + \beta_{22}inc_t + \beta_{32}p_t + \beta_{42}r_t^s - \beta_{42}r_t^b$. Ограничения на вектор β_2 принимают вид: $\beta_{52} = -\beta_{42}$, $\beta_{62} = 0$. Эти ограничения можно записать в виде $R_2\beta_2 = 0$ (неявная форма), где

$$R_2 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

или в виде $\beta_2 = H_2 \mathcal{G}_2$, где

$$H_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{G}_2 = \begin{pmatrix} \mathcal{G}_{12} \\ \mathcal{G}_{22} \\ \mathcal{G}_{32} \\ \mathcal{G}_{42} \end{pmatrix}, \quad \text{так что } \beta_2 = \begin{pmatrix} \mathcal{G}_{12} \\ \mathcal{G}_{22} \\ \mathcal{G}_{32} \\ \mathcal{G}_{42} \\ -\mathcal{G}_{42} \\ 0 \end{pmatrix}.$$

Наконец, если долгосрочная процентная ставка r_t^b определяется как функция только от m_t , p_t и t , то это подразумевает наличие долговременной связи между переменными r_t^b , m_t , p_t и t без включения в нее переменных inc_t и r_t^s , так что стационарной является линейная комбинация $z_{3t} = \beta_{13}m_t + \beta_{33}p_t + \beta_{53}r_t^b + \beta_{63}t$. Ограничения на вектор β_3 принимают вид: $\beta_{23} = 0$, $\beta_{43} = 0$. Эти ограничения можно записать в виде $R_3\beta_3 = 0$ (неявная форма), где

$$R_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

или в виде $\beta_3 = H_3 \mathcal{G}_3$, где

$$H_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{G}_3 = \begin{pmatrix} \mathcal{G}_{13} \\ \mathcal{G}_{23} \\ \mathcal{G}_{33} \\ \mathcal{G}_{43} \end{pmatrix}, \quad \text{так что } \beta_3 = \begin{pmatrix} \mathcal{G}_{13} \\ 0 \\ \mathcal{G}_{23} \\ 0 \\ \mathcal{G}_{33} \\ \mathcal{G}_{43} \end{pmatrix}.$$

Необходимое и достаточное условие идентифицируемости r , $1 < r < N$, коинтегрирующих векторов имеет в общем случае довольно сложный вид. Однако при $r = 2, 3$ его достаточно просто проверить (см., например, (Patterson, 2000, p. 635—641)).

При $r = 2$ коинтегрирующие векторы β_1, β_2 идентифицируемы тогда и только тогда, когда выполняются соотношения:

$$\text{rank}(R_1 H_2) \geq 1, \quad \text{rank}(R_2 H_1) \geq 1.$$

При $r = 3$ коинтегрирующие векторы $\beta_1, \beta_2, \beta_3$ идентифицируемы тогда и только тогда, когда выполняются соотношения:

$$\text{rank}(R_i H_j) \geq 1, \quad i \neq j, \quad i, j = 1, 2, 3 \text{ (6 соотношений),}$$

$$\text{rank}(R_1 [H_2, H_3]) \geq 2, \quad \text{rank}(R_2 [H_1, H_3]) \geq 2, \quad \text{rank}(R_3 [H_1, H_2]) \geq 2.$$

Проверим выполнение этого условия в только что рассмотренном примере, где $r = 3$. Имеем:

$$\begin{aligned} R_1 H_2 &= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, & R_1 H_3 &= \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\ R_2 H_1 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & R_2 H_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ R_3 H_1 &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, & R_3 H_2 &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Все 6 матриц имеют ранг $2 > 1$, так что первая группа условий выполняется. Далее:

$$\begin{aligned} R_1 [H_2, H_3] &= \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} R_2 [H_1, H_3] &= \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

$$R_3[H_1, H_2] = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ранги всех трех матриц равны 2, так что и эта группа условий выполнена. Таким образом, коинтегрирующие векторы β_1 , β_2 , β_3 идентифицируемы.

Если r коинтегрирующих векторов идентифицируемы, то на каждый из них накладывается не менее $(r - 1)$ линейных ограничений. В случае когда на каждый из этих векторов накладывается ровно $(r - 1)$ ограничений, имеем дело с точной идентифицируемостью. Если же хотя бы для одного из векторов количество ограничений превышает $(r - 1)$, то речь идет о сверхидентифицируемости. В последнем случае имеются «лишние» ограничения и возможность проверки гипотезы о том, что заявленные дополнительные ограничения на векторы $\beta_1, \beta_2, \dots, \beta_r$ действительно выполняются.

В рассмотренном примере можно, следуя работе (Johansen, Juselius, 1994), наложить следующие — более строгие — ограничения на векторы $\beta_1, \beta_2, \beta_3$.

Вектор β_1 нормализуется на inc_t , коэффициенты при m_t и p_t равны по абсолютной величине и противоположны по знаку, а коэффициенты при обеих процентных ставках равны нулю. Соответствующая равновесная связь интерпретируется как прокси для совокупного дохода относительно линейного тренда. Это дает матрицу вида

$$R_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

У вектора β_2 равны нулю все коэффициенты, кроме коэффициентов при r_t^s и r_t^b , которые равны по абсолютной величине и противоположны по знаку; вектор нормализуется на один из них. Интерпретация: стационарность дифференциала процентных ставок. Матрица R_2 принимает вид:

$$R_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Наконец, у вектора β_3 равны нулю коэффициенты при m_t , inc_t и r_t^s ; вектор нормализуется на r_t^b . Интерпретация: долгосрочная процентная ставка определяется как функция от цены и временного тренда. Матрица R_3 имеет вид:

$$R_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

В итоге получим следующую картину (табл. 2.1).

Таблица 2.1

Ограничения на коинтегрирующий вектор

Вектор	Количество ограничений	Количество «лишних» ограничений
β_1	3	1
β_2	5	3
β_3	3	1
ВСЕГО	11	5

Проверка гипотезы H_0 о выполнении совокупности всех сверхидентифицирующих ограничений производится с использованием асимптотического критерия хи-квадрат, указанного в работе (Johansen, Juselius, 1994). При справедливости гипотезы H_0 статистика этого критерия имеет асимптотическое распределение хи-квадрат с q степенями свободы, где q — суммарное количество дополнительных ограничений. Гипотеза H_0 отвергается при слишком больших значениях этой статистики. Однако неотвержение гипотезы H_0 отнюдь не означает, что именно указанную в гипотезе совокупность сверхидентифицирующих ограничений следует использовать, поскольку вообще можно сформировать не один, а несколько различных наборов сверхидентифицирующих ограничений, которые также могут оказаться неотвергнутыми. Приемлемость некоторого конкретного множества ограничений должна подкрепляться также другими соображениями. Среди них можно выделить следующие:

- являются ли осмысленными с точки зрения экономической теории числовые значения оценок коэффициентов, получаемых при выбранных ограничениях?
- являются ли осмысленными с точки зрения экономической теории **коэффициенты адаптации** (*adjustment coefficients*) α_{ij} ?

В нашем примере наблюдаемое значение статистики этого критерия равно 3.5; асимптотическое критическое значение, соответствующее уровню зна-

чимости 0.05, равно $\chi^2_{0.95}(5) = 11.07$. Таким образом нулевая гипотеза о выполнении всех 5 сверхидентифицирующих ограничений не отвергается.

После уточнения идентифицирующих ограничений, которые накладываются на коинтегрирующие векторы, производится оценивание параметров приведенной формы ЕСМ

$$\Delta y_t = \alpha \beta^T y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_p \Delta y_{t-p} + \varepsilon_t,$$

в результате чего находятся, в частности, и оценки коинтегрирующих векторов $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_r$. Эти оценки подставляются в уравнение структурной ЕСМ

$$\Phi \Delta y_t = \Gamma_1^* \Delta y_{t-1} + \dots + \Gamma_p^* \Delta y_{t-p} + a \beta^T y_{t-1} + \zeta_t$$

вместо неизвестных «истинных» коинтегрирующих векторов $\beta_1, \beta_2, \dots, \beta_r$. Возможность получения состоятельных оценок для параметров $\Phi, \Gamma_1^*, \dots, \Gamma_p^*, a$ структурной формы связана с идентифицируемостью этой системы. Ее можно записать также в виде:

$$(\Phi, -\Gamma_1^*, \dots, -\Gamma_p^*, -a) \begin{pmatrix} \Delta y_t \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \\ \beta^T y_{t-1} \end{pmatrix} = \zeta_t,$$

или

$$A^T Z_t = \zeta_t,$$

где $A^T = (\Phi, -\Gamma_1^*, \dots, -\Gamma_p^*, -a)$ — матрица размера $(N \times (N + pN + rN))$;

$$Z_t = \begin{pmatrix} \Delta y_t \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \\ \beta^T y_{t-1} \end{pmatrix} \text{ — вектор стационарных переменных размера } ((N + pN + rN) \times 1).$$

Каждая строка матрицы A^T , т.е. каждый столбец матрицы A , относится к отдельному уравнению. Соответственно для идентифицируемости i -го уравнения необходимо наложить на i -й столбец A_i матрицы A не менее $(N - 1)$ ограничений в виде

$$R_i A_i = 0 \text{ (неявная форма)}$$

или

$$A_i = H_i \vartheta_i \text{ (явная форма), } R_i H_i = 0.$$

Как и в разд. 1, гарантией идентифицируемости i -го структурного уравнения служит выполнение рангового условия идентифицируемости, само оценивание должно проводиться системными методами (например, FIML). На основании оцененной структурной ЕСМ (SECM) строится приведенная форма — **ЕСМ с ограничениями**, соответствующая этой SECM.

Проиллюстрируем процесс построения ЕСМ с ограничениями примером для модели IS/LM, который начали рассматривать ранее. В работе (Johansen, Juselius, 1994) такое построение проводилось по статистическим данным для Австралии (квартальные данные, период с III квартала 1975 г. по I квартал 1991 г.).

При оценивании ЕСМ с учетом указанных выше 11 линейных ограничений на коинтегрирующие векторы и того, что по результатам предварительного анализа модель VAR в уровнях имеет порядок 2, получим оценки, приведенные в табл. 2.2 (для коинтегрирующих векторов β_i) и в табл. 2.3 (для векторов коэффициентов адаптации α_i).

Таблица 2.2

Оценки для коинтегрирующих векторов

Переменная	β_1	β_2	β_3
m_t	-0.193	0	0
inc_t	1	0	0
p_t	0.193	0	-0.488
r_t^s	0	1	0
r_t^b	0	-1	1
t	-0.005	0	0.009

Таблица 2.3

Оценки для векторов коэффициентов адаптации

Уравнение для	α_1	α_2	α_3
Δm_t	0.030	0.159	-0.569
Δinc_t	-0.458	-0.001	0.405
Δp_t	0.325	-0.039	0.054
Δr_t^s	0.337	-0.008	-0.168
Δr_t^b	0.109	0.023	-0.213

Для того чтобы коинтегрирующие линейные комбинации флуктуировали вокруг нулевого уровня, к ним были добавлены константа и дамми-переменная $D84_t$, равная 1 в период с I квартала 1984 г. по I квартал 1991 г. и равная 0 на остальной части периода наблюдений, отражающая изменение правил регулирования в банковском секторе. В результате в правых частях урав-

нений ЕСМ вместо переменных $z_{1,t-1}$, $z_{2,t-1}$, $z_{3,t-1}$ используются переменные $есм1_{t-1}$, $есм2_{t-1}$, $есм3_{t-1}$, где

$$есм1_t = inc_t - 0.193(m_t - p_t) - 0.005t - 0.027D84_t - 8.43,$$

$$есм2_t = (r_t^s - r_t^b) + 0.00967D84_t + 0.03,$$

$$есм3_t = r_t^b - 0.488(p_t - 0.019t) - 0.008D84_t - 0.52.$$

Как говорилось ранее, необходимость в построении структурной ЕСМ возникает из-за наличия коррелированности между переменными, стоящими в левых частях ЕСМ. В рассматриваемом примере в левых частях ЕСМ находятся переменные Δm_t , Δinc_t , Δp_t , Δr_t^s и Δr_t^b . Следовательно, надо выяснить, имеется ли между ними заметная корреляция. Вычисленные выборочные коэффициенты корреляции между этими переменными приведены в табл. 2.4.

Таблица 2.4

Коэффициенты корреляции между переменными

	Δm_t	Δinc_t	Δp_t	Δr_t^s	Δr_t^b
Δm_t	1	0.29	0.20	-0.10	-0.10
Δinc_t	0.29	1	0.35	-0.18	0
Δp_t	0.20	0.35	1	-0.12	-0.10
Δr_t^s	-0.10	-0.18	-0.12	1	0.65
Δr_t^b	-0.10	0	-0.10	0.65	1

Ориентируясь на табл. 2.4, Йохансен и Юселиус делают вывод о наличии коррелированности переменных, что требует системного оценивания. Если исходить из того, что модель VAR в уровнях имеет порядок 2, то в правую часть ЕСМ могут входить значения приращений переменных, запаздывающие не более чем на 1 шаг. Соответственно в правых частях уравнений структурной ЕСМ потенциально могут присутствовать следующие переменные:

$$\Delta m_t, \Delta inc_t, \Delta p_t, \Delta r_t^s, \Delta r_t^b, \Delta m_{t-1}, \Delta inc_{t-1}, \Delta p_{t-1}, \Delta r_{t-1}^s, \Delta r_{t-1}^b, \\ есм1_{t-1}, есм2_{t-1}, есм3_{t-1},$$

так что в каждом из 5 уравнений (для Δm_t , Δinc_t , Δp_t , Δr_t^s и Δr_t^b) потенциально может быть 12 коэффициентов, подлежащих оцениванию. Однако при таком количестве неизвестных (неспецифицированных) коэффициентов уравнения не могут быть идентифицируемыми. Для их идентифицируемости необходимо наложить на коэффициенты каждого уравнения как минимум 4 ($5 - 1 = 4$) ограничения. И здесь, в отличие от выбора идентифицирующих ограничений на коинтегрирующие векторы, ориентирующегося главным образом на экономические представления, приходится опираться по большей части на ста-

тистическую информацию, содержащуюся в самих данных, т.е. на эмпирическую картину адаптации, а не на строгое априорное обусловливание.

Первоначально Йохансен и Юселиус берут ровно по 4 ограничения на каждое уравнение (табл. 2.5), что обеспечивает точную идентифицируемость системы.

Таблица 2.5

Ограничения, обеспечивающие идентифицируемость системы

Уравнение для	Переменные, коэффициенты при которых зануляются			
Δm_t	Δr_t^s	Δinc_t	Δr_t^b	Δinc_{t-1}
Δinc_t	Δr_t^s	Δp_t	Δr_t^b	Δr_{t-1}^b
Δp_t	Δr_t^b	Δp_{t-1}	Δr_{t-1}^s	Δr_{t-1}^b
Δr_t^s	Δr_{t-1}^s	Δinc_{t-1}	Δp_{t-1}	Δr_{t-1}^b
Δr_t^b	Δinc_t	Δm_t	Δp_t	Δr_{t-1}^b

Однако в оцененной с такими ограничениями модели оказались статистически незначимыми (по t -статистикам) почти все оцененные коэффициенты. В связи с этим были проведены эксперименты с различными наборами ограничений. В итоге Йохансен и Юселиус пришли к разбиению системы 5 уравнений на два блока, один из которых содержит уравнения для Δm_t , Δinc_t , Δr_t^s , а другой — уравнения для Δp_t и Δr_t^b . При этом первый блок носит системный характер, а второй — характер приведенной формы, т.е. в правых частях уравнений первого блока имеются эндогенные переменные, порождаемые в рамках системы трех уравнений, а в рамках второго блока — нет. Специфицированные коэффициенты уравнений приведены в табл. 2.6 и 2.7.

Таблица 2.6

Специфицированные коэффициенты

	Δm_t	Δinc_t	Δr_t^s	Δp_t	Δr_t^b	Δm_{t-1}	Δinc_{t-1}	Δr_{t-1}^s	Δp_{t-1}	Δr_{t-1}^b
Δm_t	-1	0	0		0		0	0		0
Δinc_t		-1	0		0	0	0		0	0
Δr_t^s	0	0	-1	0			0			
Δp_t	0	0	0	-1	0		0	0		0
Δr_t^b	0	0	0	0	-1		0	0	0	0

В системе из трех первых уравнений помимо нормализующих накладывается еще 19 ($7 + 7 + 5 = 19$) ограничений (зануляются 19 коэффициентов). Однако необходимым минимумом для каждого из трех уравнений является наличие 2 ($3 - 1 = 2$) ограничений. Таким образом, «избыточными» здесь

Таблица 2.7

Специфицированные коэффициенты (продолжение)

	$есм1_{t-1}$	$есм2_{t-1}$	$есм3_{t-1}$
Δm_t	0		
Δinc_t		0	
Δr_t^s			0
Δp_t			
Δr_t^b		0	

являются: 5 ограничений в первом уравнении, 5 ограничений во втором и 3 ограничения в третьем, всего 13 ограничений. Результаты оценивания SECM с такими ограничениями приведены в табл. 2.8 и 2.9 (в скобках указаны значения t -статистик для оцененных коэффициентов).

Таблица 2.8

Результаты оценивания SECM с ограничениями

	Δm_t	Δinc_t	Δr_t^s	Δp_t	Δr_t^b	Δm_{t-1}	Δr_{t-1}^s	Δp_{t-1}	Δr_{t-1}^b
Δm_t	-1			0.35 (2.6)		0.31 (2.9)		0.41 (3.4)	
Δinc_t	0.25 (1.1)	-1		0.31 (2.0)			0.17 (1.3)		
Δr_t^s			-1		1.10 (6.5)	0.21 (3.0)	0.34 (3.2)	-0.24 (2.8)	-0.45 (2.1)
Δp_t				-1		-0.08 (0.8)		-0.13 (1.0)	
Δr_t^b					-1	0.08 (1.7)			

Таблица 2.9

Результаты оценивания SECM с ограничениями (продолжение)

	$есм1_{t-1}$	$есм2_{t-1}$	$есм3_{t-1}$
Δm_t		0.20 (2.3)	-0.55 (3.6)
Δinc_t	-0.44 (4.1)		0.28 (1.3)
Δr_t^s	0.19 (2.8)	-0.28 (4.7)	
Δp_t	0.20 (2.2)	-0.12 (1.4)	0.48 (3.4)
Δr_t^b	0.12 (3.0)		-0.09 (1.3)

Иначе говоря, оцененная SECM имеет вид:

$$\begin{aligned}\Delta m_t &= 0.35\Delta p_t + 0.31\Delta m_{t-1} + 0.41\Delta p_{t-1} + 0.20есм2_{t-1} - 0.55есм3_{t-1}, \\ \Delta inc_t &= 0.25m_t + 0.31\Delta p_t + 0.17\Delta r_{t-1}^s - 0.44есм1_{t-1} + 0.28есм3_{t-1}, \\ \Delta r_t^s &= 1.10\Delta r_t^b + 0.21\Delta m_{t-1} + 0.34\Delta r_{t-1}^s - 0.24\Delta p_{t-1} - 0.45r_{t-1}^b + \\ &\quad + 0.19есм1_{t-1} - 0.28есм2_{t-1}, \\ \Delta p_t &= -0.08\Delta m_{t-1} - 0.13\Delta p_{t-1} + 0.20есм1_{t-1} - 0.12есм2_{t-1} + 0.48есм3_{t-1}, \\ \Delta r_{t-1}^b &= 0.08\Delta m_{t-1} + 0.12есм1_{t-1} - 0.09есм3_{t-1}.\end{aligned}$$

При этом оцененные коэффициенты по большей части статистически значимы.

Поскольку система оценивалась с наложением количества ограничений больше необходимого для идентифицируемости уравнений, имеется возможность проверки гипотезы о действительном выполнении «лишних» ограничений. Статистика соответствующего критерия принимает значение 4.82, что намного меньше критического значения $\chi_{0.95}^2(13) = 22.36$, так что указанная гипотеза не отвергается.

Заметим, что оцененной SECM соответствуют следующие матрицы Φ и a :

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & -0.35 & 0 \\ -0.25 & 1 & 0 & -0.31 & 0 \\ 0 & 0 & 1 & 0 & -1.1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad a = \begin{pmatrix} 0 & 0.2 & -0.55 \\ -0.44 & 0 & 0.28 \\ 0.19 & -0.28 & 0 \\ 0.20 & -0.12 & 0.48 \\ 0.12 & 0 & -0.09 \end{pmatrix}.$$

При переходе от структурной формы к приведенной матрица коэффициентов адаптации находится по формуле $\alpha = \Phi^{-1}a$. В результате получим коэффициенты адаптации, указанные в табл. 2.10.

Таблица 2.10

Коэффициенты адаптации при переходе от структурной формы к приведенной

Уравнение для	Коэффициенты адаптации					
	в приведенной ЕСМ с ограничениями			в ЕСМ (для сравнения)		
Δm_t	0.070	0.158	-0.382	0.030	0.159	-0.569
Δinc_t	-0.360	0.002	0.333	-0.458	-0.001	0.405
Δr_t^s	0.322	-0.280	-0.100	0.325	-0.039	0.054
Δp_t	0.200	-0.120	0.480	0.337	-0.008	-0.168
Δr_t^b	0.120	0.000	-0.090	0.109	0.023	-0.213

Хотя в первых трех уравнениях SECM при некоторых из $est1_t$, $est2_t$, $est3_t$ были нулевые коэффициенты, в приведенной форме соответствующие им коэффициенты α_{ij} отличны от нуля и показывают, как влияние отклонений от равновесия распространяется на все переменные системы.

В работе (Johansen, Juselius, 1994) проведен детальный экономический анализ результирующих уравнений, исходя из которого сделан вывод о том, что полученная аппроксимация изучаемой экономической структуры, согласованная с имеющимися наблюдениями, вполне удовлетворительна.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Чем отличаются структурные модели VAR и ECM от моделей VAR и ECM, которые были рассмотрены во второй части учебника?
2. Как получается приведенная форма ECM, соответствующая структурной ECM?
3. Означает ли отсутствие некоторой корректирующей составляющей в одном из уравнений SECM то, что эта составляющая будет отсутствовать и в соответствующем уравнении приведенной ECM?
4. Когда возникает вопрос об идентифицируемости коинтегрирующих векторов? В чем состоит порядковое условие идентифицируемости коинтегрирующих векторов? Гарантирует ли оно идентифицируемость коинтегрирующих векторов?
5. Какую полезную гипотезу можно проверить в ситуации, когда некоторые ограничения на коинтегрирующие векторы, имеющие экономическую интерпретацию, оказываются «лишними» с точки зрения возможности идентификации коинтегрирующих векторов?
6. Как производятся построение и оценивание структурной ECM? Что является гарантией идентифицируемости i -го структурного уравнения?
7. Чем отличается приведенная ECM с ограничениями от «обычной» ECM, которая была рассмотрена во второй части учебника?

Р а з д е л 3

ПАНЕЛЬНЫЕ ДАННЫЕ

Тема 3.1

ПАНЕЛЬНЫЕ ДАННЫЕ: МОДЕЛЬ ПУЛА, МОДЕЛЬ КОВАРИАЦИОННОГО АНАЛИЗА, МОДЕЛЬ КАЖУЩИХСЯ НЕСВЯЗАННЫМИ РЕГРЕССИЙ

Панельные данные — модель пула

Пусть имеем данные $\{y_{it}, x_{it}; i = 1, \dots, N, t = 1, \dots, T\}$ о значениях переменных y и x для N субъектов (индивидов, фирм, стран, регионов и т.п.) в T последовательных моментов (периодов) времени (в этом случае говорят, что имеем дело с **панельными данными** (*panel data*)) и хотим оценить модель линейной связи между переменными y и x , считая y объясняемой, а x — объясняющей переменной. В общем случае x является вектором конечной размерности p , и наиболее общей формой линейной модели наблюдений для такой ситуации являлась бы спецификация

$$y_{it} = x_{it}^T \theta_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где θ_{it} измеряет частное влияние x_{it} в момент (период) t для субъекта i .

Однако такая модель слишком общая, чтобы быть полезной, и приходится накладывать какую-то структуру на коэффициенты θ_{it} .

Простейшей в этом отношении является модель **пула** (*pool*) с $\theta_{it} \equiv \theta$:

$$y_{it} = x_{it}^T \theta_{it} + u_{it},$$

в которой предполагается, что

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

и что

$$E(x_{it} u_{js}) = 0 \quad \text{для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

так что x является экзогенной переменной. В этом случае имеем дело с обычной линейной регрессией с NT наблюдениями, удовлетворяющей предполо-

жениям классической нормальной линейной модели. Для получения эффективных оценок вектора коэффициентов θ достаточно использовать обычный метод наименьших квадратов (*OLS*). При соответствующих предположениях о поведении значений объясняющих переменных, когда $N \rightarrow \infty$ или/и $T \rightarrow \infty$, эта оценка является также состоятельной оценкой этого вектора.

ПРИМЕР 3.1.1

Рассмотрим один популярный объект статистических исследований — данные о размере инвестиций (*invest*), о рыночной цене (*mvalue*) и об акционерном капитале (*kstock*) 10 крупных компаний США за период с 1935 по 1954 г. При анализе этих и других данных по большей части используем пакет статистического анализа *Stata8* и приведем протоколы оценивания, иногда с некоторыми сокращениями. Оценив по указанным статистическим данным модель пула, получим результаты, приведенные в табл. 3.1¹.

Таблица 3.1

Оценка модели пула по статистическим данным
Cross-sectional time-series FGLS regression; Coefficients: generalized least squares;
Panels: homoskedastic; Correlation: no autocorrelation

<i>Estimated covariances</i> = 1		<i>Number of obs</i> = 200		
<i>Estimated autocorrelations</i> = 0		<i>Number of groups</i> = 10		
<i>Estimated coefficients</i> = 3		<i>Time periods</i> = 20		
<i>Wald chi2(2)</i> = 866.14		<i>Prob > chi2</i> = 0.0000		
<i>Log likelihood</i> = -1191.802				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1155622	0.0057918	19.95	0.00
<i>kstock</i>	0.2306785	0.025284	9.12	0.00
<i>cons</i>	-42.71437	9.440069	-4.52	0.00

Приведенный протокол оценивания показывает, что оценивается модель пула, в которой отсутствуют гетероскедастичность и автокоррелированность ошибок. Для проверки гипотезы о незначимости регрессии в целом (т.е. гипотезы о нулевых значениях коэффициентов при объясняющих переменных *mvalue* и *kstock*) здесь используется критерий Вальда, основанный на статистике $Wald = qF$, где F — обычная F -статистика для проверки этой гипотезы; q — количество линейных ограничений на параметры модели (в данном при-

¹ В четвертой графе табл. 3.1 приведены значения отношений оценок коэффициентов к стандартным ошибкам, рассчитанным по асимптотическому нормальному распределению оценок максимального правдоподобия. В связи с этим здесь и в последующих таблицах указанное отношение называется не t -статистикой, а z -статистикой. P -значения, приводимые в пятой графе, соответствуют стандартному нормальному распределению.

мере $q = 2$). Статистика критерия Вальда имеет асимптотическое распределение хи-квадрат с q степенями свободы. Вычисленный на основе этого асимптотического распределения наблюдаемый уровень значимости (P -значение), соответствующий наблюдаемому значению 866.14, равен 0.0000. Таким образом, гипотеза о нулевых значениях коэффициентов при объясняющих переменных $mvalue$ и $kstock$ отвергается. Во второй графе таблицы приведены оценки коэффициентов, в третьей — оценки для стандартных ошибок этих оценок. В четвертой графе — значения t -статистик для раздельной проверки гипотез о равенстве нулю отдельных коэффициентов, а в пятой — соответствующие им P -значения, вычисляемые на основе нормального приближения распределения Стьюдента (отсюда обозначение z вместо обычного t в заголовке четвертой графы). Полученные P -значения говорят о высокой статистической значимости оценок коэффициентов. ■

В такой упрощенной модели, собственно, и не возникает никаких особенностей статистического анализа, связанных с панельным представлением данных. Однако положение изменится, если предположить, что в той же модели $y_{it} = x_{it}^T \theta + u_{it}$ ошибки u_{it} , оставаясь статистически независимыми между собой, имеют разные дисперсии для различных субъектов: $D(u_{it}) = \sigma_{ui}^2$. В этом случае OLS -оценки коэффициентов остаются несмещенными, но возникает смещение при оценивании дисперсий этих оценок, что отражается на оцененных значениях стандартных ошибок оценок, используемых при построении доверительных интервалов для коэффициентов и при проверке гипотез о значениях коэффициентов (например, при проверке их статистической значимости). Хотя и здесь особенность панельного характера данных отражается лишь в структуре весов: при применении **взвешенного метода наименьших квадратов** (WLS — *weighted least squares*) веса, приписываемые различным наблюдениям, не изменяются в пределах наблюдений одного субъекта.

ПРИМЕР 3.1.1 (продолжение)

Рассмотрим результаты оценивания, представленные в табл. 3.2. Здесь протокол оценивания указывает на применение взвешенного метода наименьших квадратов (*Panels: heteroskedastic*¹).

По сравнению с предыдущим результатом существенно снизилось значение оцененного коэффициента при переменной $kstock$ и произошло двукратное уменьшение оцененной стандартной ошибки для этого коэффициента. Соответственно изменился и вычисленный 95%-й интервал для данного коэффициента: теперь это интервал (0.129, 0.178), тогда как ранее это был интервал (0.181, 0.280). ■

¹ В англоязычной литературе используется двойное написание термина «гетероскедастичность»: и как *heteroscedasticity*, и как *heteroskedasticity*.

Таблица 3.2

Оценка с помощью взвешенного метода наименьших квадратов
Cross-sectional time-series FGLS regression; Coefficients: generalized least squares;
Panels: heteroskedastic; Correlation: no autocorrelation

<i>Wald chi2(2) = 669.69</i>		<i>Prob > chi2 = 0.0000</i>		
<i>Log likelihood = -1037.152</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1116328	0.0049823	22.41	0.00
<i>kstock</i>	0.1537718	0.0125707	12.23	0.00
<i>cons</i>	-21.44348	3.901219	-5.50	0.00

Заметим, что, как и в обычной модели регрессии, вместо взвешенного метода наименьших квадратов можно использовать оценки коэффициентов, полученные обычным методом наименьших квадратов (*OLS*), но при этом следует произвести **коррекцию стандартных ошибок** этих оценок. При использовании такого подхода в пакете *Stata 8* получим результаты, приведенные в табл. 3.3.

ПРИМЕР 3.1.1 (продолжение)

Таблица 3.3

Оценка с использованием коррекции стандартных ошибок
.xtpcse invest mvalue kstock, hetonly casewise;
Linear regression, heteroskedastic panels corrected standard errors

<i>Estimated covariances = 10</i>		<i>R-squared = 0.8124</i>		
<i>Wald chi2(2) = 567.87</i>		<i>Prob > chi2 = 0.0000</i>		
<i>Het-corrected</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1155622	0.0070863	16.31	0.00
<i>kstock</i>	0.2306785	0.029747	7.75	0.00
<i>cons</i>	−42.71437	7.131515	−5.99	0.00

Следующим шагом в усложнении модели может быть дополнительное предположение о наличии корреляционной связи между ошибками в уравнениях для разных субъектов в один и тот же момент времени (*cross-sectional correlation*):

$$\text{Cov}(u_{it}, u_{js}) = \begin{cases} 0 & \text{для } t \neq s \\ \sigma_{ij} (\neq 0) & \text{для } t = s, \end{cases}$$

так что матрица ковариаций $\Sigma = (\sigma_{ij})$ не является диагональной.

В этом случае приходится использовать обобщенный метод наименьших квадратов (*GLS*), учитывающий это обстоятельство. Деталей коснемся

несколько позже — вместе с моделями кажущихся несвязанными регрессий, а сейчас только посмотрим, что дает применение этого метода к анализируемым данным.

ПРИМЕР 3.1.1 (продолжение)

Результаты *GLS*-оценивания представлены в табл. 3.4. Заметим, что в этом случае помимо 3 коэффициентов модели приходится оценивать еще 10 дисперсий случайных ошибок в уравнениях для 10 предприятий, а также 45 ковариаций σ_{ij} , $i \neq j$. Если не накладывать никаких дополнительных ограничений на структуру матрицы ковариаций $\Sigma = (\sigma_{ij})$, то оценивание каждой ковариации (или дисперсии) производится на основе всего 20 наблюдений и потому может быть весьма неточным.

Таблица 3.4

GLS-оценивание модели пула

*Cross-sectional time-series FGLS regression; Coefficients: generalized least squares;
Panels: heteroskedastic with cross-sectional correlation; Correlation: no autocorrelation*

<i>Estimated covariances</i> = 55					<i>Number of obs</i> = 200				
<i>Estimated autocorrelations</i> = 0					<i>Number of groups</i> = 10				
<i>Estimated coefficients</i> = 3					<i>Time periods</i> = 20				
<i>Wald chi2(2)</i> = 3738.07					<i>Prob > chi2</i> = 0.0000				
<i>Log likelihood</i> = −879.4274									
Переменная		Коэффициент		Стандартная ошибка		z-статистика		P-значение	
<i>mvalue</i>		0.1127515		0.0022364		50.42		0.00	
<i>kstock</i>		0.2231176		0.0057363		38.90		0.00	
<i>cons</i>		−39.84382		1.717563		−23.20		0.00	

И здесь можно оставить *OLS*-оценки для коэффициентов, скорректировав оценки стандартных ошибок. Результаты приведены в табл. 3.5. ■

Таблица 3.5

OLS-оценивание модели пула

*.xtpcse invest mvalue kstock, casewise;
Linear regression, correlated panels corrected standard errors (PCSEs)*

<i>Estimated covariances</i> = 55					<i>R-squared</i> = 0.8124				
<i>Wald chi2(2)</i> = 637.41					<i>Prob > chi2</i> = 0.0000				
<i>Het-corrected</i>									
Переменная		Коэффициент		Стандартная ошибка		z-статистика		P-значение	
<i>mvalue</i>		0.1155622		0.0072124		16.02		0.00	
<i>kstock</i>		0.2306785		0.0278862		8.27		0.00	
<i>cons</i>		−42.71437		6.780965		−6.30		0.00	

Следующим шагом в усложнении модели является снятие предположения о взаимной независимости ошибок в пределах одного субъекта — например, путем предположения о том, что последовательность ошибок при наблюдении i -го субъекта следует процессу авторегрессии первого порядка AR(1) с нулевым средним. Поясним это на примере модели $y_{it} = x_{it}^T \theta + u_{it}$, в которой

$$u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it},$$

где $|\rho_i| < 1$;

$\varepsilon_{i1}, \dots, \varepsilon_{iT}$ — случайные величины, являющиеся гауссовскими инновациями, так что они взаимно независимы и имеют одинаковое распределение $N(0, \sigma_{\varepsilon i}^2)$, и, кроме того, ε_{it} не зависит от значений $u_{i,t-k}$, $k \geq 1$.

Коэффициент ρ_i можно оценить различными способами. Например, можно оценить (методом наименьших квадратов) модель $y_{it} = x_{it}^T \theta + u_{it}$ без учета автокоррелированности ошибок, получить последовательность остатков $\hat{u}_{i1}, \hat{u}_{i2}, \dots, \hat{u}_{iT}$, вычислить значение статистики Дарбина — Уотсона

$$d_i = \frac{\sum_{t=2}^T (\hat{u}_{it} - \hat{u}_{i,t-1})^2}{\sum_{t=1}^T \hat{u}_{it}^2}$$

и, используя приближенное соотношение $\rho_i \cong 1 - \frac{d_i}{2}$, получить приближен-

ную оценку $\hat{\rho}_{i,DW} \cong 1 - \frac{d_i}{2}$.

Можно поступить иначе: получив последовательность остатков $\hat{u}_{i1}, \hat{u}_{i2}, \dots, \hat{u}_{iT}$, использовать оценку наименьших квадратов, вычисленную при оценивании уравнения регрессии

$$\hat{u}_{it} = \rho_i \hat{u}_{i,t-1} + \eta_{it}.$$

Искомая оценка вычисляется по формуле:

$$\hat{\rho}_i = \frac{\sum_{t=2}^T \hat{u}_{it} \hat{u}_{i,t-1}}{\sum_{t=1}^T \hat{u}_{it}^2}.$$

(В пакете *Stata* 8 эта оценка обозначена как $\hat{\rho}_{iscorr}$.)

После вычисления оценок для ρ_i , $i = 1, \dots, N$, в уравнениях для каждого субъекта производится известное преобразование Прайса — Уинстена пере-

менных для получения модели с независимыми ошибками. Объединив преобразованные уравнения в систему, можем произвести в ней *OLS*-оценивание коэффициентов.

Если предполагается, что уравнения имеют общий *AR*-параметр, т.е. $\rho_1 = \rho_2 = \dots = \rho_N = \rho$, то это общее значение ρ оценивается величиной $\hat{\rho} = \frac{\rho_1 + \rho_2 + \dots + \rho_N}{N}$ и в преобразовании Прайса — Уинстена используется одна эта оценка.

ПРИМЕР 3.1.1 (продолжение)

Будем предполагать, что дисперсии ошибок для разных субъектов могут быть различными. Если при этом предполагается отсутствие перекрестной коррелированности ошибок между уравнениями и ошибки в уравнениях для разных субъектов следуют одинаковым *AR*(1)-моделям (с общим ρ), то оценивание такой модели дает результаты, представленные в табл. 3.6. При этом использовалась оценка, вычисляемая по статистике Дарбина — Уотсона.

Таблица 3.6

Оценка ρ по статистике Дарбина — Уотсона

.xtpcse invest mvalue kstock, correlation(ar1) hetonly rho type(dw)
 (note: estimates of rho outside [-1, 1] bounded to be in the range [-1, 1])
 Prais-Winsten regression, heteroskedastic panels corrected standard errors

<i>Estimated covariances</i> = 10		<i>R-squared</i> = 0.5468		
<i>Estimated autocorrelations</i> = 1				
<i>Estimated coefficients</i> = 3				
<i>Wald chi2(2)</i> = 91.72		<i>Prob</i> > <i>chi2</i> = 0.0000		
<i>Het-corrected</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.0972395	0.0126259	7.70	0.000
<i>kstock</i>	0.306441	0.0561245	5.46	0.000
<i>cons</i>	−42.07116	21.02442	−2.00	0.045
<i>rho</i>	0.8678619			

Если использовать второй вариант оценивания коэффициента ρ , описанный выше, это приведет к результатам, указанным в табл. 3.7. Оцененное значение ρ существенно изменилось.

Если допускается перекрестная коррелированность ошибок между уравнениями и ошибки в уравнениях для разных субъектов следуют одинаковым *AR*(1)-моделям (с общим ρ), то оценивание такой модели (по *DW*-варианту) дает результаты, представленные в табл. 3.8. И опять при использовании второго варианта оценивания коэффициента ρ получим несколько отличные результаты (табл. 3.9).

Таблица 3.7

OLS-оценивание коэффициента ρ *.xtpcse invest mvalue kstock, correlation(ar1) hetonly rhotype(tscorr) casewise**Prais-Winsten regression, heteroskedastic panels corrected standard errors*

<i>Estimated covariances = 10</i>		<i>R-squared = 0.6904</i>		
<i>Estimated autocorrelations = 1</i>				
<i>Estimated coefficients = 3</i>				
<i>Wald chi2(2) = 192.41</i>		<i>Prob > chi2 = 0.0000</i>		
<i>Het-corrected</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1032102	0.0112252	9.19	0.000
<i>kstock</i>	0.2947986	0.0459298	6.42	0.000
<i>cons</i>	−45.78767	13.97367	−3.28	0.001
<i>rho</i>	0.7563511			

Таблица 3.8

DW-оценка ρ при допущении перекрестной коррелированности ошибок*.xtpcse invest mvalue kstock, correlation(ar1) rhotype(dw)**(note: estimates of rho outside [-1, 1] bounded to be in the range [-1, 1])**Prais-Winsten regression, heteroskedastic panels corrected standard errors;**Panels: heteroskedastic (balanced)*

<i>Estimated covariances = 55</i>		<i>R-squared = 0.5468</i>		
<i>Estimated autocorrelations = 1</i>				
<i>Estimated coefficients = 3</i>				
<i>Wald chi2(2) = 120.05</i>		<i>Prob > chi2 = 0.0000</i>		
<i>Panel-corrected</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.0972395	0.0124362	7.82	0.000
<i>kstock</i>	0.306441	0.054533	5.62	0.000
<i>cons</i>	−42.07116	24.09387	−1.75	0.081
<i>rho</i>	0.8678619			

Таблица 3.9

OLS-оценка ρ при допущении перекрестной коррелированности ошибок*.xtpcse invest mvalue kstock, correlation(ar1) rhotype(tscorr) casewise Prais-Winsten regression,**correlated panels corrected standard errors (PCSEs); Panels: correlated (balanced)*

<i>Estimated covariances</i> = 55	<i>R-squared</i> = 0.6904
<i>Estimated autocorrelations</i> = 1	
<i>Estimated coefficients</i> = 3	
<i>Wald chi2(2)</i> = 215.52	<i>Prob > chi2</i> = 0.0000
<i>Panel-corrected</i>	

Окончание табл. 3.9

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1032102	0.0108656	9.50	0.000
<i>kstock</i>	0.2947986	0.0432809	6.81	0.000
<i>cons</i>	−45.78767	15.24513	−3.00	0.003
<i>rho</i>	0.7563511			

Посмотрим, что дает оценивание модели с перекрестной коррелированностью ошибок между уравнениями, когда ошибки в уравнениях для разных субъектов могут следовать разным AR(1)-моделям (с разными ρ_i). Результаты оценивания такой модели (по *DW*-варианту) приведены в табл. 3.10, а оценивания по второму варианту — в табл. 3.11.

Таблица 3.10

Оценка ρ по Дарбину — Уотсону при перекрестной коррелированности ошибок

.xtpcse invest mvalue kstock, correlation(pсар1) rhotype(dw) casewise
Prais-Winsten regression, correlated panels corrected standard errors (PCSEs);
Panels: correlated (balanced); Autocorrelation: panel-specific AR(1)

<i>Estimated covariances</i> = 55		<i>R-squared</i> = 0.7570		
<i>Estimated autocorrelations</i> = 10				
<i>Estimated coefficients</i> = 3				
<i>Wald chi2(2)</i> = 211.38		<i>Prob > chi2</i> = 0.0000		
<i>Panel-corrected</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1013946	0.0108632	9.33	0.000
<i>kstock</i>	0.3449446	0.0478113	7.21	0.000
<i>cons</i>	−41.18685	19.33078	−2.13	0.033
<i>rhos</i>	0.7427231, 0.8831453, 0.9741851, 0.7277056, 0.9564705, ..., 0.9343119			

Таблица 3.11

OLS-оценка ρ при перекрестной коррелированности ошибок

.xtpcse invest mvalue kstock, correlation(pсар1) rhotype(tscorr) casewise
Prais-Winsten regression, correlated panels corrected standard errors (PCSEs);
Panels: correlated (balanced); Autocorrelation: panel-specific AR(1)

<i>Estimated covariances</i> = 55	<i>R-squared</i> = 0.8670
<i>Estimated autocorrelations</i> = 10	
<i>Estimated coefficients</i> = 3	
<i>Wald chi2(2)</i> = 444.53	<i>Prob > chi2</i> = 0.0000
<i>Panel-corrected</i>	

Окончание табл. 3.11

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>mvalue</i>	0.1052613	0.0086018	12.24	0.000
<i>kstock</i>	0.3386743	0.0367568	9.21	0.000
<i>cons</i>	-58.18714	12.63687	-4.60	0.000
<i>rhos</i>	0.5135627, 0.8701700, 0.9023497, 0.6336800, 0.8571502, ..., 0.8752707			

Заметим, что и здесь приходится оценивать значительное количество дисперсий, ковариаций и автокорреляций, используя всего 20 наблюдений, растянутых во времени.

Для удобства объединим полученные результаты в одну таблицу, дополнительно указав в последней графе 95%-е доверительные интервалы для коэффициентов (табл. 3.12).

Таблица 3.12

Сравнение оценок модели пула

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение	95%-й интервал
Независимые одинаково распределенные ошибки					
<i>mvalue</i>	0.116	0.0058	19.95	0.00	0.104 0.127
<i>kstock</i>	0.231	0.0253	9.12	0.00	0.181 0.280
Гетероскедастичность — WLS					
<i>mvalue</i>	0.112	0.0050	22.41	0.00	0.102 0.122
<i>kstock</i>	0.154	0.0126	12.23	0.00	0.129 0.178
SUR — GLS					
<i>mvalue</i>	0.113	0.0022	50.42	0.00	0.108 0.117
<i>kstock</i>	0.223	0.0057	38.90	0.00	0.212 0.234
AR(1) — common rho (Durbin — Watson): est rho = 0.8678619					
<i>mvalue</i>	0.097	0.0126	7.70	0.00	0.072 0.122
<i>kstock</i>	0.306	0.0561	5.46	0.00	0.196 0.416
AR(1) — common rho (OLS): est rho = 0.7563511					
<i>mvalue</i>	0.103	0.0112	9.19	0.00	0.081 0.125
<i>kstock</i>	0.295	0.0460	6.42	0.00	0.205 0.385
SUR&AR(1) — common rho (D-W): est rho = 0.8678619					
<i>mvalue</i>	0.097	0.0124	7.82	0.00	0.073 0.122
<i>kstock</i>	0.306	0.0545	5.62	0.00	0.200 0.413

Окончание табл. 3.12

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение	95%-й интервал
SUR&AR(1) — <i>common rho (OLS): est rho = 0.7563511</i>					
<i>mvalue</i>	0.103	0.0109	9.50	0.00	0.082 0.125
<i>kstock</i>	0.295	0.0433	6.81	0.00	0.210 0.388
SUR&AR(1) — <i>different rhos (D-W)</i>					
<i>mvalue</i>	0.101	0.0109	9.33	0.00	
<i>kstock</i>	0.345	0.0478	7.21	0.00	
<i>rhos</i>	0.7427231, 0.8831453, 0.9741851, 0.7277056, ..., 0.9343119				
SUR&AR(1) — <i>different rhos (TSCORR)</i>					
<i>mvalue</i>	0.105	0.0086	12.24	0.00	
<i>kstock</i>	0.339	0.0368	9.21	0.00	
<i>rhos</i>	0.5135627, 0.8701700, 0.9023497, 0.6336800, ..., 0.8752707				

Здесь SUR&AR(1) означает наличие корреляции между ошибками в разных уравнениях в совпадающие моменты времени и AR(1)-модель для ошибок в пределах каждого предприятия. ■

Панельные данные — модель ковариационного анализа

Обратимся теперь к модели ковариационного анализа (*covariance analysis model*):

$$M_0 : y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где α_i и β_i — неизвестные постоянные;

x — переменная, для простоты ее будем рассматривать пока как скалярную переменную;

u_{it} — случайные ошибки.

Если предполагать, что

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

и что

$$E(x_{it}u_{js}) = 0 \text{ для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

так что x является экзогенной переменной, то мы имеем дело с N **не связанными между собой** (*unrelated*) линейными регрессиями, удовлетворяющими предположениям классической нормальной линейной регрессии. Для получения оценок параметров α_i и β_i эти регрессии можно оценивать в этом случае порознь. Таким образом, оценки наименьших квадратов для α_i и β_i имеют вид:

$$\hat{\beta}_i = \frac{\sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)}{\sum_{t=1}^T (x_{it} - \bar{x}_i)^2},$$

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta}_i \bar{x}_i, \quad i = 1, \dots, N,$$

где

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}.$$

Эти оценки при фиксированных значениях y_{it} и x_{it} имеют нормальное распределение и являются наилучшими линейными несмещенными оценками (*BLUE*).

Если ошибки u_{it} независимы между собой и имеют нормальные распределения с нулевыми средними, но дисперсии их различны для разных субъектов, так что

$$u_{it} \sim i.i.d. N(0, \sigma_{ui}^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

то следует использовать взвешенный метод наименьших квадратов, приписывая каждому наблюдению для i -го субъекта вес $w_i = \frac{1}{\sigma_{ui}}$. Поскольку дисперсии σ_{ui}^2 в реальных исследованиях не известны, приходится использовать **доступную (*feasible*) версию** этого метода, в которой вместо весов $w_i = \frac{1}{\sigma_{ui}}$

берутся их оценки $w_i = \frac{1}{\hat{\sigma}_{ui}}$, где $\hat{\sigma}_{ui}^2$ — подходящие оценки неизвестных дисперсий. В качестве таковых можно брать, например, несмещенные оценки дисперсий σ_{ui}^2 :

$$\hat{\sigma}_{ui}^2 = \frac{RSS^{(i)}}{N - p},$$

где $RSS^{(i)}$ — сумма квадратов остатков, получаемых при оценивании модели регрессии для i -го субъекта;

p — количество объясняющих переменных в уравнениях регрессии (для парной регрессии с константой $p = 2$).

Несколько более сложная модель возникает, если предположить коррелированность ошибок для разных субъектов в совпадающие моменты времени. Это **модель кажущихся несвязанными регрессий**¹ (*SUR* — *seemingly*

¹ В некоторых руководствах по эконометрике такую модель называют **системой внешне не связанных уравнений**.

unrelated regressions). При наличии такой коррелированности следует использовать уже не взвешенный, а обобщенный метод наименьших квадратов (*GLS*). Если представить уравнение для i -го субъекта в векторно-матричной форме:

$$y^{(i)} = X^{(i)}\theta^{(i)} + u^{(i)},$$

где

$$y^{(i)} = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix}, \quad X^{(i)} = \begin{pmatrix} 1 & x_{i1} \\ \vdots & \vdots \\ 1 & x_{iT} \end{pmatrix}, \quad \theta^{(i)} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad u^{(i)} = \begin{pmatrix} u_{i1} \\ \vdots \\ u_{iT} \end{pmatrix},$$

то модель *SUR* можно записать в следующем виде:

$$\begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(N)} \end{pmatrix} = \begin{pmatrix} X^{(1)} & 0 & \cdots & 0 \\ 0 & X^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X^{(N)} \end{pmatrix} \begin{pmatrix} \theta^{(1)} \\ \vdots \\ \theta^{(N)} \end{pmatrix} + \begin{pmatrix} u^{(1)} \\ \vdots \\ u^{(N)} \end{pmatrix},$$

или (в очевидных обозначениях)

$$y = X\theta + u.$$

При сделанных предположениях ковариационная матрица $(NT \times 1)$ -вектора u равна:

$$\Omega = \text{Cov}(u) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1N} \\ \Sigma_{12} & \Sigma_{22} & \cdots & \Sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{1N} & \Sigma_{2N} & \cdots & \Sigma_{NN} \end{pmatrix},$$

где $(T \times T)$ -матрица Σ_{ij} имеет вид:

$$\Sigma_{ij} = \begin{pmatrix} \sigma_{ij} & 0 & \cdots & 0 \\ 0 & \sigma_{ij} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{ij} \end{pmatrix}, \quad \sigma_{ij} = \text{Cov}(u_{it}, u_{jt}).$$

Обобщенная оценка наименьших квадратов (*GLS*-оценка) вектора θ находится по формуле:

$$\hat{\theta}_{SUR} = \hat{\theta}_{GLS} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} y.$$

Заметим, что подлежащая обращению матрица Ω имеет размер $(NT \times NT)$. Однако такого обращения можно избежать, используя соотношение:

$$\Omega^{-1} = \begin{pmatrix} \Sigma_{11}^{(-1)} & \Sigma_{12}^{(-1)} & \cdots & \Sigma_{1N}^{(-1)} \\ \Sigma_{12}^{(-1)} & \Sigma_{22}^{(-1)} & \cdots & \Sigma_{2N}^{(-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{1N}^{(-1)} & \Sigma_{2N}^{(-1)} & \cdots & \Sigma_{NN}^{(-1)} \end{pmatrix},$$

где

$$\Sigma_{ij}^{(-1)} = \begin{pmatrix} \sigma_{ij}^{(-1)} & 0 & \cdots & 0 \\ 0 & \sigma_{ij}^{(-1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{ij}^{(-1)} \end{pmatrix},$$

а $\sigma_{ij}^{(-1)}$ — элементы матрицы Σ^{-1} , обратной к матрице

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{1N} & \cdots & \sigma_{NN} \end{pmatrix}.$$

Благодаря этому достаточно произвести обращение матрицы размера $(N \times N)$.

Учет коррелированности ошибок в различных уравнениях позволяет ожидать определенного выигрыша в точности оценивания каждого из $\theta^{(i)}$ за счет информации, идущей от других уравнений через указанную коррелированность. Однако реальный выигрыш зависит от целого ряда факторов. Например, если $\sigma_{ij} = \sigma^2 \rho$ для $i \neq j$, то предпочтительность *SUR*-оценки возрастает с ростом $\rho \rightarrow 1$, когда T велико. В то же время если $\rho = 0$, то *SUR*-оценка и *OLS*-оценка совпадают. Кроме того, непосредственная реализация *SUR*-оценивания на практике невозможна из-за того, что значения σ_{ij} исследователю не известны. **Доступный** вариант *SUR*-оценивания предусматривает использование адаптивной оценки $\hat{\theta}_{FGLS}$ вектора θ , при вычислении которой неизвестные значения σ_{ij} заменяются их состоятельными оценками $\hat{\sigma}_{ij}$.

Пусть $e^{(i)} = y^{(i)} - X^{(i)}\hat{\theta}^{(i)}$ — вектор остатков, получаемый при *OLS*-оценивании уравнения для i -го субъекта. Тогда естественной оценкой для σ_{ij} является

$$\hat{\sigma}_{ij} = \frac{(e^{(i)})^T e^{(j)}}{T}.$$

При $j = i$ это есть просто $\frac{RSS^{(i)}}{T}$, и, как известно, такая оценка для дисперсии ошибки в i -м уравнении имеет смещение, а несмещенной оценкой для этой дисперсии является $\frac{RSS^{(i)}}{T - p}$, где p — количество объясняющих перемен-

ных в уравнении регрессии. (Конечно, при этом должно выполняться условие $T > p$.) При соответствующих условиях, налагающихся на матрицу X и в классической модели линейной регрессии, обе оценки — $\hat{\theta}_{GLS}$ и $\hat{\theta}_{FGLS}$ — при $T \rightarrow \infty$ состоятельны.

ПРИМЕР 3.1.2

В табл. 3.13 воспроизведены приведенные в (Greene, 2003, р. 335) ежегодные данные об объемах инвестиций (y) и прибыли (x) 3 предприятий ($N = 3$) за десятилетний период ($T = 10$).

Таблица 3.13

Объем инвестиций (y) и прибыль (x) предприятий за 10 лет

t	Первое предприятие		Второе предприятие		Третье предприятие	
	$Y1$	$X1$	$Y2$	$X2$	$Y3$	$X3$
1	13.32	12.85	20.30	22.93	8.85	8.65
2	26.30	25.69	17.47	17.96	19.60	16.55
3	2.62	5.48	9.31	9.16	3.87	1.47
4	14.94	13.79	18.01	18.73	24.19	24.91
5	15.80	15.41	7.63	11.31	3.99	5.01
6	12.20	12.59	19.84	21.15	5.73	8.34
7	14.93	16.64	13.76	16.13	26.68	22.70
8	29.82	26.45	10.00	11.61	11.49	8.36
9	20.32	19.64	19.51	19.55	18.49	15.44
10	4.77	5.43	18.32	17.06	20.84	17.87

На рис. 3.1—3.3 приведены графики изменения объемов инвестиций и прибыли по каждому из 3 предприятий.

Раздельное оценивание уравнений регрессии (в пакете *EViews*) дает результаты, представленные в табл. 3.14.

Матрица $\Sigma = (\sigma_{ij})$ оценивается следующим образом:

$$\hat{\Sigma} = \begin{pmatrix} 1.2549 & -0.0099 & -0.9101 \\ -0.0099 & 1.9628 & 1.0351 \\ -0.9101 & 1.0351 & 4.3279 \end{pmatrix}.$$

Соответствующая корреляционная матрица имеет вид:

$$\begin{pmatrix} 1 & -0.0063 & -0.3905 \\ -0.0063 & 1 & 0.3552 \\ -0.3905 & 0.3552 & 1 \end{pmatrix}.$$

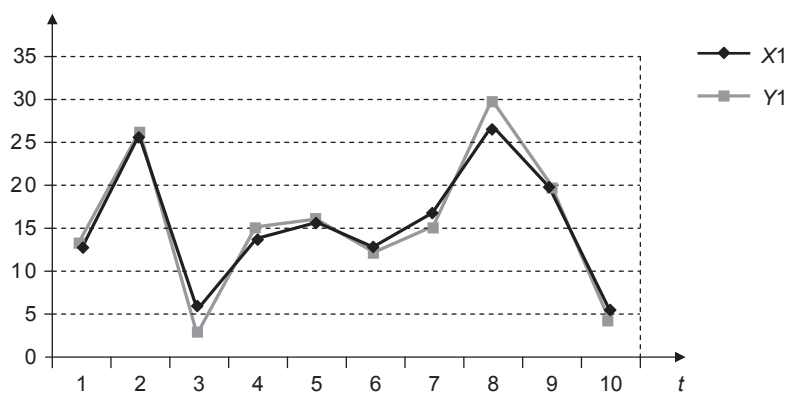


Рис. 3.1

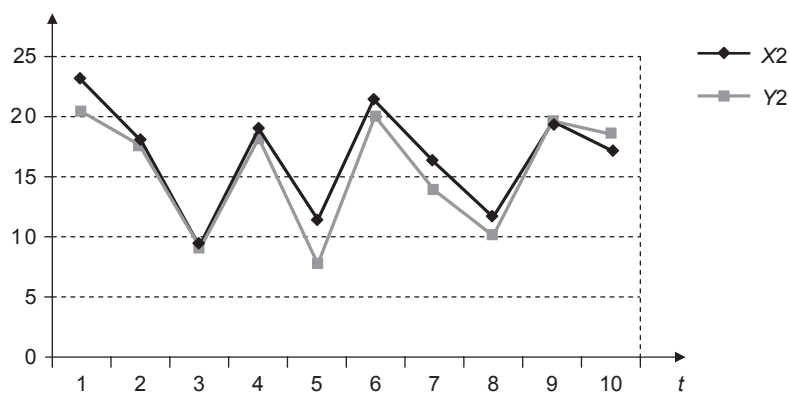


Рис. 3.2

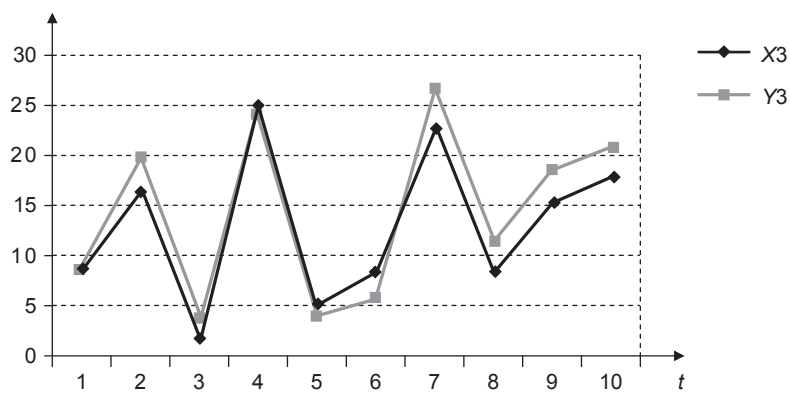


Рис. 3.3

Таблица 3.14

Результаты раздельного оценивания уравнений регрессии

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
Первое предприятие				
C	−2.468913	0.980426	−2.518205	0.0359
X1	1.167170	0.058250	20.03737	0.0000
R-squared	0.9805			
Второе предприятие				
C	−1.384797	1.972680	−0.701988	0.5026
X2	1.014542	0.115314	8.798102	0.0000
R-squared	0.9063			
Третье предприятие				
C	0.455479	1.491604	0.305362	0.7679
X3	1.076374	0.100360	10.72516	0.0000
R-squared	0.9350			

При использовании доступного *GLS* получим результаты, приведенные в табл. 3.15. ■

Таблица 3.15

Результаты *GLS*-оценивания уравнений регрессии

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
Первое предприятие				
C	−2.857213	0.812548	−3.52	0.002
X1	1.192389	0.047490	25.11	0.000
R-squared	0.9800			
Второе предприятие				
C	−2.11701	1.66034	−1.28	0.214
X2	1.05876	0.09663	10.96	0.000
R-squared	0.9046			
Третье предприятие				
C	0.721196	1.199687	0.60	0.553
X3	1.055824	0.077589	13.61	0.000
R-squared	0.9346			

Оцененные коэффициенты несколько отличаются от результатов раздельного оценивания уравнений, и вопрос в том, сколь значимым является это различие. В связи с этим представляет интерес проверка гипотезы

$$H_0: \sigma_{ij} = 0 \text{ для } i \neq j.$$

В предположении нормальности ошибок эта гипотеза соответствует статистической независимости ошибок в разных уравнениях. Для проверки этой гипотезы можно использовать **критерий Бройша — Пагана** (*Breusch — Pagan test*). Статистика этого критерия равна:

$$\lambda = T \sum_{i=2}^N \sum_{j=1}^{i-1} r_{ij}^2,$$

где $r_{ij} = \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii} \hat{\sigma}_{jj}}}$ — оцененная корреляция между ошибками в i -м и j -м уравнениях.

При гипотезе H_0 эта статистика имеет асимптотическое распределение хи-квадрат с числом степеней свободы, равным $N(N-1)/2$ (заметим, что гипотеза H_0 накладывает именно столько ограничений, поскольку $\sigma_{ij} = \sigma_{ji}$).

В нашем примере накладывается 3 ограничения, статистика критерия принимает значение 2.787. Соответствующее ему P -значение, вычисленное на основе распределения $\chi^2(3)$, равно 0.4256, так что если ориентироваться на это P -значение, то гипотеза независимости не отвергается.

Следует также обратить внимание на то, что различие между оценками коэффициентов при переменных x_1, x_2, x_3 довольно невелико, так что возникает вопрос о проверке гипотезы совпадения этих коэффициентов:

$$H_0: \beta_1 = \beta_2 = \beta_3.$$

В рамках модели SUR для проверки этой гипотезы используются две формы критерия Вальда: одна основана на F -статистике и P -значении, рассчитанном исходя из соответствующего F -распределения, а другая — на статистике qF (q — количество линейных ограничений) и P -значении, рассчитанном исходя из асимптотического распределения $\chi^2(q)$ этой статистики. Результаты использования этих двух форм приведены в табл. 3.16.

Таблица 3.16

Результаты использования двух форм критерия Вальда

Wald Test:			
<i>F-statistic</i>	1.342317	<i>Probability</i>	0.278120
<i>Chi-square</i>	2.684634	<i>Probability</i>	0.261240

Разница в двух P -значениях весьма мала и не приводит к различию в статистических выводах — гипотеза $H_0: \beta_1 = \beta_2 = \beta_3$ не отвергается.

Поскольку ранее на основании применения критерия Бройша — Пагана не была отвергнута гипотеза независимости ошибок в разных уравнениях, естественно попытаться проверить гипотезу $H_0: \beta_1 = \beta_2 = \beta_3$ в условиях такой независимости.

Заметим, что модель SUR в нашем примере записывается следующим образом:

$$\begin{pmatrix} y_{11} \\ \vdots \\ y_{1T} \\ y_{21} \\ \vdots \\ y_{2T} \\ y_{31} \\ \vdots \\ y_{3T} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1T} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & x_{21} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & x_{2T} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_{31} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & x_{3T} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} u_{11} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{2T} \\ u_{31} \\ \vdots \\ u_{3T} \end{pmatrix}.$$

Это означает, что ее можно рассматривать как модель линейной регрессии переменной y , принимающей значения $y_{11}, y_{12}, \dots, y_{1T}, y_{21}, y_{22}, \dots, y_{2T}, y_{31}, y_{32}, \dots, y_{3T}$, на следующие 6 переменных: три **дамми-переменных** (*dummy variables*):

d_1 со значениями $\underbrace{1, \dots, 1}_T, \underbrace{0, \dots, 0}_T, \underbrace{0, \dots, 0}_T$;

d_2 со значениями $\underbrace{0, \dots, 0}_T, \underbrace{1, \dots, 1}_T, \underbrace{0, \dots, 0}_T$;

d_3 со значениями $\underbrace{0, \dots, 0}_T, \underbrace{0, \dots, 0}_T, \underbrace{1, \dots, 1}_T$;

и три комбинированных переменных d_1x , d_2x и d_3x , построенных на основании указанных дамми-переменных и переменной x , принимающей значения $x_{11}, x_{12}, \dots, x_{1T}, x_{21}, x_{22}, \dots, x_{2T}, x_{31}, x_{32}, \dots, x_{3T}$. Переменные d_1x , d_2x и d_3x принимают значения:

$d_1x: \underbrace{x_{11}, x_{12}, \dots, x_{1T}}_T, \underbrace{0, 0, \dots, 0}_T, \underbrace{0, 0, \dots, 0}_T$;

$d_2x: \underbrace{0, 0, \dots, 0}_T, \underbrace{x_{21}, x_{22}, \dots, x_{2T}}_T, \underbrace{0, 0, \dots, 0}_T$;

$d_3x: \underbrace{0, 0, \dots, 0}_T, \underbrace{0, 0, \dots, 0}_T, \underbrace{x_{31}, x_{32}, \dots, x_{3T}}_T$.

Если случайные ошибки в разных уравнениях статистически независимы, можно получить эффективные оценки коэффициентов, применяя *OLS*-оцени-

вание, и проверить гипотезу $H_0: \beta_1 = \beta_2 = \beta_3$ с использованием обычного F -критерия. При этом получим результаты, приведенные в табл. 3.17. Гипотеза $H_0: \beta_1 = \beta_2 = \beta_3$ не отвергается и в предположении независимости ошибок.

Таблица 3.17

OLS-оценивание с помощью обычного F -критерия

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение
<i>D1</i>	−2.468913	1.388033	−1.778714	0.0880
<i>D2</i>	−1.384797	2.233064	−0.620133	0.5410
<i>D3</i>	0.455479	1.137109	0.400559	0.6923
<i>D1</i> *X	1.167170	0.082467	14.15324	0.0000
<i>D2</i> *X	1.014542	0.130535	7.772208	0.0000
<i>D3</i> *X	1.076374	0.076508	14.06874	0.0000
<i>R-squared</i>	0.950532			
<i>Wald Test:</i>				
<i>F</i> -statistic	0.592788	<i>Probability</i>		0.560676

Будем говорить о модели

$$M_0: y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

как о модели без ограничений. Обозначим остаточную сумму квадратов (RSS) в этой модели как S_0 :

$$S_0 = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\alpha}_i - \hat{\beta}_i x_{it})^2.$$

В рамках модели без ограничений рассмотрим две гипотезы:

- 1) $H_1: \beta_i$ одинаковы для всех i ,
- 2) $H_2: \beta_i$ и α_i одинаковы для всех i .

Гипотеза $H_1: \beta_i$ одинаковы для всех i . Этой гипотезе соответствует модель

$$M_1: y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Остаточную сумму квадратов (RSS) в модели M_1 обозначим через S_1 :

$$S_1 = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\alpha}_i - \hat{\beta} x_{it})^2.$$

Модель M_1 можно записать в виде:

$$y_{it} = \sum_{j=1}^N \alpha_j d_{ij} + \beta x_{it} + u_{it},$$

где $d_{ij} = 1$, если $j = i$, и $d_{ij} = 0$ в противном случае, так что здесь в правой части имеем N дамми-переменных. Обозначим:

$$y = (y_{11}, y_{12}, \dots, y_{1T}, y_{21}, y_{22}, \dots, y_{2T}, \dots, y_{N1}, y_{N2}, \dots, y_{NT})^T,$$

$$x = (x_{11}, x_{12}, \dots, x_{1T}, x_{21}, x_{22}, \dots, x_{2T}, \dots, x_{N1}, x_{N2}, \dots, x_{NT})^T,$$

$$u = (u_{11}, u_{12}, \dots, u_{1T}, u_{21}, u_{22}, \dots, u_{2T}, \dots, u_{N1}, u_{N2}, \dots, u_{NT})^T,$$

$$d_1 = \left(\underbrace{1, 1, \dots, 1}_T, \underbrace{0, 0, \dots, 0}_{NT-T} \right)^T,$$

$$d_2 = \left(\underbrace{0, 0, \dots, 0}_T, \underbrace{1, 1, \dots, 1}_T, \underbrace{0, 0, \dots, 0}_{NT-2T} \right)^T,$$

.....

$$d_N = \left(\underbrace{0, 0, \dots, 0}_{NT-T}, \underbrace{1, 1, \dots, 1}_T \right)^T,$$

и пусть $X = [d_1 \ d_2 \ \dots \ d_N \ x]$ — матрица размера $(NT \times (N + 1))$, столбцами которой являются векторы d_1, d_2, \dots, d_N, x . В этих обозначениях модель M_1 принимает вид:

$$y = X\theta + u,$$

где $\theta = (\alpha_1, \alpha_2, \dots, \alpha_N, \beta)^T$.

Соответственно оценка наименьших квадратов $\hat{\theta} = (\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_N, \hat{\beta})^T$ для вектора θ вычисляется по формуле:

$$\hat{\theta} = (X^T X)^{-1} X^T y.$$

Будем предполагать далее, что

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

и что

$$E(x_{it} u_{js}) = 0 \text{ для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

так что x является экзогенной переменной.

При этих предположениях и при фиксированной матрице X оценка $\hat{\theta}$ имеет $(N+1)$ -мерное нормальное распределение, причем

$$E(\hat{\theta}) = \theta,$$

т.е. $\hat{\theta}$ является несмещенной оценкой для θ , а ковариационная матрица случайного вектора $\hat{\theta}$ имеет вид:

$$\text{Cov}(\hat{\theta}) = \sigma_u^2 (X^T X)^{-1}.$$

Интересно, что численно *те же самые* значения оценок параметров $\alpha_1, \alpha_2, \dots, \alpha_N, \beta$ можно получить иным способом. Пусть

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \bar{u}_i = \frac{1}{T} \sum_{t=1}^T u_{it} \quad \text{— средние по времени значения}$$

переменных y, x и ошибки для i -го субъекта исследования.

Усреднив по времени обе части уравнений

$$y_{it} = \alpha_i + \beta x_{it} + u_{it},$$

получим

$$\bar{y}_i = \alpha_i + \beta \bar{x}_i + \bar{u}_i, \quad i = 1, \dots, N.$$

Из двух последних уравнений находим:

$$y_{it} - \bar{y}_i = \beta (x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T$$

(«модель, скорректированная на индивидуальные средние»).

В правой части полученной модели оказались исключенными параметры $\alpha_1, \alpha_2, \dots, \alpha_N$. В рамках этой модели оценка параметра β , который обычно представляет первоочередной интерес, вычисляется по формуле:

$$\hat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2},$$

и о ней говорят как о **«внутригрупповой» оценке** («*within-group estimate*»), имея в виду, что она строится только на основании отклонений значений переменных от их средних по времени и тем самым учитывает только изменчивость в пределах каждого субъекта, не обращая внимания на изменчивость между субъектами. Точнее, конечно, следовало бы говорить о **«внутрисубъектной»** оценке, но используемая здесь терминология заимствована из теории дисперсионного анализа, где субъекты исследования часто объединяются в некоторое количество групп, так что индекс i относится не к отдельному субъекту, а к группе субъектов. Впрочем, в последнее время в эконометрической литературе об указанной оценке чаще стали говорить просто как

о «внутри»-оценке (*within-estimator*). Соответственно и мы часто будем использовать этот термин.

Определив в последней модели оценку наименьших квадратов $\hat{\beta}$, оценки для параметров $\alpha_1, \alpha_2, \dots, \alpha_N$ можно вычислить следующим образом:

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta} \bar{x}_i, \quad i = 1, \dots, N.$$

Полученные в итоге этих двух шагов оценки $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_N, \hat{\beta}$ численно совпадают с оценками наименьших квадратов в модели $y_{it} = \sum_{j=1}^N \alpha_j d_{ij} + \beta x_{it} + u_{it}$, в чем можно убедиться непосредственно. Следует только учитывать, что стандартные ошибки оценок $\hat{\alpha}_i$, получаемые в результате двухшаговой процедуры, нельзя вычислять по формулам для стандартных ошибок оценок наименьших квадратов.

Гипотеза H_2 : β_i и α_i одинаковы для всех i . Ей соответствует модель

$$M_2: y_{it} = \alpha + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T \text{ (пул — pool)}.$$

Оценки наименьших квадратов для параметров α и β вычисляются по формулам:

$$\hat{\beta} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2},$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x},$$

$$\text{где } \bar{y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it}, \quad \bar{x} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{it}.$$

Обозначим остаточную сумму квадратов в модели M_2 через S_2 :

$$S_2 = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\alpha} - \hat{\beta} x_{it})^2.$$

Рассмотрим задачу проверки гипотез H_1 и H_2 в рамках модели M_0 при сделанных ранее предположениях.

Проверка гипотезы H_2 . Используем F -статистику

$$F_2 = \frac{\frac{S_2 - S_0}{2(N-1)}}{\frac{S_0}{NT - 2N}};$$

если гипотеза H_2 верна, то $F_2 \sim F_{2(N-1), NT-2N}$.

Если значение F_2 статистически незначимо, следует объединить данные в пул. Если же значение F_2 статистически значимо, следует искать источник гетерогенности параметров.

Проверка гипотезы H_1 . Используем F -статистику

$$F_1 = \frac{\frac{S_1 - S_0}{N-1}}{\frac{S_0}{NT-2N}};$$

если гипотеза H_1 верна, то $F_1 \sim F_{N-1, NT-2N}$.

Если значение F_1 статистически значимо, проверка прекращается. Если же значение F_1 статистически незначимо, гипотеза H_1 (о совпадении всех β_i) не отвергается.

Можно также применить условный критерий гетерогенности α_i , а именно проверить гипотезу

$$H_3: \alpha_1 = \dots = \alpha_N$$

при условии $\beta_1 = \dots = \beta_N$, т.е. в рамках модели

$$M_1: y_{it} = \alpha_i + \beta x_{it} + u_{it}.$$

Для этой цели используем статистику

$$F_3 = \frac{\frac{S_2 - S_1}{N-1}}{\frac{S_1}{NT-N-1}};$$

которая при гипотезе H_3 имеет распределение $F_{N-1, NT-N-1}$.

ПРИМЕР 3.1.2 (продолжение)

Продолжим рассмотрение данных об инвестициях и прибыли 3 предприятий. Выше в рамках модели M_0 мы проверили гипотезу H_1 : « β_i одинаковы для всех i » и не отвергли ее. В рамках той же модели M_0 проверим гипотезу H_2 : « β_i и α_i одинаковы для всех i ». Эта гипотеза отвергается (табл. 3.18).

Таблица 3.18

Проверка гипотезы H_2

Wald Test:			
<i>F-statistic</i>	3.595209	<i>Probability</i>	0.019644

В рамках модели $M_1: y_{it} = \alpha_i + \beta x_{it} + u_{it}$ проверим гипотезу $H_3: \alpha_1 = \dots = \alpha_N$. Эта гипотеза также отвергается (табл. 3.19).

Таблица 3.19

Проверка гипотезы H_3

Wald Test:			
<i>F-statistic</i>	6.810977	<i>Probability</i>	0.004183

Полученные результаты говорят в пользу модели M_1 с одинаковыми β_i , но различными α_i . ■

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что понимается под панельными данными?
2. В чем состоит модель пула?
 - Как можно получить эффективные оценки параметров такой модели?
 - Как следует действовать в ситуации, когда ошибки в разных уравнениях имеют разные дисперсии?
 - Как следует действовать в ситуации, когда существует корреляционная связь между ошибками в разных уравнениях в один и тот же момент времени?
 - Как следует действовать в ситуации, когда существует корреляционная связь между ошибками в пределах одного уравнения?
3. В чем состоит модель кажущихся несвязанными регрессий?
 - Как производится оценивание такой модели?
 - Как можно проверить гипотезу статистической независимости ошибок в разных уравнениях?
4. В чем состоит модель ковариационного анализа? Как в рамках этой модели можно проверить гипотезу о совпадении коэффициентов при некоторой переменной во всех уравнениях?

Тема 3.2

МОДЕЛИ С ФИКСИРОВАННЫМИ И СЛУЧАЙНЫМИ ЭФФЕКТАМИ

Фиксированные эффекты

Обратимся теперь к методам анализа панельных данных, предназначенным в основном для анализа данных $\{y_{it}, x_{it}; i = 1, \dots, N, t = 1, \dots, T\}$, в которых количество субъектов исследования N велико, а количество наблюдений T над каждым субъектом мало. Вследствие малости T в таких ситуациях затруднительно использовать технику, интерпретирующую $y_{1t}, y_{2t}, \dots, y_{NT}$ как N временных рядов длины T (например, технику векторных авторегрессий и моделей коррекции ошибок для нестационарных временных рядов). Основная направленность методов, предполагающих малость T , — получение по возможности наиболее эффективных оценок коэффициентов.

Сначала сфокусируем внимание на модели, соответствующей гипотезе H_1 со скалярной объясняющей переменной x :

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

т.е.

$$y_{it} = \sum_{j=1}^N \alpha_i d_{ij} + \beta x_{it} + u_{it},$$

где $d_{ij} = 1$, если $i = j$, и $d_{ij} = 0$ в противном случае, так что в правой части имеем N дамми-переменных. Здесь α_i трактуются как неизвестные фиксированные параметры (**фиксированные эффекты** — *fixed effects*). Как и ранее, будем предполагать¹, что в этой модели

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

и что

$$E(x_{it} u_{js}) = 0 \quad \text{для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

так что x является экзогенной переменной.

Альтернативные названия этой модели:

- 1) **OLS-дамми-модель** (*LSDV* — *least squares dummy variables*);
- 2) **модель с фиксированными эффектами** (*FE* — *fixed effects*);
- 3) **модель ковариационного анализа** (*CV* — *covariance analysis*).

В этой модели оценка наименьших квадратов, как было отмечено выше, имеет вид:

$$\hat{\beta}_{CV} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2},$$

при этом

$$D(\hat{\beta}_{CV}) = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2}.$$

Альтернативные названия этой оценки:

- 1) **«внутригрупповая» оценка** («внутри»-оценка — *within-estimator*);
- 2) **оценка фиксированных эффектов** (*FE-оценка*);
- 3) **ковариационная оценка**.

¹ Если интересоваться только асимптотическими свойствами оценок, то здесь и далее предположение о нормальности распределения ошибок является излишним: достаточно предполагать, что $u_{it} \sim i.i.d.$ с $E(u_{it}) = 0$ и $D(u_{it}) = \sigma_u^2$, $0 < \sigma_u^2 < \infty$ (см., например, (Hsiao, 2003)).

Часто для этой оценки используют также обозначения $\hat{\beta}_W$ (индекс W — от *within*) и $\hat{\beta}_{FE}$. Как было отмечено выше, эта оценка имеет одно и то же значение при двух альтернативных методах ее получения: в рамках статистической модели $y_{it} = \sum_{j=1}^N \alpha_i d_{ij} + \beta x_{it} + u_{it}$ с дамми-переменными и в рамках модели в отклонениях от групповых средних $y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i)$, $i = 1, \dots, N$, $t = 1, \dots, T$. Однако если количество субъектов анализа N велико, то в первой модели приходится обращаться матрицу весьма большого порядка $(N + 1)$, тогда как во второй модели такая проблема не возникает.

Оценки для фиксированных эффектов вычисляются по формуле:

$$\hat{\alpha}_i = \bar{y}_i - \hat{\beta} \bar{x}_i, \quad i = 1, \dots, N.$$

При сделанных предположениях $\hat{\beta}_{CV}$ является **наилучшей линейной несмещенной оценкой** (*BLUE* — *best linear unbiased estimate*) для коэффициента β ,

$$p \lim_{T \rightarrow \infty} \hat{\beta}_{CV} = \beta, \quad p \lim_{N \rightarrow \infty} \hat{\beta}_{CV} = \beta, \quad p \lim_{T \rightarrow \infty} \hat{\alpha}_i = \alpha_i,$$

но $p \lim_{N \rightarrow \infty} \hat{\alpha}_i \neq \alpha_i$, хотя $E(\hat{\alpha}_i) = \alpha_i$.

Таким образом, $\hat{\beta}_{CV}$ является состоятельной оценкой и когда $N \rightarrow \infty$, и когда $T \rightarrow \infty$, в то время как $\hat{\alpha}_i$ состоятельна только тогда, когда $T \rightarrow \infty$. Последнее есть следствие того, что оценивание каждого α_i производится фактически лишь по T наблюдениям, так что при фиксированном T с ростом N происходит лишь увеличение количества параметров α_i , но это не приводит к повышению точности оценивания каждого конкретного α_i .

Заметим, что если нас интересует только состоятельность оценки $\hat{\beta}_{CV}$, но не ее эффективность (т.е. свойство *BLUE*), то для этого не требуется строгая экзогенность x (т.е. не требуется, чтобы $E(x_{it} u_{js}) = 0$ для любых $i, j = 1, \dots, N$, $t, s = 1, \dots, T$). В этом случае достаточно выполнения соотношений $E(x_{it} u_{is}) = 0$ для любых $t, s = 1, \dots, T$ и $i = 1, \dots, N$ (т.е. требуется лишь экзогенность x в рамках каждого отдельного субъекта исследования).

В модели с фиксированными эффектами полученные выводы являются условными по отношению к значениям эффектов α_i в выборке. Такая интерпретация наиболее подходит для случаев, когда субъектами исследования выступают страны, крупные компании или предприятия, т.е. каждый субъект «имеет свое лицо».

Сами эффекты α_i , по существу, отражают наличие у субъектов исследования некоторых индивидуальных характеристик, не изменяющихся со временем в процессе наблюдений, которые трудно или даже невозможно наблю-

дать или измерить. Если значения таких характеристик не наблюдаются, то эти характеристики невозможно непосредственно включить в правые части уравнений регрессии в качестве объясняющих переменных. Но тогда мы имеем дело с «пропущенными переменными» — с ситуацией, которая может приводить к смещению оценок наименьших квадратов. Чтобы исключить такое смещение, в правые части уравнений вместо значений ненаблюдаемых индивидуальных характеристик как раз и вводятся ненаблюдаемые эффекты α_i . Проиллюстрируем возникновение указанного смещения на следующем примере.

ПРИМЕР 3.2.1

На рис. 3.4 представлено облако рассеяния точек (x_{it}, y_{it}) , порожденных моделью

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, 2, \quad t = 1, \dots, 100,$$

в которой $\alpha_1 = 150$, $\alpha_2 = 250$, $\beta = 0.6$, $u_{it} \sim i.i.d. N(0, 10^2)$. Значения x_{it} заданы (неслучайны); при $i = 1$ значения x_{1t} меньше 150, а при $i = 2$ значения x_{2t} больше 150.

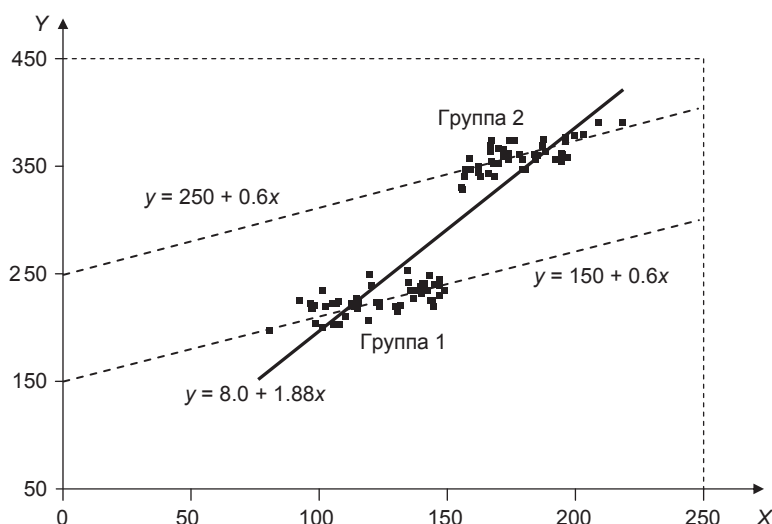


Рис. 3.4

Облако точек распадается на две группы точек: в группе 1 объединяются точки, соответствующие $i = 1$, а в группе 2 — точки, соответствующие $i = 2$. Точки группы 1 располагаются вдоль (теоретической) прямой $y = 150 + 0.6x$ (нижняя пунктирная линия на графике), точки группы 2 — вдоль (теоретической) прямой $y = 250 + 0.6x$.

Если по имеющимся 100 наблюдениям оценивать статистическую модель

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \quad i = 1, 2, \quad t = 1, \dots, 100, \text{ (пул),}$$

не принимающую во внимание возможное наличие индивидуальных эффектов, то оцененная модель принимает вид:

$$\hat{y}_{it} = 8.00 + 1.88x_{it}.$$

Таким образом, оценка коэффициента β оказывается завышенной втрое по сравнению со значением, использованным при моделировании.

В то же время, если перейти от переменных x_{it} , y_{it} к отклонениям от средних значений в группах $(x_{it} - \bar{x}_i)$ и $(y_{it} - \bar{y}_i)$, то для новых переменных облако рассеяния (рис. 3.5) концентрируется вокруг начала координат (соответствующие точки изображены черными квадратами) и вытянуто в правильном направлении.

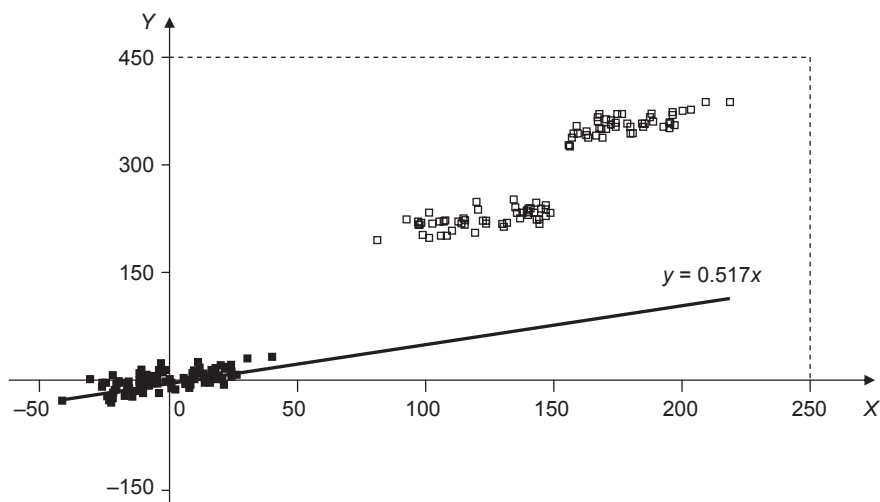


Рис. 3.5

Оцененная модель в отклонениях от средних в группах имеет вид:

$$\hat{y}_{it} = 0.517x_{it},$$

и на этот раз оценка коэффициента β оказывается близкой к значению $\beta = 0.6$, использованному при моделировании.

Если оценивать модель с дамми-переменными

$$y_{it} = \alpha_1 d_{i1} + \alpha_2 d_{i2} + \beta x_{it} + u_{it},$$

где $d_{ij} = 1$, если $j = i$, и $d_{ij} = 0$ в противном случае, то получим результаты оценивания, приведенные в табл. 3.20.

Таблица 3.20

Объясняемая переменная Y

Переменная	Коэффициент	Стандартная ошибка	t -статистика	P -значение
$D1$	161.4615	7.260153	22.23940	0.0000
$D2$	264.8593	10.46430	25.31074	0.0000
X	0.517319	0.058229	8.884227	0.0000

Полученные оценки $\hat{\alpha}_1 = 161.46$, $\hat{\alpha}_2 = 264.86$, $\hat{\beta} = 0.517$ близки к значениям параметров, использованным при моделировании. ■

Случайные эффекты

Запишем модель

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

соответствующую гипотезе H_1 , в равносильном виде:

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it},$$

где $\sum_{i=1}^N \alpha_i = 0$.

При таком условии α_i называют **дифференциальными эффектами** (*differential effects*). В ряде ситуаций N субъектов, для которых имеются статистические данные, рассматриваются как случайная выборка из некоторой более широкой совокупности (популяции), и исследователя интересуют не конкретные субъекты, попавшие в выборку, а обезличенные субъекты, имеющие заданные характеристики. Соответственно в таких ситуациях предполагается, что α_i являются случайными величинами, и тогда речь идет о модели

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it}$$

как о **модели со случайными эффектами** (*random effects model*). В такой модели α_i уже не интерпретируются как значения некоторых фиксированных параметров и не подлежат оцениванию. Вместо этого оцениваются параметры распределения случайных величин α_i .

Обозначив $v_{it} = \alpha_i + u_{it}$, получим другую запись этой модели:

$$y_{it} = \mu + \beta x_{it} + (\alpha_i + u_{it}) = \mu + \beta x_{it} + v_{it}.$$

В такой форме модели ошибка v_{it} состоит из двух компонент — α_i и u_{it} . Как и в модели с фиксированными эффектами, случайные эффекты α_i также

отражают наличие у субъектов исследования некоторых индивидуальных характеристик, не изменяющихся со временем в процессе наблюдений, которые трудно или даже невозможно наблюдать или измерить. Однако теперь значения этих характеристик встраиваются в состав случайной ошибки, как это делается в классической модели регрессии, в которой наличие случайных ошибок интерпретируется как недостаточность включенных в модель объясняющих переменных для полной интерпретации изменений объясняемой переменной.

К прежним предположениям о том, что

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

и что

$$E(x_{it} u_{js}) = 0 \quad \text{для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

добавим следующие предположения:

$$E(\alpha_i) = 0 \quad (\text{так что и } E(v_{it}) = 0),$$

$$E(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2, & \text{если } i = j, \\ 0, & \text{если } i \neq j \end{cases}$$

(так что последовательность значений $\alpha_1, \dots, \alpha_N$ представляет *случайную выборку* из распределения $N(0, \sigma_\alpha^2)$),

$$E(x_{it} \alpha_j) = 0, \quad i, j = 1, \dots, N, \quad t = 1, \dots, T$$

(так что $E(x_{it} v_{js}) = 0$, и в модели со случайными ошибками v_{it} переменная x является экзогенной).

Если предположить еще, что

$$E(u_{it} \alpha_i) = 0,$$

то условная относительно x_{it} дисперсия случайной величины y_{it} равна:

$$D(y_{it} | x_{it}) = D(v_{it} | x_{it}) = D(v_{it}) = D(\alpha_i + u_{it}) = \sigma_\alpha^2 + \sigma_u^2.$$

Таким образом, дисперсия y_{it} складывается из двух некоррелированных компонент — их называют **компонентами дисперсии** (*variance components*). Саму модель называют

- **моделью компонент дисперсии** (*variance components model*);
- **стандартной моделью со случайными эффектами** (**RE-модель** — *random effects model*)

В векторной форме эта модель имеет вид:

$$y_i = [ex_i] \delta + v_i,$$

где

$$y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix}, \quad e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, \quad x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iT} \end{pmatrix}, \quad \delta = \begin{pmatrix} \mu \\ \beta \end{pmatrix}, \quad v_i = \begin{pmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iT} \end{pmatrix}.$$

Заметим, что

$$E(v_{it} v_{is}) = E(\alpha_i + u_{it}, \alpha_i + u_{is}) = \begin{cases} \sigma_\alpha^2 + \sigma_u^2, & \text{если } t = s, \\ \sigma_\alpha^2, & \text{если } t \neq s, \end{cases}$$

так что случайные величины v_{it} и v_{is} коррелированы, даже если не коррелированы ошибки u_{it} , и ковариационная матрица случайного вектора v_i имеет вид:

$$V = E(v_i v_i^T) = \sigma_u^2 I_T + \sigma_\alpha^2 e e^T.$$

Например, при $T = 3$

$$V = \begin{bmatrix} \sigma_u^2 + \sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_u^2 + \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \sigma_u^2 + \sigma_\alpha^2 \end{bmatrix}.$$

При этом

$$\text{Corr}(v_{it}, v_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} = \rho \quad \text{для всех } t \neq s$$

(предположение равной коррелированности в модели компонент дисперсии).

Оценивание. В *RE*-модели оценка

$$\hat{\beta}_{CV} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2}$$

остается несмещенной и состоятельной оценкой для β . Однако она перестает быть эффективной оценкой (*BLUE*), как это было в модели с фиксированными эффектами, поскольку не учитывает коррелированность v_{it} во времени для субъекта i .

Можно ожидать, что обобщенная оценка наименьших квадратов (*GLS*-оценка), учитывающая такую коррелированность, будет более эффективной. Заметим, что *GLS*-оценка для δ имеет вид:

$$\hat{\delta}_{GLS} = \left[\sum_{i=1}^N \begin{bmatrix} e^T \\ x_i^T \end{bmatrix} V^{-1} \begin{bmatrix} e x_i \end{bmatrix} \right]^{-1} \left[\sum_{i=1}^N \begin{bmatrix} e^T \\ x_i^T \end{bmatrix} V^{-1} y_i \right]$$

и что

$$V^{-1} = \frac{1}{\sigma_u^2} \left[I_T - \frac{ee^T}{T} + \Psi \frac{ee^T}{T} \right],$$

где $\Psi = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}$.

Чтобы не возникало путаницы с другими *GLS*-оценками, для *GLS*-оценки в стандартной модели со случайными эффектами используется также обозначение $\hat{\beta}_{RE}$ (**RE-оценка** — *random effect estimator*). Заметим, что диагональные элементы матрицы V^{-1} равны $\frac{1 - \frac{1 - \Psi}{T}}{\sigma_u^2}$, а недиагональные элементы

равны $-\frac{1 - \Psi}{T\sigma_u^2}$.

Практически $\hat{\beta}_{RE}$ можно получить следующим образом. Усредняя по t обе части уравнения

$$y_{it} = \mu + \beta x_{it} + v_{it},$$

получаем соотношение

$$\bar{y}_i = \mu + \beta \bar{x}_i + \bar{v}_i.$$

Обозначив $\theta = 1 - \sqrt{\Psi}$, произведем преобразование:

$$y_{it}^* = y_{it} - \theta \bar{y}_i, \quad x_{it}^* = x_{it} - \theta \bar{x}_i, \quad v_{it}^* = v_{it} - \theta \bar{v}_i, \quad \mu^* = (1 - \theta)\mu.$$

В результате получим преобразованную модель

$$y_{it}^* = \mu^* + \beta x_{it}^* + v_{it}^*$$

с экзогенной переменной x_{it}^* , в которой ковариационная матрица вектора ошибок v_{it}^* имеет вид:

$$\text{Cov}(v_{it}^*) = \sigma_u^2 I_T.$$

Поэтому применение *OLS* к преобразованной модели дает *BLUE*-оценку:

$$\hat{\beta}_{GLS} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it}^* - \bar{x}^*)(y_{it}^* - \bar{y}^*)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it}^* - \bar{x}^*)^2},$$

где

$$\bar{y}^* = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it}^*, \quad \bar{x}^* = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{it}^*.$$

Это выражение можно представить в виде:

$$\hat{\beta}_{GLS} = w\hat{\beta}_b + (1-w)\hat{\beta}_{CV},$$

где

$$\hat{\beta}_b = \frac{\sum_{i=1}^T (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})}{\sum_{i=1}^T (\bar{x}_i - \bar{x})^2},$$

$\hat{\beta}_b$ — «межгрупповая» оценка («между»-оценка — *between estimator*), соответствующая регрессии средних значений \bar{y}_i на константу и средние значения \bar{x}_i , т.е.

$$\bar{y}_i = \mu + \beta \bar{x}_i + \bar{v}_i$$

(«модель для групповых средних»), и игнорирующая внутригрупповую изменчивость,

$$w = \frac{\Psi \sum_{i=1}^N (\bar{x}_i - \bar{x})^2}{\sum_{i=1}^N \sum_{t=1}^T (\bar{x}_{it} - \bar{x}_i)^2 + \Psi \sum_{i=1}^N (\bar{x}_i - \bar{x})^2}.$$

Таким образом, обобщенная оценка наименьших квадратов $\hat{\beta}_{GLS}$ в *RE*-модели учитывает и внутригрупповую, и межгрупповую изменчивость. Она является взвешенным средним «межгрупповой» оценки $\hat{\beta}_b$ (учитывающей только межгрупповую изменчивость) и «внутригрупповой» оценки $\hat{\beta}_{CV}$ (учитывающей только внутригрупповую изменчивость), а w измеряет вес, придаваемый межгрупповой изменчивости. При сделанных предположениях обе оценки — $\hat{\beta}_b$ и $\hat{\beta}_{CV}$ — состоятельны, следовательно, состоятельна и сама $\hat{\beta}_{GLS}$.

Если $T \rightarrow \infty$, то $\Psi \rightarrow 0$, $w \rightarrow 0$ и $\hat{\beta}_{GLS} \rightarrow \hat{\beta}_{CV}$, так что при больших T оценки для β , получаемые в рамках моделей фиксированных и случайных эффектов, эквивалентны.

Если $\sigma_\alpha^2 \rightarrow 0$, то $\Psi \rightarrow 1$ и $V = E(v_i v_i^T) = \sigma_u^2 I_T + \sigma_\alpha^2 e e^T \rightarrow \sigma_u^2 I_T$. Соответственно при этом *GLS*-оценка переходит в *OLS*-оценку, т.е.

$$\hat{\beta}_{GLS} \rightarrow \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2} = \hat{\beta}_{OLS}$$

(в пределе нет никаких эффектов).

Заметим далее, что

$$D(\hat{\beta}_{GLS}) = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 + \Psi \sum_{i=1}^N (\bar{x}_i - \bar{x})^2}.$$

В то же время

$$D(\hat{\beta}_{CV}) = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2}.$$

Поскольку $\Psi > 0$, из двух последних соотношений следует, что

$$D(\hat{\beta}_{GLS}) < D(\hat{\beta}_{CV}),$$

т.е. *GLS*-оценка эффективнее. Она эффективнее оценки $\hat{\beta}_{CV}$ именно потому, что использует информацию как о внутригрупповой изменчивости, так и о межгрупповой изменчивости.

Чтобы реализовать эту *GLS*, т.е. получить **доступную *GLS*-оценку** (*FGLS* — *feasible GLS*, или *EGLS* — *estimated GLS*), надо подставить в выражения для Ψ (и θ) подходящие оценки для σ_u^2 и $\sigma_u^2 + T\sigma_\alpha^2$.

Оценить σ_u^2 можно, используя **внутригрупповые остатки** (*within residuals*)

$$(y_{it} - \bar{y}_i) - \hat{\beta}_{CV}(x_{it} - \bar{x}_i),$$

которые получены при оценивании модели, скорректированной на индивидуальные средние:

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T [(y_{it} - \bar{y}_i) - \hat{\beta}_{CV}(x_{it} - \bar{x}_i)]^2}{N(T-1) - 1}$$

(в этой модели эффективное количество наблюдений равно $N(T-1)$ из-за наличия N линейных связей между уравнениями).

Оценить дисперсию σ_α^2 случайных эффектов $\sigma_\alpha^2 = D(\alpha_i)$ можно, заметив, что при оценивании модели $\bar{y}_i = \mu + \beta \bar{x}_i + \bar{v}_i$, приводящей к «межгрупповой» оценке

$$\hat{\beta}_b = \frac{\sum_{i=1}^T (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})}{\sum_{i=1}^T (\bar{x}_i - \bar{x})^2},$$

дисперсия остатка для i -й группы равна:

$$D(\bar{y}_i - \hat{\mu}_b - \hat{\beta}_b \bar{x}_i) = \frac{\sigma_u^2}{T} + \sigma_\alpha^2.$$

Состоятельной оценкой для $\frac{\sigma_u^2}{T} + \sigma_\alpha^2$ является

$$\frac{\sum_{i=1}^N (\bar{y}_i - \hat{\mu}_b - \hat{\beta}_b \bar{x}_i)^2}{N - 2}.$$

Поэтому состоятельной оценкой для σ_α^2 служит

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^N (\bar{y}_i - \hat{\mu}_b - \hat{\beta}_b \bar{x}_i)^2}{N - 2} - \frac{\hat{\sigma}_u^2}{T},$$

а

$$\hat{\sigma}_u^2 + T\hat{\sigma}_\alpha^2 = T \frac{\sum_{i=1}^N (\bar{y}_i - \hat{\mu}_b - \hat{\beta}_b \bar{x}_i)^2}{N - 2}$$

является состоятельной оценкой для $\sigma_u^2 + T\sigma_\alpha^2$. Эти две оценки используют межгрупповые остатки. Они являются также оценками максимального правдоподобия соответствующих дисперсий.

Следует отметить, что — особенно при небольших значениях N и T — значение вычисленной указанным образом оценки дисперсии σ_α^2 может оказаться отрицательным.

Как было отмечено выше, $\hat{\beta}_{GLS}$ можно представить в виде:

$$\hat{\beta}_{GLS} = w\hat{\beta}_b + (1 - w)\hat{\beta}_{CV},$$

так что $\hat{\beta}_{GLS}$ является линейной комбинацией «внутри»-оценки и «между»-оценки. Эта линейная комбинация оптимальна. Поэтому, например, оценка $\hat{\beta}_{OLS}$, также являющаяся линейной комбинацией этих двух оценок (при $\Psi = 1$), хотя и состоятельна, но менее эффективна.

Критерий Бройша — Пагана для индивидуальных эффектов

Данный критерий используется для проверки в рамках RE -модели (со стандартными предположениями) гипотезы

$$H_0: \sigma_\alpha^2 = 0 \text{ (сведение к модели пула).}$$

Идея критерия основана на тождестве

$$\sum_{i=1}^N \left(\sum_{t=1}^T u_{it} \right)^2 = \sum_{i=1}^N \sum_{t=1}^T u_{it}^2 + \sum_{i=1}^N \sum_{s \neq t} u_{is} u_{it},$$

из которого следует, что

$$\frac{\sum_{i=1}^N \left(\sum_{t=1}^T u_{it} \right)^2}{\sum_{i=1}^N \sum_{t=1}^T u_{it}^2} - 1 = \frac{\sum_{i=1}^N \sum_{s \neq t} u_{is} u_{it}}{\sum_{i=1}^N \sum_{t=1}^T u_{it}^2}.$$

При отсутствии автокоррелированности случайных ошибок u_{it} правая часть последнего равенства мала. Поэтому статистику критерия можно основывать на выражении, стоящем в левой части, в которое вместо ненаблюдаемых значений u_{it} подставляются остатки \hat{u}_{it} , полученные при OLS -оценивании модели пула. Против гипотезы H_0 говорят «слишком большие» значения

$$\left[\frac{\sum_{i=1}^N \left(\sum_{t=1}^T \hat{u}_{it} \right)^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right]^2.$$

Статистика критерия Бройша — Пагана

$$BP = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N \left(\sum_{t=1}^T \hat{u}_{it} \right)^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right]^2 = \frac{NT}{2(T-1)} \left[1 - \frac{T^2 \sum_{i=1}^N \bar{\hat{u}}_i^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} \right]^2$$

при гипотезе H_0 имеет асимптотическое распределение $\chi^2(1)$. Соответственно, гипотеза H_0 отвергается, если наблюдаемое значение статистики BP превышает критическое значение, рассчитанное по распределению $\chi^2(1)$.

Коэффициенты детерминации, разложение полной суммы квадратов

При анализе панельных данных возникают некоторые проблемы с определением коэффициента детерминации R^2 . Так, во многих руководствах по эконометрике и монографиях, специально посвященных анализу панельных данных, вообще не упоминается о коэффициенте детерминации. В то же время в некоторых пакетах статистического анализа предусмотрено вычисление коэффициентов детерминации и для панельных данных.

Проблема с определением коэффициента детерминации в случае панельных данных связана с неопределенностью в отношении того, что считать полной суммой квадратов, подлежащей разложению на объясненную регрессией и остаточную суммы квадратов. Здесь имеем соотношение:

$$\frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y})^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)^2 + \frac{1}{N} \sum_{i=1}^N (\bar{y}_i - \bar{y})^2,$$

и в качестве полной суммы квадратов может использоваться каждая из трех сумм квадратов, входящих в это выражение. Соответствующие этим полным суммам регрессионные модели объясняют:

- отклонения наблюдаемых значений y_{it} от их среднего по всем NT наблюдениям;
- отклонения наблюдаемых значений y_{it} в группах от их средних по группе;
- отклонения средних по группам от среднего по всем NT наблюдениям.

Если используем оценку пул, то она получается в результате применения метода наименьших квадратов к уравнению

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

При этом коэффициент детерминации равен квадрату (выборочного) коэффициента корреляции между переменными y_{it} и

$$\hat{y}_{it} = \hat{\alpha} + \hat{\beta}_{OLS} x_{it},$$

где $\hat{\beta}_{OLS}$ — OLS-оценка коэффициента β в модели пула.

Об этом коэффициенте детерминации говорят как о **R^2 -полном** (R^2 -overall):

$$R_{overall}^2 = \text{corr}^2(y_{it}, \hat{\alpha} + \hat{\beta}_{OLS} x_{it}) = \text{corr}^2(y_{it}, \hat{\beta}_{OLS} x_{it}).$$

Если используем оценку «между», то она получается в результате применения метода наименьших квадратов к уравнению

$$\bar{y}_i = \mu + \beta \bar{x}_i + \bar{u}_i, \quad i = 1, \dots, N.$$

При этом коэффициент детерминации равен квадрату (выборочного) коэффициента корреляции между переменными \bar{y}_i и

$$\hat{\bar{y}}_i = \hat{\mu} + \hat{\beta}_b \bar{x}_i,$$

где $\hat{\beta}_b$ — «между»-оценка для коэффициента β .

Об этом коэффициенте детерминации говорят как о **R^2 -между** (R^2 -between):

$$R^2_{\text{between}} = \text{corr}^2(\hat{\mu} + \hat{\beta}_b \bar{x}_i, \bar{y}_i) = \text{corr}^2(\hat{\beta}_b \bar{x}_i, \bar{y}_i).$$

Если используем оценку «внутри», то она получается в результате применения метода наименьших квадратов к уравнению

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

В правой части последнего уравнению отсутствует константа. А при *OLS*-оценивании уравнений вида $z_i = \beta w_i + v_i$ коэффициент детерминации в общем случае не равен квадрату выборочного коэффициента корреляции между переменными $\hat{z}_i = \hat{\beta} w_i$ и z_i . Однако если переменные z_i и w_i центрированы, так что $\bar{z} = \bar{w} = 0$, то такое равенство обеспечивается. В нашем случае переменные $y_{it} - \bar{y}_i$ и $x_{it} - \bar{x}_i$ центрированы, так что коэффициент детерминации, получаемый при оценивании уравнения в отклонениях от средних по группам равен квадрату (выборочного) коэффициента корреляции между переменными $\tilde{y}_{it} = y_{it} - \bar{y}_i$ и

$$\hat{\tilde{y}}_{it} = \hat{\beta}_{CV} (x_{it} - \bar{x}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где $\hat{\beta}_{CV}$ — «внутри»-оценка для коэффициента β .

Об этом коэффициенте детерминации говорят как о **R^2 -внутри** (R^2 -within):

$$R^2_{\text{within}} = \text{corr}^2(\hat{\beta}_{CV} (x_{it} - \bar{x}_i), y_{it} - \bar{y}_i).$$

Каждый из этих 3 вариантов R^2 является обычным коэффициентом детерминации в соответствующей модели регрессии. В то же время при анализе различных моделей панельных данных часто сообщаются вычисленные значения всех 3 вариантов R^2 , несмотря на то что в модели с фиксированными эффектами используется оценка $\hat{\beta}_{CV}$, в модели со случайными эффектами — оценка $\hat{\beta}_{GLS}$, а в модели пула — оценка $\hat{\beta}_{OLS}$.

Более точно, при анализе панельных данных под названиями R^2_{within} , R^2_{between} , R^2_{overall} принято сообщать значения:

$$R_{within}^2 = \text{corr}^2(y_{it} - \bar{y}_i, \hat{\beta}(x_{it} - \bar{x}_i)),$$

$$R_{between}^2 = \text{corr}^2(\bar{y}_i, \hat{\beta} \bar{x}_i),$$

$$R_{overall}^2 = \text{corr}^2(y_{it}, \hat{\beta} x_{it})$$

независимо от того, каким образом была получена оценка $\hat{\beta}$. Если β является p -мерным вектором, то соответственно:

$$R_{within}^2 = \text{corr}^2(y_{it} - \bar{y}_i, \hat{\beta}(x_{it} - \bar{x}_i)^T),$$

$$R_{between}^2 = \text{corr}^2(\bar{y}_i, \hat{\beta} \bar{x}_i^T),$$

$$R_{overall}^2 = \text{corr}^2(y_{it}, \hat{\beta} x_{it}^T).$$

При этом:

- приводимое значение R_{within}^2 является коэффициентом детерминации в обычном смысле, если $\hat{\beta} = \hat{\beta}_{within}$;
- приводимое значение $R_{between}^2$ является коэффициентом детерминации в обычном смысле, если $\hat{\beta} = \hat{\beta}_{between}$;
- приводимое значение $R_{overall}^2$ является коэффициентом детерминации в обычном смысле, если $\hat{\beta} = \hat{\beta}_{OLS}$.

ПРИМЕР 3.2.2

Приведем результаты оценивания в пакете *Stata8* моделей с фиксированными (табл. 3.21, 3.22) и случайными (табл. 3.23) эффектами для данных о 3 предприятиях. При этом заметим, что в рамках этого пакета приняты обозначения, отличающиеся от используемых нами: индивидуальные эффекты обозначаются как u_i , а случайные ошибки — как e_{it} . Чтобы избежать путаницы, в приводимые далее протоколы оценивания внесены соответствующие изменения.

Таблица 3.21

Fixed-effects (within) regression

R-SQ:	
<i>within</i>	= 0.9478
<i>between</i>	= 0.8567
<i>overall</i>	= 0.9209

Проверка значимости регрессии в целом:

$$F(1,26) = 472.26, \quad \text{Prob} > F = 0.0000.$$

Оцененное значение коэффициента корреляции между индивидуальным эффектом и предсказанием $Corr(\alpha_i, Xb) = -0.2311$. Остальная часть протокола оценивания приведена в табл. 3.22.

Таблица 3.22

Оценивание модели с фиксированными эффектами для 3 предприятий

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение <i>P</i> > <i>t</i>
<i>x</i>	1.102192	0.0507186	21.73	0.000
<i>cons</i>	−1.394325	0.8230266	−1.69	0.102
<i>sigma_alfa</i>	1.480319			
<i>sigma_u</i>	1.745136			
<i>rho</i>	0.4184474 (<i>fraction of variance due to α_i</i>)			
<i>F</i> test that all <i>A_i</i> = 0:				
<i>F</i> (2, 26) = 6.81, <i>Prob</i> > <i>F</i> = 0.0042				

Критерий в последней строке соответствует гипотезе с двумя линейными ограничениями: поскольку в модель включена постоянная составляющая, одно линейное ограничение накладывается заранее как идентифицирующее и не подлежащее проверке.

Таблица 3.23

Random-effects GLS regression

R^2 :	
$within$	= 0.9478
$between$	= 0.8567
$overall$	= 0.9209
Random effects: $u_i \sim \text{Gaussian}$	
$corr(\alpha_i, X) = 0$ (предполагается)	
Критерий значимости регрессии в целом:	
$Wald \chi^2(1) = 325.94, \quad Prob > \chi^2 = 0.0000$	
σ_{α}	0
σ_u	1.7451362
ρ	0 (fraction of variance due to α_i)

Здесь полученная оценка для σ_{α}^2 оказалась отрицательной, поэтому ее значение полагается равным нулю. Однако тогда модель со случайными эффектами редуцируется к модели пула (табл. 3.24).

Таблица 3.24

Оценивание модели пула для 3 предприятий

у	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z$
x	1.0589590	0.586557	18.05	0.000
cons	-0.7474755	0.955953	-0.78	0.434

В то же время если в рамках модели со случайными эффектами применить критерий Бройша — Пагана для проверки гипотезы об отсутствии таковых эффектов, т.е. гипотезы $H_0: \sigma_\alpha^2 = 0$, то полученное значение статистики критерия равно 8.47. Этому значению соответствует рассчитанное по асимптотическому распределению хи-квадрат с 1 степенью свободы P-значение 0.0036. Но это говорит против гипотезы H_0 . И опять это можно объяснить малым количеством наблюдений — ведь распределение хи-квадрат здесь только асимптотическое.

В пакете Stata 8 есть возможность оценить модель со случайными эффектами, не прибегая к GLS-оцениванию, а используя метод максимального правдоподобия. Это дает результаты, указанные в табл. 3.25. По критерию отношения правдоподобий гипотеза $H_0: \sigma_\alpha^2 = 0$ отвергается.

Таблица 3.25

Оценивание модели с помощью метода максимального правдоподобия

Random-effects ML regression; Random effects: $\alpha_i \sim \text{Gaussian}$; Log likelihood = -61.09837;
Критерий значимости регрессии в целом: LR $\chi^2(1) = 121.60$, Prob > $\chi^2 = 0.0000$

у	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z$
x	1.092893	0.0501518	21.79	0.000
cons	-1.255205	1.019264	-1.23	0.218
sigma_α	1.064836	0.5552752	1.92	0.055
sigma_u	1.713621	0.2334960	7.34	0.000
rho	0.2785682	0.2205921		
Likelihood-ratio test of sigma_alfa = 0:				
chibar2(01) = 4.70, Prob > chibar2 = 0.015				

При оценивании *between*-регрессии (табл. 3.26) она оказывается статистически незначимой, а близкое к 1 значение коэффициента детерминации R^2_{between} не должно вводить в заблуждение: для оценивания двух коэффициентов имеется всего 3 наблюдения. ■

Таблица 3.26

Between-регрессия (регрессия на групповые средние)

<i>R-SQ</i> :				
<i>within</i> = 0.9478				
<i>between</i> = 0.8567				
<i>overall</i> = 0.9209				
Проверка значимости регрессии в целом:				
<i>F</i> (1, 1) = 5.98, <i>Prob</i> > <i>F</i> = 0.2471				
<i>y</i>	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение <i>P</i> > <i>t</i>
<i>x</i>	0.3137715	0.1283133	2.45	0.247
<i>cons</i>	10.40202	1.9296160	5.39	0.117

Выбор между моделями с фиксированными или случайными эффектами

Прежде всего напомним отмеченные ранее особенности моделей с фиксированными или случайными эффектами.

FE: получаемые выводы — *условные* по отношению к значениям эффектов α_i в выборке. Это соответствует ситуациям, когда эти значения нельзя рассматривать как случайную выборку из некоторой более широкой совокупности (популяции). Такая интерпретация наиболее подходит для случаев, когда субъектами исследования являются страны, крупные компании или предприятия, т.е. каждый субъект «имеет свое лицо».

RE: получаемые выводы — *безусловные* относительно совокупности всех эффектов α_i . Исследователя не интересуют конкретные субъекты в выборке — для него это обезличенные субъекты, выбранные случайным образом из более широкой совокупности (так что набор $\alpha_1, \alpha_2, \dots, \alpha_N$, трактуется как случайная выборка из некоторого распределения).

Заметим в связи с этим, что:

$$\text{в FE-модели } E(y_{it} | x_{it}) = E(\alpha_i + \beta x_{it} + u_{it} | x_{it}) = \alpha_i + \beta x_{it};$$

$$\text{в RE-модели } E(y_{it} | x_{it}) = E(\mu + \alpha_i + \beta x_{it} + u_{it} | x_{it}) = \mu + \beta x_{it}.$$

Напомним, что RE-модель предполагает, в частности, что $E(\alpha_i | x_{it}) = 0$. Чтобы избавиться от этого условия ортогональности, предположим, что

$$\alpha_i = a\bar{x}_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2), \quad E(\varepsilon_i | \varepsilon_{it}) = 0.$$

Это приводит к модели Мундлака (*Mundlak model*):

$$y_{it} = \mu + \beta x_{it} + a\bar{x}_i + \varepsilon_i + u_{it},$$

она также является моделью компонент ошибки, но отличается от предыдущей модели тем, что в правую часть добавляется переменная \bar{x}_i , которая изменяется только от субъекта к субъекту и отражает неоднородность субъектов. Эта переменная в отличие от α_i наблюдаема. Заметим, что в модели Мундлака

$$\begin{aligned} E(\alpha_i x_{it}) &= E((a\bar{x}_i + \varepsilon_i)x_{it}) = aE(\bar{x}_i x_{it}) = \\ &= \frac{a}{T} \sum_{s=1}^T E(x_{is}x_{it}) = \frac{a}{T} \left(E(x_{it}^2) + \sum_{s=1, s \neq t}^T E(x_{is}x_{it}) \right), \end{aligned}$$

так что если $a \neq 0$, то условие $E(\alpha_i x_{it}) = 0$ в общем случае не выполняется ни для одного $i = 1, \dots, N$.

Применение *GLS* к этой модели дает *BLUE*-оценки для β и a :

$$\hat{\beta}_{GLS}^* = \hat{\beta}_{CV}, \quad a_{GLS}^* = \hat{\beta}_b - \hat{\beta}_{CV},$$

и

$$\hat{\mu}_{GLS}^* = \bar{y} - \bar{x}\hat{\beta}_b.$$

Иначе говоря, *BLUE*-оценкой для β в этой модели является ковариационная (внутригрупповая) оценка, и из $E(\hat{a}_{GLS}^*) = a$, $E(\hat{a}_{GLS}^*) = E(\hat{\beta}_b) - E(\hat{\beta}_{CV})$ и $E(\hat{\beta}_{CV}) = \beta$ получаем:

$$E(\hat{\beta}_b) = a + \beta.$$

Как было показано выше, в *RE*-модели (предполагающей выполнение условия $E(\alpha_i x_{it}) = 0$)

$$\hat{\beta}_{RE} = w\hat{\beta}_b + (1-w)\hat{\beta}_{CV}.$$

Если использовать эту же оценку в модели Мундлака, то для нее получим:

$$E(\hat{\beta}_{RE}) = wE(\hat{\beta}_b) + (1-w)E(\hat{\beta}_{CV}) = w(\beta + a) + (1-w)\beta = \beta + wa,$$

так что если $a \neq 0$, то $\hat{\beta}_{RE}$ — смещенная оценка.

Критерии спецификации

Речь здесь идет о том, совпадает или нет условное распределение α_i при заданном x_i с безусловным распределением α_i . Если не совпадает — наилучшей оценкой является $\hat{\beta}_{CV}$ (*FE*), если совпадает — наилучшей оценкой является $\hat{\beta}_{GLS}$ (*RE*).

Критерий 1. Используя формулировку Мундлака, проверяем гипотезу $H_0: a = 0$ против альтернативы $H_1: a \neq 0$.

Критерий 2 — критерий Хаусмана (Hausman). Проверяемая гипотеза: $H_0: E(\alpha_i | x_{it}) = 0$, альтернативная гипотеза: $H_1: E(\alpha_i | x_{it}) \neq 0$.

Идея критерия 2 основывается на следующих фактах:

- при гипотезе H_0 и оценка $\hat{\beta}_{GLS}$, соответствующая RE -модели, и оценка $\hat{\beta}_{CV}$, соответствующая FE -модели, состоятельны;
- при гипотезе H_1 оценка $\hat{\beta}_{GLS}$ несостоятельна, а оценка $\hat{\beta}_{CV}$ состоятельна.

Соответственно если гипотеза H_0 верна, то между оценками $\hat{\beta}_{GLS}$ и $\hat{\beta}_{CV}$ не должно наблюдаться систематического расхождения, и эта гипотеза должна отвергаться при «слишком больших» по абсолютной величине значениях разности $\hat{\beta}_{CV} - \hat{\beta}_{GLS}$ (больших — в сравнении со стандартной ошибкой этой разности).

Пусть $\hat{q} = \hat{\beta}_{CV} - \hat{\beta}_{GLS}$, тогда из общей формулы для дисперсии суммы двух случайных величин следует:

$$D(\hat{q}) = D(\hat{\beta}_{CV} - \hat{\beta}_{GLS}) = D(\hat{\beta}_{CV}) + D(\hat{\beta}_{GLS}) - 2Cov(\hat{\beta}_{CV}, \hat{\beta}_{GLS}).$$

Если выполняются предположения стандартной RE -модели, то, как было указано выше, $\hat{\beta}_{GLS}$ является эффективной оценкой, а $\hat{\beta}_{CV}$ — неэффективной. Хаусман показал, что эффективная оценка не коррелирована с разностью ее и неэффективной оценки, так что если гипотеза H_0 верна, то

$$Cov(\hat{\beta}_{GLS}, \hat{\beta}_{GLS} - \hat{\beta}_{CV}) = D(\hat{\beta}_{GLS}) - Cov(\hat{\beta}_{GLS}, \hat{\beta}_{CV}) = 0,$$

$$Cov(\hat{\beta}_{GLS}, \hat{\beta}_{CV}) = D(\hat{\beta}_{GLS}),$$

и

$$D(\hat{q}) = D(\hat{\beta}_{CV}) - D(\hat{\beta}_{GLS}).$$

Как уже говорилось выше,

$$D(\hat{\beta}_{GLS}) = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 + \Psi \sum_{i=1}^N (\bar{x}_i - \bar{x})^2},$$

$$D(\hat{\beta}_{CV}) = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2}.$$

Заменяя в этих выражениях неизвестные параметры их состоятельными оценками, указанными ранее, получим состоятельную оценку $\hat{D}(\hat{q})$ для $D(\hat{q})$.

Статистика критерия Хаусмана

$$m = \frac{\hat{q}^2}{\hat{D}(\hat{q})}$$

имеет при гипотезе H_0 асимптотическое ($N \rightarrow \infty$) распределение $\chi^2(1)$.

Для K регрессоров при гипотезе H_0 статистика

$$m = \hat{q}^T [C\hat{ov}(\hat{q})]^{-1} \hat{q}$$

имеет асимптотическое распределение $\chi^2(K)$.

Численно идентичный критерий для проверки гипотезы H_0 получается при использовании расширенной модели регрессии

$$y_{it}^* = x_{it}^* \beta + (x_{it} - \bar{x}_i) \gamma + \varepsilon_{it},$$

где $y_{it}^* = y_{it} - \theta \bar{y}_i$, $x_{it}^* = x_{it} - \theta \bar{x}_i$,

$$\theta = 1 - \sqrt{\Psi} = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}}.$$

Гипотеза H_0 означает в этой регрессии, что $\gamma = 0$. Можно показать, что здесь

$$\hat{\beta}_{OLS} = \hat{\beta}_b, \quad \hat{\gamma}_{OLS} = \hat{\beta}_{CV} - \hat{\beta}_b.$$

Гипотезу H_0 можно также проверить, используя любую из следующих разностей:

$$\hat{q}_1 = \hat{\beta}_{GLS} - \hat{\beta}_{CV},$$

$$\hat{q}_2 = \hat{\beta}_{GLS} - \hat{\beta}_b,$$

$$\hat{q}_3 = \hat{\beta}_{CV} - \hat{\beta}_b,$$

$$\hat{q}_4 = \hat{\beta}_{GLS} - \hat{\beta}_{OLS}.$$

Это вытекает из соотношения

$$\hat{\beta}_{GLS} = w\hat{\beta}_b + (1-w)\hat{\beta}_{CV}.$$



Замечание 3.2.1. Все входящие в эти разности оценки параметра β состоятельны при гипотезе H_0 , поэтому все эти разности при гипотезе H_0 должны сходиться к нулю.

ПРИМЕР 3.2.2 (продолжение)

Применим критерий Хаусмана (табл. 3.27).

Таблица 3.27

Результаты применения критерия Хаусмана

	Коэффициент		$(b - B)$	$S.E.$
	$(b) \text{ fix}$	(B)		
X	1.102192	1.058959	0.0432328	—

Здесь $b = \hat{\beta}_{CV}$, $B = \hat{\beta}_{GLS}$, $(b - B) = \hat{q} = \hat{\beta}_{CV} - \hat{\beta}_{GLS}$.

Статистика критерия Хаусмана равна:

$$chi2(1) = (b - B)'[(V_b - V_B)^{-1}](b - B) = -2.15,$$

где V_b и V_B — состоятельные оценки дисперсий оценок $\hat{\beta}_{CV}$ и $\hat{\beta}_{GLS}$, соответственно.

Поскольку значение этой статистики оказалось отрицательным, критерий Хаусмана применить не удастся. ■

ПРИМЕР 3.2.3 (размер заработной платы)

Статистические данные (из *National Longitudinal Survey, Youth Sample*, США) содержат сведения о 545 полностью занятых мужчинах, которые окончили школу до 1980 г. и за которыми велось наблюдение в течение 1980—1987 гг. В 1980 г. эти мужчины были в возрасте от 17 до 23 лет и включились в рынок труда совсем недавно, так что на начало периода их трудовой стаж составлял в среднем около 3 лет. Логарифмы среднечасовой заработной платы ($WAGE_LN$) зависят здесь от длительности школьного обучения ($SCHOOL$), трудового стажа ($EXPER$) и его же в квадрате ($EXPER2$), а также от дамми-переменных, указывающих на членство в профсоюзе ($UNION$), работу в государственном секторе (PUB), семейный статус (состоит ли в браке — MAR), а также на цвет кожи (чернокожий или нет — $BLACK$) и испаноязычность ($HISP$).

Результаты оценивания в пакете *Stata8* приведены в табл. 3.28 (в скобках указаны значения t -статистики (или z -статистики) для проверки равенства соответствующего коэффициента нулю).

Таблица 3.28

Результаты оценивания для статистических данных США

Переменная	Оценка			
	<i>Between</i>	<i>FE</i>	<i>OLS</i>	<i>RE</i>
<i>SCHOOL</i>	0.095 (8.68)	—	0.099 (21.24)	0.101 (11.32)
<i>EXPER</i>	−0.050 (−1.00)	0.116 (13.81)	0.089 (8.82)	0.112 (13.52)
<i>EXPER2</i>	0.005 (1.59)	−0.0043 (−7.08)	−0.0028 (−4.03)	−0.0041 (−6.85)

Окончание табл. 3.28

Переменная	Оценка			
	<i>Between</i>	<i>FE</i>	<i>OLS</i>	<i>RE</i>
<i>UNION</i>	0.274 (5.82)	0.081 (4.20)	0.180 (10.46)	0.106 (5.96)
<i>MAR</i>	0.145 (3.50)	0.045 (2.46)	0.108 (6.86)	0.063 (3.73)
<i>BLACK</i>	-0.139 (-2.84)	—	-0.144 (-6.11)	-0.144 (-3.02)
<i>HISP</i>	0.005 (0.13)	—	0.016 (0.75)	0.020 (0.46)
<i>PUB</i>	-0.056 (-0.52)	0.035 (0.90)	0.004 (0.09)	0.030 (0.83)
<i>Within R²</i>	0.0470	0.1782	—	0.1776
<i>Between R²</i>	0.2196	0.0006	—	0.1835
<i>Overall R²</i>	0.1371	0.0642	—	0.1808

Если выполнены предположения модели со случайными эффектами, то все четыре оценки состоятельны (если, конечно, объясняющие переменные не коррелированы с ошибкой), и при этом *RE*-оценка имеет наибольшую эффективность. Если, однако, индивидуальные эффекты α_i коррелированы хотя бы с одной из объясняющих переменных, то состоятельной остается только *FE*-оценка. Поэтому встает вопрос о проверке гипотезы H_0 о том, что модель является *RE*-моделью. Для этого можно сравнивать оценки «внутри» (*FE*) и «между» или оценки «внутри» (*FE*) и *RE* (соответствующие критерии равносильны). Проще сравнивать вторую пару, используя критерий Хаусмана (табл. 3.29), описанный ранее.

Таблица 3.29

Результаты применения критерия Хаусмана

	Коэффициент		$(b - B)$	$\text{sqrt}(\text{diag}(V_b - V_B))$
	(b)	(B)		
	<i>fe</i>	<i>re</i>	<i>Difference</i>	<i>S.E.</i>
<i>EXPER</i>	0.116457	0.1117851	0.0046718	0.0016345
<i>EXPER2</i>	-0.0042886	0.0040575	-0.0002311	0.0001269
<i>UNION</i>	0.081203	0.1064134	-0.0252104	0.0073402
<i>MAR</i>	0.0451061	0.0625465	-0.0174403	0.0073395
<i>PUB</i>	0.0349267	0.0301555	0.0047713	0.0126785
Test: Ho: difference in coefficients not systematic $\chi^2(5) = (b - B)[(V_b - V_B)^{-1}](b - B) = 31.75$ $\text{Prob} > \chi^2 = 0.0000$				

Вычисленное значение статистики критерия (табл. 3.29) равно 31.75 и отражает различия в *FE*- и *RE*-оценках коэффициентов при 5 переменных: *EXPER*, *EXPER2*, *UNION*, *MAR*, *PUB*. Для распределения $\chi^2(5)$ значение 31.75 соответствует *P*-значению $6.6 \cdot 10^{-6}$, так что нулевая гипотеза (*RE*-модель) заведомо отвергается. ■

Автокоррелированные ошибки

Во всех рассмотренных выше ситуациях предполагалось, что случайные составляющие u_{it} — взаимно независимые случайные величины, имеющие одинаковое распределение $N(0, \sigma_u^2)$. Между тем вполне возможно, что для *i*-го субъекта последовательные ошибки $u_{i1}, u_{i2}, \dots, u_{iT}$ не являются независимыми, а следуют, скажем, стационарному процессу авторегрессии первого порядка с нулевым средним.

Точнее говоря, пусть имеем дело с моделью

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

в которой

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it},$$

где $|\rho| < 1$;

$\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{iT}$ — случайные величины, являющиеся гауссовскими инновациями, так что они взаимно независимы и имеют одинаковое распределение $N(0, \sigma_\varepsilon^2)$ и, кроме того, ε_{it} не зависит от значений $u_{i,t-k}$, $k \geq 1$.

Общий для всех субъектов коэффициент ρ можно оценить различными способами. При этом в большинстве случаев сначала переходят к модели, скорректированной на групповые средние:

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i),$$

т.е.

$$\tilde{y}_{it} = \beta \tilde{x}_{it} + \tilde{u}_{it},$$

а затем поступают по-разному.

Можно, например, оценить (методом наименьших квадратов) последнюю модель без учета автокоррелированности ошибок, получить последовательность остатков $\hat{\tilde{u}}_{i1}, \hat{\tilde{u}}_{i2}, \dots, \hat{\tilde{u}}_{iT}$, вычислить значение статистики Дарбина — Уотсона

$$d = \frac{\sum_{i=1}^N \sum_{t=2}^T (\hat{\tilde{u}}_{it} - \hat{\tilde{u}}_{i,t-1})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{\tilde{u}}_{it}^2}$$

и, используя приближенное соотношение $\rho \cong 1 - d/2$, получить оценку $\hat{\rho}_{DW} \cong 1 - d/2$.

Можно поступить иначе: получив последовательность остатков $\hat{u}_{i1}, \hat{u}_{i2}, \dots, \hat{u}_{iT}$, использовать оценку наименьших квадратов, получаемую при оценивании уравнения регрессии

$$\hat{u}_{it} = \rho \hat{u}_{i,t-1} + \eta_{it}.$$

Искомая оценка вычисляется по формуле:

$$\hat{\rho}_{iscorr} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{i,t-1}}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2}.$$

После получения оценки для ρ производится преобразование переменных для получения модели с независимыми ошибками. Наконец, в рамках преобразованной модели производится обычный анализ на фиксированные или случайные эффекты.

ПРИМЕР 3.2.4

В примере с тремя предприятиями для модели с фиксированными эффектами получаем (в пакете *Stata*) следующие результаты:

- при использовании *DW*-оценки — результаты, приведенные в табл. 3.30. Оценка коэффициента β по сравнению со значением 1.102192, полученным ранее без учета возможной автокоррелированности ошибок, практически не изменилась. И это согласуется со значением статистики Дарбина — Уотсона. Вывод об отсутствии индивидуальных эффектов также не изменяется;
- при использовании *tscorr*-оценки — результаты, приведенные в табл. 3.31. Оцененное значение коэффициента автокорреляции ρ на этот раз почти в 2 раза меньше. Оценка коэффициента β практически не изменилась. ■

Таблица 3.30

Результаты использования *DW*-оценки

`.xtregar y x, fe rhotype(dw) lbi` — **FE (within) regression with AR(1) disturbances;**
Number of obs = 27

<i>R-SQ</i> :	
<i>within</i>	= 0.9569
<i>between</i>	= 0.1111
<i>overall</i>	= 0.9252
$F(1, 23) = 510.54, \quad Prob > F = 0.0000$	
$corr(\hat{\alpha}_i, Xb) = -0.1625$	

Окончание табл. 3.30

<i>y</i>	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение <i>P</i> > <i>t</i>
<i>x</i>	1.105796	0.0489398	22.60	0.000
<i>cons</i>	−1.317209	0.6964634	−1.89	0.071
<i>rho_ar</i>	0.170171			
<i>sigma_α</i>	1.423608			
<i>sigma_u</i>	1.773845			
<i>rho_fov</i>	0.3917622	(доля дисперсии, соответствующей индивидуальным эффектам α_i)		
<i>F</i> test that all <i>alfa_i</i> = 0:				
<i>F</i> (2,23) = 3.82, <i>Prob</i> > <i>F</i> = 0.0370				
<i>modified Bhargava et al. Durbin-Watson</i> = 1.664958				

Таблица 3.31

Результаты использования *tscorr*-оценки
.xtregar y x, fe rhotype(tscorr); FE (within) regression with AR(1) disturbances;
 Number of obs = 27

<i>R-SQ:</i>				
<i>within</i> = 0.9540				
<i>between</i> = 0.1111				
<i>overall</i> = 0.9252				
$F(1, 23) = 476.47$, $Prob > F = 0.0000$				
$corr(\alpha_i, Xb) = -0.1626$				
<i>y</i>	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение $P > t$
<i>x</i>	1.109267	0.050818	21.83	0.000
<i>cons</i>	-1.392025	0.7698403	-1.81	0.084
<i>rho_ar</i>	0.09213053			
<i>sigma_α</i>	1.4281087			
<i>sigma_u</i>	1.7701594			
<i>rho_fov</i>	0.39426073	(доля дисперсии, соответствующей индивидуальным эффектам α_i)		
<i>F</i> test that all $\alpha_i = 0$:				
$F(2,23) = 4.66$, $Prob > F = 0.0199$				

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В каких ситуациях для анализа панельных данных используется модель с фиксированными эффектами? Как оценивается такая модель? Каковы свойства получаемых оценок?
2. В каких ситуациях для анализа панельных данных используется модель со случайными эффектами? Как оценивается такая модель? Каковы свойства получаемых оценок?
3. Каким образом производится выбор между моделями с фиксированными и случайными эффектами? Какие соображения лежат в основе построения критерия Хаусмана?
4. Каковы особенности модели Мундлака? Как оценивается такая модель?
5. Каковы особенности анализа панельных данных в случае автокоррелированности случайных ошибок?
6. Какие варианты коэффициента детерминации используются при анализе панельных данных?

Тема 3.3

ДВУНАПРАВЛЕННЫЕ МОДЕЛИ

Двунаправленные модели с фиксированными эффектами

Рассмотрим модель, в которую помимо **индивидуальных эффектов** (*individual effects*) α_i включаются также **временные эффекты** (*time effects*) λ_t :

$$y_{it} = \mu + \alpha_i + \lambda_t + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где $\sum_{i=1}^N \alpha_i = 0$, $\sum_{t=1}^T \lambda_t = 0$, так что α_i и λ_t — дифференциальные эффекты. При этом и α_i , и λ_t интерпретируются как неизвестные постоянные.

Обозначив

$$\bar{y}_t = \frac{1}{N} \sum_{i=1}^N y_{it} \text{ и т.д. (средние по субъектам),}$$

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it} \text{ и т.д. (средние по времени),}$$

$$\bar{y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it} \text{ и т.д. (средние по всем наблюдениям),}$$

получим:

$$(y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}) = (x_{it} - \bar{x}_i - \bar{x}_t + \bar{x})\beta + (u_{it} - \bar{u}_i - \bar{u}_t + \bar{u}).$$

Оценка наименьших квадратов для коэффициента β в этом уравнении (двунаправленная внутригрупповая оценка — *two-way within estimator*) имеет вид:

$$\hat{\beta}_{CV} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i - \bar{y}_t + \bar{y})(x_{it} - \bar{x}_i - \bar{x}_t + \bar{x})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i - \bar{x}_t + \bar{x})^2} = \frac{W_{xy}}{W_{xx}}.$$

На основании соотношений

$$(\bar{y}_i - \bar{y}) = \alpha_i + (\bar{x}_i - \bar{x})\beta + (\bar{u}_i - \bar{u}),$$

$$(\bar{y}_t - \bar{y}) = \lambda_t + (\bar{x}_t - \bar{x})\beta + (\bar{u}_t - \bar{u})$$

можно получить оценки для α_i и λ_t :

$$\hat{\alpha}_i = (\bar{y}_i - \bar{y}) - \hat{\beta}_{CV}(\bar{x}_i - \bar{x}),$$

$$\hat{\lambda}_t = (\bar{y}_t - \bar{y}) - \hat{\beta}_{CV}(\bar{x}_t - \bar{x}).$$

Для оценивания двунаправленной модели с фиксированными эффектами в пакете *Stata* применяется процедура *xtreg, fe* с включением дамми-переменных для временных периодов.

ПРИМЕР 3.3.1 (продолжение примера 3.2.3 — размер заработной платы)

Результаты оценивания двунаправленной модели с фиксированными эффектами приведены в табл. 3.32.

Таблица 3.32

Оценка двунаправленной модели с фиксированными эффектами

Fixed-effects (within) regression

<i>R-SQ:</i>				
<i>within</i> = 0.1808, <i>between</i> = 0.0005, <i>overall</i> = 0.0638				
<i>F</i> (11, 3804) = 76.30, <i>Prob</i> > <i>F</i> = 0.0000				
<i>corr</i> (<i>a_i</i> , <i>Xb</i>) = -0.1203				
Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение <i>P</i> > <i>t</i>
<i>BLACK</i> (<i>dropped</i>)				
<i>EXPER</i>	0.1317361	0.0098356	3.39	0.000
<i>EXPER2</i>	-0.0051704	0.0007047	7.34	0.000
<i>HISP</i> (<i>dropped</i>)				
<i>MAR</i>	0.0464781	0.0183123	0.54	0.011
<i>PUB</i>	0.0347278	0.0385989	0.90	0.368
<i>SCHOOL</i> (<i>dropped</i>)				

Окончание табл. 3.32

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение <i>P</i> > <i>t</i>
<i>UNION</i>	0.0791253	0.0193354	0.09	0.000
<i>Y80 (dropped)</i>				
<i>Y81</i>	0.0193045	0.0203652	0.95	0.343
<i>Y82</i>	−0.0112773	0.0202281	0.56	0.577
<i>Y83</i>	−0.0419533	0.0203211	0.06	0.039
<i>Y84</i>	−0.0383904	0.0203151	1.89	0.059
<i>Y85</i>	−0.0428743	0.0202506	2.12	0.034
<i>Y86</i>	−0.0275581	0.0203878	1.35	0.177
<i>cons</i>	1.0283830	0.0299620	4.32	0.000

Заметим, что значимыми здесь оказываются оцененные коэффициенты при тех же переменных, что и в однонаправленной модели с фиксированными эффектами. Остальная часть протокола представлена в табл. 3.33. Результаты последнего теста не выявляют значимого влияния временных эффектов. ■

Таблица 3.33

Оценка двунаправленной модели с фиксированными эффектами (продолжение)

<i>sigma_α</i>	0.40078197
<i>sigma_u</i>	0.3509988
<i>rho</i>	0.56593121 (<i>fraction of variance due to α_i</i>)
<i>F</i> test that all <i>A_i</i> = 0:	
<i>F</i> (544, 3804) = 7.97, <i>Prob</i> > <i>F</i> = 0.0000	
<i>test</i> <i>Y81</i> = <i>Y82</i> = <i>Y83</i> = <i>Y84</i> = <i>Y85</i> = <i>Y86</i> = 0 (здесь <i>Y81</i> — дамми-переменная, равная 1 в 1981 г. и равная 0 в остальные годы; аналогично определяются <i>Y82</i> , ..., <i>Y86</i>)	
<i>F</i> (6, 3804) = 1.96, <i>Prob</i> > <i>F</i> = 0.0680	

Двунаправленные модели со случайными эффектами

Пусть теперь в двунаправленной модели

$$y_{it} = \mu + \alpha_i + \lambda_t + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

предполагается, что α_i и λ_t — случайные величины и что

$$E(\alpha_i) = E(\lambda_t) = E(u_{it}) = 0,$$

$$E(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2, & \text{если } i = j, \\ 0 & \text{если } i \neq j, \end{cases}$$

$$E(\lambda_t \lambda_s) = \begin{cases} \sigma_\lambda^2, & \text{если } t = s, \\ 0 & \text{если } t \neq s, \end{cases}$$

$$E(u_{it} u_{js}) = \begin{cases} \sigma_u^2, & \text{если } i = j \text{ и } t = s, \\ 0 & \text{в противном случае,} \end{cases}$$

$$E(\alpha_i \lambda_t) = 0, \quad E(\alpha_i u_{it}) = 0, \quad E(\lambda_t u_{it}) = 0,$$

$$E(\alpha_i x_{it}) = E(\lambda_t x_{it}) = E(u_{it} x_{it}) = 0.$$

Определим $v_{it} = \alpha_i + \lambda_t + u_{it}$. Тогда

$$D(v_{it}) = D(y_{it} | x_{it}) = \sigma_\alpha^2 + \sigma_\lambda^2 + \sigma_u^2,$$

так что ошибка состоит из трех компонент. При этом

$$\begin{aligned} \text{Cov}(v_{it}, v_{js}) &= \begin{cases} \sigma_\alpha^2, & \text{если } i = j, t \neq s, \\ \sigma_\lambda^2, & \text{если } t = s, i \neq j, \end{cases} \\ \text{Corr}(v_{it}, v_{js}) &= \begin{cases} \frac{\sigma_\alpha^2}{D(v_{it})}, & \text{если } i = j, t \neq s, \\ \frac{\sigma_\lambda^2}{D(v_{it})}, & \text{если } t = s, i \neq j, \\ 1, & \text{если } t = s, i = j, \\ 0, & \text{если } t \neq s, i \neq j, \end{cases} \end{aligned}$$

«Межсубъектная» оценка определяется следующим образом:

$$\hat{\beta}_{bi} = \frac{\sum_{i=1}^N (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})}{\sum_{i=1}^N (\bar{x}_i - \bar{x})^2} = \frac{B_{xy}}{B_{xx}},$$

а «межвременная» оценка равна:

$$\hat{\beta}_{bt} = \frac{\sum_{t=1}^T (\bar{x}_t - \bar{x})(\bar{y}_t - \bar{y})}{\sum_{t=1}^T (\bar{x}_t - \bar{x})^2} = \frac{C_{xy}}{C_{xx}}.$$

GLS -оценка равна:

$$\hat{\beta}_{GLS} = \omega_1 \hat{\beta}_{CV} + \omega_2 \hat{\beta}_{bi} + \omega_3 \hat{\beta}_{bt},$$

где

$$\omega_1 = \frac{W_{xx}}{W_{xx} + \phi_2^2 B_{xx} + \phi_3^2 C_{xx}} = \frac{W_{xx}}{T_{xx}},$$

$$\omega_2 = \frac{\phi_2^2 B_{xx}}{T_{xx}}, \quad \omega_3 = \frac{\phi_3^2 C_{xx}}{T_{xx}},$$

$$\phi_2^2 = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}, \quad \phi_3^2 = \frac{\sigma_u^2}{\sigma_u^2 + N\sigma_\lambda^2},$$

$$T_{xx} = W_{xx} + \phi_2^2 B_{xx} + \phi_3^2 C_{xx}.$$

Иначе говоря, GLS -оценка является взвешенным средним одной «внутри»- и двух «между»-оценок с весами, отражающими источники изменчивости.

- Если $\sigma_\alpha^2 = \sigma_\lambda^2 = 0$ (так что все α_i и λ_t равны нулю), то

$$\phi_2^2 = \phi_3^2 = 1, \quad \omega_2 = \frac{B_{xx}}{T_{xx}}, \quad \omega_3 = \frac{C_{xx}}{T_{xx}}, \quad T_{xx} = W_{xx} + B_{xx} + C_{xx},$$

$$\begin{aligned} \hat{\beta}_{GLS} &= \frac{W_{xx}}{T_{xx}} \frac{W_{xy}}{W_{xx}} + \frac{B_{xx}}{T_{xx}} \frac{B_{xy}}{B_{xx}} + \frac{C_{xx}}{T_{xx}} \frac{C_{xy}}{C_{xx}} = \frac{W_{xy} + B_{xy} + C_{xy}}{W_{xx} + B_{xx} + C_{xx}} = \\ &= \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2} = \hat{\beta}_{OLS} \quad (\text{как для пула}). \end{aligned}$$

- При $T \rightarrow \infty$ и $N \rightarrow \infty$ имеем $\phi_2^2 \rightarrow 0$ и $\phi_3^2 \rightarrow 0$, и

$$\hat{\beta}_{GLS} = \hat{\beta}_{CV} \quad (\text{как для модели с фиксированными эффектами}).$$

Можно также комбинировать фиксированные временные эффекты и случайные индивидуальные эффекты с включением дамми-переменных для временных периодов.

Критерии для индивидуальных и временных эффектов

Критерий Бройша — Пагана. Это критерий для проверки гипотезы

$$H_0: \sigma_\alpha^2 = \sigma_\lambda^2 = 0 \quad (\text{сведение к модели пула}).$$

Здесь опять можно использовать OLS -оценки параметров для получения OLS -остатков \hat{u}_{it} .

Статистика критерия $BP = LM_1 + LM_2$, где

$$LM_1 = \frac{NT}{2(T-1)} \left[\frac{T^2 \sum_{i=1}^N \bar{\hat{u}}_i^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right]^2,$$

$$LM_2 = \frac{NT}{2(T-1)} \left[\frac{N^2 \sum_{t=1}^T \bar{\hat{u}}_t^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2} - 1 \right]^2.$$

При гипотезе H_0 статистика BP имеет асимптотическое распределение $\chi^2(2)$.
При гипотезе $H_0^\alpha: \sigma_\alpha^2 = 0$ $LM_1 \sim \chi^2(1)$; при гипотезе $H_0^b: \sigma_\lambda^2 = 0$ $LM_2 \sim \chi^2(1)$.

В качестве альтернативных могут быть использованы F -критерии, как при проверке гипотез для фиксированных эффектов:

- для проверки гипотезы

$$H_0^\alpha: \sigma_\alpha^2 = 0 \quad (\text{при условии } \lambda_t = 0)$$

используем применявшуюся ранее статистику:

$$F_3 = \frac{\frac{S_2 - S_1}{N-1}}{\frac{S_1}{NT - N - 1}},$$

где S_2 — RSS от OLS -регрессии (пул);

S_1 — RSS от однонаправленной «внутригрупповой» регрессии (основанной на однонаправленной CV -оценке);

- для проверки гипотезы

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_{N-1} = 0, \quad \lambda_1 = \lambda_2 = \dots = \lambda_{T-1} = 0$$

используем статистику:

$$F_{2\text{-way}} = \frac{\frac{S_3 - S_{2w}}{N+T-2}}{\frac{S_{2w}}{(N-1)(T-1)-1}},$$

где S_{2w} — RSS от двунаправленной «внутригрупповой» регрессии;

- для проверки гипотезы

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_{N-1} = 0 \quad \text{при } \lambda_t \neq 0$$

используем статистику:

$$F_4 = \frac{\frac{S_4 - S_{2w}}{N-1}}{\frac{S_{2w}}{(N-1)(T-1)-1}},$$

где S_4 — RSS от регрессии

$$(y_{it} - \bar{y}_t) = (x_{it} - \bar{x}_t)\beta + (u_{it} - \bar{u}_t);$$

- для проверки гипотезы

$$H_0 : \lambda_1 = \lambda_2 = \dots = \lambda_{T-1} = 0 \quad \text{при} \quad \alpha_i \neq 0$$

используем статистику:

$$F_5 = \frac{\frac{S_1 - S_{2w}}{T-1}}{\frac{S_{2w}}{(N-1)(T-1)-1}}.$$

Может иметь смысл трактовка α_i как случайных, а λ_i как фиксированных эффектов.

Для рассматриваемого примера. В рамках модели с фиксированными индивидуальными и временными эффектами проверка гипотезы об отсутствии индивидуальных эффектов дает результат, представленный в табл. 3.34. Полученное значение F -статистики приводит к отвержению этой гипотезы на любом разумном уровне значимости.

Таблица 3.34

Проверка гипотезы об отсутствии индивидуальных эффектов

F test that all $A_i = 0$:
$F(544, 3804) = 7.97, \quad Prob > F = 0.0000$

Проверка в рамках той же модели гипотезы об отсутствии временных эффектов дает результат, приведенный в табл. 3.35. Гипотеза об отсутствии фиксированных временных эффектов не отвергается F -критерием на 5%-м уровне значимости.

Таблица 3.35

Проверка гипотезы об отсутствии временных эффектов

test: $Y81 = Y82 = Y83 = Y84 = Y85 = Y86 = 0$
$F(6, 3804) = 1.96, \quad Prob > F = 0.0680$

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Как производится оценивание двунаправленных моделей с фиксированными эффектами?
2. Как производится оценивание двунаправленных моделей со случайными эффектами?

Тема 3.4

НЕСБАЛАНСИРОВАННЫЕ ПАНЕЛИ, ЭНДОГЕННЫЕ ОБЪЯСНЯЮЩИЕ ПЕРЕМЕННЫЕ, МОДЕЛИ С ИНДИВИДУАЛЬНО-СПЕЦИФИЧЕСКИМИ ПЕРЕМЕННЫМИ

Несбалансированные панели

Сосредоточимся на однофакторной модели. Такая модель могла бы включать временные дамми как фиксированные эффекты.

До сих пор мы предполагали, что

$$i = 1, \dots, N, \quad t = 1, \dots, T,$$

так что в каждый из T моментов времени имеются данные обо всех N субъектах, участвующих в анализе. В таких случаях говорят о **сбалансированной панели** (*balanced panel*). Теперь рассмотрим модель

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T_i, \quad \sum_{i=1}^N \alpha_i = 0,$$

в которой количество наблюдений для разных субъектов может быть различным. В этом случае говорят о **несбалансированной панели** (*unbalanced panel*).

Основными результатами для несбалансированных панелей являются следующие:

- 1) *OLS*-оценка та же, что и раньше, и она является *BLUE*, если $\sigma_\alpha^2 = 0$;
- 2) внутригрупповая (*CV*) оценка в основных чертах та же, что и ранее, хотя \bar{y}_i и \bar{x}_i вычисляются по периодам времени разной длины для разных субъектов;
- 3) «между»-оценка также в основном сохраняется, только средние вычисляются по T_i наблюдениям для субъекта i ;
- 4) преобразование переменных для получения *GLS*-оценки здесь имеет вид:

$$y_{it}^* = y_{it} - \theta_i \bar{y}_i,$$

$$x_{it}^* = x_{it} - \theta_i \bar{x}_i,$$

$$\text{где } \bar{y}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} y_{it}, \quad \bar{x}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} x_{it}, \quad \theta_i = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T_i \sigma_\alpha^2}},$$

т.е. θ_i изменяется от субъекта к субъекту.

Эндогенные объясняющие переменные

Рассмотрим модель с несколькими объясняющими переменными, часть из которых являются эндогенными:

$$y_{2it} = y_{1it}\gamma + x_{it}\beta + \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где y_{1it} — вектор-строка g_1 эндогенных переменных;

x_{it} — вектор-строка k_1 экзогенных переменных;

γ и β — векторы-столбцы размерностей g_1 и k_1 .

Пусть z_{it} — вектор-строка k_2 инструментальных переменных, $k_2 \geq g_1$, так что $E(z_{it} u_{it}) = 0$.

Обозначим:

$$\begin{aligned} \bar{y}_{2i} &= \frac{1}{T} \sum_{t=1}^T y_{2it}, & \bar{y}_{1i} &= \frac{1}{T} \sum_{t=1}^T y_{1it}, & \bar{x}_i &= \frac{1}{T} \sum_{t=1}^T x_{it}, \\ \bar{y}_2 &= \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{2it}, & \bar{y}_1 &= \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{1it}, & \bar{x} &= \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{it}. \end{aligned}$$

Оценив методом инструментальных переменных (метод IV) «внутри»-регрессию

$$(y_{2it} - \bar{y}_{2i}) = (y_{1it} - \bar{y}_{1i})\gamma + (x_{it} - \bar{x}_i)\beta + (u_{it} - \bar{u}_i),$$

получим **IV-«внутри»-оценки** (IV within estimators) $\hat{\gamma}_{IVW}$, $\hat{\beta}_{IVW}$. Оценив методом инструментальных переменных «между»-регрессию

$$(\bar{y}_{2i} - \bar{y}_2) = (\bar{y}_{1i} - \bar{y}_1)\gamma + (\bar{x}_i - \bar{x})\beta + (\bar{u}_i - \bar{u}),$$

получим **IV-«между»-оценку** (IV between estimator) $\hat{\gamma}_{IVB}$.

Например, если $k_2 = g_1 = 1$, $k_1 = 0$, так что $y_{2it} = y_{1it}\gamma + \alpha_i + u_{it}$, то

$$\begin{aligned} \hat{\gamma}_{IVW} &= \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{2it} - \bar{y}_{2i})(z_{it} - \bar{z}_i)}{\sum_{i=1}^N \sum_{t=1}^T (y_{1it} - \bar{y}_{1i})(z_{it} - \bar{z}_i)}, \\ \hat{\gamma}_{IVB} &= \frac{\sum_{i=1}^N \bar{y}_{2i} \bar{z}_i}{\sum_{i=1}^N \bar{y}_{1i} \bar{z}_i}. \end{aligned}$$

Если $E(Y_{1it} \alpha_i) = 0$, то более эффективна оценка со случайными эффектами. Чтобы получить ее, используем преобразование переменных:

$$\begin{aligned}\tilde{y}_{2it} &= y_{2it} - \hat{\theta} \bar{y}_{2i}, & \tilde{y}_{1it} &= y_{1it} - \hat{\theta} \bar{y}_{1i}, \\ \tilde{u}_{it} &= u_{it} - \hat{\theta} \bar{u}_i, & \tilde{z}_{it} &= z_{it} - \hat{\theta} \bar{z}_i,\end{aligned}$$

где

$$\begin{aligned}\hat{\theta} &= 1 - \sqrt{\frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + T\hat{\sigma}_\alpha^2}}, \\ \hat{\sigma}_u^2 &= \frac{\sum_{i=1}^N \sum_{t=1}^T ((y_{2it} - \bar{y}_{2i}) - (y_{1it} - \bar{y}_{1i}) \hat{\gamma}_{IYW})^2}{NT - N}, \\ \hat{\sigma}_u^2 + T\hat{\sigma}_\alpha^2 &= \frac{T \sum_{i=1}^N (\bar{y}_{2i} - \bar{y}_{1i} \hat{\gamma}_{IWB})^2}{N}.\end{aligned}$$

Применим метод IV к уравнению $\tilde{y}_{2it} = \tilde{y}_{1it} \gamma + \tilde{u}_{it}$, используя в качестве инструментов $z_{it} - \bar{z}_i$ и \bar{z}_i или \tilde{z}_{it} .

Более общим образом, пусть $\delta = [\gamma \beta]$ — вектор-строка размерности $(g_1 + k_1)$. На практике приходится применять метод IV (2SLS) трижды:

- 1) для получения $\hat{\delta}_{IYW}$;
- 2) для получения $\hat{\delta}_{IWB}$;

в результате этих двух шагов получают оценки $\hat{\sigma}_u^2$ и $\hat{\sigma}_u^2 + T\hat{\sigma}_\alpha^2$, которые используются для преобразования модели;

- 3) реализуя метод IV для преобразованной модели.

ПРИМЕР 3.4.1

Для исследования зависимости заработной платы женщин от различных факторов были взяты данные из *National Longitudinal Survey, Youth Sample*, США, по $N = 4134$ молодым женщинам, находившимся в 1968 г. в возрасте от 14 до 26 лет. Наблюдения проводились с 1968 по 1988 г. Однако данные неполные: по отдельным субъектам количество наблюдений изменялось от 1 до 12 (в среднем 4.6 наблюдения для одного субъекта).

Рассмотрим сначала модель с фиксированными эффектами:

$$\begin{aligned}\ln w_{it} &= \beta_1 tenure + \beta_2 age + \beta_3 age2 + \beta_4 notsmsa + \\ &+ \beta_5 union + \beta_6 south + \mu + \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,\end{aligned}$$

где w_{it} — размер заработной платы;
 $tenure$ — продолжительность (стаж) работы на наблюдаемом рабочем месте;
 age — возраст;
 $notsmsa$ — проживание вне столичных регионов;
 $union$ — принадлежность к профсоюзу;
 $south$ — проживание на юге страны.

При оценивании указанной модели с использованием «внутри»-оценки получим результаты, приведенные в табл. 3.36. Все оцененные коэффициенты имеют высокую статистическую значимость и ожидаемые знаки. Значительная часть изменчивости (70%) объясняется индивидуальными эффектами (табл. 3.37).

Таблица 3.36

Результаты применения «внутри»-оценки
`.xtreg n_w tenure age age_2 not_smsa union south, fe i(idcode)`
Fixed-effects (within) regression

<i>R-SQ:</i>				
<i>within</i> = 0.1333, <i>between</i> = 0.2375, <i>overall</i> = 0.2031				
<i>F</i> (6, 14867) = 381.19, <i>Prob</i> > <i>F</i> = 0.0000				
<i>corr</i> (α_i , Xb) = 0.2074				
<i>ln_wage</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение <i>P</i> > <i>z</i>
<i>tenure</i>	0.0176205	0.0008099	21.76	0.000
<i>age</i>	0.0311984	0.0033902	9.20	0.000
<i>age_2</i>	−0.0003457	0.0000543	−6.37	0.000
<i>not_smsa</i>	−0.0972535	0.0125377	−7.76	0.000
<i>union</i>	0.0975672	0.0069844	13.97	0.000
<i>south</i>	−0.0620932	0.0133270	−4.66	0.000
<i>cons</i>	1.091612	0.0523126	20.87	0.000

Таблица 3.37

Применение «внутри»-оценки (продолжение)

<i>sigma_α</i>	0.3910683
<i>sigma_u</i>	0.25545969
<i>rho</i>	0.70091004 (<i>fraction of variance due to α_i</i>)
<i>F test that all $\alpha_i = 0$:</i>	
<i>F</i> (4133, 14867) = 8.31, <i>Prob</i> > <i>F</i> = 0.0000	

В то же время если считать, например, что стаж работы на наблюдаемом рабочем месте зависит от принадлежности к профсоюзу и от региона проживания (юг — не юг) и ошибки в уравнении для такой связи коррелированы

с ошибками в уравнении для логарифма заработной платы, тогда переменная $tenure_{it}$ в уравнении

$$\ln w_{it} = \beta_1 tenure_{it} + \beta_2 age_{it} + \beta_3 age2_{it} + \beta_4 notsmsa_{it} + \mu + \alpha_i + u_{it}$$

коррелирована с ошибкой u_{it} , и для получения состоятельных оценок коэффициентов этого уравнения приходится прибегать к методу инструментальных переменных. При сделанных предположениях в качестве инструментов для $tenure_{it}$ можно использовать *union* и *south*. Полный список инструментов, обеспечивающих однозначную идентификацию коэффициентов μ и $\beta_1, \beta_2, \beta_3, \beta_4$, включает 5 переменных: *union, south, age, age2, notsmsa*. При использовании этих переменных в качестве инструментов получим результаты, приведенные в табл. 3.38 и 3.39. Оцененные коэффициенты и здесь имеют ожидаемые знаки. Однако на этот раз оказались статистически незначимыми оцененные коэффициенты при переменных *age* и *notsmsa*. ■

Таблица 3.38

Результаты применения инструментальных переменных
`.xtivreg ln_w age age_2 not_smsa (tenure = union south), fe i(idcode)`
Fixed-effects (within) IV regression

<i>R</i> -SQ:				
<i>within</i> = 0.1333, <i>between</i> = 0.2375, <i>overall</i> = 0.2031				
Wald $\chi^2(4) = 147926.58$ <i>Prob</i> > $\chi^2 = 0.0000$				
$\text{corr}(\alpha_i, Xb) = -0.6843$				
<i>ln_wage</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z$
<i>tenure</i>	0.2403531	0.0373419	6.44	0.000
<i>age</i>	0.0118437	0.0090032	1.32	0.188
<i>age_2</i>	-0.0012145	0.0001968	-6.17	0.000
<i>not_smsa</i>	-0.0167178	0.0339236	-0.49	0.622
<i>cons</i>	1.678287	0.1626657	10.32	0.000

Таблица 3.39

Применение инструментальных переменных (продолжение)

<i>sigma_a</i>	0.70661941
<i>sigma_u</i>	0.63029359
<i>rho</i>	0.55690561 (<i>fraction of variance due to α_i</i>)
<i>F</i> test that all $\alpha_i = 0$:	
$F(4133, 14869) = 1.44$, <i>Prob</i> > $F = 0.0000$	
Instrumented: <i>tenure</i>	
Instruments: <i>age age_2 not_smsa union south</i>	

Применение «внутри»-оценки предпочтительно при интерпретации индивидуальных эффектов как фиксированных эффектов. Если эти эффекты рассматривать как случайные и некоррелированные с остальными объясняющими переменными, то предпочтительнее использовать инструментальное GLS-оценивание, как это было описано выше. Такой подход приводит к результатам, представленным в таблицах 3.40 и 3.41.

Таблица 3.40

Применение инструментального GLS-оценивания
.xtivreg ln_w age age2 not_smsa black (tenure = union birth_yr south black), re i(idcode)
G2SLS random-effects IV regression

<i>R-SQ:</i>				
<i>within</i> = 0.0664, <i>between</i> = 0.2098, <i>overall</i> = 0.1463				
<i>Wald chi2(4)</i> = 1446.37 <i>Prob > chi2</i> = 0.0000				
<i>corr(α_i, Xb)</i> = 0 (assumed)				
<i>ln_wage</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение <i>P > z</i>
<i>tenure</i>	0.1391798	0.0078756	17.67	0.000
<i>age</i>	0.0279649	0.0054182	5.16	0.000
<i>age_2</i>	-0.0008357	0.0000871	-9.60	0.000
<i>not_smsa</i>	-0.2235103	0.0111371	-20.07	0.000
<i>black</i>	-0.2078613	0.0125803	-16.52	0.000
<i>cons</i>	1.337684	0.0844988	15.83	0.000

Таблица 3.41

Применение инструментального GLS-оценивания (продолжение)

<i>sigma_a</i>	0.36582493			
<i>sigma_u</i>	0.63031479			
<i>rho</i>	0.25197078 (<i>fraction of variance due to a_i</i>)			
<i>Instrumented: tenure</i>				
<i>Instruments: age age_2 not_smsa black union birth_yr south</i>				

Модели с индивидуально-специфическими переменными

Оценивание в RE- и FE-моделях. До сих пор в модели с фиксированными эффектами неоднородность субъектов исследования характеризовалась наличием ненаблюдаемых характеристик, влияние которых отражалось в модели посредством параметров α_i . Однако неоднородность субъектов может выражаться также в различных значениях для этих субъектов некоторых наблю-

даемых характеристик, не изменяющихся для каждого субъекта в процессе наблюдений. Например, в исследованиях, касающихся зависимости размера заработной платы индивида от различных факторов, такими характеристиками могут быть пол, базовое образование и т.п. В связи с этим рассмотрим модель

$$y_{it} = \mu + \alpha_i + \beta x_{it} + \gamma z_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T_i,$$

где z_i — переменная, специфическая только в отношении субъекта.

Если эта модель трактуется как *RE*-модель, в которой эффекты не коррелированы с x_{it} и z_i , то проблем с оцениванием коэффициентов β и γ не возникает: в этом случае *BLUE* являются *GLS*-оценки для β и γ .

Если же эта модель трактуется как *FE*-модель или если α_i случайны и $E(\alpha_i z_i) = 0$, но $E(\alpha_i x_{it}) \neq 0$, то *GLS*-оценки (строящиеся, как в *RE*-модели) несостоятельны, и приходится искать другие оценки. Усредним обе части уравнения по t :

$$\bar{y}_i = \mu + \alpha_i + \beta \bar{x}_i + \gamma z_i + \bar{u}_i, \quad i = 1, \dots, N.$$

Тогда

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N.$$

В применении к последнему уравнению метод наименьших квадратов приводит опять к оценке с фиксированными эффектами ($\hat{\beta}_{CV}$). Но при таком подходе из исходного уравнения вымечаются не только α_i , но и z_i . Однако если α_i фиксированы или $E(\alpha_i x_{it}) \neq 0$, но $E(\alpha_i z_i) = 0$, то тогда все же можно построить состоятельную оценку коэффициента γ .

Для этого заметим, что

$$\bar{y}_i - \beta \bar{x}_i = \mu + \gamma z_i + (\alpha_i + \bar{u}_i).$$

Если считать значение β известным, можно оценить эту модель, минимизируя $\sum_{i=1}^N (\alpha_i + \bar{u}_i)^2$, и получить *between*-оценки:

$$\hat{\gamma} = \frac{\sum_{i=1}^N ((y_i - \bar{y}) - (\bar{x}_i - \bar{x})\beta)(z_i - \bar{z})}{\sum_{i=1}^N (z_i - \bar{z})^2},$$

$$\hat{\mu} = \bar{y} - \bar{x}\beta - \bar{z}\hat{\gamma}.$$

Подставив $\hat{\beta}_{CV}$ вместо β в эти два выражения, получим оценки $\tilde{\gamma}$ и $\tilde{\mu}$ для γ и μ . При этом $\text{plim}_{N \rightarrow \infty} \tilde{\gamma} = \gamma$, т.е. $\tilde{\gamma}$ — состоятельная оценка.

ПРИМЕР 3.4.2 (продолжение примера 3.2.3 — размер заработной платы)

При оценивании модели с фиксированными эффектами оказались выметенными переменные *SCHOOL*, *BLACK*, *HISP*. Попробуем все же получить оценки коэффициентов при этих переменных. Использование процедуры, имеющейся в пакете *Stata* 8, приводит к результату, представленному в табл. 3.42. Полученные оценки практически совпадают с оценками для соответствующих коэффициентов в модели со случайными эффектами (при *GLS*-оценивании). ■

Таблица 3.42

Between-регрессия (регрессия на групповые средние)

<i>R-SQ</i> :				
<i>within</i> = 0.0000				
<i>between</i> = 0.2119				
<i>overall</i> = 0.1264				
$F(3, 541) = 48.47$, $Prob > F = 0.0000$				
Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение $P > t$
<i>SCHOOL</i>	0.1015825	0.0089372	11.37	0.000
<i>BLACK</i>	−0.1442396	0.0484007	−2.98	0.003
<i>HISP</i>	0.0210173	0.0435069	0.48	0.629

Модель Хаусмана — Тейлора. Рассмотрим модель

$$y_{it} = X_{it}\beta + Z_i\gamma + \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где X_{it} — строка k наблюдаемых переменных, изменяющихся и от субъекта к субъекту, и во времени;

Z_i — строка g наблюдаемых переменных, инвариантных относительно времени на периоде наблюдений.

Предполагается, что выполнены обычные предположения *RE*-модели (в частности, что все объясняющие переменные экзогенны в обычном смысле), за исключением того, что теперь ненаблюдаемые индивидуальные эффекты α_i могут быть коррелированными с одними объясняющими переменными и не коррелированными с другими объясняющими переменными.

Хаусман и Тейлор предложили в таком случае производить разбиение:

$$X_{it} = [X_{1it} \ X_{2it}], \quad Z_i = [Z_{1i} \ Z_{2i}],$$

в котором k_1 переменных, входящих в состав X_{1it} , и g_1 переменных, входящих в состав Z_{1i} , экзогенны по отношению к α_i в том смысле, что

$$E(X_{1it} \alpha_i) = E(Z_{1i} \alpha_i) = 0,$$

а k_2 переменных, входящих в состав X_{2it} , и g_2 переменных, входящих в состав Z_{2i} , эндогенны по отношению к α_i в том смысле, что

$$E(X_{2it} \alpha_i) \neq 0, \quad E(Z_{2i} \alpha_i) \neq 0.$$

При таком разбиении модель записывается в виде:

$$y_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\gamma_1 + Z_{2i}\gamma_2 + \alpha_i + u_{it},$$

и переход к модели, скорректированной на индивидуальные средние,

$$y_{it} - \bar{y}_i = (X_{1it} - \bar{X}_{1i})\beta_1 + (X_{2it} - \bar{X}_{2i})\beta_2 + u_{it},$$

приводит к возможно неэффективной, но состоятельной оценке $\hat{\beta}_{CV}$ для $\beta = (\beta_1^T, \beta_2^T)^T$. Однако при этом опять вместе с α_i вымываются переменные, инвариантные относительно времени.

Получив оценку $\hat{\beta}_{CV} = (\hat{\beta}_{1,CV}^T, \hat{\beta}_{2,CV}^T)^T$ для β , вычислим для каждого субъекта остатки от оцененной «внутри»-регрессии:

$$d_{it} = (y_{it} - \bar{y}_i) - (X_{1it} - \bar{X}_{1i})\hat{\beta}_{1,CV} - (X_{2it} - \bar{X}_{2i})\hat{\beta}_{2,CV}$$

и получим состоятельную оценку для дисперсии случайных ошибок $D(u_{it}) = \sigma_u^2$:

$$\hat{\sigma}_u^2 = \frac{RSS_{CV}}{NT - (k_1 + k_2)}.$$

Далее по аналогии с приведенным выше заметим, что

$$\bar{y}_i - \bar{X}_{1i}\beta_1 - \bar{X}_{2i}\beta_2 = Z_{1i}\gamma_1 + Z_{2i}\gamma_2 + (\alpha_i + u_{it}).$$

Только на этот раз в правой части $E(Z_{2i}\alpha_i) \neq 0$, так что *OLS*-оценки для γ_1 и γ_2 , получаемые по этой модели (*between*-оценки), смещены и несостоятельны. Поэтому здесь для получения состоятельных оценок для γ_1 и γ_2 применяем метод инструментальных переменных (*2SLS*), используя инструменты $[Z_{1i} \ X_{1it}]$. При этом количество экзогенных переменных в X_{1it} (k_1) должно быть не меньше числа эндогенных переменных в Z_{2i} (g_2). Полученные оценки для γ_1 и γ_2 обозначим $\hat{\gamma}_{1,IV}$ и $\hat{\gamma}_{2,IV}$.

Используя все четыре полученные оценки, образуем «остатки»

$$e_{it} = y_{it} - X_{1it}\hat{\beta}_{1,CV} - X_{2it}\hat{\beta}_{2,CV} + Z_{1i}\hat{\gamma}_{1,IV} + Z_{2i}\hat{\gamma}_{2,IV}$$

и на их основе определим статистику

$$S^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T e_{it}^2,$$

которая является состоятельной оценкой для суммы $(\sigma_u^2 + T\sigma_\alpha^2)$. Тогда состоятельная оценка для σ_α^2 получается как

$$\hat{\sigma}_\alpha^2 = \frac{S^2 - \hat{\sigma}_u^2}{T}.$$

Следующим шагом является выполнение стандартного *GLS*-преобразования всех переменных, используемого в *RE*-модели:

$$y_{it}^* = y_{it} - \theta \bar{y}_i \text{ и т.п.}$$

Для реализации этого преобразования в качестве оценки параметра θ берется

$$\hat{\theta} = 1 - \sqrt{\hat{\Psi}},$$

где $\hat{\Psi} = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + T\hat{\sigma}_\alpha^2}$;

$\hat{\sigma}_u^2$ и $\hat{\sigma}_\alpha^2$ — полученные выше состоятельные оценки для σ_u^2 и σ_α^2 .

Это приводит к преобразованному уравнению

$$y_{it}^* = X_{1it}^* \beta_1 + X_{2it}^* \beta_2 + Z_{1i}^* \gamma_1 + Z_{2i}^* \gamma_2 + v_{it}^*.$$

К этому уравнению применим метод IV (*2SLS*), используя инструменты $[X_{1it} - \bar{X}_{1i}, X_{2it} - \bar{X}_{2i}, \bar{X}_{1i}, Z_{1i}]$, в результате получим для вектора $\delta = [\beta^T, \gamma^T]^T$ инструментальную оценку Хаусмана — Тейлора: $\hat{\delta}_{HT} = [\hat{\beta}_{HT}^T, \hat{\gamma}_{HT}^T]^T$

✓ Замечание 3.4.1

1. В процедуре Хаусмана — Тейлора инструменты берутся внутри самой модели.
2. X_{1it} используется как инструмент дважды: как среднее и как отклонение от среднего.
3. Если $k_1 < g_2$, то параметр γ неидентифицируем. В этом случае $\hat{\beta}_{HT} = \hat{\beta}_{CV}$ и $\hat{\gamma}_{HT}$ не существует.
4. Если $k_1 = g_2$, то $\hat{\beta}_{HT} = \hat{\beta}_{CV}$ и $\hat{\gamma}_{HT} = \hat{\gamma}_{IV} = (\hat{\gamma}_{1,IV}^T, \hat{\gamma}_{2,IV}^T)^T$ (случай точной идентифицируемости).
5. Если $k_1 > g_2$, то уравнение сверхидентифицируемо. В этом случае матрица $Cov(\hat{\beta}_{CV}) - Cov(\hat{\beta}_{HT})$ положительно определена и оценка Хаусмана — Тейлора более эффективна, чем «внутри»-оценка.

Влияние метода Хаусмана — Тейлора на прикладные исследования относительно мало из-за трудности нахождения экзогенных переменных X_1 , которые можно было бы уверенно рассматривать как не коррелированные с α_i (так что $E(X_{1it} \alpha_i) = 0$).

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Каковы особенности оценивания несбалансированных панелей?
2. Каковы особенности оценивания моделей, в которых некоторые объясняющие переменные являются эндогенными?
3. Каковы особенности оценивания моделей, в которых некоторые объясняющие переменные специфичны только в отношении субъекта?

Тема 3.5

ДИНАМИЧЕСКИЕ МОДЕЛИ

Динамическая модель: несостоятельность «внутри»-оценки

Ранее была рассмотрена двунаправленная модель, в которую помимо индивидуальных эффектов α_i включались также временные эффекты λ_t :

$$y_{it} = \mu + \alpha_i + \lambda_t + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

При этом индивидуальные эффекты α_i оставались постоянными во времени, а временные эффекты λ_t не изменялись от субъекта к субъекту. Между тем довольно часто панельные данные используются для оценивания связей, которые имеют динамический характер. В таких ситуациях естественно рассматривать модели, содержащие в правых частях уравнений запаздывающие значения объясняемой переменной; в простейшем случае таковой является динамическая модель

$$y_{it} = \gamma y_{i,t-1} + \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

в которой $|\gamma| < 1$, $u_{it} \sim i.i.d. N(0, \sigma_u^2)$ — инновации (так что $E(u_{it} y_{i,t-k}) = 0$ для $k > 0$). Будем предполагать, что значения y_{it} наблюдаются для $t = 0, 1, \dots, T$. «Внутри»-оценка (ковариационная оценка) для γ имеет вид:

$$\hat{\gamma}_{CV} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(y_{i,t-1} - \bar{y}_{i,-1})}{\sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{i,-1})^2},$$

где

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{y}_{i,-1} = \frac{1}{T} \sum_{t=1}^T y_{i,t-1}.$$

В соответствии с определением модели,

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T (\gamma y_{i,t-1} + \alpha_i + u_{it}) = \gamma \bar{y}_{i,-1} + \alpha_i + \bar{u}_i,$$

где $\bar{u}_i = \frac{1}{T} \sum_{t=1}^T u_{it}$, так что

$$y_{it} - \bar{y}_i = \gamma(y_{i,t-1} - \bar{y}_{i,-1}) + (u_{it} - \bar{u}_i)$$

и

$$\hat{\gamma}_{CV} = \gamma + \frac{\frac{\sum_{i=1}^N \sum_{t=1}^T (u_{it} - \bar{u}_i)(y_{i,t-1} - \bar{y}_{i,-1})}{NT}}{\frac{\sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{i,-1})^2}{NT}}.$$

Такая оценка существует, если знаменатель дроби в правой части последнего выражения отличен от нуля.

Можно показать (см., например, (Hsiao C., 2003, p. 72)), что предел по вероятности числителя дроби в правой части последнего выражения равен:

$$\begin{aligned} p \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (u_{it} - \bar{u}_i)(y_{i,t-1} - \bar{y}_{i,-1}) &= \\ &= p \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{i=1}^N \left(\sum_{t=1}^T u_{it} y_{i,t-1} - \bar{u}_i \sum_{t=1}^T y_{i,t-1} - \bar{y}_{i,-1} \sum_{t=1}^T u_{it} + T \bar{u}_i \bar{y}_{i,-1} \right) = \\ &= p \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (-\bar{u}_i \bar{y}_{i,-1}) = -\frac{\sigma_u^2}{T^2} \left[\frac{T-1-T\gamma+\gamma^T}{(1-\gamma)^2} \right], \end{aligned}$$

а предел по вероятности знаменателя этой дроби равен:

$$p \lim_{N \rightarrow \infty} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (y_{i,t-1} - \bar{y}_{i,-1})^2 = \frac{\sigma_u^2}{1-\gamma^2} \left[1 - \frac{1}{T} - 2\gamma \frac{(T-1)-T\gamma+\gamma^T}{T^2(1-\gamma)^2} \right].$$

Отсюда вытекает, что если не только $N \rightarrow \infty$, но и $T \rightarrow \infty$, то первый предел по вероятности стремится к нулю, а второй — к отличному от нуля значению

$$\frac{\sigma_u^2}{1-\gamma^2} \neq 0,$$

так что $p \lim_{N \rightarrow \infty} \hat{\gamma}_{CV} = \gamma$, и $\hat{\gamma}_{CV}$ является состоятельной оценкой параметра γ .

Если же значение T фиксированное, тогда первый предел не равен нулю и $p \lim_{N \rightarrow \infty} \hat{\gamma}_{CV} \neq \gamma$, так что оценка $\hat{\gamma}_{CV}$ несостоятельна.

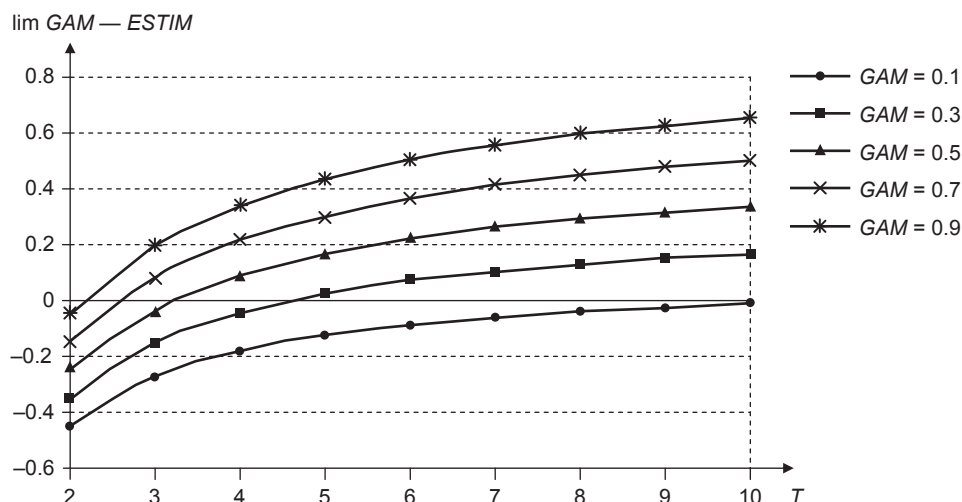


Рис. 3.6

Асимптотическое смещение оценки $\hat{\gamma}_{CV}$ является следствием «внутри»-преобразования, выметающего индивидуальные эффекты α_i из каждого наблюдения, что порождает корреляцию между остатками $(u_{it} - \bar{u}_i)$ в преобразованной модели и объясняющей переменной $(y_{i,t-1} - \bar{y}_{i,-1})$. Когда T велико, эта корреляция близка к нулю. Для малых значений T смещение отрицательно, если $\gamma > 0$, и не стремится к нулю при $\gamma \rightarrow 0$. Поскольку в типичной панели T мало, возникающее смещение трудно игнорировать. Например, для $T = 2$ асимптотическое смещение равно $-\frac{1+\gamma}{2}$. На рис. 3.6 показано, как асимптотическое смещение оценки $\hat{\gamma}_{CV}$ изменяется с ростом T .

Получение состоятельной оценки: обобщенный метод моментов

«Внутри»-оценка остается несостоятельной при малых T и когда в правую часть уравнения модели добавляются экзогенные объясняющие переменные.

Для преодоления этой проблемы можно воспользоваться другим преобразованием, выметающим α_i : вместо вычитания средних по времени перейти к первым разностям временных рядов для каждого субъекта. При этом получаем:

$$y_{i,t} - y_{i,t-1} = \gamma(y_{i,t-1} - y_{i,t-2}) + (u_{it} - u_{i,t-1}), \quad i = 1, \dots, N, \quad t = 2, \dots, T,$$

или

$$\Delta y_{i,t} = \gamma \Delta y_{i,t-1} + \Delta u_{it},$$

где обозначено:

$$\Delta y_{i,t} = y_{i,t} - y_{i,t-1}, \quad \Delta u_{it} = u_{it} - u_{i,t-1}.$$

Здесь

$$\text{Cov}(\Delta y_{i,t-1}, \Delta u_{it}) = \text{Cov}(y_{i,t-1} - y_{i,t-2}, u_{it} - u_{i,t-1}) = -\text{Cov}(y_{i,t-1}, u_{i,t-1}) \neq 0.$$

Поэтому *OLS*-оценка для γ в преобразованном («продифференцированном») уравнении оказывается несостоятельной, даже если $T \rightarrow \infty$. Однако к преобразованному уравнению можно применить метод инструментальных переменных. Для этого достаточно заметить, что если взять переменную $y_{i,t-2}$, то для нее

$$\text{Cov}(y_{i,t-2}, \Delta u_{it}) = \text{Cov}(y_{i,t-2}, u_{it} - u_{i,t-1}) = 0,$$

$$\text{Cov}(y_{i,t-2}, \Delta y_{i,t-1}) = \text{Cov}(y_{i,t-2}, y_{i,t-1} - y_{i,t-2}) \neq 0,$$

а это означает, что данная переменная может использоваться в качестве инструмента для $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$, и это приводит к оценке:

$$\hat{\gamma}_{IV} = \frac{\sum_{i=1}^N \sum_{t=2}^T (y_{i,t} - y_{i,t-1}) y_{i,t-2}}{\sum_{i=1}^N \sum_{t=2}^T (y_{i,t-1} - y_{i,t-2}) y_{i,t-2}}.$$

Необходимое условие состоятельности этой оценки:

$$p \lim_{N \rightarrow \infty} \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T (u_{it} - u_{i,t-1}) y_{i,t-2} = 0$$

при $T \rightarrow \infty$ или/и $N \rightarrow \infty$.

В качестве инструмента для $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ вместо $y_{i,t-2}$ можно использовать, например, разность $(y_{i,t-2} - y_{i,t-3})$, что приводит к другой оценке:

$$\hat{\gamma}_{IV} = \frac{\sum_{i=1}^N \sum_{t=3}^T (y_{i,t} - y_{i,t-1})(y_{i,t-2} - y_{i,t-3})}{\sum_{i=1}^N \sum_{t=3}^T (y_{i,t-1} - y_{i,t-2})(y_{i,t-2} - y_{i,t-3})},$$

для состоятельности которой необходимо, чтобы

$$p \lim_{N \rightarrow \infty} \frac{1}{N(T-2)} \sum_{i=1}^N \sum_{t=3}^T (u_{it} - u_{i,t-1})(y_{i,t-2} - y_{i,t-3}) = 0.$$

Состоятельность обеих оценок гарантируется отсутствием автокоррелированности $u_{i,t}$.

Заметим теперь, что

$$\begin{aligned} p \lim_{N \rightarrow \infty} \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=2}^T (u_{it} - u_{i,t-1}) y_{i,t-2} &= E[(u_{it} - u_{i,t-1}) y_{i,t-2}] = 0, \\ p \lim_{N \rightarrow \infty} \frac{1}{N(T-2)} \sum_{i=1}^N \sum_{t=3}^T (u_{it} - u_{i,t-1}) (y_{i,t-2} - y_{i,t-3}) &= \\ &= E[(u_{it} - u_{i,t-1}) (y_{i,t-2} - y_{i,t-3})] = 0. \end{aligned}$$

Это условия на моменты совместных распределений пар случайных величин $(u_{it} - u_{i,t-1}), y_{i,t-2}$ и $(u_{it} - u_{i,t-1}), (y_{i,t-2} - y_{i,t-3})$. Если оба эти **условия ортогональности, моментные условия** (*orthogonality conditions, moment conditions*) выполняются, то применение сразу двух инструментов приводит к повышению эффективности оценок (используется большее количество информации). Заметим, что можно найти и другие подходящие инструменты. Например, каждая из переменных $y_{i,t-2-j}$, $j = 0, 1, \dots$ удовлетворяет условиям

$$\begin{aligned} E[(u_{it} - u_{i,t-1}) y_{i,t-2-j}] &= 0, \\ E[(y_{i,t-1} - y_{i,t-2}) y_{i,t-2-j}] &\neq 0, \end{aligned}$$

так что для $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ и эти переменные годятся в качестве инструментов.

С учетом того, что у нас $t = 2, \dots, T$, можно, следуя работе Ареллано и Бонда (*Arellano, Bond, 1991*), создать список инструментов, поступая следующим образом.

Предположим, что $T = 2$, так что имеем только уравнения

$$y_{i,2} - y_{i,1} = \gamma(y_{i,1} - y_{i,0}) + (u_{i2} - u_{i,1}), \quad i = 1, \dots, N.$$

В силу сказанного выше в качестве инструмента для $\Delta y_{i,1} = y_{i,1} - y_{i,0}$ годится переменная $y_{i,0}$.

Если $T = 3$, так что для каждого $i = 1, \dots, N$ имеем 2 уравнения:

$$\begin{aligned} y_{i,2} - y_{i,1} &= \gamma(y_{i,1} - y_{i,0}) + (u_{i2} - u_{i,1}), \\ y_{i,3} - y_{i,2} &= \gamma(y_{i,2} - y_{i,1}) + (u_{i3} - u_{i,2}), \end{aligned}$$

то в качестве инструмента для $\Delta y_{i,1} = y_{i,1} - y_{i,0}$ опять годится переменная $y_{i,0}$, а в качестве инструмента для $\Delta y_{i,2} = y_{i,2} - y_{i,1}$ — как $y_{i,0}$, так и $y_{i,1}$. Соответственно соотношение

$$E[(u_{i2} - u_{i1}) y_{i0}] = 0$$

рассматривается как моментное условие для $t = 2$, а для $t = 3$ имеем пару моментных условий:

$$E[(u_{i3} - u_{i2}) y_{i1}] = 0, \quad E[(u_{i3} - u_{i2}) y_{i0}] = 0.$$

Если $T = 4$, то для $t = 2$ и $t = 3$ годятся те же условия, что и в случае $t = 3$, а для $t = 4$ — три моментных условия:

$$E[(u_{i4} - u_{i3})y_{i0}] = 0,$$

$$E[(u_{i4} - u_{i3})y_{i1}] = 0,$$

$$E[(u_{i4} - u_{i3})y_{i2}] = 0.$$

Всю эту совокупность $\frac{T(T-1)}{2}$ моментных условий можно использовать в рамках **обобщенного метода моментов** (*GMM — Generalized Method of Moments*).

Для произвольного T определим $(T-1) \times 1$ -вектор:

$$\Delta u_i = \begin{pmatrix} u_{i2} - u_{i1} \\ \vdots \\ u_{iT} - u_{i,T-1} \end{pmatrix}$$

и $(T-1) \times \frac{T(T-1)}{2}$ -матрицу:

$$Z_i = \begin{pmatrix} [y_{i0}] & 0 & \dots & 0 \\ 0 & [y_{i0}, y_{i1}] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & [y_{i0}, \dots, y_{iT-2}] \end{pmatrix}.$$

Каждая строка матрицы Z_i содержит инструменты, подходящие для соответствующего момента времени. В этих обозначениях указанная совокупность $\frac{T(T-1)}{2}$ моментных условий записывается в виде:

$$E[Z_i^T \Delta u_i] = 0.$$

Если определить еще

$$\Delta y_i = \begin{pmatrix} y_{i2} - y_{i1} \\ \vdots \\ y_{iT} - y_{i,T-1} \end{pmatrix}, \quad \Delta y_{i,-1} = \begin{pmatrix} y_{i1} - y_{i0} \\ \vdots \\ y_{i,T-1} - y_{i,T-2} \end{pmatrix},$$

то последнее соотношение записывается также в виде:

$$E[Z_i^T (\Delta y_i - \gamma \Delta y_{i,-1})] = 0.$$

В отличие от метода наименьших квадратов, количество моментных условий здесь больше числа условий, необходимых для определения с их помо-

щью значения γ , так что использование разных условий приводит к различным оценкам. Следовательно, нет возможности получить значение γ , при котором выполняются все указанные моментные условия. Вместо этого приходится ограничиваться требованием в каком-то смысле «наилучшего» приближения ко всем моментным условиям сразу. Чтобы использовать всю совокупность моментных условий, в *GMM* минимизируется квадратичная форма от выборочных аналогов моментных условий:

$$Q(\gamma) = \left[\frac{1}{N} \sum_{i=1}^N Z_i^T (\Delta y_i - \gamma \Delta y_{i,-1}) \right]^T W_N \left[\frac{1}{N} \sum_{i=1}^N Z_i^T (\Delta y_i - \gamma \Delta y_{i,-1}) \right],$$

где W_N — симметричная положительно определенная взвешивающая матрица.

Искомый минимум достигается при значении, равном

$$\hat{\gamma}_{GMM} = \left(\left(\sum_{i=1}^N \Delta y_{i,-1}^T Z_i \right) W_N \left(\sum_{i=1}^N Z_i^T \Delta y_{i,-1} \right) \right)^{-1} \times \left(\sum_{i=1}^N \Delta y_{i,-1}^T Z_i \right) W_N \left(\sum_{i=1}^N Z_i^T \Delta y_i \right).$$

Это и есть *GMM*-оценка параметра γ . Свойства этой оценки зависят от выбора взвешивающей матрицы W_N . При положительной определенности матрицы W_N — в частности, для единичной матрицы $W_N = I_N$ — *GMM*-оценка состоятельна.

Однако желательно выбирать матрицу W_N таким образом, чтобы *GMM*-оценка была по возможности наиболее эффективной, — о такой матрице говорят как об **оптимальной взвешивающей матрице** (*optimal weighting matrix*). Такая матрица должна удовлетворять условию:

$$p \lim_{N \rightarrow \infty} W_N = (\text{Cov}(Z_i^T \Delta u_i))^{-1} = [E(Z_i^T \Delta u_i (\Delta u_i)^T Z_i)]^{-1}.$$

Если на ковариационную матрицу вектора ошибок наблюдений для i -го субъекта не накладывается никаких ограничений, можно поступить следующим образом.

Сначала полагаем $W_N = I_N$ и производим *GMM*-оценивание коэффициента γ в модели $\Delta y_{i,t} = \gamma \Delta y_{i,t-1} + \Delta u_{it}$ с такой взвешивающей матрицей, определяя предварительную оценку $\hat{\gamma}^{(1)}$ для γ . При этом получаем остатки

$$\Delta \hat{u}_{it} = \Delta y_{i,t} - \hat{\gamma}^{(1)} \Delta y_{i,t-1}$$

и составляем из них векторы

$$\Delta \hat{u}_i = \begin{pmatrix} \Delta \hat{u}_{i2} \\ \vdots \\ \Delta \hat{u}_{iT} \end{pmatrix}.$$

Искомая матрица определяется после этого соотношением

$$W_N^{opt} = \left(\frac{1}{N} \sum_{i=1}^N Z_i^T \Delta \hat{u}_i (\Delta \hat{u}_i)^T Z_i \right)^{-1}.$$

Если $u_{it} \sim i.i.d.$, то положение значительно упрощается. В этом случае

$$E(\Delta u_i (\Delta u_i)^T) = \sigma_u^2 G = \sigma_u^2 \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{pmatrix},$$

и поэтому не требуется предварительного оценивания γ . Оптимальная матрица определяется соотношением

$$W_N^{opt} = \left(\frac{1}{N} \sum_{i=1}^N Z_i^T G Z_i \right)^{-1},$$

и GMM -оценивание производится за один шаг.

В общем случае GMM -оценка $\hat{\gamma}_{GMM}$ имеет асимптотически нормальное распределение с ковариационной матрицей

$$p \lim_{N \rightarrow \infty} \left(\left(\frac{1}{N} \sum_{i=1}^N \Delta y_{i,-1}^T Z_i \right) \left(\frac{1}{N} \sum_{i=1}^N Z_i^T \Delta u_i (\Delta u_i)^T Z_i \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N Z_i^T \Delta y_{i,-1} \right) \right)^{-1}.$$

Если $u_{it} \sim i.i.d.$, то средняя составляющая редуцируется к

$$\sigma_u^2 W_N^{opt} = \sigma_u^2 \left(\frac{1}{N} \sum_{i=1}^N Z_i^T G Z_i \right)^{-1}.$$

Рассмотрим теперь динамическую модель с экзогенными переменными:

$$y_{it} = \gamma y_{i,t-1} + x_{it}^T \beta + \alpha_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Здесь дифференцирование приводит к модели

$$\Delta y_{i,t} = \gamma \Delta y_{i,t-1} + \Delta x_{it}^T \beta + \Delta u_{it}.$$

Если предполагать, что все p объясняющих переменных, входящих в состав вектора x_{it} , строго экзогенны — в том смысле, что они не коррелированы с каждым u_{is} , тогда

$$E(x_{it} \Delta u_{is}) = 0 \text{ для всех } t, s,$$

так что в указанные ранее списки инструментов для каждого момента (периода) времени можно добавить x_{i1}, \dots, x_{iT} . Тогда для момента t список инструментов принимает вид: $[y_{i0}, y_{i1}, \dots, y_{i,t-2}, x_{i1}, \dots, x_{iT}]$. Такой список может быть весьма длинным, если p и T не очень малы. В то же время при строгой экзогенности x_{it} имеем также:

$$E(\Delta x_{it} \Delta u_{it}) = 0 \text{ для каждого } t,$$

так что Δx_{it} могут выступать в качестве инструментов для самих себя. При таком подходе список инструментов для момента t имеет вид: $[y_{i0}, y_{i1}, \dots, y_{i,t-2}, \Delta x_{i1}, \dots, \Delta x_{it}]$. Этот список существенно короче, если панель достаточно длинна.

Предположим теперь, что переменные в x_{it} не являются строго экзогенными, но являются *предetermined* — в том смысле, что

$$E(x_{it} u_{is}) = 0 \text{ для всех } s > t.$$

В этом случае уже не все x_{i1}, \dots, x_{iT} годятся в качестве инструментов для продифференцированного уравнения в момент t , а только $x_{i1}, \dots, x_{i,t-1}$, и соответственно накладываются моментные условия

$$E(x_{i,t-j} \Delta u_{it}) = 0 \text{ для } j = 1, \dots, t-1.$$

Разумеется, если в состав x_{i1} входят как строго эндогенные, так и predetermined переменные, то списки инструментов соответствующим образом корректируются.

✓ **Замечание 3.5.1.** Указанная выше «оптимальная» взвешивающая матрица является оптимальной в отношении выбранного множества инструментов. В то же время возникает вопрос об «оптимальном» выборе самих инструментов.

Привлечение большего количества инструментов подразумевает получение более эффективных оценок. Однако здесь возникают две опасности:

- некоторые из переменных, привлеченных в качестве инструментов, в действительности могут быть коррелированными с ошибками; для предотвращения таких ситуаций необходимо проверять гипотезу о выполнении соответствующих условий ортогональности;
- оценки коэффициентов могут иметь значительное смещение вследствие оценивания взвешивающей матрицы W_N .

Проверка гипотез о правильности спецификации динамической модели

ПРИМЕР 3.5.1

Вернемся к приведенным ранее данным об объемах инвестиций y и прибыли x трех предприятий ($N = 3$) за десятилетний период ($T = 10$) и рассмотрим на этот раз динамическую модель первого порядка

$$y_{it} = \mu + \alpha_i + \gamma y_{i,t-1} + \beta_1 x_{it} + \beta_2 x_{i,t-1} + u_{it}, \quad i = 1, 2, 3, \quad t = 2, \dots, 10.$$

Дифференцирование приводит к модели для разностей:

$$\Delta y_{it} = \gamma \Delta y_{i,t-1} + \beta_1 \Delta x_{it} + \beta_2 \Delta x_{i,t-1} + \Delta u_{it}, \quad i = 1, 2, 3, \quad t = 3, \dots, 10.$$

В программе *xtabond* в пакете *Stata 8* в качестве инструментов используются переменные, указанные в табл. 3.43, а также $\Delta x_{i3} + \dots + \Delta x_{i10}$ и $\Delta x_{i2} + \dots + \Delta x_{i9}$.

Таблица 3.43

Переменные, используемые в качестве инструментов

t	Инструменты	Количество
3	y_{i1}	1
4	y_{i1}, y_{i2}	2
5	y_{i1}, y_{i2}, y_{i3}	3
6	$y_{i1}, y_{i2}, y_{i3}, y_{i4}$	4
7	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}$	5
8	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}$	6
9	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}, y_{i7}$	7
10	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}, y_{i7}, y_{i8}$	8
ВСЕГО		36

Это дает всего 38 ($36 + 2$) моментных условий. Поскольку модель в разностях содержит 3 неизвестных коэффициента, для оценивания этих коэффициентов достаточно использовать только 3 моментных условия, а остальные 35 ($38 - 3$) условий — избыточные. Их можно было бы не использовать, но это снизило бы эффективность получаемых оценок. ■

Вместе с тем наличие избыточных условий позволяет проверять адекватность сделанных в отношении модели предположений. Точнее говоря, возникает возможность проверки гипотезы H_0 о том, что избыточные условия (выведенные на основании исходных предположений о рассматриваемой

модели) действительно выполняются. Для проверки этой гипотезы используется **статистика Саргана** (*Sargan statistic*):

$$S = NQ(\hat{\theta}_{GMM}),$$

где $\hat{\theta}_{GMM}$ — GMM-оценка вектора коэффициентов θ модели (в нашем примере $\theta = (\gamma, \beta_1, \beta_2)^T$);

$Q(\hat{\theta}_{GMM})$ — значение при $\theta = \hat{\theta}_{GMM}$ квадратичной формы, минимизируемой в методе GMM.

Если гипотеза H_0 справедлива, то статистика Саргана имеет асимптотическое (при $N \rightarrow \infty$) распределение хи-квадрат с числом степеней свободы, равным количеству избыточных моментных условий (в нашем примере оно равно 35).

Приведем теперь (табл. 3.44) результаты применения программы *xtabond* при использовании взвешивающей матрицы

$$W_N^{opt} = \left(\frac{1}{N} \sum_{i=1}^N Z_i^T G Z_i \right)^{-1}$$

(так что оценивание производится за один шаг).

Таблица 3.44

Результаты использования взвешивающей матрицы
.xtabond y x l1(x), lags(1); Arellano-Bond dynamic panel-data estimation

<i>Number of obs</i> = 24 <i>Number of groups</i> = 3 <i>Obs per group: min</i> = 8 <i>One-step results</i>				
Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z$
Δy_{-1}	0.0830295	0.2100682	0.40	0.693
Δx	1.132349	0.0606134	18.68	0.000
Δx_{-1}	-0.0423772	0.2375232	-0.18	0.858
<i>cons</i>	0.1032841	0.1361505	0.76	0.448
<i>Sargan test of over-identifying restrictions:</i> <i>chi2(35) = 21.81</i> <i>Prob > chi2 = 0.9600</i>				

Результаты применения критерия Саргана говорят в пользу гипотезы о выполнении избыточных предположений. В то же время коэффициенты при запаздывающей разности значений объясняемой переменной и запаздывающей разности объясняющей переменной статистически незначимы, что возвращает нас к статической модели регрессии.

В программе *xtabond* пакета *Stata8* используется еще один критерий проверки адекватности модели. Он основан на следующем обстоятельстве. Если ошибки u_{i1}, \dots, u_{iT} взаимно независимы, то:

- соседние разности $\Delta u_{i1}, \Delta u_{i,t-1}$ коррелированы, так как

$$\text{Corr}(\Delta u_{it}, \Delta u_{i,t-1}) = \text{Corr}(u_{it} - u_{i,t-1}, u_{i,t-1} - u_{i,t-2}) = -\sigma_u^2;$$

- отстоящие на большее количество периодов времени разности $\Delta u_{i1}, \Delta u_{i,t-s}$, $s = 2, 3, \dots$, напротив, не являются коррелированными.

Соответственно для дополнительной проверки адекватности оцененной модели проверяется наличие автокоррелированности первого порядка и отсутствие автокоррелированности второго порядка. Результаты такой проверки для только что оцененной модели выглядят следующим образом:

Arellano-Bond test that average autocovariance in residuals of order 1 is 0:

$$H_0: \text{no autocorrelation } z = -2.56, \quad Pr > z = 0.0106.$$

Arellano-Bond test that average autocovariance in residuals of order 2 is 0:

$$H_0: \text{no autocorrelation } z = 0.77, \quad Pr > z = 0.4427.$$

Полученные результаты подтверждают правильность сделанных предположений об ошибках.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какая проблема возникает при попытке использования «внутри»-оценки для оценивания коэффициентов динамической модели с фиксированными эффектами? Какое преобразование помогает обойти эту проблему?
2. В чем состоит обобщенный метод моментов? Как он используется для оценивания коэффициентов динамической модели с фиксированными эффектами?
3. Как можно проверить гипотезу о правильности спецификации динамической модели, оцененной обобщенным методом моментов?

Р а з д е л 4

МОДЕЛИ С ДИСКРЕТНЫМИ И ОГРАНИЧЕННЫМИ ОБЪЯСНЯЕМЫМИ ПЕРЕМЕННЫМИ

Тема 4.1

МОДЕЛИ, В КОТОРЫХ ОБЪЯСНЯЕМАЯ ПЕРЕМЕННАЯ ПРИНИМАЕТ ТОЛЬКО ДВА РАЗЛИЧНЫХ ЗНАЧЕНИЯ

Ситуации, когда в модели объясняемая переменная принимает только два различных значения, возникают при исследовании влияния тех или иных субъективных и объективных факторов на наличие или отсутствие некоторого признака у отдельных домашних хозяйств (наличие или отсутствие в семье автомобиля), у отдельных индивидуумов (занятый — безработный), у отдельных фирм (обанкротилась или нет в течение определенного периода) и т.п. Если исследование затрагивает n субъектов, т.е. если имеем n наблюдений, то факт наличия или отсутствия такого признака в i -м наблюдении удобно индексировать числами 1 (наличие признака) и 0 (отсутствие признака). Тем самым определяется **индикаторная (дихотомическая, бинарная — *indicator, dichotomic, binary variable*) переменная** y_i , которая принимает в i -м наблюдении значение y_i . При этом $y_i = 1$ при наличии рассматриваемого признака у i -го субъекта и $y_i = 0$ — при отсутствии рассматриваемого признака у i -го субъекта.

Если пытаться объяснить наличие или отсутствие рассматриваемого признака значениями (точнее, сочетанием значений) некоторых факторов (объясняющих переменных), то, следуя идеологии классической линейной модели, можно рассмотреть модель наблюдений

$$y_i = \theta_1 x_{i1} + \dots + \theta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

где x_{i1}, \dots, x_{ip} — значения p объясняющих переменных в i -м наблюдении;

$\theta_1, \dots, \theta_p$ — неизвестные параметры;

$\varepsilon_1, \dots, \varepsilon_n$ — случайные ошибки, отражающие влияние на наличие или отсутствие рассматриваемого признака у i -го субъекта каких-то неучтенных дополнительных факторов.

Однако попытка оценить такую модель методом наименьших квадратов наталкивается на определенные трудности.

При обычном предположении $E(\varepsilon_i | x_i) = 0$, $i = 1, \dots, n$ получаем

$$E(y_i | x_i) = \theta_1 x_{i1} + \dots + \theta_p x_{ip} = x_i^T \theta,$$

где $\theta = (\theta_1, \dots, \theta_p)^T$ — вектор-столбец (неизвестных) коэффициентов (верхний индекс T указывает на транспонирование вектора или матрицы);

$x_i^T = (x_{i1}, \dots, x_{ip})$ — вектор-строка (известных) значений объясняющих переменных в i -м наблюдении.

В то же время, поскольку y_i — случайная величина, принимающая только два значения — 0 и 1, ее условное математическое ожидание (при заданном значении x_i) равно:

$$E(y_i | x_i) = 1 \cdot P\{y_i = 1 | x_i\} + 0 \cdot P\{y_i = 0 | x_i\} = P\{y_i = 1 | x_i\}.$$

Таким образом,

$$\theta_1 x_{i1} + \dots + \theta_p x_{ip} = P\{y_i = 1 | x_i\},$$

т.е. $\theta_1 x_{i1} + \dots + \theta_p x_{ip}$ — вероятность, а значит, должно выполняться соотношение

$$0 \leq \theta_1 x_{i1} + \dots + \theta_p x_{ip} \leq 1,$$

Это первая из трудностей, с которыми сталкиваемся при обращении к таким моделям.

Далее, при $y_i = 1$ получаем $\varepsilon_i = 1 - x_i^T \theta$, а при $y_i = 0$ имеем $\varepsilon_i = -x_i^T \theta$, так что (при фиксированном x_i) ε_i может принимать в i -м наблюдении только два значения. Тогда (условные) вероятности этих значений равны:

$$P\{\varepsilon_i = 1 - x_i^T \theta | x_i\} = P\{y_i = 1 | x_i\} = x_i^T \theta,$$

$$P\{\varepsilon_i = -x_i^T \theta | x_i\} = P\{y_i = 0 | x_i\} = 1 - x_i^T \theta,$$

Соответственно случайная величина ε_i имеет условное математическое ожидание:

$$\begin{aligned} E(\varepsilon_i | x_i) &= (1 - x_i^T \theta) P\{\varepsilon_i = 1 - x_i^T \theta | x_i\} + (-x_i^T \theta) P\{\varepsilon_i = -x_i^T \theta | x_i\} = \\ &= (1 - x_i^T \theta) x_i^T \theta - x_i^T \theta (1 - x_i^T \theta) = 0 \end{aligned}$$

и условную дисперсию:

$$\begin{aligned} D(\varepsilon_i | x_i) &= E(\varepsilon_i^2 | x_i) - (E(\varepsilon_i | x_i))^2 = E(\varepsilon_i^2 | x_i) = \\ &= (1 - x_i^T \theta)^2 x_i^T \theta + (-x_i^T \theta)^2 (1 - x_i^T \theta) = \\ &= x_i^T \theta (1 - x_i^T \theta) [x_i^T \theta + (1 - x_i^T \theta)] = x_i^T \theta (1 - x_i^T \theta). \end{aligned}$$

Таким образом, здесь возникает также проблема гетероскедастичности, осложненная еще и тем, что в выражения для дисперсий ε_i входит и (неизвестный) вектор параметров θ .

Предположим, что y_i индексирует наличие или отсутствие собственного автомобиля у i -й семьи, а x_i — средний ежемесячный доход, приходящийся на каждого члена этой семьи (в условных единицах). Естественно предполагать, что вероятность наличия автомобиля увеличивается с ростом x_i . Если использовать линейную модель

$$y_i = \alpha + \beta x_i + \varepsilon_i, \quad i = 1, \dots, n,$$

то

$$E(y_i | x_i) = P\{y_i = 1 | x_i\} = \alpha + \beta x_i.$$

Таким образом, если значение x_i увеличить на 1, вероятность наличия автомобиля увеличится на величину, равную

$$(\alpha + \beta(x_i + 1)) - (\alpha + \beta x_i) = \beta,$$

независимо от того, сколь большим или малым является среднедушевой доход x_i .

Между тем такое положение вряд ли можно считать оправданным. Скорее, можно предположить, что для семей с малыми доходами наличие автомобиля — большая редкость, и некоторое увеличение среднедушевого дохода лишь ненамного увеличит вероятность приобретения автомобиля такой семьей. Для семей с весьма высокими доходами возрастание вероятности наличия автомобиля также не может быть существенным, поскольку такие семьи, как правило, уже имеют автомобиль. Большее влияние увеличения дохода на возрастание вероятности наличия автомобиля должно наблюдаться для семей со средними доходами, т.е. в «переходной зоне» от доходов, *еще* не позволяющих обзавестись собственным автомобилем, к доходам, *уже* обеспечившим возможность приобретения собственного автомобиля.

Возьмем прямоугольную систему координат, в которой по оси абсцисс будем откладывать размеры среднедушевых семейных доходов. Пусть

$$x_{(1)} = \min\{x_1, \dots, x_n\}, \quad x_{(n)} = \max\{x_1, \dots, x_n\},$$

так что $x_{(1)} \leq x \leq x_{(n)}$ — интервал значений среднедушевых доходов рассматриваемых семей. Разобьем этот интервал на некоторое количество m подынтер-

валов одинаковой длины $l = \frac{x_{(n)} - x_{(1)}}{m}$. Над каждым таким подынтервалом

построим прямоугольник, нижнее основание которого совпадает с этим подынтервалом. Пусть в пределы j -го подынтервала ($j = 1, \dots, m$) попадают среднечисловые доходы n_j семей, и при этом лишь у $n_{j,1}$ из этих семей имеется автомобиль. (Для определенности значения x_i , лежащие на границе двух соседних подынтервалов, будем относить к подынтервалу, расположенному левее.) Тогда высоту прямоугольника, построенного над j -м подынтервалом, положим равной

$$h_j = \frac{n_{j,1}}{n_j}.$$

При этом предполагаем, что общее количество рассматриваемых семей n достаточно велико, так что можно взять не слишком малое количество подынтервалов m и при этом все еще иметь достаточное количество значений x_i в каждом подынтервале.

Построим теперь ломаную с концами в точках $(x_{(1)}, 0)$ и $(x_{(n)}, 1)$, узлы которой совпадают с серединами верхних сторон построенных прямоугольников. Эта ломаная является графиком некоторой кусочно-линейной функции $G_n(x)$. И если $P\{y_i = 1 | x_i = x\} = G(x)$, то функция $G_n(x)$ в какой-то мере «оценивает» функцию $G(x)$. Правда, если функцию $G(x)$ естественно считать неубывающей (возрастающей) по x , то в силу случайных причин функция $G_n(x)$ вполне может иметь и участки убывания. Тем не менее при большом количестве наблюдений и достаточном количестве подынтервалов график функции $G_n(x)$ отражает в общих чертах форму «истинной» функции $G(x)$, так что по поведению функции $G_n(x)$ можно судить о совместимости или о несовместимости линейной модели с данными наблюдений.

Рассмотрим (искусственно смоделированную) выборку, состоящую из 1000 семей со среднечисловыми месячными доходами от 100 до 2100 условных единиц, среди которых 510 семей имеют собственный автомобиль.

Построенная по этим данным ломаная (график функции $G_n(x)$) (рис. 4.1) указывает на то, что «истинная» функция $G(x)$ имеет, скорее, не линейную, а S-образную форму.

Если тем не менее исходить из линейной модели наблюдений, то метод наименьших квадратов дает для параметров такой модели следующие оценки: $\hat{\alpha} = -0.237628$, $\hat{\beta} = 0.000680$, так что условная вероятность $P\{y_i = 1 | x_i\}$ оценивается как

$$\hat{P}\{y_i = 1 | x_i\} = -0.237628 + 0.000680x_i.$$

При $x_i \leq 349$ правая часть принимает отрицательные значения, а при $x_i \geq 1821$ — значения, превышающие 1, что выходит за пределы интервала возможных значений вероятности.

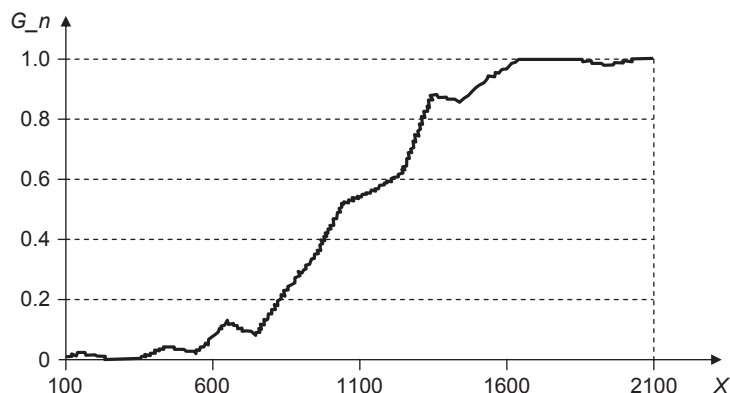


Рис. 4.1

Заметим, что в число функций, имеющих S -образную форму и значения в пределах от 0 до 1, входит целый ряд функций распределения, используемых в теории вероятностей и математической статистике (например, нормальные функции распределения).

Если использовать функцию нормального распределения $N(\mu, \sigma^2)$, имеющего математическое ожидание μ и дисперсию σ^2 , то

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{(z-\mu)^2}{2\sigma^2}\right\} dz.$$

Замена переменной $\frac{z-\mu}{\sigma} = t$ приводит это соотношение к виду:

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-t^2/2} dt = \Phi\left(\frac{x-\mu}{\sigma}\right),$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$ — функция стандартного нормального распределения $N(0, 1)$, математическое ожидание которого равно 0, а дисперсия равна 1.

Соотношение $G(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ можно записать также в виде:

$$G(x) = \Phi(\alpha + \beta x),$$

где $\alpha = -\frac{\mu}{\sigma}$, $\beta = \frac{1}{\sigma}$.

Таким образом, используя для аппроксимации $G(x)$ функцию нормального распределения, приходим к модели

$$y_i = \Phi(\alpha + \beta x_i) + \varepsilon_i, \quad i = 1, \dots, n.$$

Оценив параметры α и β этой модели, получим также оценки параметров функции нормального распределения, аппроксимирующей функцию $G(x)$:

$$\hat{\mu} = -\frac{\hat{\alpha}}{\hat{\beta}}, \quad \hat{\sigma} = \frac{1}{\hat{\beta}}.$$

Проблема, однако, в том, каким образом производить оценивание.

Заметим, что функция $G(x) = \Phi(\alpha + \beta x)$ нелинейна по параметрам, так что здесь имеем дело с нелинейной моделью регрессии. Следуя принципу наименьших квадратов, для получения оценок $\hat{\alpha}$ и $\hat{\beta}$ надо минимизировать по α и β сумму квадратов

$$Q(\alpha, \beta) = \sum_{i=1}^n (y_i - \Phi(\alpha + \beta x_i))^2.$$

Однако, в отличие от линейной модели, здесь нормальные уравнения нелинейны, не имеют решения в явном виде, и для получения приближенных значений оценок $\hat{\alpha}$ и $\hat{\beta}$ приходится использовать итерационные процедуры. Как и в рассмотренном ранее случае линейной модели, здесь возникает также проблема гетероскедастичности: условные дисперсии ошибок равны:

$$D(\varepsilon_i | x_i) = \Phi(\alpha + \beta x_i)(1 - \Phi(\alpha + \beta x_i)).$$

Соответственно для учета различия этих дисперсий при разных i следует использовать взвешенный метод наименьших квадратов, т.е. минимизировать по α и β сумму квадратов

$$Q(\alpha, \beta) = \sum_{i=1}^n w_i (y_i - \Phi(\alpha + \beta x_i))^2,$$

где веса w_i определяются соотношением

$$w_i = \frac{1}{D(\varepsilon_i | x_i)} = [\Phi(\alpha + \beta x_i)(1 - \Phi(\alpha + \beta x_i))]^{-1}.$$

К сожалению, эти веса зависят не только от x_i , но и от значений параметров α и β , которые нам не известны и которые как раз и подлежат оцениванию. Поэтому для реализации итерационной процедуры оценивания необходимы некоторые начальные оценки весов \hat{w}_i^0 , $i = 1, \dots, n$, а для этого, в свою очередь, необходимы начальные оценки \hat{G}_i^0 значений $G_i = G(x_i) = \Phi(\alpha + \beta x_i)$, которые дали бы оценки весов в виде

$$\hat{w}_i^0 = [\hat{G}_i^0(1 - \hat{G}_i^0)]^{-1}.$$

Поскольку у нас $y_i = 0$ или $y_i = 1$, то единственная разумная возможность — положить $\hat{G}_i^0 = 1$, если $y_i = 1$, и $\hat{G}_i^0 = 0$, если $y_i = 0$. Однако в обоих случаях вес \hat{w}_i^0 не определен (знаменатель равен нулю).

Ввиду отмеченных выше трудностей применения метода наименьших квадратов к рассмотренным моделям используем альтернативный метод оценивания, широко распространенный в прикладных исследованиях, а именно — метод максимального правдоподобия.

Однако, прежде чем перейти к этому методу, следует заметить, что в качестве объясняющих факторов в моделях рассмотренного типа могут выступать несколько переменных, и тогда получим модель вида:

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i, \quad i = 1, \dots, n,$$

которую обычно называют **моделью бинарного выбора** (*binary choice model*).

Использование метода максимального правдоподобия для оценивания моделей бинарного выбора

Итак, пусть наша задача состоит в оценивании параметров модели бинарного выбора:

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i, \quad i = 1, \dots, n,$$

где $G(z)$ — S -образная функция распределения, имеющего плотность $g(z) = G'(z)$.

В соответствии с введенными выше обозначениями $\theta_1 x_{i1} + \dots + \theta_p x_{ip} = x_i^T \theta$, так что $G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) = G(x_i^T \theta)$. Предположим, что при фиксированных значениях объясняющих переменных в n наблюдениях, что соответствует фиксированным значениям векторов x_i , случайные ошибки $\varepsilon_1, \dots, \varepsilon_n$ статистически независимы и $E(\varepsilon_i | x_i) = 0$, так что $P\{y_i = 1 | x_i\} = E(y_i | x_i) = G(x_i^T \theta)$. Тогда при фиксированных x_i статистически независимы и случайные величины $G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i$, $i = 1, \dots, n$, т.е. статистически независимы y_1, \dots, y_n . В силу этого (условная при фиксированных x_i , $i = 1, \dots, n$) совместная вероятность получения конкретного набора наблюдений y_1, \dots, y_n (конкретного набора нулей и единиц) равна произведению:

$$\prod_{i=1}^n (P\{y_i = 1 | x_i\})^{y_i} (P\{y_i = 0 | x_i\})^{1-y_i} = \prod_{i=1}^n (G(x_i^T \theta))^{y_i} (1 - G(x_i^T \theta))^{1-y_i}.$$

Правая часть этого выражения при фиксированных x_i , $i = 1, \dots, n$ является функцией от вектора неизвестных параметров θ :

$$L(\theta) = L(\theta | x_1, \dots, x_n) = \prod_{i=1}^n (G(x_i^T \theta))^{y_i} (1 - G(x_i^T \theta))^{1-y_i}$$

и интерпретируется как функция правдоподобия параметров $\theta_1, \dots, \theta_p$. При различных наборах значений $\theta_1, \dots, \theta_p$ получаются различные $L(\theta)$, т.е. при фиксированных x_i , $i = 1, \dots, n$, вероятность наблюдать конкретный набор зна-

чений y_1, \dots, y_n может быть более высокой или более низкой в зависимости от значения θ . Метод максимального правдоподобия предлагает в качестве оценки вектора параметров θ использовать значение $\theta = \hat{\theta}$, максимизирующее функцию правдоподобия, так что

$$L(\hat{\theta}) = \max_{\theta} L(\hat{\theta}) = \max_{\theta} \prod_{i=1}^n (G(x_i^T \theta))^{y_i} (1 - G(x_i^T \theta))^{1-y_i}.$$

Опираясь на свойство монотонного возрастания функции $\ln(z)$, то же значение $\hat{\theta}$ можно найти, максимизируя логарифмическую функцию правдоподобия $\ln L(\theta)$. В нашем случае

$$\ln L(\theta) = \sum_{i=1}^n y_i \ln G(x_i^T \theta) + \sum_{i=1}^n (1 - y_i) \ln(1 - G(x_i^T \theta)).$$

Не будем углубляться в технические детали соответствующих процедур максимизации — такие процедуры «встроены» во многие прикладные пакеты статистических программ для персональных компьютеров, и читатель при необходимости может ими воспользоваться. Заметим только: если не имеет место чистая мультиколлинеарность объясняющих переменных (т.е. если матрица $X = (x_{ij})$ значений p объясняющих переменных в n наблюдениях имеет ранг p , так что ее столбцы линейно независимы), то функция $L(\theta)$ имеет единственный локальный максимум, являющийся и глобальным максимумом, что гарантирует сходимость соответствующих итерационных процедур к оценке максимального правдоподобия.

Рассмотрим результаты применения метода максимального правдоподобия для оценивания параметров α и β моделей

$$y_i = G(\alpha + \beta x_i) + \varepsilon_i, \quad i = 1, \dots, n,$$

по упомянутым выше смоделированным данным. При этом используем предусмотренную в пакете *Econometric Views (EViews)* возможность выбора в качестве $G(z)$ следующих функций:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt \text{ — функция стандартного нормального распределения } N(0, 1) \text{ (**пробит-модель** — } probit \text{ model});$$

$$\Lambda(z) = \frac{e^z}{1 + e^z} \text{ — функция стандартного логистического распределения (**логит-модель** — } logit \text{ model);}$$

$$G(z) = 1 - \exp(-e^z) \text{ — функция стандартного распределения экстремальных значений (минимума) I типа (**распределение Гомпертца, гомпит-модель** — } gompit \text{ model).}$$

Заметим, что функции плотности первых двух распределений являются четными (графики этих плотностей симметричны относительно оси ординат),

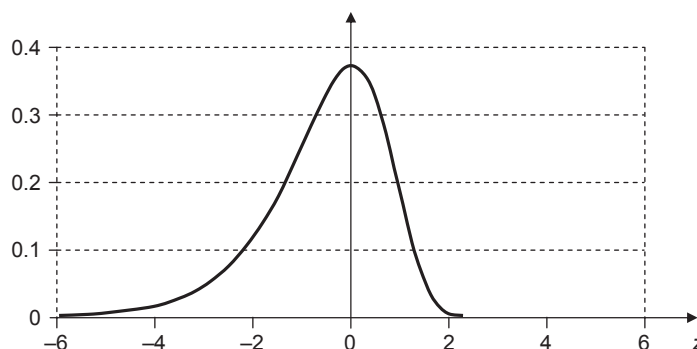


Рис. 4.2

тогда как функция плотности последнего из трех распределений не обладает таким свойством. Ее график асимметричен и скошен в сторону отрицательных значений аргумента (рис. 4.2).

Результаты оценивания указанных трех моделей по смоделированным данным (1000 наблюдений) с использованием пакета *EViews* приведены в табл. 4.1¹.

Таблица 4.1

Оценка моделей по смоделированным данным (1000 наблюдений)

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
<i>Пробит-модель</i>				
<i>C</i>	-3.503812	0.200637	-17.46343	0.0000
<i>X</i>	0.003254	0.000178	18.25529	0.0000
<i>Логит-модель</i>				
<i>C</i>	-6.357013	0.411837	-15.43576	0.0000
<i>X</i>	0.005892	0.000368	16.01461	0.0000
<i>Гомпит-модель</i>				
<i>C</i>	-3.022612	0.162178	-18.63764	0.0000
<i>X</i>	0.003344	0.000168	19.93322	0.0000

Полученные значения оценок параметров α и β в первой модели ($\hat{\alpha} = -3.503812$, $\hat{\beta} = 0.003254$) соответствуют оценкам $\hat{\mu} = 1076.77$ и $\hat{\sigma} = 307.31$ параметров функции нормального распределения, сглаживающей построенную ранее

¹ В четвертой графе приведены значения отношений оценок коэффициентов к стандартным ошибкам, рассчитанным по асимптотическому нормальному распределению оценок максимального правдоподобия. В связи с этим здесь и в последующих таблицах указанное отношение называется не *t*-статистикой, а *z*-статистикой. *P*-значения, приводимые в пятой графе, соответствуют стандартному нормальному распределению.

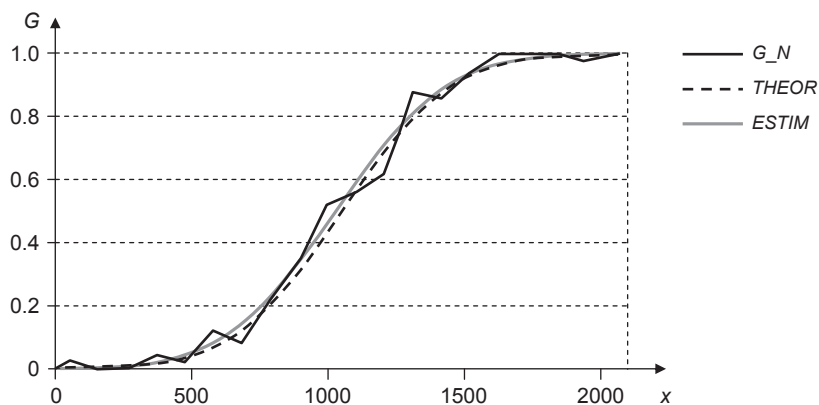


Рис. 4.3

функцию $G_n(x)$, график которой представляет собой ломаную. Заметим, что в действительности при моделировании данных использовалась в качестве $G(x)$ функция нормального распределения с параметрами $\mu = 1100$ и $\sigma = 300$. Рисунок 4.3 позволяет сравнить поведение:

- кусочно-линейной функции $G_n(x)$;
- теоретической функции $G(x)$, соответствующей нормальному распределению $N(1100, 300^2)$;
- оцененной функции $\hat{G}(x)$, соответствующей нормальному распределению $N(1076.77, 307.31^2)$.

На рис. 4.4 для сравнения добавлены оцененные функции $\hat{G}(x)$ для логит- и гомпит-моделей. Кривые, получаемые по пробит- и логит-моделям, очень мало отличаются как друг от друга, так и от теоретической кривой. В то же

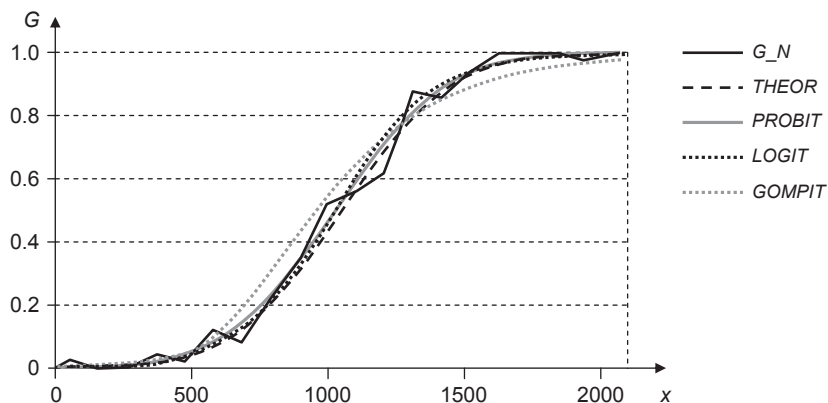


Рис. 4.4

время кривая, полученная по гомпит-модели, представляется менее удовлетворительной. Разумеется, хотелось бы иметь некоторые количественные критерии для сравнения разных моделей и для проверки адекватности каждой из рассматриваемых моделей данным наблюдений. Займемся этой проблемой.

Показатели качества модели, критерии согласия, сравнение альтернативных моделей

Прежде всего обратим внимание на следующее обстоятельство. Пусть методом наименьших квадратов оценивается обычная линейная модель

$$y_i = \theta_1 x_{i1} + \dots + \theta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

с $x_{i1} \equiv 1$ (модель с константой), в которой объясняемая переменная y может принимать непрерывный ряд значений. В таком случае простейшим показателем качества оцененной модели является коэффициент детерминации R^2 :

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

$$\text{где } \hat{y}_i = \hat{\theta}_1 x_{i1} + \dots + \hat{\theta}_p x_{ip}, \quad \bar{y} = \frac{y_1 + \dots + y_n}{n};$$

TSS — «полная» сумма квадратов;

RSS — «остаточная» сумма квадратов.

Если оценивать *тривиальную модель*, в правую часть которой включается единственная объясняющая переменная $x_{i1} \equiv 1$, т.е. модель

$$y_i = \theta_1 + \varepsilon_i, \quad i = 1, \dots, n,$$

то для такой модели $\hat{\theta}_1 = \bar{y}$, $\hat{y}_i = \hat{\theta}_1 = \bar{y}$, так что $RSS = 0$ и $R^2 = 0$. При добавлении в правую часть модели дополнительных объясняющих переменных коэффициент R^2 возрастает, и этот коэффициент будет тем больше, чем более выраженной является линейная связь объясняемой переменной с совокупностью объясняющих переменных, включенных в правую часть. Своего максимального значения ($R^2 = 1$) коэффициент детерминации достигает в предельном случае, когда для всех $i = 1, \dots, n$ выполняются точные соотношения

$$y_i = \theta_1 x_{i1} + \dots + \theta_p x_{ip}.$$

Поскольку теперь имеем дело с нелинейными моделями

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i, \quad i = 1, \dots, n,$$

не можем пользоваться обычным коэффициентом детерминации R^2 . В этом случае желательно определить какую-то другую меру качества подобранной модели.

Одна из имеющихся возможностей в этом отношении — сравнение количеств неправильных предсказаний, получаемых по выбранной модели и по модели, в которой в качестве единственной объясняющей переменной выступает константа (тривиальная модель).

Естественным представляется при $G(x_i^T \hat{\theta}) > 1/2$ предсказывать значение $y_i = 1$. Для симметричных распределений это равносильно условию $x_i^T \hat{\theta} > 0$, так что прогнозные значения равны:

$$\hat{y}_i = \begin{cases} 1, & \text{если } x_i^T \hat{\theta} > 0, \\ 0, & \text{если } x_i^T \hat{\theta} \leq 0. \end{cases}$$

Количество неправильных предсказаний по выбранной модели равно:

$$n_{\text{wrong},1} = \sum_{i=1}^n |y_i - \hat{y}_i| = \sum_{i=1}^n (y_i - \hat{y}_i)^2;$$

доля неправильных предсказаний по выбранной модели равна:

$$v_{\text{wrong},1} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

В то же время, если рассмотреть тривиальную модель, то для нее значение $y_i = 1$ предсказывается для всех $i = 1, \dots, n$, когда $G(\hat{\theta}_1) > 1/2$, т.е. когда $\bar{y} > 1/2$ (значения $y_i = 1$ наблюдаются более чем в половине наблюдений). Соответственно значение $y_i = 0$ предсказывается для всех $i = 1, \dots, n$, когда $G(\hat{\theta}_1) \leq 1/2$, т.е. когда $\bar{y} \leq 1/2$ (значения $y_i = 1$ наблюдаются не более чем в половине наблюдений). При этом доля неправильных предсказаний по тривиальной модели равна:

$$v_{\text{wrong},0} = \begin{cases} 1 - \bar{y}, & \text{если } \bar{y} > 1/2, \\ \bar{y}, & \text{если } \bar{y} \leq 1/2. \end{cases}$$

За показатель качества модели можно было бы взять коэффициент

$$R_{\text{predict}}^2 = 1 - \frac{v_{\text{wrong},1}}{v_{\text{wrong},0}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

Проблема, однако, в том, что выбранная модель может дать предсказание хуже, чем тривиальная, так что $v_{\text{wrong},1} > v_{\text{wrong},0}$, и тогда $R_{\text{predict}}^2 < 0$. Отметим также, что вообще $v_{\text{wrong},0} \leq 0.5$, так что тривиальная модель может неправильно предсказать не более половины наблюдений. А если оказывается, что

в выборке значения y_i равны 1 для 90% наблюдений, тогда $v_{wrong,0} = 0.1$, и, чтобы получить $R^2_{predict} > 0$, необходимо, чтобы альтернативная модель давала более 90% правильных предсказаний. Это означает, что большая доля правильных предсказаний $1 - v_{wrong,1}$ сама по себе не говорит еще о качестве модели. Эта доля может быть большой и для плохой модели.

Рассмотрим теперь альтернативный подход к построению аналога коэффициента R^2 для моделей бинарного выбора. Поскольку для оценивания таких моделей мы использовали метод максимального правдоподобия, то естественным представляется сравнение максимумов функций правдоподобия (или максимумов логарифмических функций правдоподобия) для выбранной и тривиальной моделей.

Пусть L_1 — максимум функции правдоподобия для выбранной модели, а L_0 — максимум функции правдоподобия для тривиальной модели. Заметим, что при этом $L_0 \leq L_1 \leq 1$, так что и $\ln L_0 \leq \ln L_1 \leq 0$. В рамках этого подхода среди множества других были предложены следующие показатели качества моделей бинарного выбора:

$$pseudoR^2 = 1 - \frac{1}{1 + \frac{2(\ln L_1 - \ln L_0)}{n}},$$

$$McFaddenR^2 = 1 - \frac{\ln L_1}{\ln L_0}.$$

Последний показатель часто обозначают как **LRI** — **индекс отношения правдоподобий** (*likelihood ratio index*).

Оба показателя изменяются в пределах от 0 до 1. Если для выбранной модели $\hat{\theta}_1 = \dots = \hat{\theta}_p = 0$, то $L_0 = L_1$, и оба показателя равны 0. Второй показатель может оказаться равным 1, если $\ln L_1 = 0$, т.е. $L_1 = 1$. Такая модель дает точное предсказание, так что $\hat{y}_i = y_i$ для всех $i = 1, \dots, n$. Но при этом для рассмотренных выше моделей (пробит, логит и гомпит) оказывается невозможным доведение до конца итерационной процедуры оценивания вектора параметров θ из-за взрывного возрастания абсолютной величины $x_i^T \theta$ в процессе итераций. Это связано с тем, что у таких моделей при конечных значениях $x_i^T \theta$ выполняются строгие неравенства $0 < G(x_i^T \theta) < 1$, поэтому функция правдоподобия не может достигать значения 1.

ПРИМЕР 4.1.1

Продолжая начатый выше статистический анализ смоделированного множества данных, вычислим значения альтернативных вариантов коэффициента R^2 для трех оцененных моделей бинарного выбора. Величины, необходимые для вычисления этих значений, приведены в табл. 4.2. (Напомним: в смоделирован-

ной выборке количество семей, имеющих собственный автомобиль, равно 510, что составляет более половины семей. Поэтому тривиальная модель дает для всех 1000 наблюдений прогноз $y_i = 1$, что приводит к 49% ошибок.)

Таблица 4.2

Данные для моделей бинарного выбора

Модель	$v_{wrong,1}$	$\ln L_1$
Пробит	0.125	-275.7686
Логит	0.124	-275.4592
Гомпит	0.121	-292.6808
	$v_{wrong,0}$	$\ln L_0$
Тривиальная	0.490	-692.9472

Соответственно для различных вариантов коэффициента R^2 получаем:

пробит-модель

$$R^2_{predict} = 1 - \frac{v_{wrong,1}}{v_{wrong,0}} = 1 - \frac{0.125}{0.490} = 0.745,$$

$$pseudoR^2 = 1 - \frac{1}{1 + \frac{2(\ln L_1 - \ln L_0)}{n}} = 1 - \frac{1}{1 + \frac{2(-275.7686 + 692.9472)}{1000}} = 0.4548,$$

$$McFaddenR^2 = 1 - \frac{\ln L_1}{\ln L_0} = 1 - \frac{-275.7686}{-692.9472} = 0.6020;$$

логит-модель

$$R^2_{predict} = 1 - \frac{v_{wrong,1}}{v_{wrong,0}} = 1 - \frac{0.124}{0.490} = 0.7470,$$

$$pseudoR^2 = 1 - \frac{1}{1 + \frac{2(\ln L_1 - \ln L_0)}{n}} = 1 - \frac{1}{1 + \frac{2(-275.4592 + 692.9472)}{1000}} = 0.4550,$$

$$McFaddenR^2 = 1 - \frac{\ln L_1}{\ln L_0} = 1 - \frac{-275.4592}{-692.9472} = 0.6025;$$

гомпит-модель

$$R^2_{predict} = 1 - \frac{v_{wrong,1}}{v_{wrong,0}} = 1 - \frac{0.121}{0.490} = 0.7531,$$

$$pseudoR^2 = 1 - \frac{1}{1 + \frac{2(\ln L_1 - \ln L_0)}{n}} = 1 - \frac{1}{1 + \frac{2(-292.6808 + 692.9472)}{1000}} = 0.4446,$$

$$McFaddenR^2 = 1 - \frac{\ln L_1}{\ln L_0} = 1 - \frac{-275.4592}{-692.9472} = 0.5776.$$

Сведем полученные значения в общую таблицу (табл. 4.3).

Таблица 4.3

Модели бинарного выбора: разные варианты коэффициента R^2

Модель	$R^2_{predict}$	$pseudoR^2$	$McFaddenR^2$
Пробит	0.7450	0.4548	0.6020
Логит	0.7470	0.4550	0.6025
Гомпит	0.7531	0.4446	0.5776

Отметим близость всех вариантов коэффициента R^2 для пробит- и логит-моделей. Гомпит-модель дает несколько лучшее предсказание, в то время как логит-модель несколько лучше двух других с точки зрения коэффициентов $pseudoR^2$ и $McFaddenR^2$.

Представим теперь, что в нашем примере вместо смоделированных значений y_i наблюдались бы следующие значения:

$$y_i = 0 \text{ для } x_i \leq 1100, \quad y_i = 1 \text{ для } x_i > 1100.$$

Тогда 100%-е точное предсказание этих значений дала бы модель

$$P\{y_i = 1\} = \begin{cases} 0, & \text{если } x_i \leq 1100, \\ 1, & \text{если } x_i > 1100. \end{cases}$$

Вместе с тем в рамках пробит-, логит- и гомпит-моделей оценки максимального правдоподобия в такой ситуации не определены, так как максимум функции правдоподобия не достигается при конечных значениях параметров. ■

Как и в случае обычных линейных моделей, сравнивать качество нескольких альтернативных моделей бинарного выбора с разным количеством объясняющих переменных можно, опираясь на значения информационных критериев Акаике (AIC) и Шварца (SC):

$$AIC = -\frac{2\ln L_k}{n} + \frac{2p}{n}, \quad SC = -\frac{2\ln L_k}{n} + \frac{p \ln n}{n},$$

а также информационного критерия Хеннана — Куинна:

$$HQ = -\frac{2\ln L_k}{n} + \frac{2p \ln(\ln n)}{n},$$

где L_k — максимальное значение функции правдоподобия для k -й из альтернативных моделей;

p — количество объясняющих переменных в этой модели.

При этом среди нескольких альтернативных моделей выбирается та, которая минимизирует значение статистики критерия. Заметим, что эти три критерия различаются размерами «штрафа», который приходится платить за включение в модель большего количества объясняющих переменных.

В рассмотренном выше примере во всех трех моделях использовались одни и те же объясняющие переменные (константа и среднедушевой доход семьи), так что по каждому информационному критерию в качестве наилучшей будет выбрана модель, для которой максимум функции правдоподобия наибольший. В табл. 4.4 приведены полученные при оценивании значения информационных критериев.

Таблица 4.4

Информационные критерии для моделей бинарного выбора

Модель	AIC	SC	HQ
Пробит	0.555537	0.565353	0.559268
Логит	0.554918	0.564734	0.558649
Гомпит	0.589362	0.599177	0.593092

По всем трем критериям наилучшей признается логит-модель. Она имеет наибольший среди трех моделей максимум функции правдоподобия. Вместе с тем отметим, что преимущество логит-модели над пробит-моделью весьма мало.

Для проверки адекватности подобранной модели имеющимся данным существует ряд статистических критериев согласия, одним из них является **критерий Хосмера — Лемешоу** (*Hosmer-Lemeshow test*)¹. Не будем описывать его детально, воспользуемся тем, что этот критерий реализован в некоторых пакетах статистического анализа, в том числе в пакете *EViews*. Отметим только, что этот критерий основан на сравнении количеств предсказываемых моделью и действительно наблюдаемых случаев с $y_i = 1$ в нескольких группах, на которые разбивается множество наблюдений.

Сопоставим результаты применения критерия Хосмера — Лемешоу к подобранным выше моделям бинарного выбора. В табл. 4.5 приведены P -значения, соответствующие статистике Хосмера — Лемешоу (рассчитанные по асим-

¹ Подробнее об этом критерии см., например, в (*Hosmer, Lemeshow, 1989*).

птотическому распределению хи-квадрат с соответствующим числом степеней свободы) при разбиении множества наблюдений на 10 групп. Если ориентироваться на эти P -значения, то гомпит-модель следует признать неудовлетворительной.

Таблица 4.5

Критерий Хосмера — Лемешоу в моделях бинарного выбора

Модель	Пробит	Логит	Гомпит
P -значение	0.1509	0.5511	0.0000

В заключение рассмотрим пример подбора модели бинарного выбора с несколькими объясняющими переменными. В этом примере фигурируют следующие финансовые показатели 66 фирм на конец одного и того же года:

$$X_1 = \frac{\text{Оборотный капитал}}{\text{Общая сумма активов}},$$

$$X_2 = \frac{\text{Нераспределенная прибыль}}{\text{Общая сумма активов}},$$

$$X_3 = \frac{\text{Доходы до вычета процентов и налогов}}{\text{Общая сумма активов}},$$

$$X_4 = \frac{\text{Рыночная стоимость активов за вычетом задолженности}}{\text{Балансовая стоимость общей суммы обязательств}},$$

$$X_5 = \frac{\text{Объем продаж}}{\text{Общая сумма активов}}.$$

В течение последующих двух лет половина из этих фирм обанкротилась. Фирмы занумерованы от 1 до 66 так, что первые 33 фирмы в этом списке обанкротились. Введем индикаторную переменную y_i , полагая

$$y_i = \begin{cases} 0 & \text{для } i = 1, \dots, 33, \\ 1 & \text{для } i = 34, \dots, 66, \end{cases}$$

т.е. $y_i = 1$, если фирма выжила в течение двух лет.

Попробуем сначала подобрать к указанным данным пробит-модель

$$y_i = \Phi(\alpha + \beta_1 x_{i1} + \dots + \beta_5 x_{i5}) + \varepsilon_i, \quad i = 1, \dots, 66.$$

При попытке оценить параметры такой модели наталкиваемся на указанное ранее затруднение, которое связано с расходимостью итерационного процесса. Поэтому приходится отказаться от желания включить в правую часть модели сразу все имеющиеся в распоряжении показатели и перейти к рассмотрению редуцированных моделей.

При оценивании большинства моделей, в которых используются только 4 из 5 финансовых показателей, опять наталкиваемся на ту же проблему. Итерационный процесс сходится только для двух моделей, включающих в качестве объясняющих переменных (помимо константы) наборы показателей (X_1, X_2, X_4, X_5) и (X_1, X_3, X_4, X_5) соответственно. Однако каждый из оцененных коэффициентов этих моделей имеет P -значение, превышающее 0.10, что указывает на необходимость дальнейшей редукции моделей.

Среди моделей, использующих только 3 финансовых показателя, лучшей по $McFaddenR^2(LRI)$ является модель с набором объясняющих переменных $(1, X_2, X_4, X_5)$, но и в ней все оцененные коэффициенты имеют P -значения, превышающие 0.184.

Вообще множество моделей, в которых оценки коэффициентов при всех включенных в их правые части финансовых показателей статистически значимы (при 5%-м пороге), исчерпывается 6 моделями, которые в качестве объясняющих переменных имеют наборы

$$(1, X_1, X_4), (1, X_3, X_4), \\ (1, X_1), (1, X_2), (1, X_3), (1, X_4).$$

В табл. 4.6 приведены результаты, характеризующие сравнительное качество этих моделей. В первой графе указаны финансовые показатели, включенные в модель.

Таблица 4.6

Характеристика финансовых показателей моделей

Набор переменных	LRI	AIC	SC	HQ	Количество неправильных предсказаний	Хосмер — Лемешоу (5 групп) P -значения
X_1, X_4	0.645	0.582	0.682	0.621	6	0.4955
X_3, X_4	0.785	0.389	0.488	0.427	3	0.6499
X_1	0.441	0.835	0.902	0.861	12	0.4820
X_2	0.829	0.298	0.364	0.324	3	0.6916
X_3	0.668	0.520	0.587	0.547	7	0.0525
X_4	0.460	0.809	0.875	0.835	10	0.0004

Критерий Хосмера — Лемешоу признает неадекватной последнюю модель и близкой к неадекватной предпоследнюю модель. Среди остальных 4 моделей по всем показателям лучшей оказывается модель, использующая единственный финансовый показатель X_2 . Она дает следующую оценку вероятности выживания фирмы:

$$\hat{P}\{y_i = 1 | x_i\} = \Phi(-0.6625 + 0.0987x_{i2}).$$

Оцененная модель правильно предсказывает банкротство 31 из 33 обанкротившихся и выживание 32 из 33 выживших фирм. Это соответствует 95.45% правильных предсказаний, тогда как тривиальная модель дает в данном случае только 50% правильных предсказаний.

Таким образом, согласно полученным результатам вероятность выживания фирмы определяется в основном отношением размера нераспределенной прибыли к общей сумме активов фирмы и возрастает с ростом этого отношения.

Интерпретация коэффициентов

Поскольку модели логит, пробит и гомпит являются нелинейными, оцененные коэффициенты в этих моделях имеют интерпретацию, отличающуюся от интерпретации коэффициентов в линейной модели.

Все эти модели имеют вид:

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

при этом

$$P\{y_i = 1 | x_i\} = E(y_i | x_i) = G(x_i^T \theta).$$

Пусть k -я объясняющая переменная является непрерывной переменной. Тогда **предельный эффект** (*marginal effect*) этой переменной определяется как производная

$$\frac{\partial P\{y_i = 1 | x_i\}}{\partial x_{ik}} = \frac{\partial G(x_i^T \theta)}{\partial x_{ik}},$$

и в отличие от линейной модели этот эффект зависит от значений объясняющих переменных для i -го субъекта $x_i = (x_{i1}, \dots, x_{ip})^T$. Малое изменение Δx_{ik} k -й объясняющей переменной приводит (при неизменных значениях остальных объясняющих переменных) к изменению вероятности $P\{y_i = 1 | x_i\}$ на величину, приближенно равную

$$\Delta P\{y_i = 1 | x_i\} \cong \frac{\partial P\{y_i = 1 | x_i\}}{\partial x_{ik}} \cdot \Delta x_{ik} = \frac{\partial G(x_i^T \theta)}{\partial x_{ik}} \cdot \Delta x_{ik}.$$

Заметим, что, поскольку модель нелинейна, при интерпретации значений предельного эффекта надо иметь в виду отклик интересующей нас вероятности именно на малые приращения объясняющей переменной.

В случае когда сама объясняющая переменная принимает только два значения — 0 и 1 (дамми-переменная), указывающие на наличие (1) или отсутствие (0) у субъекта определенного признака, малые изменения переменной, о которых говорилось выше, попросту невозможны. В этом случае предельный эффект определяют просто как разность

$$P\{y_i = 1 | x_i^*, d_i = 1\} - P\{y_i = 1 | x_i^*, d_i = 0\},$$

где d_i — рассматриваемая дамми-переменная;

x_i^* — вектор значений остальных объясняющих переменных.

В пробит-модели $P\{y_i = 1 | x_i\} = \Phi(x_i^T \theta) = \Phi(\theta_1 x_{i1} + \dots + \theta_p x_{ip})$ малое изменение Δx_{ik} k -й объясняющей переменной приводит (при неизменных значениях остальных объясняющих переменных) к изменению вероятности $P\{y_i = 1 | x_i\}$ на величину, приближенно равную:

$$\begin{aligned} \Delta P\{y_i = 1 | x_i\} &\cong \frac{\partial \Phi(\theta_1 x_{i1} + \dots + \theta_p x_{ip})}{\partial x_{ik}} \cdot \Delta x_{ik} = \\ &= \varphi(x_i^T \theta) \theta_k \Delta x_{ik}, \end{aligned}$$

где $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ — функция плотности стандартного нормального распределения $N(0, 1)$, математическое ожидание которого равно 0, а дисперсия равна 1.

Предельный эффект k -й объясняющей переменной равен $\varphi(x_i^T \theta) \theta_k$ (а не θ_k — как в линейной модели).

В логит-модели $P\{y_i = 1 | x_i\} = \Lambda(x_i^T \theta) = \Lambda(\theta_1 x_{i1} + \dots + \theta_p x_{ip})$ малое изменение Δx_{ik} k -й объясняющей переменной приводит (при неизменных значениях остальных объясняющих переменных) к изменению вероятности $P\{y_i = 1 | x_i\}$ на величину, приближенно равную:

$$\Delta P\{y_i = 1 | x_i\} \cong \frac{\partial P\{y_i = 1 | x_i\}}{\partial x_{ik}} \cdot \Delta x_{ik} = \frac{\partial \Lambda(x_i^T \theta)}{\partial x_{ik}} \cdot \Delta x_{ik}.$$

Учитывая явный вид функции $\Lambda(z)$, найдем отсюда:

$$\Delta P\{y_i = 1 | x_i\} \cong \{\Lambda(x_i^T \theta)(1 - \Lambda(x_i^T \theta))\theta_k\} \Delta x_{ik}.$$

Выражение, заключенное в фигурные скобки, представляет собой предельный эффект для k -й объясняющей переменной в логит-модели.

Заметим теперь следующее. Пусть $p = P(A)$ — вероятность некоторого события A , $0 < p < 1$. Отношение $\frac{p}{1-p}$ часто называют **шансами** (*odds*) этого

события. Например, если $p = \frac{2}{3}$, то $\frac{p}{1-p} = \frac{\frac{2}{3}}{\frac{1}{3}} = 2$, и шансы на то, что собы-

тие A произойдет, равны $2:1$ («два к одному», или «в 2 раза выше»). Логарифм отношения $\frac{p}{1-p}$ называют **логитом** (*logit*), $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$. Если

$\text{logit}(p) = 0$, то $p = 1 - p = 0.5$, т.е. шансы для события A равны «50 на 50». Если $\text{logit}(p) > 0$, то больше шансов, что событие A произойдет. Если $\text{logit}(p) < 0$, то больше шансов, что событие A не произойдет.

Пусть теперь $p = P\{y_i = 1 \mid x_i\}$. В логит-модели $p = \Lambda(x_i^T \theta) = \frac{\exp(x_i^T \theta)}{1 + \exp(x_i^T \theta)}$,

$1 - p = \frac{1}{1 + \exp(x_i^T \theta)}$, так что $\text{logit}(p) = x_i^T \theta$, т.е. логит-модель линейна в отно-

шении логита. Отсюда вытекает, что изменение значения k -й объясняющей переменной на величину Δx_{ik} приводит (при неизменных значениях осталь-

ных объясняющих переменных) к изменению значения $\ln\left(\frac{p}{1-p}\right)$ на $\theta_k \Delta x_{ik}$,

что при малых значениях Δx_{ik} означает изменение значения отношения $\frac{p}{1-p}$

приблизительно на $100\theta_k \Delta x_{ik} \%$.

Иначе говоря, шансы на то, что $y_i = 1$, изменяются приблизительно на $100\theta_k \Delta x_{ik} \%$.

Проверка выполнения стандартных предположений

При анализе обычных линейных моделей регрессии проверка выполнения стандартных предположений осуществляется посредством графического анализа и различных статистических критериев, призванных выявить наличие таких особенностей статистических данных, которые могут говорить не в пользу гипотезы о выполнении стандартных предположений.

Проанализируем, однако, график остатков для пробит-модели, оцененной по рассмотренному выше множеству данных о наличии (отсутствии) собственных автомобилей у 1000 семей (рис. 4.5). Этот график по форме разительно отличается от тех, с которыми приходится сталкиваться при анализе обычных моделей регрессии с непрерывной объясняемой переменной. И это неудивительно, так как, если вспомнить свойства случайных ошибок в моделях бинарного выбора, при заданных значениях объясняющих переменных случайная величина ε_i может принимать в i -м наблюдении только два значения. Следовательно, привычный графический анализ остатков не дает здесь полной информации, и более полезным является непосредственное использование подходящих статистических критериев.

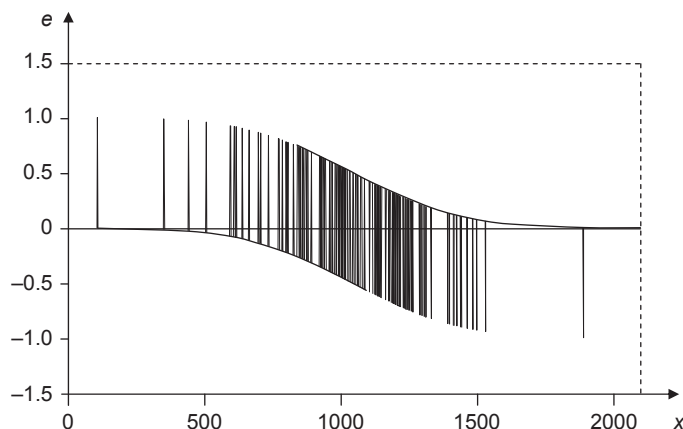


Рис. 4.5

Поскольку для оценивания модели бинарного выбора используется метод максимального правдоподобия, естественным представляется сравнение максимумов функций правдоподобия, получаемых при оценивании модели с выполненными стандартными предположениями и при оценивании модели, в которой эти предположения не выполняются. При этом предполагается, что эти две модели — **гнездовые** (*nested models*), т.е. первая вложена во вторую, так что вторая модель является более сложной, а первая — частным случаем второй модели.

Здесь следует отметить, что сравнение максимумов правдоподобий в двух гнездовых моделях фактически уже использовалось выше. Действительно, на таком сравнении основаны определения коэффициентов:

$$pseudoR^2 = 1 - \frac{1}{1 + \frac{2(\ln L_1 - \ln L_0)}{n}}$$

и

$$McFaddenR^2 = 1 - \frac{\ln L_1}{\ln L_0}.$$

В этом случае в качестве гнездовых рассматриваются основная модель (с одной или несколькими объясняющими переменными, помимо константы) и вложенная в нее тривиальная модель (в правой части объясняющей переменной является только константа).

Сосредоточимся на некоторых статистических критериях проверки гипотез о выполнении стандартных предположений, но, прежде чем перейти к рассмотрению и применению подобных критериев, изучим процесс порождения данных, приводящий к пробит-модели.

Предположим, что переменная y_i^* характеризует «полезность» наличия некоторого предмета длительного пользования для i -й семьи, и эта полезность определяется соотношением

$$y_i^* = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

где x_{i1}, \dots, x_{ip} — значения p объясняющих переменных для i -й семьи;

$\varepsilon_1, \dots, \varepsilon_n$ — случайные ошибки, отражающие влияние на полезность наличия указанного предмета для i -й семьи каких-то неучтенных дополнительных факторов.

Пусть i -я семья приобретает этот предмет длительного пользования, если $y_i^* > \gamma_i$, где γ_i — **пороговое значение** (*threshold value*), и индикаторная переменная y_i отмечает наличие ($y_i = 1$) или отсутствие ($y_i = 0$) данного предмета у i -й семьи. Тогда

$$\begin{aligned} P\{y_i = 1 | x_i\} &= P\{y_i^* > \gamma_i | x_i\} = P\{\beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i > \gamma_i | x_i\} = \\ &= P\{\varepsilon_i > \gamma_i - \beta_1 x_{i1} - \dots - \beta_p x_{ip} | x_i\}, \end{aligned}$$

и если $x_{i1} \equiv 1$, то

$$P\{y_i = 1 | x_i\} = P\{\varepsilon_i > (\gamma_i - \beta_1) - (\beta_2 x_{i2} + \dots + \beta_p x_{ip}) | x_i\}.$$

Если предположить, что ошибки $\varepsilon_1, \dots, \varepsilon_n$ — независимые в совокупности (и независимые от $x_{ij}, j = 1, \dots, p$) случайные величины, имеющие одинаковое нормальное распределение $\varepsilon_i \sim N(0, \sigma^2)$, тогда

$$\begin{aligned} P\{y_i = 1 | x_i\} &= 1 - \Phi\left(\frac{\gamma_i - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right) = \\ &= \Phi\left(\frac{-\gamma_i + \beta_1}{\sigma} + \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right). \end{aligned}$$

(Здесь использовали вытекающее из симметрии стандартного нормального распределения соотношение $1 - \Phi(x) = \Phi(-x)$.) Обозначив

$$\theta_1 = \frac{-\gamma_i + \beta_1}{\sigma}, \quad \theta_j = \frac{\beta_j}{\sigma}, \quad j = 2, \dots, p,$$

получим:

$$P\{y_i = 1 | x_i\} = \Phi(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) = \Phi(x_i^T \theta).$$

Но именно таким образом и определяется пробит-модель.

Пусть имеются в наличии только значения $y_i, x_{i1}, \dots, x_{ip}$, а значения y_i^* не доступны наблюдению. В таком случае переменную y_i^* называют **латент-**

ной (скрытой) (*latent variable*). Применив метод максимального правдоподобия, получим оценки параметров пробит-модели $\hat{\theta}_1, \dots, \hat{\theta}_p$, но по ним нельзя однозначно восстановить значения параметров β_1, \dots, β_p , если неизвестны значения σ и $\gamma_1, \dots, \gamma_n$. Действительно, если оценки $\hat{\sigma}, \hat{\gamma}_1, \dots, \hat{\gamma}_n, \hat{\beta}_1, \dots, \hat{\beta}_p$ таковы, что

$$\hat{\theta}_1 = \frac{-\hat{\gamma}_1 + \hat{\beta}_1}{\hat{\sigma}}, \quad \hat{\theta}_j = \frac{\hat{\beta}_j}{\hat{\sigma}}, \quad j = 2, \dots, p,$$

то к тем же значениям $\hat{\theta}_1, \dots, \hat{\theta}_p$ приводят оценки $k\hat{\sigma}, k\hat{\gamma}_1, \dots, k\hat{\gamma}_n, k\hat{\beta}_1, \dots, k\hat{\beta}_p$, где k — произвольное число, $-\infty < k < \infty$.

Таким образом, в рассмотренной ситуации для однозначной идентификации коэффициентов β_1, \dots, β_p необходима некая нормализация функции полезности. В стандартной модели предполагается, что $\sigma = 1$ и $\gamma_1 = \dots = \gamma_n = 0$, так что

$$\hat{\beta}_1 = \hat{\theta}_1, \dots, \hat{\beta}_p = \hat{\theta}_p,$$

и именно такую модель будем теперь рассматривать.

Прежде всего заметим, что получение оценок параметров β_1, \dots, β_p в такой модели методом максимального правдоподобия принципиально опирается на предположение о нормальности ошибок $\varepsilon_1, \dots, \varepsilon_n$: $\varepsilon_i \sim N(0, 1)$. Поэтому важной является задача проверки этого предположения, т.е. проверка гипотезы

$$H_0: \varepsilon_1, \dots, \varepsilon_n \sim i.i.d., \quad \varepsilon_i \sim N(0, 1).$$

Наряду со стандартной моделью (модель 1) рассмотрим модель 2, отличающуюся от стандартной тем, что в ней

$$P\{\varepsilon_i \leq t\} = \Phi(t + \omega_1 t^2 + \omega_2 t^3),$$

так что

$$P\{y_i = 1 | x_i\} = \Phi(x_i^T \theta + \omega_1 (x_i^T \theta)^2 + \omega_2 (x_i^T \theta)^3).$$

При этом модель 1 является частным случаем модели 2 (при $\omega_1 = \omega_2 = 0$), таким образом модели 1 и 2 — гнездовые, и в рамках более общей модели 2 гипотеза H_0 принимает вид

$$H_0: \omega_1 = \omega_2 = 0.$$

Класс распределений вида $P\{\varepsilon_i \leq t\} = \Phi(t + \omega_1 t^2 + \omega_2 t^3)$ допускает асимметрию и положительный эксцесс (островершинность) распределения. График на рис. 4.6 позволяет сравнить поведение функции стандартного нормального распределения $\Phi(t)$ (серая линия) и функции $\Phi(t + 0.5t^2 + 0.5t^3)$ (черная линия).

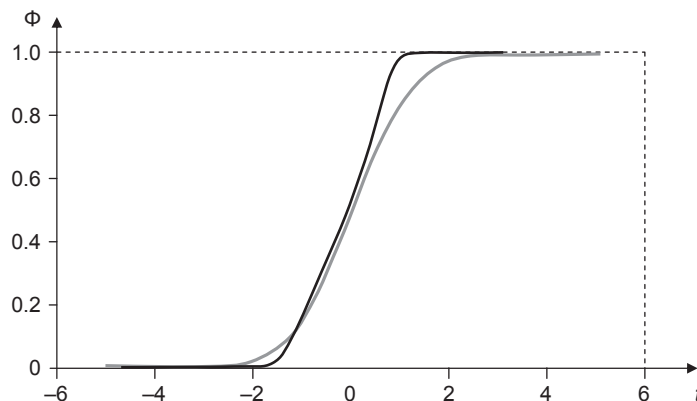


Рис. 4.6

Пусть L_j — максимум функции правдоподобия в модели j ($j = 1, 2$) и

$$LR = -2 \ln \frac{L_1}{L_2} = 2(\ln L_2 - \ln L_1).$$

Критерий отношения правдоподобий (*likelihood ratio test* — *LR test*) отвергает гипотезу H_0 , если наблюдаемое значение статистики LR превышает критическое значение LR_{crit} , соответствующее выбранному уровню значимости α . Этот критерий асимптотический: критическое значение LR_{crit} вычисляется на основе распределения, к которому стремится при $n \rightarrow \infty$ распределение статистики LR , если гипотеза H_0 верна. Этим предельным распределением является распределение хи-квадрат с двумя степенями свободы. Итак, в соответствии с критерием отношения правдоподобий гипотеза H_0 отвергается, если

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

где $\chi^2_{1-\alpha}(2)$ — квантиль уровня $(1 - \alpha)$ распределения хи-квадрат с двумя степенями свободы.

Обратимся опять к смоделированным данным о наличии или об отсутствии собственных автомобилей у 1000 домохозяйств.

Оценив пробит-модель (модель 1) по этим данным, получим результаты, приведенные в табл. 4.7.

Оценивание модели 2 дает результаты, представленные в табл. 4.8.

Соответственно здесь:

$$LR = 2(\ln L_2 - \ln L_1) = 2(275.7686 - 274.6286) = 2.28.$$

Поскольку $\chi^2_{0.95}(2) = 5.99$, критерий отношения правдоподобий не отвергает гипотезу H_0 при уровне значимости 0.05. Заметим еще, что значению $LR = 2.28$

Таблица 4.7

Оценка пробит-модели (модели 1)

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
α	-3.503812	0.200637	-17.46343	0.0000
β	0.003254	0.000178	18.25529	0.0000
$\ln L_1$	-275.7686	Akaike info criterion		0.555537
		Schwarz criterion		0.565353
		Hannan-Quinn criter		0.559268

Таблица 4.8

Оценка общей модели (модели 2)

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
α	-3.851178	0.324895	-11.85359	0.0000
β	0.003540	0.000292	12.11708	0.0000
ω_1	0.022954	0.025086	0.915039	0.3602
ω_2	-0.017232	0.010178	-1.693097	0.0904
$\ln L_2$	-274.6286	Akaike info criterion		0.557257
		Schwarz criterion		0.576888
		Hannan-Quinn criter		0.564718

соответствует (вычисляемое по асимптотическому распределению $\chi^2(2)$) P-значение 0.6802. Таким образом, критерий отношения правдоподобий не отвергает гипотезу H_0 при любом разумном уровне значимости.

Еще одним «стандартным предположением» является предположение об одинаковом распределении случайных ошибок ε_i в процессе порождения данных. В сочетании с предположением нормальности этих ошибок данное условие сводится к совпадению дисперсий всех этих ошибок. Нарушение этого условия приводит к гетероскедастичной модели и к несостоятельности оценок максимального правдоподобия, получаемых на основании стандартной модели. Для проверки гипотезы совпадения дисперсий можно опять рассмотреть какую-нибудь более общую модель с наличием гетероскедастичности, частным случаем которой является стандартная пробит-модель.

В примере с автомобилями можно допустить, что дисперсии случайных ошибок в процессе порождения данных возрастают с увеличением значений x_i , например, как

$$D(\varepsilon_i | x_i) = \exp(k x_i), \quad k > 0,$$

так что (модель 3)

$$P\{y_i = 1 | x_i\} = \Phi\left(\frac{\alpha + \beta x_i}{\sqrt{\exp(kx_i)}}\right).$$

Здесь имеем две гнездовые модели — модель 3, допускающую гетероскедастичность в указанной форме, и модель 1 (стандартную пробит-модель) как ее частный случай. В рамках модели 3 выполнение стандартных предположений соответствует гипотезе $H_0: k = 0$. Оценивание модели 3 по смоделированным данным дает результаты, приведенные в табл. 4.9.

Таблица 4.9

Оценка общей модели, допускающей гетероскедастичность (модели 3)

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
α	−3.141966	0.317695	−9.889867	0.0000
β	0.002883	0.000316	9.132687	0.0000
k	0.000236	0.000186	−1.269192	0.2044
$\ln L_3$	−275.2619	Akaike info criterion		0.556524
		Schwarz criterion		0.571247
		Hannan-Quinn criter		0.562120

При сравнении с моделью 1 получаем:

$$LR = 2(\ln L_3 - \ln L_1) = 2(275.7686 - 275.2619) = 1.013.$$

Это значение меньше критического значения 3.84, соответствующего уровню значимости 0.05 и вычисленного как квантиль уровня 0.95 асимптотического распределения хи-квадрат с одной степенью свободы. Следовательно, гипотеза $H_0: k = 0$ не отвергается.

Отметим, что решения, принятые на основании критерия отношения правдоподобий, согласуются с решениями, принимаемыми в рассмотренном примере на основании информационных критериев (табл. 4.10).

Таблица 4.10

Критерии принятия решений

Модель	AIC	SC	HQ
Модель 1 (пробит)	0.555537	0.565353	0.559268
Модель 2 (Пирсон)	0.557257	0.576888	0.564718
Модель 3 (гетеро)	0.556524	0.571247	0.562120

По всем трем критериям стандартная пробит-модель предпочтительнее альтернативных моделей.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Когда приходится использовать модели, в которых объясняемая переменная принимает только два различных значения? Почему в подобных ситуациях линейные модели оказываются непригодными?
2. Как используется метод максимального правдоподобия для оценивания параметров модели бинарного выбора?
3. Чем различаются модели пробит, логит и гомпит? В каких ситуациях оценки максимального правдоподобия параметров этих моделей не определены?
4. Как интерпретируются оцененные коэффициенты в моделях пробит, логит и гомпит?
5. Какие варианты коэффициента детерминации используются при подборе моделей бинарного выбора?
6. Как можно сравнить качество нескольких альтернативных моделей бинарного выбора с разным количеством объясняющих переменных?
7. Для какой цели используется критерий Хосмера — Лемешоу?
8. Каким образом линейная модель с латентной объясняющей переменной приводит к пробит-модели? Как проверяется выполнение стандартных предположений об ошибках в латентной модели?

Тема 4.2

МОДЕЛИ, В КОТОРЫХ ОБЪЯСНЯЕМАЯ ПЕРЕМЕННАЯ ПРИНИМАЕТ НЕСКОЛЬКО РАЗЛИЧНЫХ ЗНАЧЕНИЙ

Порядковая пробит-модель

В примере с наличием или отсутствием у семьи собственного автомобиля значение $y_i = 1$ говорило о том, что i -я семья имеет собственный автомобиль, но не уточняло, сколько в действительности автомобилей имеет семья — один, два или, быть может, еще больше. Обращаясь к процессу порождения данных, ориентирующемуся на значения функции полезности и сравнение ее с пороговыми значениями, можно предположить наличие не одного, а двух пороговых значений для каждой семьи, так что при превышении первого порога семья имеет в наличии один автомобиль, а при превышении второго (более высокого) порога — два автомобиля или более.

Обобщая эту ситуацию, рассмотрим процесс порождения данных, в котором имеется некоторая ненаблюдаемая (латентная) переменная y_i^* , значения которой связаны со значениями x_{i1}, \dots, x_{ip} объясняющих переменных для i -го субъекта исследования следующим образом:

$$y_i^* = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

где ε_i — случайная ошибка, отражающая влияние на значение y_i^* неучтенных дополнительных факторов.

Вместе со значениями x_{i1}, \dots, x_{ip} наблюдается также переменная y_i , которая может принимать K различных значений в соответствии со следующей схемой:

$$y_i = \begin{cases} 1, & \text{если } y_i^* \leq \gamma_{i,1}, \\ \dots & \\ k, & \text{если } \gamma_{i,k-1} < y_i^* \leq \gamma_{i,k}, \\ \dots & \\ K, & \text{если } y_i^* > \gamma_{i,K-1}, \end{cases}$$

где $\gamma_{i,1} < \dots < \gamma_{i,k} < \dots < \gamma_{i,K-1}$ — пороговые значения, вообще говоря, ненаблюдаемые.

Предполагая, что ошибки $\varepsilon_1, \dots, \varepsilon_n$ — независимые в совокупности (и независимые от $x_{ij}, j = 1, \dots, p$) случайные величины, имеющие одинаковое нормальное распределение $\varepsilon_i \sim N(0, \sigma^2)$, получаем **порядковую пробит-модель** (*ordered probit model*).

Рассмотрим частный случай, когда $K = 3$ и пороговые значения одинаковы для всех субъектов исследования, так что $\gamma_{i,1} \equiv \gamma_1, \gamma_{i,2} \equiv \gamma_2$ и

$$y_i = \begin{cases} 1, & \text{если } y_i^* \leq \gamma_1, \\ 2, & \text{если } \gamma_1 < y_i^* \leq \gamma_2, \\ 3, & \text{если } y_i^* > \gamma_2. \end{cases}$$

При этом

$$\begin{aligned} P\{y_i = 1 | x_i\} &= P\{y_i^* \leq \gamma_1 | x_i\} = P\{\beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i \leq \gamma_1 | x_i\} = \\ &= P\{\varepsilon_i \leq (\gamma_1 - \beta_1) - (\beta_2 x_{i2} + \dots + \beta_p x_{ip}) | x_i\} = \\ &= \Phi\left(\frac{\gamma_1 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right); \end{aligned}$$

$$\begin{aligned} P\{y_i = 2 | x_i\} &= P\{\gamma_1 < y_i^* \leq \gamma_2 | x_i\} = P\{\gamma_1 < \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i \leq \gamma_2 | x_i\} = \\ &= \Phi\left(\frac{\gamma_2 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right) - \Phi\left(\frac{\gamma_1 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right); \end{aligned}$$

$$\begin{aligned} P\{y_i = 3 | x_i\} &= P\{y_i^* \leq \gamma_2 | x_i\} = P\{\beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i > \gamma_2 | x_i\} = \\ &= P\{\varepsilon_i > (\gamma_2 - \beta_1) - (\beta_2 x_{i2} + \dots + \beta_p x_{ip}) | x_i\} = \\ &= 1 - \Phi\left(\frac{\gamma_2 - \beta_1}{\sigma} - \frac{\beta_2 x_{i2} + \dots + \beta_p x_{ip}}{\sigma}\right). \end{aligned}$$

Пусть в наличии имеются только значения $y_i, x_{i1}, \dots, x_{ip}, i = 1, \dots, n$. Применив метод максимального правдоподобия, как и в случае пробит-модели с двумя исходами, нельзя однозначно восстановить значения параметров β_1, \dots, β_p , если не известны σ, γ_1 и γ_2 . Поэтому и здесь для однозначной идентификации коэффициентов β_1, \dots, β_p необходима какая-то нормализация. В стандартной модели предполагается, что $\sigma = 1$ и $\gamma_1 = 0$, хотя возможны и другие нормализации.

Используя стандартную нормализацию и обозначив $\gamma_2 = \gamma$, получим в модели с тремя исходами:

$$P\{y_i = 1 | x_i\} = P\{y_i^* \leq 0 | x_i\} = \Phi(-x_i^T \beta);$$

$$P\{y_i = 2 | x_i\} = P\{0 < y_i^* \leq \gamma | x_i\} = \Phi(\gamma - x_i^T \beta) - \Phi(-x_i^T \beta);$$

$$P\{y_i = 3 | x_i\} = P\{y_i^* > \gamma | x_i\} = 1 - \Phi(\gamma - x_i^T \beta).$$

При этом коэффициент β_j допускает двойное истолкование. В соответствии с моделью для y_i^* положительное значение этого коэффициента означает, что переменная y_i^* возрастает с возрастанием j -й объясняющей переменной. В соответствии с приведенными выражениями для вероятностей получения значений $y_i = 1, y_i = 2$ и $y_i = 3$ последнее приводит к возрастанию вероятности $P\{y_i = 3 | x_i\}$ и к убыванию вероятности $P\{y_i = 1 | x_i\}$. Что же касается вероятности $P\{y_i = 2 | x_i\}$, то здесь возможно как возрастание, так и убывание этой вероятности в зависимости от конкретной ситуации.

Прогнозирование по оцененной модели производится в соответствии со следующим соглашением. Прогнозное значение \hat{y}_i полагается равным k_0 , если $\hat{P}\{y_i = k_0 | x_i\} = \max_{k=1, \dots, K} \hat{P}\{y_i = k | x_i\}$.

ПРИМЕР 4.2.1

Рассмотрим выборку, состоящую из 1000 семей со среднедушевым месячным доходом от 100 до 2100 у.е., среди которых 499 семей не имеют собственного автомобиля, 369 семей имеет один автомобиль, 132 семьи — два автомобиля. Выборка получена посредством моделирования. При этом был использован процесс порождения данных в виде:

$$y_i^* = x_i + \varepsilon_i, \quad i = 1, \dots, 1000,$$

где ε_i — независимые в совокупности (и независимые от x_i) случайные величины, имеющие одинаковое нормальное распределение $\varepsilon_i \sim N(0, 300^2)$, т.е. $\sigma = 300$.

Здесь $K = 3$ и границы γ_1 и γ_2 были выбраны равными соответственно: $\gamma_1 = 1100$ и $\gamma_2 = 1850$.

В результате получаем порядковую пробит-модель:

$$y_i = \begin{cases} 1, & \text{если } y_i^* \leq \gamma_1, \\ 2, & \text{если } \gamma_1 < y_i^* \leq \gamma_2, \\ 3, & \text{если } y_i^* > \gamma_2, \end{cases}$$

где $y_1 = 1$, если i -я семья не имеет автомобиля;

$y_1 = 2$, если i -я семья имеет один автомобиль;

$y_1 = 3$, если i -я семья имеет два (или более) автомобиля.

На рис. 4.7 показана зависимость полученных значений y_i^* от x_i . Горизонтальные линии соответствуют разделительным порогам $LEVEL1 = 1100$ и $LEVEL2 = 1850$.

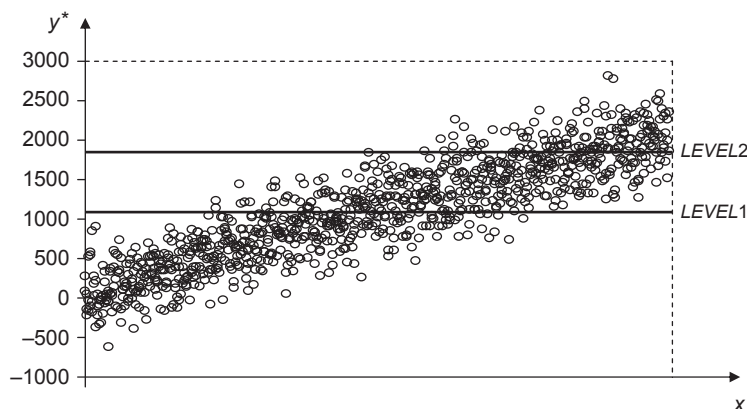


Рис. 4.7

Наблюдения с $y_i^* \leq 1100$ встречаются в группе семей с доходами от 200 до 1600 у.е.; наблюдения с $1100 < y_i^* \leq 1850$ — в группе семей с доходами от 548 до 2094 у.е.; наблюдения с $y_i^* > 1850$ — в группе семей с доходами от 1318 у.е. и выше. Важно отметить, что эти группы пересекаются, и это связано как раз с наличием случайной составляющей в уравнении полезности. Если бы этой составляющей не было, получилась бы картина, изображенная на рис. 4.8.

В таком случае произошло бы разбиение на три непересекающиеся группы.

- для всех семей с доходами, не превышающими 1100 у.е. ($y_i = 1$);
- для всех семей с доходами, превышающими 1100 у.е., но не превышающими 1850 у.е. ($y_i = 2$);
- для всех семей с доходами, превышающими 1850 у.е. ($y_i = 3$).

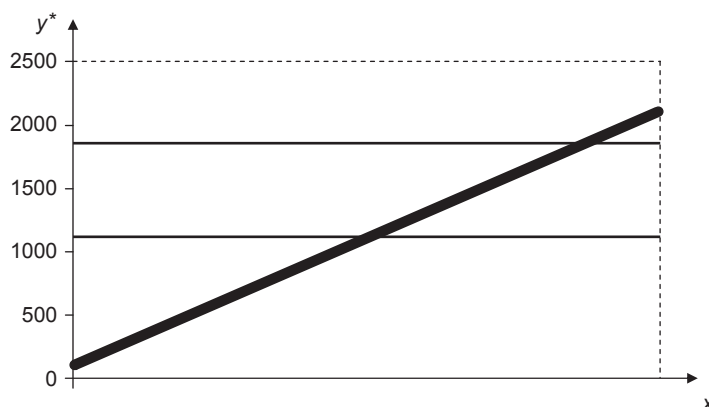


Рис. 4.8

Представим теперь, что имеем только выборочные данные, т.е. пары (x_i, y_i) , $i = 1, \dots, 1000$. Оценивание методом максимального правдоподобия порядковой пробит-модели с нормализацией $\sigma = 1$ (именно такая нормализация используется в пакете *EViews*) дает результаты, представленные в табл. 4.11.

Таблица 4.11

Метод максимального правдоподобия для пробит-модели с нормализацией

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
X	0.003361	0.000158	21.31648	0.0000
Границы				
γ_1	3.693723	0.185109	19.95431	0.0000
γ_2	6.306692	0.279737	22.54510	0.0000

Иначе говоря, нормализованная модель оценивается как

$$y_i^* = 0.003361x_i + u_i,$$

где $u_i \sim N(0, 1)$, и

$$y_i = \begin{cases} 1, & \text{если } y_i^* \leq 3.693723, \\ 2, & \text{если } 3.693723 < y_i^* \leq 6.306692, \\ 3, & \text{если } y_i^* > 6.306692. \end{cases}$$

Если учесть, что мы сами смоделировали выборку и поэтому знаем значение σ , то переход к модели с $\sigma = 300$ соответствует оцененной модели

$$y_i^* = 300 \cdot 0.003361x_i + 300u_i = 1.0083x_i + \varepsilon_i,$$

где $\varepsilon_i \sim N(0, 300^2)$, и

$$\gamma_1 = 300 \cdot 3.693723 = 1108.1169,$$

$$\gamma_2 = 300 \cdot 6.306692 = 1892.0076,$$

Как видно, параметры оцененной модели очень близки к параметрам истинной модели. Результаты прогнозов по оцененной модели приведены в табл. 4.12. Диаграмма на рис. 4.9 иллюстрирует эти результаты.

Таблица 4.12

Результаты прогнозов по оцененной модели

y_i	Количество наблюдений	\hat{y}_i	Количество наблюдений	Ошибка
1	499	1	500	-1
2	369	2	387	-18
3	132	3	113	19

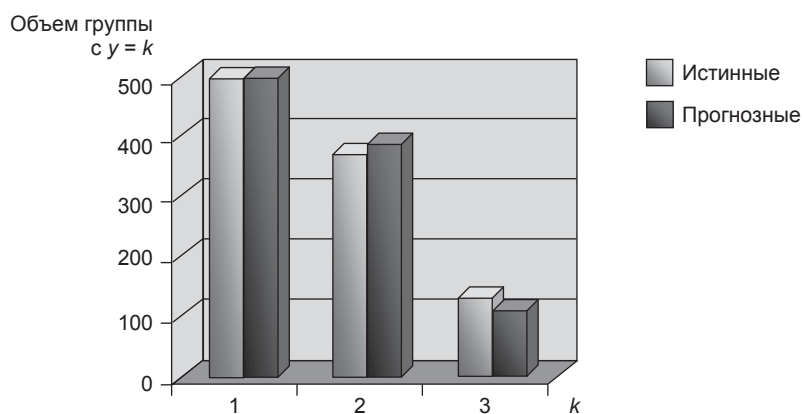


Рис. 4.9

Для сравнения в табл. 4.13 приведены результаты прогнозов по тривиальной модели, не учитывающей в уравнении для y_i^* влияние доходов i -й семьи.

Таблица 4.13

Результаты прогнозов по тривиальной модели

y_i	Количество наблюдений	\hat{y}_i	Количество наблюдений	Ошибка
1	499	1	1000	-501
2	369	2	0	369
3	132	3	0	132

Приведем также сводную таблицу количества правильных и неправильных прогнозов для значений $y_i = 1, 2, 3$ (табл. 4.14).

Таблица 4.14

Сравнение результатов прогнозов моделей

Значение	$\hat{y}_i = 1$	$\hat{y}_i = 2$	$\hat{y}_i = 3$
$y_i = 1$	438	61	0
$y_i = 2$	62	265	42
$y_i = 3$	0	61	71

Таким образом, из 1000 прогнозов правильными оказались 774, т.е. 77.4%. При этом значения $y_i = 1$ правильно прогнозируются в 438 случаях из 499, т.е. в 87.8% случаев, значения $y_i = 2$ — в 71.8% случаев, значения $y_i = 3$ — в 53.8% случаев. ■

Мультиномиальная модель

В целом ряде случаев не существует естественного упорядочения альтернатив, благодаря которому и возникает монотонная связь между непрерывной латентной переменной и наблюдаемой переменной, принимающей конечное количество значений.

Пусть имеем K таких альтернатив (занумеруем их в произвольном порядке числами $1, \dots, K$) и пусть i -й субъект исследования приписывает k -й альтернативе полезность u_{ik} , так что

$$u_{ik} = \beta_1 x_{i1,k} + \dots + \beta_p x_{ip,k} + \varepsilon_{ik} = x_{ik}^T \beta + \varepsilon_{ik}, \quad i = 1, \dots, n, \quad k = 1, \dots, K,$$

где $x_{ik} = (x_{i1,k}, \dots, x_{ip,k})^T$;

ε_{ik} — независимые в совокупности (и независимые от x_{ik}) случайные величины, имеющие одинаковое распределение.

Предположим, что i -й субъект выбирает альтернативу k , если для него эта альтернатива имеет максимальную полезность. В этом случае полагаем $y_i = k$. Тогда (условная при заданных значениях x_{ik} , $k = 1, \dots, K$) вероятность того, что i -й субъект выберет альтернативу k , равна:

$$P\{y_i = k\} = P\left\{u_{ik} = \max_{j=1, \dots, K} u_{ij}\right\} = P\left\{x_{ik}^T \beta + \varepsilon_{ik} > \max_{j=1, \dots, K, j \neq k} (x_{ij}^T \beta + \varepsilon_{ij})\right\}.$$

Выразить такую вероятность в явном виде весьма проблематично. Однако если предположить, что общим для всех случайных величин ε_{ik} является

стандартное распределение экстремальных значений (максимума) I типа с функцией распределения

$$G(z) = \exp(-e^{-z}), \quad -\infty < z < \infty,$$

(это распределение часто называют также распределением Гумбеля), то формула для вычисления вероятности $P\{y_i = k\}$ принимает достаточно простой вид:

$$P\{y_i = k\} = \frac{\exp(x_{ik}^T \beta)}{\exp(x_{i1}^T \beta) + \exp(x_{i2}^T \beta) + \dots + \exp(x_{iK}^T \beta)}.$$

Заметим, однако, что если и числитель, и знаменатель правой части последнего выражения разделить на $\exp(x_{i1}^T \beta)$, то получим

$$P\{y_i = k\} = \frac{\exp(x_{ik}^T \beta - x_{i1}^T \beta)}{1 + \exp(x_{i2}^T \beta - x_{i1}^T \beta) + \dots + \exp(x_{iK}^T \beta - x_{i1}^T \beta)}.$$

Следовательно, каким бы ни было значение линейной комбинации $x_{i1}^T \beta$, вероятность $P\{y_i = k\}$ будет зависеть только от разностей $(x_{i2}^T \beta - x_{i1}^T \beta)$, ..., $(x_{iK}^T \beta - x_{i1}^T \beta)$. Это обстоятельство приводит к естественной нормализации, при которой полагают

$$x_{i1}^T \beta = 0, \quad i = 1, \dots, n,$$

тогда

$$P\{y_i = k\} = \frac{\exp(x_{ik}^T \beta)}{1 + \exp(x_{i2}^T \beta) + \dots + \exp(x_{iK}^T \beta)}.$$

Такую модель разные авторы называют по-разному. Так, в книгах (Verbeek, 2000) и (Amemiya, 1985) об этой модели говорится как о **мультиномиальной логит-модели** (*multinomial logit model*). В книгах (Greene, 2003) и (Davidson, MacKinnon, 1993) она называется **условной логит-моделью** (*conditional logit model*), а под мультиномиальной логит-моделью подразумевается модель

$$P\{y_i = k\} = \frac{\exp(x_i^T \beta^k)}{\exp(x_i^T \beta^1) + \exp(x_i^T \beta^2) + \dots + \exp(x_i^T \beta^K)},$$

в которой объясняющие переменные специфичны только в отношении самих субъектов исследования (но не в отношении альтернатив), а специфичными в отношении альтернатив являются коэффициенты модели. Соответственно здесь $\beta^k = (\beta_{1,k}, \dots, \beta_{p,k})^T$ — вектор коэффициентов при объясняющих переменных в представлении функции полезности для k -й альтернативы:

$$u_{ik} = \beta_{1,k} x_{i1} + \dots + \beta_{p,k} x_{ip} + \varepsilon_{ik} = x_i^T \beta^k + \varepsilon_{ik}, \quad i = 1, \dots, n.$$

Последняя модель под названием мультиномиальной логит-модели появляется и в пакете *EViews*. Поскольку в этой модели x_i не зависят от альтернативы, являясь собственными атрибутами субъекта, то

$$P\{y_i = k\} = \frac{\exp(x_i^T (\beta^k - \beta^1))}{1 + \exp(x_i^T (\beta^2 - \beta^1)) + \dots + \exp(x_i^T (\beta^K - \beta^1))},$$

так что эта вероятность зависит только от разностей $\beta^2 - \beta^1, \dots, \beta^K - \beta^1$, а для нормализации можно положить вектор β^1 равным нулевому вектору. При такой нормализации

$$P\{y_i = k\} = \frac{\exp(x_i^T \beta^k)}{1 + \exp(x_i^T \beta^2) + \dots + \exp(x_i^T \beta^K)}.$$

В этом случае (условная при фиксированных $x_{ij}, j = 1, \dots, p, i = 1, \dots, n$) совместная вероятность получения конкретного набора наблюдений y_1, \dots, y_n (конкретного набора значений $1, \dots, K$) равна произведению:

$$\prod_{i=1}^n \prod_{k=1}^K (P\{y_i = k\})^{d_{ik}} = \prod_{i=1}^n \prod_{k=1}^K \left(\frac{\exp(x_i^T \beta^k)}{1 + \exp(x_i^T \beta^2) + \dots + \exp(x_i^T \beta^K)} \right)^{d_{ik}},$$

где

$$d_{ik} = \begin{cases} 1, & \text{если } y_i = k, \\ 0, & \text{если } y_i \neq k. \end{cases}$$

Правая часть этого выражения при фиксированных $x_i, i = 1, \dots, n$, является функцией от вектора неизвестных параметров $\beta, \beta = (\beta^1, \dots, \beta^K)^T$:

$$L(\beta) = L(\beta | x_1, \dots, x_n) = \prod_{i=1}^n \prod_{k=1}^K \left(\frac{\exp(x_i^T \beta^k)}{1 + \exp(x_i^T \beta^2) + \dots + \exp(x_i^T \beta^K)} \right)^{d_{ik}},$$

и эта функция как функция правдоподобия является объектом максимизации по β . Результатом такой максимизации являются оценки максимального правдоподобия для векторов коэффициентов $\hat{\beta}^k = (\hat{\beta}_{1,k}, \dots, \hat{\beta}_{p,k})^T, k = 1, \dots, K$.

ПРИМЕР 4.2.2

Рассмотрим смоделированную ситуацию, в которой, как и в последней модели, переменные специфичны только в отношении самих субъектов исследования.

Пусть $x_{i1} \equiv 1, x_{i2}$ — типичное количество посещений продуктового магазина в неделю i -й семьей (от 1 до 7), x_{i3} — среднемесячный доход на одного члена i -й семьи (от 50 до 250 у.е.). Выбранная модель порождения данных имитирует поведение 1000 семей, проживающих в одном и том же многоэтажном доме и приобретающих продукты в трех продуктовых магазинах,

ближайших к этому дому. Каждая семья отдает предпочтение одному из трех магазинов, так что здесь имеем 3 альтернативы. Магазины различаются тремя сравнительными характеристиками: ассортиментом (богатый, бедный, промежуточный), удаленностью от дома (наибольшая, наименьшая, средняя) и уровнем цен (максимальный, минимальный, средний). Альтернативы были занумерованы 1, 2, 3 произвольным образом (табл. 4.15).

Таблица 4.15

Характеристики магазинов

k	Характеристики k -го магазина		
	Ассортимент	Удаленность	Уровень цен
1	Богатый	Наибольшая	Средний
2	Бедный	Наименьшая	Минимальный
3	Промежуточный	Средняя	Максимальный

Предполагается, что i -я семья приписывает k -й альтернативе полезность u_{ik} :

$$u_{ik} = \beta_{k1}x_{i1} + \beta_{k2}x_{i2} + \beta_{k3}x_{i3} + \varepsilon_{ik}, \quad i = 1, \dots, 1000, \quad k = 1, 2, 3,$$

где ε_{ik} — независимые в совокупности (и независимые от x_{ij}) случайные величины, имеющие одинаковое распределение с функцией распределения

$$G(z) = \exp(-e^{-z}), \quad -\infty < z < \infty,$$

При этом используем нормализацию

$$\beta_{11} = 0, \quad \beta_{12} = 0, \quad \beta_{13} = 0.$$

Остальные коэффициенты выбраны следующим образом:

$$\beta_{21} = -0.8, \quad \beta_{22} = 1.0, \quad \beta_{23} = -0.0032,$$

$$\beta_{31} = -0.4, \quad \beta_{32} = 0.3, \quad \beta_{33} = 0.0032,$$

так что функции полезности для трех альтернатив имеют вид:

$$u_{i1} = \varepsilon_{i1},$$

$$u_{i2} = -0.8 + x_{i2} - 0.0032x_{i3} + \varepsilon_{i2},$$

$$u_{i3} = -0.4x_{i1} + 0.3x_{i2} + 0.0032x_{i3} + \varepsilon_{i3}.$$

Поведение этих функций иллюстрирует рис. 4.10.

В соответствии с моделью порождения данных i -я семья выбирает альтернативу k , если для этой семьи альтернатива k имеет максимальную полезность. В этом случае полагаем $y_i = k$.

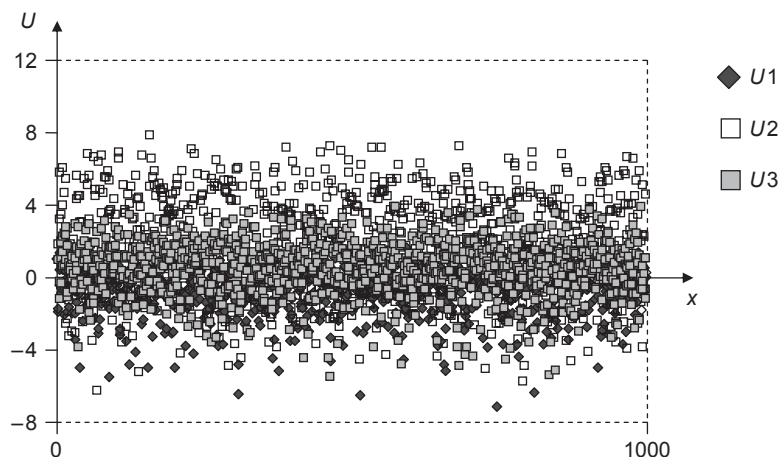


Рис. 4.10

Результаты оценивания методом максимального правдоподобия приведены в табл. 4.16. Все оцененные коэффициенты, за исключением $\hat{\beta}_{23}$, имеют высокую статистическую значимость.

Сравним истинные и оцененные значения коэффициентов (табл. 4.17).

Таблица 4.16

Оценивание методом максимального правдоподобия

Коэффициент	Оценка	Стандартная ошибка	z-статистика	P-значение
β_{21}	-1.655130	0.358914	-4.611496	0.0000
β_{22}	1.270612	0.097636	13.01381	0.0000
β_{23}	-0.001778	0.002134	-0.833304	0.4047
β_{31}	-1.031242	0.327444	-3.149372	0.0016
β_{32}	0.439590	0.087563	5.020273	0.0000
β_{33}	0.006283	0.001957	3.211368	0.0013

Таблица 4.17

Истинные и оцененные значения коэффициентов

Коэффициент	Истинное значение	Оценка
β_{21}	-0.8	-1.655130
β_{22}	1.0	1.270612
β_{23}	-0.0032	-0.001778
β_{31}	-0.4	-1.031242
β_{32}	0.3	0.439590
β_{33}	0.0032	0.006283

Знаки оцененных коэффициентов соответствуют знакам истинных коэффициентов. Кроме того, соблюдается упорядочение значений соответственных коэффициентов, имеющих одинаковые знаки:

$$\beta_{21} < \beta_{31} \quad \text{и} \quad \hat{\beta}_{21} < \hat{\beta}_{31},$$

$$\beta_{22} > \beta_{32} \quad \text{и} \quad \hat{\beta}_{22} > \hat{\beta}_{32}.$$

На основании полученных оценок коэффициентов можно вычислить прогнозные значения вероятностей $P\{y_i = k\}$ предпочтения альтернатив $k = 1, 2, 3$, полагая

$$\hat{P}\{y_i = k\} = \frac{\exp(x_i^T \hat{\beta}^k)}{1 + \exp(x_i^T \hat{\beta}^2) + \exp(x_i^T \hat{\beta}^3)}$$

и используя эти прогнозные значения, дать предсказание номера альтернативы, которую предпочтет семья из рассматриваемого дома с заданной частотой посещения продуктового магазина и заданным уровнем месячного дохода на одного члена семьи. Можно, например, предсказывать для i -й семьи в качестве предпочтительной альтернативу k , если

$$\hat{P}\{y_i = k\} > \hat{P}\{y_i = l\}, \quad l \neq k.$$

Применив такое правило к нашему примеру, получим результаты, приведенные в табл. 4.18.

Таблица 4.18

Предсказание предпочтительных альтернатив

Показатель	Альтернатива (k)		
	1	2	3
Истинный объем группы k	146	603	251
Прогноз объема группы k	101	664	235

Здесь под группой k подразумевается группа семей (среди рассматриваемых 1000 семей), отдающих предпочтение альтернативе k .

Диаграмма на рис. 4.11 иллюстрирует данные табл. 4.18.

Предсказанные объемы групп правильно воспроизводят упорядочение между наблюдаемыми размерами групп: в обоих случаях максимальное количество семей предпочитает альтернативу 2 и минимальное количество семей — альтернативу 1.

Несмотря на то что индивидуальные прогнозы не являются главной целью в подобных исследованиях, все же приведем сводную таблицу количества правильных и неправильных прогнозов для значений $y_i = 1, 2, 3$ (табл. 4.19).

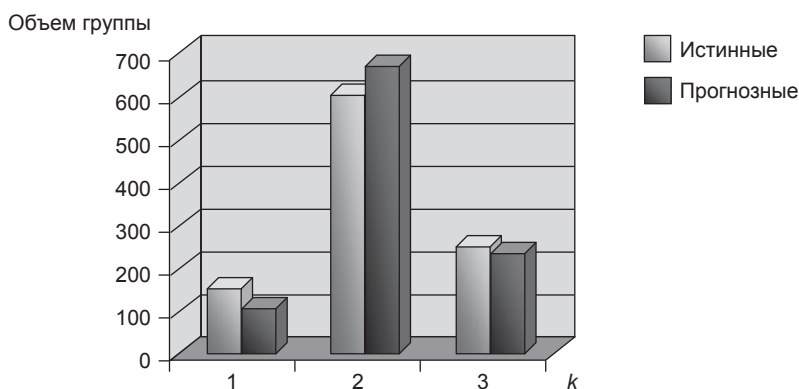


Рис. 4.11

Таблица 4.19

Данные индивидуальных прогнозов

Значение	$\hat{y}_i = 1$	$\hat{y}_i = 2$	$\hat{y}_i = 3$
$y_i = 1$	48	26	72
$y_i = 2$	11	550	42
$y_i = 3$	42	88	121

Таким образом, из 1000 прогнозов правильными оказались 719, т.е. 71.9%. При этом значения $y_i = 1$ правильно прогнозируются в 48 случаях из 146, т.е. только в 32.9% случаев, тогда как значения $y_i = 2$ правильно прогнозируются в 91.2% случаев, значения $y_i = 3$ правильно прогнозируются в 48.2% случаев. ■

ПРИМЕР 4.2.3

В следующей ситуации (в отличие от предыдущих примеров) одна из переменных специфична только в отношении альтернатив, а другая зависит и от альтернативы, и от субъекта.

Пусть $stores_k$ — количество магазинов в k -м торговом центре, $dist_{ik}$ — расстояние от места проживания i -й семьи до k -го торгового центра. Выбранная модель порождения данных имитирует поведение 1000 семей, предпочитающих совершать покупки в трех торговых центрах, каждая семья отдает предпочтение одному из них, так что имеем 3 альтернативы. Альтернативы были занумерованы 1, 2, 3 произвольным образом.

Здесь переменная $stores_k$ специфична только в отношении альтернатив, тогда как значения переменной $dist_{ik}$ зависят и от альтернативы, и от конкретной семьи.

Предполагается, что i -я семья приписывает k -й альтернативе полезность u_{ik} :

$$u_{ik} = \beta_1 stores_k + \beta_2 dist_{ik} + \varepsilon_{ik}, \quad i = 1, \dots, 1000, \quad k = 1, 2, 3,$$

где ε_{ik} — независимые в совокупности (и независимые от $stores_k$ и от $dist_{ik}$) случайные величины, имеющие одинаковое распределение с функцией распределения $G(z) = \exp(-e^{-z})$, $-\infty < z < \infty$.

Коэффициенты выбраны следующим образом:

$$\beta_1 = 0.6, \quad \beta_2 = -1.0,$$

так что функции полезности для трех альтернатив имеют вид:

$$u_{i1} = 0.6 stores_1 - dist_{i1} + \varepsilon_{i1},$$

$$u_{i2} = 0.6 stores_2 - dist_{i2} + \varepsilon_{i2},$$

$$u_{i3} = 0.6 stores_3 - dist_{i3} + \varepsilon_{i3}.$$

В соответствии с моделью порождения данных i -я семья выбирает альтернативу k , если для этой семьи альтернатива k имеет максимальную полезность. В этом случае полагаем $y_i = k$.

Результаты оценивания методом максимального правдоподобия приведены в табл. 4.20.

Таблица 4.20

Оценивание методом максимального правдоподобия

Коэффициент	Оценка	Стандартная ошибка	z-статистика	P-значение
β_1	0.685420	0.043382	15.79956	0.0000
β_2	-1.101152	0.069472	-15.85038	0.0000

Будем опять предсказывать для i -й семьи в качестве предпочтительной альтернативу k , если

$$\hat{P}\{y_i = k\} > \hat{P}\{y_i = l\}, \quad l \neq k.$$

Применив такое правило к нашему примеру, получим результаты, приведенные в табл. 4.21.

Таблица 4.21

Предсказание предпочтенных альтернатив

Показатель	Альтернатива (k)		
	1	2	3
Истинный объем группы k	674	275	51
Прогноз объема группы k	688	266	46

Диаграмма на рис. 4.12 иллюстрирует эти результаты. ■

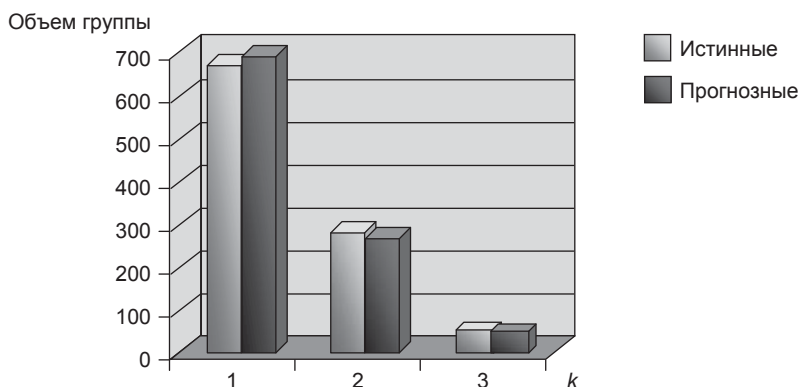


Рис. 4.12

- ✓ **Замечание 4.2.1.** Как было отмечено выше, в рассмотренной мультиномиальной логит-модели, в которой объясняющие переменные специфичны только в отношении самих субъектов исследования,

$$P\{y_i = k\} = \frac{\exp(x_i^T (\beta^k - \beta^1))}{1 + \exp(x_i^T (\beta^2 - \beta^1)) + \dots + \exp(x_i^T (\beta^k - \beta^1))}.$$

Отсюда вытекает, что

$$\frac{P\{y_i = k\}}{P\{y_i = m\}} = \frac{\exp(x_i^T (\beta^k - \beta^1))}{\exp(x_i^T (\beta^m - \beta^1))} = \exp(x_i^T (\beta^k - \beta^m)),$$

т.е. отношение вероятностей выбора альтернатив k и m определяется только параметрами уравнений для полезностей этих двух альтернатив и собственными атрибутами i -го субъекта и не зависит от параметров уравнений для полезностей остальных $(K - 2)$ альтернатив.

- ✓ **Замечание 4.2.2.** Если рассматривается условная логит-модель (с постоянными значениями коэффициентов во всех K уравнениях полезности), в которой объясняющие переменные специфичны в отношении альтернатив, то, как говорилось выше, в такой ситуации

$$P\{y_i = k\} = \frac{\exp(x_{ik}^T \beta)}{\exp(x_{i1}^T \beta) + \dots + \exp(x_{ik}^T \beta)},$$

так что здесь

$$\frac{P\{y_i = k\}}{P\{y_i = m\}} = \frac{\exp(x_{ik}^T \beta)}{\exp(x_{im}^T \beta)} = \exp((x_{ik}^T - x_{im}^T) \beta),$$

т.е. отношение вероятностей выбора альтернатив k и m определяется только общим параметром уравнений для полезностей различных альтернатив и значениями в i -м наблюдении объясняющих переменных, соответствующих k -й и m -й альтернативам. Это отношение не зависит от значений в i -м наблюдении объясняющих переменных, соответствующих остальным $(K - 2)$ альтернативам. Такое свойство независимости оказывается нежелательным во многих ситуациях.

✓ **Замечание 4.2.3.** Пусть среди объясняющих переменных в условной логит-модели (с постоянными значениями коэффициентов во всех K уравнениях полезности) имеются переменные, специфичные только в отношении субъектов (т.е. значения этих переменных для i -го субъекта не зависят от альтернативы).

Пусть соответственно

$$x_{ik}^T = (v_{ik}^T, w_i^T),$$

где v_{ik}^T — вектор значений для i -го субъекта переменных, которые зависят от альтернативы;

w_i^T — вектор значений для i -го субъекта переменных, которые не зависят от альтернативы.

Соответственно разбивается и вектор коэффициентов:

$$\beta^T = (\gamma^T, \delta^T).$$

Тогда

$$\begin{aligned} P\{y_i = k\} &= \frac{\exp(v_{ik}^T \gamma + w_i^T \delta)}{\exp(v_{i1}^T \gamma + w_i^T \delta) + \dots + \exp(v_{iK}^T \gamma + w_i^T \delta)} = \\ &= \frac{\exp(v_{ik}^T \gamma)}{\exp(v_{i1}^T \gamma) + \dots + \exp(v_{iK}^T \gamma)}, \end{aligned}$$

так что эта вероятность не зависит от значений переменных, специфичных только в отношении субъектов.

Чтобы (в рамках модели с постоянным вектором коэффициентов) учесть возможное влияние таких переменных на вероятности $P\{y_i = k\}$, модель надо модифицировать. Одним из возможных способов модификации является создание группы дамми-переменных для альтернатив (*DUMMY* для альтернативы k принимает значение 1, если $y_i = k$, и значение 0 — в противном случае) и умножение каждой из них на переменные, не зависящие от альтернатив. Тем самым достигается изменение коэффициентов при этих переменных в зависимости от альтернатив.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Каким образом линейная модель с латентной объясняющей переменной приводит к порядковой пробит-модели?
2. Каким методом оценивается порядковая пробит-модель? Что необходимо для однозначной идентификации коэффициентов латентной модели, лежащей в основе пробит-модели? Как производится прогнозирование по оцененной модели?
3. Что понимается под мультиномиальной логит-моделью? В каких ситуациях используется мультиномиальная модель? Каким образом оцениваются параметры мультиномиальной логит-модели? Как производится прогнозирование по оцененной модели?

Тема 4.3

ЦЕНЗУРИРОВАННЫЕ МОДЕЛИ РЕГРЕССИИ (ТОБИТ-МОДЕЛИ)

Модель тобит-I

Развивая пример с наличием или отсутствием у семьи собственного автомобиля, представим, что имеются следующие данные: для семей, обладающих автомобилем, известна стоимость этого автомобиля s_i (если в семье несколько автомобилей, то s_i — суммарная стоимость этих автомобилей). Таким образом, здесь наблюдаются пары $(x_i, price_observed_i)$, где x_i — среднедушевой месячный доход i -й семьи:

$$price_observed_i = \begin{cases} s_i, & \text{если } i\text{-я семья имеет автомобиль,} \\ 0, & \text{если } i\text{-я семья не имеет автомобиля.} \end{cases}$$

Обратимся к смоделированной выборке, состоящей из 1000 семей со среднедушевым месячным доходом от 100 до 1600 у.е. Для удобства наблюдения упорядочены в соответствии с возрастанием x_i , так что $x_1 \leq x_2 \leq \dots \leq x_{1000}$.

Диаграмма рассеяния для этих данных имеет весьма специфический вид (рис. 4.13).

Большое количество точек (их 418), расположенных на оси абсцисс, показывает, что 418 семей из 1000 рассматриваемых не имеют собственного автомобиля. В то же время у семей, владеющих автомобилем, минимальная цена автомобиля равна 2002 у.е., и это может означать, что на рынке, в том числе на вторичном, просто нет автомобилей с ценой менее 2000 у.е.

Как проводить статистический анализ подобных данных? Можно попытаться, например, использовать все 1000 наблюдений и по этим наблюдениям методом наименьших квадратов оценить линейную статистическую модель

$$price_observed_i = \alpha + \beta x_i + \varepsilon_i.$$

Результаты такого оценивания приведены в табл. 4.22.

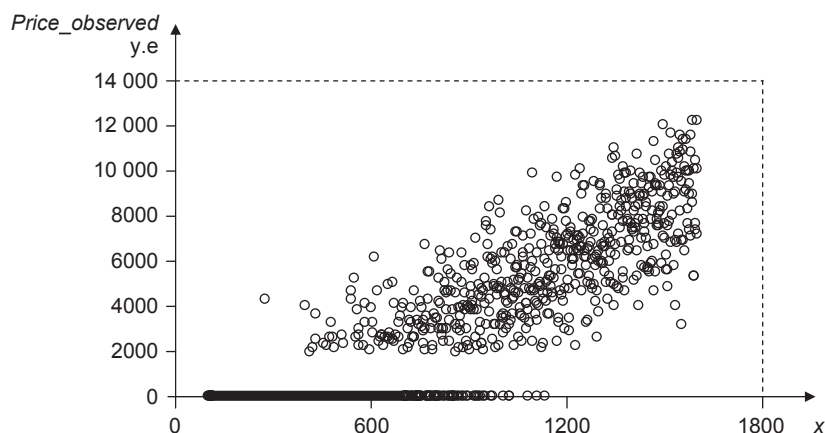


Рис. 4.13

Таблица 4.22

Оценка линейной статистической модели (1000 наблюдений)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
<i>C</i>	−2427.821	121.0156	−20.06205	0.0000
<i>X</i>	6.915595	0.126948	54.47591	0.0000
<i>R-squared</i>	0.748337			

В то же время можно проигнорировать наблюдения с $price_observed_i = 0$ и произвести оценивание той же линейной модели только по остальным наблюдениям (в количестве 582). При таком подходе получим результаты, представленные в табл. 4.23.

Таблица 4.23

Оценка линейной статистической модели (582 наблюдения)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
<i>C</i>	−1037.189	274.4903	−3.778599	0.0002
<i>X</i>	6.119677	0.233812	26.173530	0.0000
<i>R-squared</i>	0.541521	Mean dependent var		5919.6700

Рисунок 4.14 позволяет сравнить значения $price_observed_i$, прогнозные значения, полученные по первой модели (по 1000 наблюдениям), т.е.

$$pricef_1000_i = \hat{\alpha} + \hat{\beta}x_i = -2427.821 + 6.915595x_i,$$

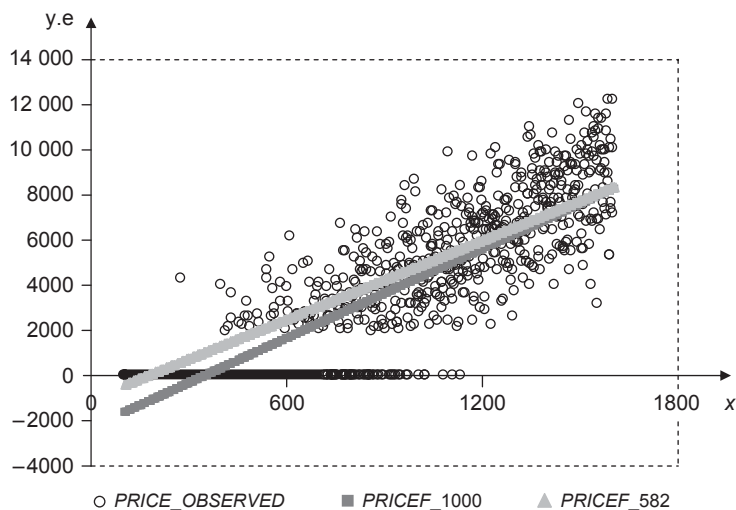


Рис. 4.14

и прогнозные значения, полученные по второй модели (по 582 наблюдениям), т.е.

$$pricef_582_i = \hat{\alpha} + \hat{\beta}x_i = -1037.189 + 6.119677x_i.$$

Конечно, по такой картине вряд ли можно говорить об адекватном представлении данных этими двумя моделями. Желательно построить модель процесса, который мог породить такого рода данные. Для этой цели можно опять использовать идею латентной переменной, но в данной ситуации, скорее, следовало бы говорить о частично наблюдаемой переменной.

Обращаясь к той же выборке, состоящей из 1000 семей, рассмотрим линейную модель наблюдений

$$price_i^* = \alpha + \beta x_i + \sigma \varepsilon_i, \quad i = 1, \dots, n,$$

где $price_i^*$ — цена, которую уплатила за покупку автомобиля (автомобилей) i -я семья, имеющая автомобиль, или цена, которую уплатила бы за покупку автомобиля i -я семья, не имеющая автомобиля, если бы эта семья решила приобрести его.

Естественно предполагать, что при этом $\beta > 0$, так что возрастание x_i приводит в среднем к возрастанию $price_i^*$. Однако существенное влияние других ненаблюдаемых факторов, объединяемых в случайную составляющую, может приводить к значительным отклонениям переменной $price_i^*$ от «средней линии» $price^* = \alpha + \beta x$. Возможные отрицательные значения $price_i^*$ свидетельствуют о наличии факторов, в той или иной степени препятствующих планированию каких бы то ни было расходов на покупку автомобиля.

Предположим, что i -я семья покупает автомобиль по цене $price_i^*$, если последняя превышает минимально возможную стоимость γ автомобиля на рынке (первичном и вторичном), т.е. если $price_i^* > \gamma$.

В такой модели значения переменной $price_i^*$ доступны лишь для части наблюдений — только для семей, имеющих автомобиль. Для остальных семей известно только, что $price_i^* \leq \gamma$. Такие данные называют **цензурированными** (*censored data*). В данном случае данные **цензурированы слева на уровне γ** , а модель получения этих данных называют **цензурированной линейной моделью** (*censored linear model*). При этом наблюдается цензурированная переменная

$$price_censored_i = \begin{cases} price_i^* & \text{если } price_i^* > \gamma, \\ \gamma & \text{если } price_i^* \leq \gamma. \end{cases}$$

Диаграмма рассеяния переменных x_i , $price_censored_i$ в нашем примере представлена на рис. 4.15.

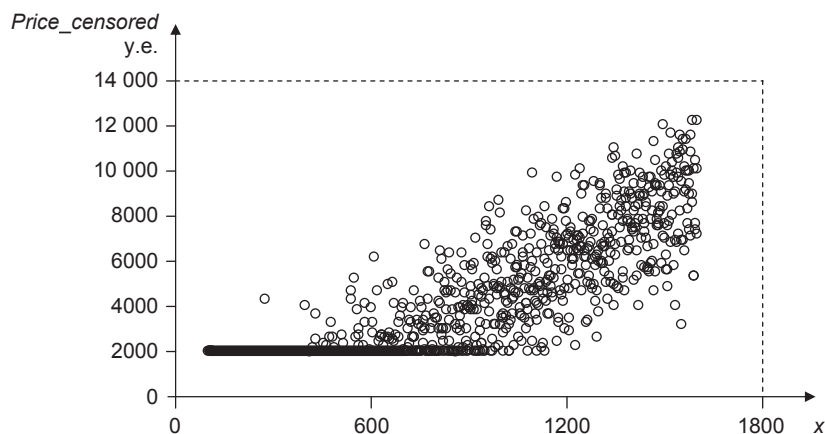


Рис. 4.15

Если значение γ известно, то вместо переменной $price_i^*$ можно рассмотреть переменную $y_i = price_i^* - \gamma$. Значения последней также наблюдаются только для семей, имеющих автомобиль. Для остальных семей положим $y_i = 0$, так что

$$y_i = \begin{cases} price - \gamma, & \text{если } price_i^* > \gamma, \\ 0, & \text{если } price_i^* \leq \gamma. \end{cases}$$

Диаграмма рассеяния переменных x_i , y_i в нашем примере изображена на рис. 4.16.

Теперь можно поставить вопрос о подходящем методе оценивания параметров цензурированных линейных моделей.

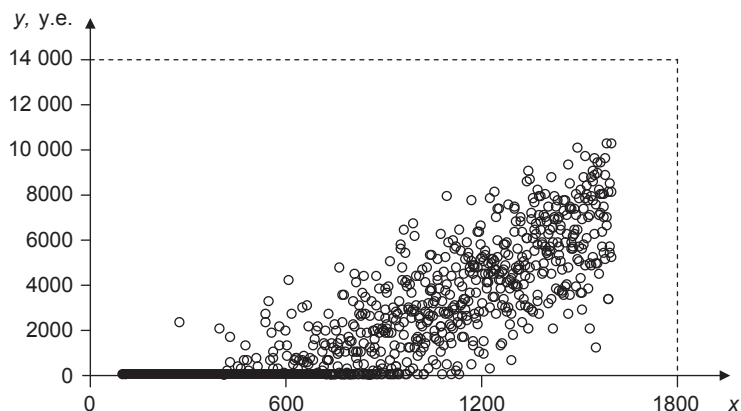


Рис. 4.16

Обычно при рассмотрении подобных ситуаций опираются на предположение о нормальности распределения ошибок ε_i . (Впрочем, имеющиеся пакеты статистических программ позволяют проводить статистический анализ и для других распределений ошибок. Например, в пакете *EViews* допускается вместо нормального распределения ошибок использовать логистическое распределение и распределение экстремальных значений первого типа.)

Предположим, что имеем дело с некоторым показателем y_i^* , значения которого наблюдаются только при условии $y_i^* > 0$ (в нашем примере в качестве такого показателя выступала переменная $(price_i^* - 2000)$). Пусть в правую часть модели для этого показателя включаются p объясняющих переменных (показателей, характеризующих i -й субъект), т.е.

$$y_i^* = \theta_1 x_{i1} + \dots + \theta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

где $\varepsilon_1, \dots, \varepsilon_n$ — ошибки, независимые в совокупности (и независимые от x_{ij} , $j = 1, \dots, p$) случайные величины, имеющие одинаковое нормальное распределение $\varepsilon_i \sim N(0, \sigma^2)$.

Наблюдаемыми являются значения x_{ij} , $j = 1, \dots, p$, $i = 1, \dots, n$, и значения переменной y_i ,

$$y_i = \begin{cases} y_i^*, & \text{если } y_i^* > 0, \\ 0, & \text{если } y_i^* \leq 0. \end{cases}$$

О такой цензурированной модели регрессии говорят как о **стандартной тобит-модели** (*tobit model*).

В стандартной тобит-модели для фиксированных значений x_{ij} , $j = 1, \dots, p$, имеем

$$y_i^* \sim N(\theta_1 x_{i1} + \dots + \theta_p x_{ip}, \sigma^2),$$

$$E(y_i^* | x_{ij}, j=1, \dots, p) = \theta_1 x_{i1} + \dots + \theta_p x_{ip},$$

т.е.

$$E(y_i^* | x_i) = x_i^T \theta,$$

где, как и ранее, обозначено $x_i = (x_{i1}, \dots, x_{ip})^T$, $\theta = (\theta_1, \dots, \theta_p)^T$. В нашем примере значение коэффициента θ_j определяет изменение ожидаемой суммы расходов на (возможную) покупку автомобиля для семьи с вектором показателей $x_i = (x_{i1}, \dots, x_{ip})^T$ при увеличении на единицу значения j -го показателя.

Если для оценивания коэффициентов θ_j использовать только наблюдения с $y_i > 0$, получим **усеченную модель регрессии** (*truncated regression*):

$$y_i = \theta_1 x_{i1} + \dots + \theta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n_1,$$

где n_1 — количество семей, имеющих автомобиль (среди всех n рассматриваемых семей).

Конечно, при переходе к усеченной модели придется заново нумеровать используемые n_1 наблюдений. В такой модели для значений $w > 0$ имеем

$$P\{y_i \leq w\} = P\{y_i^* \leq w | y_i^* > 0\} = \frac{P\{0 < y_i^* \leq w\}}{P\{y_i^* > 0\}},$$

где

$$P\{0 < y_i^* \leq w\} = P\left\{\frac{-x_i^T \theta}{\sigma} < \frac{y_i^* - x_i^T \theta}{\sigma} \leq \frac{w - x_i^T \theta}{\sigma}\right\} =$$

$$= \Phi\left(\frac{w - x_i^T \theta}{\sigma}\right) - \Phi\left(\frac{-x_i^T \theta}{\sigma}\right)$$

и

$$P\{y_i^* > 0\} = 1 - P\left\{\frac{y_i^* - x_i^T \theta}{\sigma} \leq \frac{-x_i^T \theta}{\sigma}\right\} = 1 - \Phi\left(\frac{-x_i^T \theta}{\sigma}\right) = \Phi\left(\frac{x_i^T \theta}{\sigma}\right).$$

Если взять производную $\frac{dP\{y_i \leq w\}}{dw}$, получим функцию плотности распределения случайной величины y_i (условного при заданном x_i):

$$p_{y_i}(w) = \frac{1}{\sigma} \frac{\varphi\left(\frac{w - x_i^T \theta}{\sigma}\right)}{\Phi\left(\frac{x_i^T \theta}{\sigma}\right)}.$$

Отсюда получаем выражение для условного математического ожидания:

$$E(y_i|x_i) = \int_0^{\infty} w p_{y_i}(w) dw = x_i^T \theta + \sigma \lambda \left(\frac{x_i^T \theta}{\sigma} \right),$$

где обозначено $\lambda(z) = \frac{\varphi(z)}{\Phi(z)}$.

Таким образом, $E(y_i|x_i)$ — нелинейная функция от x_i и θ , причем

$$E(y_i|x_i) > x_i^T \theta.$$

Рассмотрим теперь другой подход к оцениванию коэффициентов исходной модели

$$y_i^* = \theta_1 x_{i1} + \dots + \theta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n,$$

при котором неполные наблюдения не отбрасываются, а учитываются при оценивании. В рамках этого подхода в качестве объясняемой берем переменную

$$y_i = \begin{cases} x_i^T \theta + \varepsilon_i, & \text{если } x_i^T \theta + \varepsilon_i > 0, \\ 0, & \text{если } x_i^T \theta + \varepsilon_i \leq 0. \end{cases}$$

В этом случае

$$P\{y_i = 0 | x_i\} = P\{\varepsilon_i \leq -x_i^T \theta\} = \Phi \left(-\frac{x_i^T \theta}{\sigma} \right) = 1 - \Phi \left(\frac{x_i^T \theta}{\sigma} \right),$$

а для $w > 0$

$$P\{y_i \leq w | x_i\} = P \left\{ \frac{y_i - x_i^T \theta}{\sigma} \leq \frac{w - x_i^T \theta}{\sigma} \right\} = \Phi \left(\frac{w - x_i^T \theta}{\sigma} \right).$$

Это приводит к следующему выражению для условного математического ожидания y_i :

$$\begin{aligned} E(y_i|x_i) &= 0 \left(1 - \Phi \left(\frac{x_i^T \theta}{\sigma} \right) \right) + \int_0^{\infty} w \varphi \left(\frac{w - x_i^T \theta}{\sigma} \right) dw \cdot \Phi \left(\frac{x_i^T \theta}{\sigma} \right) = \\ &= \left(x_i^T \theta + \sigma \lambda \left(\frac{x_i^T \theta}{\sigma} \right) \right) \cdot \Phi \left(\frac{x_i^T \theta}{\sigma} \right). \end{aligned}$$

Оно отличается от выражения для $E(y_i|x_i)$ в усеченной модели умножением последнего на $\Phi \left(\frac{x_i^T \theta}{\sigma} \right)$, т.е. на величину, меньшую 1. Раскрыв скобки в правой части, получим

$$\begin{aligned}
 E(y_i|x_i) &= x_i^T \theta \Phi\left(\frac{x_i^T \theta}{\sigma}\right) + \sigma \lambda\left(\frac{x_i^T \theta}{\sigma}\right) \cdot \Phi\left(\frac{x_i^T \theta}{\sigma}\right) = \\
 &= x_i^T \theta \Phi\left(\frac{x_i^T \theta}{\sigma}\right) + \sigma \varphi\left(\frac{x_i^T \theta}{\sigma}\right).
 \end{aligned}$$

Предельный эффект изменения переменной x_{ij} равен:

$$\frac{\partial E(y_i|x_i)}{\partial x_{ij}} = \theta_j \Phi\left(\frac{x_i^T \theta}{\sigma}\right),$$

т.е. меньше значения коэффициента θ_j в исходной модели: он получается умножением этого коэффициента на вероятность того, что $y_i^* > 0$.

Заметим в связи с этим, что если $\tilde{E}(y_i|x_i)$ — условное математическое ожидание значения y_i в усеченной модели, то для него

$$\frac{\partial \tilde{E}(y_i|x_i)}{\partial x_{ij}} = \theta_j [1 - z\lambda(z) - \lambda^2(z)],$$

где $\lambda(z) = \frac{\varphi(z)}{\Phi(z)}$, $z = \frac{x_i^T \theta}{\sigma}$.

Продолжим рассмотрение смоделированной выборки, состоящей из 1000 семей, 582 из которых имеют автомобиль. Подберем к тем же данным усеченную и цензурированную модели.

Заметим, что если переменная $y_i^* = price_i^* - 2000$ порождается моделью $y_i^* = \alpha + \beta x_i + \varepsilon_i$, $i = 1, \dots, 1000$, то сама переменная $price_i^*$ порождается моделью $price_i^* = (\alpha + 2000) + \beta x_i + \varepsilon_i$. Поэтому достаточно произвести оценивание коэффициентов модели $y_i^* = \alpha + \beta x_i + \varepsilon_i$, опираясь на данные (x_i, y_i) . Такое оценивание дает результаты, представленные в табл. 4.24 (для усеченной модели) и табл. 4.25 (для цензурированной модели).

Таблица 4.24

Оценка усеченной модели

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
C	-5710.678	480.1485	-11.89357	0.0000
X	8.103471	0.376079	21.54728	0.0000
Error Distribution				
σ	1822.273	66.21537	27.52040	0.0000

Таблица 4.25

Оценка цензурированной модели

Переменная	Коэффициент	Стандартная ошибка	z-статистика	P-значение
C	-6041.883	233.5302	-25.87195	0.0000
X	8.363125	0.209276	39.96215	0.0000
<i>Error Distribution</i>				
σ	1823.565	53.95272	33.79933	0.0000

Получаем следующие оцененные модели для прогноза значений переменной $price_i^*$:

$$price_i^* = -3710.678 + 8.103471x_i \text{ (усеченная модель),}$$

$$price_i^* = -4041.883 + 8.363125x_i \text{ (цензурированная модель).}$$

Дисперсии случайных составляющих оцениваются соответственно как 1822.273^2 и 1823.565^2 . Заметим, что «теоретическая» модель, по которой генерировались данные, имела вид:

$$price_i^* = -3600 + 8x_i + 1800u_i,$$

где u_1, \dots, u_{1000} — независимые случайные величины, имеющие одинаковое стандартное нормальное распределение $N(0, 1)$.

На рис. 4.17 для сравнения показан разброс значений переменной $price_i^*$ и прогнозных значений для этой переменной, полученных по оцененной усеченной модели ($price_starf_trun$) и по оцененной цензурированной модели ($price_starf_cens$). Отметим, что прогнозные значения, полученные по двум оцененным моделям, весьма близки.

На рис. 4.18 проиллюстрирован разброс значений переменной y_i и ожидаемых значений переменной y_i , рассчитанных по двум оцененным моделям.

Отметим, что для $x_i \geq 1330$ ожидаемые значения y_i , рассчитанные по цензурированной модели, больше ожидаемых значений y_i , рассчитанных по усеченной модели. Однако это различие практически незаметно. В то же время для $x_i < 1330$ ожидаемые значения y_i , рассчитанные по цензурированной модели, меньше ожидаемых значений y_i , рассчитанных по усеченной модели, причем это различие становится весьма заметным при уменьшении значений x_i .

Заметим также, что ожидаемые значения y_i , рассчитанные и по усеченной, и по цензурированной моделям, положительны для всех 1000 наблюдений, тогда как это не выполняется для линейных моделей, подобранных методом наименьших квадратов.

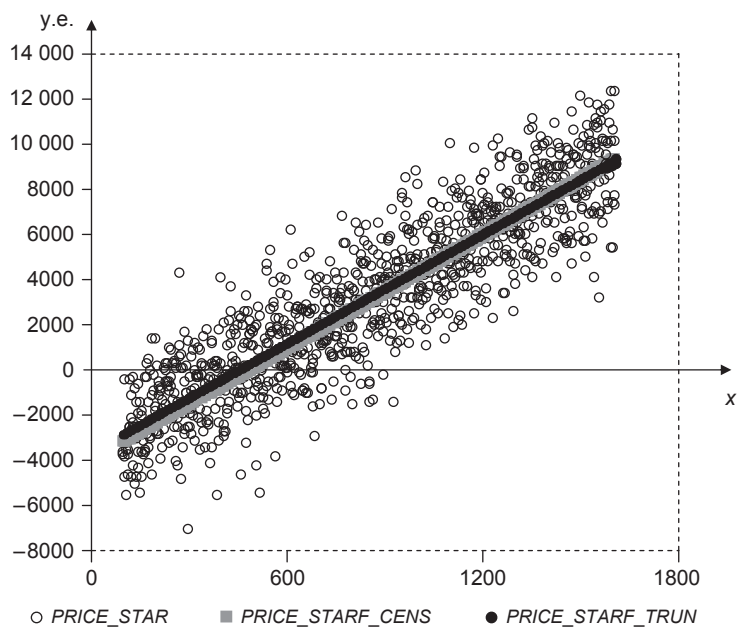


Рис. 4.17

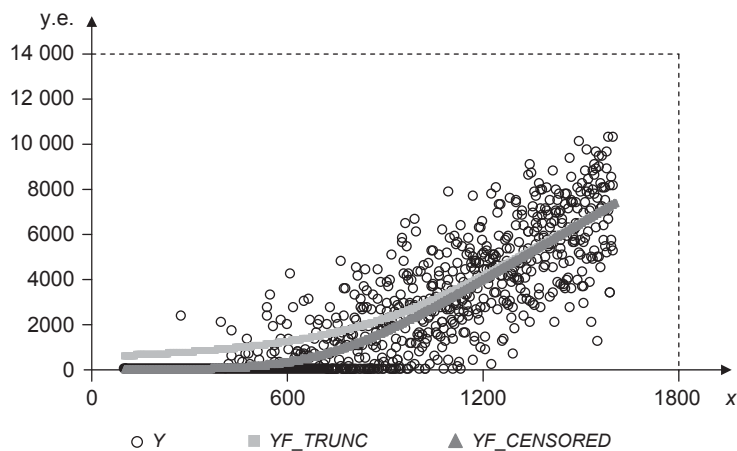


Рис. 4.18

Так, оценивание модели $y_i = \alpha + \beta x_i + \varepsilon_i$ по всем 1000 наблюдениям обычным методом наименьших квадратов дает результаты, представленные в табл. 4.26.

Для $x_i \leq 470$ подобранная модель прогнозирует отрицательные значения объясняемой переменной. При подгонке такой модели методом наименьших квадратов по 582 наблюдениям получаем результаты, приведенные в табл. 4.27.

Таблица 4.26

Оценка линейной модели (1000 наблюдений)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	-2075.806	104.7679	-19.81338	0.0000
X	5.130473	0.109904	46.68158	0.0000

Таблица 4.27

Оценка линейной модели (582 наблюдения)

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	-3037.189	274.4903	-11.06483	0.0000
X	6.119667	0.233812	26.17353	0.0000

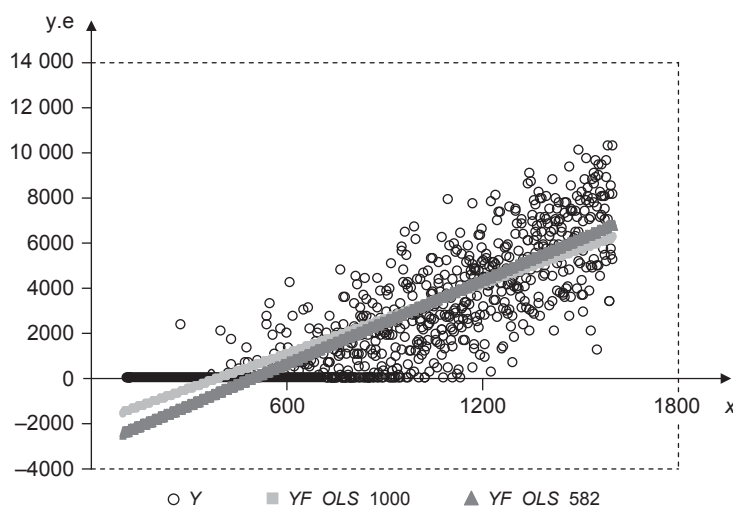


Рис. 4.19

Оцененная модель прогнозирует отрицательные значения объясняемой переменной для значений $x_i \leq 498$. Это положение иллюстрирует рис. 4.19.

Как было сказано ранее (во второй части), одним из показателей качества прогноза произвольного временного ряда z_i , $i = 1, \dots, n$, является его **средняя абсолютная процентная ошибка** (*MAPE* — *mean squared absolute error*). Она определяется следующим образом: если \hat{z}_i — прогнозное значение для z_i , то

$$MAPE = \frac{1}{n} \sum_{i=1}^n 100 \left| \frac{z_i - \hat{z}_i}{z_i} \right|.$$

Сравним качество полученных альтернативных прогнозов для y_i с точки зрения средней абсолютной процентной ошибки (табл. 4.28). Как видно из этой таблицы, наилучшее качество имеют прогнозы, полученные с использованием цензурированной модели регрессии.

Таблица 4.28

Прогнозы, полученные с помощью разных моделей регрессии

Модель	<i>OLS</i> _582	<i>OLS</i> _1000	<i>Truncated</i>	<i>Censored</i>
MAPE %	118.46	99.86	126.69	71.96

Обратим внимание еще на одно обстоятельство. Выше было отмечено, что

$$E(y_i|x_i) = \left(x_i^T \theta + \sigma \lambda \left(\frac{x_i^T \theta}{\sigma} \right) \right) \cdot \Phi \left(\frac{x_i^T \theta}{\sigma} \right) = \tilde{E}(y_i|x_i) \cdot \Phi \left(\frac{x_i^T \theta}{\sigma} \right),$$

где $\tilde{E}(y_i|x_i)$ — условное математическое ожидание значения y_i в усеченной модели.

Отсюда получаем следующее разложение:

$$\frac{\partial E(y_i|x_i)}{\partial x_{ij}} = \Phi \left(\frac{x_i^T \theta}{\sigma} \right) \frac{\partial \tilde{E}(y_i|x_i)}{\partial x_{ij}} + \tilde{E}(y_i|x_i) \frac{\partial \Phi \left(\frac{x_i^T \theta}{\sigma} \right)}{\partial x_{ij}}.$$

Первое слагаемое отражает изменение в ожидаемых значениях $y_i > 0$, взвешенное с весом $\Phi \left(\frac{x_i^T \theta}{\sigma} \right) = P\{y_i > 0\}$, а второе — изменение вероятности

$P\{y_i > 0\}$, взвешенное с весом, равным $\tilde{E}(y_i|x_i)$. Заметим в связи с этим, что

$$\frac{\partial P\{y_i > 0\}}{\partial x_{ij}} = \frac{\partial \Phi \left(\frac{x_i^T \theta}{\sigma} \right)}{\partial x_{ij}} = \frac{1}{\sigma} \varphi \left(\frac{x_i^T \theta}{\sigma} \right) \theta_j.$$

В нашем примере изменение $\tilde{E}(y_i|x_i)$ показано на рис. 4.20 (по оси абсцисс на этом и следующих 5 графиках откладываются значения среднемесячного дохода на одного члена семьи).

Изменение производной $\frac{\partial P\{y_i > 0\}}{\partial x_{ij}}$ показано на рис. 4.21.

Входящие в разложение для $\frac{\partial E(y_i|x_i)}{\partial x_{ij}}$ слагаемые изображены на рис. 4.22

и 4.23.

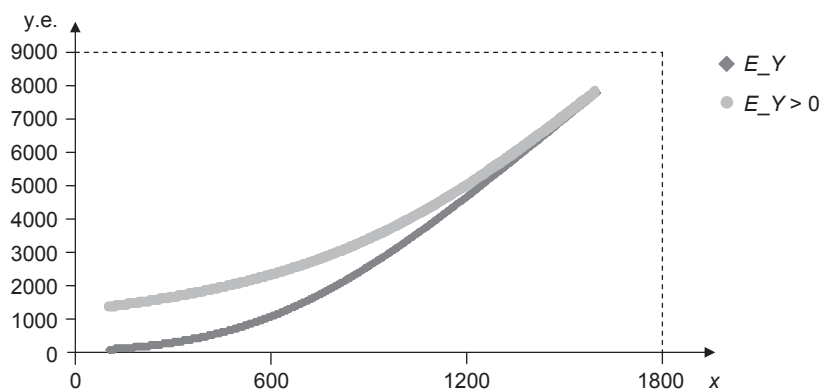


Рис. 4.20

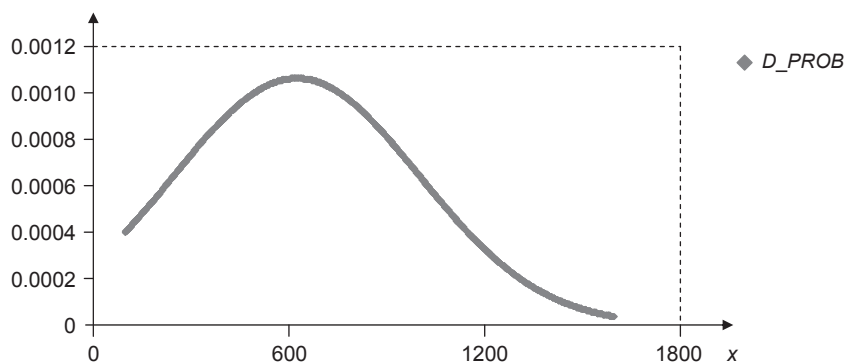


Рис. 4.21

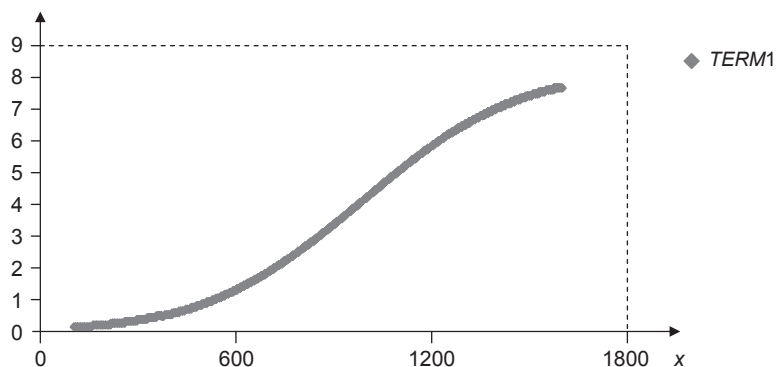


Рис. 4.22

В сумме они дают функцию $\frac{\partial E(y_i|x_i)}{\partial x_{ij}} = \theta_j \Phi\left(\frac{x_i^T \theta}{\sigma}\right)$, которая представлена на рис. 4.24.

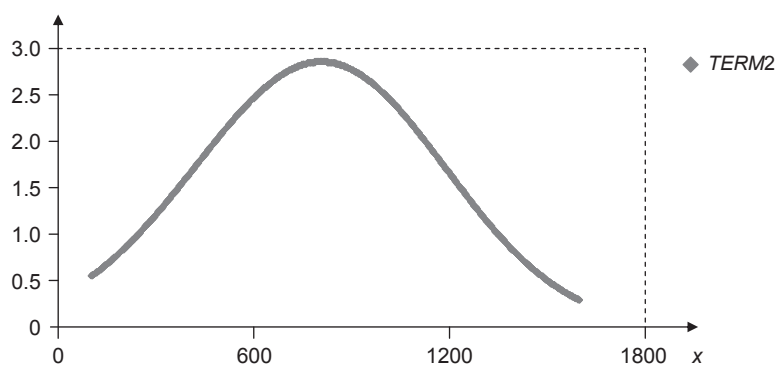


Рис. 4.23

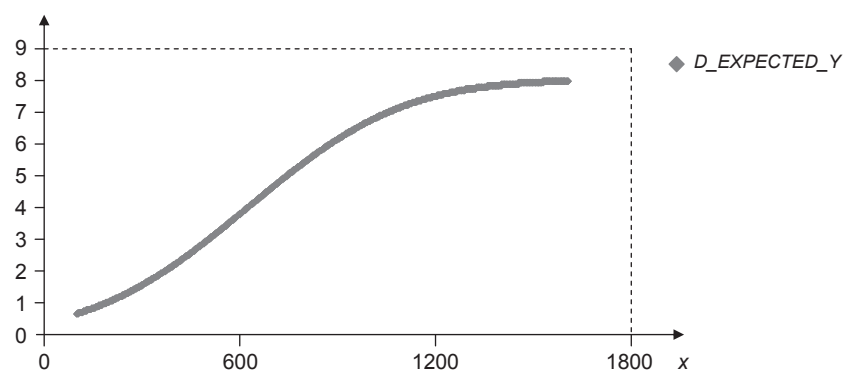


Рис. 4.24

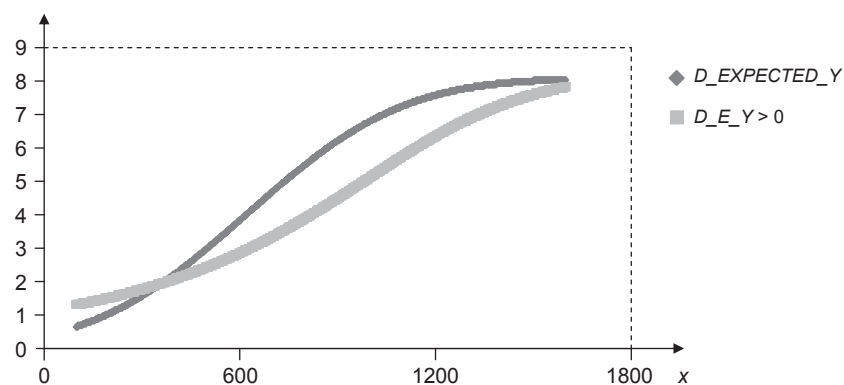


Рис. 4.25

Рисунок 4.25 позволяет сравнить влияние единичного возрастания дохода на ожидаемые значения y_i во всей популяции ($D_EXPECTED_Y$) и среди семей с $y_i > 0$ ($D_E_Y > 0$).

Модель тобит-II

Выше была рассмотрена линейная модель наблюдений

$$price_i^* = \alpha + \beta x_i + \sigma \varepsilon_i, \quad i = 1, \dots, n,$$

где $price_i^*$ — цена, которую уплатила за покупку автомобиля (автомобилей) i -я семья, имеющая автомобиль, или цена, которую уплатила бы за покупку автомобиля i -я семья, не имеющая автомобиля, если бы эта семья решила приобрести его.

При этом предполагали, что i -я семья покупает автомобиль по цене $price_i^*$, если $price_i^* > \gamma$. Таким образом, в этой модели решение о приобретении или неприобретении автомобиля определяется самой ценой, по которой предполагается приобрести автомобиль. В то же время можно рассмотреть и другую модель, в которой процесс принятия решения о стоимости покупаемого автомобиля отделен от процесса принятия решения о покупке автомобиля.

Пусть имеем дело с некоторым показателем y_i^* , значения которого наблюдаются не для всех i . Значение y_i^* наблюдается, если выполнено условие $h_i^* > 0$, где h_i^* — некоторая функция полезности. Предположим, что:

$$y_i^* = x_{1i}^T \theta_1 + \varepsilon_{1i}, \quad i = 1, \dots, n,$$

$$h_i^* = x_{2i}^T \theta_2 + \varepsilon_{2i}, \quad i = 1, \dots, n,$$

где $x_{1i} = (x_{11,i}, \dots, x_{1p_1,i})^T$ — вектор значений p_1 объясняющих переменных в уравнении для y_i^* ;

$\theta_1 = (\theta_{11}, \dots, \theta_{1p_1})^T$ — вектор коэффициентов при этих переменных;

$x_{2i} = (x_{21,i}, \dots, x_{2p_2,i})^T$ — вектор значений p_2 объясняющих переменных в уравнении для h_i^* ;

$\theta_2 = (\theta_{21}, \dots, \theta_{2p_2})^T$ — вектор коэффициентов при этих переменных.

Случайные составляющие ε_{1i} и ε_{2i} могут быть *коррелированными*, так что $Cov(\varepsilon_{1i}, \varepsilon_{2i}) \neq 0$. Следуя обычной практике, будем предполагать, что двумерные случайные векторы $(\varepsilon_{1i}, \varepsilon_{2i})^T$, $i = 1, \dots, n$, независимы в совокупности и имеют одинаковое двумерное нормальное распределение $N_2(0, \Sigma)$ с нулевым вектором математических ожиданий и ковариационной матрицей

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

т.е.

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim i.i.d. N_2 \left(0, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right).$$

Для нормализации функции полезности полагаем $\sigma_2 = 1$.

Наблюдаемыми являются:

- значения объясняющих переменных $x_{1j,i}, j = 1, \dots, p_1, x_{2j,i}, j = 1, \dots, p_2, i = 1, \dots, n$;
- значения переменной h_i :

$$h_i = \begin{cases} 1, & \text{если } h_i^* > 0, \\ 0, & \text{если } h_i^* \leq 0; \end{cases}$$

- значения переменной y_i :

$$y_i = \begin{cases} y_i^* & \text{если } h_i = 1, \\ 0, & \text{если } h_i = 0. \end{cases}$$

Определенную таким образом модель называют **стандартной тобит-II моделью** (*Tobit II model*). Соответственно о модели, рассмотренной в предыдущем разделе, в этом контексте говорят как о **стандартной тобит-I модели** (*standard Tobit I model*).

✓ **Замечание 4.3.1.** Объясняющие переменные в уравнениях для y_i^* и h_i^* могут быть как одинаковыми, так и различными. В ряде ситуаций экономическая аргументация указывает на необходимость включения в правую часть уравнения для h_i^* («уравнение выбора») всех переменных, включенных в правую часть уравнения для y_i^* . При этом коэффициенты при одной и той же переменной в уравнениях для y_i^* и h_i^* могут быть различными.

Если предположить, что $x_{1i}^T \theta_1 = x_{2i}^T \theta_2$ и $\varepsilon_{1i} = \varepsilon_{2i}$, то возвращаемся к стандартной тобит-I модели, рассмотренной в предыдущем разделе.

Обращаясь опять к примеру с автомобилями, можно рассмотреть, например, модели, в которых значение $price_i^*$ определяется по той же формуле:

$$price_i^* = \alpha + \beta x_i + \sigma \varepsilon_i, \quad i = 1, \dots, n,$$

но наличие автомобиля соответствует выполнению соотношения $h_i^* > 0$, в котором $h_i^* = \gamma + \delta x_i + u_i$, или, например, $h_i^* = \gamma + \delta x_i + \kappa d_{man} + u_i$, где $d_{man} = 1$, если главой семьи является мужчина, и $d_{man} = 0$, если главой семьи является женщина.

Прежде всего заметим, что (при фиксированных значениях x_{1i}, x_{2i}):

$$\begin{aligned} E\{y_i | h_i = 1\} &= x_{1i}^T \theta_1 + E\{\varepsilon_{1i} | h_i = 1\} = x_{1i}^T \theta_1 + E\{\varepsilon_{1i} | \varepsilon_{2i} > -x_{2i}^T \theta_2\} = \\ &= x_{1i}^T \theta_1 + \frac{\sigma_{12}}{\sigma_2^2} E\{\varepsilon_{2i} | \varepsilon_{2i} > -x_{2i}^T \theta_2\} = x_{1i}^T \theta_1 + \sigma_{12} \lambda(x_{2i}^T \theta_2), \end{aligned}$$

где, как и ранее, $\lambda(z) = \frac{\varphi(z)}{\Phi(z)}$.

Если $\sigma_{12} = 0$, то

$$E\{y_i | h_i = 1\} = x_{1i}^T \theta_1.$$

Это означает, что если ε_{1i} и ε_{2i} не коррелированы между собой, то можно, игнорируя уравнение для h_i^* , производить непосредственное оценивание уравнения регрессии

$$y_i = x_{1i}^T \theta_1 + \varepsilon_{1i}$$

методом наименьших квадратов по наблюдениям с $h_i = 1$. Это приводит к состоятельному оцениванию значений $x_{1i}^T \theta_1$.

Однако если $\sigma_{12} \neq 0$, то при таком оценивании возникает смещение оценки $x_{1i}^T \theta_1$, пропорциональное величине $\lambda(x_{2i}^T \theta_2)$, которую называют в этом контексте **лямбдой Хекмана** (*Heckman's lambda*).

Получить состоятельные и асимптотически эффективные оценки параметров модели тобит-II можно, используя метод максимального правдоподобия, при котором соответствующая функция правдоподобия максимизируется по всем возможным значениям параметров модели θ_1 , θ_2 , σ_1 , σ_2 . Однако чаще такую модель оценивают, используя **двухшаговую процедуру Хекмана** (*Heckman's two-stage procedure*). Она проста в вычислительном отношении и дает хорошие стартовые значения для итерационной процедуры максимизации функции правдоподобия.

Идея Хекмана состоит в использовании приведенного выше соотношения

$$E\{y_i | h_i = 1\} = x_{1i}^T \theta_1 + \sigma_{12} \lambda(x_{2i}^T \theta_2)$$

и построении на его основе модели регрессии

$$y_i = x_{1i}^T \theta_1 + \sigma_{12} \lambda_i + v_i,$$

где λ_i — переменная, определяемая соотношением:

$$\lambda_i = \lambda(x_{2i}^T \theta_2) = \frac{\varphi(x_{2i}^T \theta_2)}{\Phi(x_{2i}^T \theta_2)}.$$

Если ε_{1i} не коррелирована с x_{1i} и x_{2i} , то v_i не коррелирована с x_{1i} и λ_i , так что эту модель регрессии можно оценивать методом наименьших квадратов. Проблема, однако, в том, что значения λ_i не наблюдаются, поскольку неизвестен вектор коэффициентов θ_2 в модели выбора.

Оценивание вектора θ_2 производится в рамках пробит-модели бинарного выбора. При этом получаем оцененные значения $\hat{\lambda}_i = \lambda(x_{2i}^T \hat{\theta}_2)$ (первый шаг процедуры Хекмана). Эти оцененные значения используются затем на втором шаге процедуры вместо λ_i . Модель $y_i = x_{1i}^T \theta_1 + \sigma_{12} \hat{\lambda}_i + v_i$ оценивается методом наименьших квадратов, в результате получаем состоятельные (хотя и неэф-

фективные) оценки для θ_1 и σ_{12} . Используя эти оценки, получаем оцененное ожидаемое значение y_i при заданных x_{1i} , x_{2i} и $h_i = 1$ в виде:

$$\hat{E}\{y_i | x_{1i}, x_{2i}, h_i = 1\} = x_{1i}^T \hat{\theta}_1 + \hat{\sigma}_{12} \lambda(x_{2i}^T \hat{\theta}_2).$$

Если же нас интересует ожидаемое значение y_i при заданных x_{1i} , x_{2i} без условия $h_i = 1$, то оно оценивается величиной:

$$\hat{E}\{y_i | x_{1i}, x_{2i}\} = x_{1i}^T \hat{\theta}_1.$$

✓ **Замечание 4.3.2.** Поскольку смещение при оценивании уравнения для y_i^* методом наименьших квадратов вызывается коррелированностью ε_{1i} и ε_{2i} , представляет интерес проверка гипотезы

$$H_0: \sigma_{12} = 0$$

об отсутствии такой коррелированности в рамках модели, оцененной на втором шаге. Отметим только, что при проверке этой гипотезы следует производить коррекцию значений стандартных ошибок оценок, учитывающую гетероскедастичность модели и тот факт, что вместо переменной λ_i на втором шаге используется предварительно оцененная переменная $\hat{\lambda}_i$.

Заметим, наконец, что в стандартной тобит-II модели функция правдоподобия имеет вид:

$$L(\theta_1, \theta_2, \sigma_1, \sigma_{12}) = \prod_{i=1}^n (P\{h_i = 0\})^{1-h_i} (P\{h_i = 1\} f(y_i | h_i = 1))^{h_i},$$

где $f(y_i | h_i = 1)$ — условная плотность распределения случайной величины y_i при $h_i = 1$.

Здесь

$$P\{h_i = 0\} = 1 - \Phi(x_{2i}^T \theta_2),$$

$$P\{h_i = 1\} f(y_i | h_i = 1) = P\{h_i = 1 | y_i\} f(y_i),$$

$$P\{h_i = 1 | y_i\} = \Phi \left(\frac{x_{2i}^T \theta_2 + \frac{\sigma_{12}}{\sigma_1^2} (y_i - x_{1i}^T \theta_1)}{\sqrt{1 - \frac{\sigma_{12}^2}{\sigma_1^2}}} \right),$$

$$f(y_i) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp \left(-\frac{(y_i - x_{1i}^T \theta_1)^2}{2\sigma_1^2} \right).$$

Для начала итерационной процедуры в качестве стартовых можно взять оценки параметров, полученные в процессе реализации двухшаговой процедуры Хекмана.

ПРИМЕР 4.3.1

Пусть в примере с автомобилями наличие у семьи собственного автомобиля определяется условием $w_i^* > 2000$, где

$$w_i^* = -3600 + 8x_i + 1800\varepsilon_{2i}, \quad \varepsilon_{21}, \dots, \varepsilon_{2,1000} \sim i.i.d. N(0, 1).$$

Обозначив $h_i^* = w_i^* - 2000$, запишем это условие в виде $h_i^* > 0$, где

$$h_i^* = -5600 + 8x_i + 1800\varepsilon_{2i},$$

и нормализуем функцию полезности, разделив обе части последнего равенства на 1800:

$$h_i^* = -3.111 + 0.00445x_i + \varepsilon_{2i}.$$

Пусть «потенциальная цена» автомобиля для i -й семьи определяется уравнением:

$$price_i^* = 4000 + 6x_i + \varepsilon_{1i}, \quad \varepsilon_{11}, \dots, \varepsilon_{1,1000} \sim i.i.d. N(0, 1000^2).$$

В смоделированной выборке пары $(\varepsilon_{11}, \varepsilon_{21}), \dots, (\varepsilon_{1,1000}, \varepsilon_{2,1000})$ взаимно независимы, но $Cov(\varepsilon_{1i}, \varepsilon_{2i}) = 707$, так что коэффициент корреляции случайных величин $\varepsilon_{1i}, \varepsilon_{2i}$ равен: $\rho_{12} = 0.707$.

В принятых выше общих обозначениях модели тобит-II получаем:

$$y_i^* = \theta_{11}x_{11,i} + \theta_{12}x_{12,i} + \varepsilon_{1i}, \quad h_i^* = \theta_{21}x_{21,i} + \theta_{22}x_{22,i} + \varepsilon_{2i},$$

где $x_{11,i} = x_{21,i} = 1$, $x_{12,i} = x_{22,i} = x_i$, $\theta_{11} = 4000$, $\theta_{12} = 6$, $\theta_{21} = -3.111$, $\theta_{22} = 0.00445$. При этом $\sigma_1 = 1000$, $\sigma_2 = 1$, $\sigma_{12} = 707$.

Применив к смоделированным данным двухшаговую процедуру Хекмана, получим на первом шаге оцененное уравнение:

$$h_i^* = -3.450 + 0.00476x_i,$$

а на втором шаге — оцененное уравнение:

$$price_i^* = 3936.2 + 5.995x_i.$$

Используя полученные оценки параметров в качестве стартовых значений итерационной процедуры максимального правдоподобия, приходим к уравнениям:

$$h_i^* = -3.483 + 0.00480x_i,$$

$$price_i^* = 4159.3 + 5.828x_i.$$

При этом получим также $\hat{\sigma}_1 = 1010.7$, $\hat{\rho}_{12} = 0.598$.

Как видим, оцененные значения параметров достаточно близки к значениям, при которых производилось порождение данных. Графики, иллюстрирующие полученные результаты, приведены на рис. 4.26 — 4.30. ■

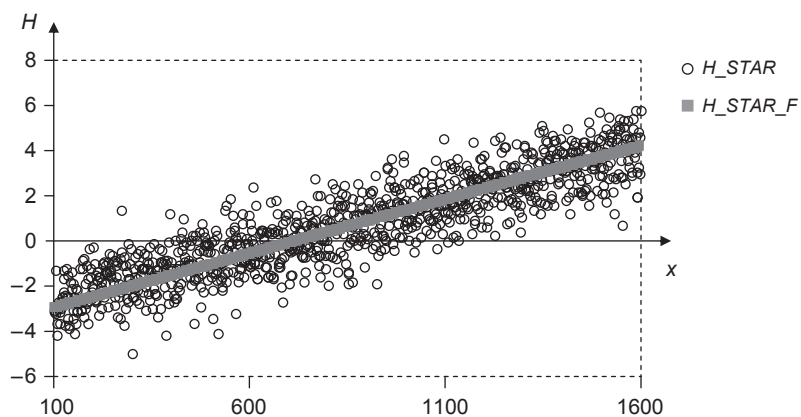


Рис. 4.26

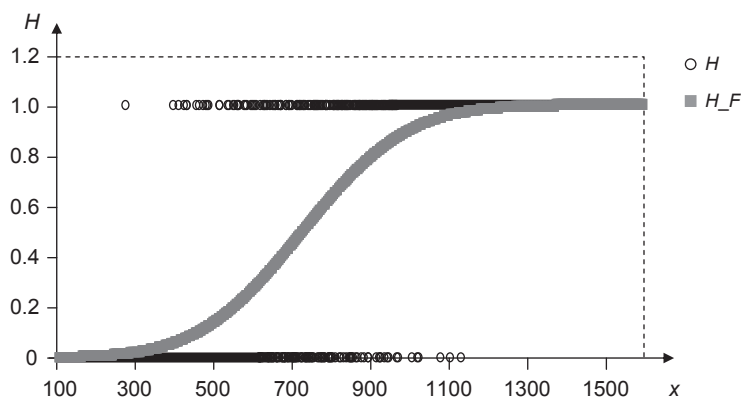


Рис. 4.27

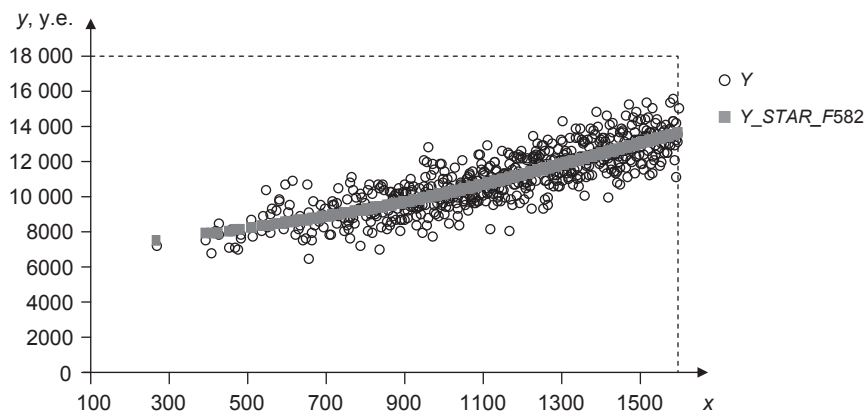


Рис. 4.28

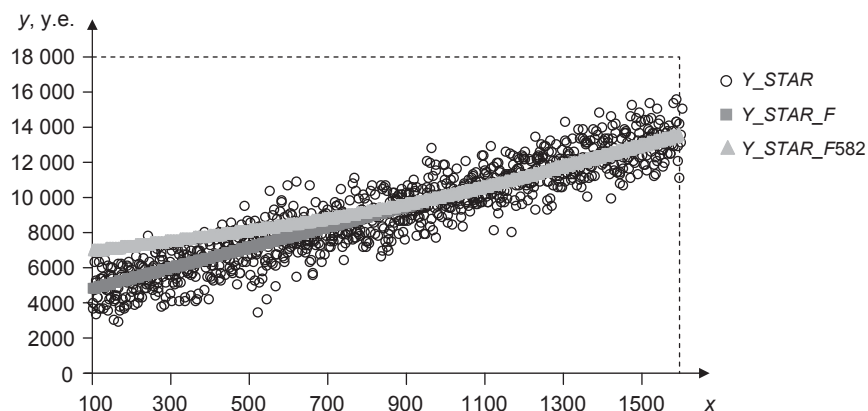


Рис. 4.29

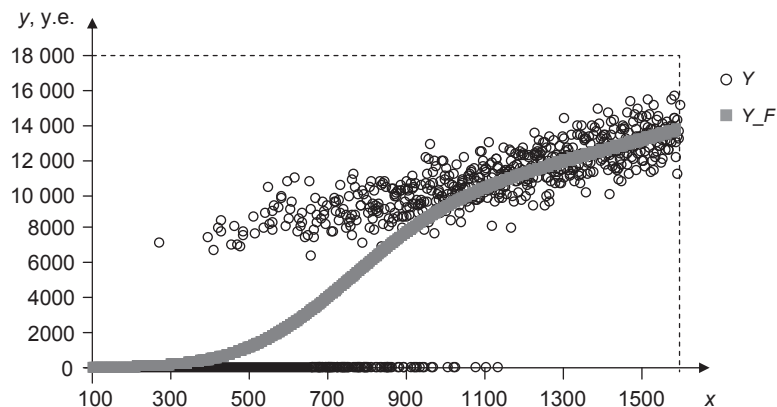


Рис. 4.30

ПРИМЕР 4.3.2

В условиях примера 4.3.1 смоделируем данные с измененной функцией полезности, полагая теперь

$$h_i^* = -4 + 0.003x_i + 2(d_{man})_i + \varepsilon_{2i},$$

где $d_{man} = 1$, если главой семьи является мужчина, и $d_{man} = 0$, если главой семьи является женщина.

Применив к новым смоделированным данным двухшаговую процедуру Хекмана, получим на первом шаге оцененное уравнение:

$$h_i^* = -4.280 + 0.00297x_i + 2.347(d_{man})_i,$$

а на втором шаге — оцененное уравнение:

$$price_i^* = 3879.97 + 6.124x_i.$$

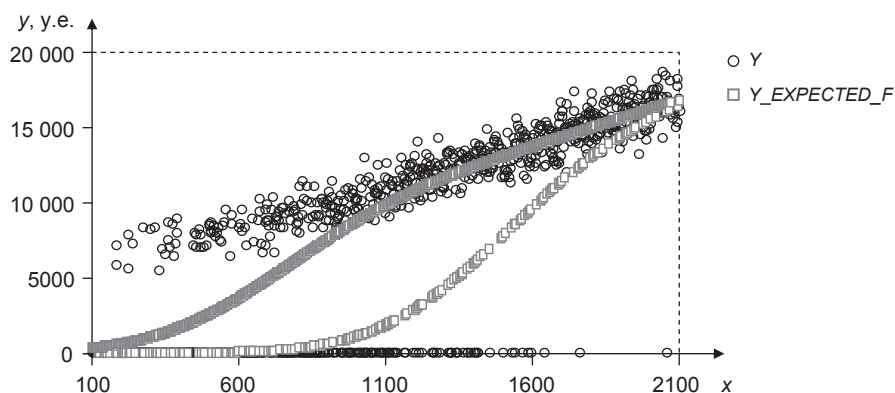


Рис. 4.31

При этом получим также $\hat{\sigma}_1 = 984.2$, $\hat{\rho}_{12} = 0.643$.

Как видим, и здесь оцененные значения параметров достаточно близки к значениям, при которых производилось порождение данных. На рис. 4.31 приведены для сравнения наблюдаемые значения переменной y_i и оцененные ожидаемые значения этой переменной ($Y_EXPECTED_F$).

Обратим внимание на две ветви графика оцененных ожидаемых значений y_i . Верхняя ветвь соответствует семьям, которые возглавляют мужчины, а нижняя — семьям, которые возглавляют женщины. ■

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какие статистические данные называют цензурированными? Какую модель называют цензурированной линейной?
2. Какую модель называют усеченной моделью регрессии?
3. Какую модель называют стандартной тобит-моделью? Как интерпретируются коэффициенты в стандартной тобит-модели? Как производится оценивание коэффициентов?
4. Какую модель называют стандартной тобит-II? Каковы особенности оценивания коэффициентов такой модели? Как устроена двухшаговая процедура Хекмана?

Тема 4.4

МОДЕЛИ БИНАРНОГО ВЫБОРА ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ

Выше уже рассматривались модели такого рода, но только для данных, относящихся к одному-единственному моменту (периоду) времени (*cross-section data* — данные в сечениях).

Перейдем теперь к панельным данным. И здесь модель бинарного выбора обычно связывается с наличием некоторой ненаблюдаемой (латентной) переменной y_{it}^* и наблюдаемой индикаторной переменной y_{it} , такой, что

$$y_{it} = \begin{cases} 1, & \text{если } y_{it}^* > 0, \\ 0, & \text{если } y_{it}^* \leq 0. \end{cases}$$

Например, y_{it} может указывать на то, работает i -й индивид в период t или нет. Типичной является модель, в которой

$$y_{it}^* = x_{it}^T \beta + \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где x_{it}^T — вектор-строка значений объясняющих переменных для i -го субъекта в период t .

Предположим, что случайные ошибки ε_{it} независимы, одинаково распределены в обоих направлениях и имеют симметричное распределение с функцией распределения G и что объясняющие переменные строго экзогенны.

Если трактовать α_i как фиксированные неизвестные параметры, то это соответствует включению в модель N дамми-переменных. Логарифмическая функция правдоподобия тогда принимает вид:

$$\ln L(\beta, \alpha_1, \dots, \alpha_N) = \sum_{i,t} y_{it} \ln G(\alpha_i + x_{it}^T \beta) + \sum_{i,t} (1 - y_{it}) \ln(1 - G(\alpha_i + x_{it}^T \beta)).$$

И здесь возникает ситуация, когда оценка максимального правдоподобия для β и α_i , даже при выполнении стандартных предположений, состоятельна только если $T \rightarrow \infty$, а при конечном T и $N \rightarrow \infty$ она несостоятельна. Такая ситуация уже встречалась в рамках *OLS*-оценивания линейной модели с фиксированными эффектами. Только там при несостоятельности оценок для α_i оценка для β оставалась все же состоятельной, тогда как здесь несостоятельность оценок для α_i в общем случае переносится и на оценку для β . В работе (Chamberlain, 1980) приведено несколько примеров, иллюстрирующих такое положение, причем в одном из примеров (с $T = 2$) оценка максимального правдоподобия для β сходится по вероятности при $N \rightarrow \infty$ не к истинному значению β , а к значению 2β . В этой же работе был предложен альтернативный подход к оцениванию параметра β в модели бинарного выбора для панельных данных — **метод условного максимального правдоподобия** (*CML* — *conditional maximum likelihood*), который может быть применен к логит-модели (но не годится для пробит-модели).

Логит-модель с фиксированными эффектами

Логит-модели с фиксированными эффектами соответствует латентное уравнение

$$y_{it}^* = x_{it}^T \beta + \alpha_i + \varepsilon_{it},$$

где ε_{it} имеют логистическое распределение $G(z) = \Lambda(z) = \frac{e^z}{1 + e^z}$. При этом:

$$P\{y_{it} = 1 | x_{it}, \alpha_i, \beta\} = \frac{\exp\{x_{it}^T \beta + \alpha_i\}}{1 + \exp\{x_{it}^T \beta + \alpha_i\}},$$

$$P\{y_{it} = 0 | x_{it}, \alpha_i, \beta\} = 1 - \frac{\exp\{x_{it}^T \beta + \alpha_i\}}{1 + \exp\{x_{it}^T \beta + \alpha_i\}} = \frac{1}{1 + \exp\{x_{it}^T \beta + \alpha_i\}}.$$

Следуя работе (Chamberlain, 1980), рассмотрим логит-модель с $T = 2$. В этом случае сумма $(y_{i1} + y_{i2})$ есть просто общее количество периодов безработицы (суммарная продолжительность пребывания в состоянии безработицы) для i -го субъекта, и для каждого субъекта есть 4 возможные последовательности состояний (y_{i1}, y_{i2}) : (0, 0), (0, 1), (1, 0), (1, 1). Соответственно сумма $(y_{i1} + y_{i2})$ может принимать только 3 возможных значения: 0, 1, 2 (у i -го субъекта не было периодов безработицы, был один период безработицы, был безработным в течение обоих периодов).

Вычислим условные вероятности указанных 4 последовательностей при различных значениях суммы $(y_{i1} + y_{i2})$.

Если $y_{i1} + y_{i2} = 0$, то возможна только последовательность (0, 0), так что

$$P\{(0,0) | y_{i1} + y_{i2} = 0, \alpha_i, \beta\} = 1,$$

а условные вероятности 3 других последовательностей при $y_{i1} + y_{i2} = 0$ равны 0.

Если $y_{i1} + y_{i2} = 2$, то возможна только последовательность (1, 1), так что

$$P\{(1,1) | y_{i1} + y_{i2} = 2, \alpha_i, \beta\} = 1,$$

а условные вероятности 3 других последовательностей при $y_{i1} + y_{i2} = 2$ равны 0.

Если же $y_{i1} + y_{i2} = 1$, то возможны две последовательности: (0, 1) и (1, 0), и они имеют условные вероятности, отличные от 0 и 1, тогда как условные вероятности последовательностей (0, 0) и (1, 1) равны 0. При этом:

$$P\{(0,1) | y_{i1} + y_{i2} = 1, x_{it}, \alpha_i, \beta\} = \frac{P\{(0,1) | x_{it}, \alpha_i, \beta\}}{P\{y_{i1} + y_{i2} = 1 | x_{it}, \alpha_i, \beta\}}$$

$$= \frac{P\{y_{i1} = 0 | x_{it}, \alpha_i, \beta\} P\{y_{i2} = 1 | x_{it}, \alpha_i, \beta\}}{P\{y_{i1} = 0 | x_{it}, \alpha_i, \beta\} P\{y_{i2} = 1 | x_{it}, \alpha_i, \beta\} + P\{y_{i1} = 1 | x_{it}, \alpha_i, \beta\} P\{y_{i2} = 0 | x_{it}, \alpha_i, \beta\}} =$$

$$= \frac{1}{1 + \exp\{x_{i1}^T \beta + \alpha_i\}} \frac{\exp\{x_{i2}^T \beta + \alpha_i\}}{1 + \exp\{x_{i2}^T \beta + \alpha_i\}} =$$

$$= \frac{1}{1 + \exp\{x_{i1}^T \beta + \alpha_i\}} \frac{\exp\{x_{i2}^T \beta + \alpha_i\}}{1 + \exp\{x_{i2}^T \beta + \alpha_i\}} + \frac{\exp\{x_{i1}^T \beta + \alpha_i\}}{1 + \exp\{x_{i1}^T \beta + \alpha_i\}} \frac{1}{1 + \exp\{x_{i2}^T \beta + \alpha_i\}} =$$

$$= \frac{\exp\{x_{i2}^T \beta + \alpha_i\}}{\exp\{x_{i1}^T \beta + \alpha_i\} + \exp\{x_{i2}^T \beta + \alpha_i\}} = \frac{\exp\{x_{i2}^T \beta\}}{\exp\{x_{i1}^T \beta\} + \exp\{x_{i2}^T \beta\}},$$

так что индивидуальные эффекты вымываются, и

$$P\{(0, 1) | y_{i1} + y_{i2} = 1, x_{it}, \alpha_i, \beta\} = \frac{\exp\{(x_{i2} - x_{i1})^T \beta\}}{1 + \exp\{(x_{i2} - x_{i1})^T \beta\}}.$$

Соответственно

$$\begin{aligned} P\{(1, 0) | y_{i1} + y_{i2} = 1, x_{it}, \alpha_i, \beta\} &= 1 - P\{(0, 1) | y_{i1} + y_{i2} = 1, x_{it}, \alpha_i, \beta\} = \\ &= \frac{1}{1 + \exp\{(x_{i2} - x_{i1})^T \beta\}}. \end{aligned}$$

Таким образом, все условные вероятности

$$P\{(y_{i1}, y_{i2}) = (r, s) | y_{i1} + y_{i2} = l, x_{it}, \alpha_i, \beta\},$$

где $r, s = 0, 1$, $l = 0, 1, 2$, не зависят от индивидуальных эффектов α_i , так что максимизация по β и α_i условной (при условии заданных фиксированных значений x_{it} и $(y_{i1} + y_{i2})$) функции правдоподобия сводится к максимизации этой функции только по параметру β , что приводит к состоятельной оценке для β .

Полученные результаты означают, что при $T = 2$ можно оценивать логит-модель с фиксированными эффектами, используя стандартную логит-модель, в которой в качестве объясняющей переменной выступает $(x_{i2} - x_{i1})$, а в качестве наблюдаемой бинарной переменной — переменная, отражающая изменение значения переменной y_{it} при переходе от первого ко второму наблюдению (1 — при возрастании y_{it} , 0 — при убывании y_{it}). Соответственно функцию правдоподобия можно записать в виде:

$$\begin{aligned} L &= \prod_{i=1}^N P\{(0, 0) | \cdot\} P\{(0, 1) | \cdot\} P\{(1, 0) | \cdot\} P\{(1, 1) | \cdot\} = \\ &= \prod_{i=1}^N 1^{S_{00}} \left(\frac{\exp\{(x_{i2} - x_{i1})^T \beta\}}{1 + \exp\{(x_{i2} - x_{i1})^T \beta\}} \right)^{S_{01}} \left(\frac{1}{1 + \exp\{(x_{i2} - x_{i1})^T \beta\}} \right)^{S_{10}} 1^{S_{11}}, \end{aligned}$$

где $S_{rs} = 1$, если $y_{i1} = r, y_{i2} = s$, и $S_{rs} = 0$, если $y_{i1} \neq r$ и/или $y_{i2} \neq s$.

Заметим, что субъекты, для которых $y_{i1} = y_{i2} = 0$ или $y_{i1} = y_{i2} = 1$, не вносят дополнительный вклад в функцию правдоподобия и фактически игнорируются при оценивании. Для оценивания существенны только те субъекты, которые хотя бы однажды изменяют свой статус в отношении y_{it} .

Проверку гипотезы H_0 об отсутствии индивидуальных эффектов можно осуществить с помощью критерия типа критерия Хаусмана, основанного на разности между оценкой $\hat{\beta}_{CML}$ для β , полученной методом условного максимального правдоподобия, и обычной логит- ML -оценкой $\hat{\beta}_{ML}$, игнорирующей

индивидуальные эффекты (при вычислении последней константа исключается). Логит- MLE состоятельна и эффективна только при гипотезе H_0 и несостоятельна при альтернативе. Условная же MLE состоятельна и при гипотезе H_0 , и при альтернативе, но неэффективна, так как использует не все данные. Таким образом, положение здесь соответствует тому, при котором реализуется схема Хаусмана (см. разд. 1, тема 1.2). Статистика критерия

$$H = (\hat{\beta}_{CML} - \hat{\beta}_{ML})^T (C\hat{ov}(\hat{\beta}_{CML}) - C\hat{ov}(\hat{\beta}_{ML}))^{-1} (\hat{\beta}_{CML} - \hat{\beta}_{ML})$$

имеет при гипотезе H_0 асимптотическое распределение хи-квадрат с числом степеней свободы, равным размерности вектора β .

✓ **Замечание 4.4.1.** Существенным недостатком метода условного правдоподобия является то, что он предполагает (условную) независимость y_{i1}, \dots, y_{iT} при фиксированных x_{it}, α_i, β . Метод неприменим к моделям с запаздывающими значениями объясняющей переменной в правой части.

Пробит-модель со случайными эффектами

Если α_i статистически независимы от объясняющих переменных, то естественно рассмотреть модель со случайными эффектами. В этом случае можно использовать и логит-, и пробит-модели, проще применять пробит-модель. Здесь исходим из спецификации:

$$y_{it} = \begin{cases} 1, & \text{если } y_{it}^* > 0, \\ 0, & \text{если } y_{it}^* \leq 0, \end{cases}$$

$$y_{it}^* = x_{it}^T \beta + u_{it},$$

$$u_{it} = \alpha_i + \varepsilon_{it},$$

где α_i — случайные индивидуальные эффекты, $\alpha_i \sim i.i.d. N(0, \sigma_\alpha^2)$;

ε_{it} — случайные ошибки, $\varepsilon_{it} \sim i.i.d. N(0, \sigma_\varepsilon^2)$,

и предполагаем, что индивидуальные эффекты и ошибки статистически независимы. В этих условиях совместное распределение случайных величин u_{i1}, \dots, u_{iT} нормально, причем $E(u_{it}) = 0$, $D(u_{it}) = \sigma_\alpha^2 + \sigma_\varepsilon^2$, $Cov(u_{it}, u_{is}) = Cov(\alpha_i + \varepsilon_{it}, \alpha_i + \varepsilon_{is}) = \sigma_\alpha^2$ для $s \neq t$, так что значение коэффициента корреляции $\rho = \frac{Cov(u_{it}, u_{is})}{\sqrt{D(u_{it})}\sqrt{D(u_{is})}} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}$ между ошибками u_{it} и u_{is} внутри одной

группы (субъекта) одинаково для любых $s \neq t$. В стандартной спецификации $D(u_{it}) = 1$ и $\rho = \sigma_\alpha^2$. Это соответствует предположению о том, что $\sigma_\varepsilon^2 = 1 - \sigma_\alpha^2$.

Совместная вероятность получения набора y_{i1}, \dots, y_{iT} при заданных x_{i1}, \dots, x_{iT} определяется следующим образом:

$$\begin{aligned} P(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \beta) &= \int_{-\infty}^{\infty} P(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \alpha_i, \beta) p(\alpha_i) d\alpha_i = \\ &= \int_{-\infty}^{\infty} \left[\prod_{t=1}^T P(y_{it} | x_{it}, \alpha_i, \beta) \right] p(\alpha_i) d\alpha_i, \end{aligned}$$

где

$$P(y_{it} | x_{it}, \alpha_i, \beta) = \begin{cases} \Phi\left(\frac{x_{it}^T \beta + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}}\right), & \text{если } y_{it} = 1, \\ 1 - \Phi\left(\frac{x_{it}^T \beta + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}}\right), & \text{если } y_{it} = 0, \end{cases}$$

$$p(\alpha_i) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left\{-\frac{1}{2} \frac{\alpha_i^2}{\sigma_\alpha^2}\right\}.$$

Итоговый интеграл вычисляется численными методами. Соответствующие программы имеются в некоторых пакетах (например, в *Stata*).

При гипотезе H_0 об отсутствии индивидуальных эффектов $\rho = \sigma_\alpha^2 = 0$, так что проверка отсутствия индивидуальных эффектов сводится к проверке гипотезы $\rho = 0$.

ПРИМЕР 4.4.1

Рассмотрим результаты исследования панельных данных по Германии (*German Socio-Economic Panel*) за 5-летний период с 1985 по 1989 г., проводившихся с целью выяснения влияния состояния безработицы на степень удовлетворенности жизнью (*well-being*) (см.: (*Winkelmann L., Winkelmann R., 1995*)).

Исходные данные были представлены в порядковой шкале и индексировались от 0 до 10. Ввиду сложности построения порядковых логит- и пробит-моделей для панельных данных данные были сжаты до бинарных. Индивиды, у которых удовлетворенность жизнью индексировалась от 0 до 4, считались неудовлетворенными жизнью, тогда как индивиды с индексом 5 и выше считались более или менее удовлетворенными жизнью.

В качестве объясняющих переменных первоначально были привлечены следующие характеристики индивидов:

- *непрерывная переменная*:
LOGINCOME — логарифм (располагаемого) дохода (см. ниже);
- *дамми-переменные*:
UNEMP (безработный), *NOPARTIC* (не включен в рынок труда), *SELFEMP* (самозанятый), *PARTTIME* (частично занятый) — указывают текущий

статус на рынке труда (категория «полностью занятый» не снабжается дамми-переменной);

MALE — индивид мужского пола;

OKHEALTH — хорошее состояние здоровья;

AGE — возраст и *AGESQUARED* — квадрат возраста;

VOCATIONAL D. — наличие профессиональной степени;

UNIVERSITY D. — наличие университетской степени;

MARRIED — состояние в браке.

Для правильной интерпретации влияния безработицы на удовлетворенность жизнью надо учитывать целый ряд обстоятельств.

Например, безработица обычно ведет к сокращению дохода, что, в свою очередь, может снизить удовлетворенность. Однако если доход включается в число объясняющих переменных, то коэффициент при безработице фактически измеряет влияние безработицы *при прочих равных*, т.е. при сохранении дохода постоянным. Это могло быть в реальности, если бы страховое возмещение при безработице достигало 100%. При высоком (отрицательном) коэффициенте корреляции между безработицей и доходом оценки остаются несмещенными, но их точность может значительно уменьшаться.

Удовлетворенность индивида может снижаться вследствие безработицы также и потому, что доля его вклада в общий семейный бюджет уменьшается. Поскольку, в отличие от доходов домохозяйства, имеющиеся в панели данные в отношении индивидуального дохода ограничены прошлыми доходами, в уравнение включается не индивидуальный доход, а доход домохозяйства.

Включение в модель взаимодействий факторов позволяет выделять дифференциальные эффекты влияния безработицы в группах, имеющих разные атрибуты (например, в разных возрастных группах). В связи с этим в модель помимо собственно *UNEMP* включены также переменные:

$$UNEMP*AGE<29$$

$$UNEMP*AGE30-49.$$

Важным также является взаимодействие между безработицей и продолжительностью текущего периода пребывания в состоянии безработного. Влияние продолжительности безработицы на психическое состояние индивида достаточно хорошо документировано в психологической литературе. В соответствии с этим в уравнение включаются переменные:

UNEMPDUR — продолжительность периода безработицы, продолжающегося в настоящее время;

*UNEMPDURSQ** 10^{-2} — для учета возможной нелинейности.

К сожалению, в данных полностью отсутствовал региональный аспект, так что не удалось выяснить влияние локального уровня безработицы.

Следует отметить очень существенное взаимодействие безработицы и пола индивида. По этой причине обработка данных производилась также раздельно по мужчинам и женщинам.

Наличие протяженных наблюдений позволяет, в принципе, включать в модель элементы динамики. Однако выше уже говорилось о проблемах с оцениванием динамических моделей бинарного выбора, в частности, с включением в правую часть запаздывающих значений объясняемой переменной. В данном анализе предполагается устойчивость индивидуальных откликов во времени, и она учитывается введением в модель индивидуальных эффектов. Для проверки этого свойства в модель помимо переменной *UNEMPDUR* включается также переменная *PREVDUR*, указывающая на суммарную продолжительность периодов безработицы за 10 лет, предшествующих анализируемому периоду.

Кроме того, в модели в качестве объясняющей используется еще и переменная *CHANGEINC* — относительное изменение дохода домохозяйства по сравнению с предыдущим годом. Это связано с некоторой неясностью в отношении того, что именно влияет на удовлетворенность — собственно уровень доходов или его изменение.

Наконец, нельзя исключать возможного наличия и обратной причинной связи между безработицей и уровнем удовлетворенности, т.е. того, что внутренне менее удовлетворенные индивиды скорее и оказываются безработными. Если это так, то переменная *UNEMP* коррелирует со случайным эффектом в модели случайных эффектов и оценка максимального правдоподобия, не учитывающая этого, оказывается несостоятельной. Поскольку далее приводятся результаты статистического анализа и по модели с фиксированными эффектами, и по модели со случайными эффектами, можно оценить устойчивость результатов к этой допустимой эндогенности.

Основная задача состоит в проверке гипотезы о негативном влиянии безработицы на удовлетворенность при очистке от влияния прочих факторов, включая влияние снижения доходов.

В табл. 4.29 и 4.30 приведены результаты оценивания порядковой пробит-модели, не учитывающей панельную структуру данных (данные усреднены по времени), но включающей временные дамми. В этой порядковой пробит-модели предполагается, что уравнение полезности имеет вид:

$$S_{it}^* = x_{it}^T \beta + \varepsilon_{it},$$

и наблюдаемое значение индекса удовлетворенности определяется соотношениями:

$$S_{it} = \begin{cases} 0, & \text{если } S_{it}^* \leq 0, \\ 1, & \text{если } 0 < S_{it}^* \leq \gamma_1, \\ \vdots & \vdots \\ 10, & \text{если } S_{it}^* > \gamma_9. \end{cases}$$

В табл. 4.31 и 4.32 приведены результаты оценивания *FE* логит-модели и *RE* пробит-модели с бинарным откликом (1 — если уровень удовлетворенности равен 5 или выше, 0 — если уровень удовлетворенности ниже 5).

Таблица 4.29

Оценка порядковой пробит-модели (объединенная выборка)

Переменная	Коэффициент	Стандартная ошибка
<i>Constant</i>	1.661	(0.168)
<i>MALE</i>	−0.059	(0.015)
<i>UNEMP</i>	−0.265	(0.050)
<i>UNEMP*AGE<29</i>	−0.164	(0.078)
<i>UNEMP*AGE30-49</i>	−0.221	(0.065)
<i>UNEMPDUR</i>	−0.011	(0.007)
<i>UNEMPDURSQ*10^{−2}</i>	0.013	(0.020)
<i>PREVDUR</i>	−0.009	(0.001)
<i>NOPARTIC</i>	−0.042	(0.018)
<i>SELFEMP</i>	−0.107	(0.033)
<i>PARTTIME</i>	−0.060	(0.028)
<i>MARRIED</i>	0.140	(0.017)
<i>OKHEALTH</i>	0.428	(0.014)
<i>AGE</i>	−0.050	(0.004)
<i>AGESQ*10^{−2}</i>	0.062	(0.005)
<i>VOCATIONAL D.</i>	0.083	(0.013)
<i>UNIVERSITY D.</i>	0.106	(0.033)
<i>LOGINCOME</i>	0.208	(0.017)
<i>CHANGEINC</i>	0.018	(0.022)
1986	−0.027	(0.020)
1987	−0.127	(0.020)
1988	−0.191	(0.020)
1989	−0.219	(0.021)
<i>Observations</i>	24055	

Таблица 4.30

Оценка порядковой пробит-модели (раздельные выборки)

Переменная	Мужчины		Женщины	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
<i>Constant</i>	2.273	(0.246)	1.259	(0.234)
<i>UNEMP</i>	−0.402	(0.063)	−0.118	(0.089)
<i>UNEMP*AGE<29</i>	−0.321	(0.105)	−0.064	(0.126)
<i>UNEMP*AGE30-49</i>	−0.245	(0.085)	−0.175	(0.109)
<i>UNEMPDUR</i>	−0.003	(0.010)	−0.010	(0.018)
<i>UNEMPDURSQ*10^{−2}</i>	−0.013	(0.028)	0.042	(0.064)
<i>PREVDUR</i>	−0.010	(0.001)	−0.008	(0.001)

Окончание табл. 4.30

Переменная	Мужчины		Женщины	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
<i>NOPARTIC</i>	-0.273	(0.035)	0.056	(0.024)
<i>SELFEMP</i>	-0.087	(0.038)	-0.155	(0.069)
<i>PARTTIME</i>	-0.180	(0.129)	-0.001	(0.031)
<i>MARRIED</i>	0.133	(0.024)	0.110	(0.026)
<i>OKHEALTH</i>	0.409	(0.021)	0.426	(0.021)
<i>AGE</i>	-0.063	(0.006)	-0.049	(0.006)
<i>AGESQ</i> *10 ⁻²	0.078	(0.007)	0.062	(0.007)
<i>VOCATIONAL D.</i>	0.066	(0.019)	0.093	(0.019)
<i>UNIVERSITY D.</i>	0.086	(0.040)	0.174	(0.059)
<i>LOGINCOME</i>	0.165	(0.024)	0.253	(0.025)
<i>CHANGEINC</i>	0.014	(0.030)	0.015	(0.031)
1986	-0.018	(0.028)	-0.037	(0.029)
1987	-0.104	(0.028)	-0.152	(0.030)
1988	-0.175	(0.028)	-0.206	(0.030)
1989	-0.178	(0.029)	-0.258	(0.030)
<i>Observations</i>	12605		11450	

Таблица 4.31

Оценка RE пробит-модели и FE логит-модели (мужчины)

Переменная	<i>Random Effects</i>		<i>Fixed Effects</i>	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
<i>Constant</i>	1.849	(0.682)	—	—
<i>UNEMP</i>	-0.849	(0.145)	-1.257	(0.358)
<i>UNEMP*AGE</i> <29	-0.089	(0.225)	0.115	(0.518)
<i>UNEMP*AGE</i> 30-49	-0.013	(0.187)	0.368	(0.419)
<i>UNEMPDUR</i>	-0.026	(0.025)	-0.030	(0.047)
<i>UNEMPDURSQ</i> *10 ⁻²	0.027	(0.072)	0.109	(0.145)
<i>PREVDUR</i>	-0.011	(0.004)	—	—
<i>NOPARTIC</i>	-0.511	(0.084)	-0.287	(0.243)
<i>SELFEMP</i>	-0.185	(0.107)	0.643	(0.418)
<i>PARTTIME</i>	-0.274	(0.274)	-0.162	(0.554)
<i>MARRIED</i>	0.294	(0.068)	0.575	(0.254)
<i>OKHEALTH</i>	0.448	(0.052)	0.409	(0.128)
<i>AGE</i>	-0.073	(0.019)	0.009	(0.114)
<i>AGESQ</i> *10 ⁻²	0.086	(0.023)	-0.112	(0.131)

Окончание табл. 4.31

Переменная	Random Effects		Fixed Effects	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
VOCATIONAL D.	0.098	(0.060)	—	—
UNIVERSITY D.	0.057	(0.115)	—	—
LOGINCOME	0.135	(0.068)	0.676	(0.245)
CHANGEINC	0.158	(0.072)	0.086	(0.174)
ρ	0.382	(0.040)	—	—
No. of individuals	2521		556	

Таблица 4.32

Оценка RE пробит-модели и FE логит-модели (женщины)

Переменная	Random Effects		Fixed Effects	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
Constant	−0.160	(0.639)	—	—
UNEMP	0.003	(0.209)	−0.192	(0.484)
UNEMP*AGE<29	−0.255	(0.299)	−0.069	(0.641)
UNEMP*AGE30-49	−0.550	(0.240)	−0.821	(0.548)
UNEMPDUR	−0.017	(0.039)	0.014	(0.074)
UNEMPDURS Q *10 ^{−2}	0.047	(0.144)	−0.022	(0.270)
PREVDUR	−0.009	(0.003)	—	—
NOPARTIC	0.007	(0.069)	−0.180	(0.225)
SELFEMP	−0.057	(0.191)	−0.268	(0.497)
PARTTIME	0.022	(0.087)	−0.206	(0.241)
MARRIED	0.220	(0.074)	0.561	(0.261)
OKHEALTH	0.529	(0.053)	0.620	(0.126)
AGE	−0.035	(0.019)	−0.179	(0.108)
AGES Q *10 ^{−2}	0.042	(0.023)	0.066	(0.125)
VOCATIONAL D.	0.149	(0.063)	—	—
UNIVERSITY D.	0.280	(0.185)	—	—
LOGINCOME	0.278	(0.068)	0.356	(0.233)
CHANGEINC	0.159	(0.073)	0.265	(0.160)
ρ	0.411	(0.039)	—	—
No. of individuals	2290		538	

Результаты, приведенные в таблицах, подтверждают сделанные ранее выводы. В частности, влияние безработицы выглядит отрицательным. Однако следует помнить, что в нелинейных моделях оцененные коэффициенты

не представляют предельные эффекты. Более того, они несравнимы для разных спецификаций модели.

Тем не менее в рамках оценивания порядковой пробит-модели без разделения по полу можно отметить следующие моменты:

- влияние пола оказывается весьма значимым (оценка коэффициента равна -0.059 при стандартной ошибке 0.015), что говорит в пользу раздельного анализа данных по мужчинам и женщинам;
- подтверждаются результаты более ранних исследований, указывающие на U-образную форму степени негативного влияния безработицы на удовлетворенность: оцененный коэффициент выше по абсолютной величине в наиболее активных возрастных группах;
- уровень доходов имеет существенный положительный эффект при измерении его в уровнях, но не в приращениях.

Положительное влияние имеют наличие профессионального или университетского диплома, а также состояние в браке.

При раздельном оценивании по мужчинам и женщинам обнаруживаются большие дифференциальные эффекты, так что публикация результатов только по объединенной выборке, как это делают некоторые авторы, может вводить в заблуждение. Наиболее существенное различие заключается в том, что на мужчин состояние безработного действует в значительно большей степени. Так, для мужчин среднего возраста эффект безработицы равен:

$$-0.402 - 0.245 = -0.647,$$

тогда как для женщин он равен только:

$$-0.118 - 0.175 = -0.293.$$

Это соответствует росту неудовлетворенности на 12% для мужчин и только на 4% для женщин.

Отметим также, что пребывание вне рынка рабочей силы негативно влияет на мужчин, но не на женщин.

При переходе к анализу собственно панельными методами следует заметить, что в *FE*-модели вымечаются все регрессоры, значения которых не изменяются во времени (*Constant*, *PREVDUR*, *VOCATIONAL D.*, *UNIVERSITY D.*). Кроме того, количество индивидов, по которым производится оценивание, сокращается, поскольку в этом случае в выборку надо включать только тех, у кого значение бинарного отклика изменялось на 5-летнем периоде исследования. Переход к *FE*- и *RE*-моделям не очень сильно повлиял на выводы, сделанные ранее на основе порядковой модели. ■

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какие проблемы возникают при оценивании моделей бинарного вывода для панельных данных?
2. Какие модели бинарного вывода со случайными эффектами можно оценить методом максимального правдоподобия?

Тема 4.5

ТОБИТ-МОДЕЛИ ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ

Тобит-модель со случайными эффектами
для панельных данных

В пробит-модели со случайными эффектами исходили из спецификации:

$$y_{it} = \begin{cases} 1, & \text{если } y_{it}^* > 0, \\ 0, & \text{если } y_{it}^* \leq 0. \end{cases}$$

В тобит-модели со случайными эффектами исходим из спецификации:

$$y_{it} = \begin{cases} y_{it}^*, & \text{если } y_{it}^* > 0, \\ 0, & \text{если } y_{it}^* \leq 0. \end{cases}$$

Сохранив остальные предположения теми же, что и в пробит-модели со случайными эффектами, получим следующее выражение для условной совместной плотности распределения y_{i1}, \dots, y_{iT} при заданных x_{i1}, \dots, x_{iT} :

$$p(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \beta) = \int_{-\infty}^{\infty} \left[\prod_{t=1}^T p(y_{it} | x_{it}, \alpha_i, \beta) \right] p(\alpha_i) d\alpha_i,$$

где

$$p(\alpha_i) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left\{-\frac{1}{2} \frac{\alpha_i^2}{\sigma_\alpha^2}\right\},$$

$$P(y_{it} | x_{it}, \alpha_i, \beta) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left\{-\frac{1}{2} \frac{(y_{it} - x_{it}^T \beta - \alpha_i)^2}{\sigma_\varepsilon^2}\right\}, & \text{если } y_{it} > 0, \\ 1 - \Phi\left(\frac{x_{it}^T \beta + \alpha_i}{\sigma_\varepsilon}\right), & \text{если } y_{it} = 0. \end{cases}$$

ПРИМЕР 4.5.1

Аукционы репо (кредит под залог ценных бумаг) являются основным инструментом управления рынком со стороны Европейского центрального банка (ЕЦБ). На этих аукционах банки подают заявки — биды (*bids*) на получение кредита на определенный срок (*maturity*), в которых указываются величина кредита и проценты по кредиту, по которым покупатель кредита предполагает получить соответствующий кредит. Центральный банк на каждом аукционе выделяет для предоставления кредитов некоторую сумму и последовательно удовлетворяет полученные заявки банков, начиная с тех, в кото-

рых указан максимальный среди поданных заявок процент по кредиту. Отличительной чертой аукционов репо, проводимых ЕЦБ, является предварительное объявление Центральным банком минимального процента по бидам — нижней границы процентной ставки по кредиту, по которой еще принимаются заявки. Эта ставка весьма определенно указывает на направленность политики ЕЦБ и обычно устанавливает нижнюю границу для краткосрочных процентных ставок в зоне евро. Однако когда банки ожидают снижения процентных ставок, текущая минимальная ставка представляется им завышенной, и банки начинают воздерживаться от участия в аукционах. Такое поведение банков серьезно затрудняет управление ликвидностью со стороны ЕЦБ и ведет к повышению волатильности процентных ставок. В связи с этим представляет интерес выяснение роли минимальной ставки. Для этой цели можно использовать данные об индивидуальных заявках банков на аукционах без установления минимальной ставки. Такие аукционы репо проводил Бундесбанк Германии. Важно, что в остальном аукционы репо, проводимые Бундесбанком и ЕЦБ, следовали одним и тем же правилам. Это дает уникальную возможность исследовать действительную роль минимальной ставки.

Используя панельные данные, можно оценить модели для вероятности неучастия банка в аукционе и для размера индивидуальной заявки банка. Последняя переменная цензурирована слева, поскольку она наблюдается только в случае участия банка в аукционе, и это обстоятельство можно учесть, используя панельную тобит-модель. Это сделано в работе (*Linzert, Nautz, Breitung, 2003*), где для анализа были взяты данные о заявках 275 банков, поданных в Центробанке земли Гессен. Эта выборка достаточно репрезентативна, поскольку в этой земле находится Франкфурт — финансовый центр Германии. Аукционы репо проводились еженедельно. Но фактически для анализа подходит только один достаточно долгий период, когда Бундесбанк не изменял формат аукциона. Это период с апреля по ноябрь 1995 г. Соответственно в этом периоде были проведены 33 аукциона со стандартным сроком кредита около двух недель.

Во многих отношениях поведение банков на этих аукционах сходно с поведением банков на аукционах, проводимых ЕЦБ. Во-первых, крупные покупатели участвуют в аукционах чаще, но их заявки относительно малы по сравнению с общим объемом заявок. Во-вторых, банки обычно подают на одном аукционе не более трех заявок (разные заявки ориентируются на разные проценты). В-третьих, довольно много покупателей участвует в аукционах редко. Из 275 банков только 175 хотя бы раз участвовали в аукционах в рассматриваемом периоде. Подобная картина имеет место и в аукционах репо ЕЦБ.

Заметим, что в данном периоде наблюдалось падение процентных ставок. В частности, в августе 1995 г. Бундесбанк понизил ломбардную процентную ставку на 50 базисных пунктов. Таким образом, есть возможность посмотреть, как это повлияло на поведение банков на аукционах.

На спрос банков на аукционах репо должна влиять стоимость альтернативных возможностей рефинансирования. В связи с этим в число объясняющих переменных включается переменная *spread*, определяемая как разность между ожидаемой ставкой отсечения на аукционе и ставкой овернайт. Ожидаемая ставка отсечения была рассчитана на основании уравнения коррекции ошибок, соответствующего выявленной коинтеграционной связи между ставкой отсечения и ставкой овернайт на межбанковском рынке.

Следующая объясняющая переменная — *term spread* (временной спред) определяется как разность между одномесечной процентной ставкой и овернайт-ставкой. Отрицательное значение временного спреда означает ожидание снижения процентных ставок. Для аукционов репо ЕЦБ ожидаемые изменения ключевых процентных ставок ЕЦБ оказывают весьма сильное влияние на поведение банков. В частности, когда банки ожидают снижения процентных ставок, неучастие их в аукционах затрудняет управление ликвидностью со стороны центрального банка. Ориентируясь на движение временного спреда, банки также должны были ожидать снижения Бундесбанком ломбардной ставки в августе 1995 г. Для учета возможного «забастовочного» поведения банков в число объясняющих переменных включается дамми-переменная *underbidding*, принимающая значение 1 для последнего аукциона, предшествовавшего понижению Бундесбанком ломбардной ставки.

Неопределенность, ощущаемую в день аукциона, отражает переменная *volatility*. Она оценивается на основе модели¹ *EGARCH*(1, 1) для ежедневных наблюдений значений ставки овернайт, в которой условная дисперсия ε_t относительно прошлого удовлетворяет соотношению:

$$\ln \sigma_t^2 = \omega + \beta \ln \sigma_{t-1}^2 + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}}.$$

Влияние такой неопределенности может быть двояким. С одной стороны, при возрастании неопределенности банки становятся более осторожными. С другой — если дело касается возможного проигрыша в аукционе, то чем больше неопределенность, тем большей может оказаться заявка банка и под более высокие проценты.

Переменная *reserve fulfillment* измеряет потребность банковского сектора в ликвидности. Она определялась как отношение резервных остатков всех банков Германии перед проведением аукциона к агрегированному требованию минимального резерва. Если резервные остатки малы, то банки имеют более сильные побуждения к участию в аукционе. Поскольку Бундесбанк допускает усреднение по (обычно месячному) периоду между проверками выполнения требований в отношении резервов, этот эффект может быть особенно выражен на последнем аукционе данного периода, и для учета этого

¹ См., например, (*Verbeek*, 2000, ch. 8: sect. 8.10.1 (русский перевод: *Вербик М.* Путеводитель по современной эконометрике. М.: Научная книга, 2008)).

вводится дамми-переменная *end of period*, равная 1, если аукцион является последним на указанном периоде.

Еще одна переменная — *maturing allotment* учитывает тот факт, что банки часто используют двухнедельный репо на возобновляемой основе. Эта переменная определяется как логарифм размера репо, полученного банком двумя неделями ранее.

Дамми-переменные *large*, *medium* и *small* характеризуют размер банка, отражающийся в среднем размере заявки. Эти переменные комбинируются со всеми другими объясняющими переменными для изучения поведения банков в зависимости от их размеров.

Эмпирические результаты

Решение банка об участии в аукционе. Здесь объясняемая переменная y_{it} равна 1, если i -й банк участвует в t -м аукционе, проводимом в течение рассматриваемого периода, и равна 0 — в противном случае. В табл. 4.33 приведены результаты оценивания панельной логит-модели

$$P\{y_{it} = 1 | x_{it}\} = \Lambda(x_{it}^T \beta),$$

в которую включены все указанные выше объясняющие переменные.

В табл. 4.33 приведены результаты и для *RE*-, и для *FE*-модели. Однако *RE*-модель предпочтительнее, поскольку критерий типа критерия Хаусмана не отвергает гипотезу о некоррелированности индивидуальных эффектов с объясняющими переменными.

Напомним (см. разд. 4, тема 4.1), что, поскольку логит-модель является нелинейной моделью, оцененные коэффициенты имеют интерпретацию, отличающуюся от интерпретации коэффициентов в линейной модели. В связи с этим во второй графе табл. 4.33 приведены значения предельного эффекта для переменных со статистически значимыми оценками коэффициентов, вычисленные при средних значениях объясняющих переменных на рассмотренном периоде. Так, значение 0.060 предельного эффекта для дамми-переменной *end of period* означает, что если аукцион проводится в конце периода между проверками выполнения требований в отношении резервов, то (при неизменных значениях остальных объясняющих переменных) шансы на то, что банк примет участие в аукционе, против того, что банк не примет участия в аукционе, возрастают в среднем приблизительно на 6%.

Таблица 4.33 показывает, что ожидания в отношении процентных ставок имеют статистически значимое влияние на решение банка об участии в аукционе. В полном согласии с тем, что наблюдается на аукционах ЕЦБ, вероятность участия банка в аукционах репо Бундесбанка уменьшается, когда отрицательное значение переменной *term spread* указывает на ожидание снижения процентных ставок.

В то же время статистическая незначимость оцененного коэффициента при переменной *underbidding* подчеркивает, что ожидаемый уровень отсечки

Таблица 4.33

Оценка панельной логит-модели

Переменная	<i>Random Effects</i>		<i>Conditional FE</i>
	Оценка коэффициента	Предельный эффект	Оценка коэффициента
<i>Term spread</i>	8.19 (5.16)	0.879	8.70 (5.48)
<i>Underbidding</i>	0.36 (1.20)	—	0.30 (0.97)
<i>Spread</i>	−11.57 (−6.82)	−1.242	−12.14 (−7.14)
<i>Volatility</i>	−0.19 (−2.08)	−0.020	−0.20 (−2.18)
<i>Reserve fulfillment</i>	0.68 (0.55)	—	1.11 (0.90)
<i>Maturing allotment</i>	0.13 (22.24)	0.014	0.11 (19.15)
<i>End of period</i>	0.74 (6.28)	0.060	0.76 (6.41)
<i>Size dummies</i>			
<i>large</i>	2.51 (1.57)		
<i>medium</i>	−8.98 (−3.43)		
<i>small</i>	−12.70 (−4.82)		
<i>Pseudo-R</i> ²	0.1142		0.1048
<i>No. of observations</i>	8525		4495
<i>No. of groups</i>	275		145
Примечание. В скобках приведены значения <i>t</i> -статистик. Для статистически значимых оценок приведены значения предельных эффектов. При оценивании <i>FE</i> -модели методом условного максимального правдоподобия учитываются только те банки, которые принимают участие в аукционах хотя бы дважды. Псевдо- <i>R</i> ² вычислен по формуле, приведенной в разд. 4 (тема 4.1).			

на аукционе Бундесбанка не имеет определяющего влияния на решения банков об участии в аукционе. Оцененный коэффициент при переменной *spread* имеет высокую статистическую значимость и отрицательное значение. Это означает, что, если ожидаемая ставка репо центрального банка выше ставки на рынке денег, количество банков, принимающих решение об участии в аукционе репо, сокращается.

Оцененный коэффициент при переменной *maturing allotment* имеет высокую статистическую значимость и положительное значение, подтверждая то, что банки используют репо на возобновляемой основе.

Высокая статистическая значимость и положительность оцененного коэффициента при переменной *end of period* указывают на увеличение вероятности участия банка в аукционе, проводимом на последней неделе периода между проверками выполнения требований в отношении резервов.

Коэффициенты при дамми-переменных, связанных с размерами банков, отражают тот очевидный факт, что большие банки участвуют в аукционах

чаще, чем малые. Большие банки используют аукционы репо не только для своей потребности в ликвидности, но и для перепродажи и активной торговли резервами на вторичном рынке.

Влияние переменной *reserve fulfillment* не выявлено: оцененный коэффициент при этой переменной статистически незначим. Что касается переменной *volatility*, то для нее значение *t*-статистики лишь ненамного превосходит 5%-й критический уровень.

Для выяснения вопроса о том, зависит ли отклик банка на изменение того или иного фактора от размера банка, в правую часть уравнения добавляются взаимодействия факторов с размером банка, т.е. переменные, которые являются произведениями объясняющих переменных на дамми, соответствующие возможным размерам банка. Результаты оценивания расширенной *RE*-модели приведены в первой графе табл. 4.34. Во второй графе этой таблицы — *P*-значения статистик критерия для проверки гипотезы об отсутствии эффекта размера для отдельных объясняющих переменных.

Отметим значимое влияние размера банка на отклик банка в отношении ожидаемых процентных ставок (*term spread*) и ожидаемой альтернативной стоимости (*spread*). Для обеих переменных влияние оказывается наиболее слабым для банков среднего размера. Коэффициенты при переменной *maturing allotment* показывают, что сезонный характер участия в аукционах более выражен для банков малого и среднего размера.

Что касается остальных объясняющих переменных, то здесь не обнаруживается значимого влияния размера банка.

В последней графе табл. 4.34 приведены *P*-значения критериев совместной значимости для каждой группы переменных. Отметим, что расширенная модель подтверждает сомнения в значимости волатильности для принятия банком решения об участии в аукционе.

Размер заявки отдельного банка. Исследуем теперь, каким образом перечисленные выше факторы влияют на размер заявки отдельного банка. Поскольку эту переменную можно наблюдать только в том случае, если банк принимает решение об участии в аукционе, она цензурирована слева (при неучастии банка в аукционе размер его заявки считаем равным нулю), и игнорирование этого обстоятельства может выразиться в смещении получаемых оценок. Соответственно исследование здесь проводится с привлечением панельной тобит-модели.

Предельный эффект *k*-й объясняющей переменной в тобит-модели вычисляется по формуле:

$$\frac{\partial E\{y_{it} | x_{it}\}}{\partial x_{k,it}} = \Phi\left(\frac{x_{it}^T \beta}{\sigma}\right) \beta_k.$$

В табл. 4.35 приведены оцененные коэффициенты и вычисленные значения предельных эффектов. Для сравнения в последней графе таблицы приведены результаты *GLS*-оценивания *RE*-модели, которая пренебрегает информацией, содержащейся в нулевых заявках.

Таблица 4.34

Оценивание логит-модели и размер банка

Переменная	Оценка коэффициента	H_0 : нет эффекта размера в группе (P -значение)	H_0 : равны нулю все коэффициенты в группе (P -значение)
<i>Term spread</i>			
<i>large banks</i>	20.36 (0.98)	0.0002	0.0000
<i>medium banks</i>	0.74 (0.31)		
<i>small banks</i>	14.02 (6.45)		
<i>Underbidding</i>			
<i>large banks</i>	0.19 (0.08)	0.7285	0.5392
<i>medium banks</i>	0.12 (0.26)		
<i>small banks</i>	0.60 (1.45)		
<i>Spread</i>			
<i>large banks</i>	−20.71 (−0.92)	0.0018	0.0000
<i>medium banks</i>	−4.65 (−1.84)		
<i>small banks</i>	−16.81 (−7.17)		
<i>Volatility</i>			
<i>large banks</i>	−1.24 (−1.06)	0.5788	0.1260
<i>medium banks</i>	−0.13 (−0.96)		
<i>small banks</i>	−0.24 (−1.92)		
<i>Reserve fulfillment</i>			
<i>large banks</i>	1.60 (0.12)	0.3998	0.5626
<i>medium banks</i>	−1.23 (−0.67)		
<i>small banks</i>	2.17 (1.26)		
<i>Maturing allotment</i>			
<i>large banks</i>	0.02 (0.40)	0.0187	0.0000
<i>medium banks</i>	0.12 (13.67)		
<i>small banks</i>	0.15 (17.27)		
<i>End of period</i>			
<i>large banks</i>	1.07 (0.71)	0.8521	0.0000
<i>medium banks</i>	0.66 (3.59)		
<i>small banks</i>	0.78 (5.00)		
<i>Size dummies</i>			
<i>large banks</i>	−6.23 (−0.40)	0.1123	0.0000
<i>medium banks</i>	−1.75 (−0.79)		
<i>small banks</i>	−8.12 (−3.89)		
<i>Pseudo-R²</i>	0.1094		
<i>No. of observations</i>	8525		
<i>No. of groups</i>	275		
<i>Примечание.</i> В скобках приведены значения t -статистик.			

Таблица 4.35

Оценка панельной тобит-модели и GLS-оценивание RE-модели

Переменная	Tobit		Random Effects GLS
	Оценка коэффициента	Предельный эффект	Оценка коэффициента
<i>Term spread</i>	11.59 (4.23)	1.00	1.43 (4.27)
<i>Underbidding</i>	0.32 (0.61)	—	0.05 (1.97)
<i>Spread</i>	−16.55 (−7.82)	−1.43	−1.94 (−5.40)
<i>Volatility</i>	−0.24 (−1.50)	—	−0.01 (0.38)
<i>Reserve fulfillment</i>	1.32 (0.61)	—	−0.67 (2.39)
<i>Maturing allotment</i>	0.23 (20.41)	0.02	0.01 (7.39)
<i>End of period</i>	0.93 (4.61)	0.08	−0.08 (−1.13)
<i>Size dummies</i>			
<i>large</i>	3.95 (1.50)		—
<i>medium</i>	−8.98 (−3.43)		−3.46 (−5.10)
<i>small</i>	−12.70 (−4.82)		−4.80 (−7.14)
<i>Constant</i>	—		21.01 (28.61)
<i>Pseudo-R²</i>	0.161		0.062
<i>No. of observations</i>	8525		2625
<i>No. of groups</i>	275		275
Примечание. В скобках приведены значения <i>t</i> -статистик. Для статистически значимых оценок приведены значения предельных эффектов. При <i>RE GLS</i> -оценивании все 5900 цензурированных слева наблюдений отбрасывались.			

Можно отметить определенное сходство оценок предельных эффектов в тобит-модели с оценками коэффициентов в указанной RE-модели.

Исследование не обнаруживает значимого влияния на размер заявок переменных *volatility* и *underbidding*, но обнаруживает значимое влияние переменных *term spread*, *spread*, *maturing allotment* и дамми-переменной *end of period*. Незначимость оцененного коэффициента при переменной *underbidding* показывает, что даже за неделю до ожидаемого понижения ставок Бундесбанк не имел трудностей с поставкой на аукцион репо соответствующего объема резервов. Таким образом, в отличие от аукционов ЕЦБ, ожидаемые процентные ставки не препятствуют Бундесбанку в управлении денежным рынком.

В табл. 4.36 приведены результаты оценивания расширенной тобит-модели, которая включает взаимодействия объясняющих переменных и дамми-переменных, относящихся к размеру банка.

Таблица 4.36

Оценивание расширенной тобит-модели и размер банка

Переменная	Оценка коэффициента	H_0 : нет эффекта размера в группе (P -значение)	H_0 : равны нулю все коэффициенты в группе (P -значение)
<i>Term spread</i>			
<i>large banks</i>	5.19 (0.74)	0.0004	0.0000
<i>medium banks</i>	1.82 (0.89)		
<i>small banks</i>	12.91 (6.63)		
<i>Underbidding</i>			
<i>large banks</i>	−0.70 (−0.55)	0.5847	0.7014
<i>medium banks</i>	−0.01 (−0.04)		
<i>small banks</i>	0.42 (1.06)		
<i>Spread</i>			
<i>large banks</i>	−0.64 (−0.75)	0.0076	0.0000
<i>medium banks</i>	−6.15 (−2.82)		
<i>small banks</i>	−15.30 (−7.37)		
<i>Volatility</i>			
<i>large banks</i>	−0.23 (−0.57)	0.2357	0.0428
<i>medium banks</i>	−0.36 (−0.30)		
<i>small banks</i>	−0.31 (−2.78)		
<i>Reserve fulfillment</i>			
<i>large banks</i>	0.09 (0.02)	0.5056	0.6187
<i>medium banks</i>	−0.57 (−0.36)		
<i>small banks</i>	2.02 (1.29)		
<i>Maturing allotment</i>			
<i>large banks</i>	0.03 (1.48)	0.0000	0.0000
<i>medium banks</i>	0.13 (17.90)		
<i>small banks</i>	0.20 (23.57)		
<i>End of period</i>			
<i>large banks</i>	0.09 (0.18)	0.5445	0.0000
<i>medium banks</i>	0.59 (3.91)		
<i>small banks</i>	0.67 (4.77)		
<i>Size dummies</i>			
<i>large banks</i>	4.78 (0.74)	0.0103	0.0001
<i>medium banks</i>	−1.99 (−1.04)		
<i>small banks</i>	−8.34 (−4.40)		
<i>Pseudo-R</i> ²	0.228		
<i>No. of observations</i>	8525		
<i>No. of groups</i>	275		
Примечание. В скобках приведены значения <i>t</i> -статистик. При оценивании учитывались все 5900 цензурированных слева наблюдений.			

Полученные оценки явно указывают на эффект размера банка в отношении влияния ожидаемых процентных ставок (*term spread*), ожидаемой альтернативной стоимости (*spread*) и переменной *maturing allotment*. ■

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Чем отличается тобит-модель со случайными эффектами для панельных данных от пробит-модели со случайными эффектами?
2. Как оценивается тобит-модель со случайными эффектами для панельных данных?

Задания для семинарских занятий, работы в компьютерном классе и для самостоятельной работы

Методические указания к заданиям, предполагающим использование пакетов компьютерных программ статистического анализа данных, ориентированы на использование пакетов *EViews* 6 и *Stata* 10.

К разделу 1

Задание 1. *Диаграммы рассеяния в моделях одновременных уравнений*

1. Используя объект *Model* пакета *EViews*, постройте реализацию модели:

$$DGP: Q_t = 20 - 0.5P_t + u_t, \quad (\text{спрос})$$

$$Q_t = 10 + 0.5P_t + v_t, \quad (\text{предложение})$$

$$u_t, v_t \sim i.i.d. N(0, 1), \quad t = 1, 2, \dots, 100.$$

- Постройте диаграмму рассеяния для переменных Q_t, P_t .
- Вокруг какой точки происходит рассеяние облака точек?
- Выявляются ли на диаграмме направления прямых спроса и предложения?
- Можно ли в данной ситуации восстановить коэффициенты структурной формы по коэффициентам приведенной формы?

2. Используя объект *Model*, постройте реализацию модели:

$$DGP: Q_t = 20 - 0.5P_t + 0.1Y_t + u_t,$$

$$Q_t = 10 + 0.5P_t + v_t,$$

$$u_t, v_t \sim i.i.d. N(0, 1), \quad t = 1, 2, \dots, 100,$$

$$Y_t \sim i.i.d., \text{ равномерно распределены в интервале } (0, 1).$$

- Постройте диаграмму рассеяния для переменных Q_t, P_t .
- Вокруг какой точки происходит рассеяние облака точек?
- Выявляются ли на диаграмме направления прямых спроса и предложения?
- Можно ли в данной ситуации восстановить коэффициенты структурной формы по коэффициентам приведенной формы?

3. Используя объект *Model*, постройте реализацию модели:

$$\begin{cases} Q_t = a_0 + a_1 P_t + a_2 Y_t + u_t, \\ Q_t = b_0 + b_1 P_t + b_2 R_t + v_t \end{cases}$$

с теми же значениями коэффициентов a_0, a_1, a_2, b_0, b_1 и теми же последовательностями Y_t, u_t и v_t , что и в предыдущем пункте, и с различными значениями коэффициента b_2 : 40, 10, 6, 2.

- Постройте диаграммы рассеяния для переменных Q_t, P_t по смоделированным данным.
- Выявляются ли визуально направления прямых спроса и предложения в этих 4 ситуациях?

Задание 2. Проверка выполнения условий идентифицируемости уравнений структурной формы системы

1. Проведите самостоятельно проверку выполнения условий идентифицируемости уравнений структурной формы систем, рассмотренных в теме 1.1.
2. Проверьте выполнение рангового условия идентифицируемости уравнений структурной формы этих систем, используя соответствующие таблицы.

Задание 3. Проверка выполнения условий идентифицируемости уравнений структурной формы системы

1. Проанализируйте модель одновременных уравнений:

$$y_{t1} = ay_{t2} + bx_t + \varepsilon_{t1},$$

$$y_{t2} = cy_{t1} + dx_t + \varepsilon_{t2},$$

в которой $a, b, c, d \neq 0$,

$$\begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \sim i.i.d. N(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

Переменная x_t рассматривается как экзогенная для этой системы.

- Специфицируйте приведенную форму модели.
- Все ли коэффициенты структурных уравнений идентифицируемы?
- Пусть экономическая теория предполагает наличие следующих соотношений:

$$a + c = 1, \quad b - d = 0.$$

Будет ли структурная форма идентифицируемой при выполнении этих ограничений? Покажите, как при выполнении этих ограничений определяются коэффициенты структурных уравнений по коэффициентам приведенной формы.

2. Рассмотрите модель:

$$\begin{cases} P_t = aS_t + bX_t + \varepsilon_{t1}, \\ S_t = cS_{t-1} + \varepsilon_{t2}, \end{cases}$$

$$\begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix} \sim i.i.d. N(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

где P_t — объем производства некоторого товара;

S_t — объем продаж этого товара;

X_t — некоторая переменная, важная для принятия решений об объемах производства, экзогенная по отношению к рассматриваемой системе одновременных уравнений.

- Как характеризуется каждое из двух уравнений системы (недоидентифицировано, точно идентифицируемо, сверхидентифицируемо)?
- Как можно оценить коэффициенты структурной формы и проверить гипотезу $H_0: b = 0$?
- Влияет ли решение, принятое на основании этой проверки, на идентифицируемость модели? Как характеризуется каждое из двух уравнений новой системы (недоидентифицировано, точно идентифицируемо, сверхидентифицируемо)?
- Как можно проверить гипотезу $H_0: c = 0$ в исходной модели? Влияет ли решение, принятое на основании этой проверки, на идентифицируемость модели?

Предположим, что переменная X_t связана с переменными P_t и S_t тождеством $X_t - X_{t-1} = P_t - S_t$, так что переменную X_t можно рассматривать как инвентарные запасы товара, а $X_t - X_{t-1}$ — как изменение запасов по сравнению с предыдущим периодом.

- Влияет ли эта информация на выводы относительно идентифицируемости модели?
- Влияет ли эта информация на выбор процедуры оценивания?

Задание 4. Примеры, в которых порядковое условие выполнено, а ранговое условие не выполнено

Рассмотрите пример, приведенный в Замечании 1.1.4, постройте другие подходящие примеры.

Задание 5. Рекурсивные системы, оценивание рекурсивных систем

1. Дана модель одновременных уравнений:

$$\begin{cases} y_{t1} + a_{12}y_{t2} + a_{13}y_{t3} = a_{14}x_{t1} + a_{15}x_{t2} + \varepsilon_{t1}, \\ y_{t2} + a_{23}y_{t3} = a_{24}x_{t1} + a_{25}x_{t2} + \varepsilon_{t2}, \\ y_{t3} = a_{34}x_{t1} + a_{35}x_{t2} + \varepsilon_{t3}, \end{cases}$$

в которой

$$\begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \\ \varepsilon_{t3} \end{pmatrix} \sim i.i.d. N(0, \Sigma), \quad \Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix},$$

значения σ_1^2 , σ_2^2 , σ_3^2 неизвестны. Переменные x_{t1} и x_{t2} рассматриваются как экзогенные для этой системы.

- Как взаимодействуют между собой переменные y_{t1} , y_{t2} , y_{t3} ?
- Все ли коэффициенты системы идентифицируемы?

Пусть теперь $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}$, т.е. на (неизвестные) элементы ковариационной матрицы вектора ошибок не накладывается дополнительных ограничений.

- Как изменяется положение с идентифицируемостью коэффициентов в этой ситуации по сравнению с предыдущей?
- Выполнено ли ранговое условие для третьего уравнения?
- Какие дополнительные ограничения следует наложить на коэффициенты системы, чтобы второе уравнение стало идентифицируемым?

- Используя объект *Model* пакета *EViews*, постройте реализацию модели:

$$\begin{aligned} DGP: \quad Q_t &= 125 - 0.5P_t + u_t, \\ P_t &= 0.5Q_{t-1} + w_t, \\ u_t, w_t &\sim i.i.d. N(0, 1), \quad t = 1, 2, \dots, 100. \end{aligned}$$

В качестве стартового значения для Q используйте стационарное значение этой переменной.

- Постройте диаграмму рассеяния для переменных Q_t , P_t . Вокруг какой точки на этой диаграмме происходит рассеяние?
- Можно ли производить *OLS*-оценивание коэффициентов рассматриваемой системы отдельно для каждого из двух уравнений? Если да, проведите такое оценивание.

- Используя объект *Model*, постройте реализацию модели:

$$\begin{aligned} DGP: \quad Q_t &= 125 - 0.5P_t + u_t, \\ P_t &= 0.5Q_{t-1} + 2.2v_t, \\ v_t &= u_t + w_t, \\ u_t, w_t &\sim i.i.d. N(0, 1), \quad t = 1, 2, \dots, 100. \end{aligned}$$

В качестве стартового значения для Q используйте стационарное значение этой переменной.

- Постройте диаграмму рассеяния для переменных Q_t , P_t .
- Произведите *OLS*-оценивание коэффициентов рассматриваемой системы отдельно для каждого из двух уравнений. Объясните полученные результаты.
- Произведите *2SLS*-оценивание коэффициентов рассматриваемой системы, выбрав соответствующие инструменты:
 - а) выполните непосредственно оба шага процедуры.
 - б) используйте встроенную процедуру *TSLS*-пакета *EViews*.
- Произведите *3SLS*- и *FIML*-оценивание коэффициентов рассматриваемой системы.

Задание 6. Оценивание систем одновременных уравнений

Повторите исследование, проведенное в примерах 1.2.1 и 1.2.2.

Методические указания. Предваряя исследование, проведенное в примере 1.2.1:

- постройте диаграмму рассеяния для переменных *Price* и *Quantity*, а также оцените модель линейной связи между этими двумя переменными;
- постройте диаграмму рассеяния для переменных Y_1 и Y_2 , а также оцените модель линейной связи между этими двумя переменными.

Задание 7. Прогнозирование по оцененной системе одновременных уравнений

Повторите исследование, проведенное в конце разд. 1.

Задание 8. Оценивание, диагностика и коррекция моделей систем одновременных уравнений (при отсутствии корреляции между ошибками, входящими в разные уравнения)

В табл. П-1 Приложения¹ содержатся данные об объеме продаж Y и цене P некоторого товара в индексной форме.

1. Методом наименьших квадратов оцените отдельно уравнение предложения и уравнение спроса. Что можно сказать о результатах оценивания?
2. Рассмотрите систему одновременных уравнений:

$$\begin{cases} Y_t = \alpha_0 + \alpha_1 P_t + \varepsilon_{1t} \\ P_t = \beta_0 + \beta_1 Y_t + \varepsilon_{2t} \end{cases}$$

Постройте приведенную форму этой системы и оцените коэффициенты приведенной формы. Можно ли на основании этих оценок однозначно оценить коэффициенты структурных уравнений?

¹ В этом и остальных заданиях к разделам данной части учебника даются ссылки на Приложение, содержащее необходимые статистические данные и их описание, помещенное после задания 45.

3. Рассмотрите расширенную систему:

$$\begin{cases} Y_t = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_{1t} \\ P_t = \beta_0 + \beta_1 Y_t + \beta_2 Z_t + \varepsilon_{2t}. \end{cases}$$

Значения экзогенных переменных X и Z также указаны в табл. П-1.

Постройте приведенную форму для расширенной системы, оцените коэффициенты приведенной формы. Можно ли по этим оцененным коэффициентам однозначно определить коэффициенты структурных уравнений расширенной системы? Почему? Найдите соответствующие оценки.

4. Оцените расширенную систему двухшаговым методом наименьших квадратов. Сравните полученные оценки с оценками, найденными при выполнении п. 3. Объясните результаты сравнения оценок.
5. Что можно сказать о коррелированности ошибок в разных уравнениях в совпадающие моменты времени? Имеет ли здесь смысл применять метод *3SLS* с точки зрения теории? Сравните результаты, полученные методами *2SLS*, *3SLS* и *FIML*.
6. Обратите внимание на значения статистики Дарбина — Уотсона для каждого оцененного структурного уравнения. О чем говорят эти значения? Для какого уравнения явно возникает проблема автокоррелированности ошибок?
7. Рассмотрите возможность использования для улучшения спецификации второго уравнения экзогенной переменной $W1$.
8. Рассмотрите расширенную систему, во второе уравнение которой добавлена переменная $W1$:

$$\begin{cases} Y_t = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_{1t} \\ P_t = \beta_0 + \beta_1 Y_t + \beta_2 Z_t + \beta_3 W1_t + \varepsilon_{2t}. \end{cases}$$

Остаются ли оба уравнения системы точно идентифицируемыми? Оцените эту систему двухшаговым методом наименьших квадратов ($W1$ дополняет список инструментов). Как изменились значения коэффициента детерминации и статистики Дарбина — Уотсона при добавлении новой переменной? Как изменилась оценка дисперсии случайной ошибки во втором уравнении системы? Можно ли удовлетвориться полученными результатами?

9. После оценивания модели, указанной в п. 8, проверьте следующую гипотезу:

$$H_0: \alpha_0 = -99, \quad \alpha_1 = \alpha_2 = 1, \quad \beta_0 = 2, \quad \beta_1 = -1, \quad \beta_2 = \beta_3 = 1.$$

Используйте статистику Вальда, основанную на выявлении значимости отклонений от выполнения указанных условий. Сколько линейных ограничений содержится в этой гипотезе? Можно ли согласиться с тем, что статистические данные порождены следующей моделью:

$$\begin{cases} Y_t = -99 + P_t + X_t + \varepsilon_{1t} \\ P_t = 2 - Y_t + Z_t + W1_t + \varepsilon_{2t}, \end{cases}$$

где ε_{1t} и ε_{2t} — не коррелированные между собой последовательности независимых случайных величин, имеющих одинаковое нормальное распределение с нулевым математическим ожиданием?

Методические указания (к пунктам задания).

1. В пакете *Econometric Views (EViews)* оцените парные модели регрессии переменной Y на константу и переменную P , переменной P на константу и переменную Y . Используйте объект *Equation*.
3. Для оценивания уравнений приведенной системы используйте объект *Equation*. Проверьте выполнение условий идентифицируемости для расширенной системы.
4. Создайте новый объект *System*, пусть он имеет название *sys1*. В меню этого объекта выберите:

Spec (Спецификация системы) →

$y=c(1)+c(2)*p+c(3)*x$ (спецификация первого структурного уравнения)

$p=c(4)+c(5)*y+c(6)*z$ (спецификация второго структурного уравнения)

inst x z (список инструментальных переменных)

→ *Estimate* → *Two-Stage Least Squares* → *OK*

В протоколе оценивания $C(1)$, $C(2)$, $C(3)$ соответствуют коэффициентам первого уравнения, $C(4)$, $C(5)$, $C(6)$ — коэффициентам второго уравнения (в порядке вхождения констант и переменных в эти уравнения).

5. После оценивания системы можно увидеть корреляционную матрицу остатков: *View* → *Residuals* → *Correlation Matrix*. Для применения 3SLS в п. 4 вместо *Estimate* → *Two-Stage Least Squares* закажите: *Estimate* → *Three-Stage Least Squares*. Для применения FIML в п. 4 вместо *Estimate* → *Two-Stage Least Squares* закажите: *Estimate* → *FIML*.
6. Дополнительно можно посмотреть на коррелограммы рядов остатков для каждого уравнения, автоматически генерируемых при выполнении команды *Estimate*.
7. Найдите коэффициент корреляции между переменной $W1$ и рядом остатков для второго уравнения.
8. Создайте новый объект *System*, пусть он имеет название *sys2*. В меню этого объекта выберите:

Spec →

$y=c(1)+c(2)*p+c(3)*x$

$p=c(4)+c(5)*y+c(6)*z+c(7)*w1$

inst x z w1

→ *Estimate* → *Two-Stage Least Squares* → *OK*

9. В меню объекта *sys2* выберите:

View → *Wald Coefficient Tests...* → $c(1)=-99$, $c(2)=c(3)=1$,
 $c(4)=2$, $c(5)=-1$, $c(6)=c(7)=1$

В протоколе оценивания приводятся P -значения, соответствующие асимптотическому распределению хи-квадрат с числом степеней свободы, равным количеству линейных ограничений в проверяемой гипотезе.

Задание 9. *Оценивание, диагностика и коррекция подобранных моделей систем одновременных уравнений (при наличии корреляции между ошибками, входящими в разные уравнения)*

Оцените двухшаговым методом наименьших квадратов систему одновременных уравнений:

$$\begin{cases} Y1_t = \alpha_0 + \alpha_1 P_t + \alpha_2 X_t + \varepsilon_{1t} \\ P_t = \beta_0 + \beta_1 Y_t + \beta_2 Z_t + \beta_3 W1_t + \varepsilon_{2t}. \end{cases}$$

1. Что можно сказать о наличии корреляции между ошибками в разных уравнениях системы? Следует ли в этом случае применять для оценивания системы 3SLS и FIML?
2. Сравните результаты, получаемые при оценивании системы методами 2SLS, 3SLS и FIML.

Методические указания. Данные возьмите из табл. П-1 Приложения. При оценивании системы методом 3SLS после спецификации системы используйте:

→ Estimate → Three-Stage Least Squares → OK

К разделу 2

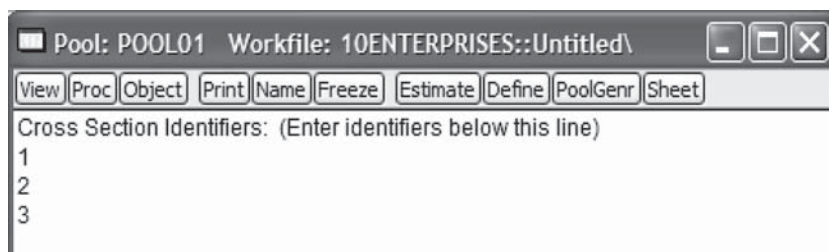
Задание 10. *Использование объекта Pool пакета Eviews 6 для анализа панельных данных*

В табл. 3.2 (пример 3.1.2) приведены ежегодные данные об объемах инвестиций y и прибыли x 3 предприятий ($N = 3$) за десятилетний период ($T = 10$). Ряды Y_i , X_i содержат данные по i -му предприятию ($i = 1, 2, 3$).

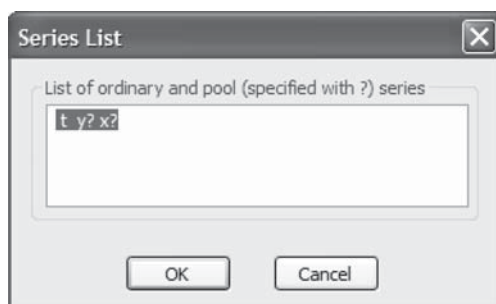
1. Используя эти статистические данные, оцените модель пула, выражающую зависимость объемов инвестиций от размеров прибыли предприятий. Знакома ли оцененная регрессия?

Методические указания. В пакете EViews создайте рабочий файл **10enterprises.wfl** на 10 наблюдений: *File/New/Workfile*. В окне *Workfile structure type* выберите *Unstructured/Undated*, в окне *Data range* укажите число наблюдений, т.е. 10. Скопируйте в этот файл ряды $Y_1, X_1, Y_2, X_2, Y_3, X_3$.

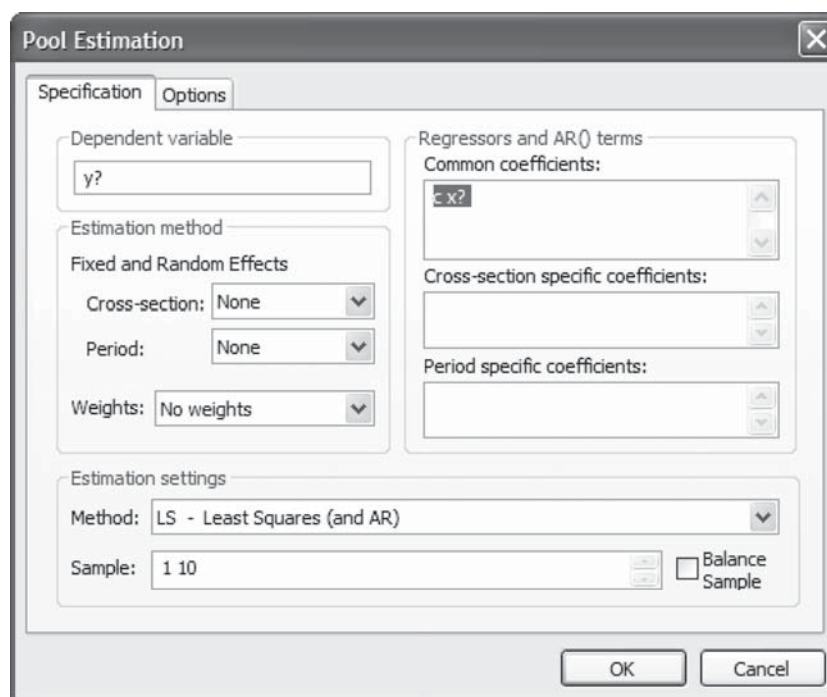
Создайте новый объект *Pool* и нажмите кнопку *Define*. На открывшейся странице перечислите идентификаторы предприятий: 1, 2, 3 — каждый на своей строке.



Нажмите на кнопку *Sheet*, в открывшемся окне укажите задействованные переменные, снабжая переменные, имеющие панельную структуру, знаком вопроса.

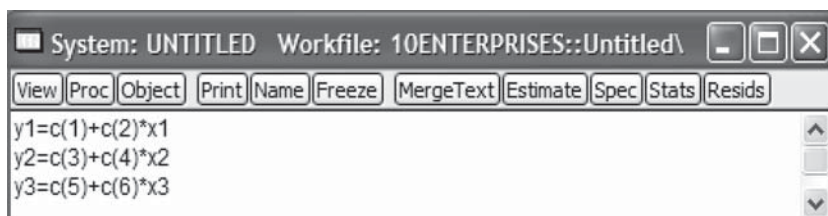


(Мы добавили здесь переменную $t @trend+1$, нумерующую наблюдения, хотя можно обойтись и без нее.) В результате получаем *stacked by cross-section*-представление данных, в котором они расположены блоками друг над другом: данные для 1-го предприятия приводятся в первых $T = 10$ строках, затем 10 строк занимают данные для 2-го предприятия, и последние 10 строк занимают данные для 3-го предприятия. Нажмите клавишу *Estimate* и специфицируйте уравнение регрессии, соответствующее модели пула.



2. Оцените модели парной регрессии отдельно по 3 предприятиям и сравните оценки дисперсий ошибок в уравнениях регрессии. Затем, допустив различие дисперсий ошибок для разных субъектов, используйте взвешенный метод наименьших квадратов (*WLS*). Изменились ли статистические выводы по сравнению с моделью пула?

Методические указания. Для отдельного оценивания моделей парной регрессии методом наименьших квадратов можно использовать объект *System*.



Чтобы применить взвешенный метод наименьших квадратов (*WLS*), укажите на использование весов для разных уравнений:

Estimation method: Weighted L.S. (equation weights)

3. Вместо *WLS* примените *OLS* с коррекцией стандартных ошибок оценок для отдельных уравнений. Изменились ли статистические выводы по сравнению с предыдущими?
4. а) Оценив систему, использованную в п. 2, получите корреляционную матрицу остатков уравнений системы.
б) Оцените ту же модель пула, допуская не только различие дисперсий ошибок для разных субъектов, но и перекрестную коррелированность этих ошибок.

Методические указания.

- а) После оценивания системы примените:

View /Residuals /Correlation Matrix

- б) В спецификации уравнения (объект *Pool*) укажите:

Weights: Cross-section SUR

5. Оцените модель пула, допуская различие дисперсий ошибок для разных субъектов и одинаковую *AR(1)*-структуру ошибок для каждого субъекта. Изменились ли статистические выводы по сравнению с предыдущими?

Методические указания. Специфицируя уравнение в объекте *Pool*, добавьте в окно *Common coefficients* составляющую *ar(1)*.

6. Оцените ту же модель, допустив различие дисперсий ошибок для разных субъектов и разную *AR(1)*-структуру ошибок для разных субъектов.

Методические указания. Специфицируя уравнение в объекте *Pool*, составляющую *ar(1)* поместите в окно

Cross-section specific coefficients

Задание 11. *Оценивание модели SUR (в пакете EViews 6)*

В файле **grunfeld.dta** (см. описание в Приложении) содержатся статистические данные о размере инвестиций (*invest*), рыночной цене (*mvalue*) и об акционерном капитале (*kstock*) 10 крупных корпораций США за 20-летний период — с 1935 по 1954 г.

Используя эти данные, оцените модель *SUR*, выражающую зависимость переменной *invest* от следующих факторов: *mvalue*, *kstock*.

Методические указания. Создайте новый объект *Pool* и укажите в нем идентификаторы корпораций (через *Define*). Специфицируйте модель, оценивая ее сначала методом наименьших квадратов, не учитывающим перекрестную коррелированность ошибок: *Weights: No weights*. Обратите внимание на то, что на сей раз, в отличие от модели пула, общих коэффициентов для разных субъектов нет: они могут быть разными для разных корпораций. Соответственно в окне *cross-section specific coefficients* поместите: *c x? y?*

Получите корреляционную матрицу ошибок (между уравнениями):

View/Residuals/Correlation Matrix.

Наличие заметной корреляции между ошибками в уравнениях для некоторых пар предполагает использование обобщенного метода наименьших квадратов *GLS* в модели *SURE*. Заказываем применение этого метода: *Weights: Cross-section SUR*.

Постройте ряды остатков (*Proc/Make Residuals*), составьте группу из этих остатков и рассмотрите таблицу описательных статистик для этих рядов (*View/Descriptive Stats/Common Sample*). Гипотеза нормальности отвергается только для 4-й корпорации.

Рассмотрите коррелограммы рядов остатков. Для этого в рамках объекта *Pool* используйте *View/Residuals/Graphs*. Форма данных графиков указывает на наличие автокоррелированности в этих рядах (по крайней мере, для некоторых из них). Следовательно, имеет смысл учесть такую автокоррелированность.

Оценивание модели с общей *AR(1)* не устраняет автокоррелированность остатков, которая заметно выражена для первого предприятия. Оценивание модели с различными *AR(1)* не устраняет эту автокоррелированность. Однако спецификация с общей *AR(1)* и разными *AR(2)* решает эту проблему.

Проверим теперь гипотезу о том, что коэффициенты при переменных *mvalue* и *kstock* не изменяются от корпорации к корпорации:

View/Coefficient Tests/Wald — Coefficient Restrictions.

Эта гипотеза отвергается.

Задание 12. *Проверка гипотез в модели SUR*

Обратитесь к данным задания 10, оцените для них модель *SUR*, допуская разную *AR(1)*-структуру ошибок для разных субъектов, исследуйте возможность сведения этой модели к более простым моделям.

Методические указания.

1. После оценивания модели *SUR* с разной *AR(1)*-структурой ошибок для разных субъектов, используя переходы

View/Coefficient Tests/Wald — Coefficient Restrictions:

- проверьте гипотезу о равенстве угловых коэффициентов при переменной x для всех предприятий;
 - проверьте возможность редукции модели *SUR* к модели пула.
2. Проверьте возможность редукции к модели пула модели с фиксированными эффектами. При спецификации этой модели укажите:

Common coefficients: x? ; *Cross-section specific coefficients: c ar(1)*

Приведенные ниже задания 13—27 выполняются с использованием пакета *Stata*, методические указания ориентированы на версию *Stata 10*.

Задание 13. *Оценивание модели пула — часовая оплата труда мужчин*

1. Используя статистические данные, содержащиеся в файле **males.dta**, описание которого имеется в Приложении, оцените модель пула, выражающую зависимость переменной *WAGE_LN* от следующих факторов:

SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB,

где $EXPER2 = EXPER^2$.

Значима ли оцененная регрессия в целом? Какие из указанных факторов значимо влияют на размер заработной платы? Как выражается влияние соответствующих факторов на размер заработной платы?

2. Оцените ту же модель, допустив различие дисперсий ошибок для разных субъектов и используя взвешенный метод наименьших квадратов (*WLS*). Изменились ли статистические выводы по сравнению с моделью пула?
3. В рамках п. 2 вместо использования *WLS* примените *OLS* с коррекцией стандартных ошибок оценок. Изменились ли статистические выводы по сравнению с предыдущими?
4. Оцените ту же модель, допустив не только различие дисперсий ошибок для разных субъектов, но и перекрестную коррелированность этих ошибок. Почему такая постановка задачи нереальна для анализируемой выборки?
5. Оцените ту же модель, допустив различие дисперсий ошибок для разных субъектов и одинаковую *AR(1)*-структуру ошибок для каждого субъекта. Используйте для оценивания *AR*-коэффициента два варианта: а) статистику Дарбина — Уотсона, б) регрессионную оценку. Изменились ли статистические выводы по сравнению с предыдущими?
6. Оцените ту же модель, допустив различие дисперсий ошибок для разных субъектов и разную *AR(1)*-структуру ошибок для разных субъектов. Используйте для оценивания *AR*-коэффициента два варианта: а) статистику Дарбина — Уотсона, б) регрессионную оценку. Изменились ли статистические выводы по сравнению с предыдущими?

Методические указания.

- Запустите пакет *Stata 10*.
- Задайте достаточный объем памяти для расчетов: в окне команд напечатайте команду *set memory 10M*.
- В окне команд напечатайте команду *set matsize 545* — это информирует программу о количестве субъектов в выборке.

- Откройте файл **males.dta**.
- Откройте редактор данных (*Data/Data Editor*). Появится таблица статистических данных. Обратите внимание на структуру этой таблицы. Закройте редактор данных. (Если данные изменены в процессе просмотра, то для сохранения этих изменений нажмите *Preserve*. Если изменения не надо сохранять, нажмите *Restore*.)

1. В пакете *Stata* используйте команду

```
xtgls WAGE_LN SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB.
```

(Обратите внимание на то, что в пакете *Stata* заглавные и строчные буквы различаются.) Для просмотра результатов, не уместившихся в окне, нажмите на *more*.

2. Используйте ту же команду, но с опцией *panels(het)*:

```
xtgls WAGE_LN SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB, panels(het).
```

(Для упрощения набора однотипных команд удобно использовать окно *Review*. Достаточно выбрать в этом окне нужную или близкую к нужной команду и щелкнуть по ней. Выделенная команда копируется в окно команд и там может быть выполнена, быть может, после соответствующего дополнения или редактирования.)

3. Используйте команду *xtpcse* (*panel corrected standard errors*) с опцией *het*:

```
xtpcse WAGE_LN SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB, het.
```

Перед применением этой команды укажите переменные — идентификаторы субъектов и времени:

```
xtset NR YEAR.
```

4. В этом случае используется либо команда *xtgls* — с опцией *panels(correlated)*, либо команда *xtpcse* — без опций. Однако прежде чем механически использовать эту команду, обратите внимание на количество парных корреляций, которое необходимо оценить.

5. (а) Используйте команду *xtpcse* с опциями *correlation(ar1) rhotype(dw) het*:

```
xtpcse WAGE_LN SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB,  
correlation(ar1) rhotype(dw) het.
```

- (б) Используйте команду *xtpcse* с опциями *correlation(ar1) rhotype(tscorr) het*.

6. (а) Используйте команду *xtpcse* с опциями *correlation(pсар1) rhotype(dw) het*.

- (б) Используйте команду *xtpcse* с опциями *correlation(pсар1) rhotype(tscorr) het*.

В пп. 5 и 6 можно применить команду *xtgls*, только тогда вместо опции *het* (в *xtpcse*) надо использовать опцию *panels(het)*.

Задание 14. Оценивание модели пула с перекрестной коррелированностью ошибок

1. Используя статистические данные, содержащиеся в файле **grunfeld.dta**, описание которого имеется в Приложении, оцените модель пула, выражающую зависимость переменной *invest* от следующих факторов: *mvalue*, *kstock*. Выполните все пункты задания 13 для этого случая.

2. Почему в этом случае можно использовать команду *xtgls* с опцией *panels (correlated)* и команду *xtpcse*?
3. Используйте эти команды в п. 5 и 6 задания 13.

Методические указания.

1. После открытия файла здесь можно использовать команду *list* для просмотра таблицы данных (в задании 13 такой просмотр слишком затянулся бы из-за большого количества субъектов).
2. Сравните количество оцениваемых парных корреляций с количеством наблюдений.

Задание 15. Оценивание модели SUR — данные Grunfeld

Используя статистические данные, содержащиеся в файле **grunfeld.dta**, оцените модель SUR, выражающую зависимость переменной *invest* от следующих факторов: *mvalue*, *kstock*.

Методические указания. В пакете *Stata* произвести GLS-оценивание модели SUR можно следующим образом. Сначала необходимо изменить формат представления данных (от «длинного» представления перейти к «широкому»):

```
reshape wide invest mvalue kstock, i(year) j(company),
```

(вернуться к исходному представлению можно командой

```
reshape long invest mvalue kstock, i(year) j(company)).
```

В таком представлении для каждого года выделена своя строка (всего 20 строк). В первой графе указаны годы, входящие в период наблюдений: 1935—1954. В следующих трех графах приведены значения новых переменных: *invest1*, *mvalue1*, *kstock1*, представляющих значения переменных *invest*, *mvalue*, *kstock* для компании с кодом 1. Далее идут три графы, в которых приведены значения новых переменных: *invest2*, *mvalue2*, *kstock2*, представляющих значения переменных *invest*, *mvalue*, *kstock* для компании с кодом 2, и т.д. В последних трех графах приведены значения переменных *invest10*, *mvalue10*, *kstock10*, представляющих значения переменных *invest*, *mvalue*, *kstock* для компании с кодом 10.

После перехода к новому формату выбираем в главном меню:

Statistics → *Linear regression and related* → *Multiple equation models* →
Seemingly Unrelated Regressions

Для каждой компании выделяется свое уравнение (*Equation*), для которого следует указать переменные, находящиеся в левой и правой частях уравнения:

Equation 1 invest1 mvalue1 kstock1

Equation 2 invest2 mvalue2 kstock2

.....

Equation 10 invest10 mvalue10 kstock10

→ OK

Указанные действия равносильны использованию команды *sureg*:

```
sureg (invest1 = mvalue1 kstock1) (invest2 = mvalue2 kstock2) ... (invest10 =  
= mvalue10 kstock10).
```

В протоколе оценивания приводятся результаты раздельного оценивания 10 уравнений регрессии:

переменной *invest1* на константу и переменные *mvalue1*, *kstock1*,
 переменной *invest2* на константу и переменные *mvalue2*, *kstock2*,

 переменной *invest10* на константу и переменные *mvalue10*, *kstock10*.

Задание 16. Проверка гипотез в модели SUR — данные Grunfeld

После оценивания модели SUR в задании 15 можно перейти к проверке некоторых гипотез.

1. Проверьте гипотезу отсутствия корреляции между ошибками в разных уравнениях, относящимися к одному и тому же моменту времени.
2. Проверьте гипотезу о том, что коэффициенты при переменных *mvalue* и *kstock* не изменяются от субъекта к субъекту (сведение к модели ковариационного анализа).
3. Можно ли считать случайные ошибки в модели регрессии в рамках каждой отдельной корпорации некоррелированными?
4. Имея в виду ответ на вопрос п. 3, можно ли полностью доверять результатам проверки гипотез в п. 1 и 2?
5. Имея в виду результаты, полученные в п. 1—4, выскажите свое мнение о корректности использования модели пула в задании 14.

Методические указания.

1. В команде *sureg* примените опцию *corr* (это равносильно указанию в меню опции *Perform Breush-Pagan test*).

Ввиду небольшого количества данных можно выполнить коррекцию стандартных ошибок и корреляций (опция *small corr*).

2. После оценивания модели SUR примените команду:

```
test (mvalue1 = mvalue2 = mvalue3 = mvalue4 = mvalue5 =  
= mvalue6 = mvalue7 = mvalue8 = mvalue9 = mvalue10) (kstock1 =  
= kstock2 = kstock3 = kstock4 = kstock5 = kstock6 =  
= kstock7 = kstock8 = kstock9 = kstock10)
```

3. Скопируйте данные в их широком представлении в новый *dta*-файл, образуйте в новом файле переменную $t = \text{year} - 1934$ (*Data/Create or Change Variables/Create New Variable*) и объявите эту переменную как переменную, индексирующую время: *tsset t*. После этого можно для каждой корпорации оценить регрессию *invest* на *mvalue* и *kstock* (все три переменные берутся с соответствующим номером): *regress invest1 mvalue1 kstock1* и произвести проверку на автокоррелированность ошибок в соответствующей модели регрессии. Для выполнения такой проверки можно использовать команды: *estat dwatson*, *estat durbinalt*, *estat bgodfrey*.

✓ **Замечание.** При оценивании некоторых моделей требуется предварительно увеличить значение *matsize*, на что указывает появляющееся сообщение на черном экране.

Задание 17. *Оценивание межгрупповой регрессии — часовая оплата труда женщин*

Используя статистические данные из файла **nlswork.dta**, опишите, как описано в Приложении, оцените коэффициенты межгрупповой (*between*) регрессии, выражающей зависимость переменной *ln_wage* от переменных *tenure*, *age*, *age_2*, *not_smsa*, где $age_2 = age^2$.

Как изменяется с возрастом размер часовой заработной платы, после какого возраста это влияние убывает? Как изменяется размер часовой заработной платы с увеличением продолжительности работы на последнем месте работы? Как влияет на размер заработной платы проживание вне столичных регионов?

Методические указания. Для порождения новых переменных используйте команду *generate*. Например:

generate age_2 = age^2

Для оценивания модели используйте команду *xtreg* с опцией *be*:

xtreg ln_w tenure age age_2 not_smsa, be i(idcode).

Задание 18. *Оценивание межгрупповой регрессии — часовая оплата труда мужчин*

Используя статистические данные из файла **males.dta**, оцените коэффициенты межгрупповой (*between*) регрессии, выражающей зависимость переменной *WAGE_LN* от следующих факторов: *SCHOOL*, *EXPER*, *EXPER2*, *UNION*, *MAR*, *BLACK*, *HISP*, *PUB*, где $EXPER2 = EXPER^2$.

Какие из указанных факторов значимо влияют на размер заработной платы? Как выражается влияние соответствующих факторов на размер заработной платы? Сравните полученные результаты с результатами, полученными в задании 13, где рассматривалась модель пула.

Методические указания. Для оценивания модели используйте команду *xtreg* с опцией *be*.

Задание 19. *Оценивание моделей с фиксированными и случайными эффектами — часовая оплата труда мужчин*

1. Используя статистические данные из файла **males.dta**, оцените модель с фиксированными эффектами для переменной *WAGE_LN* и объясняющими переменными *SCHOOL*, *EXPER*, *EXPER2*, *UNION*, *MAR*, *BLACK*, *HISP*, *PUB*. Объясните результаты, касающиеся оценок коэффициентов при переменных *SCHOOL*, *BLACK* и *HISP*.
2. Используя статистические данные из файла **males.dta**, оцените модель со случайными эффектами для переменной *WAGE_LN* и объясняющими переменными *SCHOOL*, *EXPER*, *EXPER2*, *UNION*, *MAR*, *BLACK*, *HISP*, *PUB*.

Методические указания.

1. В пакете *Stata* используйте команду *xtreg* с опцией *fe*.
2. В пакете *Stata* используйте команду *xtreg* с опцией *re*.

Задание 20. Сравнение результатов оценивания модели пула, межгрупповой регрессии, моделей с фиксированными и случайными эффектами — часовая оплата труда мужчин

1. Сравните результаты, полученные при оценивании модели пула, межгрупповой регрессии, моделей с фиксированными и случайными эффектами в заданиях 13, 18 и 19.
2. Какие из полученных оценок коэффициентов являются состоятельными? Примените критерий Хаусмана для проверки гипотезы о том, что модель, порождающая данные, является моделью со случайными эффектами. Объясните результат применения этого критерия.
3. Проверьте гипотезу об отсутствии (фиксированных) индивидуальных эффектов (о возможности сведения модели с фиксированными эффектами к модели пула).
4. Проверьте гипотезу однородности дисперсий между субъектами. Если эта гипотеза отвергается, проведите новое оценивание, учитывающее неоднородность дисперсий между субъектами.

Методические указания (к пунктам задания).

2. Если выполнены предположения модели со случайными эффектами, то все полученные оценки состоятельны; в противном случае состоятельной является только *FE*-оценка. Для применения критерия Хаусмана используйте следующую последовательность команд:

xtreg с опцией *fe* → *estimates store fe* → *xtreg* с опцией *re* →
estimates store re → *hausman fe re*

3. Результат проверки указанной гипотезы приведен в последней строке протокола оценивания *FE*-модели.
4. Для проверки этой гипотезы можно использовать модифицированный критерий Вальда: в пакете *Stata* используется команда *xttest3*, применяемая вслед за применением команды *xtreg* с опцией *fe*. (Эта команда скачивается с сайта *stata.com*.)

Задание 21. Проверка статистической значимости временных эффектов

Используя статистические данные из файла **males.dta**, оцените двунаправленную модель с фиксированными эффектами для переменной *WAGE_LN* с объясняющими переменными *SCHOOL*, *EXPER*, *EXPER2*, *UNION*, *MAR*, *BLACK*, *HISP*, *PUB*. Являются ли временные эффекты статистически значимыми?

Методические указания. Используйте команду *xtreg fe*:

xtreg WAGE_LN SCHOOL EXPER EXPER2 UNION MAR BLACK HISP PUB
IYEAR_2 IYEAR_3 IYEAR_4 IYEAR_5 IYEAR_6 IYEAR_7 IYEAR_8, fe

После ее выполнения используйте команду

test IYEAR_2 = IYEAR_3 = IYEAR_4 = IYEAR_5 = IYEAR_6 =
= IYEAR_7 = IYEAR_8 = 0

Задание 22. *Оценивание модели с эндогенными объясняющими переменными — часовая оплата труда женщин*

1. Используя статистические данные из файла **nlswork.dta**, описание которого имеется в Приложении, оцените модель с фиксированными эффектами, выражающую зависимость переменной \ln_wage от переменных

$$tenure\ age\ age_2\ not_smsa,$$

где $age_2 = age^2$.

2. Все ли оцененные коэффициенты статистически значимы? Согласуются ли их знаки со знаками оценок коэффициентов, полученных в задании 20? Какая часть изменчивости объясняется индивидуальными эффектами?
3. Стаж работы на последнем рабочем месте может зависеть от принадлежности к профсоюзу и от региона проживания (юг — не юг). При этом ошибки в уравнении, описывающем такую связь, могут быть коррелированными с ошибками в уравнении для логарифма заработной платы, так что переменная $tenure_{it}$ в уравнении

$$\ln w_{it} = \beta_1 tenure_{it} + \beta_2 age_{it} + \beta_3 age2_{it} + \beta_4 not_smsa_{it} + \mu + \alpha_i + u_{it}$$

коррелирована с ошибкой u_{it} , и для получения состоятельных оценок коэффициентов этого уравнения приходится прибегать к методу инструментальных переменных. Проведите оценивание этим методом в рамках моделей с фиксированными и случайными эффектами. Сравните полученные результаты.

Методические указания (к пунктам задания).

1. Для порождения новых переменных используйте команду *generate*. В данном случае:

$$generate\ age_2 = age^2.$$

Для оценивания модели с фиксированными эффектами используйте команду *xtreg* с опцией *fe*:

$$xtreg\ ln_w\ tenure\ age\ age_2\ not_smsa, fe\ i(idcode)$$

3. При сделанных предположениях в качестве инструментов для $tenure_{it}$ можно использовать *union* и *south*. Полный список инструментов, обеспечивающий однозначную идентификацию коэффициентов μ и $\beta_1, \beta_2, \beta_3, \beta_4$, включает 5 переменных: *union, south, age, age2, not_smsa*. Используйте эти переменные в качестве инструментов. Для этого примените команду

$$xtivreg\ ln_w\ age\ age_2\ not_smsa\ (tenure = union\ south), fe\ i(idcode)$$

(в модели с фиксированными эффектами) или команду

$$xtivreg\ ln_w\ age\ age_2\ not_smsa\ (tenure = union\ south), re\ i(idcode).$$

(в модели со случайными эффектами).

Задание 23. *Оценивание модели с индивидуально-специфическими переменными*

В задании 19 при оценивании модели с фиксированными эффектами для размера часовой оплаты труда мужчин (файл **males.dta**) оказались выметенными переменные *SCHOOL, BLACK* и *HISP*. Для получения оценок влияния этих переменных используйте методику, изложенную в теме 3.4 разд. 3 лекционной части.

Методические указания. Оцените модель с фиксированными эффектами для переменной *WAGE_LN* с объясняющими переменными *EXPER*, *EXPER2*, *UNION*, *MAR*, *PUB*:

```
xtreg WAGE_LN EXPER EXPER2 UNION MAR PUB, fe
```

После этого выполните команду

```
predict residfe, ue
```

(опция *ue* сохраняет значения комбинированных остатков: сумма эффекта и случайной ошибки), а затем оцените *between* модель для переменной *residfe* с объясняющими переменными *SCHOOL*, *BLACK* и *HISP*:

```
xtreg residfe SCHOOL BLACK HISP, be.
```

Задание 24. Автокоррелированные ошибки

1. Используя те же данные (файл **males.dta**), примените метод оценивания, допускающий автокоррелированность ошибок в модели с фиксированными эффектами.

Методические указания. Используйте команду

```
xtregar WAGE_LN SCHOOL EXPER EXPER2 UNION MAR  
BLACK HISP PUB, fe rhotype(dw) lbi
```

Здесь *lbi* — значение статистики критерия *Baltagi-Wu*, заменяющего критерий Дарбина — Уотсона при несбалансированных данных. (Заметим, что таблиц критических значений этой статистики не существует.)

Вместо *dw* в *rhotype* можно использовать другие варианты оценивания автокорреляции первого порядка.

2. Сравните значения статистик критериев *DW* и *Baltagi-Wu*, применяя тот же метод оценивания к данным из файла **nlswork**.

Задание 25. Оценивание динамической модели для панельных данных обобщенным методом моментов

Используя данные из задания 10, постройте динамическую модель первого порядка для объемов инвестиций 3 предприятий, когда в качестве экзогенной переменной используется прибыль. Примените обобщенный метод моментов с выбором инструментов, предложенным Ареллано и Бондом. О чем говорят полученные результаты?

Методические указания. Здесь имеется в виду модель:

$$y_{it} = \mu + \alpha_i + \gamma y_{i,t-1} + \beta_1 x_{it} + \beta_2 x_{i,t-1} + u_{it}, \quad i = 1, 2, 3, \quad t = 3, \dots, 10.$$

Дифференцирование приводит к модели для разностей:

$$\Delta y_{it} = \gamma \Delta y_{i,t-1} + \beta_1 \Delta x_{it} + \beta_2 \Delta x_{i,t-1} + \Delta u_{it}, \quad i = 1, 2, 3, \quad t = 3, \dots, 10.$$

Если следовать описанию программы *xtabond* в пакете *Stata* 8, то в этой программе в качестве инструментов используются переменные, указанные в таблице, а также $\Delta x_{i3} + \dots + \Delta x_{i10}$ и $\Delta x_{i2} + \dots + \Delta x_{i9}$. Это дает всего 38 (36 + 2) моментных условий. Поскольку модель в разностях содержит 3 неизвестных коэффициента,

то для оценивания этих коэффициентов достаточно использовать только 3 моментных условия, а остальные 35 ($38 - 3$) условий — избыточные.

t	Инструменты	Количество
3	y_{i1}	1
4	y_{i1}, y_{i2}	2
5	y_{i1}, y_{i2}, y_{i3}	3
6	$y_{i1}, y_{i2}, y_{i3}, y_{i4}$	4
7	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}$	5
8	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}$	6
9	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}, y_{i7}$	7
10	$y_{i1}, y_{i2}, y_{i3}, y_{i4}, y_{i5}, y_{i6}, y_{i7}, y_{i8}$	8
	ВСЕГО	36

Используйте команду

xtabond y x l1(x), lags(1)

Здесь $l1(x)$ указывает на включение в правую часть уравнения запаздывающего значения переменной x ; *lags(1)* указывает на включение в правую часть уравнения запаздывающего значения переменной y . В протоколе оценивания приведены оценки коэффициентов в продифференцированном уравнении и соответствующие им P -значения, значение статистики критерия Саргана для проверки гипотезы о выполнении избыточных ограничений и соответствующее ему P -значение, а также (в версии *Stata* 8) значения статистик критериев Ареллано — Бонда для проверки гипотезы об отсутствии в ряде остатков автокорреляций первого и второго порядков и соответствующие им P -значения.

В протоколе:

D1 — разность первого порядка;

LD — запаздывающая разность первого порядка.

Sargan test of over-identifying restrictions:

$\chi^2(35) = 21.81$ $Prob > \chi^2 = 0.9600$

Arellano-Bond test that average autocovariance in residuals of order 1 is 0:

H_0 : no autocorrelation $z = -2.56$ $Pr > z = 0.0106$

Arellano-Bond test that average autocovariance in residuals of order 2 is 0:

H_0 : no autocorrelation $z = 0.77$ $Pr > z = 0.4427$.

На правильность выбора инструментов указывает P -значение для статистики Саргана.

На правильность сделанных предположений об ошибках указывают отклонение гипотезы об отсутствии автокоррелированности первого порядка и неотклонение гипотезы об отсутствии автокоррелированности второго порядка.

Задание 26. Оценивание моделей для панельных данных о производстве молока

На сайте <http://pages.stern.nyu.edu/~wgreene/Econometrics/PanelDataSets.htm> в файле **Spanish Dairy Farm Production** содержатся данные о производстве молока на 247 фермах Испании за период с 1993 по 1998 г.

Переменные:

FARM — идентификатор фермы;

YEAR — год, 93, 94, ..., 98.

Объясняющие переменные:

COWS — количество коров, $X1$ — отклонения $\ln(COWS)$ от его среднего по всем фермам;

LAND — земельные угодья, $X2$ — определяется аналогично;

LABOR — затраты труда, $X3$ — определяется аналогично;

FEED — корм, $X4$ — определяется аналогично;

$X11, X22, X33, X44, X12, X13, X14, X23, X24, X34$ — квадраты и произведения переменных $X1, X2, X3, X4$;

YEAR93, ..., YEAR98 — дамми-переменные для 1993, ..., 1998 г.

Объясняемые переменные:

MILK — объем производства молока на ферме;

$YIT = \ln(MILK)$

Используя эти данные, оцените различные модели, применявшиеся в предыдущих задачах, с целью выяснения влияния указанных факторов на объем производства молока на ферме.

Задание 27. Оценивание моделей для панельных данных о спросе на бензин

В файле <http://www.wiley.com/legacy/wileychi/baltagi/supp/Gasoline.dat> приведены данные о спросе на бензин в 18 странах *OECD* (*Organization for Economic Cooperation and Development*) в период с 1960 по 1978 г.

Переменные:

COUNTRY — название страны (*AUSTRIA, BELGIUM, CANADA, DENMARK, FRANCE, GERMANY, GREECE, IRELAND, ITALY, JAPAN, NETHERLA, NORWAY, SPAIN, SWEDEN, SWITZERL, TURKEY, UK, USA*);

YEAR — год, 1960—1978.

Объясняющие переменные (в логарифмах уровней):

LINCOME_P — доход на душу населения;

LRPMG — реальная цена бензина;

LCARPCAP — количество автомобилей на душу населения;

Объясняемая переменная (в логарифмах уровней):

LGASPCAR — потребление на один автомобиль.

Используя эти данные, оцените различные модели, использованные в предыдущих задачах, с целью выяснения зависимости потребления бензина в расчете на один автомобиль от указанных факторов.

Методические указания. Сначала рассмотрите модель зависимости *LGASPCAR* от *LRPMG*. После этого оцените полную модель и сравните полученные результаты.

Задание 28. *Интерпретация результатов оценивания моделей с фиксированными и случайными эффектами*

Статистические данные содержат сведения о 5107 индивидах, выбранных случайным образом из граждан США, имевших в 1980 г. возраст от 16 до 23 лет. Опросы производились ежегодно в период с 1980 по 1994 г., а также в 1996, 1998 и 2000 гг. (всего 18 опросов).

В банке данных содержится информация по следующим переменным:

Количественные переменные	
<i>earnings</i>	недельная заработная плата (средняя, в долларах)
<i>s</i>	количество лет, затраченных на образование
<i>exper</i>	общий стаж работы (в годах)
Дамми-переменные	
<i>male</i>	= 1 для мужчин, = 0 для женщин
<i>married</i>	= 1 для проживающих с партнером, = 0 для одиноких
<i>union</i>	= 1 для членов профсоюза
<i>ethblack</i>	= 1 для чернокожего населения, = 0 для других
<i>ethhisp</i>	= 1 для индивидов, происходящих из Латинской Америки, = 0 для других

Для исследования факторов, влияющих на размер заработной платы, оценивались модели линейной регрессии логарифма заработной платы на константу и переменные *s*, *exper*, *expsq* ($=\text{exper}^2$), *union*, *married*, *ethblack*, *ethhisp*:

- в форме модели пула;
- в форме модели с фиксированными эффектами;
- в форме модели со случайными эффектами.

Были получены следующие результаты (отдельно для мужчин и для женщин) (см. табл.).

Переменная	Мужчины			Женщины		
	<i>OLS</i>	<i>FE</i>	<i>RE</i>	<i>OLS</i>	<i>FE</i>	<i>RE</i>
<i>const</i>	5.500 (0.030)	6.176 (0.100)	5.702 (0.042)	5.328 (0.031)	6.137 (0.097)	5.512 (0.045)
<i>s</i>	0.089 (0.001)	0.045 (0.008)	0.080 (0.003)	0.089 (0.002)	0.036 (0.003)	0.081 (0.003)
<i>exper</i>	0.041 (0.004)	0.037 0.003	0.035 0.003	0.051 (0.004)	0.037 (0.003)	0.041 (0.003)

Окончание таблицы

Переменная	Мужчины			Женщины		
	<i>OLS</i>	<i>FE</i>	<i>RE</i>	<i>OLS</i>	<i>FE</i>	<i>RE</i>
<i>expersq</i>	−0.0006 (0.0001)	−0.0004 (0.0001)	−0.0004 (0.0001)	−0.0008 (0.0001)	−0.0004 (0.0001)	−0.0006 (0.0001)
<i>union</i>	0.210 (0.010)	0.118 (0.011)	0.042 (0.010)	0.131 (0.012)	0.068 (0.013)	0.085 (0.012)
<i>married</i>	0.153 (0.008)	0.042 (0.011)	0.074 (0.010)	−0.012 (0.009)	0.004 (0.012)	−0.013 (0.010)
<i>ethblack</i>	−0.167 (0.012)	—	−0.205 (0.025)	−0.115 (0.012)	—	−0.118 (0.024)
<i>ethhisp</i>	−0.034 (0.015)	—	−0.070 (0.031)	(0.024) (0.016)	—	0.013 (0.031)

- Что означает отсутствие оценок коэффициентов при некоторых переменных для моделей с фиксированными эффектами? Почему этого не наблюдается для количества лет, затраченных на образование, для членства в профсоюзе и для семейного статуса?
- Как интерпретируется оцененная константа в *FE*-модели?
- Чем может объясняться существенное расхождение статистически значимых оценок некоторых коэффициентов в *FE*- и *RE*-моделях? Являются ли состоятельными в данном случае *OLS* для пула, *FE*-оценка и *RE*-оценка? Какой критерий может помочь в решении этого вопроса? На чем основана статистика соответствующего критерия? Какое асимптотическое распределение имеет эта статистика?
- Что дает (количественно) для работающего членство в профсоюзе?
- Что дает (количественно) для работающего наличие семьи? Имеется ли в этом отношении разница между мужчинами и женщинами?
- Как влияет на размер заработной платы этническое происхождение работающего? Одинаково ли такое влияние для мужчин и женщин?

К разделу 3

Задание 29. Построение пробит-модели банкротства предприятий, котирующихся на бирже (в пакете *EViews*)

В табл. П-2 Приложения приведены данные на конец одного и того же года о финансовых показателях 66 предприятий, акции которых котируются на бирже. В течение двух последующих лет половина из этих предприятий обанкротилась (для них бинарная переменная Y принимает значение 0). Постройте по этим данным пробит-модель выживания предприятия. В качестве потенциальных объясняющих переменных используйте переменные X_1 — X_5 , участвующие в построении известного Z -индекса Альтмана кредитоспособности (Z -счет):

$$X_1 = \frac{\text{Чистый оборотный капитал}}{\text{Общая сумма активов}},$$

$$X_2 = \frac{\text{Нераспределенная прибыль}}{\text{Общая сумма активов}},$$

$$X_3 = \frac{\text{Прибыль до вычета процентов и налогов}}{\text{Общая сумма активов}},$$

$$X_4 = \frac{\text{Рыночная стоимость активов за вычетом задолженности}}{\text{Балансовая стоимость общей суммы обязательств}},$$

$$X_5 = \frac{\text{Выручка}}{\text{Общая сумма активов}}.$$

1. Начните рассмотрение с модели, включающей все потенциальные объясняющие переменные.
2. Оставьте для дальнейшего рассмотрения только те модели, для которых итерационный процесс вычисления оценок максимального правдоподобия сходится.
3. Из оставшихся моделей выберите те, для которых оценки всех коэффициентов статистически значимы.
4. Среди оставшихся после выполнения п. 1 и 2 моделей выберите лучшую, используя для сравнения альтернативных моделей критерии *LRI*, *AIC*, *SC*, *HQ*, количество неправильных предсказаний, а также критерий Хосмера — Лемешоу.
5. Дайте интерпретацию коэффициентов выбранной модели.

Методические указания. В пакете *EViews* создайте новый объект *Equation* — пусть он называется *eq1*:

Object → *New Object* → *eq1* →
Equation specification: $y \text{ с } x_1 \ x_2 \ x_3 \ x_4 \ x_5$
Estimation Settings: Method: *BINARY* →
Binary estimation method: *Probit* → *OK*.

Оценки коэффициентов и соответствующие *P*-значения, а также значения критериев *LRI*, *AIC*, *SC*, *HQ* приведены непосредственно в протоколе оценивания. Для определения количества правильных и неправильных прогнозов по оцененной модели выполните следующие действия (в рамках объекта *eq1*):

View → *Expectation-Prediction Table* → *Success if prob is greater than 0.5* → *OK*.

Для применения критерия Хосмера – Лемешоу:

View → *Goodness-of-Fit Test (Hosmer-Lemeshow)* → *OK*.

По умолчанию производится разбиение множества наблюдений на 10 групп в соответствии с предсказанными вероятностями получения значения 1.

Разбиение на группы может производиться также по значениям одной из объясняющих переменных — этот подход используется в критерии *Andrews*. Для этого следует использовать:

View → Goodness-of-Fit Test (Hosmer-Lemeshow) →

В открывшемся окне отметить *Series or Series Expression* и указать ряд, используемый для группирования.

Задание 30. Построение логит- и гомпит-моделей банкротства предприятий, kotiрующихся на бирже (в пакете *EViews*)

1. В постановке задания 29 подберите по тем же данным логит- и гомпит-модели. Сравните наилучшие пробит-, логит- и гомпит-модели.
2. Вычислите для каждого предприятия значение Z-индекса Альтмана. Как классифицируются предприятия по этому индексу?
3. Дайте интерпретацию коэффициентов выбранных моделей. Каковы особенности интерпретации коэффициентов в логит-модели?

Методические указания (к пунктам задания).

1. Вместо *Binary estimation method: Probit* выберите:

Binary estimation method: Logit

или *Binary estimation method: Gompit* соответственно.

2. Значение индекса Альтмана вычисляется по формуле:

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.99X_5.$$

Предприятия, для которых $Z > 2.99$, причисляются к числу финансово устойчивых, предприятия, для которых $Z < 1.81$, причисляются к безусловно-несостоятельным. Интервал $[1.81, 2.99]$ составляет зону неопределенности.

Задание 31. Построение моделей банкротства предприятий, kotiрующихся на бирже (в пакете *Stata*)

Проведите анализ данных, рассмотренных в заданиях 29 и 30, используя пакет *Stata*.

Методические указания. Создайте новый файл в формате *Stata* и перенесите в него данные из табл. П-2, приведенной в Приложении (данные о финансовых показателях 66 предприятий).

Для построения пробит-модели воспользуйтесь командой *probit*. После выполнения этой команды получите результаты классификации и процент правильно классифицированных исходов (команда *lstat*), результаты применения критерия согласия Хосмера — Лемешоу (команда *lfit*), значения информационных критериев Акаике и Шварца — команда *estimates table, stats(aic bic)*. Сравните результаты, полученные в пакете *Stata*, с результатами, полученными в заданиях 29 и 30 при применении пакета *EViews*.

Задание 32. Построение моделей, объясняющих наличие или отсутствие у семьи автомобиля величиной ее среднедушевого дохода (в пакете *EViews*)

По образцу примера, рассмотренного в теме 4.1, сгенерируйте данные о среднедушевых месячных доходах (в условных единицах) 1000 семей — переменная X . Значение переменной Y равно 1, если семья имеет собственный автомобиль, и равно 0 — в противном случае.

1. Постройте модели, объясняющие наличие или отсутствие автомобиля у семей величиной их среднедушевого дохода. Дайте интерпретацию коэффициентов подобранных моделей. Как интерпретируется значение коэффициента при переменной X в логит-модели? Выберите лучшую среди этих моделей, используя для сравнения альтернативных моделей критерии LRI , AIC , SC , HQ , количество неправильных предсказаний, а также критерий Хосмера — Лемешоу.
2. Проверьте предположение о гомоскедастичности ошибок в модели для латентной переменной, соответствующей пробит-модели.
3. Проверьте предположение о нормальности ошибок в модели для латентной переменной.

Методические указания (к пунктам задания).

2. Наряду со стандартной пробит-моделью рассмотрите модель с гетероскедастичностью в форме

$$D(\varepsilon_i | x_i) = \exp(k x_i).$$

Сравните максимумы логарифмов правдоподобий, полученные при оценивании этих двух моделей. Для стандартной пробит-модели соответствующая величина указывается в протоколе оценивания в строке *Log likelihood*. Для оценивания расширенной модели создайте объект *Logl*:

Object → *New Object* → *Logl*.

Запишите в открывшемся окне выражение для логарифмического правдоподобия в расширенной модели:

@logl logl2 (указывается название ряда значений ряда *logl*)

$u=c(1)+c(2)*x2$ ряд значений $x_i^T \theta$

$v=@sqrt(\exp(c(3)*x2))$ (выражение для корня из дисперсии)

@param c(1) -3 c(2) 0.003 c(3) 0.0001 (задаются стартовые значения параметров; $c(1)$ и $c(2)$ — близкие к оценкам в оцененной стандартной пробит-модели, $c(3)$ — близкое к нулю)

$logl2=y*log(@cnorm(u/v))+(1-y)*log(1-@cnorm(u/v))$ (выражение для логарифмического правдоподобия)

→ *Estimate*

Возьмите из протокола оценивания значение *Log likelihood* (пусть это будет значение *LNL2*) и сравните его со значением *Log likelihood*, полученным при оценивании стандартной пробит-модели (пусть это будет значение *LNL1*), образовав статистику $LR = 2*(LNL2 - LNL1)$. Сравните полученное значение LR с квантилью уровня 0.95 распределения хи-квадрат с одной степенью свободы — вычисляется как @qchisq(0.95, 1).

3. Наряду со стандартной пробит-моделью рассмотрите модель с

$$G(t) = \Phi(t + \omega_1 t^2 + \omega_2 t^3).$$

Сравните максимумы логарифмов правдоподобий, полученные при оценивании этих двух моделей. Для стандартной пробит-модели соответ-

ствующая величина указывается в протоколе оценивания в строке *Log likelihood*. Для оценивания расширенной модели создайте объект *Logl*:

Object → *New Object* → *Logl*

Запишите в открывшемся окне выражение для логарифмического правдоподобия в расширенной модели:

@logl logl1

$u=c(1)+c(2)*x$

@temp u

@param c(1) -3 c(2) 0.003 c(3) 0.1 c(4) 0.1

$\log l1=y*\log(@cnorm(u+c(3)*u^2+c(4)*u^3))+(1-y)*\log(1-$
 $-@cnorm(u+c(3)*u^2+c(4)*u^3))$

→ *Estimate*

Возьмите из протокола оценивания значение *Log likelihood* (пусть это будет значение *LNL2*) и сравните его со значением *Log likelihood*, полученным при оценивании стандартной пробит-модели (пусть это будет значение *LNL1*), образовав статистику $LR = 2*(LNL2 - LNL1)$. Сравните полученное значение *LR* с квантилью уровня 0.95 распределения хи-квадрат с одной степенью свободы — вычисляется как @qchisq(0.95, 1).

Задание 33. Интерпретация протоколов оценивания моделей логит и пробит

Статистические данные представляют собой информацию о 2578 индивидах, выбранных случайным образом из граждан Дании, имевших в 1990 г. возраст от 20 до 59 лет. Информация дана по следующим переменным:

Количественные переменные	
<i>hourwage</i>	среднечасовая заработная плата (в датских кронах)
<i>age</i>	возраст
<i>education</i>	количество лет, затраченных на образование
<i>exper</i>	стаж работы (в годах)
Дамми-переменные	
<i>sex</i>	= 1 для женщин, = 0 для мужчин
<i>province</i>	= 1 для живущих в провинции, = 0 для живущих в столичном регионе
<i>public</i>	= 1 для работающих в государственном секторе, = 0 для работающих в частном секторе
<i>unem90</i>	= 1 для имевших периоды безработицы в 1990 г., = 0 для не имевших периоды безработицы в 1990 г.
<i>unem89</i>	= 1 для имевших периоды безработицы в 1989 г., = 0 для не имевших периоды безработицы в 1989 г.

- Объясните основные различия между моделями, в которых дамми-переменные входят только в качестве объясняющих переменных, и моделями, в которых объясняемая переменная является дамми-переменной (например, переменная *unem90*).
- Какие сложности возникают при рассмотрении линейных моделей регрессии с бинарной объясняемой переменной?

Согласно теории человеческого капитала размер заработной платы индивидов зависит от квалификации и знаний, которые можно приобрести как результат образования, от стажа работы и прохождения специальных курсов обучения специфической деятельности. Простая модель заработной платы, известная как модель Минсера, имеет следующий вид:

$$\ln \text{hourwage}_i = \beta_0 + \beta_1 \text{education}_i + \beta_2 \text{exper}_i + \beta_3 \text{exper}_i^2 + \varepsilon_i, \quad i = 1, \dots, N.$$

Расширенная форма этого уравнения включает дополнительные объясняющие переменные:

$$\ln \text{hourwage}_i = \beta_0 + \beta_1 \text{education}_i + \beta_2 \text{exper}_i + \beta_3 \text{exper}_i^2 + \beta_4 \text{sex}_i + \beta_5 \text{province}_i + \varepsilon_i.$$

- Как следует интерпретировать коэффициенты в $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ в последней модели?

Оцениваются логит- и пробит-модели с бинарной объясняемой переменной *unem90*, в которых в качестве объясняющих переменных выступают переменные *sex, age, province, education, exper, public, unem89*.

- Каким методом оцениваются эти модели? Дайте интерпретацию приведенных ниже результатов оценивания.

Результаты оценивания логит-модели:

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
<i>constant</i>	−1.046900	0.40957	−2.556	0.0106
<i>sex</i>	−0.218060	0.13225	−1.649	0.0993
<i>age</i>	−0.002385	0.0084483	−0.282	0.7777
<i>province</i>	0.388170	0.13191	2.943	0.0033
<i>education</i>	−0.081989	0.024224	−3.385	0.0007
<i>exper</i>	−0.042591	0.011483	−3.709	0.0002
<i>public</i>	0.154820	0.14011	1.105	0.2693
<i>unem89</i>	3.211500	0.12310	26.087	0.0000

$$\log l = -908.2405471.$$

Проверка значимости регрессии в целом: $LR = 1068.5493$.

$$\text{McFadden's } R^2 = 0.3704.$$

Результаты оценивания пробит-модели:

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение
<i>constant</i>	−0.621740	0.21867	−2.843	0.0045
<i>sex</i>	−0.123230	0.071690	−1.719	0.0857
<i>age</i>	−0.0021156	0.0045947	−0.460	0.6452
<i>province</i>	0.216000	0.071017	3.042	0.0024
<i>education</i>	−0.044325	0.013165	−3.367	0.0008
<i>exper</i>	−0.023321	0.0062433	−3.735	0.0002
<i>public</i>	0.078649	0.075868	1.037	0.3000
<i>unem89</i>	1.889600	0.068603	27.543	0.0000

$$\log l = -905.6507406.$$

Проверка значимости регрессии в целом: $LR = 1073.7290$.

$$McFadden's R^2 = 0.3722.$$

- Как можно проверить нулевую гипотезу о равенстве нулю коэффициентов при переменных *age* и *public* (одновременно)?

Высокая статистическая значимость оценок коэффициентов при переменной *unem89* указывает на наличие определенной перманентности пребывания в положении безработного.

- Как можно проверить предположение о том, что перманентный характер пребывания в состоянии безработного в большей степени выражен у женщин?

Задание 34. Оценивание порядковой пробит-модели в пакете *EViews*

Используя пакет *EViews*, сгенерируйте данные о среднедушевом месячном доходе семьи (*x*) и количестве автомобилей (*y*), имеющихся у семьи, по аналогии с примером 4.2.1 из разд. 4 лекционной части.

1. Подберите подходящую порядковую пробит-модель, объясняющую количество автомобилей в семье величиной среднедушевого месячного дохода семьи.
2. Проанализируйте прогнозные качества подобранной модели. Сравните их с качеством прогнозов по тривиальной модели, не учитывающей влияние доходов семьи.

Методические указания.

1. Используйте процесс порождения данных в виде:

$$y_i^* = x_i + \varepsilon_i, \quad i = 1, \dots, 1000,$$

где x_1, \dots, x_{1000} — независимые, одинаково распределенные случайные величины, имеющие равномерное распределение в интервале (100, 2100);

ε_i — независимые в совокупности (и независимые от x_i) случайные величины, имеющие одинаковое нормальное распределение $\varepsilon_i \sim N(0, 300^2)$, т.е. $\sigma = 300$.

Как и в примере 4.2.1, положите $K = 3$, $\gamma_1 = 1100$ и $\gamma_2 = 1850$.

2. В объекте *Equation* в качестве метода оценивания выберите *ORDERED* — *Ordered Choice*,
а в качестве распределения ошибок выберите *Normal*.
3. Рассмотрите таблицу, получаемую как *View* → *Prediction Evaluation*.

Задание 35. Оценивание порядковой пробит-модели в пакете Stata

Проведите анализ данных, сгенерированных в задании 34, используя пакет *Stata*.

Методические указания. Для построения порядковой пробит-модели воспользуйтесь командой *oprobit*:

```
oprobit y x
```

После оценивания коэффициентов модели используйте команды:

```
predict float p1 p2 p3, p (p1, p2, p3 — прогнозные значения вероятностей)
predict xb, xb (xb — прогнозные значения полезности)
```

Классификация семей по *xb*:

```
generate float number = 1 if xb < _cut1 (вместо _cut1 подставьте оцененное
значение нижнего граничного уровня, указанное в протоколе оцени-
вания)
```

```
replace number = 2 if _cut1 < xb & xb <= _cut2 (вместо _cut1 подставьте оце-
ненное значение нижнего граничного уровня, указанное в протоколе
оценивания, а вместо _cut2 — оцененное значение верхнего гранич-
ного уровня)
```

```
replace number = 3 if xb > _cut2 (вместо _cut2 подставьте оцененное значе-
ние верхнего граничного уровня)
```

Вычисляем количество семей, отнесенных прогнозом в каждую группу:

```
summarize number if number == 1
summarize number if number == 2
summarize number if number == 3
```

Вычисляем действительное количество семей в каждой группе:

```
summarize y if y == 1
summarize y if y == 2
summarize y if y == 3.
```

Задание 36. *Оценивание в пакете Stata мультиномиальной логит-модели, в которой объясняющие переменные являются собственными атрибутами субъекта*

Используя пакет *EViews*, сгенерируйте (по аналогии с примером 4.2.2 из разд. 4 лекционной части) данные о типичном для семьи количестве посещений продуктового магазина в неделю (X_2), среднемесячном доходе на одного члена семьи (X_3) и о предпочтительном выборе семьей одного из трех магазинов ($D_k = 1$, если семья отдает предпочтение k -му продуктового магазину, $k = 1, 2, 3$, $D_k = 0$ в противном случае).

1. Оцените в пакете *Stata* параметры соответствующей мультиномиальной логит-модели.
2. На основании полученных оценок коэффициентов вычислите прогнозные значения вероятностей предпочтения альтернатив для каждой семьи, дайте прогнозы предпочтительной альтернативы для каждой семьи, найдите прогнозные размеры групп семей, предпочитающих k -й магазин, $k = 1, 2, 3$. Сравните прогнозные размеры групп семей с имеющимися данными.
3. Получите количества правильных и неправильных прогнозов по каждой группе.

Методические указания.

1. Для оценивания мультиномиальной логит-модели, в которой объясняющие переменные являются собственными атрибутами субъекта, воспользуйтесь командой *mlogit*. Предварительно создайте новую переменную *number* (объясняемая переменная), равную k , если $D_k = 1$:

```
generate int number1 = 1 (int определяет целочисленную переменную)
replace number1 = 2 if d2 == 1
replace number1 = 3 if d3 == 1
mlogit number x1 x2, basecategory(1)
```

(Последняя опция выделяет базовую категорию, для которой значения всех коэффициентов в функции полезности полагаются равными 0.)

2. После оценивания коэффициентов модели используйте команду

```
predict float p1 p2 p3, p (p1, p2, p3 — прогнозные значения вероятностей)
```

На основании полученных прогнозных значений вероятностей принадлежности семей группам 1, 2 и 3 строим прогнозные значения номеров предпочитаемых альтернатив (переменная *number_forecast*):

```
generate int number_forecast = 1
replace number_forecast = 2 if p2 >= p1 & p2 >= p3
replace number_forecast = 3 if p3 >= p1 & p3 >= p2
```

Вычисляем количество семей, отнесенных прогнозом в каждую группу:

```
summarize number_forecast if number_forecast == 1
summarize number_forecast if number_forecast == 2
summarize number_forecast if number_forecast == 3
```

Вычисляем действительное количество семей в каждой группе:

```
summarize number if number == 1
summarize number if number == 2
summarize number if number == 3
```

3. Получаем количества правильных и неправильных прогнозов по каждой группе:

```
generate int d11 = 1 if number == 1 & number_forecast == 1
generate int d12 = 1 if number == 1 & number_forecast == 2
generate int d13 = 1 if number == 1 & number_forecast == 3
generate int d21 = 1 if number == 2 & number_forecast == 1
generate int d22 = 1 if number == 2 & number_forecast == 2
generate int d23 = 1 if number == 2 & number_forecast == 3
generate int d31 = 1 if number == 3 & number_forecast == 1
generate int d32 = 1 if number == 3 & number_forecast == 2
generate int d33 = 1 if number == 3 & number_forecast == 3
summarize d11
summarize d12
summarize d13
summarize d21
summarize d22
summarize d23
summarize d31
summarize d32
summarize d33
```

Задание 37. *Оценивание в пакете EViews мультиномиальной логит-модели, в которой объясняющие переменные являются собственными атрибутами субъекта*

Выполните задания, перечисленные в задании 36, используя пакет EViews.

Методические указания. Составьте соответствующую рассматриваемой ситуации программу в пакете EViews, ориентируясь на имеющуюся в пакете программу *mlogit1.prg*.

Задание 38. *Оценивание в пакете EViews мультиномиальной логит-модели, в которой объясняющие переменные могут зависеть не только от субъекта, но и от альтернативы (условная логит-модель)*

Используя пакет EViews, сгенерируйте (по аналогии с примером 4.2.3 из разд. 4 лекционной части) данные для последующего оценивания:

$stores_k$ — количество магазинов в k -м торговом центре, $k = 1, 2, 3$;
 $dist_{ik}$ — расстояние от места проживания i -й семьи до k -го торгового центра;
 $y_i = k$, если i -я семья предпочитает делать покупки в k -м торговом центре.

Здесь переменная $stores_k$ специфична только в отношении альтернатив, тогда как значения переменной $dist_{ik}$ зависят и от альтернативы, и от конкретной семьи. В примере 4.2.3 были взяты значения $stores_1 = 23$, $stores_2 = 19$, $stores_3 = 13$, значения $dist_{ik}$ моделировались как реализации независимых, одинаково распределенных случайных величин, имеющих равномерное распределение в интервале $(0, 10)$.

1. Используя метод максимального правдоподобия, оцените параметры соответствующей мультиномиальной логит-модели.
2. На основании полученных оценок коэффициентов найдите прогнозные размеры групп семей, предпочитающих k -й торговый центр, $k = 1, 2, 3$.

Методические указания. Составьте соответствующую рассматриваемой ситуации программу в пакете *EViews*, ориентируясь на имеющуюся в пакете программу *clogit1.prg*.

Задание 39. Оценивание модели тобит (тобит-I) в пакете *EViews*

Сгенерируйте значения переменных, использованных при рассмотрении модели тобит-I в теме 4.3 разд. 4 лекционной части:

x — среднедушевой месячный доход семьи (*i.i.d.* случайные величины, равномерно распределенные в интервале $(1, 1600)$);
 $price_observed$ — переменная, равная 0 для семей, не имеющих автомобиля, и равная цене, по которой был куплен автомобиль (автомобили), для семей, имеющих автомобиль.

1. Методом наименьших квадратов оцените линейную модель зависимости $price_observed$ от среднедушевого месячного дохода семьи, используя для этой цели:
 - а) все 1000 наблюдений (модель 1);
 - б) данные только по семьям, которые имеют автомобиль (модель 2).

Какой общий недостаток имеют обе построенные модели?

2. Образуйте переменную y , равную 0 для семей, не имеющих автомобиля, и равную $y = (price_observed - 2000)$ для семей, имеющих автомобиль. Методом наименьших квадратов оцените линейную модель зависимости переменной y от среднедушевого месячного дохода семьи, используя для этой цели:
 - а) все 1000 наблюдений (модель 3);
 - б) данные только по семьям, которые имеют автомобиль (модель 4).

Какой общий недостаток имеют обе построенные модели?

3. Оцените стандартную тобит-модель зависимости переменной y от среднедушевого месячного дохода семьи
 - а) по всем 1000 наблюдений (модель 5);
 - б) только по наблюдениям для семей, имеющих автомобиль (модель 6).
4. Сравните характеристики прогнозов (в пределах выборки), получаемых по моделям 3, 4, 5, 6.

Методические указания (к пунктам задания).

1б. При спецификации уравнения в окне *Sample* укажите:

if price_observed > 0.

2б. При спецификации уравнения в окне *Sample* укажите:

if y > 0.

3а. При спецификации уравнения в окне *Method* выберите:

ORDERED.

В открывшемся меню задайте спецификацию уравнения: $Y \text{ с } X$ и выберите:

Distribution: Normal.

Установите флажок для *Truncated sample*.

3б. Действуйте, как и в п. 3а, но снимите (не устанавливайте) флажок для *Truncated sample*.

4. Для моделей 3 и 4 действуйте обычным образом. Для моделей 5 и 6 в процедуре *Forecast* выберите:

Forecast of: Expected dependent variable.

Сравните средние абсолютные процентные ошибки прогнозов.

Задание 40. Оценивание модели тобит-II в пакете *EViews*

Сгенерируйте значения переменных, использованных в примере 4.3.1 раздела 4 лекционной части:

x — среднедушевой месячный доход семьи;

h — индикаторная переменная, равная 1 при наличии у семьи автомобиля;

y — переменная, равная 0 для семей, не имеющих автомобиля, и равная цене, по которой был куплен автомобиль, для семей, имеющих автомобиль.

Используя двухшаговую процедуру Хекмана, оцените параметры модели тобит-II, порождающей данные.

Методические указания. На первом шаге оцените пробит-модель бинарного выбора с объясняемой переменной h и объясняющей переменной x . Создайте объект *Equation* (пусть это будет *eq1*). В спецификации этого объекта укажите:

Estimation settings: Method: BINARY

Equation specification: h с x

Binary estimation method: Probit → OK.

После оценивания пробит-модели выберите в меню *eq1*:

Forecast → Forecast of: Index → Forecast name: x_theta_est → OK.

Образуйте новую переменную:

lambda_est = @dnorm(x_theta_est)/@cnorm(x_theta_est)

и оцените уравнение *eq2*:

Equation Specification: y с x lambda_est

Method: LS

Sample: 1 1000 if h = 1

Задание 41. Подбор логит-модели с фиксированными эффектами в пакете Stata

Используя статистические данные файла **union.dta**, описание которого приведено в Приложении, подберите логит-модель с фиксированными эффектами для членства женщин в профсоюзе, объясняющую членство в профсоюзе следующими факторами: возрастом (*age*), образованием (*grade*), проживанием вне столичных регионов (*not_smsa*), проживанием в южных штатах США (*south*). Возможное изменение влияния последней переменной на периоде наблюдения учитывается вводом дополнительной переменной $south_t = south * (year - 1970)$.

Методические указания. Используйте команду *xtlogit fe*. Если в результате оценивания оказались статистически незначимыми оценки при двух или более переменных, то для проверки гипотезы об одновременном обращении в нуль коэффициентов при этих переменных можно применить критерий Вальда. Если, например, статистически незначимы коэффициенты при переменных *var1* и *var2*, то достаточно после выполнения команды *xtlogit fe* использовать команду:

$$test\ var1 = var2 = 0.$$
Задание 42. Подбор пробит-модели со случайными эффектами в пакете Stata

1. Используя статистические данные из файла **union.dta**, описание которого приведено в Приложении, подберите пробит-модель со случайными эффектами для членства в профсоюзе, объясняющую членство в профсоюзе следующими факторами: возрастом (*age*), образованием (*grade*), проживанием вне столичных регионов (*not_smsa*), проживанием в южных штатах США (*south*). Возможное изменение влияния последней переменной на периоде наблюдения учитывается вводом дополнительной переменной $south_t = south * (year - 1970)$.
2. Является ли подобранная модель статистически значимой? Все ли оценки коэффициентов статистически значимы? Являются ли статистически значимыми случайные эффекты?
3. Проинтерпретируйте полученные результаты.

Методические указания (к пунктам задания).

1. Используйте команду *xtprobit re* (будьте готовы к тому, что она выполняется достаточно долго):

$$xtprobit\ union\ age\ grade\ not_smsa\ south\ south_t,\ re.$$

2. *P*-значение, получаемое при проверке значимости регрессии в целом, приведено в протоколе оценивания перед таблицей, а *P*-значение, получаемое при проверке значимости случайных эффектов, — после таблицы.

Задание 43. Подбор тобит-модели со случайными эффектами в пакете Stata

Используя статистические данные из файла **nlswork.dta**, описание которого приведено в Приложении, подберите тобит-модель со случайными эффектами для логарифмов заработной платы в ситуации, когда размер заработной платы

известен, только если логарифм реальной заработной платы не превышает 1.9. Возможное изменение влияния проживания в южных штатах на периоде наблюдения учитывается включением в уравнение дополнительной переменной $south_t = south * (year - 1970)$.

Методические указания. Используйте команду *xttobit* с опцией *ul(1.9)*, ограничивающей сверху фиксируемое значение логарифма часовой заработной платы величиной 1.9 (что соответствует 6.686 долл./ч).

Задание 44. Оценивание модели тобит (тобит-I) в пакете Stata

Проведите анализ данных, сгенерированных в задании 39, используя пакет Stata.

Методические указания. Для оценивания тобит-модели воспользуйтесь командой *tobit*. Предварительно создайте переменную *price_censored*, значения которой равны значениям переменной *price_observed*, если *price_observed* > 0, и равны 2000 — в противном случае.

Задание 45. Оценивание модели тобит-II

Проведите анализ данных, полученных в задании 40, используя пакет Stata.

Методические указания. Для оценивания тобит-модели II воспользуйтесь командой *heckman*.

Приложение

СТАТИСТИЧЕСКИЕ ДАННЫЕ К ЗАДАНИЯМ

Таблица П-1

Объем продаж *Y* и цена *P* товара

<i>Y</i>	<i>P</i>	<i>X</i>	<i>Z</i>	<i>W1</i>	<i>Z1</i>		<i>Y</i>	<i>P</i>	<i>X</i>	<i>Z</i>	<i>W1</i>	<i>Z1</i>
97.5	105.5	90.36711	101.8214	101	101		108.5	100.5	107.1181	101.8458	106	109
99.0	106.0	92.04135	100.3622	102	102		109.5	100.5	108.6194	105.2693	103	110
97.5	102.5	95.36486	92.32723	105	102		101.5	91.5	108.0453	88.30700	102	109
101.0	103.0	97.14663	95.13061	105	104		100.0	92.0	108.3275	88.62322	101	107
102.5	103.5	97.68223	101.1366	102	105		108.5	102.5	104.7296	108.9392	100	105
100.5	98.5	100.2795	93.70041	103	108		105.5	102.5	102.2678	104.1693	103	106
102.5	98.5	102.9282	96.34262	104	106		99.0	97.0	100.2017	92.02088	102	107
101.5	96.5	103.7268	90.63259	106	107		107.0	105.0	99.28349	108.0477	101	109
106.5	103.5	100.6989	100.3094	107	105		107.0	100.0	105.7602	100.4600	104	110
103.0	99.0	102.0795	90.93913	108	104		107.0	99.0	107.7204	98.20371	106	108
101.5	95.5	104.0013	90.13066	105	106							

Таблица II-2

Финансовые показатели 66 предприятий

Y	X1	X2	X3	X4	X5		Y	X1	X2	X3	X4	X5
0	36.7	-62.8	-89.5	54.1	1.7		1	35.2	43.0	16.4	99.1	1.3
0	24.0	3.3	-3.5	20.9	1.1		1	38.8	47.0	16.0	126.5	1.9
0	-61.6	-120.8	-103.2	24.7	2.5		1	14.0	-3.3	4.0	91.7	2.7
0	-1.0	-18.1	-28.8	36.2	1.1		1	55.1	35.0	20.8	72.3	1.9
0	18.9	-3.8	-50.6	26.4	0.9		1	59.3	46.7	12.6	724.1	0.9
0	-57.2	-61.2	-56.2	11.0	1.7		1	33.6	20.8	12.5	152.8	2.4
0	3.0	-20.3	-17.4	8.0	1.0		1	52.8	33.0	23.6	475.9	1.5
0	-5.1	-194.5	-25.8	6.5	0.5		1	45.6	26.1	10.4	287.9	2.1
0	17.9	20.8	-4.3	22.6	1.0		1	47.4	68.6	13.8	581.3	1.6
0	5.4	-106.1	-22.9	23.8	1.5		1	40.0	37.3	33.4	228.8	3.5
0	23.0	-39.4	-35.7	69.1	1.2		1	69.0	59.0	23.1	406.0	5.5
0	-67.6	-164.1	-17.7	8.7	1.3		1	34.2	49.6	23.8	126.6	1.9
0	-185.1	-308.9	-65.8	35.7	0.8		1	47.0	12.5	7.0	53.4	1.8
0	13.5	7.2	-22.6	96.1	2.0		1	15.4	37.3	34.1	570.1	1.5
0	5.7	-118.3	-34.2	21.7	1.5		1	56.9	35.3	4.2	240.3	0.9
0	72.4	-185.9	-280.0	12.5	6.7		1	43.8	49.5	25.1	115.0	2.6
0	17.0	-34.6	-19.4	35.5	3.4		1	20.7	18.1	13.5	63.1	4.0
0	-31.2	-27.9	6.3	7.0	1.3		1	33.8	31.4	15.7	144.8	1.9
0	14.1	-48.2	6.8	16.6	1.6		1	35.3	21.5	-14.4	90.0	1.0
0	-60.6	-49.2	-17.2	7.2	0.3		1	24.4	8.5	5.8	149.1	1.5
0	26.2	-19.2	-36.7	90.4	0.8		1	48.9	40.6	5.8	82.0	1.8
0	7.0	-18.1	-6.5	16.5	0.9		1	49.9	34.6	26.4	310.0	1.8
0	-53.1	-98.0	-20.8	26.6	1.7		1	54.8	19.9	26.7	239.9	2.3
0	-17.2	-129.0	-14.2	267.9	1.3		1	39.0	17.4	12.6	60.5	1.3
0	32.7	-4.0	-15.8	177.4	2.1		1	53.0	54.7	14.6	771.7	1.7
0	26.7	-8.7	-36.3	32.5	2.8		1	20.1	53.5	20.6	307.5	1.1
0	-7.7	-59.2	-12.8	21.3	2.1		1	53.7	35.9	26.4	289.5	2.0
0	18.0	-13.1	-17.6	14.6	0.9		1	46.1	39.4	30.5	700.0	1.9
0	2.0	-38.0	1.6	7.7	1.2		1	48.3	53.1	7.1	164.4	1.9
0	-35.3	-57.9	0.7	13.7	0.8		1	46.7	39.8	13.8	229.1	1.2
0	5.1	-8.8	-9.1	100.9	0.9		1	60.3	59.5	7.0	226.6	2.0
0	0.0	-64.7	-4.0	0.7	0.1		1	17.9	16.3	20.4	105.6	1.0
0	25.2	-11.4	4.8	7.0	0.9		1	24.7	21.7	-7.8	118.6	1.6

ФАЙЛЫ В ФОРМАТЕ STATA

Файл **grunfeld.dta** на сайте <http://www.statapress.com/data/r8/xt.html> содержит данные о размере инвестиций (*invest*), рыночной цене (*mvalue*) и акционерном капитале (*kstock*) 10 крупных корпораций США за 20-летний период — с 1935 по 1954 г.

Кодировка корпораций в файле:

- 1 — *GM (GENERAL MOTORS)*
- 2 — *USSTEEL (AMERICAN STEEL)*
- 3 — *GE (GENERAL ELECTRIC)*
- 4 — *CHRYSLER*
- 5 — *ARCO (ATLANTIC RICHFIELD)*
- 6 — *IBM*
- 7 — *UNION OIL*
- 8 — *WESTINGHOUSE*
- 9 — *GOODYEAR*
- 10 — *DIAMOND MATCH*

Файл **males.dta** на сайте <http://www.econ.kuleuven.be/gme/data.htm> содержит данные, использованные в статье (Vella, Verbeek, 1998) и представляющие результаты статистического обследования (*National Longitudinal Survey, Youth Sample*, США) о 545 полностью занятых мужчинах, окончивших школу до 1980 г. и наблюдавшихся в течение 8-летнего периода — с 1980 по 1987 г. В 1980 г. эти мужчины были в возрасте от 17 до 23 лет. Панель сбалансирована: по всем субъектам количество наблюдений равно 8. Исследование проводилось с целью выявления зависимости размера заработной платы мужчин (переменная *wage*) от различных факторов. В файле 44 переменные:

- NR* — идентификатор субъекта,
- YEAR* — год наблюдения,
- SCHOOL* — количество лет обучения,
- EXPER* — количество лет пребывания на рынке труда,
- $EXPER2 = EXPER^2$,
- HOURS* — количество проработанных часов за год,
- WAGE* — натуральный логарифм часовой заработной платы в долларах США.

Бинарные (0 — 1) дамми-переменные:

- BLACK* = 1 для чернокожих,
- HISP* = 1 для испаноязычных,
- HLTH* = 1 при наличии проблем со здоровьем,
- MAR* = 1 для женатых,
- NC* = 1 для проживающих на центральном севере США,
- NE* = 1 для проживающих на северо-востоке США,
- RUR* = 1 для проживающих в сельской местности,

$S = 1$ для проживающих на юге США,
 $UNION = 1$ для членов профсоюза,
 $IYEAR_2 = 1$ для 1981 г., $IYEAR_3 = 1$ для 1982 г., $IYEAR_4 = 1$ для 1983 г.,
 $IYEAR_5 = 1$ для 1984 г., $IYEAR_6 = 1$ для 1985 г., $IYEAR_7 = 1$ для 1986 г.,
 $IYEAR_8 = 1$ для 1987 г.

Кроме того, имеется еще 12 дамми-переменных, характеризующих область деятельности (AG — сельское хозяйство, MIN — добывающая промышленность, CON — строительство, $TRAD$ — торговля, TRA — транспорт, FIN — финансы, BUS — сфера обслуживания и ремонта, PER — обслуживающий персонал, ENT — сфера развлечений, MAN — обрабатывающая промышленность, PRO — услуги профессионального характера, PUB — государственное управление), и 9 дамми-переменных ($OCC1$ — $OCC9$), характеризующих профессию (характер выполняемой работы).

Файл **nlswork.dta** (<http://www.stata-press.com/data/r10/nlswork.dta>) содержит результаты статистического обследования (*National Longitudinal Survey, Youth Sample*, США), в котором участвовали молодые женщины, имевшие в 1968 г. возраст от 14 до 26 лет. Наблюдения проводились с 1968 по 1988 г. В файле содержатся данные по 5159 субъектам обследования. Панель не сбалансирована: по отдельным субъектам количество наблюдений изменялось от 1 до 12 (в среднем 4.6 наблюдения для одного субъекта исследования). Исследование проводилось с целью выявления зависимости размера заработной платы женщин (переменная w_{it}) от различных факторов. В файле задействованы следующие переменные (для некоторых субъектов отсутствуют данные по одной или нескольким переменным):

$idcode$ — идентификатор субъекта обследования,
 $year$ — год интервьюирования,
 $birth_yr$ — год рождения,
 age — возраст,
 $race$ — переменная, равная 1 для белого населения, 2 — для чернокожего населения, 3 — для остального населения,
 msp — дамми-переменная, равная 1, если женщина состоит в браке,
 nev_mar — дамми-переменная, равная 1, если женщина никогда не состояла в браке,
 $grade$ — законченное количество лет учебы,
 $collgrade$ — дамми-переменная, равная 1, если женщина окончила колледж,
 $nots_msa$ — дамми-переменная, равная 1, если проживает вне столичных регионов,
 c_city — дамми-переменная, равная 1, если проживает в центральном городе,
 $south$ — дамми-переменная, равная 1, если проживает на юге страны,
 ind_code — область занятости,
 occ_code — код профессии,

<i>union</i>	— дамми-переменная, равная 1, если женщина состоит в профсоюзе,
<i>wks_ue</i>	— количество недель без работы за последний год,
<i>ttl_exp</i>	— суммарный стаж работы,
<i>tenure</i>	— продолжительность (стаж) работы на наблюдаемом рабочем месте,
<i>hours</i>	— обычное количество рабочих часов,
<i>wks_work</i>	— количество проработанных недель за последний год,
<i>ln_wage</i>	— логарифм реальной заработной платы (долларов в час).

Файл **union.dta** (<http://www.stata-press.com/data/r10/union.dta>) построен на базе файла **nlswork** и содержит данные только по 4434 субъектам обследования, которые дали ответ на вопрос о членстве в профсоюзах. В файл не включены некоторые переменные, содержащиеся в файле **nlswork**.

Литература

1. *Верби́к М.* (2008). Путеводитель по современной эконометрии. М.: Научная книга.
2. *Носко В.П.* (2004). Эконометрика. Элементарные методы и введение в регрессионный анализ временных рядов. М.: ИЭПП.
3. *Носко В.П.* (2005). Эконометрика для начинающих (Дополнительные главы). М.: ИЭПП.
4. *Amemiya T.* (1985). *Advanced Econometrics*. Oxford: Blackwell.
5. *Arellano M., Bond S.* (1991). Some Tests of Specification for Panel Data: Monte-Carlo Evidence and an Application to Employment Equations // *Review of Economic Studies*. Vol. 58. P. 277—294.
6. *Chamberlain G.* (1980). Analysis of Covariance with Qualitative Data // *Review of Economic Studies*. Vol. XLVII (1). No. 146. P. 225—238.
7. *Davidson R., MacKinnon J.G.* (1993). *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.
8. *Dwivedi T., Srivastava K.* (1978). Optimality of Least Squares in Seemingly Unrelated Regression Model // *Journal of Econometrics*. Vol. 7. P. 391—395.
9. *Godfrey L.G., Hutton J.* (1994). Discriminating between errors-in-variables / simultaneity and misspecification in linear regression models // *Economic Letters*. Vol. 44. P. 359—364.
10. *Greene W. H.* (2003). *Econometric analysis*. 5th edition. Upper Saddle River, NJ. Prentice Hall.
11. *Hausman J.A.* (1978). Specification tests in econometrics // *Econometrica*. Vol. 46. P. 1251—1271.
12. *Hosmer D.W., Lemeshow S.* (1989). *Applied Logistic Regression*. New York: Wiley.
13. *Hsiao C.* (2003). *Analysis of Panel Data*. Cambridge: Cambridge University Press.
14. *Johansen S., Juselius K.* (1994). Identification of the Long-run and Short-run Structure. An Application to the IS/LM Model // *Journal of Econometrics*. Vol. 63. P. 7—36.
15. *Linzert T., Nautz D., Breitung J.* (2003). Bidder Behavior in Repo Auctions without Minimum Bid Rate: Evidence from the Bundesbank. Deutsche Bundesbank, Research Centre, Discussion Paper Series 1: Economic Studies: 13.
16. *Patterson K.* (2000). *An Introduction to Applied Econometrics: A Time Series Approach*. New York: St's Martin Press.
17. *Schmidt P.* (1976). *Econometrics*. New York: Marcel Dekker.
18. *Spencer D., Berk K.* (1981). A Limited Information Specification Test // *Econometrica*. 49. P. 1079—1085.

19. *Vella F., Verbeek M.* (1998). Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men // *Journal of Applied Econometrics*. Vol. 13. P. 163—183.
20. *Verbeek M.* (2000). *A Guide to Modern Econometrics*. Chichester: Wiley.
21. *Winkelmann L., Winkelmann R.* (1995). Unemployment: where does it hurt? Centre for Economic Policy Research Discussion Paper, 1093.

Интернет-ресурсы

<http://www.feweb.vu.nl/econometriclinks/>

На этом сайте имеется большое количество ссылок на различные интернет-ресурсы, посвященные изучению, разработке и применению эконометрических методов.

<http://www.stanford.edu/~clint/bench/>

На этом сайте выложено большое количество статистических данных, которые можно использовать для практического освоения эконометрических методов анализа моделей одновременных уравнений, панельных данных, моделей с дискретной объясняющей переменной (*Probit*, *Logit* и др.).

Глоссарий

К разделу 1

Двухшаговый метод наименьших квадратов (*2SLS — two-step least squares*) — метод получения оценок параметров i -го уравнения структурной формы, содержащего в правой части эндогенные переменные. На первом шаге производится очистка эндогенных переменных, включенных в правую часть i -го уравнения структурной формы, что должно обеспечивать получение переменных, не коррелированных с ошибкой в i -м уравнении. На втором шаге производится оценивание уравнения, полученного заменой эндогенных переменных в правой части i -го уравнения их очищенными вариантами.

Идентифицируемость системы структурных уравнений (*identifiability of a system of simultaneous equations*) **в целом** (на основании приведенной формы системы) — идентифицируемость всех коэффициентов системы и идентифицируемость ковариационной матрицы случайных ошибок, входящих в правые части уравнений системы. При этом при восстановлении коэффициентов и ковариационной матрицы ошибок в структурной форме используются не только коэффициенты приведенной формы, но и ковариационная матрица ошибок в приведенной форме.

Исключающие ограничения (*exclusion restrictions*) — линейные ограничения, накладываемые на коэффициенты i -го структурного уравнения (помимо условия нормировки), состоящие в исключении из i -го уравнения некоторых эндогенных и предопределенных переменных.

Концентрированная логарифмическая функция правдоподобия (*concentrated log-likelihood function*) — функция, полученная исключением из логарифмической функции правдоподобия слагаемых, не зависящих от неизвестных параметров.

Косвенный метод наименьших квадратов (*ILS — indirect least squares*) — метод получения оценок параметров i -го стохастического уравнения структурной формы через оценки наименьших квадратов коэффициентов уравнений приведенной формы. Метод применим в случае точной идентифицируемости i -го структурного уравнения.

Критерий Дарбина — Ву — Хаусмана (*Durbin — Wu — Hausman test*) — одна из версий критерия Хаусмана.

Пусть $y_i = Y_i\alpha_i + X_i\theta_i + u_i$, X_i — матрица значений инструментальных переменных, Y_i — матрица значений тех объясняющих переменных в i -м структурном уравнении, которые не входят в состав инструментальных переменных и не являются линейными комбинациями последних.

Проверяемая гипотеза H_0 : в i -м уравнении отсутствует проблема эндогенности, т.е. все объясняющие переменные в составе Y_i не коррелированы с u_i . Если эта гипотеза выполнена, то оценивание i -го уравнения можно производить обычным методом наименьших квадратов (*OLS*). В противном случае надо применять метод инструментальных переменных.

Критерий Хаусмана (*Hausman test*) — критерий, используемый для проверки правильности спецификации системы одновременных уравнений в отношении подразделения включенных в систему переменных на эндогенные и экзогенные и основанный на свойствах трехшаговой оценки наименьших квадратов (*3SLS*) и двухшаговой оценки наименьших квадратов (*2SLS*) вектора коэффициентов. Если все структурные уравнения специфицированы правильно, то *3SLS* состоятельна и эффективна; если же хотя бы одно из уравнений специфицировано неправильно, то *3SLS* перестает быть состоятельной оценкой. Соответственно статистика критерия основана на расхождении *3SLS*- и *2SLS*-оценок.

Кронекеровское произведение (*Kronecker product*) — результат алгебраической операции, известной как тензорное умножение матрицы на матрицу.

Недоидентифицированность (*underidentification*) **структурной формы системы** — отсутствие возможности восстановления всех коэффициентов структурной формы системы по коэффициентам приведенной формы.

Неидентифицируемое (недоопределенное — *underidentified*) уравнение — отсутствие возможности восстановления всех коэффициентов отдельного структурного уравнения по коэффициентам приведенной формы системы одновременных уравнений.

Нормированное уравнение (*normalized equation*) — структурное уравнение, в котором коэффициент при одной из эндогенных переменных, входящих в уравнение, равен 1.

Ограничения неисключающего типа (*non-excluding conditions*) — линейные ограничения, накладываемые на коэффициенты i -го структурного уравнения (помимо условия нормировки), не сводящиеся к исключению из i -го уравнения некоторых эндогенных и predetermined переменных.

Оценка максимального правдоподобия с ограниченной информацией (*LIML — limited information maximum likelihood*) — оценка максимального правдоподобия параметров смешанной системы

$$\begin{cases} y_1 = Y_1\alpha_1 + X_1\theta_1 + u_1, \\ Y_1 = X\Pi_1 + W_1, \end{cases}$$

состоящей из структурного уравнения для одной из эндогенных переменных и приведенной формы для остальных эндогенных переменных. Используется в ситуациях, в которых одно из стохастических структурных уравнений, например,

$$y_1 = Y_1\alpha_1 + X_1\theta_1 + u_1,$$

идентифицируемо, а остальные уравнения либо неидентифицируемы либо имеются сомнения в правильности их спецификации. При этом известны список всех предопределенных переменных, включаемых в систему, и значения этих предопределенных переменных в n наблюдениях.

Оценка максимального правдоподобия с полной информацией (*FIML — full information maximum likelihood*) — оценка максимального правдоподобия вектора параметров структурной формы системы одновременных уравнений, учитывающая полную информацию о структуре модели одновременных уравнений.

Порядковое условие идентифицируемости (*order condition for identification*) — необходимое условие идентифицируемости i -го структурного уравнения: количество предопределенных переменных в системе, не включенных в i -е структурное уравнение, должно быть не меньше количества эндогенных переменных, включенных в i -е уравнение, уменьшенного на единицу. Если в левой части i -го структурного уравнения находится единственная эндогенная переменная, то $(g - g_i^*) - 1$ есть просто количество эндогенных переменных, включенных в правую часть этого уравнения.

Предопределенная переменная (*predetermined variable*) — переменная, входящая в правую часть i -го структурного уравнения, значение которой в момент t не коррелировано с ошибками в этом уравнении, соответствующими моментам $t, t + 1, \dots$

Приведенная форма одновременных уравнений (*reduced form of simultaneous equations*) — получаемая на основе структурной формы одновременных уравнений совокупность уравнений для каждой эндогенной переменной, в правых частях которых не содержатся эндогенные переменные.

Ранговое условие идентифицируемости (*rank condition for identification*) — необходимое и достаточное условие того, что неспецифицированные коэффициенты i -го структурного уравнения системы одновременных уравнений

$$y_t\Gamma = x_tB + u_t, \quad t = 1, \dots, n,$$

с g эндогенными и K предопределенными переменными определяются по матрице коэффициентов приведенной формы (после применения правила нормировки) однозначным образом.

Это условие имеет вид: $\text{rank}(\Phi_i A) = g - 1$, где $A = \begin{bmatrix} \Gamma \\ B \end{bmatrix}$, Φ_i — матрица

размера $R_i \times (g + K)$, определяющая R_i однородных линейных ограничений $\Phi_i \alpha_i = 0$, которые (в дополнение к условию нормировки) накладыва-

ются на элементы i -го столбца матрицы A (представляющие коэффициенты при эндогенных и экзогенных переменных в i -м структурном уравнении системы).

Рекурсивные системы (*recursive systems*) — системы одновременных уравнений, в которых эндогенные переменные последовательно связаны друг с другом, а ошибки в различных уравнениях не коррелированы между собой. Для оценивания коэффициентов уравнений в этой ситуации можно использовать обычный метод наименьших квадратов.

Сверхидентифицируемое (переопределенное — *overidentified*) уравнение — структурное уравнение, для которого выполнено ранговое условие идентифицируемости и для восстановления коэффициентов которого имеется количество соотношений, большее минимально необходимого.

Если $\hat{\Pi}$ — оценка матрицы коэффициентов приведенной формы $\Pi = B\Gamma^{-1}$, полученная свободным оцениванием (без учета ограничений на коэффициенты структурных уравнений системы), то в такой ситуации существует несколько вариантов восстановления по матрице $\hat{\Pi}$ коэффициентов i -го структурного уравнения, приводящих к различным результатам.

Структурная форма одновременных уравнений (*structural form of simultaneous equations*) — форма записи системы одновременных уравнений, представляющая в явном виде взаимные связи между входящими в модель переменными, показывающая, как эти переменные взаимодействуют друг с другом.

Точно идентифицируемое (*just identified*) уравнение — структурное уравнение, для которого выполнено ранговое условие идентифицируемости и для восстановления коэффициентов которого имеется количество соотношений, равное минимально необходимому.

Если $\hat{\Pi}$ — оценка матрицы коэффициентов приведенной формы $\Pi = B\Gamma^{-1}$, полученная свободным оцениванием (без учета ограничений на коэффициенты структурных уравнений системы), то в такой ситуации коэффициенты i -го структурного уравнения восстанавливаются по матрице $\hat{\Pi}$ однозначным образом.

Трехшаговая оценка наименьших квадратов, 3SLS-оценка (*three-stage least squares*) — доступная обобщенная оценка наименьших квадратов коэффициентов структурных уравнений системы, принимающая во внимание наличие коррелированности между ошибками в разных структурных уравнениях.

Частично идентифицируемая система (*partially identified system*) — система одновременных уравнений, в которой идентифицируемы не все структурные уравнения.

Экзогенные переменные (*exogenous variables*) — переменные, значения которых определяются вне системы одновременных уравнений, которые могут выступать в уравнениях системы только как объясняющие переменные. Экзогенная переменная, входящая в i -е структурное уравнение системы, не коррелирована с ошибкой в этом уравнении.

Порядковое условие (*order condition*) **идентифицируемости r коинтегрирующих векторов** — необходимое условие идентифицируемости r коинтегрирующих векторов. Если векторы уже нормализованы, то необходимым условием идентифицируемости r коинтегрирующих векторов является наложение на каждый из r векторов не менее $(r - 1)$ линейных ограничений.

При выполнении этого порядкового условия идентифицируемости полученные r векторов могут все же оказаться линейно зависимыми, так что, например, вектор β_1 нельзя отличить от некоторой линейной комбинации векторов β_1, \dots, β_r . Поэтому, в принципе, следует производить еще и проверку линейной независимости полученных r векторов.

Приведенная форма ЕСМ (*reduced form error correction model*) — модель коррекции ошибок, у которой в правые части уравнений для приращений отдельных переменных входят только запаздывающие значения приращений других переменных, получаемая на основе структурной ЕСМ.

Структурная ЕСМ (*SECM — structural error correction model*) — модель ЕСМ, у которой в правые части уравнений для приращений отдельных переменных входят не только запаздывающие, но и текущие значения приращений других переменных.

Структурная VAR (*SVAR — structural vector autoregression*) — модель VAR, у которой в правые части уравнений для отдельных переменных входят не только запаздывающие, но и текущие значения других переменных.

Явная форма линейных ограничений (*direct restrictions*) — представление r_i линейных ограничений на i -й коинтегрирующий вектор β_i в виде: $\beta_i = H_i \mathcal{Q}_i$, где H_i — матрица размера $N \times (N - r_i)$ ранга $(N - r_i)$; \mathcal{Q}_i — вектор размера $(N - r_i) \times 1$.

К разделу 3

Взвешенный метод наименьших квадратов (*WLS — weighted least squares*) — вариант метода наименьших квадратов, в котором наблюдениям могут придаваться различные веса.

«Внутригрупповая» оценка, «within»-оценка, «внутри»-оценка (*«within-group» estimate*) — оценка коэффициентов при объясняющих переменных в модели с фиксированными эффектами

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

полученная в рамках модели для отклонений от индивидуальных средних:

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Внутригрупповые остатки (*within residuals*) — остатки, получаемые в результате оценивания модели в отклонениях от индивидуальных средних:

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Временные эффекты (*time effects*) — аддитивные составляющие λ_t в уравнениях двунаправленных моделей с фиксированными или случайными индивидуальными эффектами, одинаковые для различных субъектов, но изменяющиеся от одного момента времени к другому. Например,

$$y_{it} = \mu + \alpha_i + \lambda_t + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где $\sum_{i=1}^N \alpha_i = 0$, $\sum_{t=1}^T \lambda_t = 0$, так что α_i и λ_t — дифференциальные эффекты.

Двунаправленная внутригрупповая оценка (*two-way within estimator*) — оценка вектора коэффициентов двунаправленной модели с фиксированными индивидуальными и временными эффектами.

Дифференциальные эффекты (*differential effects*) — эффекты α_i в модели со случайными эффектами, выраженные в отклонениях от среднего:

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it},$$

где $\sum_{i=1}^N \alpha_i = 0$.

Компоненты дисперсии (*variance components*) — составляющие дисперсии случайных составляющих $v_i = \alpha_i + u_{it}$ в стандартной модели со случайными эффектами

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it} = \mu + \beta x_{it} + (\alpha_i + u_{it}) = \mu + \beta x_{it} + v_{it}.$$

Критерий Бройша — Пагана (*Breusch — Pagan test*) — статистический критерий для проверки гипотезы об отсутствии корреляционной связи между ошибками в совпадающие моменты времени в разных уравнениях регрессии для панельных данных. Статистика критерия равна: $\lambda = T \sum_{i=2}^N \sum_{j=1}^{i-1} r_{ij}^2$, где

$$r_{ij} = \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii} \hat{\sigma}_{jj}}} \text{ — оцененная корреляция между ошибками в } i\text{-м и } j\text{-м уравнениях.}$$

Критерий Бройша — Пагана в двунаправленной модели (*Breusch-Pagan test in a two-way model*) — критерий для проверки в рамках двунаправленной модели с фиксированными индивидуальными и временными эффектами гипотезы об отсутствии индивидуальных и временных эффектов, т.е. возможности сведения такой модели к модели пула.

Критерий Бройша — Пагана в RE-модели (*Breusch — Pagan test in RE-model*) — критерий для проверки в рамках RE-модели (со стандартными предположениями)

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it}$$

гипотезы об отсутствии индивидуальных эффектов, $H_0: \sigma_\alpha^2 = 0$, т.е. возможности сведения такой модели к модели пула.

Критерий Хаусмана (*Hausman test*) — статистический критерий для проверки гипотезы о выполнении предположений стандартной *RE*-модели, основанный на величине расхождения между *GLS*-оценкой $\hat{\beta}_{GLS}$ в стандартной *RE*-модели и ковариационной оценкой $\hat{\beta}_{CV}$ для модели с фиксированными эффектами.

Если выполняются предположения стандартной *RE*-модели, то $\hat{\beta}_{GLS}$ является эффективной оценкой, а $\hat{\beta}_{CV}$ — неэффективной оценкой.

«Межгрупповая» оценка, «между»-оценка (*between estimator*) — оценка наименьших квадратов, соответствующая регрессии средних (по времени) значений объясняемой переменной на константу и средние (по времени) значения объясняющих переменных («модель для групповых средних»): $\bar{y}_i = \mu + \beta \bar{x}_i + \bar{v}_i$. Эта оценка игнорирует внутригрупповую изменчивость.

Модель кажущихся несвязанными регрессий¹ (*SUR, SURE — seemingly unrelated regressions*) — совокупность линейных уравнений регрессии для панельных данных, в которой имеется корреляционная связь между ошибками в разных уравнениях в совпадающие моменты времени.

Модель ковариационного анализа (*covariance analysis model*) — совокупность линейных уравнений регрессии для панельных данных, в которой коэффициенты могут быть различными в уравнениях для разных субъектов, но не изменяются во времени:

$$y_{it} = x_{it}^T \theta_i + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Модель компонент дисперсии (*variance components model*) — то же, что и стандартная модель со случайными эффектами.

Модель Мундлака (*Mundlak model*) — модель компонент ошибки, отличающаяся от стандартной модели со случайными эффектами тем, что в правую часть *i*-го уравнения добавляется наблюдаемая переменная \bar{x}_i , изменяющаяся только от субъекта к субъекту и отражающая неоднородность субъектов. Для модели с одной объясняющей переменной модель Мундлака имеет вид:

$$y_{it} = \mu + \beta x_{it} + a \bar{x}_i + \varepsilon_i + u_{it}.$$

Модель пула (пул) (*pool model*) — совокупность линейных уравнений регрессии для панельных данных, в которой коэффициенты при объясняющих переменных одинаковы для разных субъектов и для разных моментов времени:

$$y_{it} = x_{it}^T \theta + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Модель со случайными эффектами (*random effects model*) — совокупность линейных уравнений для *N* субъектов

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

¹ В некоторых руководствах по эконометрике такую модель называют *системой внешне не связанных между собой уравнений*.

в которых коэффициенты при объясняющих переменных не изменяются от субъекта к субъекту, а дифференциальные эффекты α_i , $\sum_{i=1}^N \alpha_i = 0$, рассматриваются как случайная выборка из некоторого распределения.

Модель с фиксированными эффектами (*FE — fixed effects model, LSDV — least squares dummy variables*) — совокупность линейных уравнений для N субъектов, в которых коэффициенты при объясняющих переменных не изменяются от субъекта к субъекту, а постоянные составляющие α_i различны:

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

Несбалансированная панель (*unbalanced panel*) — разное количество наблюдений для разных субъектов.

Не связанные между собой (*unrelated*) **линейные регрессии** — совокупность линейных уравнений регрессии для панельных данных, в которой отсутствует статистическая зависимость между ошибками в разных уравнениях.

Обобщенный метод моментов (*GMM — generalized method of moments*) — метод оценивания, состоящий в минимизации некоторой квадратичной формы от выборочных аналогов моментных условий (условий ортогональности).

Оптимальная взвешивающая матрица (*optimal weighting matrix*) — взвешивающая матрица минимизируемой квадратичной формы, приводящая к наиболее эффективной *GMM*-оценке.

Оценка Хаусмана — Тейлора (*Hausman-Taylor estimator*) — оценка вектора коэффициентов *RE*-модели, в которой ненаблюдаемые индивидуальные эффекты могут быть коррелированными с одними объясняющими переменными и не коррелированными с другими объясняющими переменными.

Панельные данные (*panel data*) — данные о значениях переменных для N субъектов (индивидов, фирм, стран, регионов и т.п.) в T последовательных моментов (периодов) времени.

Сбалансированная панель (*balanced panel*) — в каждый из T моментов времени имеются данные обо всех N субъектах, участвующих в анализе.

Стандартная модель со случайными эффектами (*RE-модель — random effects model*) — модель со случайными эффектами

$$y_{it} = \mu + \alpha_i + \beta x_{it} + u_{it},$$

для которой выполнены следующие предположения:

$$u_{it} \sim i.i.d. N(0, \sigma_u^2), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

$$E(x_{it} u_{js}) = 0 \quad \text{для любых } i, j = 1, \dots, N, \quad t, s = 1, \dots, T,$$

$$\alpha_i \sim i.i.d. N(0, \sigma_\alpha^2), \quad i = 1, \dots, N,$$

$$E(x_{it} \alpha_j) = E(u_{it} \alpha_j) = 0, \quad i, j = 1, \dots, N, \quad t = 1, \dots, T.$$

Статистика Саргана (*Sargan statistic*) — статистика критерия для проверки гипотезы о том, что избыточные моментные условия (выведенные на основании исходных предположений о рассматриваемой модели) действительно выполняются.

Условия ортогональности, моментные условия (*orthogonality conditions, moment conditions*) — условия отсутствия коррелированности случайных ошибок в уравнениях регрессии с переменными, которые планируется использовать в качестве инструментальных переменных. Использование в обобщенном методе моментов большего количества условий ортогональности позволяет получить большее количество инструментов и приводит к повышению эффективности оценок (используется большее количество информации).

Фиксированные эффекты (*fixed effects*) — постоянные составляющие в уравнениях линейных регрессий, различные для разных субъектов.

IV-«внутри»-оценка (*IV within estimators*) — оценка инструментальных переменных для коэффициентов внутригрупповой регрессии.

IV-«между»-оценка (*IV between estimator*) — оценка инструментальных переменных для коэффициентов межгрупповой регрессии.

«R²-внутри» (*R²-within*) — коэффициент детерминации, получаемый при оценивании методом наименьших квадратов модели линейной регрессии для отклонений от групповых средних:

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i), \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

«R²-между» (*R²-between*) — коэффициент детерминации, получаемый при оценивании модели линейной регрессии для групповых средних:

$$\bar{y}_i = \mu + \beta \bar{x}_i + \bar{v}_i, \quad i = 1, \dots, N.$$

«R²-полный» (*R²-overall*) — коэффициент детерминации, получаемый при оценивании модели пула:

$$y_{it} = \alpha + \beta x_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

RE-оценка (*random effect estimator*) — GLS-оценка вектора коэффициентов в стандартной модели со случайными эффектами. Является взвешенным средним внутригрупповой и межгрупповой оценок.

К разделу 4

Гнездовые модели (*nested models*) — две модели, одна из которых вложена в другую, является частным случаем другой модели.

Гомпит-модель (*gompit model*) — модель бинарного выбора

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

в которой в качестве функции связи $G(z)$ используется функция стандартного распределения экстремальных значений (минимума) I типа (распределение Гомпертца): $G(z) = 1 - \exp(-e^z)$.

Двухшаговая процедура Хекмана (*Heckman's two-stage procedure*) — процедура оценивания параметров модели тобит-II.

Достаточное (*sufficient*) **множество статистик** — при заданных значениях этих статистик вклад i -го субъекта в правдоподобие не зависит от эффекта α_i . Наличие такого множества статистик дает возможность использовать оценку условного максимального правдоподобия при оценивании модели бинарного выбора для панельных данных.

Индекс отношения правдоподобий (*LRI — likelihood ratio index*) — один из аналогов коэффициента детерминации для моделей бинарного выбора:

$$LRI = 1 - \frac{\ln L_1}{\ln L_0}, \text{ где } L_1 \text{ — максимум функции правдоподобия для выбранной модели, а } L_0 \text{ — максимум функции правдоподобия для тривиальной модели (в которой в качестве единственной объясняющей переменной выступает константа).}$$

Индикаторная (дихотомическая, бинарная) переменная (*indicator, dichotomic, binary variable*) — переменная, значение которой равно 1 при наличии у субъекта рассматриваемого признака и равно 0 — в противном случае.

Критерий Хосмера — Лемешоу (*Hosmer-Lemeshow test*) — статистический критерий, используемый для проверки адекватности подобранной модели бинарного выбора имеющимся данным и основанный на сравнении предсказываемых моделью и действительно наблюдаемых количеств случаев со значением индикаторной переменной, равным 1, в нескольких группах, на которые разбивается множество наблюдений.

Латентная (скрытая) переменная (*latent variable*) — ненаблюдаемая переменная, участвующая в процессе порождения данных: индикаторная объясняемая переменная принимает для некоторого субъекта значение 1, если значение латентной переменной для этого субъекта превышает «пороговый» уровень.

Логит (*logit*) — логарифм шансов. Если $p = P(A)$ — вероятность события A , то соответствующий ей логит равен:

$$\text{logit}(p) = \ln \left(\frac{p}{1-p} \right).$$

Если $\text{logit}(p) = 0$, то $p = 1 - p = 0.5$, т.е. шансы для события A равны «50 на 50». Если $\text{logit}(p) > 0$, то больше шансов, что событие A произойдет. Если $\text{logit}(p) < 0$, то больше шансов, что событие A не произойдет.

Логит-модель (*logit model*) — модель бинарного выбора

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

в которой в качестве функции связи $G(z)$ используется функция стандартного логистического распределения $\Lambda(z) = \frac{e^z}{1 + e^z}$.

Логит-модель с фиксированными эффектами (*fixed-effects logit model*) — модель бинарного выбора для панельных данных, в которой латентное уравнение имеет вид: $y_{it}^* = x_{it}^T \beta + \alpha_i + \varepsilon_{it}$, где α_i — фиксированные эффекты; ε_{it} имеют логистическое распределение с функцией распределения

$$\Lambda(z) = \frac{e^z}{1 + e^z}.$$

Лямбда Хекмана (*Heckman's lambda*) — величина, пропорциональная смещению оценки детерминированной составляющей линейной модели, возникающему при игнорировании уравнения выбора.

Модель бинарного выбора (*binary choice model*) — нелинейная регрессионная модель зависимости значений индикаторной переменной y_i от линейной комбинации объясняющих переменных:

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n.$$

Мультиномиальная логит-модель (*multinomial logit model*) — модель с K альтернативами, для которых не существует естественного упорядочения. Субъект выбирает альтернативу k , если для него эта альтернатива имеет максимальную полезность. Объясняющие переменные специфичны только в отношении самих субъектов исследования, значения коэффициентов в K уравнениях полезности различны:

$$P\{y_i = k\} = \frac{\exp(x_i^T (\beta^k - \beta^1))}{1 + \exp(x_i^T (\beta^2 - \beta^1)) + \dots + \exp(x_i^T (\beta^k - \beta^1))}.$$

Оценка условного максимального правдоподобия (*conditional maximum likelihood estimator*) — оценка, используемая при оценивании моделей бинарного выбора для панельных данных, в которых значения латентной функции порождаются моделью $y_{it}^* = x_{it}^T \beta + \alpha_i + \varepsilon_{it}$. Оценка получается максимизацией функции правдоподобия, условной относительно некоторого множества статистик t_i , которое достаточно для эффекта α_i .

Пороговое значение (*threshold value*) — значение, при превышении которого латентной переменной объясняемая индикаторная переменная принимает значение 1.

Порядковая пробит-модель (*ordered probit model*) — модель с K различными исходами, в основе которой лежит ненаблюдаемая переменная, представляющая сумму линейной комбинации объясняющих переменных и случайной ошибки, имеющей нормальное распределение.

Предельный эффект переменной (*marginal effect*) — в модели бинарного выбора

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

предельный эффект непрерывной переменной x_{ik} определяется как производная

$$\frac{\partial P\{y_i = 1 | x_i\}}{\partial x_{ik}} = \frac{\partial G(x_i^T \theta)}{\partial x_{ik}}.$$

В случае когда объясняющая переменная принимает только два значения — 0 и 1 (дамми-переменная), указывающие на наличие (1) или отсутствие (0) у субъекта определенного признака, предельный эффект определяют как разность $P\{y_i = 1 | x_i^*, d_i = 1\} - P\{y_i = 1 | x_i^*, d_i = 0\}$, где d_i — значение рассматриваемой дамми-переменной; x_i^* — вектор значений остальных объясняющих переменных.

Пробит-модель (*probit model*) — модель бинарного выбора

$$y_i = G(\theta_1 x_{i1} + \dots + \theta_p x_{ip}) + \varepsilon_i = G(x_i^T \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

в которой в качестве функции связи $G(z)$ используется функция стандартного нормального распределения.

Пробит-модель со случайными эффектами (*random-effects probit model*) — модель бинарного выбора для панельных данных, в которой латентное уравнение имеет вид: $y_{it}^* = x_{it}^T \beta + \alpha_i + \varepsilon_{it}$, где α_i — случайные эффекты; ε_{it} имеют нормальное распределение.

Стандартная тобит-модель, стандартная тобит-I модель (*standard Tobit I model*) — цензурированная линейная модель, в которой ошибки являются независимыми в совокупности (и независимыми от объясняющих переменных) случайными величинами, имеющими одинаковое нормальное распределение.

Стандартная тобит-II модель (*standard Tobit II model*) — модель, в которой значения некоторой переменной y_i^* порождаются линейной моделью

$$y_i^* = x_{1i}^T \theta_1 + \varepsilon_{1i}, \quad i = 1, \dots, n,$$

но наблюдаются лишь для тех наблюдений, для которых переменная выбора, порождаемая линейной моделью

$$h_i^* = x_{2i}^T \theta_2 + \varepsilon_{2i}, \quad i = 1, \dots, n,$$

превышает нулевой уровень.

Объясняющие переменные в уравнениях для y_i^* и h_i^* могут быть как одинаковыми, так и различными. В ряде ситуаций экономическая аргументация указывает на необходимость включения в правую часть уравнения для h_i^* всех переменных, включенных в правую часть уравнения для y_i^* . При этом коэффициенты при одной и той же переменной в уравнениях для y_i^* и h_i^* могут быть различными.

Усеченная модель регрессии (*truncated regression*) — модель, использующая только те наблюдения, для которых значения переменной, порождаемой цензурированной линейной моделью, известны.

Условная логит-модель (*conditional logit model*) — модель с K альтернативами, для которых не существует естественного упорядочения. Субъект выбирает альтернативу k , если для него эта альтернатива имеет максимальную полезность. Объясняющие переменные специфичны в отношении альтер-

натив, значения коэффициентов во всех K уравнениях полезности постоянны:

$$P\{y_i = k\} = \frac{\exp(x_{ik}^T \beta)}{\exp(x_{i1}^T \beta) + \dots + \exp(x_{iK}^T \beta)}.$$

Цензурированная линейная модель (*censored linear model*) — модель, в которой значения некоторой переменной порождаются линейной моделью, но наблюдаются лишь для части наблюдений, а для остальных наблюдений известно только, что в них указанная переменная выше (ниже) некоторого уровня.

Цензурированные данные (*censored data*) — данные, в которых значения некоторой переменной наблюдаются лишь для части наблюдений, а для остальных наблюдений известно только, что в них указанная переменная выше (ниже) некоторого уровня.

Шансы (*odds*) — шансы некоторого события A — отношение вероятности этого события к вероятности противоположного события: $Odds(A) = \frac{P(A)}{1 - P(A)}$.

Например, если $P(A) = p = 2/3$, то $\frac{p}{1-p} = \frac{2/3}{1-2/3} = 2$, и шансы за то, что

событие A произойдет, против того, что это событие не произойдет, равны $2 : 1$ (два к одному, или в 2 раза выше).

ЧАСТЬ 4

ВРЕМЕННЫЕ РЯДЫ: ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ. МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ

Р а з д е л 5

СГЛАЖИВАНИЕ И ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ

Тема 5.1

АДАПТИВНЫЕ МЕТОДЫ, МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Во второй части учебника уже говорилось о том, что под временным рядом (*time series*) в экономике понимается ряд значений некоторой переменной, измеренных в последовательные моменты времени. Для многих рядов измерения производятся через равные промежутки времени (годовые, квартальные, недельные, дневные данные). Если принять длину такого промежутка за единицу времени (год, квартал, неделя, день), то можно считать, что последовательные наблюдения x_1, \dots, x_T переменной x произведены в моменты $t = 1, \dots, T$. Для некоторых экономических и финансовых показателей производить измерения через равные промежутки времени не удастся. Например, значения биржевых индексов на момент закрытия фиксируются только в те дни, когда биржа работает. В последнем случае наблюдения x_1, \dots, x_T соответствуют T последовательным рабочим дням биржи.

В динамике уровней ряда могут наблюдаться «систематические эффекты»:

- некоторая **основная тенденция** (тренд — *trend*);
- **циклические колебания** (*cyclicity*);
- **сезонная изменчивость** (*seasonality*).

После учета этих систематических эффектов остается некоторая флуктуация, которая носит нерегулярный характер (*irregularity*). В связи с этим уровень ряда x_t часто представляется в виде:

- суммы тренда, циклической, сезонной и нерегулярной компонент (**аддитивная модель**):

$$x_t = T_t + C_t + S_t + I_t;$$

- произведения этих компонент (**мультипликативная модель**):

$$x_t = T_t C_t S_t I_t.$$

В рамках такого подхода исследование временного ряда состоит в выявлении и придании количественного выражения каждой из указанных компонент. При этом конечной целью исследования может являться использование полученных результатов для прогнозирования будущих значений ряда.

Нерегулярная составляющая ряда может быть уменьшена или удалена путем применения различных методов сглаживания ряда, так что при этом более отчетливо выявляются трендовые, сезонные и циклические составляющие ряда. В рамках темы 5.1 не делается никаких предположений о возможной вероятностной модели порождения ряда и рассматриваются методы сглаживания, не требующие явного описания такой модели. При изложении этих методов предполагается, что наблюдается некоторый временной ряд x_1, x_2, \dots, x_T (T — количество наблюдений). Для реализации рассматриваемых ниже методов используется пакет программ эконометрического анализа *Econometric Views* (версия 6). Приведены результаты применения этих методов к различным экономическим временным рядам.

Скользящие средние

Простое скользящее среднее (*SMA* — *simple moving average*) с периодом усреднения K (*K-point moving average*) формируется путем усреднения K последовательных значений ряда x_1, x_2, \dots, x_T . Если в качестве K берется нечетное число, $K = 2m + 1$, то значение x_i исходного ряда заменяется значением

$$sma_i = \frac{x_{i-m} + x_{i-m+1} + \dots + x_{i+m}}{K}.$$

Разумеется, значения sma_i определяются только для $i = m + 1, \dots, T - m$. В частном случае $K = 3$ имеем:

$$sma_i = \frac{x_{i-1} + x_i + x_{i+1}}{3}, \quad i = 2, \dots, T - 1;$$

если $K = 5$, то

$$sma_i = \frac{x_{i-2} + x_{i-1} + x_i + x_{i+1} + x_{i+2}}{5}, \quad i = 3, \dots, T - 2.$$

ПРИМЕР 5.1.1

Рассмотрим динамику значений индекса ММББ (Московской межбанковской валютной биржи) на периоде с 18 сентября 2007 г. по 14 февраля 2008 г. (ряд *IND* — рис. 5.1).

Сгладим этот ряд, используя трехпериодное скользящее среднее. Результат сглаживания приведен на рис. 5.2. Для наглядности на рис. 5.3 приведены графики исходного (*IND*) и сглаженного (*SMA_3*) рядов.

На рис. 5.4 приведен график разности $DELTA_3 = IND - SMA_3$, динамика этой разности не имеет систематического тренда.

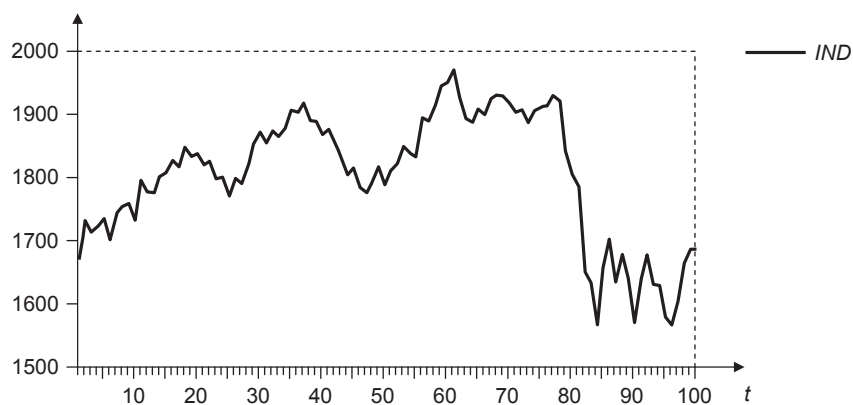


Рис. 5.1



Рис. 5.2

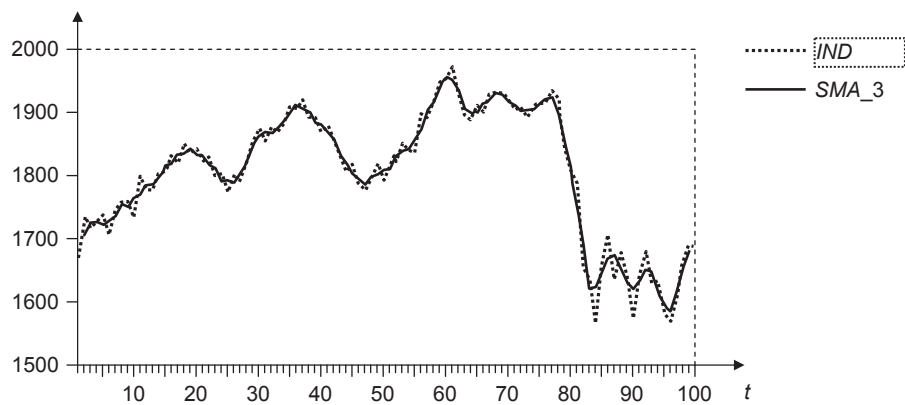


Рис. 5.3

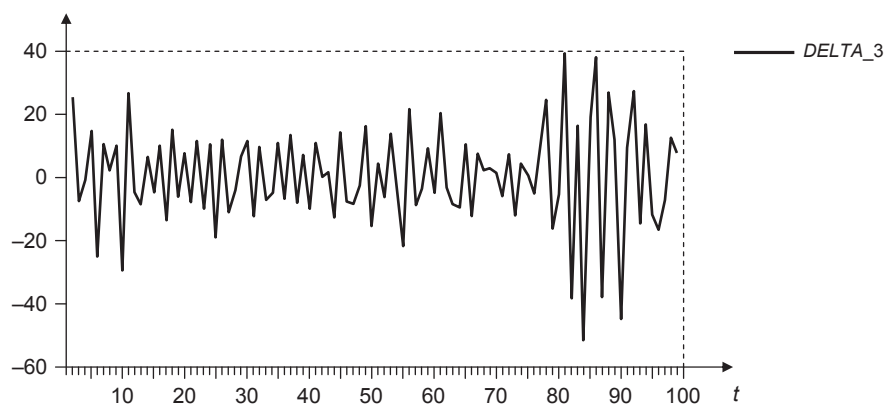


Рис. 5.4

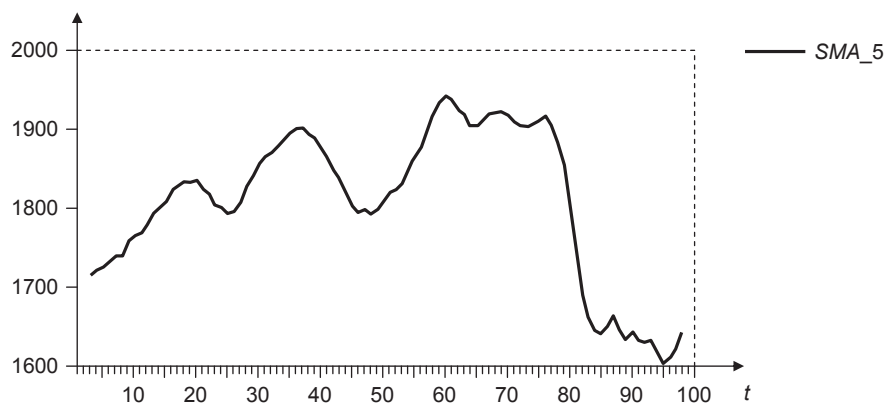


Рис. 5.5

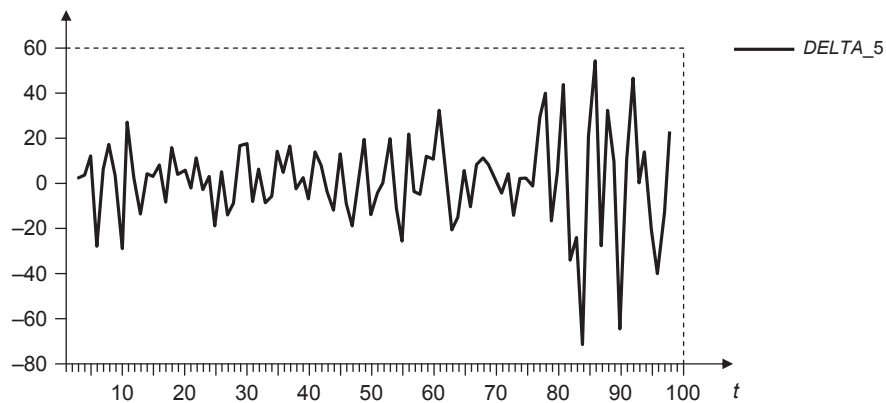


Рис. 5.6

Используем теперь 5-периодное скользящее среднее. График сглаженного ряда приведен на рис. 5.5, а график разности $DELTA_5 = IND - SMA_5$ — на рис. 5.6. ■

При рассмотрении месячных и квартальных данных в динамике рядов часто проявляются сезонные колебания, в таких случаях естественно использовать простые скользящие средние с периодами усреднения $K = 12$ и $K = 4$, соответственно. Но возникает следующий вопрос. Когда используется усреднение по нечетному количеству точек, например, по $K = 3$ точкам, то усредненное значение $\frac{x_{i-1} + x_i + x_{i+1}}{3}$ сопоставляется i -му периоду времени, т.е. центральному из трех наблюдений. Если же взять, например, $K = 4$, какому периоду времени следует сопоставить усредненное значение $\frac{x_{i-1} + x_i + x_{i+1} + x_{i+2}}{4}$?

По аналогии со случаем нечетного K это должен быть период $i + 0.5$, и это порождало бы значения $sma_{2.5}, sma_{3.5}, \dots, sma_{T-1.5}$. Чтобы обойти это затруднение, производят дополнительное усреднение полученного ряда простых скользящих средних, используя малое четное значение периода усреднения K (например, $K = 2$). В последнем случае получаем в итоге последовательность значений $sma_3^*, sma_4^*, \dots, sma_{T-2}^*$, где

$$sma_i^* = \frac{sma_{i-0.5} + sma_{i+0.5}}{2},$$

которую называют **центрированным скользящим средним** (*centered moving average*). Заметим, что при этом

$$\begin{aligned} sma_i^* &= \frac{1}{2} \left(\frac{x_{i-2} + x_{i-1} + x_i + x_{i+1}}{4} + \frac{x_{i-1} + x_i + x_{i+1} + x_{i+2}}{4} \right) = \\ &= \frac{1}{4} (0.5x_{i-2} + (x_{i-1} + x_i + x_{i+1}) + 0.5x_{i+2}). \end{aligned}$$

ПРИМЕР 5.1.2

Рассмотрим временной ряд *AGRO*, представляющий квартальные данные об индексе реального объема сельскохозяйственного производства (1993:01 = 100) за 12-летний период — с I квартала 1994 г. по IV квартал 2005 г. Динамика этого ряда имеет ярко выраженный сезонный характер (рис. 5.7).

Сгладим этот ряд, используя центрированное среднее с периодом 4. Получим сглаженный ряд *SMA_4_CENTERED* (рис. 5.8), не определенный для первых двух и для последних двух наблюдений исходного ряда.

На графике разности $DELTA = AGRO - SMA_4_CENTERED$ (рис. 5.9) видно, что динамика этой разности не имеет систематического тренда, который взял на себя сглаженный ряд. ■

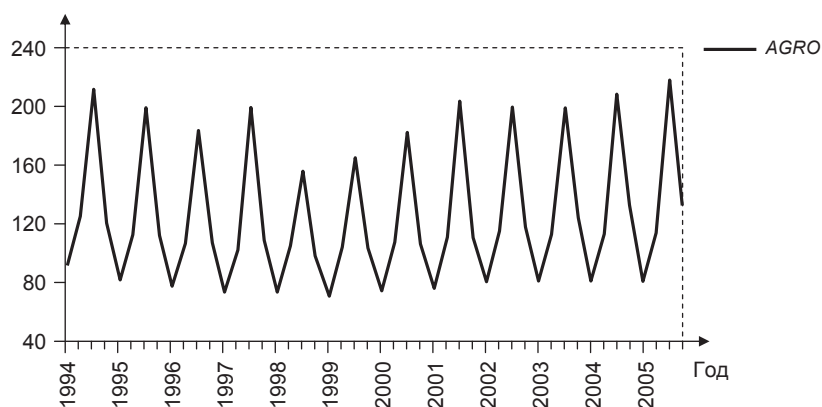


Рис. 5.7

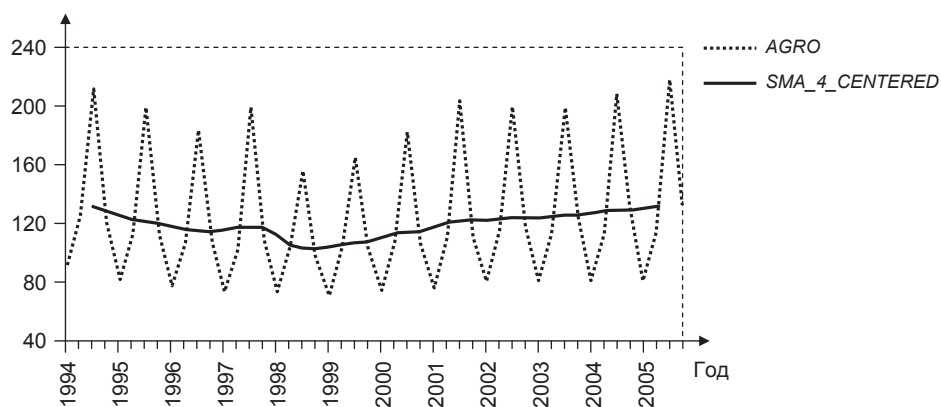


Рис. 5.8

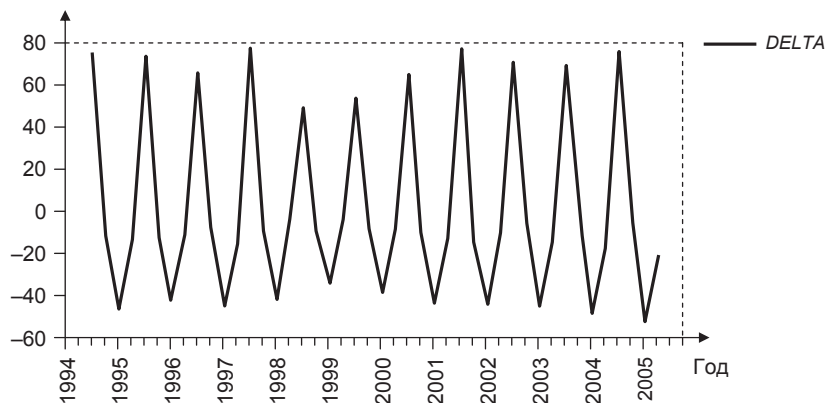


Рис. 5.9

ПРИМЕР 5.1.3

Рассмотрим временной ряд *IMP*, представляющий квартальные данные об импорте товаров и услуг в Российскую Федерацию (в млрд долл.) за 12-летний период — с I квартала 1994 г. по IV квартал 2005 г. Динамика этого ряда показана на рис 5.10.

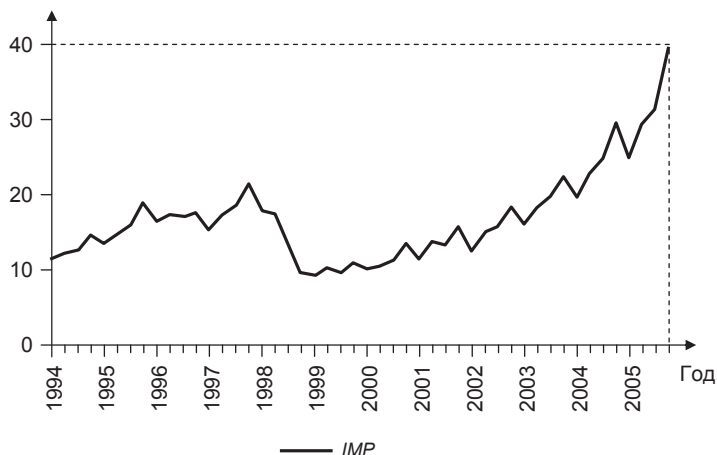


Рис. 5.10

Сгладим этот ряд, опять используя центрированное среднее с периодом 4. Для наглядности поместим исходный и сглаженный ряды на одном графике (рис. 5.11).

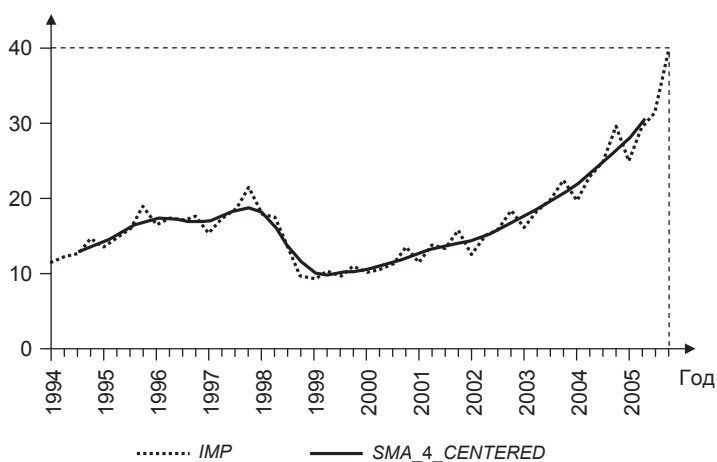


Рис. 5.11

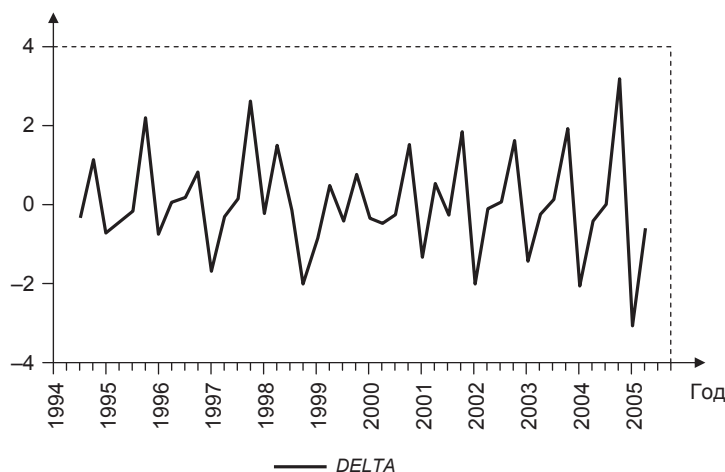


Рис. 5.12

Рассмотрим график разности $DELTA = IMP - SMA_4_CENTERED$ (рис. 5.12). Динамика этой разности не имеет систематического тренда. ■

Фильтр Ходрика — Прескотта

Этот инструмент используется для получения сглаженной оценки долговременного тренда ряда x_t . Соответствующий алгоритм подбирает сглаженный ряд s_t , для которого минимизируется сумма

$$\sum_{t=1}^T (x_t - s_t)^2 + \lambda \sum_{t=1}^T ((s_{t+1} - s_t) - (s_t - s_{t-1}))^2.$$

Параметр $\lambda > 0$ управляет гладкостью получаемого ряда. Чем больше λ , тем более гладким является ряд s_t . При $\lambda \rightarrow \infty$ ряд s_t приближается к линейному тренду. В пакете *EViews* по умолчанию используются следующие значения параметра λ :

- $\lambda = 100$ для годовых данных;
- $\lambda = 1600$ для квартальных данных;
- $\lambda = 14\,400$ для месячных данных.

ПРИМЕР 5.1.4

Применяя фильтр Ходрика — Прескотта с $\lambda = 1600$ к ряду *AGRO*, получаем ряд *HPTREND*, интерпретируемый как «долговременный тренд». График этого ряда вместе с графиком ряда *AGRO* представлен на рис. 5.13. ■

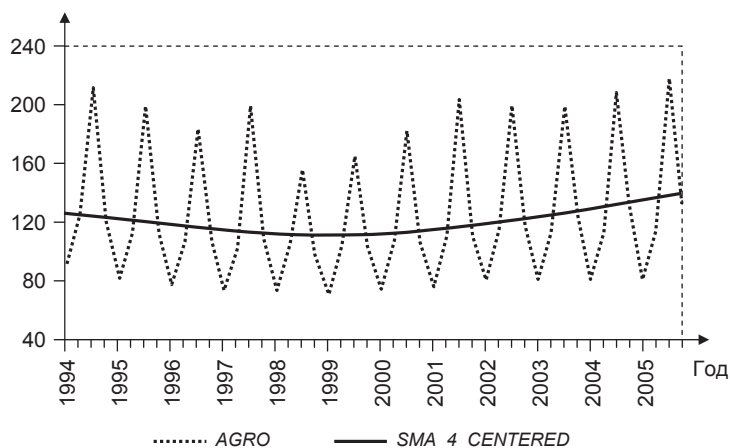


Рис. 5.13

ПРИМЕР 5.1.5

Применяя фильтр Ходрика — Прескотта с $\lambda = 1600$ к ряду *IMP*, выделяем из этого ряда «долговременный тренд» *HPTREND*. Графики рядов *IMP* и *HPTREND* представлены на рис. 5.14. ■

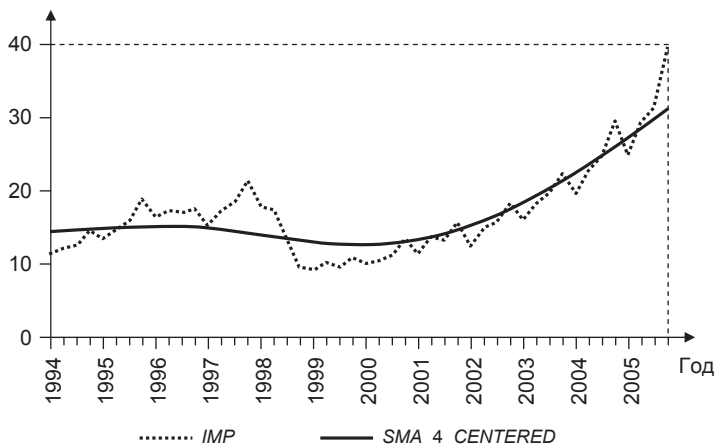


Рис. 5.14

ПРИМЕР 5.1.6

Применим фильтр Ходрика — Прескотта к ряду *IND*. Здесь данные дневные, с пропусками на выходные дни. Попробуем использовать различные значения параметра сглаживания: 100 (как для годовых данных), 1600 (как для квартальных данных), 14 400 (как для месячных данных) и 100 000. Получим результаты, отраженные на графиках, приведенных на рис. 5.15 (а, б, в, г). ■

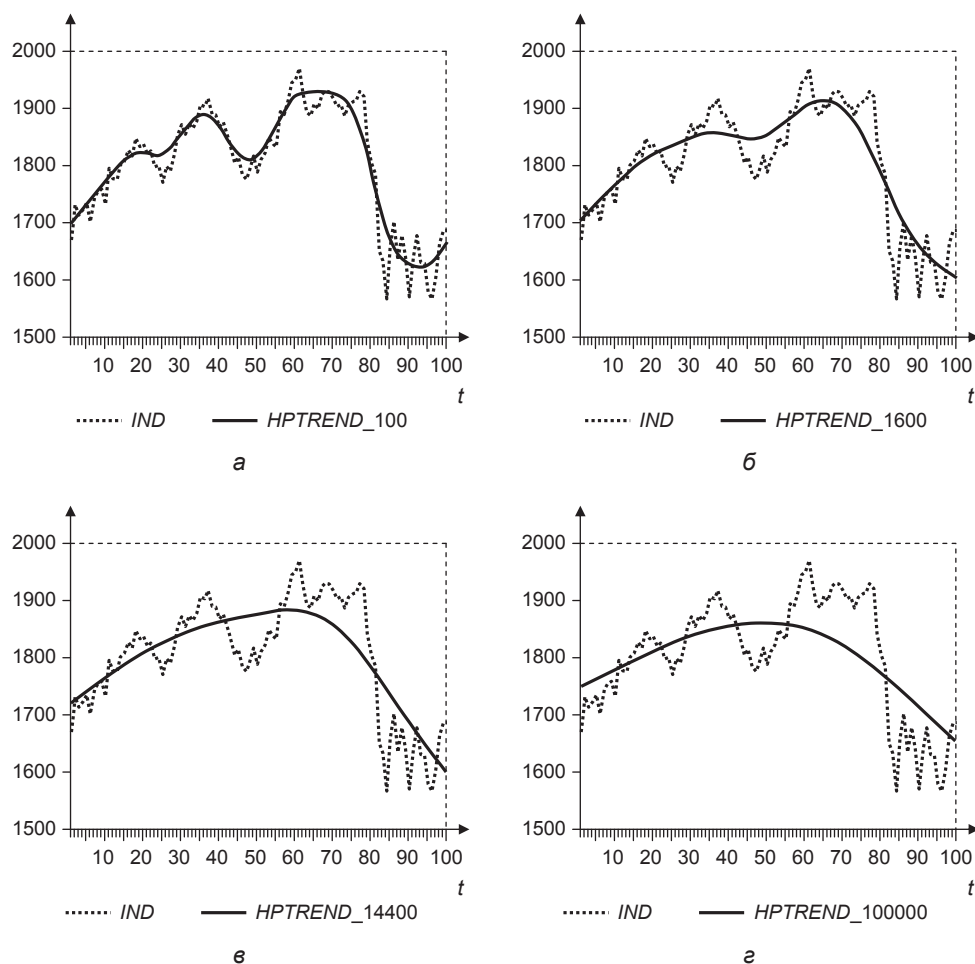


Рис. 5.15

Сглаживание ряда и прогнозирование будущих значений ряда

Прогнозирование будущих значений ряда с использованием простого скользящего среднего опирается на понятие **локального «текущего уровня»** (*current level*), вокруг которого происходят случайные (нерегулярные) флуктуации ряда. Этот уровень может изменяться с течением времени, и для периода i ($i = K, K + 1, \dots, T$) он определяется как

$$L_i = \frac{x_{i-K+1} + x_{i-K+2} + \dots + x_i}{K}$$

(т.е. среднее арифметическое K последних значений ряда).

Прогнозные значения ряда на любое количество h шагов вперед приравняются последнему текущему уровню L_T , который можно получить на основании имеющихся наблюдений x_1, x_2, \dots, x_T :

$$\hat{x}_{T+h} \equiv L_T = \frac{x_{T-K+1} + x_{T-K+2} + \dots + x_T}{K}.$$

Такой прогноз не принесет большой пользы, если ряд содержит сезонную компоненту и/или имеет выраженный возрастающий или убывающий тренд.

Взвешенное скользящее среднее (*WMA* — *Weighted Moving Average*) — метод сглаживания, в котором наблюдения усредняются с различными весами, так что наибольшие веса приписываются последним наблюдениям.

При использовании усреднения ряда x_1, x_2, \dots, x_T по трем периодам ($K = 3$)

$$wma_i = w_1 x_{i-1} + w_2 x_i + w_3 x_{i+1},$$

где w_1, w_2, w_3 — веса, приписываемые усредняемым наблюдениям: $w_3 \geq w_2 \geq w_1 \geq 0$ и $w_1 + w_2 + w_3 = 1$.

Например, можно положить

$$w_1 = 1/6, \quad w_2 = 2/6, \quad w_3 = 3/6,$$

и при таком выборе весов

$$wma_i = \frac{x_{i-1} + 2w_2 x_i + 3w_3 x_{i+1}}{6}.$$

При усреднении ряда x_1, x_2, \dots, x_T по K периодам прогнозы на h шагов вперед приравняются последнему «текущему уровню» L_T , вычисляемому по формуле:

$$L_T = w_1 x_{T-K+1} + w_2 x_{T-K+2} + \dots + w_K x_T,$$

где $w_K \geq w_{K-1} \geq \dots \geq w_1 \geq 0$, $w_1 + w_2 + \dots + w_K = 1$.

Под **экспоненциальным сглаживанием** (*Exponential Smoothing*) понимают методику, при которой наблюдения усредняются с разными весами, и при этом значения весов, приписываемых прошлым наблюдениям, убывают экспоненциальным образом по мере «старения» наблюдений.

Простое экспоненциальное сглаживание (*SES* — *Single Exponential Smoothing*) — метод, подходящий для рядов, случайным образом флуктуирующих вокруг постоянного среднего (уровня), в которых не проявляются тренды и сезонные колебания. В этом методе «текущий уровень» ряда в i -м периоде определяется рекуррентным образом:

$$L_i = \alpha x_i + (1 - \alpha)L_{i-1},$$

где α — некоторая постоянная, $0 < \alpha < 1$, так что «текущий уровень» ряда в i -м периоде есть взвешенное среднее текущего значения ряда и «текущего уровня» ряда в $(i - 1)$ -м периоде.

Из указанной формулы получаем последовательно:

$$\begin{aligned} L_i &= \alpha x_i + (1-\alpha)L_{i-1} = \alpha x_i + (1-\alpha)(\alpha x_{i-1} + (1-\alpha)L_{i-2}) = \\ &= \alpha x_i + \alpha(1-\alpha)x_{i-1} + (1-\alpha)^2 L_{i-2} = \dots = \\ &= \alpha x_i + \alpha(1-\alpha)x_{i-1} + \alpha(1-\alpha)^2 x_{i-2} + \dots + \alpha(1-\alpha)^{i-1} x_1 + (1-\alpha)^i L_0, \end{aligned}$$

т.е. в формировании значения L_i участвуют все значения x_i, x_{i-1}, \dots, x_1 , но в соответствующую сумму они входят с экспоненциально убывающими коэффициентами.

Как и при использовании скользящего среднего, прогнозные значения на h шагов вперед приравняются последнему «текущему уровню» L_T , так что прогнозное значение ряда на период $(T+h)$, получаемое на основании наблюдений x_1, x_2, \dots, x_T , определяется по формуле:

$$\hat{x}_{T+h} = L_T \quad \text{для любого } h = 1, 2, \dots$$

Заметим еще, что

$$\hat{x}_{T+1} = \hat{x}_T + \alpha(x_T - \hat{x}_T),$$

где \hat{x}_T — прогнозное значение ряда на период T , полученное на основании наблюдений x_1, x_2, \dots, x_{T-1} .

Иными словами, прогнозное значение \hat{x}_{T+1} на период $T+1$ отличается от прогнозного значения \hat{x}_T , полученного на период T , на величину, пропорциональную ошибке прогноза на период T , так что этот механизм обеспечивает адаптивный характер процедуры.

Параметр сглаживания α определяет роль последнего наблюдения в формировании «текущего уровня» ряда. Если значение α близко к 1, то влияние на «текущий уровень» всех предыдущих наблюдений оказывается незначительным. При малых значениях α «текущий уровень» определяется в большей степени прошлой динамикой ряда.

Многие специалисты рекомендуют использовать значения α в пределах 0.1—0.5, при этом аналитики большинства фирм при обработке рядов используют свои традиционные значения α . Для динамично развивающихся фирм и рынков характерны более высокие значения α , чем для более консервативных компаний и стабильных рынков; для прогнозов используют более высокие значения α , чем для анализа предшествующих тенденций.

Для реального построения прогнозных значений необходимо определить также значение L_0 . Часто в качестве него берется само x_1 .

Параметр сглаживания можно подобрать по имеющимся данным, выбрав значение, минимизирующее сумму квадратов ошибок одношаговых прогнозов, получаемых последовательным добавлением наблюдений.

ПРИМЕР 5.1.7

Применим процедуру простого экспоненциального сглаживания к ряду $x_t = IND_t$, не задавая заранее значение параметра α , а предоставляя программе самой выбрать оптимальное значение этого параметра (результаты приведены в табл. 5.1). Это приводит к значению $\alpha = 0.914$.

Таблица 5.1

Простое экспоненциальное сглаживание ряда

Sample: 1 100; Included observations: 100;

Method: Single Exponential; Original Series: IND; Forecast Series: INDSM01

Parameters:	Alpha	0.914
	Sum of Squared Residuals	127940.1
	Root Mean Squared Error	35.76871
End of Period Levels:	Mean	1687.725

В табл. 5.1:

- *Sum of Squared Residuals* — сумма квадратов расхождений между значениями исходного и сглаженного рядов:

$$SSE = \sum_{i=1}^T (x_i - s_i)^2;$$

- *Root Mean Squared Error (RMSE)* — квадратный корень из суммы квадратов расхождений, деленной на число наблюдений:

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i - s_i)^2};$$

- *Mean* — «текущий уровень» в конце периода наблюдений:

$$Mean = L_T.$$

На рис. 5.16 приведены графики исходного и сглаженного ряда.

Похоже, что график сглаженного ряда получается сдвигом вправо графика исходного ряда. Чтобы проверить это, совместим график сглаженного ряда INDSM01 и сдвинутый вправо на один период времени график исходного ряда, т.е. график ряда $IND(-1)$.

Получим практически одинаковые графики (рис. 5.17), и это означает, что прогнозные значения \tilde{x}_{t+1} практически равны текущим значениям x_t . Почему так, можно понять, построив вероятностную модель порождения ряда, что мы и сделаем в дальнейшем.

Вместе с тем посмотрим, что изменит не критический выбор одного из рекомендуемых значений 0.1—0.5, а именно $\alpha = 0.2$. В этом случае получим достаточно сглаженный ряд INDSM02 (рис. 5.18), который, однако,

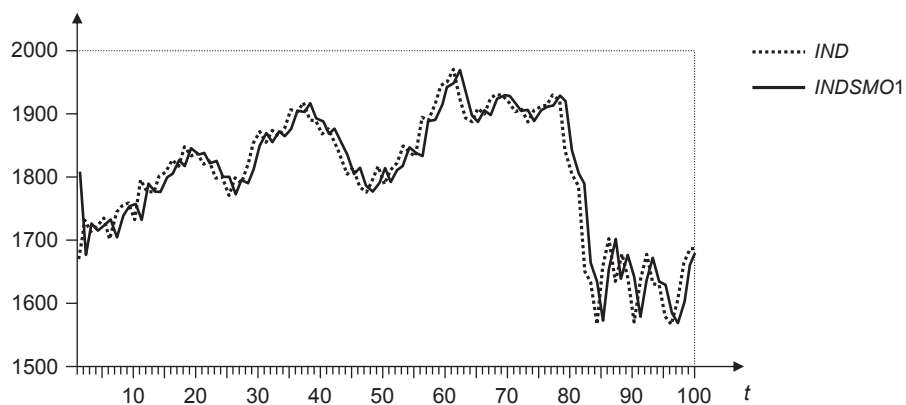


Рис. 5.16

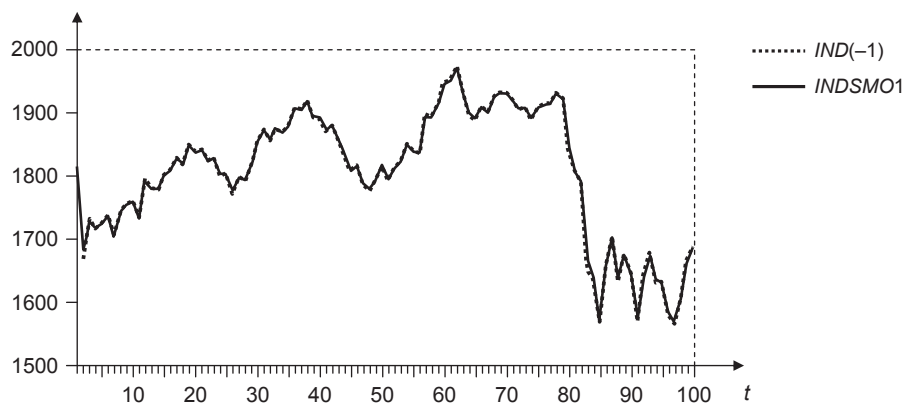


Рис. 5.17

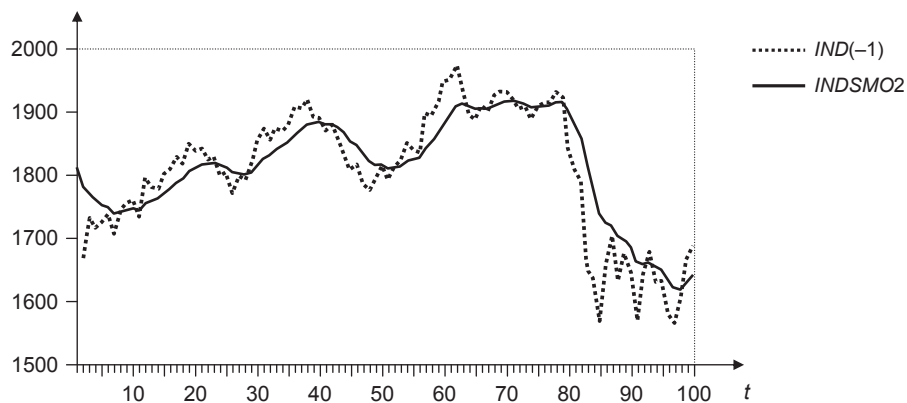


Рис. 5.18

Таблица 5.2

Простое сглаживание ряда с некритическим выбором значения α

Sample: 1 100; Included observations: 100;

Method: Single Exponential; Original Series: IND; Forecast Series: INDSM02

Parameters:	Alpha	0.2
	Sum of Squared Residuals	313564.0
	Root Mean Squared Error	55.99679
End of Period Levels:	Mean	1650.558

дает весьма плохие прогнозы (табл. 5.2): среднеквадратическая ошибка ($RMSE = 55.99679$) намного больше, чем при оцененном $\alpha = 0.914$, где $RMSE = 35.76871$. ■

Двойное экспоненциальное сглаживание, метод Брауна (*DES — Double Exponential Smoothing*) — метод, подходящий для рядов с выраженным линейным трендом, наличие которого учитывается при прогнозировании значений ряда на несколько периодов вперед.

Само название метода говорит о том, что процедура экспоненциального сглаживания здесь применяется дважды (с одним и тем же значением α): сначала — к исходному ряду, а затем — к сглаженному ряду:

$$\tilde{x}_i = \alpha x_i + (1 - \alpha)\tilde{x}_{i-1}, \quad \tilde{\tilde{x}}_i = \alpha \tilde{x}_i + (1 - \alpha)\tilde{\tilde{x}}_{i-1}.$$

При этом значение \hat{x}_{T+h} прогноза на h периодов времени вперед вычисляется по формуле:

$$\hat{x}_{T+h} = \left(2 + \frac{\alpha h}{1 - \alpha}\right) \tilde{x}_T - \left(1 + \frac{\alpha h}{1 - \alpha}\right) \tilde{\tilde{x}}_T = (2\tilde{x}_T - \tilde{\tilde{x}}_T) + \frac{\alpha}{1 - \alpha} (\tilde{x}_T - \tilde{\tilde{x}}_T) h,$$

или

$$\hat{x}_{T+h} = a + bh, \quad h = 1, 2, \dots,$$

$$\text{где } a = 2\tilde{x}_T - \tilde{\tilde{x}}_T, \quad b = \frac{\alpha}{1 - \alpha} (\tilde{x}_T - \tilde{\tilde{x}}_T).$$

Значения a и b суть параметры локального прямолинейного тренда ряда. Прогнозные значения лежат на прямой, определяемой этими параметрами.

ПРИМЕР 5.1.8

На периоде с I квартала 1994 г. по IV квартал 1997 г. количество безработных в Российской Федерации (в млн. человек) (ряд *UNEMP*) изменялось следующим образом (рис. 5.19).

Ряд имеет выраженный линейный тренд. Применим двойное экспоненциальное сглаживание для получения прогнозных значений этого ряда на 1997 г. по наблюдениям этого ряда за 1994—1996 гг. Сначала положим $\alpha = 0.2$. Результаты отражены на рис. 5.20 и в табл. 5.3.

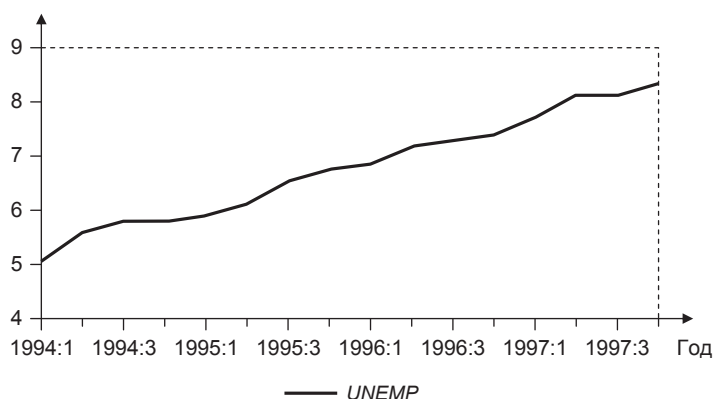


Рис. 5.19

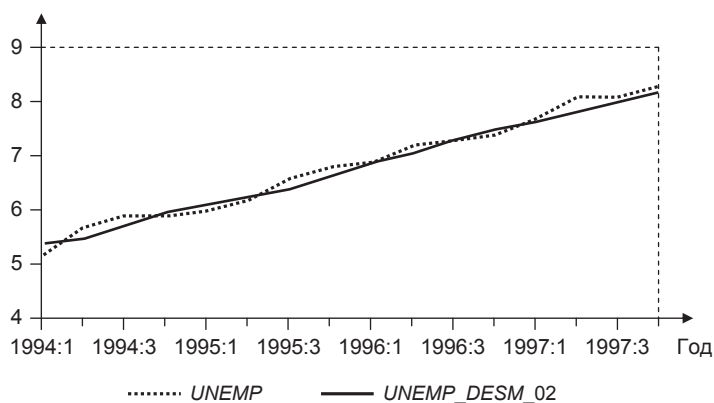


Рис. 5.20

Таблица 5.3

Двойное экспоненциальное сглаживание ряда

Sample: 1994:1 1996:4; Included observations: 12;

Method: Double Exponential; Original Series: UNEMP; Forecast Series: UNEMP_DESM_02

Parameters:	Alpha	0.2
	Sum of Squared Residuals	0.224375
	Root Mean Squared Error	0.136740
End of Period Levels:	Mean	7.257440
	Trend	0.183511

Если оставить программе подбор оптимального значения параметра α , то это приводит к значению $\alpha = 0.162$ и практически к тем же результатам (рис. 5.21 и табл. 5.4).

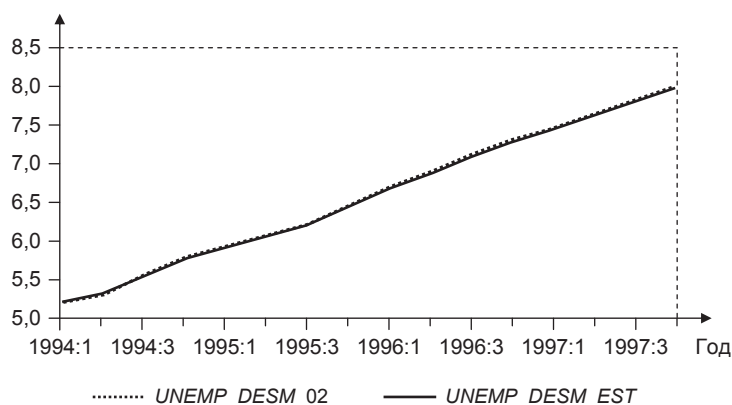


Рис. 5.21

Таблица 5.4

Двойное сглаживание ряда с подбором оптимального значения α

Sample: 1994:1 1996:4; Included observations: 12;

Method: Double Exponential; Original Series: UNEMP; Forecast Series: UNEMP_DESM_EST

Parameters:	Alpha	0.162
	Sum of Squared Residuals	0.221352
	Root Mean Squared Error	0.135816
End of Period Levels:	Mean	7.247008
	Trend	0.181239

Заметим, что если для сглаживания и прогнозирования этого ряда использовать простое экспоненциальное сглаживание с оцениваемым значением α , то это значение оценивается как $\alpha = 0.999$, и получим результаты, приведенные в табл. 5.5 и на рис. 5.22.

Таблица 5.5

Простое экспоненциальное сглаживание ряда с оцениваемым α

Sample: 1994:1 1996:4; Included observations: 12;

Method: Single Exponential; Original Series: UNEMP; Forecast Series: UNEMP_SSM

Parameters:	Alpha	0.999
	Sum of Squared Residuals	1.040402
	Root Mean Squared Error	0.294449
End of Period Levels:	Mean	7.199900

Как видно на рис. 5.23, здесь прогнозы на один шаг вперед практически совпадают с текущими значениями ряда; прогнозы на несколько шагов вперед не отличаются от прогнозов на один шаг вперед. ■

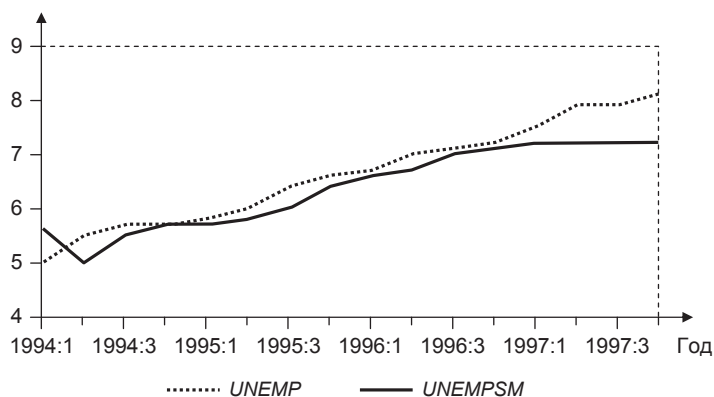


Рис. 5.22

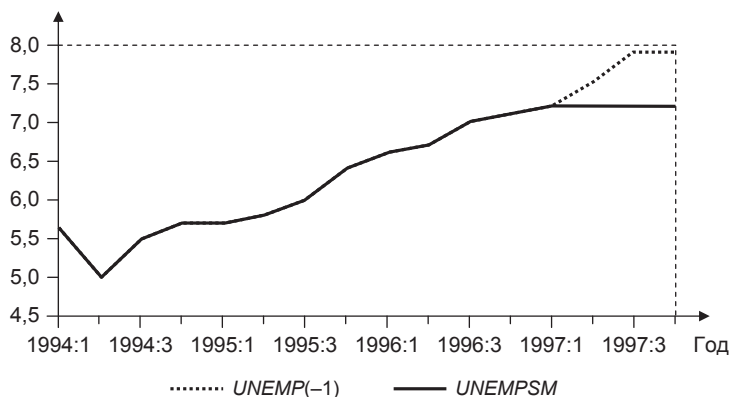


Рис. 5.23

Метод Хольта (*Holt's Linear Trend Algorithm*) используется для прогнозирования рядов с выраженным линейным трендом без сезонных составляющих. При этом прогноз осуществляется в направлении текущего локального линейного тренда, определяемого текущим уровнем ряда и текущим угловым коэффициентом локального тренда.

Пусть

L_i — текущий уровень ряда;

T_i — текущий угловой коэффициент локального тренда, соответствующие периоду i .

Эти величины вычисляются рекуррентным образом по следующим формулам:

$$L_i = \alpha x_i + (1 - \alpha)(L_{i-1} + T_{i-1}), \quad 0 < \alpha < 1,$$

$$T_i = \beta (L_i - L_{i-1}) + (1 - \beta)T_{i-1}, \quad 0 < \beta < 1.$$

Иначе говоря:

- «текущий уровень» для периода i есть взвешенное среднее значения ряда x_i в этом периоде и $(L_{i-1} + T_{i-1})$ — прогноза этого значения по локальному тренду, оцененному по предыдущим наблюдениям;
- «текущий угловой коэффициент» для периода i есть взвешенное среднее приращения «текущего уровня» $(L_i - L_{i-1})$ и значения «текущего углового коэффициента» для периода $(i - 1)$.

Соответственно значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\hat{x}_{i+h} = L_i + T_i h.$$

ПРИМЕР 5.1.9

Применим метод Хольта к ряду данных о безработице в России, предоставив компьютеру возможность выбора оптимальных значений параметров α и β (название сглаженного ряда *UNEMP_HW_EST*, предпрогнозный период 1994:1—1996:4). Результаты приведены в табл. 5.6. На рис. 5.24 изображены график ряда *UNEMP* в интервале 1994:1—1997:4, а также график сглаженного ряда *UNEMP_HW_EST* на предпрогножном периоде и сделанные по нему прогнозы на 4 квартала 1997 г.

Таблица 5.6

Метод Хольта с выбором оптимальных значений α и β

Sample: 1994:1 1996:4; Included observations: 12;

Method: Holt-Winters No Seasonal; Original Series: UNEMP;

Forecast Series: UNEMP_HW_EST

Parameters:	Alpha	1.0000
	Beta	0.0000
	Sum of Squared Residuals	0.232222
	Root Mean Squared Error	0.139111
End of Period Levels:	Mean	7.200000
	Trend	0.233333

Сравним прогнозы на 1997 г., полученные методами Брауна и Хольта (с оцениваемыми параметрами). Для этого вычислим сумму квадратов ошибок прогнозов и *RMSE* на прогножном периоде. Непосредственный подсчет этих величин (в протоколе оценивания в пакете *EViews* они не приводятся) дает результаты, указанные в табл. 5.7 (для метода Брауна) и в табл. 5.8 (для метода Хольта). Алгоритм Хольта дает большее значение угловому коэффициенту локального тренда (рис. 5.25), при этом качество прогноза несколько лучше. ■

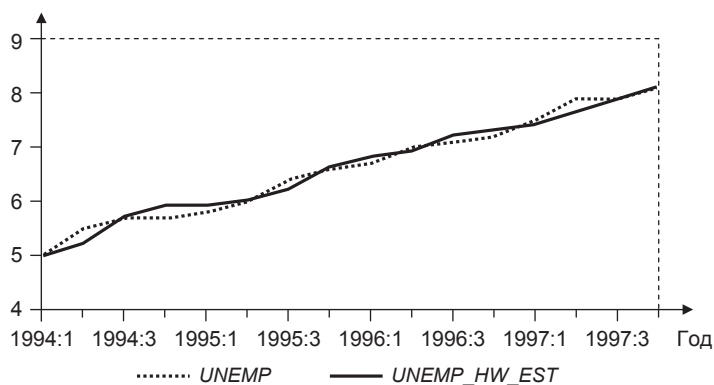


Рис. 5.24

Таблица 5.7

Прогноз, полученный с помощью метода Брауна

	<i>Sum of Squared Residuals</i>	0.117881
	<i>Root Mean Squared Error</i>	0.171667
<i>End of Period Levels:</i>	<i>Mean</i>	7.247008
	<i>Trend</i>	0.181239

Таблица 5.8

Прогноз, полученный с помощью метода Брауна

	<i>Sum of Squared Residuals</i>	0.060000
	<i>Root Mean Squared Error</i>	0.122474
<i>End of Period Levels:</i>	<i>Mean</i>	7.200000
	<i>Trend</i>	0.233333

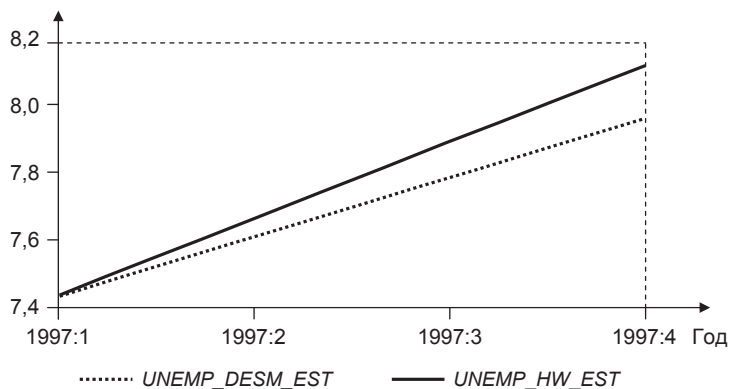


Рис. 5.25

Метод Хольта — Винтерса (*Holt — Winters's Algorithm*). Этот метод был предложен Винтерсом как обобщение метода Хольта, допускающее наличие сезонного фактора в динамике ряда. К паре уравнений Хольта, определяющих текущий уровень и текущий угловой коэффициент ряда, здесь добавляется третье уравнение, определяющее текущий сезонный фактор S_i . При этом сезонный фактор может быть как аддитивным, так и мультипликативным.

Если сезонный фактор **аддитивный**, то используется следующая система:

$$\begin{cases} L_i = \alpha(x_t - S_{t-s}) + (1 - \alpha)(L_{i-1} + T_{i-1}), & 0 < \alpha < 1, \\ T_i = \beta(L_i - L_{i-1}) + (1 - \beta)T_{i-1}, & 0 < \beta < 1, \\ S_i = \gamma(x_i - L_i) + (1 - \gamma)S_{i-s}, & 0 < \gamma < 1, \end{cases}$$

где $s = 4$ для квартальных данных и $s = 12$ для месячных данных.

Значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\begin{aligned} \hat{x}_{i+h} &= L_i + T_i h + S_{i+h-s} \quad \text{для } h = 1, 2, \dots, s, \\ \hat{x}_{i+h} &= L_i + T_i h + S_{i+h-2s} \quad \text{для } h = s + 1, s + 2, \dots, 2s, \end{aligned}$$

и т.д. Прогноз представляет собой прямую линию, на которую накладываются аддитивные сезонные факторы.

Если сезонный фактор **мультипликативный**, то используется следующая система:

$$\begin{cases} L_i = \frac{\alpha x_i}{S_{i-s}} + (1 - \alpha)(L_{i-1} + T_{i-1}), & 0 < \alpha < 1, \\ T_i = \beta(L_i - L_{i-1}) + (1 - \beta)T_{i-1}, & 0 < \beta < 1, \\ S_i = \frac{\gamma x_i}{L_i} + (1 - \gamma)S_{i-s}, & 0 < \gamma < 1. \end{cases}$$

Значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\begin{aligned} \hat{x}_{i+h} &= (L_i + T_i h) S_{i+h-s} \quad \text{для } h = 1, 2, \dots, s, \\ \hat{x}_{i+h} &= (L_i + T_i h) S_{i+h-2s} \quad \text{для } h = s + 1, s + 2, \dots, 2s, \end{aligned}$$

и т.д.

Чаще применяют вариант с мультипликативной сезонностью, поскольку в аддитивном случае сглаженный ряд имеет постоянные сезонные колебания, величина которых не зависит от общего уровня значений ряда, а в мультипликативном случае величина сезонных колебаний может изменяться в зависимости от общего уровня значений ряда. Можно, однако, применить оба варианта метода Хольта — Винтерса и выбрать из них лучший по качеству ретроспективных прогнозов.

ПРИМЕР 5.1.10

Вернемся к ряду *AGRO*, имеющему ярко выраженный сезонный характер (рис. 5.26). Применим к нему алгоритм Хольта — Винтерса с оцениваемыми параметрами. Оценивание проведем по данным до IV квартала 2003 г. включительно, а прогнозные значения вычислим на период с I квартала 2004 г. по IV квартал 2005 г. В табл. 5.9 приведены результаты применения алгоритма с аддитивной сезонностью, а в табл. 5.10 — результаты применения алгоритма с мультипликативной сезонностью.

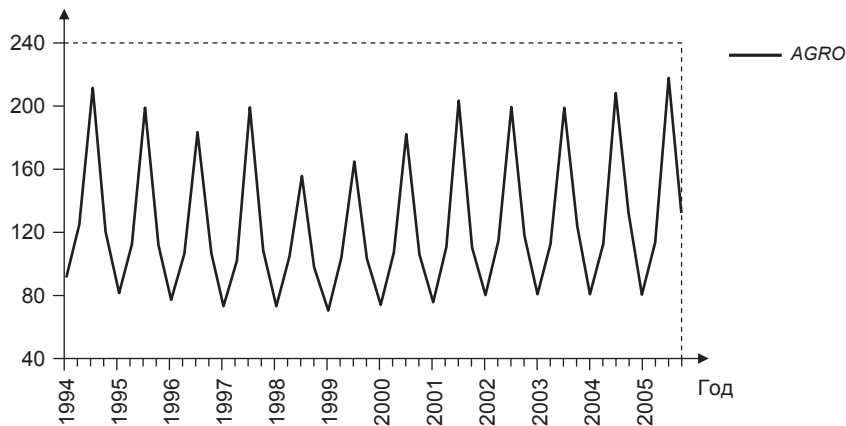


Рис. 5.26

Таблица 5.9

Применение метода Хольта — Винтерса с аддитивной сезонностью

Sample: 1994:1 2003:4; Included observations: 40;
Method: Holt-Winters Additive Seasonal; Original Series: AGRO;
Forecast Series: AGRO_HW_AD

Parameters:	Alpha		0.4600
	Beta		0.0000
	Gamma		0.0000
	Sum of Squared Residuals		2634.295
	Root Mean Squared Error		8.115255
End of Period Levels:	Mean		127.3298
	Trend		−0.201389
	Seasonals:	2003:1	−43.68208
		2003:2	−12.21069
		2003:3	66.45069
		2003:4	−10.55792

Таблица 5.10

Применение метода Хольта — Винтерса с мультипликативной сезонностью

Sample: 1994:1 2003:4; Included observations: 40;

Method: Holt-Winters Multiplicative Seasonal; Original Series: AGRO;

Forecast Series: AGRO_HW_MU

<i>Parameters:</i>	<i>Alpha</i>		0.6200
	<i>Beta</i>		0.0900
	<i>Gamma</i>		0.0000
	<i>Sum of Squared Residuals</i>		1808.309
	<i>Root Mean Squared Error</i>		6.723669
<i>End of Period Levels:</i>	<i>Mean</i>		129.8795
	<i>Trend</i>		1.073999
	<i>Seasonals:</i>	2003:1	0.627968
		2003:2	0.897011
		2003:3	1.565845
		2003:4	0.909176

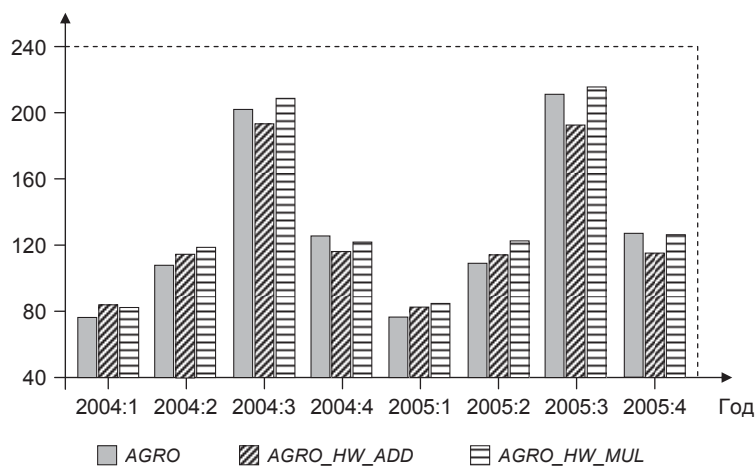


Рис. 5.27

Вариант с мультипликативной сезонностью дает лучшие результаты ретроспективных прогнозов. Диаграмма на рис. 5.27 дает возможность сравнить перспективные прогнозы по аддитивной и мультипликативной моделям. Сумма квадратов ошибок прогнозов на период с I квартала 2004 г. по IV квартал 2005 г. равна 813.75 для аддитивной модели и 489.98 для мультипликативной, так что мультипликативная модель оказалась лучшей по этому показателю и для перспективного прогнозирования. ■

Экстраполирование тренда, оцененного методом наименьших квадратов

Выше было показано, как можно получать прогнозы на основе выделения трендовых и сезонных составляющих методом сглаживания. Часто на основе визуального анализа графика ряда удается делать более определенные суждения о характере тренда. В динамике ряда на рассматриваемом периоде может явно проявляться линейный, полиномиальный или экспоненциальный характер тренда. В таких случаях прогнозирование может опираться на результаты оценивания параметров функции, описывающей тренд на всем периоде наблюдений, а само оценивание можно производить, используя метод наименьших квадратов.

ПРИМЕР 5.1.11

Ряд *UNEMP* — ряд данных о безработице в Российской Федерации на периоде с I квартала 1994 г. по IV квартал 1997 г. (квартальные данные) имеет выраженный линейный тренд (рис. 5.28).

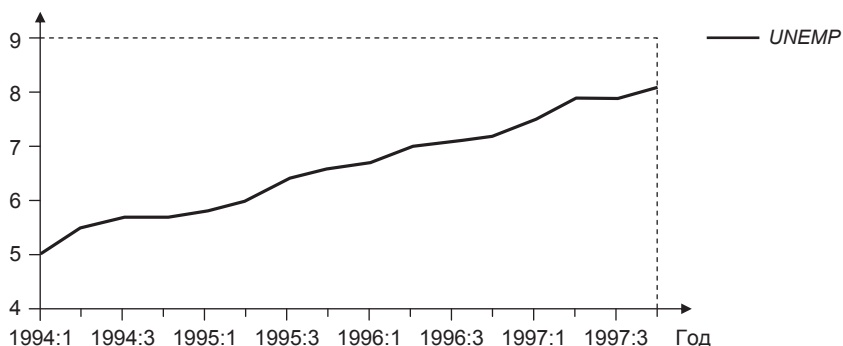


Рис. 5.28

Прогнозные значения этого ряда на 1997 г. получим по наблюдениям ряда за 1994—1996 гг., оценивая методом наименьших квадратов модель

$$UNEMP_t = \theta_1 + \theta_2 t + \varepsilon_t, \quad t = 1, \dots, 12.$$

Подобранная модель имеет вид:

$$UNEMP_t = 4.963 + 0.194t.$$

Прогнозные значения $UNEMP_t$ ряда на 1997 г. (табл. 5.11) получим по этой формуле, подставив в правую часть последовательно значения $t = 13, 14, 15, 16$. На рис. 5.29 сопоставлены наблюдаемые и прогнозные значения ряда на прогнозном периоде.

В табл. 5.12 приведены характеристики качества сделанного прогноза. ■

Таблица 5.11

Наблюдаемые и прогнозные значения ряда на 1997 г.

Период	Наблюдаемое значение	Прогнозное значение	Ошибка прогноза
1997:1	7.5	7.49	0.01
1997:2	7.9	7.68	0.22
1997:3	7.9	7.87	0.03
1997:4	8.1	8.07	0.03

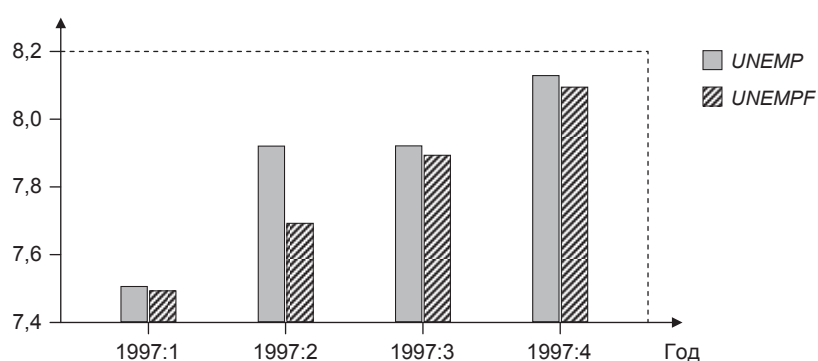


Рис. 5.29

Таблица 5.12

Характеристики качества прогноза на 1997 г.

Forecast: UNEMPF; Actual: UNEMP;

Forecast sample: 1997Q1 1997Q4; Included observations: 4

<i>Root Mean Squared Error</i>	0.111852
<i>Mean Absolute Error</i>	0.072552
<i>Mean Absolute Percent Error</i>	0.918228
<i>Theil Inequality Coefficient</i>	0.007155
<i>Bias Proportion</i>	0.420744
<i>Variance Proportion</i>	0.000077
<i>Covariance Proportion</i>	0.579178

В этой таблице:

- *Root Mean Squared Error (RMSE)* — корень из среднего квадрата ошибки прогноза ряда y_t на прогнозном периоде:

$$RMSE = \sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2};$$

- *Mean Absolute Error (MAE)* — средняя абсолютная ошибка прогноза ряда y_t на прогнозном периоде:

$$MAE = \frac{1}{h} \sum_{t=T+1}^{T+h} |\hat{y}_t - y_t|;$$

- *Mean Absolute Percent Error (MAPE)* — средняя абсолютная процентная ошибка прогноза ряда y_t на прогнозном периоде:

$$MAPE = \frac{1}{h} \sum_{t=T+1}^{T+h} \left| \frac{\hat{y}_t - y_t}{y_t} \right| \cdot 100\%;$$

- *Theil Inequality Coefficient (TIC)* — нормированный коэффициент Тейла:

$$TIC = \frac{\sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2}}{\sqrt{\frac{1}{h} \sum_{t=T+1}^{T+h} \hat{y}_t^2 + \frac{1}{h} \sum_{t=T+1}^{T+h} y_t^2}}.$$

Значения этого коэффициента изменяются от 0 до 1. Близкие к 0 значения *TIC* свидетельствуют о хорошем качестве прогноза.

В трех последних строках таблицы указаны величины, связанные со следующим разложением среднего квадрата ошибки прогнозов на прогнозном периоде длины h :

$$\frac{1}{h} \sum (\hat{y}_t - y_t)^2 = (\bar{\hat{y}} - \bar{y})^2 + (s_{\hat{y}} - s_y)^2 + 2(1-r)s_{\hat{y}}s_y,$$

где $\bar{\hat{y}}, \bar{y}, s_{\hat{y}}, s_y$ — выборочные средние и стандартные отклонения рядов \hat{y}_t и y_t на прогнозном периоде;

r — выборочный коэффициент корреляции между \hat{y} и y на прогнозном периоде.

- *Bias Proportion* — доля первого слагаемого в сумме. Большие значения этого показателя соответствуют систематическому смещению прогнозных значений по одну сторону от наблюдаемых значений ряда на прогнозном периоде.
- *Variance Proportion* — доля второго слагаемого в сумме. Большие значения этого показателя наблюдаются при несоответствии вариабельностей прогнозных и наблюдаемых значений ряда на прогнозном периоде.
- *Covariance Proportion* — доля третьего слагаемого в сумме. Малые значения этого показателя соответствуют расхождению по фазе колебаний ряда прогнозных и наблюдаемых значений ряда на прогнозном периоде.

ПРИМЕР 5.1.12

Рассмотрим статистические данные о размерах совокупной прибыли нефинансовых корпораций США (до уплаты налогов, в млрд долл., квартальные данные) за период с I квартала 1985 г. по I квартал 1989 г. — ряд *PROFITS*. График этого ряда изображен на рис. 5.30.

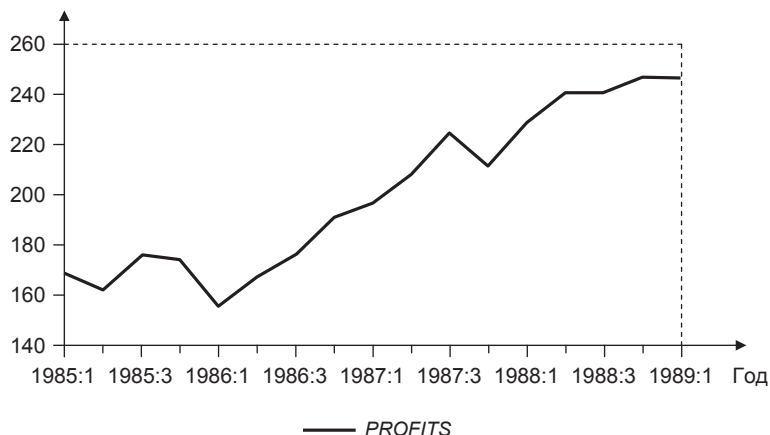


Рис. 5.30

Здесь тренд сложнее считать прямолинейным, но можно попытаться приблизить его полиномом — например, рассмотреть модель квадратичного тренда:

$$PROFITS_t = \theta_1 + \theta_2 t + \theta_3 t^2 + \varepsilon_t, \quad t = 1, \dots, 17.$$

Результаты оценивания такой модели по данным до I квартала 1988 г. включительно приведены в табл. 5.13. На рис. 5.31 приведены графики наблюдаемых и прогнозных значений ряда.

Таблица 5.13

Оценивание модели квадратичного тренда

Dependent Variable: *PROFITS*; Method: Least Squares; Sample: 1985:1 1988:1;
Included observations: 13

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
<i>C</i>	168.0573	8.912084	18.85724	0.0000
<i>T</i>	−2.087113	2.927791	−0.712863	0.4922
<i>T</i> ²	0.544056	0.203496	2.673552	0.0233
<i>R-squared</i>	0.881340	Mean dependent var		187.7231

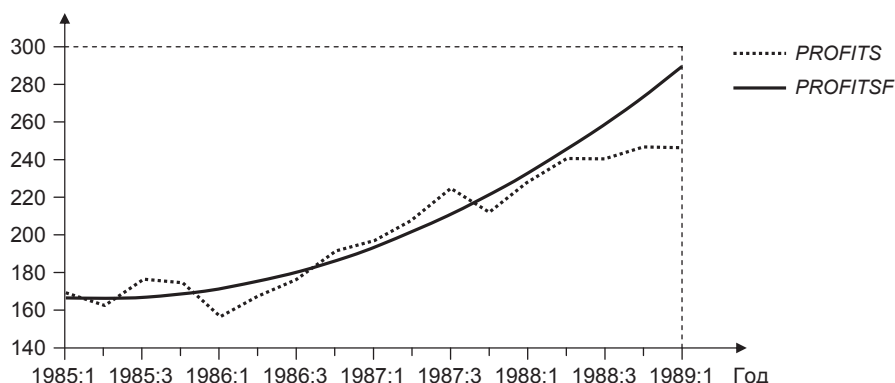


Рис. 5.31

Прогнозные значения ряда на период со II квартала 1988 г. по I квартал 1989 г. расположены выше реально наблюдаемых значений. Предупредить такое завышение прогнозных значений можно, обратившись к модели кубического тренда:

$$PROFITSt = \theta_1 + \theta_2 t + \theta_2 t^2 + \theta_3 t^3 + \varepsilon_t, \quad t = 1, \dots, 17.$$

Результаты оценивания этой модели по данным до I квартала 1988 г. включительно приведены в табл. 5.14.

Таблица 5.14

Оценивание модели кубического тренда

Dependent Variable: PROFITS

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	181.9888	13.33670	13.64572	0.0000
T	-12.20401	7.943962	-1.536263	0.1588
T^2	2.285490	1.293811	1.766479	0.1111
T^3	-0.082925	0.060904	-1.361579	0.2064
R-squared	0.901608	Mean dependent var		187.7231

Теперь (рис. 5.32) прогнозные значения на период со II квартала 1988 г. по I квартал 1989 г. оказались ниже реально наблюдаемых. В то же время $RMSE$ для этой модели (13.12275) вдвое меньше $RMSE$ для квадратичной модели (27.46526).

Сравним полученные прогнозы с прогнозами, сделанными ранее в результате применения двойного экспоненциального сглаживания с оцениваемыми параметрами (рис. 5.33). На этот раз не все прогнозные значения оказываются по одну сторону от наблюдаемых, а значение $RMSE$ на прогнозном периоде равно $RMSE = 5.244$, что лучше, чем для моделей квадратичного и кубического трендов. ■

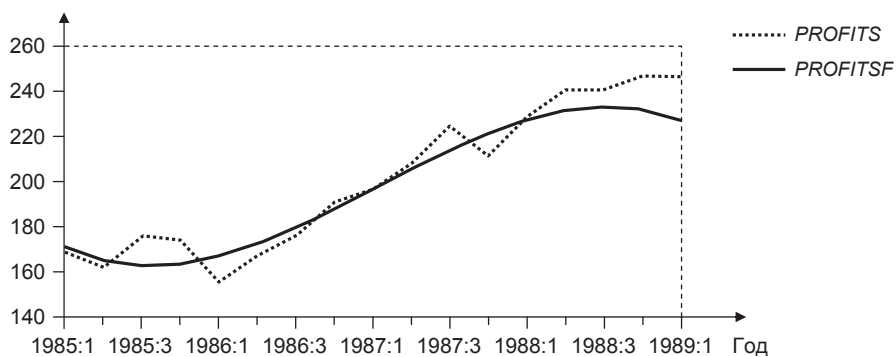


Рис. 5.32

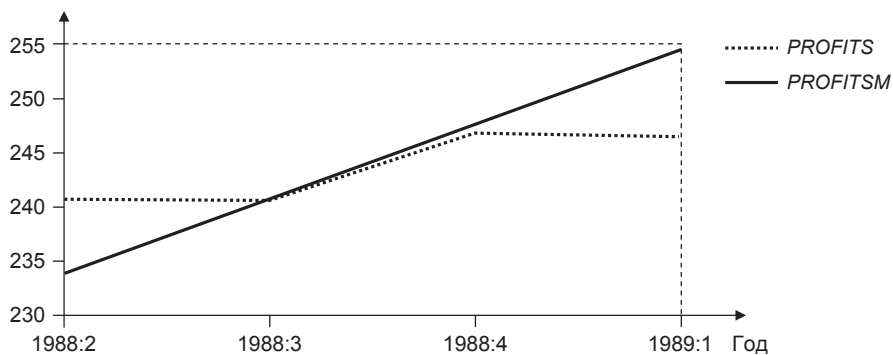


Рис. 5.33

ПРИМЕР 5.1.13

Применив фильтр Ходрика — Прескотта с $\lambda = 1600$ к ряду *AGRO*, получили ряд *HPTREND*, интерпретируемый как «долговременный тренд» (рис. 5.34).

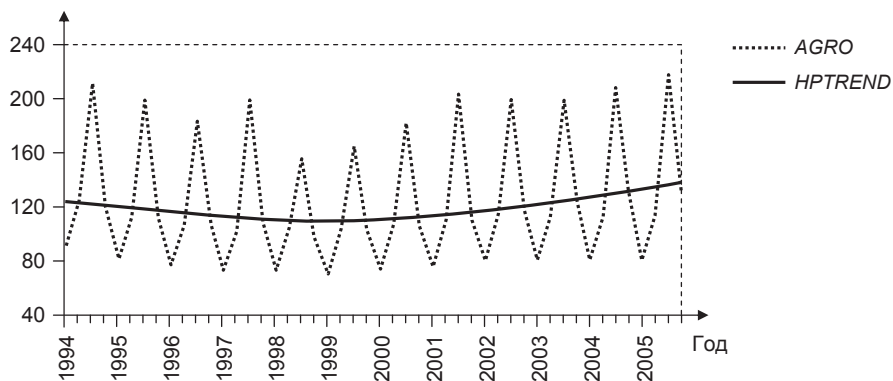


Рис. 5.34

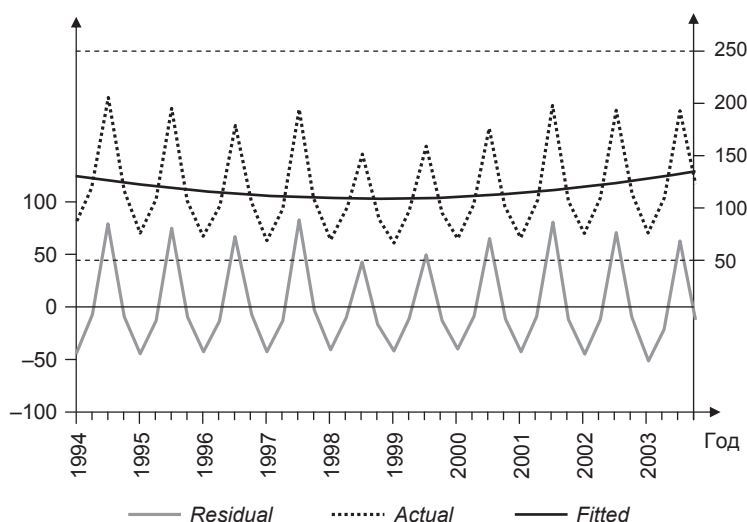


Рис. 5.35

Попробуем теперь выделить тренд регрессионными методами. Рассмотрим сначала модель квадратичного тренда:

$$AGRO_t = \theta_1 + \theta_2 t + \theta_3 t^2 + \varepsilon_t, \quad t = 1, \dots, 48,$$

и оценим ее по данным до IV квартала 2003 г. включительно.

В остатках (ряд *Residual* на рис. 5.35) видимого тренда нет, но есть явно выраженная сезонность, которую можно учесть, введя в правую часть уравнения регрессии вместо постоянной составляющей 4 сезонные дамми-переменные:

- $D1$ — переменная, равная 1 в I квартале и равная 0 в остальных трех кварталах;
- $D2$ — переменная, равная 1 во II квартале и равная 0 в остальных трех кварталах;
- $D3$ — переменная, равная 1 в III квартале и равная 0 в остальных трех кварталах;
- $D4$ — переменная, равная 1 в IV квартале и равная 0 в остальных трех кварталах.

Расширенную квадратичную модель с 4 сезонными дамми-переменными

$$AGRO_t = \theta_1 D1_t + \theta_2 D2_t + \theta_3 D3_t + \theta_4 D4_t + \theta_5 t + \theta_6 t^2 + \varepsilon_t$$

оцениваем по данным до IV квартала 2003 г. включительно. Прогнозирование по оцененной модели на 2004—2005 гг. дает весьма высокую долю смещения ($Bias\ Proportion = 0.779$), что иллюстрирует рис. 5.36, при этом $RMSE = 17.741$.

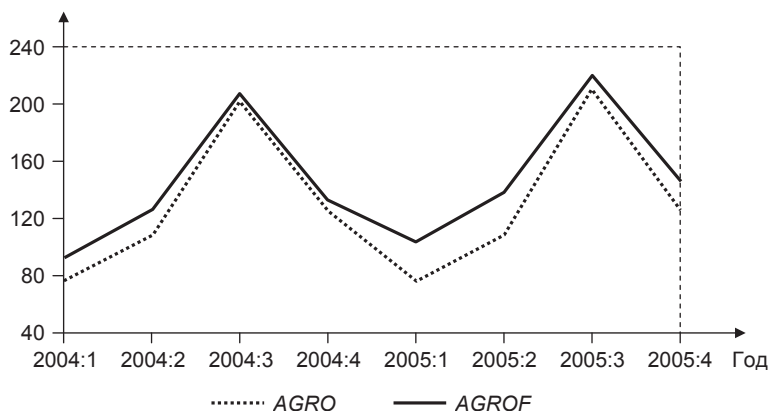


Рис. 5.36

Сравним это с результатами применения методики Хольта — Винтерса. Для аддитивной сезонности $RMSE = 10.086$, для мультипликативной сезонности $RMSE = 7.826$. Оба варианта метода Хольта — Винтерса дают лучшие прогнозы. ■

Оценим теперь расширенную квадратичную модель с 4 сезонными дамми-переменными на всем интервале с I квартала 1994 г. по IV квартал 2005 г. и получим ряд остатков $RESID_EQ1$. Наряду с ней оценим модель с 4 сезонными дамми-переменными для ряда с удаленным трендом, полученным применением фильтра Ходрика — Прескотта, т.е. для ряда $AGRO_HPTREND$. В результате оценивания второй модели получаем ряд прогнозных значений $SEASONAL$ и ряд остатков $RESID_EQ2$. На рис. 5.37 приведены графики обоих рядов остатков.

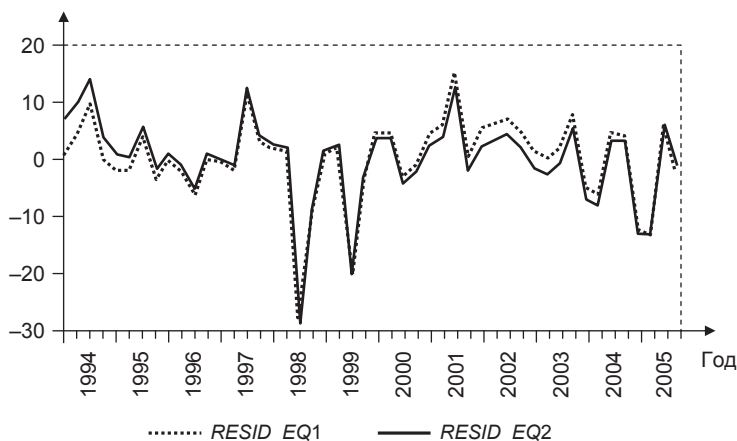


Рис. 5.37

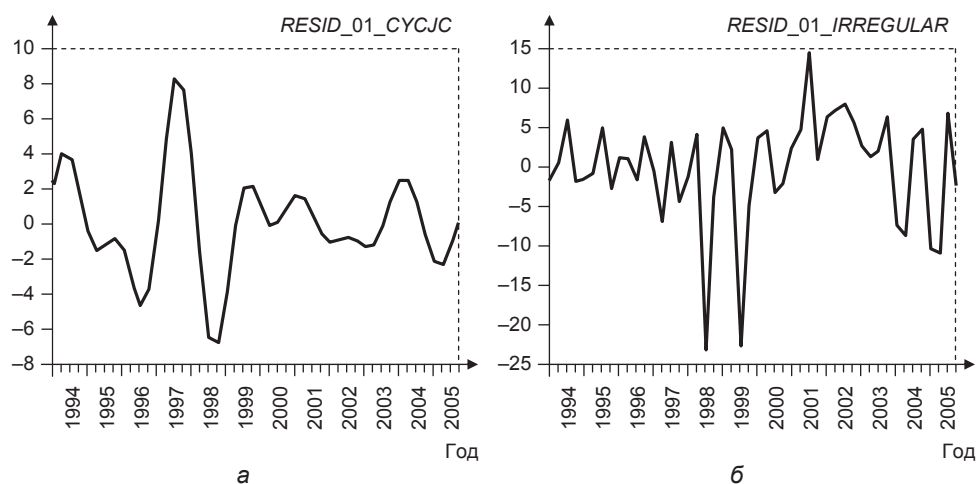


Рис. 5.38

В пакете *EViews 6* предусмотрена возможность выделения из временного ряда циклических составляющих в заданном интервале частот. Пользуясь этой возможностью, выделим из ряда остатков *RESID_EQ1* колебания с периодами от 1.5 до 4 лет (ряд *RESID_01_CYCLIC*). Очищенный от колебаний с такими частотами ряд $RESID_01_IRREGULAR = RESID_EQ1 - RESID_01_CYCLIC$ интерпретируем как нерегулярную компоненту ряда *AGRO*. Графики рядов *RESID_01_CYCLIC* и *RESID_01_IRREGULAR* представлены на рис. 5.38 (а, б).

Итоговое разложение ряда *AGRO* представлено на рис. 5.39.

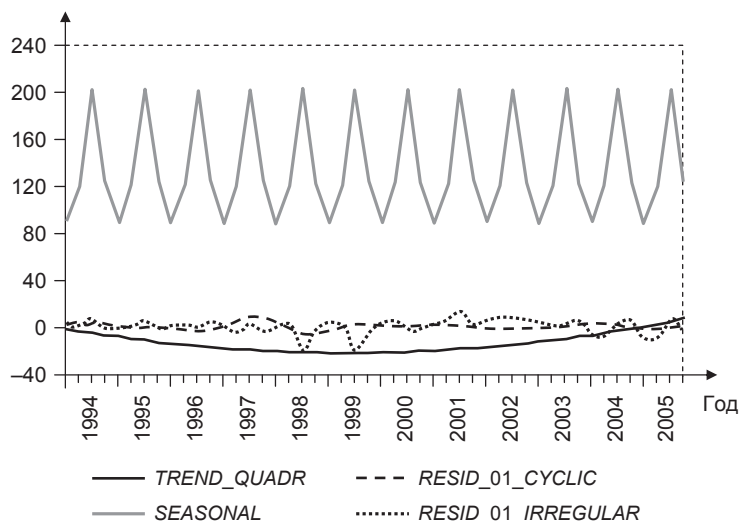


Рис. 5.39

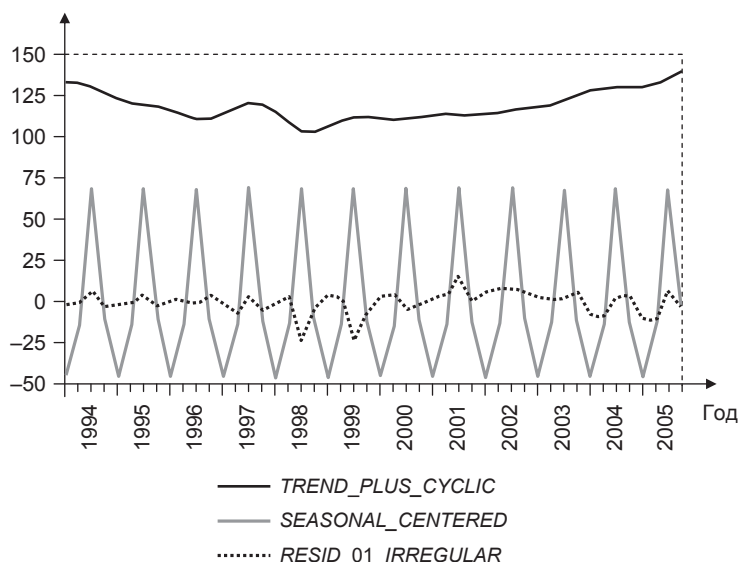


Рис. 5.40

Центрируем ряд *SEASONAL*, объединим тренд и циклическую компоненту. Соответствующее разложение ряда *AGRO* представлено на рис. 5.40.

Итак, прогнозы можно получать на основе оценивания трендовых и сезонных составляющих методом наименьших квадратов. Однако эти прогнозы не учитывают информацию о будущих значениях ряда, которая может содержаться в нерегулярной компоненте ряда. Чтобы учесть такую информацию, необходимо иметь подходящую модель и для этой компоненты. Для этой цели можно использовать стохастические (вероятностные) модели временных рядов, которые рассматривались во второй части учебника. Такие модели допускают наличие в динамике исходного ряда стохастического тренда, который может выглядеть как детерминированный, но наличие которого приводит к совершенно другим оптимальным прогнозам.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какие систематические эффекты выделяют в поведении временного ряда?
2. Как представляется временной ряд в форме мультипликативной и в форме аддитивной модели?
3. Как определяется простое скользящее среднее с периодом усреднения K ?
4. В каких случаях используется центрированное скользящее среднее и как оно определяется?
5. С какой целью используется фильтр Ходрика — Прескотта и как он определяется?
6. Как строятся прогнозы на h шагов вперед при использовании простого скользящего среднего и при использовании взвешенного скользящего среднего?

7. Как производится простое экспоненциальное сглаживание? Как в этом случае строятся прогнозные значения на h шагов вперед? Как осуществляется выбор параметра сглаживания?
8. Как производится двойное экспоненциальное сглаживание и когда оно применяется? Как строятся в этом случае прогнозные значения на h шагов вперед?
9. В чем состоит метод Хольта и когда он применяется? Как строятся в этом случае прогнозные значения на h шагов вперед?
10. Чем отличается метод Хольта — Винтерса от метода Хольта? Чем различаются два варианта метода Хольта — Винтерса?
11. Как может применяться метод наименьших квадратов для прогнозирования будущих значений временного ряда?
12. Каковы основные характеристики качества прогнозов? Как они интерпретируются?

Тема 5.2

ПРОГНОЗИРОВАНИЕ ПО МОДЕЛЯМ AR, MA, ARMA, ARIMA

Пусть имеем ряд наблюдений x_t , $t = 1, \dots, T$. Если динамика ряда не обнаруживает систематических компонент в форме долговременного тренда или сезонности, то по этим наблюдениям можно построить коррелограмму ряда и на ее основе попытаться идентифицировать модель порождения этого ряда как:

- процесс скользящего среднего MA(q);
- стационарный процесс авторегрессии AR(p);
- стационарный процесс ARMA(p, q).

Однако случайный процесс, порождающий ряд наблюдений, может и не быть стационарным. Например, он может содержать детерминированный и/или стохастический тренд. Первоначально подход Бокса — Дженкинса (Бокс, Дженкинс, 1974)) предлагал предварительно остационаривать ряд наблюдений и уже после этого применять процедуру идентификации модели. При этом остационаривание ряда наблюдений производится путем перехода от исходного ряда x_t к ряду его разностей (дифференцирование ряда):

$$\Delta x_t = x_t - x_{t-1}, \quad t = 2, 3, \dots$$

Если полученный ряд все еще нестационарный, то операция дифференцирования повторяется вплоть до получения стационарного ряда.

Некоторое осложнение в применении метода Бокса — Дженкинса вызывает то обстоятельство, что при использовании этого метода возможно получение «передифференцированного» ряда, при оценивании и при использовании которого для целей прогнозирования возникают определенные технические трудности. Это происходит, например, в случае, когда ряд x_t является рядом, стационарным относительно линейного тренда. В последнем случае для остационаривания ряда следует вычесть из него соответствующий детермини-

рованный тренд (разумеется, на практике производится оценивание этого тренда, и вычитается оцененный тренд). В свою очередь, если ряд имеет стохастический тренд, то вычитание из этого ряда оцененного детерминированного тренда не приводит к стационарному ряду. В связи с этим перед применением метода Бокса — Дженкинса обычно принимают решение о том, как производить остационирование ряда: используя дифференцирование ряда или вычитая из него оцененный тренд.

Первоначально основой для принятия одного из этих двух решений являлся анализ коррелограммы ряда. Для коррелограммы интегрированного ряда характерно достаточно медленное затухание выборочной автокорреляционной функции $r(k)$ и близкое к 1 значение этой функции на единичном запаздывании (значение $r(1)$). Однако подобный характер коррелограммы имеют и ряды с детерминированным линейным трендом, дифференцирование которых нежелательно. Поэтому в настоящее время решение принимают, основываясь на результатах проверки в отношении ряда x_t гипотезы единичного корня. Некоторые статистические критерии для проверки этой гипотезы были рассмотрены во второй части учебника. Наиболее распространенным инструментом проверки гипотезы о наличии единичного корня является критерий Дики — Фуллера, предполагающий, что рассматриваемый процесс можно представить в виде процесса $AR(p)$:

$$a(L)x_t = \varepsilon_t.$$

Критерий Дики — Фуллера имеет варианты, описывающие различные типы поведения реализаций процесса и учитывающие значение p .

Если временной ряд обладает выраженной сезонностью, то перед применением критерия Дики — Фуллера рекомендуется провести «десезонизацию» ряда, т.е. применять этот критерий к ряду остатков, полученных при оценивании уравнения регрессии ряда на все сезонные дамми-переменные (при этом константа не включается в оцениваемое уравнение).

Если гипотеза единичного корня не отвергается, то для достижения стационарности ряда его следует продифференцировать. Если для достижения стационарности ряд x_t приходится дифференцировать d раз и полученный в результате ряд оказывается стационарным типа $ARMA(p, q)$, то говорят, что ряд x_t является интегрированным (проинтегрированным) рядом порядка d ($I(d)$), и для такого ряда используется обозначение $ARIMA(p, d, q)$. Чаще всего на практике приходится иметь дело с интегрированными рядами не выше второго порядка.

Критерий Дики — Фуллера можно применять и тогда, когда ряд x_t описывается смешанной моделью авторегрессии — скользящего среднего. Дело в том, что если ряд наблюдений x_1, \dots, x_T порождается моделью $ARIMA(p, 1, q)$ с $q > 0$, то его можно аппроксимировать моделью $ARIMA(p^*, 1, 0)$ с $p^* < \sqrt[3]{T}$ и применять процедуру Дики — Фуллера к этой модели.

Пусть x_t — случайный процесс, порождающий ряд наблюдений. Наблюдаются значения x_t , $t = 1, \dots, T$, до момента T включительно и нужно построить прогноз значения этого ряда на следующий момент времени $(T + 1)$. Обозначим такой прогноз $x_{T+1|T}$. Совокупность значений $\Omega_T = \{x_t, t = 1, \dots, T\}$ можно трактовать как **информационное множество**, содержащее всю имеющуюся информацию, на основании которой строится прогнозное значение, так что это прогнозное значение есть некоторая функция от Ω_T , $x_{T+1|T} = g(\Omega_T)$. Чтобы получить в такой ситуации в некотором смысле наилучший (оптимальный) прогноз, необходим критерий, на основании которого можно было бы сравнивать разные прогнозы (с различными функциями g). Для выбора такого критерия обычно используют концепцию **функции потерь**, состоящую в принятии соглашения о том, что ошибка в прогнозе на величину Δ приводит к издержкам (потерям, *cost*) $C(\Delta)$, где $C(z)$ — некоторая функция, для которой $C(0) = 0$ и $C(z) > 0$ при $z \neq 0$. Если такая функция выбрана, то издержки прогноза $x_{T+1|T} = g(\Omega_T)$ равны $C(e_{T+1|T})$, где

$$e_{T+1|T} = x_{T+1} - g(\Omega_T) \text{ — ошибка прогноза } g(\Omega_T).$$

Естественный критерий выбора наилучшего точечного прогноза на момент $(T + 1)$ состоит в выборе такого прогноза, для которого минимизируются ожидаемые издержки $E(C(e_{T+1|T}) | \Omega_T)$. Здесь математическое ожидание соответствует условному распределению случайной величины x_{T+1} при использовании информационного множества Ω_T . Наиболее часто используется квадратичная функция потерь $C(z) = bz^2$, $b > 0$. В этом случае оптимальный прогноз имеет вид:

$$x_{T+1|T} = E(x_{T+1} | \Omega_T).$$

Если нас интересует прогноз значения ряда на момент $(T + h)$, $h = 1, 2, \dots$, то оптимальный прогноз имеет вид:

$$x_{T+h|T} = E(x_{T+h} | \Omega_T).$$

Заметим, что если временной ряд порождается моделью

$$x_t = f(t) + \varepsilon_t, \quad \varepsilon_t \sim i.i.d., \quad E(\varepsilon_t) = 0, \quad D(\varepsilon_t) = \sigma^2, \quad 0 < \sigma^2 < \infty,$$

то тогда

$$E(x_{T+h} | \Omega_T) = E(f(T+h) + \varepsilon_{T+h} | \Omega_T) = f(T+h),$$

т.е. оптимальный прогноз не зависит от содержания информационного множества. Разумеется, такой вывод предполагает знание точного значения $f(T+h)$, тогда как на практике это значение также приходится оценивать на основе имеющейся информации, и тогда большее количество информации может улучшить качество прогноза по оцененной модели.

Предположим теперь, что имеем дело с остационаренным надлежащим образом рядом x_t , и этот ряд описывается стационарной (в широком смысле) моделью $AR(p)$, $MA(q)$ или $ARMA(p, q)$. Рассмотрим процедуры построения оптимальных прогнозов по таким моделям в ситуации, когда коэффициенты модели известны. Для упрощения записи будем обозначать оптимальный точечный прогноз ряда на h шагов вперед на основе информации $\Omega_T = \{x_t, t = 1, \dots, T\}$ как $f_{T,h}$, а ошибку такого прогноза — как $e_{T,h}$.

Прогнозирование по модели $AR(p)$. Если имеем временной ряд, порождаемый процессом $AR(1)$

$$x_t = a_0 + a_1 x_{t-1} + \varepsilon_t,$$

то оптимальный точечный прогноз на один шаг вперед равен

$$f_{T,1} = E(x_{T+1} | \Omega_T) = a_0 + E(a_1 x_T | \Omega_T) + E(\varepsilon_{T+1} | \Omega_T) = a_0 + a_1 x_T.$$

Оптимальный точечный прогноз на два шага вперед равен:

$$\begin{aligned} f_{T,2} &= E(x_{T+2} | \Omega_T) = a_0 + E(a_1 x_{T+1} | \Omega_T) + E(\varepsilon_{T+2} | \Omega_T) = \\ &= a_0 + a_1 E(x_{T+1} | \Omega_T) = a_0 + a_1 (a_0 + a_1 x_T) = (1 + a_1) a_0 + a_1^2 x_T. \end{aligned}$$

И, вообще, оптимальный точечный прогноз на h шагов вперед равен:

$$f_{T,h} = E(x_{T+h} | \Omega_T) = (1 + a_1 + a_1^2 + \dots + a_1^{h-1}) a_0 + a_1^h x_T.$$

Поскольку у стационарного $AR(1)$ процесса $|a_1| < 1$, при $h \rightarrow \infty$

$$(1 + a_1 + a_1^2 + \dots + a_1^{h-1}) a_0 \rightarrow \frac{a_0}{1 - a_1} = \mu,$$

где μ — математическое ожидание процесса x_t .

По той же причине

$$a_1^h x_T \rightarrow 0 \text{ (в среднем квадратичном),}$$

так что с увеличением h влияние значения x_T на прогноз значения x_{T+h} убывает, а само прогнозное значение приближается к математическому ожиданию процесса x_t .

Заметим еще, что ошибка оптимального прогноза на один шаг равна:

$$e_{T,1} = x_{T+1} - f_{T,1} = \varepsilon_{T+1}.$$

Отсюда вытекает, что последовательность ошибок одношаговых прогнозов процесса $AR(1)$ образует процесс белого шума.

Если имеем дело со стационарным процессом авторегрессии порядка p , так что

$$x_t = a_0 + a_1 x_{t-1} + \dots + a_p x_{t-p} + \varepsilon_t,$$

то для такого процесса:

$$\begin{aligned} f_{T,1} &= E(x_{T+1} | \Omega_T) = a_0 + a_1 E(x_T | \Omega_T) + \dots + a_p E(x_{T-p+1} | \Omega_T) + E(\varepsilon_{T+1} | \Omega_T) = \\ &= a_0 + a_1 x_T + \dots + a_p x_{T-p+1}, \\ f_{T,2} &= E(x_{T+2} | \Omega_T) = a_0 + a_1 E(x_{T+1} | \Omega_T) + \dots + a_p E(x_{T-p+2} | \Omega_T) + E(\varepsilon_{T+2} | \Omega_T) = \\ &= a_0 + a_1 f_{T,1} + (a_2 x_T + \dots + a_p x_{T-p+2}), \\ f_{T,3} &= E(x_{T+3} | \Omega_T) = a_0 + a_1 E(x_{T+2} | \Omega_T) + \dots + a_p E(x_{T-p+3} | \Omega_T) + E(\varepsilon_{T+3} | \Omega_T) = \\ &= a_0 + a_1 f_{T,1} + a_2 f_{T,2} + (a_3 x_T + \dots + a_p x_{T-p+3}), \dots \end{aligned}$$

Вообще, правило получения оптимального прогноза на h шагов вперед, по сути, таково.

Берется явное представление для x_{T+h} , т.е.

$$x_{T+h} = a_0 + a_1 x_{T+h-1} + \dots + a_p x_{T+h-p} + \varepsilon_{T+h},$$

и все входящие в правую часть случайные величины x_j заменяются их оптимальными прогнозами. При этом если $j > T$, то x_j заменяется на $f_{T,j-T}$; если же $j \leq T$, то в правую часть подставляется значение, содержащееся в информационном множестве Ω_T . Значение ε_{T+h} заменяется нулем. Этот алгоритм фактически является последовательной процедурой, поскольку для получения значения $f_{T,h}$ необходимо предварительно вычислить $f_{T,1}, f_{T,2}, \dots, f_{T,h-1}$.

Прогнозирование по модели МА(q). Рассмотрим сначала скользящее среднее первого порядка

$$x_t = \varepsilon_t + b_1 \varepsilon_{t-1},$$

где ε_t — процесс белого шума с $E(\varepsilon_t) = 0$.

Оптимальный прогноз на один шаг вперед равен:

$$f_{T,1} = E(x_{T+1} | \Omega_T) = E(\varepsilon_{T+1} | \Omega_T) + b_1 E(\varepsilon_T | \Omega_T) = b_1 \varepsilon_T.$$

Ошибка этого прогноза равна:

$$e_{T,1} = x_{T+1} - f_{T,1} = \varepsilon_{T+1},$$

так что и здесь последовательность ошибок одношаговых прогнозов образует процесс белого шума. Замечая, что $\varepsilon_T = x_T - f_{T-1,1}$, получаем соотношение

$$f_{T,1} = b_1 (x_T - f_{T-1,1}),$$

дающее возможность последовательно вычислять прогнозы на один шаг вперед, начиная с некоторого начального значения $f_{0,1}$. Это значение является неопределенным, из-за этого приходится ограничиваться моделями скользя-

щего среднего, в которых влияние прошлых значений ослабевает с течением времени, так что конкретный выбор начального значения $f_{0,1}$ не играет существенной роли (если, конечно, значение T не слишком мало). В модели МА(1) такое свойство имеет место при выполнении условия $|b_1| < 1$, а это есть условие обратимости модели, о чем говорилось во второй части учебника. Учитывая, что $E(\varepsilon_t) = 0$, естественно положить $f_{0,1} = 0$.

Если $h > 1$, то оптимальный прогноз на h шагов вперед равен:

$$f_{T,h} = E(x_{T+h} | \Omega_T) = E(\varepsilon_{T+h} | \Omega_T) + b_1 E(\varepsilon_{T+h-1} | \Omega_T) = 0.$$

Заметим, что если рассматривается модель с постоянной составляющей:

$$x_t = \mu + \varepsilon_t + b_1 \varepsilon_{t-1},$$

то оптимальный прогноз на $h > 1$ шагов вперед равен μ , а оптимальный прогноз на один шаг вперед равен $f_{T,1} = \mu + b_1 \varepsilon_T$. При этом соотношение $f_{T,1} = b_1(x_T - f_{T-1,1})$ заменяется на $f_{T,1} = \mu + b_1(x_T - f_{T-1,1})$, и на этот раз естественно положить $f_{0,1} = \mu$.

Рассмотрим теперь модель МА(q):

$$x_t = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}.$$

Поскольку $x_{T+1} = \varepsilon_{T+1} + b_1 \varepsilon_T + \dots + b_q \varepsilon_{T-q+1}$, оптимальный прогноз на один шаг вперед равен:

$$f_{T,1} = b_1 \varepsilon_T + \dots + b_q \varepsilon_{T-q+1},$$

и ошибка этого прогноза равна

$$e_{T,1} = x_{T+1} - f_{T,1} = \varepsilon_{T+1}.$$

Требующиеся для вычисления $f_{T,1}$ значения $\varepsilon_T, \dots, \varepsilon_{T-q+1}$ получаем по формуле $\varepsilon_T = x_T - f_{T-1,1}$. Неопределенность значения $f_{0,1}$ преодолевается аналогично предыдущему: значения $\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{1-q}$ полагаются равными 0, что дает $f_{0,1} = 0$. И опять такой выбор не будет существенно влиять на искомый прогноз, если T достаточно велико и если для рассматриваемой модели выполнено условие обратимости.

Подобная процедура реализуется и при вычислении прогноза на h шагов вперед, где $h \leq q$. В выражении для x_{T+h} все величины, входящие в правую часть, заменяются их оптимальными прогнозами. При этом будущие значения ε_t заменяются нулями, а текущее (ε_T) и прошлые значения ε_t оцениваются по формуле $\varepsilon_T = x_T - f_{T-1,1}$ с выбором начальных значений $\varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{1-q} = 0$, так что $f_{0,1} = 0$. При $h > q$ в выражение для x_{T+h} входят только будущие (по отношению к моменту T) значения ε_t . Поэтому

$$E(x_{T+h} | \Omega_T) = E(\varepsilon_{T+h}) = 0 \text{ для } h > q.$$

Если рассматривается модель с постоянной составляющей:

$$x_t = \mu + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q},$$

тогда при том же выборе $\varepsilon_0 = \varepsilon_{-1} = \dots = \varepsilon_{1-q} = 0$ получаем естественное начальное значение $f_{0,1} = \mu$. В этом случае прогноз на большее, чем q , количество шагов вперед равен μ , т.е. математическому ожиданию процесса x_t .

Прогнозирование по модели ARMA(p, q). Прогнозирование для таких моделей производится в соответствии с алгоритмами, используемыми для моделей AR(p) и MA(q).

Рассмотрим простейший случай — модель ARMA(1, 1) с нулевым математическим ожиданием. Тогда

$$x_{T+1} = a_1 x_T + \varepsilon_{T+1} + b_1 \varepsilon_T,$$

и оптимальный прогноз на один шаг вперед равен

$$f_{T,1} = a_1 x_T + b_1 \hat{\varepsilon}_T,$$

где $\hat{\varepsilon}_T = x_T - f_{T-1,1}$.

Далее:

$$x_{T+2} = a_1 x_{T+1} + \varepsilon_{T+2} + b_1 \varepsilon_{T+1},$$

так что $f_{T,2} = a_1 f_{T,1}$, $f_{T,3} = a_1 f_{T,2} = a_1^2 f_{T,1}$ и т.д. Иначе говоря, оптимальный прогноз здесь или экспоненциально убывает к математическому ожиданию процесса (при $0 < a_1 < 1$), или представляет собой затухающую волну (при $-1 < a_1 < 0$).

✓ **Замечание 5.2.1.** Можно показать (см., например, (Tsay, 2005)), что для произвольного стационарного процесса ARMA(p, q) оптимальный прогноз на h шагов вперед сходится при $h \rightarrow \infty$ к математическому ожиданию этого процесса (свойство «возвращения к среднему» (*mean reverting*)).

Интервальные прогнозы. При использовании точечных прогнозов в реальных ситуациях может наблюдаться значительное отклонение прогнозируемых значений от наблюдаемых в действительности. В связи с этим представляет интерес построение интервальных прогнозов, обеспечивающих заданный уровень доверия.

Пусть

x_t — стационарный временной ряд с нулевым математическим ожиданием, порождаемый моделью ARMA(p, q):

$$x_t = \sum_{j=1}^p a_j x_{t-j} + \sum_{j=0}^q b_j \varepsilon_{t-j},$$

или в компактной форме:

$$a(L)x_t = b(L)\varepsilon_t.$$

Во второй части учебника уже отмечалось, что такой процесс можно представить в виде процесса $MA(\infty)$ — скользящего среднего бесконечного порядка:

$$x_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} = c(B) \varepsilon_t,$$

где $c(z) = \sum_{j=0}^{\infty} c_j z^j = \frac{b(z)}{a(z)}, \quad c_0 = 1.$

Коэффициенты c_1, c_2, \dots можно найти, приравнявая коэффициенты при одинаковых степенях z в левой и правой частях соотношения $a(z)c(z) = b(z)$.

При таком представлении процесса x_t оптимальный прогноз записывается в виде:

$$\begin{aligned} f_{T,h} &= E(X_{T+h} | \Omega_T) = E\left(\sum_{j=0}^{\infty} c_j \varepsilon_{T+h-j} | \Omega_T\right) = \\ &= E\left(\sum_{j=0}^{h-1} c_j \varepsilon_{T+h-j} | \Omega_T\right) + E\left(\sum_{j=h}^{\infty} c_j \varepsilon_{T+h-j} | \Omega_T\right) = \sum_{j=h}^{\infty} c_j \varepsilon_{T+h-j}, \end{aligned}$$

так что ошибка прогноза на h шагов вперед равна:

$$e_{T,h} = x_{T+h} - f_{T,h} = \sum_{j=0}^{h-1} c_j \varepsilon_{T+h-j}.$$

Поскольку слагаемые не коррелированы между собой и их математические ожидания равны нулю, то

$$D(e_{T,h}) = E(e_{T,h}^2) = \sigma_{\varepsilon}^2 \sum_{j=0}^{h-1} c_j^2.$$

В частности, дисперсия ошибки прогноза на один шаг вперед равна:

$$D(e_{T,1}) = \sigma_{\varepsilon}^2.$$

Если инновации ε_t имеют нормальные распределения, то

$$e_{T,h} \sim N\left(0, \sigma_{\varepsilon}^2 \sum_{j=0}^{h-1} c_j^2\right),$$

так что

$$\frac{x_{T+h} - f_{T,h}}{\sigma_{\varepsilon} \sqrt{\sum_{j=0}^{h-1} c_j^2}} \sim N(0, 1),$$

и интервальным прогнозом для x_{T+h} с уровнем доверия $(1 - \alpha)$ является интервал

$$\left(f_{T+h} - z_{1-\frac{\alpha}{2}} \sigma_{\varepsilon} \left(\sum_{j=0}^{h-1} c_j^2 \right)^{\frac{1}{2}}, f_{T+h} + z_{1-\frac{\alpha}{2}} \sigma_{\varepsilon} \left(\sum_{j=0}^{h-1} c_j^2 \right)^{\frac{1}{2}} \right).$$

Как и в случае регрессионных моделей, замена истинного неизвестного значения σ_{ε}^2 на его оценку $\hat{\sigma}_{\varepsilon}^2$ несколько изменяет распределение. При такой замене полученный интервал является приближенным, причем получаемый интервальный прогноз оказывается оптимистическим: он недооценивает действительную изменчивость значений прогноза из-за неучета того факта, что для реальных данных истинная модель не известна, а подбирается и оценивается по этим данным, что должно приводить к ухудшению качества прогноза по сравнению с прогнозом по истинной модели. Указанное ухудшение становится несущественным при оценивании модели по большому количеству наблюдений. Разумеется, производя построение прогноза на будущее, надо всегда иметь в виду, что прогноз может иметь смысл только в предположении, что модель порождения данных не изменится на прогнозном периоде (или изменится несущественно).

Отметим некоторые общие свойства ошибок прогнозов для моделей ARMA.

1. Ошибки одношаговых прогнозов $e_{T,1}, e_{T+1,1}, e_{T+2,1}, \dots$ образуют последовательность не коррелированных между собой случайных величин, тогда как ошибки прогнозов на $h \geq 2$ шагов вперед $e_{T,h}, e_{T+1,h}, e_{T+2,h}, \dots$ в общем случае коррелированы между собой.
2. Из представления $D(e_{T,h}) = \sigma_{\varepsilon}^2 \sum_{j=0}^{h-1} c_j^2$ непосредственно следует, что

$$D(e_{T,h+1}) \geq D(e_{T,h}),$$

так что точность прогноза убывает с удалением в будущее.

3. Если устремить $h \rightarrow \infty$, то получим:

$$\lim_{h \rightarrow \infty} D(e_{T,h}) = \sigma_{\varepsilon}^2 \sum_{j=0}^{\infty} c_j^2.$$

В то же время если использовать $MA(\infty)$ представление для x_t , то из него следует:

$$D(x_t) = E \left(\left(\sum_{j=0}^{\infty} c_j \varepsilon_{t-j} \right)^2 \right) = \sigma_{\varepsilon}^2 \sum_{j=0}^{\infty} c_j^2.$$

Иначе говоря, $\lim_{h \rightarrow \infty} D(e_{T,h}) = D(x_t)$, т.е. условная (при информации Ω_T) дисперсия ошибки «удаленного» прогноза близка к безусловной дисперсии самого процесса.

Прогнозирование для моделей ARIMA. Пусть x_t — процесс типа ARIMA(p, d, q), так что процесс $y_t = (1 - L)^d x_t$, полученный d -кратным дифференцированием процесса x_t , является процессом типа ARMA(p, q). Тогда при построении прогнозов значений ряда x_t естественно использовать прогнозные значения, получаемые для стационарного процесса y_t рассмотренными выше методами. Пусть $f_{T,h}^x$ и $f_{T,h}^y$ — прогнозы на h шагов вперед для процессов x_t и y_t , построенные в момент T . Тогда

$$f_{T,h}^y = (1 - L)^d f_{T,h}^x,$$

где оператор L действует на индекс h , так что, например, при $d = 1$ имеем:

$$f_{T,h}^y = (1 - L)f_{T,h}^x = f_{T,h}^x - f_{T,h-1}^x,$$

$$f_{T,h}^x = f_{T,h-1}^x + f_{T,h}^y.$$

Таким образом, прогнозирование может осуществляться в два этапа: на первом — производится прогноз для стационарного ряда y_t , а на втором — находится прогноз для нестационарного ряда x_t .

Впрочем, можно обойтись и без двухшаговой процедуры. Пусть рассматриваемый процесс удовлетворяет соотношению

$$a(L)(1 - L)^d x_t = b(L)\varepsilon_t,$$

где $a(L)$ и $b(L)$ — многочлены степеней p и q соответственно.

Тогда $A(L) = a(L)(1 - L)^d$ — многочлен степени $P = p + d$,

$$A(L) = 1 - A_1 L - A_2 L^2 - \dots - A_P L^P,$$

$$x_t - \sum_{j=1}^P A_j x_{t-j} = \sum_{j=0}^q b_j \varepsilon_{t-j}, \quad b_0 = 1,$$

так что

$$x_t = \sum_{j=1}^P A_j x_{t-j} + \sum_{j=0}^q b_j \varepsilon_{t-j}, \quad b_0 = 1,$$

и можно производить прогнозирование, используя эту модель непосредственно, как и в стационарном случае.

Дисперсия прогноза на h шагов вперед на этот раз равна:

$$D(e_{T,h}) = \sigma_\varepsilon^2 \sum_{j=0}^{h-1} C_j^2,$$

где C_j находятся приравниванием коэффициентов при одинаковых степенях z в левой и правой частях соотношения $A(z)C(z) = b(z)$, а $C(z) = \sum_{j=0}^{\infty} C_j z^j$, $C_0 = 1$.

Затруднение, связанное с необходимостью оценивать значения ε_t , можно преодолеть, как и ранее, двумя способами. Можно положить значения $\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{1-q}$ равными нулю или использовать для их оценивания процедуру обратного прогноза (*backcasting*), рассмотренную ранее в разд. 1 ч. 2. В большинстве практических ситуаций оба эти подхода дают близкие результаты. Исключениями являются случаи, когда в модели $a(L)(1-L)^d x_t = b(L)\varepsilon_t$ хотя бы один из многочленов $a(z)$ и $b(z)$ имеет корень, близкий к границе единичного круга, или когда временной ряд рассматривается на достаточно коротком периоде.

✓ **Замечание 5.2.2.** Прогнозы будущих значений ряда, получаемые с использованием моделей экспоненциального сглаживания, рассмотренных при изучении темы 5.1, являются одновременно оптимальными прогнозами для соответствующих моделей ARIMA. Например, прогноз по модели простого экспоненциального сглаживания соответствует оптимальному прогнозу по модели $\text{ARIMA}(0, 1, 1) = \text{IMA}(1, 1)$.

Во второй части учебника с использованием пакета *EViews* были построены подходящие модели для некоторых экономических временных рядов. Теперь будем применять построенные модели для прогнозирования будущих значений соответствующих показателей.

ПРИМЕР 5.2.1

В задании 26 второй части учебника предлагалось подобрать стационарную модель ARMA для ряда данных о количестве произведенных всеми отделениями фирмы *General Motors Corp* транспортных средств (грузовиков, легковых машин и автобусов) за период с 1970 по 1990 г. (ряд X_t). Критерий Дики — Фуллера отвергает гипотезу единичного корня для этого ряда (задание 41 из второй части учебника), так что подбор модели можно производить непосредственно для ряда уровней. Коррелограмма ряда указывает на модель AR(3). Диагностика не выявляет нарушений стандартных предположений об ошибках в оцененной AR(3) модели. Незначимость коэффициента при запаздывании на две единицы времени позволяет оставить в модели только запаздывания на одну и три единицы времени; эта же редуцированная модель выбирается и при использовании критерия Шварца. Именно эту модель используем для прогнозирования значений ряда на 1991 и 1992 гг.



Рис. 5.41

Средняя абсолютная процентная ошибка прогноза равна 9.1%. При этом доля ошибки, связанная со смещением прогнозных значений относительно наблюдаемых, равна 0.999, и это хорошо видно на рис. 5.41, где приведены график ряда X_t на периоде с 1970 по 1992 г. и прогнозные значения XF этого ряда. ■

ПРИМЕР 5.2.2

Рассмотрим квартальные данные об объеме валового национального продукта США за период с I квартала 1947 г. по IV квартал 1960 г. (рис. 5.42). Ряд имеет выраженный тренд, так что для идентификации модели порождения ряда его надо остационаривать, рассматривая его как TS - или DS -ряд.

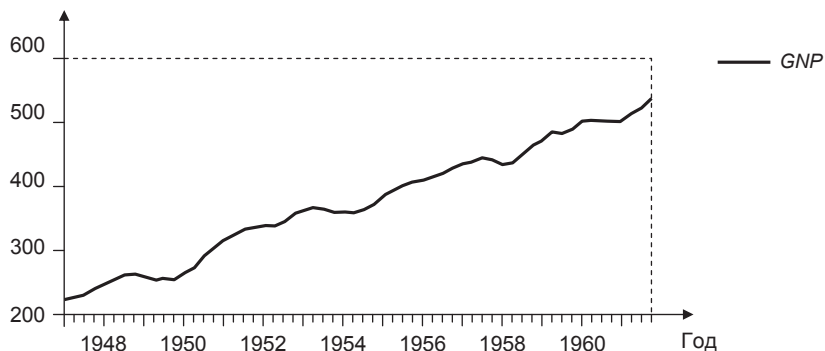


Рис. 5.42

Поскольку данные квартальные, перед применением критерия Дики — Фуллера надо бы выполнить десезонизацию ряда. Однако часто бывают доступными только уже сезонно-скорректированные данные, и именно такими

являются имеющиеся у нас данные. Поэтому пропускаем эту стадию анализа и применяем критерий Дики —Фуллера к самому ряду наблюдений. На основании коррелограммы ряда разностей последний может быть идентифицирован как AR(1) — соответственно в правую часть уравнения Дики — Фуллера включаем одну запаздывающую разность. В правую часть уравнения вводим также константу и тренд. Гипотеза единичного корня отвергается, и это приводит к решению о том, что рассматриваемый ряд стационарен относительно линейного тренда. Но тогда дифференцировать ряд не надо, и следует производить оценивание ряда в уровнях. Результаты оценивания приведены в табл. 5.15.

Таблица 5.15

Оценивание ряда в уровнях
Dependent Variable: GNP; Convergence achieved after 3 iterations

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
<i>C</i>	219.4432	5.490064	39.97097	0.0000
<i>T</i>	5.144091	0.162430	31.66964	0.0000
AR(1)	1.382216	0.112642	12.27091	0.0000
AR(2)	−0.622664	0.112637	−5.528052	0.0000
<i>R-squared</i>	0.997240	<i>Mean dependent var</i>		371.4630
<i>Adjusted R-squared</i>	0.997074	<i>S.D. dependent var</i>		82.32236
<i>S.E. of regression</i>	4.452922	<i>Akaike info criterion</i>		5.896185
<i>Sum squared resid</i>	991.4259	<i>Schwarz criterion</i>		6.043517
<i>Log likelihood</i>	−155.1970	<i>F-statistic</i>		6021.429
<i>Durbin-Watson stat</i>	2.202365	<i>Prob. (F-statistic)</i>		0.000000
<i>Inverted AR Roots</i>	0.69 − 0.38 <i>i</i>	0.69 + 0.38 <i>i</i>		

Оцененная модель проходит диагностические тесты, поэтому используем ее для прогноза (сезонно-скорректированных) объемов *GNP* на 4 квартала 1961 г. При этом средняя абсолютная ошибка прогноза оказывается существенно меньшей, чем в предыдущем примере: *MAPE* = 1.197%, хотя поведение ряда прогнозных значений оставляет желать лучшего: велики значения *Bias Proportion* (0.545) и *Variance Proportion* (0.444), а значение *Covariance Proportion* близко к нулю. Эти цифры отражает график на рис. 5.43.

Между тем естественно возникает вопрос: а что было бы, если бы последовали первоначальной методике Бокса — Дженкинса, т.е., обнаружив тренд в динамике ряда, стали бы этот ряд дифференцировать? При таком подходе получаем ряд разностей *GNP_DIF*, график которого уже не содержит тренда (рис. 5.44).

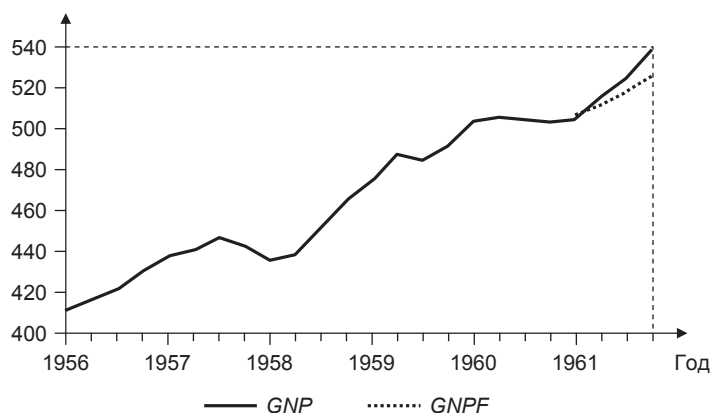


Рис. 5.43

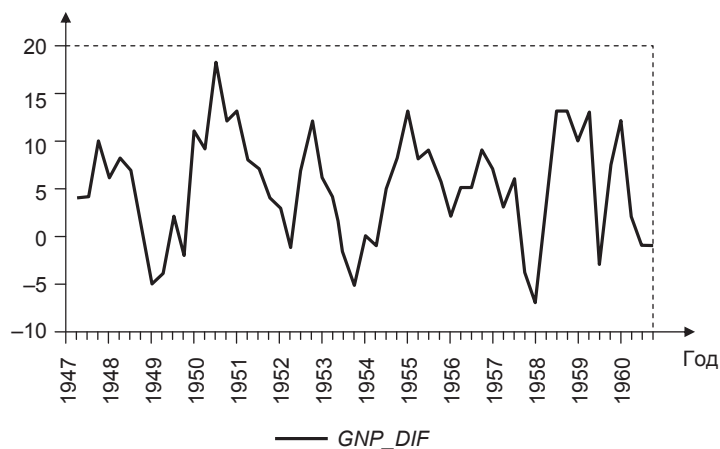


Рис. 5.44

Ряд разностей уже был идентифицирован как AR(1). Поэтому остается оценить AR(1) модель для ряда GNP_DIF . Результат оценивания приведен в табл. 5.16.

Таблица 5.16

Оценка модели для ряда разностейDependent Variable: GNP_DIF ; Sample (adjusted): 1947:3 1960:4

Переменная	Коэффициент	Стандартная ошибка	t -статистика	P -значение
C	4.999118	1.367270	3.656278	0.0006
AR(1)	0.502369	0.121622	4.130578	0.0001
R -squared	0.247050	Mean dependent var		5.092593

Оцененная модель проходит диагностические тесты, поэтому используем ее для построения прогноза ряда GNP на 4 квартала 1961 г. Для полученного в результате прогноза $MAPE = 1.865\%$ в сравнении со значением $MAPE = 1.197\%$ для прогноза, построенного ранее по модели для уровней, так что (с точки зрения $MAPE$) прогноз по самому ряду оказался примерно в 1.5 раза более точным. ■

ПРИМЕР 5.2.3

Рассмотрим данные об объемах производства тканей в Российской Федерации в период 2002—2006 гг. (месячные данные) — ряд $TKANI$ (рис. 5.45). Тренда в поведении ряда нет, он больше похож на стационарный. Если принять эту точку зрения (применение критериев единичного корня ее подтверждает), то нет необходимости в дифференцировании ряда, и можно пытаться идентифицировать модель порождения ряда непосредственно по его коррелограмме. На коррелограмме обращают на себя внимание пики автокорреляционной функции на лагах 12 и 24, отражающие наличие сезонности в представленных данных.

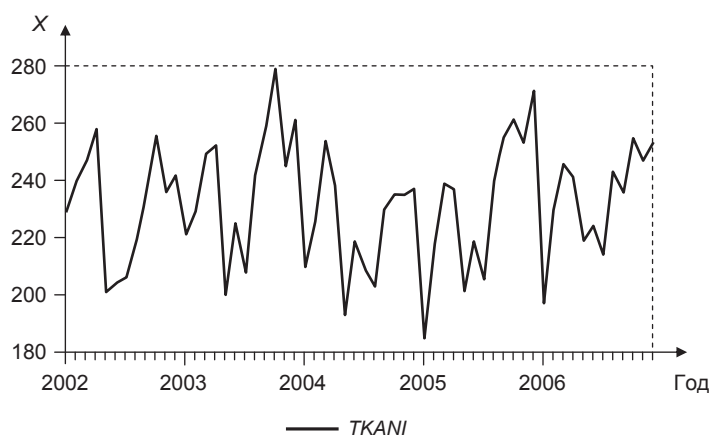


Рис. 5.45

Используя стандартную процедуру подбора стационарной модели $ARMA$ к имеющимся статистическим данным, получим в качестве одного из вариантов модель $AR(10)$ с константой и 11 сезонными дамми-переменными, из которой затем исключаются переменные со статистически незначимыми оценками коэффициентов. Результаты оценивания редуцированной модели приведены в табл. 5.17.

Используем полученную модель для построения прогнозов на 12 месяцев 2007 г. На рис. 5.46 приведены графики ряда $TKANI$ (ряд X) и ряда его прогнозных значений (ряд XF).

Таблица 5.17

Оценка редуцированной модели

Dependent Variable: X; Convergence achieved after 6 iterations

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
C	204.4995	2.522339	81.07533	0.0000
@SEAS(2)	21.45000	4.654558	4.608385	0.0000
@SEAS(3)	40.63456	5.432310	7.480163	0.0000
@SEAS(4)	38.53567	4.286232	8.990569	0.0000
@SEAS(6)	12.02075	5.086706	2.363170	0.0223
@SEAS(8)	25.42078	4.649170	5.467811	0.0000
@SEAS(9)	38.79972	5.059813	7.668212	0.0000
@SEAS(10)	50.89278	4.336347	11.73633	0.0000
@SEAS(11)	39.64055	4.903230	8.084578	0.0000
@SEAS(12)	49.27937	5.229306	9.423692	0.0000
AR(1)	0.231842	0.122870	1.886892	0.0654
AR(3)	0.279801	0.125358	2.232013	0.0304
AR(10)	-0.297668	0.093487	-3.184059	0.0026



Рис. 5.46

На этот раз значение *Covariance Proportion* весьма велико (0.829872), а значения *Bias Proportion* и *Variance Proportion* весьма малы (0.016586 и 0.153542, соответственно). Это говорит о хорошем согласовании поведения ряда прогнозов X_F с поведением наблюдаемого ряда X на прогнозном периоде, что наглядно демонстрирует рис. 5.46. ■

ПРИМЕР 5.2.4

Уровни безработицы в США в период 1987—1990 гг. (месячные данные) образуют ряд $USAUNEMP$, график которого показан на рис. 5.47. Обозначив для краткости $X = USAUNEMP$ и очистив этот ряд от сезонности (с помощью дамми-переменных), получим ряд X_{DES} , для которого производим проверку гипотезы единичного корня. Гипотеза единичного корня не отвергается. Поэтому следует подбирать модель не к самому ряду X , а к ряду его разностей $D(X)$. На коррелограмме этого ряда имеется значимая автокорреляция на лаге 5, поэтому в качестве первоначального варианта подбираем модель $AR(5)$. Диагностика оцененной модели не выявляет нарушений стандартных предположений. Поэтому можно перейти к редукции оцениваемой модели, в ходе которой из ее правой части исключаются константа и составляющая $AR(2)$. На результатах диагностики это не отражается, но в редуцированной модели все оцененные коэффициенты статистически значимы. Следовательно, можно использовать последнюю модель для прогнозирования. Построим прогноз уровней безработицы на первые 10 месяцев 1991 г. Совместный график наблюдаемых и прогнозных значений приведен на рис. 5.48. Здесь значение *Covariance Proportion* достаточно велико (0.582832), но велико и значение *Bias Proportion* (0.404464).

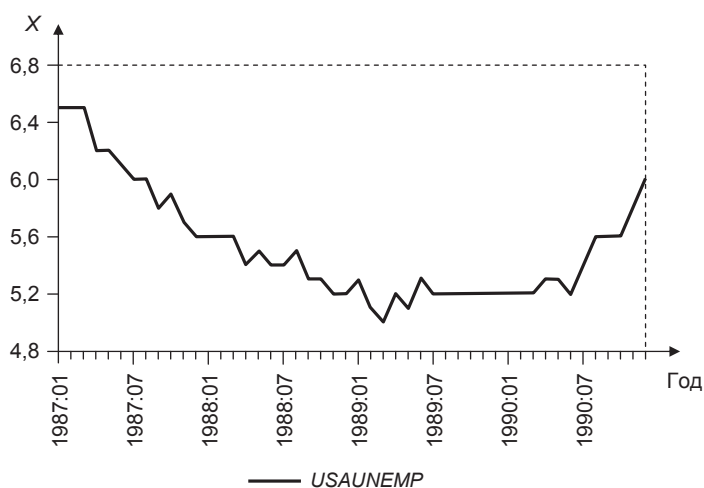


Рис. 5.47

Интересно сравнить рассчитанные прогнозы с адаптивными прогнозами, полученными методом двойного экспоненциального сглаживания и методами Хольта — Винтерса с аддитивной и мультипликативной сезонностью (выбор параметров сглаживания оставляем компьютерной программе). На рис. 5.49 представлены для сравнения график самого ряда X и графики четырех упомянутых прогнозов.

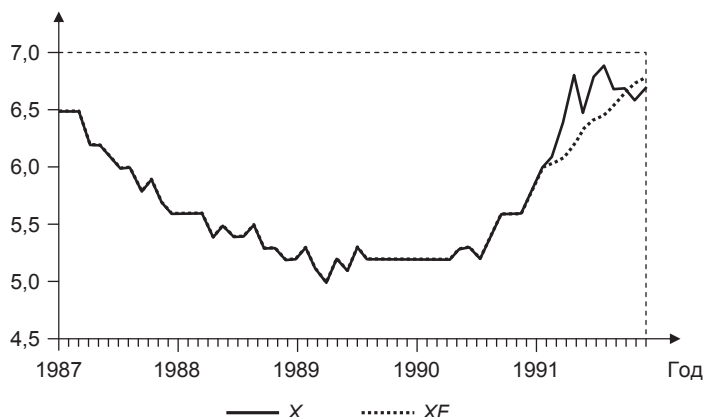


Рис. 5.48

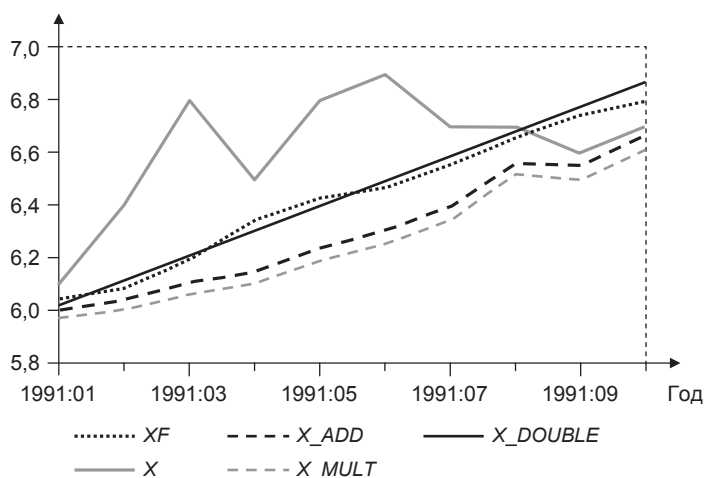


Рис. 5.49

Оба варианта метода Хольта — Винтерса дают худшие прогнозы. Двойное экспоненциальное сглаживание (прогноз по прямой) дает результаты, близкие к полученным по подобранной эконометрической модели ряда. ■

ПРИМЕР 5.2.5

При рассмотрении примера 5.1.7 (в рамках темы 5.1) применяли процедуру простого экспоненциального сглаживания к ряду $x_t = Ind_t$, не задавая заранее значение параметра α , а предоставляя программе самой выбрать оптимальное значение этого параметра. Это привело к значению $\alpha = 0.914$. Там же мы обратили внимание на то, что прогнозные значения \tilde{x}_{t+1} на один шаг вперед ока-

зались практически равными текущим значениям x_t . Попробуем подобрать для этого ряда наблюдений модель типа ARIMA.

Гипотеза единичного корня для ряда x_t не отвергается критериями Дики — Фуллера, Филлипса — Перрона и другими. В то же время гипотеза стационарности не отвергается критерием *KPSS*. Соответственно сделать определенный вывод о наличии или об отсутствии единичного корня здесь невозможно. Поэтому имеет смысл попробовать произвести построение прогнозов на основе как подобранной стационарной модели, так и подобранной *DS*-модели.

Подбор *DS*-модели приводит к модели случайного блуждания $x_t = x_{t-1} + \varepsilon_t$, но для такой модели оптимальный прогноз на любое количество шагов вперед (в том числе и на один шаг вперед), сделанный в момент T и использующий информационное множество Ω_T , совпадает со значением ряда в момент T . Именно это мы и наблюдали в примере 5.1.7.

Таблица 5.18

Результаты для модели в уровнях

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение
$X(-1)$	0.999930	0.001852	540.0148	0.0000
<i>R-squared</i>	0.891236	<i>Mean dependent var</i>		1805.1960

Если подбирать стационарную модель в уровнях, то после исключения переменных со статистически незначимыми оценками коэффициентов получим результаты, приведенные в табл. 5.18. Имеем модель, практически не отличающуюся от случайного блуждания. ■

В примере 5.2.5 оказалось невозможным сделать определенный вывод о наличии или об отсутствии единичного корня в отношении рассматриваемого ряда. Как лучше проводить прогнозирование в таких случаях? Отметим лишь некоторые исследования, проводившиеся в этом направлении.

В работе (*Campbell, Perron, 1991*) моделировались стационарные ряды, близкие к интегрированным, затем производился подбор *TS*- и *DS*-моделей по смоделированным данным, и по этим моделям строились прогнозы поведения рядов на 1 и 20 шагов вперед. Альтернативные модели сравнивались по величине среднеквадратической ошибки прогноза. Полученные результаты говорили о том, что стационарные ряды, близкие к интегрированным, лучше прогнозировать, (ошибочно) считая их интегрированными.

В работе (*Clements, Hendry, 2000*) предпринято исследование качества прогнозов по *DS*- и *TS*-моделям в зависимости от того, соответствует ли используемая для прогнозирования статистическая модель (*SM*) истинному процессу порождения данных (*DGP* — *data generating process*). Статистическое моделирование приводит к следующим результатам.

Изменение DGP в направлении «другой» модели (DS или TS) может приводить при небольших значениях T к тому, что прогнозы по статистической модели, соответствующей DGP (т.е. по «своей» модели), оказываются менее точными. Это особенно важно для практических ситуаций, в которых различие между TS - и DS -моделью бывает довольно затруднительным.

В работе (Diebold, Kilian, 2000) проводится систематическое исследование того, до какой степени предварительное тестирование на наличие единичного корня влияет на качество прогноза при различных степенях инерционности процесса, различных горизонтах прогнозирования и различных объемах выборки. Внимание концентрируется на случае одномерного процесса авторегрессии с трендом и высокой инерционностью, особенно интересном для экономики и финансов. Под **инерционностью** (*persistence*) процесса понимается сохранение влияния реализованного значения инновации в момент t на последующие значения ряда. Для невзрывного $AR(1)$ процесса

$$x_t = \rho x_{t-1} + \varepsilon_t, \quad 0 \leq \rho \leq 1,$$

степень инерционности возрастает при увеличении значения ρ от 0 до 1. При $\rho = 0$ инерционность отсутствует вовсе. При $0 < \rho < 1$ влияние реализованного значения инновации в момент t на последующие значения ряда убывает с продвижением в будущее. При $\rho = 1$ (случайное блуждание) влияние реализованного значения инновации в момент t на последующие значения ряда сохраняется неизменным.

В работе исследуется $AR(1)$ процесс с трендом, представляемый в компонентной форме как сумма линейного тренда T_t и процесса $AR(1)$:

$$y_t = T_t + x_t,$$

где $T_t = a + bt$, $x_t = \rho x_{t-1} + \varepsilon_t$.

Процесс параметризуется в соответствии с квартальными данными о послевоенном реальном ВВП в США: авторы полагают $a = 7$, $b = 0.0065$ и $\sigma = 0.0099$. Рассматриваются значения $\rho \in \{0.5, 0.9, 0.97, 0.99, 1\}$ и $T \in \{25, 30:10:80; 100:20:180; 200:40:1000\}$, включающие подходящие степени инерционности и объемы выборок для годовых, квартальных, месячных, недельных и дневных данных.

Сравниваются три модели прогнозов: $AR(1)$ в уровнях с линейным детерминированным трендом (L — для «уровней»), случайное блуждание со сносом (D — для «разностей») и модель, предлагаемая предварительными тестами Дикки — Фуллера, использующими 5%-е критические значения для конечных выборок (P — для «претестовых» моделей). Общей задачей является предсказание уровней рядов на горизонтах h от 1 до 100 периодов вперед. В качестве характеристики прогноза каждой модели берется безусловная среднеквадратическая ошибка прогноза ($PMSE$), вычисляемая по 20 000 реализаций алгоритма Монте-Карло. Для каждого значения ρ вычисляются отношения $PMSE(D)/PMSE(L)$, $PMSE(D)/PMSE(P)$ и $PMSE(P)/PMSE(L)$ для всех

комбинаций значений h и T . Ниже приведены результаты, полученные для различных пар моделей, между которыми производится выбор:

- $D\&L$ (модель в разностях или модель в уровнях). В этой связке полученные результаты указывают на предпочтительность дифференцирования в случаях, когда размер выборки небольшой или умеренный, а процесс имеет высокую инерционность, и на предпочтительность модели в уровнях в случае малоинерционных процессов или при больших выборках в случае высокоинерционного процесса.
- $D\&P$ (модель в разностях или модель, рекомендуемая на основании результатов предварительного тестирования на наличие/отсутствие единичного корня). В этой связке отношение $PMSE(D)/PMSE(P)$ либо близко к 1, либо превосходит 1, так что использование претеста предпочтительнее, чем безоглядное дифференцирование. Это означает, что при построении моделей с целью прогнозирования не рекомендуется следовать стратегии Бокса — Дженкинса неперменного дифференцирования ряда для достижения стационарности.
- $L\&P$ (модель в уровнях или модель, рекомендуемая на основании результатов предварительного тестирования на наличие/отсутствие единичного корня). В этой связке отношение $PMSE(P)/PMSE(L)$ не везде меньше 1, однако преимущество претестового подхода проявляется на большей части пространства эксперимента. Это также ставит под сомнение опирающуюся на асимптотические результаты стратегию неперменного построения прогнозируемых моделей в уровнях.

Далее были дополнительно смоделированы ряды, имитирующие:

- ВВП на душу населения США — годовые данные;
- реальный ВВП США — квартальные данные;
- индекс промышленного производства США — месячные данные;
- индекс Доу — Джонса — дневные данные.

Поскольку предварительное тестирование явно лучше, чем обязательное дифференцирование, упор делается на сравнении метода предварительного тестирования (метод P) и метода обязательного использования модели в уровнях (метод L). Предложенные рекомендации:

- для годовых данных ($T = 40 - 160$ и $h = 1 - 100$) — использовать предварительное тестирование при наличии не более 70 годовых данных, а для большего количества наблюдений — при очень высокой инерционности ряда;
- для квартальных данных ($T = 80 - 200$ и $h = 1 - 16$) — использовать предварительное тестирование для процессов с корнями 0.97 и выше и метод L при меньших значениях корней;
- для месячных данных ($T = 240 - 480$ и $h = 1 - 48$) — предварительное тестирование имеет смысл только для процессов с корнями, равными 0.99 или выше. Во всех других случаях метод L предпочтительнее;

- для дневных данных ($T = 360 - 720$ и $h = 1 - 90$) — предварительное тестирование имеет смысл, только если ряд в высшей степени инерционен, а именно имеет значения корней 0.99 и выше. В противном случае следует предпочесть метод L .

Указанные рекомендации могут показаться порочным кругом, поскольку они предполагают знание корня DGP . На практике точечные OLS -оценки корней макроэкономических рядов обычно превышают 0.97 для годовых данных, превышают 0.99 для квартальных данных и значительно превышают 0.99 для дневных данных. Более того, наличие смещения OLS -оценок при малых выборках если и приводит, то только к недооценке истинных значений корней. Отсюда можно сделать вывод, что фактически предварительное тестирование рекомендуется при проведении реального прогнозирования макроэкономических рядов, имеющих трендовое поведение.

- ✓ **Замечание 5.2.3.** При сравнении TS - и DS -моделей авторы указанных выше работ не занимались уточнением спецификации этих моделей в направлении исключения из них переменных со статистически незначимыми коэффициентами. И это имеет некоторую аргументацию.

Если анализируется качество прогнозов на один шаг вперед, то оцененная полная модель, обладая наименьшей остаточной суммой квадратов по сравнению с оцененными редуцированными моделями, полученными на ее основе, имеет наилучшие показатели качества одношаговых прогнозов в пределах периода, на котором эта модель оценивалась, по крайней мере, в отношении $RMSE$. В то же время (см., например, работу (Энгов и др., 2002)), указанное преимущество полной модели никак не гарантирует того, что полная модель обязательно даст лучшее качество одношаговых прогнозов при выходе за пределы интервала, на котором модель оценивалась. Однако (см. ту же работу) и выбор по информационному критерию Шварца не гарантирует того, что в рамках DS - или TS -моделей выбранная по этим критериям из нескольких вариантов модель обязательно даст лучшие результаты прогнозов за пределами интервала, по которому эти модели оценивались. Выбор для сравнения качества прогнозов именно «полных» TS - и DS -моделей оправдан лишь с точки зрения определенности выбора модели.

- ✓ **Замечание 5.2.4.** В практической работе при поступлении свежих данных приходится решать вопрос о том, как производить прогнозирование на будущий период: с учетом всех имеющихся на данный момент данных (включая вновь поступившие) или со сдвигом выборки (добавлением в нее «самого свежего» значения и отбрасыванием «самого старого» по времени значения). В первом случае говорят о **рекурсивном методе** (*recursive forecast*), а во втором — о **методе**

скользящего окна (*rolling forecast*). В рекурсивном методе предпрогнозный период увеличивается по мере увеличения данных. В методе скользящего окна длина предпрогнозного периода постоянна; при получении новых статистических данных старые данные исключаются, что приводит к появлению некоторого подобия окна, через которое просматриваются данные и которое последовательно смещается вперед по оси времени.

При применении рекурсивного метода не происходит потери информации. Однако в периоды экономической неустойчивости и кризисов важно быстро адаптировать модель к изменившимся условиям, и с этой точки зрения предпочтительнее выглядит метод скользящего окна. В то же время использование «слишком узкого» окна, включающего мало наблюдений, приводит к некачественному оцениванию параметров модели, что, в свою очередь, может приводить к неудовлетворительным прогнозам. Вопрос о выборе между указанными двумя методами рассматривался многими авторами. Мы отметим здесь лишь работу (*Энтов и др.*, 2002), в которой подобное исследование проводилось на примере нескольких макроэкономических рядов Российской Федерации.

Сравнение качества прогнозов, полученных по различным моделям

Пусть

$\tilde{x}_{1t}, \tilde{x}_{2t}, t = 1, \dots, h$ — прогнозные значения временного ряда x_t , полученные на основе моделей 1 и 2;

$g(x_t, \tilde{x}_{it})$ — функция потерь,

$$g(x_t, \tilde{x}_{it}) = g(e_{it}),$$

где e_{it} — ошибка прогноза, $e_{it} = x_t - \tilde{x}_{it}$, $i = 1, 2$.

Чтобы определить, предсказывает ли одна из моделей лучше другой, можно проверить нулевую гипотезу об отсутствии различий между прогнозными свойствами двух моделей $H_0: E[g(e_{1t})] = E[g(e_{2t})]$ против альтернативной гипотезы $H_A: E[g(e_{1t})] \neq E[g(e_{2t})]$. Диболд и Мариано (*Diebold, Mariano, 1995*) предложили проверять эквивалентную H_0 гипотезу о равенстве нулю среднего уровня разности между функциями потерь сравниваемых моделей:

$$H_0: E(d_t) = 0, \quad \text{где} \quad d_t = g(e_{1t}) - g(e_{2t}).$$

Предположим, что последовательность d_t стационарна и представима в форме процесса скользящего среднего $d_t = \mu + \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}$, где $b_0 = 1$, $\sum_{j=0}^{\infty} |b_j| < \infty$.

Тогда¹ при $h \rightarrow \infty$ выполняется соотношение $\sqrt{h}(\bar{d} - \mu) \xrightarrow{d} N(0, \sigma^2)$, где $\bar{d} = \frac{1}{h} \sum_{t=1}^h d_t$, $\sigma^2 = \gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_d(k)$ — долговременная дисперсия² ряда d_t , $\gamma(k) = \text{Cov}(d_t, d_{t-k})$ — автоковариации этого ряда. При этом

$$\frac{\bar{d} - \mu}{\left(\frac{\sigma^2}{h}\right)^{\frac{1}{2}}} \xrightarrow{d} N(0, 1),$$

так что для проверки гипотезы $E(d_t) = 0$ можно использовать статистику **Диболда — Мариано** (*DM — Diebold-Mariano statistics*):

$$DM = \frac{\bar{d}}{\left(\frac{\hat{\sigma}^2}{h}\right)^{\frac{1}{2}}},$$

где $\hat{\sigma}^2$ — состоятельная оценка долговременной дисперсии.

Статистика Диболда — Мариано представляет собой t -статистику для проверки гипотезы $H_0: \mu = 0$ в статистической модели $d_t = \mu + u_t$, $t = 1, 2, \dots, h$, с коррекцией стандартной ошибки оценки параметра для учета автокоррелированности ряда d_t . Для практического ее применения следует использовать критические значения, определяемые по таблице стандартного нормального распределения. Разумеется, ориентироваться на результаты проверки нулевой гипотезы здесь можно только в асимптотическом плане, т.е. для достаточно длинных последовательностей прогнозов.

Разложение Бевеиджа — Нельсона

В теме 1.1 этого раздела уже говорилось о некоторых возможных разложениях (декомпозициях) временного ряда на несколько компонент. Еще одно разложение можно получить, предполагая, что рассматриваемый ряд x_t принадлежит классу $\text{ARIMA}(p, 1, q)$.

Пусть ряд x_t порождается моделью $\text{ARIMA}(0, 1, 1)$:

$$\Delta x_t = \alpha + \varepsilon_t + \gamma \varepsilon_{t-1}, \quad t = 1, 2, \dots,$$

где $\varepsilon_t \sim i.i.d.$ случайные величины.

¹ См. (Андерсон Т., 1976, гл. 7).

² Определение долговременной дисперсии ряда было дано в разд. 10 ч. 2 учебника. Там же рассматривался вопрос о состоятельном оценивании долговременной дисперсии.

Для простоты пусть $x_0 = 0$, $\varepsilon_0 = 0$. Тогда

$$\begin{aligned} x_t &= \alpha + x_{t-1} + \varepsilon_t + \gamma \varepsilon_{t-1} = \alpha + (\alpha + x_{t-2} + \varepsilon_{t-1} + \gamma \varepsilon_{t-2}) + \varepsilon_t + \gamma \varepsilon_{t-1} = \\ &= \alpha t + \sum_{i=1}^t \varepsilon_i + \gamma \sum_{j=1}^{t-1} \varepsilon_j = \alpha t + (1 + \gamma) \sum_{i=1}^t \varepsilon_i - \gamma \varepsilon_t. \end{aligned}$$

Обозначив

$$DT_t = \alpha t, \quad ST_t = (1 + \gamma) \sum_{i=1}^t \varepsilon_i, \quad C_t = -\gamma \varepsilon_t,$$

получим представление ряда x_t в виде:

$$x_t = DT_t + ST_t + C_t = DT_t + Z_t,$$

где DT_t — детерминированный тренд в x_t ;

Z_t — стохастическая компонента ряда, являющаяся суммой двух компонент: стохастического тренда ST_t и циклической компоненты C_t .

Циклическая компонента представляет собой стационарный случайный процесс с нулевым средним. Стохастический тренд включает все случайные шоки (инновации) от ε_1 до ε_t , которые имеют перманентное влияние на уровень x_t . Сумма $x_t^p = DT_t + ST_t$ представляет итоговый тренд и является **перманентной компонентой** (*permanent component*) ряда x_t , а C_t — **транзитивной компонентой** (*transitive component*) этого ряда, так что

$$x_t = x_t^p + C_t.$$

Перманентную компоненту можно представить в виде:

$$\begin{aligned} x_t^p &= \alpha t + (1 + \gamma) \sum_{i=1}^t \varepsilon_i = \alpha + \alpha(t-1) + (1 + \gamma) \sum_{i=1}^{t-1} \varepsilon_i + (1 + \gamma) \varepsilon_t = \\ &= \alpha + x_{t-1}^p + (1 + \gamma) \varepsilon_t, \end{aligned}$$

т.е. перманентная компонента представляет собой случайное блуждание со сносом.

Сгенерируем реализацию процесса ARIMA(0, 1, 1), полагая $\alpha = 0.01$, $\gamma = 0.3$ и $\varepsilon_t \sim i.i.d. N(0, 0.05^2)$. Тогда перманентная компонента суть случайное блуждание с параметром сноса $\alpha = 0.01$ и инновациями $1.3\varepsilon_t$, а транзитивная компонента равна $-0.3\varepsilon_t$. Сгенерированная реализация и ее компоненты представлены на рис. 5.50.

Декомпозиция $Z_t = ST_t + C_t$ стохастической компоненты ряда на стохастический тренд ST_t и циклическую компоненту C_t , полученная выше, может быть обобщена на любую модель типа ARIMA(p , 1, q). Беверидж и Нельсон (*Beveridge, Nelson, 1981*) показали, что каждый такой ряд можно представить

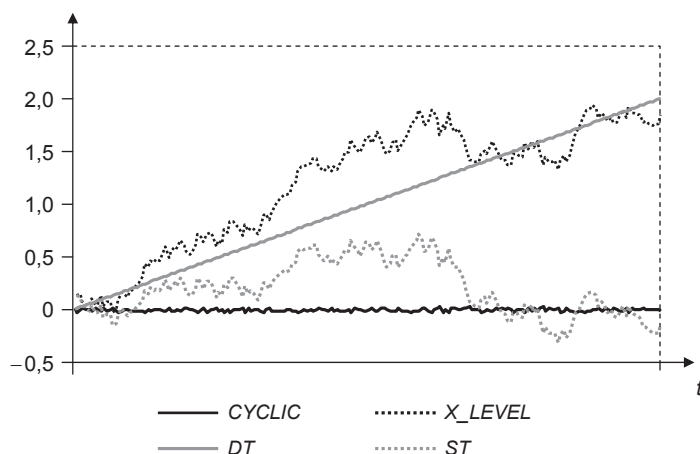


Рис. 5.50

в виде суммы стохастического тренда и стационарной компоненты. Такое представление и называется **разложением (декомпозицией) Бевриджа — Нельсона**.

Пусть стохастическая компонента ряда порождается моделью

$$A(L)Z_t = B(L)\varepsilon_t, \quad t = 1, 2, \dots,$$

где $A(L)$ и $B(L)$ — полиномы порядков p и q соответственно, причем полином $A(L)$ имеет единичный корень, так что $A(L) = (1 - L)A^*(L)$, и все корни полинома $A^*(L)$ лежат за пределами единичного круга на комплексной плоскости.

Тогда

$$(1 - L)A^*(L)Z_t = A^*(L)(1 - L)Z_t = A^*(L)\Delta Z_t,$$

и для ряда первых разностей стохастической компоненты выполняется соотношение

$$A^*(L)\Delta Z_t = B(L)\varepsilon_t.$$

Отсюда получаем:

$$\Delta Z_t = A^*(L)^{-1}B(L)\varepsilon_t = \psi(L)\varepsilon_t = (\psi(1) + (1 - L)\psi^*(L))\varepsilon_t,$$

где $\psi^*(L) = (1 - L)^{-1}(\psi(L) - \psi(1))$.

Применяя к обеим частям последнего соотношения оператор $(1 - L)^{-1}$, приходим к представлению

$$Z_t = \psi(1) \sum_{i=1}^t \varepsilon_i + \psi^*(L)\varepsilon_t = ST_t + C_t.$$

При выводе этого представления предполагается, что $Z_0 = 0$ и $\varepsilon_0 = 0$. В противном случае (если порождение ряда начинается в момент $t = 1$) надо учесть еще и начальные значения:

$$Z_t = Z_0 + \psi(1) \sum_{i=1}^t \varepsilon_i + \psi^*(L) \varepsilon_t - \psi^*(L) \varepsilon_0.$$

Передифференцированные одномерные временные ряды

В литературе по временным рядам часто присутствует предостережение о возможном получении в результате дифференцирования **передифференцированного** ряда (*overdifferenced series*). Такая ситуация возникает, например, если берутся разности ряда, который уже является стационарным.

Известно, что многие представляющие интерес модели стационарных случайных процессов обладают представлением Вольда в виде скользящего среднего (*MAR* — *moving average representation*), но также известно, что многие представляющие интерес модели стационарных случайных процессов не имеют авторегрессионного представления (*ARR* — *autoregressive representation*).

Пусть стационарный процесс x_t с нулевым математическим ожиданием имеет представление Вольда

$$x_t = \varepsilon_t + b_1 \varepsilon_{t-1} + b_2 \varepsilon_{t-2} + \dots = (1 + b_1 L + b_2 L^2 + \dots) \varepsilon_t,$$

где $\sum_{j=1}^{\infty} b_j^2 < \infty$.

Если с ростом j коэффициенты b_j достаточно быстро убывают по абсолютной величине, то процесс x_t можно аппроксимировать процессом скользящего среднего конечного (хотя, может, и достаточно большого) порядка q :

$$x_t = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q} = (1 + b_1 L + \dots + b_q L^q) \varepsilon_t = b(L) \varepsilon_t,$$

где $b(L) = \sum_{j=1}^q b_j L^j$.

При этом если все q корней уравнения $b(z) = 0$ лежат вне единичного круга $|z| \leq 1$ (условие обратимости), то существует эквивалентное представление *MA*(q) процесса x_t в виде процесса авторегрессии бесконечного порядка *AR*(∞):

$$x_t = \sum_{j=1}^{\infty} d_j x_{t-j} + \varepsilon_t,$$

или

$$d(L) x_t = \varepsilon_t,$$

где

$$d(L) = 1 - \sum_{j=1}^{\infty} d_j L^j = \frac{1}{b(L)} \quad \text{и} \quad \sum_{j=1}^{\infty} d_j^2 < \infty.$$

Отсюда вытекает, что при выполнении условия обратимости стационарный процесс $MA(q)$ можно аппроксимировать процессом авторегрессии $AR(p)$:

$$x_t = \sum_{j=1}^p d_j x_{t-j} + \varepsilon_t$$

(быть может, достаточно высокого порядка p). Прогноз значения x_{T+1} по наблюдениям x_1, x_2, \dots, x_T определяется в таком случае как условное ожидаемое значение x_{T+1} для аппроксимирующей модели $AR(T)$:

$$E(x_{T+1} | x_T, x_{T-1}, \dots, x_1) = \sum_{j=1}^T d_j x_{T+1-j}$$

и является линейной проекцией случайной величины x_{T+1} на линейное подпространство случайных величин, порожденное случайными величинами x_1, x_2, \dots, x_T .

В качестве примера рассмотрим процесс $MA(1)$

$$x_t = \varepsilon_t + b_1 \varepsilon_{t-1} = (1 + b_1 L) \varepsilon_t, \quad (5.1)$$

где $\varepsilon_t \sim i.i.d.$, с нулевым средним и конечной дисперсией.

Этот процесс стационарный, и его представление Вольда и есть уравнение (5.1). Если $|b_1| < 1$, то существует авторегрессионное представление (ARR) ряда x_t в виде:

$$x_t = \sum_{j=1}^{\infty} d_j x_{t-j} + \varepsilon_t.$$

Коэффициенты d_1, d_2, \dots можно найти, обращая оператор $(1 + b_1 L)$:

$$(1 + b_1 L)^{-1} x_t = \varepsilon_t,$$

так что

$$(1 - b_1 L + b_1^2 L^2 - \dots) x_t = \varepsilon_t,$$

откуда

$$x_t = 1 + b_1 x_{t-1} - b_1^2 x_{t-2} + \dots + \varepsilon_t.$$

При $|b_1| < 1$ коэффициенты

$$d_j = (-1)^{j+1} b_1^j$$

быстро убывают по абсолютной величине, что дает возможность аппроксимировать обратимый $MA(1)$ процесс процессом авторегрессии не очень высокого порядка.

Рассмотрим, однако, процесс

$$x_t = \varepsilon_t - \varepsilon_{t-1} = (1-L)\varepsilon_t,$$

который получается в результате дифференцирования процесса белого шума:

$$x_t = \Delta \varepsilon_t.$$

Поскольку у этого процесса $b_1 = -1$, то в качестве наивного «кандидата» на ARR в этом случае хотелось бы взять процесс $AR(\infty)$ с $d_j = -1$. Но тогда

$\sum_{j=1}^{\infty} d_j^2 = \infty$, и авторегрессионное представление для ряда x_t не определено.

Если все же использовать в качестве наивного прогноза значение, соответствующее $d_j = -1$, т.е.

$$\sum_{j=1}^T d_j x_{T+1-j} = -\sum_{j=1}^T x_{T+1-j} = -(x_1 + x_2 + \dots + x_T)$$

(заметим, что правая часть равна $\varepsilon_0 - \varepsilon_T$), то такой прогноз не будет оптимальным. Ошибка этого прогноза равна:

$$x_{T+1} - [-(x_1 + x_2 + \dots + x_T)] = (\varepsilon_{T+1} - \varepsilon_T) - (\varepsilon_0 - \varepsilon_T) = \varepsilon_{T+1} - \varepsilon_0,$$

а дисперсия ошибки прогноза равна

$$D(\varepsilon_{T+1} - \varepsilon_0) = 2\sigma^2,$$

что примерно в 2 раза больше, чем у оптимального прогноза. Сам оптимальный прогноз является линейной проекцией случайной величины x_{T+1} на линейное подпространство случайных величин, порожденное случайными величинами x_1, x_2, \dots, x_T , а эта проекция в данном случае равна¹:

$$-\left(\frac{T-1}{T}x_T + \frac{T-2}{T}x_{T-1} + \frac{T-3}{T}x_{T-2} + \dots + \frac{1}{T}x_1\right).$$

Дисперсия ошибки оптимального прогноза равна $\sigma^2 \frac{T+2}{T+1}$ и стремится к σ^2 при $T \rightarrow \infty$.

Мы предположили, что знаем значение коэффициента b_1 и оно равно $b_1 = -1$. В реальных ситуациях значение b_1 неизвестно и необходимо оценивать его по наблюдениям x_1, x_2, \dots, x_T . Рассмотрим, какие особенности возникают здесь в случае $b_1 = -1$.

Смоделируем для этого реализацию стандартного гауссовского белого шума $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{50}$ и по этой реализации построим ряд $x_t = \varepsilon_t - \varepsilon_{t-1}$ ($t = 2, \dots, 50$).

¹ См. (Hamilton, 1994, Chapter 4).

Заметим, что если дисперсия первого ряда оценивается по выборке как $(1.01)^2$, то дисперсия продифференцированного ряда оценивается как $(1.49)^2$.

Коррелограмма продифференцированного ряда явно указывает на его $MA(1)$ характер.

Sample: 1 50; Included observations: 49

<i>Autocorrelation</i>	<i>Partial Correlation</i>	<i>k</i>	<i>AC</i>	<i>PAC</i>	<i>Q-статистика</i>	<i>P-значение</i>
****	****	1	-0.494	-0.494	12.717	0.000
	***	2	-0.014	-0.341	12.727	0.002
	*	3	0.060	-0.176	12.922	0.005
*	**	4	-0.085	-0.209	13.328	0.010
*	*	5	0.095	-0.071	13.845	0.017
	*	6	-0.044	-0.058	13.957	0.030
	*	7	-0.015	-0.063	13.970	0.052
	*	8	-0.052	-0.166	14.134	0.078
*		9	0.096	-0.049	14.709	0.099
*	*	10	-0.091	-0.135	15.237	0.124
*		11	0.121	0.032	16.197	0.134
*	*	12	-0.135	-0.111	17.434	0.134

Близкое к граничному (-0.5) значение автокорреляции на лаге 1 указывает на близость коэффициента b_1 к значению -1 .

Оценивание $MA(1)$ модели с приравниванием «начального» значения ε_1 нулю дает результаты, приведенные в табл. 5.19. Оценка коэффициента b_1 существенно отличается от значения -1 .

Таблица 5.19

Оценка модели с приравниванием $\varepsilon_1 = 0$

Sample (adjusted): 2 50; Included observations: 49 after adjusting endpoints;
Convergence achieved after 6 iterations; Backcast: OFF

Переменная	Коэффициент	Стандартная ошибка	<i>t</i> -статистика	<i>P</i> -значение
MA(1)	-0.844299	0.078206	-10.79589	0.0000
<i>R-squared</i>	0.411521	<i>Mean dependent var</i>		-0.048304
<i>Adjusted R-squared</i>	0.411521	<i>S.D. dependent var</i>		1.490832
<i>S.E. of regression</i>	1.143653	<i>Akaike info criterion</i>		3.126530
<i>Sum squared resid</i>	62.78125	<i>Schwarz criterion</i>		3.165138
<i>Log likelihood</i>	-75.59997	<i>Durbin-Watson stat</i>		1.903930
<i>Inverted MA Roots</i>	0.84			

Заметим в связи с этим, что в смоделированной реализации $\varepsilon_1 = 1.540$, так что выбранное начальное значение сильно отличается от реального. Если бы в действительности было $\varepsilon_1 = 0$, то это привело бы к совершенно другому результату (табл. 5.20).

Таблица 5.20

Оценка модели при $\varepsilon_1 = 0$

Dependent Variable: X0; Sample (adjusted): 2 50; Included observations: 49 after adjusting endpoints; Convergence achieved after 9 iterations; Backcast: OFF

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
MA(1)	−1.037445	0.021529	−48.18885	0.0000
<i>R-squared</i>	0.546126	<i>Mean dependent var</i>		−0.016868
<i>Adjusted R-squared</i>	0.546126	<i>S.D. dependent var</i>		1.459325
<i>S.E. of regression</i>	0.983150	<i>Akaike info criterion</i>		2.824087
<i>Sum squared resid</i>	46.39601	<i>Schwarz criterion</i>		2.862695
<i>Log likelihood</i>	−68.19012	<i>Durbin-Watson stat</i>		2.112057
<i>Inverted MA Roots</i>	1.04			
<i>Estimated MA process is noninvertible</i>				

Полученные результаты отражают тот факт, что при значениях b_1 , близких по абсолютной величине к 1, выбор начального значения может существенно влиять на результаты оценивания.

Впрочем, результаты оценивания можно значительно улучшить, если не приравнивать начальное значение 0, а оценивать его методом «обратного прогнозирования» (*backcasting*), о чем уже говорилось во второй части учебника. При использовании этой процедуры оценивание модели по первоначальным данным (с $\varepsilon_1 = 1.540$) дает удовлетворительный результат (табл. 5.21).

Таблица 5.21

Оценка модели при $\varepsilon_1 = 1.540$

Dependent Variable: X; Sample (adjusted): 2 50; Included observations: 49 after adjusting endpoints; Convergence achieved after 7 iterations; Backcast: 1

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
MA(1)	−0.970547	0.016617	−58.40841	0.0000
<i>R-squared</i>	0.538634	<i>Mean dependent var</i>		−0.048304
<i>Adjusted R-squared</i>	0.538634	<i>S.D. dependent var</i>		1.490832
<i>S.E. of regression</i>	1.012632	<i>Akaike info criterion</i>		2.883180
<i>Sum squared resid</i>	49.22032	<i>Schwarz criterion</i>		2.921788
<i>Log likelihood</i>	−69.63790	<i>Durbin-Watson stat</i>		2.128541
<i>Inverted MA Roots</i>	0.97			

В заключение заметим, что проблемы прогнозирования временных рядов отражены в большом количестве статей, монографий и разделов монографий, посвященных статистическому анализу временных рядов. Перечислим некоторые из них, не упомянутые выше: (Box, Jenkins, Reinsel, 2008), (Clements, Hendry, 1996; 1998a; 1998b; 2001), (Dickey, Bell, Miller, 1986), (Diebold, 1998), (Enders, 1995), (Granger, Newbold, 1986), (Gregory, Smith, Yetman, 2001), (Hamilton, 1994), (Harvey, Leybourne, Newbold, 1998), (Kumar, Gill, 2000), (Snyder, Ord, Koehler, 2001), (Stock, 1996), (Stock, Watson, 1996), (Swanson, White, 1997), (Wallis, Whitley, 1991), (West, 2001).

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что понимается под оптимальным прогнозом?
2. Как производится прогнозирование по модели $AR(p)$?
3. Как производится прогнозирование по модели $MA(q)$?
4. Как производится прогнозирование по модели $ARMA(p, q)$?
5. Как строится интервальный прогноз?
6. Как изменяется точность прогноза с удалением в будущее?
7. Как производится прогнозирование по моделям $ARIMA$?
8. Как лучше проводить прогнозирование при невозможности сделать определенный вывод о наличии или об отсутствии единичного корня в отношении рассматриваемого ряда?
9. Как влияет на качество прогнозов оставление в оцененной модели запаздываний со статистически незначимыми оценками коэффициентов?
10. В чем состоят преимущества и недостатки метода рекурсивных прогнозов и метода скользящего окна?
11. Как можно с помощью статистических критериев сравнивать качество прогнозов, полученных по различным моделям?
12. Что представляет собой разложение Бевериджа — Нельсона?
13. Какие проблемы возникают при прогнозировании передифференцированного ряда?

Р а з д е л 6

МЕТОДОЛОГИЯ ВЕКТОРНЫХ АВТОРЕГРЕССИЙ

Тема 6.1

ПРОГНОЗИРОВАНИЕ ПО МОДЕЛИ ВЕКТОРНОЙ АВТОРЕГРЕССИИ, ПРОВЕРКА НАЛИЧИЯ ПРИЧИННОСТИ ПО ГРЕЙНДЖЕРУ ДЛЯ ДВУХ РЯДОВ И БОЛЕЕ

В рамках моделей векторных авторегрессий (VAR) довольно часто производят проверку гипотезы о том, что некоторая из переменных, образующих VAR, является **причиной по Грейнджеру (G-причиной)** для некоторой другой переменной, входящей в эту же VAR.

Рассмотрим пару переменных (два временных ряда) x_t и y_t . Пусть:

Ω_T^x — информационное множество, содержащее значения x_T, x_{T-1}, \dots ;

Ω_T^y — информационное множество, содержащее значения y_T, y_{T-1}, \dots ;

$\Omega_T^{x,y}$ — объединение этих двух множеств;

$\hat{y}_{T+h} | \Omega = E(y_{T+h} | \Omega)$ — оптимальный прогноз ряда y_t на h шагов вперед относительно информационного множества Ω , где в качестве Ω могут выступать информационные множества Ω_T^y или $\Omega_T^{x,y}$.

Говорят, что **переменная x не является причиной по Грейнджеру для переменной y** (x is *Granger-noncausal* for y), если

$$E \left| y_{T+h} - (\hat{y}_{T+h} | \Omega_T^y) \right|^2 \leq E \left| y_{T+h} - (\hat{y}_{T+h} | \Omega_T^{x,y}) \right|^2 \quad \text{для } h = 1, 2, \dots,$$

так что дополнительная информация о значениях x_T, x_{T-1}, \dots не улучшает оптимальный прогноз для y_{T+h} — для его построения достаточно использовать только значения y_T, y_{T-1}, \dots . Если же такая информация улучшает оптимальный прогноз, то **переменная x является причиной по Грейнджеру для переменной y** (x *Granger-causes* y).

Пусть поведение переменных x и y описывается двумерной VAR

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^{(1)} & \pi_{12}^{(1)} \\ \pi_{21}^{(1)} & \pi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} \pi_{11}^{(p)} & \pi_{12}^{(p)} \\ \pi_{21}^{(p)} & \pi_{22}^{(p)} \end{pmatrix} \begin{pmatrix} x_{t-p} \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \\ = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \Pi_1 \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \dots + \Pi_p \begin{pmatrix} x_{t-p} \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Случайные величины ε_{1t} , ε_{2t} образуют случайный вектор $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})^T$, компоненты которого не коррелированы по времени и не коррелированы с запаздывающими значениями переменных: ε_t — **вектор инноваций (обновлений) относительно информационного множества**

$$\Omega_t^{x,y} = (y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, \dots).$$

Указанную модель можно записать в виде:

$$\begin{aligned} x_t &= c_1 + \pi_{11}^{(1)} x_{t-1} + \dots + \pi_{11}^{(p)} x_{t-p} + \pi_{12}^{(1)} y_{t-1} + \dots + \pi_{12}^{(p)} y_{t-p} + \varepsilon_{1t}, \\ y_t &= c_2 + \pi_{21}^{(1)} x_{t-1} + \dots + \pi_{21}^{(p)} x_{t-p} + \pi_{22}^{(1)} y_{t-1} + \dots + \pi_{22}^{(p)} y_{t-p} + \varepsilon_{2t}. \end{aligned}$$

Отсюда видно, что оптимальные прогнозы переменных x_t , y_t на один или несколько шагов вперед задаются формулами:

$$\begin{aligned} \hat{x}_{t+1} \Big| \Omega_t^{x,y} &= c_1 + \pi_{11}^{(1)} x_t + \dots + \pi_{11}^{(p)} x_{t-p+1} + \pi_{12}^{(1)} y_t + \dots + \pi_{12}^{(p)} y_{t-p+1}, \\ \hat{y}_{t+1} \Big| \Omega_t^{x,y} &= c_2 + \pi_{21}^{(1)} x_t + \dots + \pi_{21}^{(p)} x_{t-p+1} + \pi_{22}^{(1)} y_t + \dots + \pi_{22}^{(p)} y_{t-p+1}, \\ \hat{x}_{t+2} \Big| \Omega_t^{x,y} &= c_1 + \pi_{11}^{(1)} \hat{x}_{t+1} + \dots + \pi_{11}^{(p)} x_{t-p+2} + \pi_{12}^{(1)} \hat{y}_{t+1} + \dots + \pi_{12}^{(p)} y_{t-p+2}, \\ \hat{y}_{t+2} \Big| \Omega_t^{x,y} &= c_2 + \pi_{21}^{(1)} \hat{x}_{t+1} + \dots + \pi_{21}^{(p)} x_{t-p+2} + \pi_{22}^{(1)} \hat{y}_{t+1} + \dots + \pi_{22}^{(p)} y_{t-p+2}, \\ &\dots\dots\dots \\ \hat{x}_{t+h} \Big| \Omega_t^{x,y} &= c_1 + \pi_{11}^{(1)} \hat{x}_{t+h-1} + \dots + \pi_{11}^{(p)} \hat{x}_{t+h-p} + \pi_{12}^{(1)} \hat{y}_{t+h-1} + \dots + \pi_{12}^{(p)} \hat{y}_{t+h-p}, \\ \hat{y}_{t+h} \Big| \Omega_t^{x,y} &= c_2 + \pi_{21}^{(1)} \hat{x}_{t+h-1} + \dots + \pi_{21}^{(p)} \hat{x}_{t+h-p} + \pi_{22}^{(1)} \hat{y}_{t+h-1} + \dots + \pi_{22}^{(p)} \hat{y}_{t+h-p}, \end{aligned}$$

если $h > p$. Если $h \leq p$, то вместо \hat{x}_{t+h-p} и \hat{y}_{t+h-p} следует использовать действительные значения x_{t+h-p} и y_{t+h-p} .

При этом:

- x не является G -причиной для y , если $\pi_{21}^{(1)} = \dots = \pi_{21}^{(p)} = 0$, т.е. если матрицы коэффициентов Π_j являются верхними треугольными:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^{(1)} & \pi_{12}^{(1)} \\ 0 & \pi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} \pi_{11}^{(p)} & \pi_{12}^{(p)} \\ 0 & \pi_{22}^{(p)} \end{pmatrix} \begin{pmatrix} x_{t-p} \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix};$$

- y не является G -причиной для x , если $\pi_{12}^{(1)} = \dots = \pi_{12}^{(p)} = 0$, т.е. если матрицы коэффициентов Π_j являются нижними треугольными:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^{(1)} & 0 \\ \pi_{21}^{(1)} & \pi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} \pi_{11}^{(p)} & 0 \\ \pi_{21}^{(p)} & \pi_{22}^{(p)} \end{pmatrix} \begin{pmatrix} x_{t-p} \\ y_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Если рассматриваемая двумерная VAR стабильна¹, то для проверки гипотезы H_0 : « y не является G -причиной для x » достаточно произвести OLS -оценивание модели

$$x_t = c_1 + \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + u_t$$

(с $(2p + 1)$ коэффициентами) и, используя F -критерий, проверить гипотезу

$$H_0 : \beta_1 = \dots = \beta_p = 0.$$

При нулевой гипотезе асимптотически статистика критерия имеет распределение, близкое к F -распределению $F(p, T - 2p - 1)$. Асимптотически эквивалентным этому критерию является критерий, основанный на статистике pF , которая при нулевой гипотезе имеет асимптотическое распределение $\chi^2(p)$. Соответственно для проверки гипотезы H_0 : « x не является G -причиной для y » достаточно произвести OLS -оценивание модели

$$y_t = c_1 + \alpha_1 x_{t-1} + \dots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + u_t$$

(с $(2p + 1)$ коэффициентами) и, используя F -критерий, проверить гипотезу

$$H_0 : \alpha_1 = \dots = \alpha_p = 0.$$

✓ **Замечание 6.1.1.** Результаты практического применения любого из тестов на существование причинности по Грейнджеру чрезвычайно чувствительны к выбору глубины запаздываний p и к методам, предусмотренным для случая потенциально нестационарных рядов.

Причинность по Грейнджеру не следует истолковывать как причинность в философском смысле! Это всего лишь попытка спецификации необходимого условия для причинной связи. Причинность по Грейнджеру предоставляет возможность улучшить качество прогнозов одной из переменных с привлечением информации о прошлых значениях другой переменной, в дополнение к информации о прошлых значениях самой переменной. Интерпретация причинности по Грейнджеру как необходимого условия для причинной связи наталкивается на следующее затруднение: отношение $x \xrightarrow{G} y$ (переменная x является причиной для переменной y) не является транзитивным. Например,

¹ См. разд. 8 ч. 2 данного учебника.

возможно, что $z \xrightarrow{G} x$ и $x \xrightarrow{G} y$, но переменная z не является G -причиной для переменной y .

Если в модели для пары переменных x_t и y_t обе переменные реагируют на изменение некоторой третьей переменной, не включенной в модель, то Грейнджер-причинность между переменными x_t и y_t может обнаружиться и при отсутствии G -причинного влияния одной из этих переменных на другую.

В связи с последним обстоятельством рассмотрим трехмерную VAR, описывающую совместную динамику переменных x_t , y_t и z_t :

$$\begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} \pi_{11}^{(1)} & \pi_{12}^{(1)} & \pi_{13}^{(1)} \\ \pi_{21}^{(1)} & \pi_{22}^{(1)} & \pi_{23}^{(1)} \\ \pi_{31}^{(1)} & \pi_{32}^{(1)} & \pi_{33}^{(1)} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \\ x_{t-1} \end{pmatrix} + \dots + \begin{pmatrix} \pi_{11}^{(p)} & \pi_{12}^{(p)} & \pi_{13}^{(p)} \\ \pi_{21}^{(p)} & \pi_{22}^{(p)} & \pi_{23}^{(p)} \\ \pi_{31}^{(p)} & \pi_{32}^{(p)} & \pi_{33}^{(p)} \end{pmatrix} \begin{pmatrix} y_{t-p} \\ z_{t-p} \\ x_{t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}.$$

Последнее соотношение можно записать в форме:

$$\begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} \pi_{11}^{(1)} L & \pi_{12}^{(1)} L & \pi_{13}^{(1)} L \\ \pi_{21}^{(1)} L & \pi_{22}^{(1)} L & \pi_{23}^{(1)} L \\ \pi_{31}^{(1)} L & \pi_{32}^{(1)} L & \pi_{33}^{(1)} L \end{pmatrix} \begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} + \dots + \begin{pmatrix} \pi_{11}^{(p)} L^p & \pi_{12}^{(p)} L^p & \pi_{13}^{(p)} L^p \\ \pi_{21}^{(p)} L^p & \pi_{22}^{(p)} L^p & \pi_{23}^{(p)} L^p \\ \pi_{31}^{(p)} L^p & \pi_{32}^{(p)} L^p & \pi_{33}^{(p)} L^p \end{pmatrix} \begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix},$$

или

$$\begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ \pi_{21}(L) & \pi_{22}(L) & \pi_{23}(L) \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix} \begin{pmatrix} y_t \\ z_t \\ x_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}.$$

Непосредственный перенос условия G -причинности $x \xrightarrow{G} y$ для VAR с двумя переменными на VAR с тремя переменными приводит к следующему условию отсутствия такой связи: $\pi_{13}(L) = 0$. Но если остальные элементы матрицы отличны от 0, то $x \xrightarrow{G} z$ и $z \xrightarrow{G} y$, т.е. такое определение G -причинности не является транзитивным в системах с тремя переменными и более.

Пусть, например,

$$\begin{aligned} y_t &= c_1 + \pi_{11}^{(1)} y_{t-1} + \pi_{12}^{(1)} z_{t-1} + \varepsilon_{1t}, \\ z_t &= c_2 + \pi_{22}^{(1)} z_{t-1} + \pi_{23}^{(1)} x_{t-1} + \varepsilon_{2t}, \\ x_t &= c_3 + \pi_{32}^{(1)} z_{t-1} + \pi_{33}^{(1)} x_{t-1} + \varepsilon_{3t}. \end{aligned}$$

Запаздывания переменной x не входят непосредственно в правую часть уравнения для y ($\pi_{13}^{(1)} = 0$). Однако $z_{t-1} = c_2 + \pi_{22}^{(1)} z_{t-2} + \pi_{23}^{(1)} x_{t-2} + \varepsilon_{2,t-1}$, так что

$$\begin{aligned} y_t &= c_1 + \pi_{11}^{(1)} y_{t-1} + \pi_{12}^{(1)} (c_2 + \pi_{22}^{(1)} z_{t-2} + \pi_{23}^{(1)} x_{t-2} + \varepsilon_{2,t-1}) + \varepsilon_{1t} = \\ &= (c_1 + \pi_{12}^{(1)} c_2) + \pi_{11}^{(1)} y_{t-1} + \pi_{12}^{(1)} \pi_{22}^{(1)} z_{t-2} + \pi_{12}^{(1)} \pi_{23}^{(1)} x_{t-2} + (\varepsilon_{1t} + \pi_{12}^{(1)} \varepsilon_{2,t-1}). \end{aligned}$$

Чтобы исключить влияние запаздываний переменной x на y , надо положить (дополнительно к условию $\pi_{13}^{(1)} = 0$) или $\pi_{12}^{(1)} = 0$, или $\pi_{23}(L) = 0$:

$$\pi_{12}(L) = 0 \text{ и } \pi_{13}(L) = 0 \quad (\text{случай 1})$$

$$\pi_{13}(L) = 0 \text{ и } \pi_{23}(L) = 0 \quad (\text{случай 2}).$$

В первом случае в выражение для y_t входят только прошлые значения самой этой переменной:

$$P(L) = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ \pi_{21}(L) & \pi_{22}(L) & \pi_{23}(L) \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix} = \begin{pmatrix} \pi_{11}(L) & 0 & 0 \\ \pi_{21}(L) & \pi_{22}(L) & \pi_{23}(L) \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix}.$$

Во втором случае в выражение для y_t входят еще и прошлые значения переменной z . Однако выражения для последних не содержат прошлых значений переменной x , так что и в этом случае прошлые значения переменной x не входят в выражение для y_t :

$$P(L) = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ \pi_{21}(L) & \pi_{22}(L) & \pi_{23}(L) \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix} = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & 0 \\ \pi_{21}(L) & \pi_{22}(L) & 0 \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix}.$$

В обоих случаях матрица $P(L)$ является блочно-треугольной с двумя блоками размеров (2×2) и (1×1) .

Если рассматривать пару переменных z и y , то как непосредственное, так и опосредованное (через посредство переменной x) влияние переменной y на переменную z исключается при выполнении условий:

$$\pi_{31}(L) = 0 \text{ и } \pi_{21}(L) = 0 \quad (\text{случай 3})$$

либо условий:

$$\pi_{21}(L) = 0 \text{ и } \pi_{23}(L) = 0 \quad (\text{случай 4}).$$

В случае 3 матрица $P(L)$ имеет вид:

$$P(L) = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ 0 & \pi_{22}(L) & \pi_{23}(L) \\ 0 & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix};$$

в случае 4:

$$\Pi(L) = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ 0 & \pi_{22}(L) & 0 \\ \pi_{31}(L) & \pi_{32}(L) & \pi_{33}(L) \end{pmatrix}.$$

В последней ситуации матрица $\Pi(L)$ не является блочно-треугольной, но если поставить переменную x на второе место, а переменную z — на третье место, то матрица $\Pi(L)$ примет блочно-треугольную форму:

$$\Pi(L) = \begin{pmatrix} \pi_{11}(L) & \pi_{12}(L) & \pi_{13}(L) \\ \pi_{21}(L) & \pi_{22}(L) & \pi_{23}(L) \\ 0 & 0 & \pi_{33}(L) \end{pmatrix}.$$

В подобных случаях говорят о **блочной экзогенности** (*block exogeneity*) **одной группы переменных (или некоторой переменной) в отношении другой группы переменных (или другой переменной)**. Блочная экзогенность означает, что запаздывающие значения второй группы переменных не входят в уравнения для переменных первой группы.

В рассмотренных выше четырех ситуациях (при выполнении соответствующих четырех пар условий) получаем:

- в с л у ч а е 1: переменная y блочно экзогенна в отношении переменных z и x (переменные z и x не участвуют в ее формировании);
- в с л у ч а е 2: пара (y, z) блочно экзогенна в отношении переменной x (переменная x не участвует в формировании y и z);
- в с л у ч а е 3: пара (z, x) блочно экзогенна в отношении переменной y (переменная y не участвует в формировании x и z);
- в с л у ч а е 4: переменная z блочно экзогенна в отношении переменных y и x (переменные y и x не участвуют в ее формировании).

Пусть теперь имеется N переменных (N временных рядов) $y_{1t}, y_{2t}, \dots, y_{Nt}$, составляющих векторный ряд $y_t, y_t = (y_{1t}, \dots, y_{Nt})^T$ (верхний индекс T означает транспонирование вектора). Рассмотрим разбиение вектора y_t на две части:

$$y_t = \begin{pmatrix} y_t^{(1)} \\ y_t^{(2)} \end{pmatrix},$$

где $y_t^{(1)} = (y_{1t}, \dots, y_{N_1t})^T, y_t^{(2)} = (y_{N_1+1,t}, \dots, y_{Nt})^T, 1 \leq N_1 < N$.

Пусть:

$\Omega_T^{(1)}$ — информационное множество, содержащее значения $y_{iT}, y_{i,T-1}, \dots$ ($i = 1, \dots, N_1$) первой группы переменных;

$\Omega_T^{(2)}$ — информационное множество, содержащее значения $y_{iT}, y_{i,T-1}, \dots$ ($i = N_1 + 1, \dots, N$) второй группы переменных;

$\Omega_T^{(1,2)}$ — объединение этих двух множеств;

$\hat{y}_{i,T+h}|\Omega = E(y_{i,T+h}|\Omega)$ — оптимальный прогноз i -й переменной первой группы на h шагов вперед относительно информационного множества Ω , где в качестве Ω могут выступать информационные множества $\Omega_T^{(1)}$ или $\Omega_T^{(1,2)}$.

Говорят, что **переменные второй группы не являются Грейнджер-причиной для переменных первой группы**, если

$$E\left|y_{i,T+h} - (\hat{y}_{i,T+h}|\Omega_T^{(1)})\right|^2 \leq E\left|y_{i,T+h} - (\hat{y}_{i,T+h}|\Omega_T^{(1,2)})\right|^2 \text{ для } h = 1, 2, \dots,$$

так что дополнительная информация о значениях переменных второй группы не улучшает оптимальный прогноз переменных первой группы — для его построения достаточно использовать только значения $y_{iT}, y_{i,T-1}, \dots$ ($i = 1, \dots, N_1$) первой группы переменных. Если же такая информация улучшает оптимальный прогноз, то **переменные второй группы являются причиной по Грейнджеру для переменных первой группы**.

Пусть динамика $(N \times 1)$ -вектора y_t представлена гауссовской VAR(p):

$$y_t = c + \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. N(0, \Sigma),$$

и гипотеза состоит в том, что переменные второй группы не являются Грейнджер-причиной для i -й переменной первой группы ($i = 1, 2, \dots, N_1$). Для проверки этого предположения оценивается уравнение

$$y_{it} = \omega_{1i}^T y_{i,t-1}^{(1)} + \lambda_{1i}^T y_{i,t-1}^{(2)} + \omega_{2i}^T y_{i,t-2}^{(1)} + \lambda_{2i}^T y_{i,t-2}^{(2)} + \dots + \omega_{pi}^T y_{i,t-p}^{(1)} + \lambda_{pi}^T y_{i,t-p}^{(2)} + \alpha_i + u_{it}$$

и проверяется гипотеза

$$H_0: \lambda_{1i} = \lambda_{2i} = \dots = \lambda_{pi} = 0.$$

Если рассматриваемая VAR стабильна, то использование F -статистики для проверки этой гипотезы асимптотически оправданно. Проблемы возникают в случае нестабильной VAR.

Если рассмотреть в качестве нулевой гипотезу

$$H_0: \lambda_{1i} = \lambda_{2i} = \dots = \lambda_{pi} = 0, \quad i = 1, 2, \dots, N_1,$$

охватывающую все N_1 уравнений первой группы, то такая гипотеза соответствует блочной экзогенности первой группы переменных в отношении второй группы переменных. Если эта гипотеза выполняется, то в порождении переменных первой группы не участвуют переменные второй группы. Соответственно N_2 уравнений второй группы образуют открытую VAR¹, которую можно рассматривать самостоятельно.

¹ См. разд. 8 ч. 2 данного учебника.

Последняя гипотеза накладывает $N_1 p$ ограничений на коэффициенты $\text{VAR}(p)$ для $N = N_1 + N_2$ переменных и при $N_1 > 1$ имеет перекрестный характер, затрагивая коэффициенты сразу нескольких уравнений (точнее, коэффициенты N_1 уравнений). Проверить эту гипотезу можно, применяя критерий отношения правдоподобий (LR -критерий), статистика которого равна:

$$LR = 2(l_{UNRESTR} - l_{RESTR}) = T(\ln |\hat{\Sigma}_{RESTR}| - \ln |\hat{\Sigma}_{UNRESTR}|),$$

где $l_{UNRESTR}$ — логарифм максимума функции правдоподобия при оценивании $\text{VAR}(p)$ без ограничений на коэффициенты;

l_{RESTR} — логарифм максимума функции правдоподобия при оценивании $\text{VAR}(p)$ с ограничениями на коэффициенты, соответствующими проверяемой гипотезе;

$|\hat{\Sigma}|$ — определитель оцененной ковариационной матрицы инноваций, вычисляемой двумя разными способами:

$$|\hat{\Sigma}| = \det\left(\frac{1}{T} \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_t^T\right) \quad \text{или} \quad |\hat{\Sigma}| = \det\left(\frac{1}{T-p} \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_t^T\right).$$

Статистика LR имеет асимптотическое распределение хи-квадрат с числом степеней свободы, равным количеству зануляемых коэффициентов.

Нестационарные VAR

Пусть динамика $(N \times 1)$ -вектора y_t представлена моделью $\text{VAR}(p)$:

$$y_t = \alpha + \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + \varepsilon_t,$$

$$\varepsilon_t \sim i.i.d., \quad E(\varepsilon_t) = 0, \quad \text{Cov}(\varepsilon_t) = \Sigma.$$

Последнее представление можно записать следующим образом:

$$y_t - \Pi_1 y_{t-1} - \dots - \Pi_p y_{t-p} = \alpha + \varepsilon_t,$$

или

$$\Pi(L)y_t = \alpha + \varepsilon_t,$$

где

$$\Pi(L) = I_N - \Pi_1 L - \dots - \Pi_p L^p.$$

Во второй части учебника (разд. 8, тема 8.2) уже говорилось о том, что условие стабильности такой VAR состоит в следующем:

Все корни алгебраического уравнения $\det \Pi(z) = 0$ лежат за пределами единичного круга на комплексной плоскости (т.е. модули всех этих корней больше 1).

✓ **Замечание 6.1.2**

- Если VAR рассматривается на всей оси времени, то из стабильности VAR вытекает ее стационарность как векторного процесса (стационарность в широком смысле).
- Если VAR рассматривается только с некоторого момента времени при заданных начальных условиях, то из стабильности VAR вытекает ее асимптотическая стационарность.
- Существуют (не рассматриваемые обычно) стационарные модели VAR, для которых условие стабильности не выполнено.
- Если VAR, рассматриваемая на всей оси времени, нестационарна, то она нестабильна.

Пусть компоненты вектора y_t , динамика которого представлена моделью $\text{VAR}(p)$, являются $I(1)$ -рядами (интегрированными рядами порядка 1), так что ряды их приращений (ряды разностей) являются стационарными временными рядами, а сама VAR нестабильна. Стандартную форму векторной авторегрессии порядка p

$$y_t = \alpha + \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + \varepsilon_t \quad (6.1)$$

можно переписать в виде:

$$y_t = \zeta_1 \Delta y_{t-1} + \dots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + \rho y_{t-1} + \varepsilon_t, \quad (6.2)$$

где $\rho = \Pi_1 + \dots + \Pi_p$,

$$\zeta_s = -(\Pi_{s+1} + \Pi_{s+2} + \dots + \Pi_p) \text{ для } s = 1, 2, \dots, p-1.$$

Некоинтегрированная VAR

Если $\rho = I_N$, то компоненты y_t не коинтегрированы, и модель принимает форму векторной авторегрессии $\text{VAR}(p-1)$ в разностях:

$$\Delta y_t = \zeta_1 \Delta y_{t-1} + \dots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + \varepsilon_t. \quad (6.3)$$

При оценивании статистической модели (6.2) независимо от того, $\alpha = 0$ или $\alpha \neq 0$ в DGP , совместное распределение оценок матричных коэффициентов $\zeta_1, \dots, \zeta_{p-1}$ является асимптотически нормальным, причем это асимптотическое распределение точно такое же, как и для совместного распределения оценок матричных коэффициентов $\zeta_1, \dots, \zeta_{p-1}$ в модели (6.3). Соответственно для проверки гипотез об этих коэффициентах (как в рамках уравнения (6.2), так и в рамках уравнения (6.3)) можно использовать (асимптотически) стандартные t - и F -критерии или статистику $qF \sim \chi^2(q)$, где q — количество (линейно независимых) линейных ограничений на элементы матриц $\zeta_1, \dots, \zeta_{p-1}$.

Непосредственно в рамках уравнения (6.1) возможно использование стандартных тестов для проверки гипотезы

$$H_0: DGP = \text{VAR}(p_0), \quad p_0 \geq 1,$$

против альтернативы

$$H_A: DGP = \text{VAR}(p), \quad p > p_0,$$

с целью выбора оптимальной глубины запаздываний. Но, например, F -критерий для проверки гипотез о значении суммы $\Pi_1 + \dots + \Pi_p$ (т.е. о значении ρ) в этой ситуации неприменим: оценка для этой суммы имеет нестандартное распределение. Кроме того, статистики критериев для проверки на причинность по Грейнджеру отдельных переменных или подгруппы переменных (компонент y_t) не имеют обычных асимптотических распределений. Поясним этот факт.

Пусть компоненты вектора y_t являются $I(1)$ -рядами (интегрированными рядами порядка 1), так что ряды их приращений (ряды разностей) являются стационарными. Опять рассмотрим разбиение вектора y_t на две подгруппы, включающие соответственно N_1 и $N_2 = N - N_1$ переменных, и пусть гипотеза H_0 состоит в том, что переменные второй группы не являются Грейнджер-причиной для переменных первой группы. Для проверки этой гипотезы в стационарном случае мы записывали модель в уровнях в виде:

$$y_{it} = \omega_{1i}^T y_{t-1}^{(1)} + \lambda_{1i}^T y_{t-1}^{(2)} + \omega_{2i}^T y_{t-2}^{(1)} + \lambda_{2i}^T y_{t-2}^{(2)} + \dots + \omega_{pi}^T y_{t-p}^{(1)} + \lambda_{pi}^T y_{t-p}^{(2)} + \alpha_i + \varepsilon_{it}$$

и проверяли гипотезу

$$H_0: \lambda_{1i} = \lambda_{2i} = \dots = \lambda_{pi} = 0.$$

Перепишем эту модель в виде:

$$y_{it} = \beta_{1i}^T \Delta y_{t-1}^{(1)} + \gamma_{1i}^T \Delta y_{t-1}^{(2)} + \beta_{2i}^T \Delta y_{t-2}^{(1)} + \gamma_{2i}^T \Delta y_{t-2}^{(2)} + \dots + \\ + \beta_{p-1,i}^T \Delta y_{t-p+1}^{(1)} + \gamma_{p-1,i}^T \Delta y_{t-p+1}^{(2)} + \alpha_i + \eta_i^T y_{t-1}^{(1)} + \delta_i^T y_{t-1}^{(2)} + \varepsilon_{it}.$$

При такой записи модели нулевая гипотеза принимает вид:

$$H'_0: \gamma_{1i} = \gamma_{2i} = \dots = \gamma_{p-1,i} = \delta_i = 0,$$

причем статистика F -критерия для проверки такой гипотезы численно идентична статистике F -критерия для проверки гипотезы $H_0: \lambda_{1i} = \lambda_{2i} = \dots = \lambda_{pi} = 0$. Но оценка для δ_i имеет нестандартное распределение. Моделирование показывает, что при использовании F -критерия в подобных ситуациях слишком часто определяется ложная причинность по Грейнджеру.

Осложнений, возникающих при проверке гипотезы об отсутствии причинности по Грейнджеру в рамках уравнения для уровней, не возникает при проверке аналогичной гипотезы в рамках уравнения (6.3) для рядов разностей, т.е. для VAR в разностях.

Коинтегрированная VAR

Если имеем дело с коинтегрированной VAR, компоненты которой являются $I(1)$ -рядами, модель (6.2) можно записать в форме **модели коррекции ошибок** (*error correction model* — *ECM*):

$$\Delta y_t = \zeta_1 \Delta y_{t-1} + \dots + \zeta_{p-1} \Delta y_{t-p+1} + \alpha + BA^T y_{t-1} + \varepsilon_t,$$

где B — $(N \times r)$ -матрица коэффициентов адаптации;

A — $(N \times r)$ -матрица, столбцы которой представляют r линейно независимых коинтегрирующих векторов;

r — ранг коинтеграции, $1 \leq r \leq N - 1$ (при $r = 0$ коинтеграции нет; при $r = N$ все ряды стационарны).

Если A_i — i -я строка матрицы A^T , то линейные комбинации $z_1 = A_1 y_{t-1}$, ..., $z_r = A_r y_{t-1}$ стационарны, и потенциально все эти стационарные линейные (коинтегрирующие линейные) комбинации должны быть включены в правые части уравнений системы в качестве объясняющих переменных:

$$\begin{aligned} \Delta y_{1t} = & (\zeta_{11}^{(1)} \Delta y_{1,t-1} + \dots + \zeta_{1,N}^{(1)} \Delta y_{N,t-1}) + (\zeta_{11}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{1,N}^{(2)} \Delta y_{N,t-2}) + \dots + \\ & + (\zeta_{11}^{(p-1)} \Delta y_{1,t-p+1} + \dots + \zeta_{1,N}^{(p-1)} \Delta y_{N,t-p+1}) + \\ & + \alpha_1 + B_{11} z_{1,t-1} + \dots + B_{1r} z_{r,t-1} + \varepsilon_{1t}, \end{aligned}$$

.....

$$\begin{aligned} \Delta y_{Nt} = & (\zeta_{N1}^{(1)} \Delta y_{1,t-1} + \dots + \zeta_{N,N}^{(1)} \Delta y_{N,t-1}) + (\zeta_{N1}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{N,N}^{(2)} \Delta y_{N,t-2}) + \dots + \\ & + (\zeta_{N1}^{(p-1)} \Delta y_{1,t-p+1} + \dots + \zeta_{N,N}^{(p-1)} \Delta y_{N,t-p+1}) + \\ & + \alpha_N + B_{N1} z_{1,t-1} + \dots + B_{Nr} z_{r,t-1} + \varepsilon_{Nt}. \end{aligned}$$

При этом обязательно должно выполняться условие $B \neq 0$, так что хотя бы один из коэффициентов B_{ij} , $i = 1, \dots, N$, $j = 1, \dots, r$, должен быть отличен от нуля. Если это, скажем, коэффициент B_{ij} , то в прогнозировании i -й переменной помогают, во всяком случае, прошлые значения тех переменных, у которых коэффициенты в j -й коинтегрирующей линейной комбинации отличны от нуля.

Если компоненты ряда y_t являются $I(1)$ -рядами и ранг коинтеграции равен 1, то для оценивания модели коррекции ошибок можно использовать двухступенчатую процедуру Энгла — Грейнджера. В этом случае производят *OLS*-оценивание уравнения

$$y_{1t} = \mu + \gamma_2 y_{2t} + \gamma_3 y_{3t} + \dots + \gamma_N y_{Nt} + u_t,$$

получают оценки $\hat{\mu}$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, ..., $\hat{\gamma}_N$ коэффициентов долговременного соотношения и на их основе — оценку

$$\hat{z}_{1t} = y_{1t} - \hat{\mu} - \hat{\gamma}_2 y_{2t} - \hat{\gamma}_3 y_{3t} - \dots - \hat{\gamma}_N y_{Nt}$$

стационарной линейной комбинации

$$z_{1t} = y_{1t} - \mu - \gamma_2 y_{2t} - \gamma_3 y_{3t} - \dots - \gamma_N y_{Nt},$$

имеющей нулевое среднее.

Так построенную переменную \hat{z}_{1t} подставляют вместо z_{1t} в правую часть уравнений системы в форме модели коррекции ошибок и производят *OLS*-оценивание коэффициентов последней. К оценкам всех коэффициентов этой системы, кроме коэффициента при $\hat{z}_{1,t-1}$, применимы стандартные асимптотические выводы, поскольку это коэффициенты при стационарных переменных. Вследствие наличия коинтеграции оценки $\hat{\alpha}$, $\hat{\gamma}_2$, $\hat{\gamma}_3$, ..., $\hat{\gamma}_N$ суперсостоятельны, и это приводит к тому, что получаемая в итоге оценка для коэффициента при $\hat{z}_{1,t-1}$ имеет ту же асимптотику, что и оценка коэффициента при стационарной переменной $z_{1,t-1}$ в уравнении ЕСМ с известной z_{1t} . Таким образом, при наличии единственной (с точностью до нормировки) коинтегрирующей линейной комбинации компонент вектора y_t стандартные асимптотические результаты остаются в силе для всех коэффициентов ЕСМ, оцениваемой на втором шаге процедуры.

✓ **Замечание 6.1.3.** Стандартные (асимптотические) процедуры проверки гипотез о параметрах ЕСМ могут применяться к любым гипотезам, не включающим предположения $B_{11} = \dots = B_{N1} = 0$, соответствующего некоинтегрированности компонент ряда y_t . При гипотезе $H_0: B_{11} = \dots = B_{N1} = 0$ асимптотическое распределение оценок коэффициентов B_{11} , ..., B_{N1} нестандартное и зависит от мешающих параметров.

Рассмотренная в разд. 11 ч. 2 данного учебника процедура Йохансена использует метод максимального правдоподобия и позволяет:

- оценивать ранг коинтеграции;
- проверять гипотезы о ранге коинтеграции;
- оценивать коинтегрирующие векторы, соответствующие выбранному рангу коинтеграции;
- проверять гипотезы о выполнении тех или иных ограничений на коэффициенты коинтегрирующих векторов.

В отличие от процедуры Энгла — Грейнджера, здесь ранг коинтеграции может быть любым. Однако недостатками этой процедуры являются большая чувствительность к отклонениям распределения случайных ошибок от нормального и то, что для получения удовлетворительных результатов она требует большого количества наблюдений (около 100 и даже более).

Для коинтегрированной двумерной VAR ранг коинтеграции равен $r = 1$, так что ЕСМ имеет вид:

$$\begin{aligned}\Delta y_{1t} &= (\zeta_{11}^{(1)} \Delta y_{1,t-1} + \zeta_{12}^{(1)} \Delta y_{2,t-1}) + (\zeta_{11}^{(2)} \Delta y_{1,t-2} + \zeta_{12}^{(2)} \Delta y_{2,t-2}) + \dots + \\ &\quad + (\zeta_{11}^{(p-1)} \Delta y_{1,t-p+1} + \zeta_{12}^{(p-1)} \Delta y_{2,t-p+1}) + \\ &\quad + \alpha_1 + B_{11} z_{1,t-1} + \varepsilon_{1t}, \\ \Delta y_{2t} &= (\zeta_{21}^{(1)} \Delta y_{1,t-1} + \zeta_{22}^{(1)} \Delta y_{2,t-1}) + (\zeta_{21}^{(2)} \Delta y_{1,t-2} + \zeta_{22}^{(2)} \Delta y_{2,t-2}) + \dots + \\ &\quad + (\zeta_{21}^{(p-1)} \Delta y_{1,t-p+1} + \zeta_{22}^{(p-1)} \Delta y_{2,t-p+1}) + \\ &\quad + \alpha_2 + B_{21} z_{1,t-1} + \varepsilon_{2t},\end{aligned}$$

или

$$\begin{aligned}\Delta y_{1t} &= \alpha_1 + B_{11} z_{1,t-1} + \\ &\quad + \zeta_{11}^{(1)} \Delta y_{1,t-1} + \zeta_{11}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{11}^{(p-1)} \Delta y_{1,t-p+1} + \\ &\quad + \zeta_{12}^{(1)} \Delta y_{2,t-1} + \zeta_{12}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{12}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{1t}, \\ \Delta y_{2t} &= \alpha_2 + B_{21} z_{1,t-1} + \\ &\quad + \zeta_{21}^{(1)} \Delta y_{1,t-1} + \zeta_{21}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{21}^{(p-1)} \Delta y_{1,t-p+1} + \\ &\quad + \zeta_{22}^{(1)} \Delta y_{2,t-1} + \zeta_{22}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{22}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{2t},\end{aligned}$$

где

$$z_{1,t-1} = y_{1,t-1} - \mu - \gamma_2 y_{2,t-1}, \quad \gamma_2 \neq 0.$$

Поскольку хотя бы один из коэффициентов B_{11} или B_{21} должен быть отличен от нуля (иначе не было бы коинтеграции), это означает, что в правую часть хотя бы одного из двух уравнений ЕСМ входит в качестве объясняющей переменная $z_{1,t-1} = y_{1,t-1} - \mu - \gamma_2 y_{2,t-1}$. И, следовательно, в таком уравнении имеет место причинность по Грейнджеру. Например, если в уравнении для Δy_{1t} отличен от нуля коэффициент B_{11} , то значение $y_{2,t-1}$ помогает в прогнозировании значения y_{1t} наряду со значением $y_{1,t-1}$ и запаздывающими разностями $\Delta y_{1,t-1}, \Delta y_{1,t-2}, \dots, \Delta y_{1,t-p+1}$.

Поскольку причинность по Грейнджеру в коинтегрированной двумерной VAR существует по крайней мере в одном направлении, фактически вопрос может стоять только о том, существует ли G -причинность в выбранном направлении — например, в направлении от ряда y_{2t} к ряду y_{1t} . В последнем случае гипотеза об отсутствии такой причинности сводится к гипотезе:

$$H_0: B_{11} = 0, \quad \zeta_{12}^{(1)} = \zeta_{12}^{(2)} = \dots = \zeta_{12}^{(p-1)} = 0,$$

которую можно проверить стандартными методами, даже если использовать вместо значений $z_{1,t-1}$ значения $\hat{z}_{1,t-1} = y_{1,t-1} - \hat{\mu} - \hat{\gamma}_2 y_{2,t-1}$ с $\hat{\mu}$, $\hat{\gamma}_2$, полученными на первом шаге процедуры Энгла — Грейнджера.

Методология Тода — Ямамото

Выше было показано, как можно проводить проверку на причинность по Грейнджеру, когда:

- VAR стационарна;
- VAR состоит из $I(1)$ -рядов и они не коинтегрированы (следует перейти к VAR в разностях);
- VAR состоит из $I(1)$ -рядов, и они коинтегрированы.

Применение соответствующих методов требует предварительной проверки гипотезы единичного корня для определения порядков интегрированности рядов и проверки гипотезы об их коинтегрированности/некоинтегрированности. Однако критерии для проведения такой проверки обычно обладают малой мощностью, и это ограничивает применение указанных методов.

Тода и Ямамото (*Toda, Yamamoto, 1995*) предложили процедуру, позволяющую обойти эти проверки.

Тода и Ямамото исходят из модели *DGP*:

$$y_t = \gamma_0 + \gamma_1 t + \dots + \gamma_q t^q + \Pi_1 y_{t-1} + \dots + \Pi_k y_{t-k} + \varepsilon_t,$$

где $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $Cov(\varepsilon_t) = \Sigma$, $E|\varepsilon_{it}|^{2+\delta} < \infty$ для некоторого $\delta > 0$.

Для проверки гипотезы об отсутствии причинного влияния тех или иных переменных на другие переменные предлагается проверять эту гипотезу в рамках расширенной статистической модели:

$$y_t = \gamma_0 + \gamma_1 t + \dots + \gamma_q t^q + \Pi_1 y_{t-1} + \dots + \Pi_k y_{t-k} + \dots + \Pi_p y_{t-p} + \varepsilon_t,$$

где

$$p \geq k + d, \quad d — \text{порядок интегрированности ряда } y_t.$$

В качестве d обычно берется d_{\max} — предполагаемый максимальный порядок интегрированности рядов в составе y_t .

Гипотезы отсутствия причинности по Грейнджеру в *DGP* затрагивают в такой постановке только элементы матриц Π_1, \dots, Π_k (и не затрагивают остальных матриц в SM , т.е. матриц Π_{k+1}, \dots, Π_p). Если гипотеза накладывает m линейных ограничений на элементы этих матриц, то статистика Вальда для проверки этой гипотезы имеет асимптотическое распределение $\chi^2(m)$. При этом ряд y_t может быть стационарным, $I(1)$ или $I(2)$, причем в каждом случае — еще и относительно линейного тренда, и если это $I(1)$ - или $I(2)$ -ряд, то он может быть и коинтегрированным.

Рассмотрим статистическую процедуру выбора количества запаздываний. Запишем *DGP* в виде:

$$y_t = \gamma_0 + \gamma_1 t + \dots + \gamma_q t^q + \Pi_1 y_{t-1} + \dots + \Pi_k y_{t-k} + \dots + \Pi_p y_{t-p} + \varepsilon_t,$$

где

$$\Pi_{k+1} = \dots = \Pi_p = 0, \quad p \geq k + d. \quad (6.4)$$

Оценим аналогичную SM, но без ограничения (выражение (6.4)), и проверим гипотезу

$$H_0^*: \Pi_{l+1} = \dots = \Pi_p = 0, \quad \text{где } k \leq l \leq p - 1,$$

т.е. гипотезу о невхождении в *DGP* лагов более высокого порядка, чем l . Если эта гипотеза верна, то статистика Вальда для проверки этой гипотезы имеет асимптотическое распределение $\chi^2(n^2(p-l))$, если $l \geq d$.

Заметим, что если $k \geq d$, то $l \geq k \geq d$. Следовательно, если порядки интегрированности рядов не превышают истинного количества запаздываний в *DGP*, то применима обычная процедура выбора количества лагов в VAR.

В частности, если $d = 1$, то процедура выбора порядка модели всегда асимптотически обоснованна. Если же $d = 2$, то она асимптотически обоснованна только при $k > 1$.

При выборе порядка модели можно использовать также информационные критерии.

Резюмируем сказанное: при подозрениях на возможную интегрированность или коинтегрированность рядов в составе y_t гипотезу H_0 можно проверять, не производя собственно проверки рядов на интегрированность и коинтегрированность, а лишь позаботившись о том, чтобы SM в виде VAR имела порядок $(k + d_{\max})$.

Используя стандартную асимптотическую теорию, можно проверять и другие линейные (и многие нелинейные) ограничения на первые k матриц коэффициентов.

✓ Замечание 6.1.4

- Добавление лишних лагов может значительно понизить мощность критериев, если количество рядов N велико.
- В модель можно также включать сезонные дамми-переменные.

✓ Замечание 6.1.5. Говоря о причинности по Грейнджеру в моделях нестационарных VAR, следует, вообще говоря, различать **причинность в долгосрочном плане** (*long-run*) и **причинность в краткосрочном плане** (*short-run*).

Если рассматривается некоинтегрированная VAR(p) с $I(1)$ -переменными, то, переходя к модели в разностях, получаем VAR со стационарными переменными. Для простоты пусть VAR связывает две переменные ($N = 2$), так что VAR в разностях имеет вид:

$$\begin{aligned}
\Delta y_{1t} &= \alpha_1 + \zeta_{11}^{(1)} \Delta y_{1,t-1} + \zeta_{11}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{11}^{(p-1)} \Delta y_{1,t-p+1} + \\
&\quad + \zeta_{12}^{(1)} \Delta y_{2,t-1} + \zeta_{12}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{12}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{1t}, \\
\Delta y_{2t} &= \alpha_2 + \zeta_{21}^{(1)} \Delta y_{1,t-1} + \zeta_{21}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{21}^{(p-1)} \Delta y_{1,t-p+1} + \\
&\quad + \zeta_{22}^{(1)} \Delta y_{2,t-1} + \zeta_{22}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{22}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{2t}.
\end{aligned}$$

Если в первом уравнении $\zeta_{12}^{(1)} = \zeta_{12}^{(2)} = \dots = \zeta_{12}^{(p-1)} = 0$, то y_2 не является G -причиной для y_1 в краткосрочном плане. Если во втором уравнении $\zeta_{21}^{(1)} = \zeta_{21}^{(2)} = \dots = \zeta_{21}^{(p-1)} = 0$, то y_1 не является G -причиной для y_2 в краткосрочном плане. При этом асимптотически оправданно использование F -критериев для проверки линейных гипотез о коэффициентах этой VAR, так что проверка выполнения этих соотношений может осуществляться на основе соответствующих F -критериев.

Если рассматривается коинтегрированная VAR(p) для двух $I(1)$ -переменных, то в рамках соответствующей модели коррекции ошибок:

$$\begin{aligned}
\Delta y_{1t} &= \alpha_1 + B_{11} z_{1,t-1} + \\
&\quad + \zeta_{11}^{(1)} \Delta y_{1,t-1} + \zeta_{11}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{11}^{(p-1)} \Delta y_{1,t-p+1} + \\
&\quad + \zeta_{12}^{(1)} \Delta y_{2,t-1} + \zeta_{12}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{12}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{1t}, \\
\Delta y_{2t} &= \alpha_2 + B_{21} z_{1,t-1} + \\
&\quad + \zeta_{21}^{(1)} \Delta y_{1,t-1} + \zeta_{21}^{(2)} \Delta y_{1,t-2} + \dots + \zeta_{21}^{(p-1)} \Delta y_{1,t-p+1} + \\
&\quad + \zeta_{22}^{(1)} \Delta y_{2,t-1} + \zeta_{22}^{(2)} \Delta y_{2,t-2} + \dots + \zeta_{22}^{(p-1)} \Delta y_{2,t-p+1} + \varepsilon_{2t},
\end{aligned}$$

можно проверять как гипотезы об отсутствии краткосрочной G -причинности одной из переменных в отношении другой (выражаемые соотношениями $\zeta_{12}^{(1)} = \zeta_{12}^{(2)} = \dots = \zeta_{12}^{(p-1)} = 0$ и $\zeta_{21}^{(1)} = \zeta_{21}^{(2)} = \dots = \zeta_{21}^{(p-1)} = 0$ соответственно), так и гипотезы об отсутствии долгосрочной G -причинности одной из переменных в отношении другой (выражаемые соотношениями $B_{11} = 0$ и $B_{21} = 0$ соответственно). При этом гипотеза об отсутствии G -причинности в выбранном направлении — например, в направлении от y_2 к y_1 — формулируется следующим образом:

$$H_0: B_{11} = 0, \quad \zeta_{12}^{(1)} = \zeta_{12}^{(2)} = \dots = \zeta_{12}^{(p-1)} = 0.$$

Именно в таком виде гипотеза об отсутствии G -причинности в направлении от y_2 к y_1 рассматривалась ранее, при обсуждении коинтегрированных VAR.

ПРИМЕР 6.1.1

Обратимся к рассмотренной в работе (Leeper, Sims, Zha, 1996) модели, в которой были использованы переменные:

$$LP = 100\ln(CPI_SA), \quad LY = 100\ln(RGDP_MON), \quad LM2 = 100\ln(M2_SA).$$

Оценивание модели произведем на периоде 1974:01 — 1980:03 (75 наблюдений). На этом периоде все эти ряды определяются как $I(1)$ -ряды. Кointеграция между ними не выявляется. Графики указанных рядов приведены на рис. 6.1.

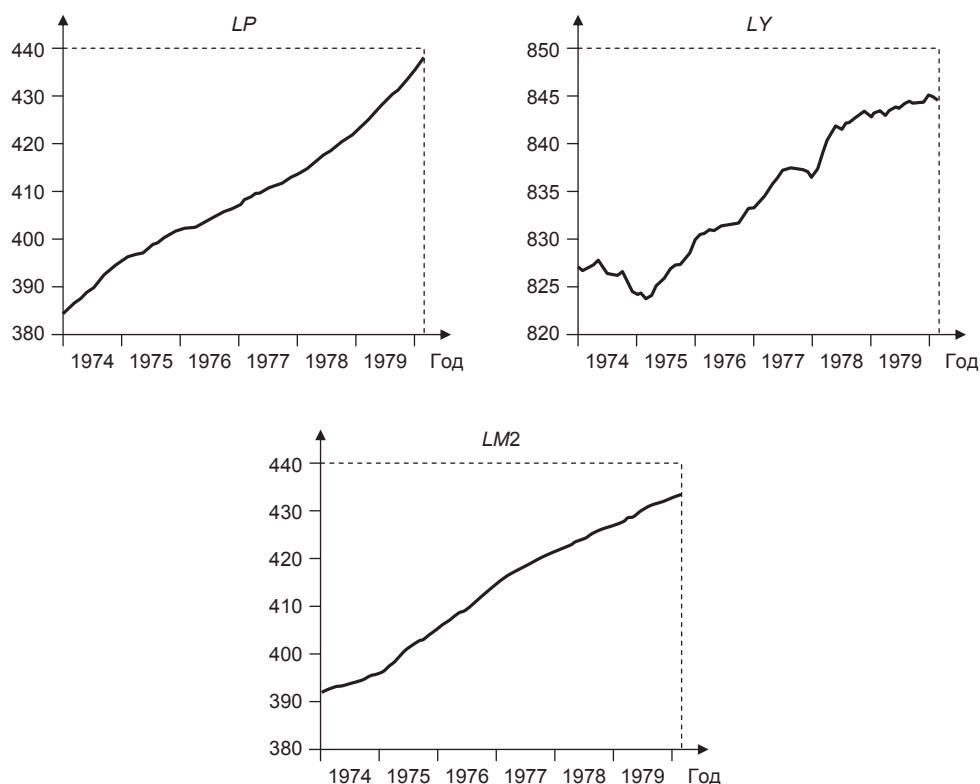


Рис. 6.1

В системе продифференцированных рядов LP_DIF , LY_DIF , $LM2_DIF$ можно провести исследование на наличие/отсутствие причинности по Грейнджеру в краткосрочном плане. Результаты такой проверки в рамках объекта *Group* пакета *EViews 6* приведены в табл. 6.1. Они уверенно указывают только на краткосрочную G -причинную связь в направлении от LY_DIF к $LM2_DIF$, на блочную экзогенность переменной LP , а также в определенной степени на то, что наиболее эндогенной является переменная $LM2$.

Таблица 6.1

Наличие причинности по Грейнджеру в краткосрочном плане

VAR Granger Causality/Block Exogeneity Wald Tests;

Sample: 1974M01 1980M03; Included observations: 73

<i>Excluded</i>	<i>Chi-sq</i>	<i>df</i>	<i>Prob.</i>
<i>Dependent variable: LP_DIF</i>			
<i>LY_DIF</i>	0.216406	2	0.8974
<i>LM2_DIF</i>	2.577506	2	0.2756
<i>All</i>	2.856723	4	0.5821
<i>Dependent variable: LY_DIF</i>			
<i>LP_DIF</i>	1.196085	2	0.5499
<i>LM2_DIF</i>	4.693268	2	0.0957
<i>All</i>	8.418789	4	0.0774
<i>Dependent variable: LM2_DIF</i>			
<i>LP_DIF</i>	4.931841	2	0.0849
<i>LY_DIF</i>	16.29977	2	0.0003
<i>All</i>	21.91118	4	0.0002

Попробуем теперь применить подход Тода и Ямамото. Для этого оценим VAR для переменных LP , LY , $LM2$. Имея в виду идентификацию рядов LP_DIF , LY_DIF , $LM2_DIF$ как стационарных рядов $AR(2)$, $AR(1)$ и $AR(4)$ соответственно (имеющих к тому же нормальные инновации), выберем порядок VAR «с запасом», полагая его равным $p = 8$. В рамках оцененной модели $VAR(8)$ исследуем возможность понижения порядка VAR. Все использованные критерии выбирают глубину запаздываний $p = 2$ (табл. 6.2). С учетом того, что все три ряда имеют порядок $I(1)$, достаточно провести проверку на причинность в рамках модели $VAR(3)$, обращая внимание только на коэффициенты при первых двух лагах. Результаты оценивания приведены в табл. 6.3.

В уравнении для LP не отвергаются ни гипотеза об одновременном занулении коэффициентов при $LY(-1)$ и $LY(-2)$, ни гипотеза об одновременном занулении коэффициентов при $LM2(-1)$ и $LM(-2)$, т.е. не отвергаются и гипотеза о том, что LY не является G -причиной для LP , и гипотеза о том, что $LM2$ не является G -причиной для LP .

В уравнении для LY не отвергаются ни гипотеза об одновременном занулении коэффициентов при $LP(-1)$ и $LP(-2)$, ни гипотеза об одновременном занулении коэффициентов при $LM2(-1)$ и $LM(-2)$, т.е. не отвергаются и гипотеза о том, что LP не является G -причиной для LY , и гипотеза о том, что $LM2$ не является G -причиной для LY .

Таблица 6.2

Сравнение критериев выбора порядка модели

VAR Lag Order Selection Criteria; Endogenous variables: LP LY LM2;
Exogenous variables: C; Sample: 1974M01 1980M03; Included observations: 67

<i>Lag</i>	<i>LogL</i>	<i>LR</i>	<i>FPE</i>	<i>AIC</i>	<i>SC</i>	<i>HQ</i>
0	-556.1126	NA	3555.689	16.68993	16.788650	16.72899
1	-26.52144	995.9476	0.000634	1.149894	1.544764	1.306145
2	3.360465	53.519830*	0.000341*	0.526553*	1.217576*	0.799993*
3	8.609312	8.930873	0.000383	0.638528	1.625704	1.029156
4	12.780580	6.723843	0.000446	0.782669	2.065998	1.290485
5	16.928370	6.314540	0.000522	0.927511	2.506993	1.552516
6	22.232650	7.600164	0.000594	1.037831	2.913465	1.780024
7	27.834510	7.524878	0.000676	1.139268	3.311055	1.998650
8	38.417400	13.268110	0.000670	1.092018	3.559957	2.068588
* indicates lag order selected by the criterion						
LR: sequential modified LR test statistic (each test at 5% level)						
FPE: Final prediction error						
AIC: Akaike information criterion						
SC: Schwarz information criterion						
HQ: Hannan-Quinn information criterion						

Таблица 6.3

Результаты оценивания VAR

Vector Autoregression Estimates; Sample (adjusted): 1974M04 1980M03;
Included observations: 72 after adjustments; Standard errors in () & t-statistics in []

Переменная	<i>LP</i>	<i>LY</i>	<i>LM2</i>
<i>LP</i> (-1)	1.349954 (0.12593) [10.7201]	0.297779 (0.26738) [1.11369]	-0.214976 (0.11509) [-1.86791]
<i>LP</i> (-2)	-0.256495 (0.21100) [-1.21563]	-0.371950 (0.44800) [-0.83024]	0.198518 (0.19284) [1.02947]
<i>LP</i> (-3)	-0.070610 (0.13284) [-0.53155]	0.037640 (0.28205) [0.13345]	0.036485 (0.12140) [0.30053]
<i>LY</i> (-1)	0.011781 (0.05927) [0.19879]	1.206699 (0.12584) [9.58944]	0.156317 (0.05416) [2.88600]

Окончание табл. 6.3

Переменная	LP	LY	$LM2$
$LY(-2)$	0.012690 (0.09186) [0.13815]	-0.331289 (0.19504) [-1.69860]	-0.312633 (0.08395) [-3.72403]
$LY(-3)$	0.012322 (0.06004) [0.20524]	-0.014698 (0.12747) [-0.11530]	0.117330 (0.05487) [2.13838]
$LM2(-1)$	-0.039687 (0.13372) [-0.29678]	-0.148995 (0.28393) [-0.52476]	1.505572 (0.12221) [12.3191]
$LM2(-2)$	-0.074510 (0.21708) [-0.34323]	0.520952 (0.46092) [1.13024]	-0.580950 (0.19840) [-2.92823]
$LM2(-3)$	0.086986 (0.12371) [0.70314]	-0.288486 (0.26267) [-1.09828]	0.075269 (0.11306) [0.66573]
C	-20.37194 (14.1050) [-1.44431]	72.28857 (29.9487) [2.41375]	25.01767 (12.8909) [1.94072]

В уравнении для $LM2$ не отвергается гипотеза об одновременном занулении коэффициентов при $LP(-1)$ и $LP(-2)$, но уверенно отвергается гипотеза об одновременном занулении коэффициентов при $LY(-1)$ и $LY(-2)$, т.е. не отвергается гипотеза о том, что LP не является G -причиной для $LM2$, и отвергается гипотеза о том, что LY не является G -причиной для $LM2$.

По результатам этих проверок уверенно обнаруживается только G -причинная связь в направлении от LY к $LM2$, что дополняет сделанный ранее вывод в отношении краткосрочной G -причинности в направлении от LY_DIF к $LM2_DIF$. ■

ПРИМЕР 6.1.2

Рассмотрим данные об объемах продаж медицинских препаратов, произведенных фармацевтической фирмой *Lydia Pinkham (Sales)*, и о расходах фирмы на рекламу этих препаратов (*Adver*) в период с 1907 по 1960 г. Для сравнения результатов различных исследований оба ряда возьмем в логарифмах уровней.

Оба ряда определяются как $I(1)$ -ряды. Проверка на наличие коинтегрированности дает положительный результат. Ниже приведены результаты оценивания долгосрочного соотношения (табл. 6.4) и модели коррекции ошибок (табл. 6.5).

Таблица 6.4

Оценка долгосрочного соотношения

<i>Cointegrating Eq:</i>	<i>CointEq1</i>
<i>SALES</i> (-1)	1.000000
<i>ADVER</i> (-1)	-0.933287 (0.11309) [-8.25291]
<i>C</i>	-1.141392

Таблица 6.5

Оценка модели коррекции ошибок

<i>Error Correction:</i>	<i>D(SALES)</i>	<i>D(ADVER)</i>
<i>CointEq1</i>	0.137508 (0.12145) [1.13218]	0.730407 (0.20471) [3.56805]
<i>D(SALES</i> (-1))	0.324113 (0.20861) [1.55365]	0.318025 (0.35161) [0.90447]
<i>D(SALES</i> (-2))	-0.000263 (0.19063) [-0.00138]	-0.481982 (0.32131) [-1.50006]
<i>D(ADVER</i> (-1))	0.054991 (0.11379) [0.48327]	0.107971 (0.19179) [0.56296]
<i>D(ADVER</i> (-2))	-0.129445 (0.10104) [-1.28108]	-0.123077 (0.17031) [-0.72268]
<i>C</i>	0.003458 (0.01633) [0.21180]	0.001031 (0.02752) [0.03746]

Корректирующая составляющая значима только в уравнении для $D(ADVER)$, что означает одностороннюю долгосрочную G -причинность в направлении от $SALES$ к $ADVER$.

Краткосрочная G -причинность не обнаруживается в обоих направлениях (табл. 6.6).

Мы провели тестирование для пары рядов — $ADVER$ и $SALES$, не принимая в расчет возможное влияние других рядов. С целью учета такого влияния в работе (Lee, Shin, Chang, 1996) выясняется наличие или отсутствие соотно-

Таблица 6.6

Наличие G-причинности в краткосрочном плане
 VEC Granger Causality/Block Exogeneity Wald Tests; Sample: 1907 1960;
 Included observations: 51

<i>Excluded</i>	<i>Chi-sq</i>	<i>df</i>	<i>Prob.</i>
<i>Dependent variable: D(SALES)</i>			
<i>D(ADVER)</i>	3.114956	2	0.2107
<i>All</i>	3.114956	2	0.2107
<i>Dependent variable: D(ADVER)</i>			
<i>D(SALES)</i>	3.422292	2	0.1807
<i>All</i>	3.422292	2	0.1807

шений $Adver \rightarrow Sales$, $Sales \rightarrow Adver$ (в смысле причинности по Грейнджеру) с дополнительным включением в тестовые уравнения совокупного располагаемого личного дохода *Income*.

Применим к этой тройке рядов подход Тода — Ямамото. Порядок VAR для этих рядов определяется как $p = 2$, так что проверку соответствующих гипотез производим в рамках расширенной VAR(3). Гипотеза об одновременном равенстве нулю коэффициентов при запаздывающих на 1 и 2 шага значениях переменной *Adver* в уравнении для *Sales* отвергается (P -значение F -статистики равно 0.014, асимптотическое P -значение статистики Вальда равно 0.007). Тем самым делается вывод о наличии связи $Adver \rightarrow Sales$, подтверждающий вывод, сделанный в (Lee, Shin, Chang, 1996).■

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Как определяется причинность по Грейнджеру для двух временных рядов?
2. Какие проблемы возникают при определении причинности по Грейнджеру между двумя рядами при наличии третьего ряда?
3. Как определяется причинность по Грейнджеру в случае N временных рядов? Как проверяется гипотеза о том, что переменные второй группы не являются Грейнджер-причиной для i -й переменной первой группы?
4. Что понимается под блочной экзогенностью? Как проверяется гипотеза блочной экзогенности одной группы переменных в отношении другой группы переменных?
5. Какие проблемы возникают при проверке гипотезы об отсутствии причинности по Грейнджеру в рамках уравнения для уровней в случае некоинтегрированной VAR?
6. Какие проблемы возникают при проверке гипотезы об отсутствии причинности по Грейнджеру в случае коинтегрированной VAR?
7. В чем состоит методология Тода — Ямамото проверки на отсутствие причинности по Грейнджеру?
8. Чем различаются понятия причинности в долгосрочном плане и причинности в краткосрочном плане для нестационарной VAR?

Тема 6.2

МЕТОДОЛОГИЯ VAR

В 1932 г. Альфред Коулс (*Alfred Cowles*), президент фирмы *Cowles and Company*, занимавшейся консультациями в области инвестиций и заинтересованной в развитии фундаментальных исследований в экономике, учредил комиссию Коулса (*The Cowles Commission for Research in Economics*). Эта комиссия состояла (и состоит) из экономистов, математиков и статистиков, деятельность которых связана с применением количественных методов в экономике и смежных дисциплинах. В том же 1932 г. было принято решение об издании журнала *Econometrica* (его первый номер вышел в январе 1933 г.), который должен был следовать двум принципам: 1) экономика — это наука; 2) в этой науке очень важен количественный аспект.

Значительный этап в деятельности комиссии начался в 1943 г. с приходом к руководству Якоба Маршака (*Jacob Marschak*). Было отмечено, что методы исследования должны быть обусловлены следующими характеристиками экономических данных и экономической теории:

- 1) теория есть система одновременных уравнений, а не отдельное уравнение;
- 2) некоторые из этих уравнений включают «случайные» составляющие, отражающие многочисленные неустойчивые причины, в дополнение к нескольким «систематическим»;
- 3) многие данные представлены в виде временных рядов, причем последующие события зависят от предыдущих;
- 4) многие опубликованные данные относятся, скорее, к агрегатам, а не к отдельным субъектам.

Развитие соответствующих математических и статистических методов не менее важно, чем немедленное получение результатов. Применение математических результатов в практических исследованиях сопровождается обратным движением: возникновение новых ситуаций в процессе практической работы ставит новые задачи перед математиками.

Значительное место в исследованиях под эгидой комиссии заняла разработка моделей и методов анализа систем одновременных уравнений. Целью таких исследований являлся количественный анализ влияния изменений в переменных, контролируемых неким монетарным полисмейкером (инструменты монетарной политики), на макроэкономические переменные, представляющие главный интерес для этого полисмейкера. Такой анализ предусматривает спецификацию и идентификацию теоретической модели, оценивание параметров и расчет динамических свойств модели с особым акцентом на долговременные свойства, симуляцию динамической модели, анализ последствий различных политик.

При этом спецификация модели производится с априорным разделением переменных на экзогенные и эндогенные. Идентификация достигается как результат накладывания большого количества ограничений. При выявлении тех или иных отклонений от стандартных предположений метода наименьших квадратов модифицируется не модель, а метод оценивания. Базовая симуляция обычно производится на основе имеющейся выборки; ее результаты сравниваются с результатами, полученными в альтернативной симуляции, основанной на модификации соответствующих экзогенных переменных. Анализ альтернативных политик базируется на динамических мультипликаторах. Собственно эконометристы остаются отделенными от выбора модели, который является прерогативой экономистов, формулирующих теоретическую модель.

Методология комиссии Коулса была подвергнута серьезной критике, известной как **критика Лукаса и Симса** ((*Lucas*, 1978), (*Sims*, 1980)).

Согласно Лукасу, традиционные структурные макромоделли бесполезны для целей симуляции политики, поскольку такие модели не принимают в расчет в явной форме ожидания экономических агентов.

Согласно Симсу, в моделях комиссии идентификация достигается за счет произвольного объявления некоторых переменных экзогенными. Однако в мире агентов, поведение которых зависит от решения некоторых вперёдсмотрящих межвременных оптимизационных моделей, никакая из переменных не может считаться экзогенной. Ошибочная трактовка переменной как экзогенной ведет к потере обратной связи и к возникновению кажущейся причинной связи. Кроме того, в теоретическую априорную модель может быть включено недостаточное количество переменных (и тогда возникает эффект пропущенных переменных) и недостаточное количество запаздываний.

Для преодоления указанных недостатков методологии комиссии были предложены:

- методология Лондонской школы экономики (*LSE*);
- методология VAR (векторных авторегрессий).

Методология *LSE* (*Sargan*, *Hendry*)

В этом подходе акценты смещаются с методов оценивания (априорно заданной модели) на получение адекватной данным спецификации модели и на ее идентифицируемость.

Сначала строится достаточно широкая базовая модель в виде векторной *ADL* в приведенной форме с достаточно большим количеством переменных и достаточно большим числом запаздываний (если, конечно, это позволяют данные). Затем эта модель редуцируется путем упрощения динамики (отбрасывания незначимых лагов) и уменьшения размерности (отбрасывания уравнений для тех переменных, для которых не отвергается гипотеза экзогенности). Далее накладываются ограничения на матрицу, определяющую долго-

временное равновесие, и производится идентификация коинтегрирующих векторов. Это приводит к статистической модели для данных с возможным разделением краткосрочной динамики и долгосрочного равновесия; эта модель идентифицируется и оценивается. Если система идентифицируется точно, то на этом все заканчивается. Если же система сверхидентифицирована, то проверяется выполнение «лишних» ограничений.

Методология VAR (Sims)

В связи с критикой Лукаса ключевой чертой новой классической макроэкономики стало разделение изменений в монетарной политике на два типа:

- изменения, которые агенты экономики предвидят правильно;
- изменения, которые являются неожиданными для агентов экономики.

Изменения первого типа должны производить нейтральные эффекты: пропорциональные изменения цен и других номинальных переменных и отсутствие влияния на реальные переменные. Неожиданные изменения (второго типа), напротив, могут отражаться и на реальных переменных.

В методологии Симса акцент делается на исследование откликов системы экономических показателей на **неожиданные (шоковые) воздействия**, которым подвергаются отдельные переменные.

В основе методологии VAR (как и методологии *LSE*) лежит общая структура вида:

$$Ay_t = C(L)y_{t-1} + \zeta_t = C_1y_{t-1} + \dots + C_p y_{t-p} + \zeta_t, \quad y_t = (y_{1t}, \dots, y_{kt})^T,$$

где ζ_t — **инновационная последовательность *i.i.d.* случайных ($k \times 1$)-векторов с нулевым математическим ожиданием.**

Приведенная форма этой структуры имеет вид:

$$y_t = \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t,$$

где $\Pi_j = A^{-1}C_j$, $u_t = A^{-1}\zeta_t$, или

$$\Pi(L)y_t = u_t, \quad \Pi(L) = I_k - \Pi_1 L - \dots - \Pi_p L^p.$$

При этом u_t также является **инновационной последовательностью *i.i.d.* случайных векторов, имеющих нулевое математическое ожидание и ковариационную матрицу Σ .**

Если рассмотрим простую VAR(1) для двух рядов ($k = 2, p = 1$):

$$\begin{aligned} y_t &= \Pi_1 y_{t-1} + u_t, \\ y_{1t} &= \pi_{11} y_{1,t-1} + \pi_{12} y_{2,t-1} + u_{1t}, \\ y_{2t} &= \pi_{21} y_{1,t-1} + \pi_{22} y_{2,t-1} + u_{2t}, \end{aligned}$$

то непосредственно в рамках такой приведенной системы можно выяснить, на сколько изменяются значения $y_{1,t+s}$, $y_{2,t+s}$ при изменении инновации u_{1t} или u_{2t} на одно стандартное отклонение. Вычислив эти изменения последовательно для значений $s = 0, 1, \dots$, получим функции откликов переменных y_{1t} и y_{2t} на шоки инноваций.

Пусть, например, $\pi_{11} = \frac{5}{8}$, $\pi_{12} = \frac{1}{2}$, $\pi_{21} = \frac{1}{4}$, $\pi_{22} = \frac{5}{8}$, так что

$$\begin{aligned} y_{1t} &= \frac{5}{8}y_{1,t-1} + \frac{1}{2}y_{2,t-1} + u_{1t}, \\ y_{2t} &= \frac{1}{4}y_{1,t-1} + \frac{5}{8}y_{2,t-1} + u_{2t}. \end{aligned}$$

Эту VAR можно записать в виде:

$$\begin{aligned} (1 - \frac{5}{8}L)y_{1t} + (-\frac{1}{2}L)y_{2t} &= u_{1t}, \\ (-\frac{1}{4}L)y_{1t} + (1 - \frac{5}{8}L)y_{2t} &= u_{2t}, \end{aligned}$$

или

$$\begin{pmatrix} (1 - \frac{5}{8}L) & -\frac{1}{2}L \\ -\frac{1}{4}L & (1 - \frac{5}{8}L) \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}.$$

При этом

$$\begin{aligned} \det \begin{pmatrix} (1 - \frac{5}{8}z) & -\frac{1}{2}z \\ -\frac{1}{4}z & (1 - \frac{5}{8}z) \end{pmatrix} &= \left(1 - \frac{5}{8}z\right)^2 - \frac{1}{8}z^2 = \left(1 - \frac{5}{8}z\right)^2 - \left(\frac{\sqrt{2}}{4}z\right)^2 = \\ &= \left(1 - \frac{5}{8}z + \frac{\sqrt{2}}{4}z\right) \left(1 - \frac{5}{8}z - \frac{\sqrt{2}}{4}z\right) = \left(1 - \left(\frac{5+2\sqrt{2}}{8}\right)z\right) \left(1 - \left(\frac{5-2\sqrt{2}}{8}\right)z\right), \end{aligned}$$

корни уравнения:

$$\det \begin{pmatrix} (1 - \frac{5}{8}z) & -\frac{1}{2}z \\ -\frac{1}{4}z & (1 - \frac{5}{8}z) \end{pmatrix} = 0$$

лежат за пределами единичного круга, так что рассматриваемая VAR стабильна (а значит, стационарна), и можно записать:

$$\begin{aligned}
\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} &= \begin{bmatrix} \left(1 - \frac{5}{8}L\right) & -\frac{1}{2}L \\ -\frac{1}{4}L & \left(1 - \frac{5}{8}L\right) \end{bmatrix}^{-1} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \\
&= \frac{1}{\left(1 - \left(\frac{5+2\sqrt{2}}{8}\right)L\right)\left(1 - \left(\frac{5-2\sqrt{2}}{8}\right)L\right)} \begin{bmatrix} \left(1 - \frac{5}{8}L\right) & \frac{1}{2}L \\ \frac{1}{4}L & \left(1 - \frac{5}{8}L\right) \end{bmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \\
&= \left(1 + \left(\frac{5+2\sqrt{2}}{8}\right)L + \left(\frac{5+2\sqrt{2}}{8}\right)^2 L^2 + \dots\right) \times \\
&\times \left(1 + \left(\frac{5-2\sqrt{2}}{8}\right)L + \left(\frac{5-2\sqrt{2}}{8}\right)^2 L^2 + \dots\right) \begin{bmatrix} \left(1 - \frac{5}{8}L\right) & \frac{1}{2}L \\ \frac{1}{4}L & \left(1 - \frac{5}{8}L\right) \end{bmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \\
&= \left(1 + \frac{5}{4}L + \dots\right) \begin{bmatrix} \left(1 - \frac{5}{8}L\right) & \frac{1}{2}L \\ \frac{1}{4}L & \left(1 - \frac{5}{8}L\right) \end{bmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \\
&= \begin{pmatrix} 1 + \frac{5}{8}L + \dots & \frac{1}{2}L + \dots \\ \frac{1}{4}L + \dots & 1 + \frac{5}{8}L + \dots \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} + \begin{pmatrix} \frac{5}{8} & \frac{1}{2} \\ \frac{1}{4} & \frac{5}{8} \end{pmatrix} \begin{pmatrix} u_{1,t-1} \\ u_{2,t-1} \end{pmatrix} + \dots
\end{aligned}$$

или

$$y_t = u_t + \Psi_1 u_{t-1} + \Psi_2 u_{t-2} + \Psi_3 u_{t-3} + \dots \equiv \Psi(L)u_t,$$

где $\Psi_1 = \begin{pmatrix} \frac{5}{8} & \frac{1}{2} \\ \frac{1}{4} & \frac{5}{8} \end{pmatrix}$ и т.д.

Тем самым есть возможность построить функции откликов обеих переменных на шоки инноваций. Однако из-за того что в общем случае $Cov(u_{1,t}, u_{2,t}) \neq 0$, возникают затруднения с интерпретацией этих функций. Из-за коррелированности инноваций u_{1t} и u_{2t} в приведенной форме невозможно полностью изолировать шок для u_{1t} от u_{2t} , т.е. нельзя произвольно изменять значение u_{1t} , сохраняя при этом значения u_{2t} , $u_{2,t+1}$, ..., $u_{2,t+s}$, ... неизменными.

Для преодоления этого затруднения предполагают, что система изменяется благодаря воздействию не коррелированных между собой **фундаментальных инноваций** (*fundamental innovations*) $\varepsilon_{1t}, \dots, \varepsilon_{kt}$. Обычно предполагается, что все они имеют единичные дисперсии, так что $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})^T$ — *i.i.d.* случайные векторы с нулевым математическим ожиданием и единичной ковариационной матрицей I_k . При этом предполагается, что инновации u_{1t}, \dots, u_{kt} являются линейными комбинациями фундаментальных инноваций, так что

$$u_t = D\varepsilon_t.$$

В примере с двумерной $VAR(1)$: поскольку $u_t = D\varepsilon_t$, то

$$\begin{aligned} y_t &= \Pi_1 y_{t-1} + D\varepsilon_t, \\ y_{1t} &= (\pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1}) + (d_{11}\varepsilon_{1t} + d_{12}\varepsilon_{2t}), \\ y_{2t} &= (\pi_{21}y_{1,t-1} + \pi_{22}y_{2,t-1}) + (d_{21}\varepsilon_{1t} + d_{22}\varepsilon_{2t}). \end{aligned}$$

Отсюда видно, что изменение u_{1t} на одно стандартное отклонение складывается из изменений ε_{1t} и ε_{2t} , которые, в свою очередь, вызывают одновременное изменение u_{2t} .

С точки зрения экономической теории первоочередной интерес представляют реакции значений $y_{1,t+s}, y_{2,t+s}$ на единичные импульсные изменения отдельных фундаментальных инноваций $\varepsilon_{1t}, \dots, \varepsilon_{kt}$ при фиксированных значениях всех остальных фундаментальных инноваций во все моменты времени. Именно на построение таких **функций импульсного отклика** (*impulse response function* — *IRF*) нацелены алгоритмы, реализуемые в пакетах статистических программ.

Если матрица $\Sigma = \text{Cov}(u_t) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$ известна, то как найти матрицу D ?

Заметим, что:

$$\begin{aligned} \sigma_{11} &= D(u_{1t}) = d_{11}^2 + d_{12}^2, & \sigma_{22} &= D(u_{2t}) = d_{21}^2 + d_{22}^2, \\ \sigma_{12} &= \text{Cov}(u_{1t}, u_{2t}) = d_{11}d_{21} + d_{12}d_{22}, \end{aligned}$$

так что для нахождения четырех неизвестных $d_{11}, d_{12}, d_{21}, d_{22}$ имеется всего три уравнения. Следовательно, матрицу D идентифицировать невозможно, если не накладывать априорных ограничений на ее структуру.

Симс предложил упорядочивать инновации в системе. Поясним это на примере двумерной $VAR(1)$. Пусть в этой VAR фундаментальная инновация ε_{2t} воздействует только на y_{2t} , а фундаментальная инновация ε_{1t} воздействует и на y_{1t} , и на y_{2t} . Тогда фундаментальная инновация ε_{2t} воздействует только на u_{2t} , а фундаментальная инновация ε_{1t} — и на u_{1t} , и на u_{2t} (рис. 6.2).

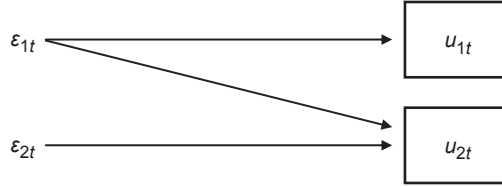


Рис. 6.2

При этом $d_{12} = 0$, и есть три уравнения для нахождения d_{11} , d_{21} , d_{22} :

$$\sigma_{11} = d_{11}^2, \quad \sigma_{22} = d_{21}^2 + d_{22}^2, \quad \sigma_{12} = d_{11}d_{21}.$$

Решая эти уравнения, находим:

$$D = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}} & \sqrt{\sigma_{22} - \frac{\sigma_{12}^2}{\sigma_{11}}} \end{pmatrix}.$$

Представим матрицу D в виде произведения двух матриц, одна из которых является нижней треугольной, а другая — диагональна:

$$D = P^* B,$$

где

$$P^* = \begin{pmatrix} 1 & 0 \\ \frac{d_{21}}{d_{11}} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} d_{11} & 0 \\ 0 & d_{22} \end{pmatrix}.$$

Тогда $\text{VAR}(1)$ записывается в виде:

$$y_t = \Pi_1 y_{t-1} + P^* B \varepsilon_t.$$

Обозначив $(P^*)^{-1} = A$, заметим, что

$$(P^*)^{-1} = \begin{pmatrix} 1 & 0 \\ -\frac{d_{21}}{d_{11}} & 1 \end{pmatrix},$$

и умножим обе части $\text{VAR}(1)$ на матрицу A :

$$A y_t = A \Pi_1 y_{t-1} + A P^* B \varepsilon_t = C y_{t-1} + B \varepsilon_t.$$

Таким образом, получим структурную модель

$$A y_t = C(L) y_{t-1} + \zeta_t,$$

в которой $\zeta_t = B \varepsilon_t$.

Пусть в двумерной VAR(1) ковариационная матрица $\Sigma = \text{Cov}(u_t)$ имеет вид:

$$\Sigma = \text{Cov}(u_t) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \text{так что } D = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Тогда структурная VAR имеет вид:

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{11} + \pi_{21} & \pi_{12} + \pi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

т.е.

$$\begin{aligned} y_{1t} &= \pi_{11}y_{1,t-1} + \pi_{12}y_{2,t-1} + \varepsilon_{1t}, \\ y_{1t} + y_{2t} &= (\pi_{11} + \pi_{21})y_{1,t-1} + (\pi_{12} + \pi_{22})y_{2,t-1} + \varepsilon_{2t}. \end{aligned} \quad (6.5)$$

Изменим теперь упорядочение инноваций (рис. 6.3).

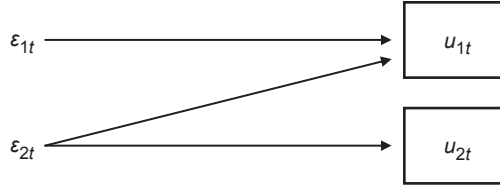


Рис. 6.3

При этом $d_{21} = 0$,

$$\begin{cases} u_{1t} = d_{11}\varepsilon_{1t} + d_{12}\varepsilon_{2t}, \\ u_{2t} = d_{22}\varepsilon_{2t}, \end{cases}$$

и есть три уравнения для нахождения d_{11} , d_{12} , d_{22} :

$$\sigma_{11} = d_{11}^2 + d_{12}^2, \quad \sigma_{22} = d_{22}^2, \quad \sigma_{12} = d_{12}d_{22}.$$

Решая эти уравнения, находим:

$$D = \begin{pmatrix} \sqrt{\sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{22}}} \\ 0 & \sqrt{\sigma_{22}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \sqrt{2} \end{pmatrix}.$$

Представим матрицу D в виде произведения:

$$D = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}.$$

Структурная VAR имеет вид:

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{2}}{2} \varepsilon_{1t} \\ \sqrt{2} \varepsilon_{2t} \end{pmatrix},$$

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \pi_{11} + \frac{1}{2} \pi_{21} & \pi_{12} + \frac{1}{2} \pi_{22} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{2}}{2} \varepsilon_{1t} \\ \sqrt{2} \varepsilon_{2t} \end{pmatrix},$$

т.е.

$$\begin{aligned} y_{1t} + \frac{1}{2} y_{2t} &= \left(\pi_{11} + \frac{1}{2} \pi_{21} \right) y_{1,t-1} + \left(\pi_{12} + \frac{1}{2} \pi_{22} \right) y_{2,t-1} + \frac{\sqrt{2}}{2} \varepsilon_{1t}, \\ y_{2t} &= \pi_{21} y_{1,t-1} + \pi_{22} y_{2,t-1} + \sqrt{2} \varepsilon_{2t}. \end{aligned} \quad (6.6)$$

В обоих случаях выбранная форма матрицы A приводит к рекурсивной системе, так как при соответствующей нумерации переменных:

- в первое уравнение с текущим значением входит только одна переменная y_{1t} , т.е. y_{1t} объясняется только запаздывающими значениями переменных y_{1t}, y_{2t} ;
- во второе уравнение с текущими значениями входят обе переменные y_{1t} и y_{2t} , так что y_{2t} объясняется с помощью y_{1t} и y_{t-1}, y_{t-2}, \dots

Такой порядок вхождения переменных интерпретируется как последовательное включение переменных в порядке возрастания их эндогенности, так что последней в систему включается «наиболее эндогенная» переменная.

Вернемся к двумерной VAR(1):

$$\begin{aligned} y_{1t} &= \frac{5}{8} y_{1,t-1} + \frac{1}{2} y_{2,t-1} + u_{1t}, \\ y_{2t} &= \frac{1}{4} y_{1,t-1} + \frac{5}{8} y_{2,t-1} + u_{2t}, \end{aligned}$$

и сравним функции откликов переменных y_{1t} и y_{2t} на импульсный шок фундаментальной инновации ε_{1t} .

При первом упорядочении:

$$\begin{aligned} y_{1t} &= \frac{5}{8} y_{1,t-1} + \frac{1}{2} y_{2,t-1} + \varepsilon_{1t}, \\ y_{2t} &= \frac{1}{4} y_{1,t-1} + \frac{5}{8} y_{2,t-1} + \varepsilon_{2t} - \varepsilon_{1t}. \end{aligned}$$

Пусть в момент $t = 1$ имеет место шок фундаментальной инновации для переменной y_{1t} (в первом уравнении рекурсивной системы), так что

$$\tilde{\varepsilon}_{1t} = \begin{cases} \varepsilon_{1t} + 1 & \text{для } t = 1, \\ \varepsilon_{1t} & \text{для } t > 1, \end{cases}$$

а ε_{2t} не изменяется ни при каком t . При этом получим измененную реализацию

$\tilde{y}_t = \begin{pmatrix} \tilde{y}_{1t} \\ \tilde{y}_{2t} \end{pmatrix}$, для которой $\tilde{y}_{1,0} = y_{1,0}$, $\tilde{y}_{2,0} = y_{2,0}$ и

$$\begin{aligned} \tilde{y}_{1t} &= \frac{5}{8}\tilde{y}_{1,t-1} + \frac{1}{2}\tilde{y}_{2,t-1} + \tilde{\varepsilon}_{1t}, \\ \tilde{y}_{2t} &= \frac{1}{4}\tilde{y}_{1,t-1} + \frac{5}{8}\tilde{y}_{2,t-1} + \varepsilon_{2t} - \tilde{\varepsilon}_{1t}. \end{aligned}$$

Для простоты пусть $y_{1,0} = y_{2,0} = 0$. Тогда при $t = 1$:

$$\begin{aligned} y_{11} &= \varepsilon_{11}, & \tilde{y}_{11} &= \varepsilon_{11} + 1 = y_{11} + 1, \\ y_{21} &= \varepsilon_{21} - \varepsilon_{11}, & \tilde{y}_{21} &= \varepsilon_{21} - \varepsilon_{11} - 1 = y_{21} - 1, \end{aligned}$$

а при $t = 2$:

$$\begin{aligned} \tilde{y}_{12} &= \frac{5}{8}\tilde{y}_{1,1} + \frac{1}{2}\tilde{y}_{2,1} + \tilde{\varepsilon}_{12} = y_{12} + \frac{1}{8}, \\ \tilde{y}_{22} &= \frac{1}{4}\tilde{y}_{1,1} + \frac{5}{8}\tilde{y}_{2,1} + \varepsilon_{22} - \tilde{\varepsilon}_{12} = y_{22} - \frac{3}{8}. \end{aligned}$$

Таким образом, получили значения функций импульсного отклика на единичный шок инновации ε_{1t} (табл. 6.7).

Таблица 6.7

Отклик переменных на шок инновации (первое упорядочение)

t	1	2
$Imp(y_{1t} \varepsilon_{11})$	1	0.125
$Imp(y_{2t} \varepsilon_{11})$	-1	-0.375

При втором упорядочении:

$$\begin{aligned} y_{2t} &= \frac{1}{4}y_{1,t-1} + \frac{5}{8}y_{2,t-1} + \sqrt{2}\varepsilon_{2t}, \\ y_{1t} &= \frac{5}{8}y_{1,t-1} + \frac{1}{2}y_{2,t-1} + \frac{\sqrt{2}}{2}\varepsilon_{1t} - \frac{\sqrt{2}}{2}\varepsilon_{2t}. \end{aligned}$$

При $t = 1$:

$$\begin{aligned} y_{21} &= \sqrt{2} \varepsilon_{21}, & \tilde{y}_{21} &= \sqrt{2} \varepsilon_{21} = y_{21}, \\ y_{11} &= \frac{\sqrt{2}}{2} \varepsilon_{11} - \frac{\sqrt{2}}{2} \varepsilon_{21}, & \tilde{y}_{11} &= y_{11} + \frac{\sqrt{2}}{2}. \end{aligned}$$

При $t = 2$:

$$\begin{aligned} \tilde{y}_{22} &= \frac{1}{4} \tilde{y}_{1,1} + \frac{5}{8} \tilde{y}_{2,1} + \sqrt{2} \varepsilon_{22} = y_{22} + \frac{\sqrt{2}}{8}, \\ \tilde{y}_{12} &= \frac{5}{8} \tilde{y}_{1,1} + \frac{1}{2} \tilde{y}_{2,1} + \frac{\sqrt{2}}{2} \varepsilon_{12} - \frac{\sqrt{2}}{2} \varepsilon_{22} = y_{12} + \frac{5\sqrt{2}}{8}. \end{aligned}$$

В этом случае получили другие значения функций импульсного отклика на единичный шок инновации ε_{1t} (табл. 6.8).

Таблица 6.8

Отклик переменных на шок инновации (второе упорядочение)

t	1	2
$Imp(y_{1t} \varepsilon_{11})$	0.707000	0.441875
$Imp(y_{2t} \varepsilon_{11})$	0	0.17675

Принимая различный порядок последовательного вхождения переменных, получили и различное поведение функций импульсного отклика.

Проиллюстрируем графически поведение этих функций (рис. 6.4 соответствует первому упорядочению, рис. 6.5 — второму упорядочению).

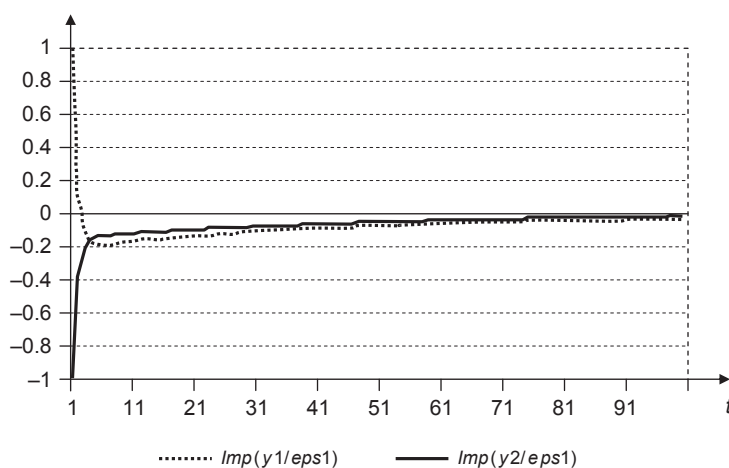


Рис. 6.4

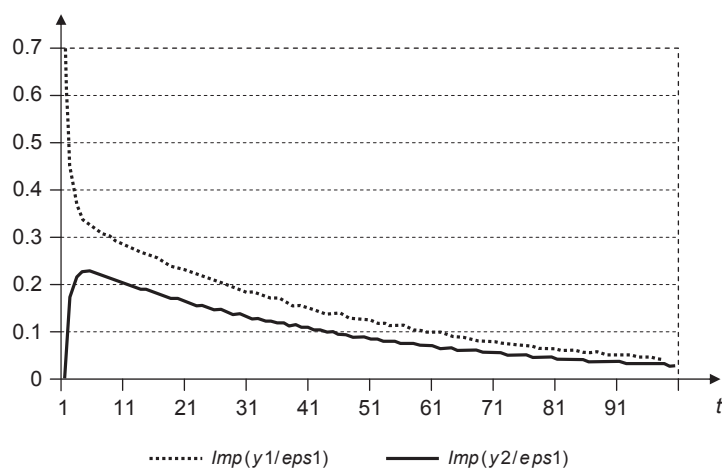


Рис. 6.5

Заметим, что мы построили функции импульсного отклика, опираясь на известные коэффициенты приведенной формы и известную ковариационную матрицу вектора инноваций в приведенной форме. Посмотрим, что получится, если считать эти параметры неизвестными и использовать для построения функций отклика их оценки, вычисленные по смоделированным реализациям длины 100. Напомним, что в нашем примере $\pi_{11} = \pi_{22} = \frac{5}{8} = 0.625$,

$$\pi_{12} = \frac{1}{2} = 0.5, \quad \pi_{21} = \frac{1}{4} = 0.25.$$

При первом упорядочении $Y1 \rightarrow Y2$ получим результаты, приведенные в табл. 6.9 и 6.10; при втором упорядочении $Y2 \rightarrow Y1$ — результаты, приведенные в табл. 6.11 и 6.12.

Таблица 6.9

Оценки коэффициентов при первом упорядочении
Vector Autoregression Estimates; Первый вариант упорядочения $Y1 \rightarrow Y2$

Переменная	Y1	Y2
Y1(-1)	0.610340 (0.04187) [14.5766]	0.206764 (0.06525) [3.16883]
Y2(-1)	0.497220 (0.04764) [10.4360]	0.664915 (0.07425) [8.95554]

Таблица 6.10

Импульсные отклики при первом упорядочении

t	Отклик Y1			Отклик Y2	
	Оценка	Точное значение		Оценка	Точное значение
1	0.958889	1.0000000		-0.947857	-1.0000000
2	0.113956	0.1250000		-0.431981	-0.3750000
3	-0.145238	-0.1093750		-0.263669	-0.2031250
4	-0.219746	-0.1699218		-0.205347	-0.1542969
5	-0.236222	-0.1833496		-0.181974	-0.1389160
6	-0.234657	-0.1840515		-0.169839	-0.1326600
7	-0.227668	-0.1813621		-0.161447	-0.1289253
8	-0.219229	-0.1778140		-0.154422	-0.1259189
9	-0.210586	-0.1740932		-0.148006	-0.1231530
10	-0.202121	-0.1703850		-0.141953	-0.1204940
11	-0.193944	-0.1667380		-0.136178	-0.1179048
12	-0.186082	-0.1631632		-0.130647	-0.1153740

Таблица 6.11

Импульсные отклики при втором упорядочении

t	Отклик Y1			Отклик Y2	
	Оценка	Точное значение		Оценка	Точное значение
1	0.674760	0.707000		0.000000	0.000000
2	0.457393	0.441875		0.103987	0.176750
3	0.353585	0.364547		0.146912	0.220937
4	0.301190	0.338311		0.162461	0.229223
5	0.272183	0.326055		0.165814	0.227842
6	0.253924	0.317706		0.163807	0.223915
7	0.240707	0.310523		0.159519	0.219373
8	0.229952	0.303764		0.154330	0.214739
9	0.220490	0.297222		0.148860	0.210153
10	0.211785	0.290840		0.143381	0.205652
11	0.203591	0.284601		0.138013	0.201242
12	0.195789	0.278496		0.132805	0.196927

Таблица 6.12

Значения переменных при втором упорядочении
Vector Autoregression Estimates; Первый вариант упорядочения $Y2 \rightarrow Y1$

Переменная	$Y1$	$Y2$
$Y1(-1)$	0.677859 (0.03880) [17.4726]	0.154110 (0.06036) [2.55303]
$Y2(-1)$	0.418677 (0.04350) [9.62382]	0.734929 (0.06769) [10.8573]

Сравним теперь функции импульсного отклика, построенные по VAR, оцененной по полученным реализациям, и по VAR, породившей эти реализации. Первому упорядочению $Y1 \rightarrow Y2$ соответствуют рис. 6.6 и 6.7; второму упорядочению $Y2 \rightarrow Y1$ — рис. 6.8 и 6.9.

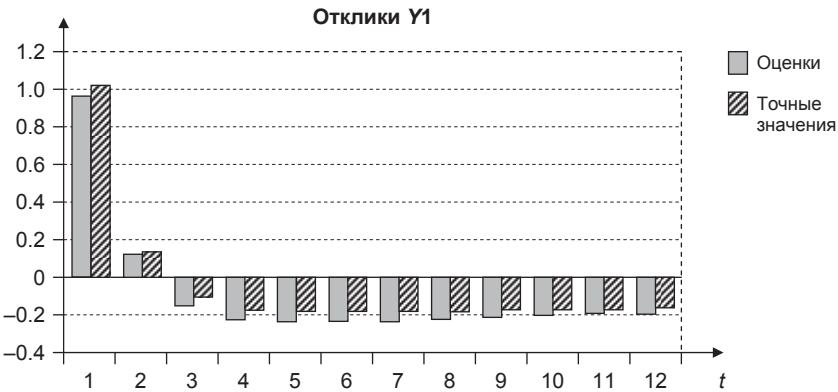


Рис. 6.6



Рис. 6.7

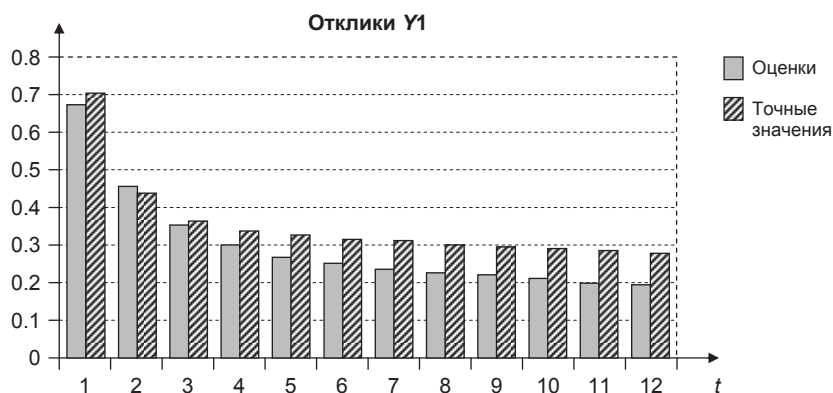


Рис. 6.8



Рис. 6.9

Следует обратить внимание на построенные на основании асимптотических распределений доверительные интервалы для (отдельных!) значений импульсных откликов: рис. 6.10 соответствует упорядочению $Y1 \rightarrow Y2$; рис. 6.11 — упорядочению $Y2 \rightarrow Y1$.

Заметим для дальнейшего, что столь существенное различие в поведении функций импульсного отклика при альтернативных упорядочениях связано с существенной перекрестной коррелированностью инноваций в приведенной форме VAR. В сгенерированных данных эта корреляция равна -0.628 , а для остатков от оцененной VAR она равна -0.634 .

В общем случае предполагают, что структурная модель VAR имеет вид:

$$Ay_t = C(L)y_{t-1} + B\varepsilon_t = C_1y_{t-1} + \dots + C_p y_{t-p} + B\varepsilon_t,$$

т.е.

$$\zeta_t = B\varepsilon_t,$$

где $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})^T$ — *i.i.d.* случайные векторы с нулевым математическим ожиданием и единичной ковариационной матрицей I_k .

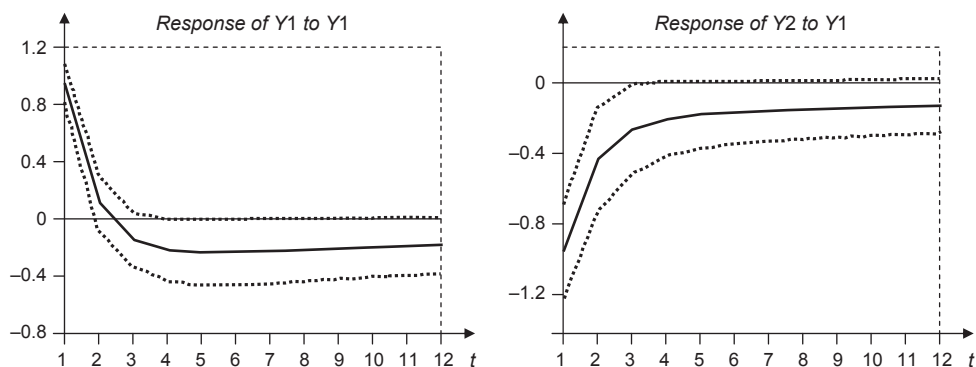


Рис. 6.10

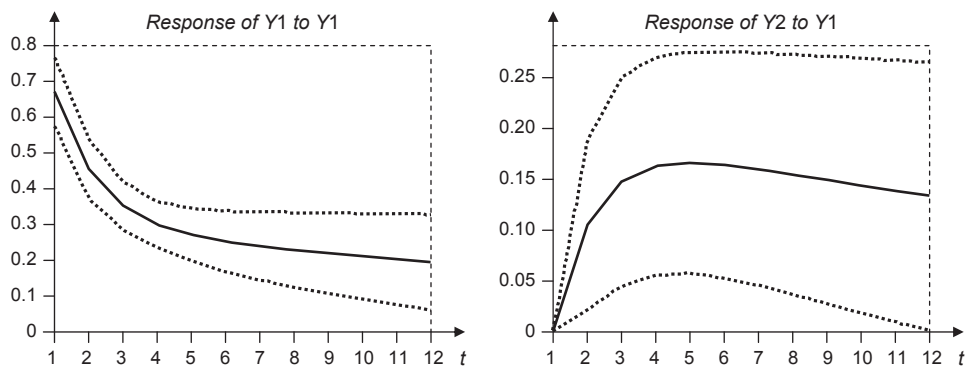


Рис. 6.11

Так что инновации $\zeta_{1t}, \dots, \zeta_{k,t}$ в структуре являются линейными комбинациями **ортогонализированных (фундаментальных) инноваций** $\varepsilon_{1t}, \dots, \varepsilon_{kt}$. Соответственно приведенная форма имеет вид:

$$y_t = A^{-1}C(L)y_{t-1} + A^{-1}B\varepsilon_t = \Pi^*(L)y_{t-1} + u_t = \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t,$$

где $\Pi_j = A^{-1}C_j$, $\Pi^*(L) = \Pi_1 + \dots + \Pi_p L^{p-1}$,

$$u_t = A^{-1}B\varepsilon_t = D\varepsilon_t,$$

так что

$$Ay_t = A\Pi^*(L)y_{t-1} + B\varepsilon_t. \quad (6.7)$$

Можно без труда оценить VAR в приведенной форме, используя *OLS*. Однако, получив оценки матриц Π_1, \dots, Π_p , не всегда можно получить через них оценку представления (6.7), поскольку не известны матрицы A и B . Это

связано с общей проблемой идентификации структурных уравнений по приведенной форме системы (см. разд. 1 ч. 3 данного учебника). Заметим, что в рассматриваемой ситуации

$$\Sigma = \text{Cov}(u_t) = \text{Cov}(A^{-1}B\varepsilon_t) = A^{-1}B\text{Cov}(\varepsilon_t)(A^{-1}B)^T = A^{-1}BB^T(A^{-1})^T.$$

Оценив приведенную VAR, можно получить и оценку $\hat{\Sigma}$ ковариационной матрицы Σ . При этом замена Σ на $\hat{\Sigma}$ приводит к оценочному уравнению

$$\hat{\Sigma} = \hat{A}^{-1}\hat{B}\hat{B}^T(\hat{A}^{-1})^T$$

для A и B . Матрица $\hat{\Sigma}$ содержит число сочетаний из k по 2 плюс k , т.е. $\binom{k}{2} + k = \frac{k^2 + k}{2}$, различных элементов, тогда как общее количество неизвест-

ных элементов в матрицах \hat{A} и \hat{B} равно $2k^2$. (В частности, при $k = 2$ в матрице $\hat{\Sigma}$ имеется три различных элемента, а в матрицах \hat{A} и \hat{B} — 8 неизвестных элементов.) Поэтому идентификация возможна лишь при наложении на матрицы A и B (соответственно на матрицы \hat{A} и \hat{B}) достаточного количества ограничений. Этот шаг исследования исторически реализовывался различными способами. Методология Симса, которую мы уже применяли к VAR(1) с двумя переменными, фактически основана на следующем результате (**разложение Холецкого**).

Всякую положительно полуопределенную матрицу Σ можно представить в виде произведения $\Sigma = Q^T Q$, где Q — верхняя треугольная матрица с положительными диагональными элементами, причем такая матрица Q единственна.

Заметим, что определитель матрицы Q равен произведению диагональных элементов этой матрицы, а потому отличен от нуля. Следовательно, матрица Q имеет обратную, которая также является верхней треугольной. Заметим также, что транспонированная матрица $D = Q^T$ является нижней треугольной матрицей с положительными диагональными элементами:

$$D = \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ d_{21} & d_{22} & 0 & \cdots & 0 \\ d_{31} & d_{32} & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{k1} & d_{k2} & d_{k3} & \cdots & d_{kk} \end{pmatrix},$$

и при этом

$$\Sigma = DD^T.$$

В свою очередь, матрицу D можно представить в виде произведения:

$$D = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ d_{21}^* & 1 & 0 & \cdots & 0 \\ d_{31}^* & d_{32}^* & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{k1}^* & d_{k2}^* & d_{k3}^* & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{kk} \end{pmatrix} =$$

$$= D^* \begin{pmatrix} d_{11} & 0 & 0 & \cdots & 0 \\ 0 & d_{22} & 0 & \cdots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{kk} \end{pmatrix}.$$

Приравняем:

$$A^{-1} = D^*, \quad B = \begin{pmatrix} d_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_{kk} \end{pmatrix}.$$

Тогда матрица $A = D^{*-1}$ — нижняя треугольная с единицами на диагонали:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & 1 \end{pmatrix}.$$

Итак, если наложить на матрицы A и B ограничения:

A — нижняя треугольная матрица с единицами на диагонали;

B — диагональная матрица с положительными диагональными элементами,

то уравнение $\Sigma = A^{-1}B(A^{-1}B)^T$ имеет единственное решение, т.е. имеет место точная идентифицируемость матриц A и B :

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & 1 \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{kk} \end{pmatrix}.$$

При этом b_{ii} можно рассматривать как среднее квадратическое отклонение инновации в i -м уравнении структуры, а форма матрицы A соответствует рекурсивной системе:

- в первое уравнение с текущим значением входит только одна переменная y_{1t} , т.е. y_{1t} объясняется только запаздывающими значениями переменных y_{1t}, \dots, y_{kt} ;
- во второе уравнение с текущими значениями входят только переменные y_{1t} и y_{2t} , так что y_{2t} объясняется с помощью y_{1t} и y_{t-1}, y_{t-2}, \dots ;
- ...
- наконец, в k -е уравнение входят текущие значения всех переменных y_{1t}, \dots, y_{kt} , так что y_{kt} объясняется с помощью $y_{1t}, \dots, y_{k-1,t}$ и y_{t-1}, y_{t-2}, \dots

Такой порядок вхождения переменных интерпретируется как последовательное включение переменных в порядке возрастания их эндогенности, так что последней в систему включается «наиболее эндогенная» переменная. При выбранной нормализации (единицы на диагонали матрицы A) y_{1t} — «наименее эндогенная» переменная, y_{kt} — «наиболее эндогенная» переменная.

В рекурсивной структуре, полученной с использованием изложенного метода, случайные ошибки в разных уравнениях являются не коррелированными между собой. Это означает, что соответствующую систему одновременных уравнений можно оценивать, используя обычный метод наименьших квадратов (*OLS*).

Вернемся к приведенной форме k -мерной VAR(p):

$$y_t = \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t,$$

которую мы уже записывали в виде:

$$\Pi(L)y_t = u_t, \quad \Pi(L) = I_k - \Pi_1 L - \dots - \Pi_p L^p.$$

Если все корни уравнения $\det \Pi(z) = 0$ лежат за пределами единичного круга, то VAR стабильна, и можно записать:

$$\begin{aligned} y_t &= \Pi(L)^{-1} u_t = (I_k - \Pi_1 L - \dots - \Pi_p L^p)^{-1} u_t = \\ &= u_t + \Psi_1 u_{t-1} + \Psi_2 u_{t-2} + \Psi_3 u_{t-3} + \dots \equiv \Psi(L) u_t, \end{aligned}$$

где

$$\Psi(L) = I_k + \Psi_1 L + \Psi_2 L^2 + \Psi_3 L^3 + \dots$$

Это есть **векторное МА-представление** (*vector moving average* — *VMA*) k -мерного ряда y_t , **основанное на последовательности инноваций u_t приведенной формы VAR**. При этом последовательность матричных коэффициентов $\psi_1, \psi_2, \psi_3, \dots$ абсолютно суммируема, т.е. если $\psi_{ij}^{(k)}$ — элемент

матрицы ψ_k с индексом ij , то $\sum_{k=1}^{\infty} |\psi_{ij}^{(k)}| < \infty$. Отсюда непосредственно вытекает, что

$$\lim_{k \rightarrow \infty} \psi_{ij}^{(k)} = 0 \text{ для любой пары } (i, j). \quad (6.8)$$

Если $u_t = D\varepsilon_t$, где D — нижняя треугольная матрица из представления $\Sigma = DD^T$, то

$$y_t = D\varepsilon_t + \Psi_1 D\varepsilon_{t-1} + \Psi_2 D\varepsilon_{t-2} + \Psi_3 D\varepsilon_{t-3} + \dots$$

Это есть **VMA-представление** k -мерного ряда y_t , **основанное на последовательности фундаментальных инноваций** ε_t . Элемент матрицы $\psi_h D$ с индексом ij равен изменению — **импульсному отклику** (*impulse response*) i -й переменной в момент времени $(t + h)$ в ответ на единичное изменение шока j -й переменной в момент времени t при сохранении неизменными всех остальных шоков во все моменты времени. Совокупность таких откликов для значений $h = 1, 2, \dots$ образует **функцию импульсного отклика** (*impulse response function* — *IRF*).

✓ **Замечание 6.2.1.** Из соотношения (6.8) следует, что у стабильной VAR функции импульсного отклика стремятся к нулю при $h \rightarrow \infty$.

✓ **Замечание 6.2.2.** Принимая различные порядки последовательного вхождения переменных, получаем и различное поведение функций импульсного отклика, что дает возможность сравнивать альтернативные теории.

В модели двумерной VAR переменная y_{1t} могла бы представлять объем производства (*output*), а переменная y_{2t} — деньги (*money*). Тогда упорядочение $y_{1t} \rightarrow y_{2t}$ соответствует схеме, изображенной на рис. 6.12. В этой схеме шоки в объеме производства оказывают немедленное воздействие и на объем производства, и на деньги, тогда как шоки в деньгах оказывают немедленное воздействие только на деньги. Такое упорядочение соответствует представлению, согласно которому денежная политика имеет только запаздывающее влияние на объем производства.

Можно рассмотреть и альтернативную схему, соответствующую упорядочению $y_{2t} \rightarrow y_{1t}$ (рис. 6.13). В этой схеме шоки в объеме производства оказывают немедленное воздействие только на объем производства, тогда как шоки в деньгах оказывают немедленное воздействие и на деньги, и на объем производства. Это соответствует представлению о том, что деньги поставляются центральным

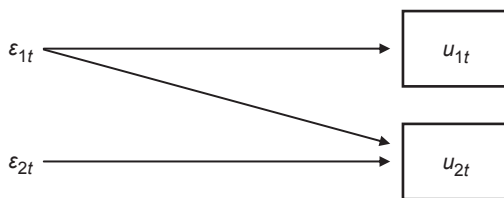


Рис. 6.12

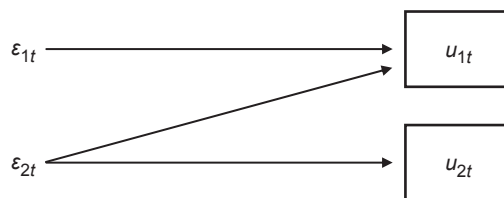


Рис. 6.13

банком, а объем производства становится известным центральному банку лишь с опозданием. Поэтому деньги не могут немедленно реагировать на шоки в объеме производства.

Наряду с функциями импульсного отклика важным инструментом анализа в методологии VAR является также **декомпозиция (разложение) дисперсий ошибок прогнозов** (*variance decomposition*) на h шагов вперед, $h = 1, 2, \dots$. Базой для построения такой декомпозиции служит указанное выше *VMA*-представление ряда y_t , основанное на последовательности фундаментальных инноваций ε_t :

$$y_t = D\varepsilon_t + \Psi_1 D\varepsilon_{t-1} + \Psi_2 D\varepsilon_{t-2} + \Psi_3 D\varepsilon_{t-3} + \dots$$

Рассмотрим для простоты прогноз на один шаг вперед. Для момента $(t+1)$ имеем:

$$y_{t+1} = D\varepsilon_{t+1} + \Psi_1 D\varepsilon_t + \Psi_2 D\varepsilon_{t-1} + \Psi_3 D\varepsilon_{t-2} + \dots$$

Прогноз значения y_{t+1} , сделанный в момент t , имеет вид:

$$\hat{y}_{t+1|t} = E(y_{t+1} | \varepsilon_t, \varepsilon_{t-1}, \dots) = \Psi_1 D\varepsilon_t + \Psi_2 D\varepsilon_{t-1} + \Psi_3 D\varepsilon_{t-2} + \dots,$$

ошибка такого прогноза равна:

$$e_{t+1|t} = y_{t+1} - \hat{y}_{t+1|t} = D\varepsilon_{t+1}.$$

Рассмотрим матрицу

$$e_{t+1|t} e_{t+1|t}^T = D\varepsilon_{t+1} (D\varepsilon_{t+1})^T = D\varepsilon_{t+1} \varepsilon_{t+1}^T D^T.$$

Математическое ожидание этой матрицы есть матрица, составленная математическими ожиданиями ее элементов, так что

$$\begin{aligned} E(e_{t+1|t} e_{t+1|t}^T) &= E(D \varepsilon_{t+1} \varepsilon_{t+1}^T D^T) = DE(\varepsilon_{t+1} \varepsilon_{t+1}^T)D^T = \\ &= DCov(\varepsilon_{t+1})D^T = DD^T. \end{aligned}$$

На диагонали этой матрицы находятся значения дисперсий ошибок прогнозов на один шаг вперед для рядов y_1 и y_2 .

Матрицу DD^T можно представить в виде:

$$DD^T = d_1 d_1^T + d_2 d_2^T,$$

где d_j — j -й столбец матрицы D .

При этом

$$d_1 = \begin{pmatrix} d_{11} \\ d_{21} \end{pmatrix}, \quad d_1^T = (d_{11} \quad d_{21}), \quad d_2 = \begin{pmatrix} d_{12} \\ d_{22} \end{pmatrix}, \quad d_2^T = (d_{12} \quad d_{22}),$$

так что дисперсия ошибки прогноза на один шаг вперед ряда y_1 равна:

$$D(e_{1,t+1|t}) = d_{11}^2 + d_{12}^2,$$

а дисперсия ошибки прогноза на один шаг вперед ряда y_2 равна:

$$D(e_{2,t+1|t}) = d_{21}^2 + d_{22}^2.$$

Таким образом, получили декомпозиции дисперсий ошибок прогноза — разложение каждой из них на две компоненты. Что представляют собой эти компоненты? Заметим, что

$$D \varepsilon_{t+1} = D \begin{pmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{pmatrix} = \begin{pmatrix} d_{11}\varepsilon_{1,t+1} + d_{12}\varepsilon_{2,t+1} \\ d_{21}\varepsilon_{1,t+1} + d_{22}\varepsilon_{2,t+1} \end{pmatrix}.$$

Отсюда вытекает следующее:

- $d_{11}\varepsilon_{1,t+1}$ — составляющая ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$;
- $d_{12}\varepsilon_{2,t+1}$ — составляющая ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$;
- $D(d_{11}\varepsilon_{1,t+1} + d_{12}\varepsilon_{2,t+1}) = d_{11}^2 + d_{12}^2$ — дисперсия ошибки прогноза на один шаг вперед ряда y_1 ;
- $D(d_{11}\varepsilon_{1,t+1}) = d_{11}^2$ — составляющая дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$;
- $D(d_{12}\varepsilon_{2,t+1}) = d_{12}^2$ — составляющая дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$;

- $\frac{d_{11}^2}{d_{11}^2 + d_{12}^2}$ — доля дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$;
- $\frac{d_{12}^2}{d_{11}^2 + d_{12}^2}$ — доля дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$;

Аналогично получаем:

- $D(d_{21}\varepsilon_{1,t+1} + d_{22}\varepsilon_{2,t+1}) = d_{21}^2 + d_{22}^2$ — дисперсия ошибки прогноза на один шаг вперед ряда y_2 ;
- $D(d_{21}\varepsilon_{1,t+1}) = d_{21}^2$ — составляющая дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$;
- $D(d_{22}\varepsilon_{2,t+1}) = d_{22}^2$ — составляющая дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$;
- $\frac{d_{21}^2}{d_{21}^2 + d_{22}^2}$ — доля дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$;
- $\frac{d_{22}^2}{d_{21}^2 + d_{22}^2}$ — доля дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$.

Вернемся к рассмотренной ранее двумерной VAR(1):

$$\begin{aligned} y_{1t} &= \frac{5}{8}y_{1,t-1} + \frac{1}{2}y_{2,t-1} + u_{1t}, \\ y_{2t} &= \frac{1}{4}y_{1,t-1} + \frac{5}{8}y_{2,t-1} + u_{2t}. \end{aligned}$$

При упорядочении $Y1 \rightarrow Y2$

$$D = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix},$$

так что

- доля дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$, равна $\frac{d_{11}^2}{d_{11}^2 + d_{12}^2} = \frac{1}{1} = 1$;
- доля дисперсии ошибки прогноза на один шаг вперед ряда y_1 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$, равна $\frac{d_{12}^2}{d_{11}^2 + d_{12}^2} = \frac{0}{1} = 0$;

- доля дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{1,t+1}$, равна $\frac{d_{21}^2}{d_{21}^2 + d_{22}^2} = \frac{1}{2}$;
- доля дисперсии ошибки прогноза на один шаг вперед ряда y_2 , обусловленная фундаментальной инновацией $\varepsilon_{2,t+1}$, равна $\frac{d_{22}^2}{d_{21}^2 + d_{22}^2} = \frac{1}{2}$.

В общем случае ошибка прогноза по VAR на h шагов вперед вычисляется по формуле:

$$\begin{aligned} y_{t+h} - \hat{y}_{t+h|t} &= u_{t+h} + \Psi_1 u_{t+h-1} + \Psi_2 u_{t+h-2} + \dots + \Psi_{h-1} u_{t+1} = \\ &= D\varepsilon_{t+h} + \Psi_1 D\varepsilon_{t+h-1} + \Psi_2 D\varepsilon_{t+h-2} + \dots + \Psi_{h-1} D\varepsilon_{t+1}, \end{aligned}$$

так что

$$\begin{aligned} E(y_{t+h} - \hat{y}_{t+h|t})(y_{t+h} - \hat{y}_{t+h|t})^T &= DD^T + \Psi_1 DD^T \Psi_1^T + \dots + \Psi_{h-1} DD^T \Psi_{h-1}^T = \\ &= \sum_{j=1}^k (d_j d_j^T + \Psi_1 d_j d_j^T \Psi_1^T + \dots + \Psi_{h-1} d_j d_j^T \Psi_{h-1}^T), \end{aligned}$$

где d_j — j -й столбец матрицы D .

При этом вклад j -й фундаментальной инновации в дисперсию прогноза ряда $y_{l,t}$ равен l -му диагональному элементу матрицы, заключенной в скобки под знаком суммы.

Обычно результат декомпозиции дисперсий прогнозов представляется как перечень долей каждого из слагаемых в общей сумме (в процентах). В пакетах программ статистического анализа строятся графики, показывающие динамику изменений каждой такой доли с изменением h , $h = 1, 2, \dots$

✓ **Замечание 6.2.3.** На практике для построения декомпозиции дисперсии прогнозов, как и для построения функции импульсного отклика, приходится использовать не теоретическую, а оцененную по имеющимся статистическим данным модель VAR. Поэтому, даже при использовании сгенерированных данных, следующих заданному процессу порождения данных (в виде приведенной формы VAR с заданными коэффициентами), декомпозиция дисперсии ошибок прогнозов, построенная по имеющимся данным, отличается от теоретической.

Ниже приведены таблицы и графики декомпозиций дисперсий ошибок прогнозов для нашего примера с двумерной VAR(1). Первому упорядочению соответствуют табл. 6.13 и рис. 6.14, второму упорядочению — табл. 6.14 и рис. 6.15.

Variance Decomposition

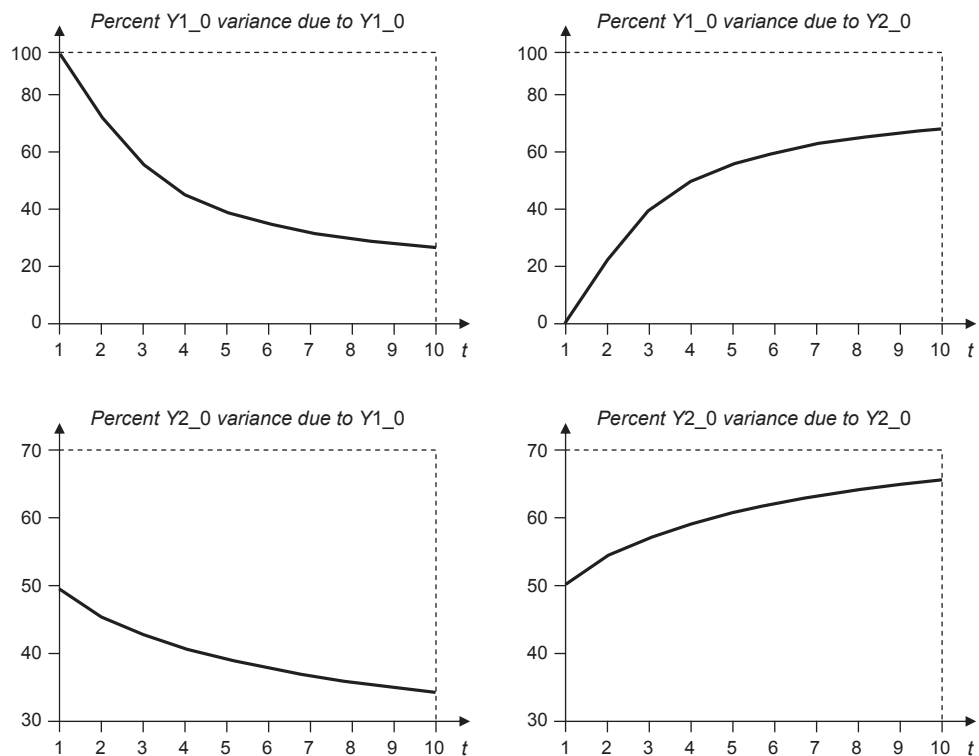


Рис. 6.14

Таблица 6.13

Декомпозиции дисперсий ошибок прогнозов (первое упорядочение)

Variance Decomposition of Y1_0:				Variance Decomposition of Y2_0:			
Period	S.E.	Y1_0	Y2_0	Period	S.E.	Y1_0	Y2_0
1	0.959222	100.0000	0.000000	1	1.401312	49.71512	50.28488
2	1.099896	76.54276	23.45724	2	1.527966	45.55110	54.44890
3	1.268521	58.29527	41.70473	3	1.589649	42.85573	57.14427
4	1.417664	47.71032	52.28968	4	1.636419	40.82702	59.17298
5	1.542655	41.18427	58.81573	5	1.676155	39.20087	60.79913
6	1.647747	36.82258	63.17742	6	1.710909	37.86380	62.13620
7	1.737110	33.72222	66.27778	7	1.741576	36.74852	63.25148
8	1.813880	31.41700	68.58300	8	1.768740	35.80842	64.19158
9	1.880371	29.64445	70.35555	9	1.792859	35.00920	64.99080
10	1.938326	28.24592	71.75408	10	1.814315	34.32484	65.67516

Variance Decomposition

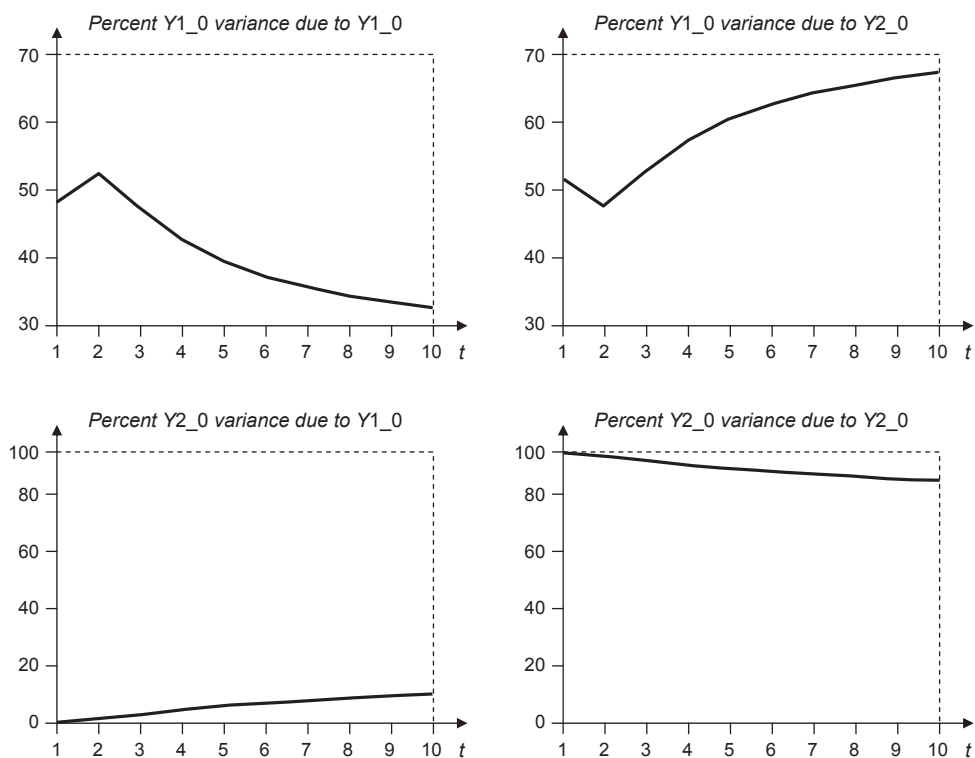


Рис. 6.15

Таблица 6.14

Декомпозиции дисперсий ошибок прогнозов (второе упорядочение)

Variance Decomposition of Y1_0:				Variance Decomposition of Y2_0:			
Period	S.E.	Y1_0	Y2_0	Period	S.E.	Y1_0	Y2_0
1	0.959222	50.28488	49.71512	1	1.401312	0.000000	100.0000
2	1.099896	53.52920	46.47080	2	1.527966	1.182727	98.81727
3	1.268521	48.33907	51.66093	3	1.589649	2.584105	97.41590
4	1.417664	44.19547	55.80453	4	1.636419	3.818615	96.18138
5	1.542655	41.44300	58.55700	5	1.676155	4.847692	95.15231
6	1.647747	39.56726	60.43274	6	1.710909	5.701919	94.29808
7	1.737110	38.22719	61.77281	7	1.741576	6.416083	93.58392
8	1.813880	37.22952	62.77048	8	1.768740	7.018400	92.98160
9	1.880371	36.46213	63.53787	9	1.792859	7.530525	92.46947
10	1.938326	35.85661	64.14339	10	1.814315	7.969058	92.03094

Методология VAR довольно быстро стала доминирующей в эмпирической монетарной экономике. Она дала не только возможность описывать широкий спектр реальных данных, но и базу для анализа альтернативных теорий и гипотез. Модели VAR не представляют истину в последней инстанции, но являются полезным инструментом для прояснения взаимодействий между различными переменными.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Каковы основные черты методологии комиссии Коулса?
2. На какие недостатки методологии комиссии Коулса указывали Лукас и Симс?
3. Каковы основные черты методологии Лондонской школы экономики?
4. Что лежит в основе методологии VAR?
5. Что понимается под фундаментальными инновациями?
6. Что представляют собой функции импульсного отклика?
7. В чем состоит проблема идентификации структурной формы VAR на основе приведенной формы? Как решается эта проблема?
8. Как строится рекурсивная система на основании приведенной формы VAR? Что понимается под наименее эндогенной и под наиболее эндогенной переменными в полученной рекурсивной системе? Можно ли оценивать полученную рекурсивную систему обычным методом наименьших квадратов?
9. Влияет ли порядок вхождения переменных на поведение функций импульсного отклика?
10. Что представляет собой декомпозиция дисперсий ошибок прогнозов?

Тема 6.3

ЭМПИРИЧЕСКИЕ ИССЛЕДОВАНИЯ

В исследованиях по эмпирической монетарной экономике рассматривались модели VAR с различным количеством переменных.

В работе (*Sims, 1980*) была применена методология VAR для сравнения динамики экономики США в период между двумя мировыми войнами (*IW*) и в период после Второй мировой войны (*AW*). Использовались месячные данные. Первоначально Симс рассмотрел VAR с тремя переменными:

$M1$ — денежный агрегат $M1$ (сезонно скорректированная);

IP — индекс промышленного производства (сезонно скорректированная);

P — индекс оптовых цен (сезонно скорректированная).

Заметим, что в VAR с 3 переменными схема влияния фундаментальных инноваций с упорядочением $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t}$ имеет вид, показанный на рис. 6.16. В качестве y_{1t} в этом упорядочении выступала переменная $M1$ (наименее эндогенная), в качестве y_{2t} — переменная IP , а в качестве y_{3t} — переменная P (наиболее эндогенная): Денежная масса \rightarrow Объем производства \rightarrow Цены.

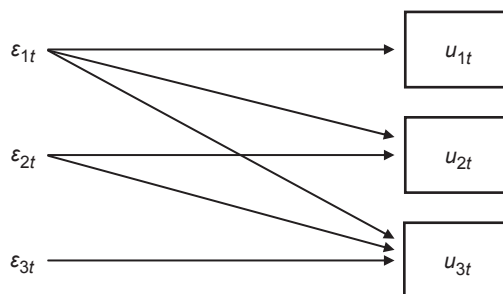


Рис. 6.16

Симс привел декомпозиции дисперсий ошибок прогнозов этих 3 переменных на 48 месяцев вперед для периода *IW* (между двумя мировыми войнами (1920—1941)) и в период *AW* (после Второй мировой войны (1948—1978)). Все переменные брались в логарифмах, в правые части уравнений включались константы и по 12 лагов каждой переменной.

Таблица 6.15

Декомпозиции дисперсий ошибок прогнозов 3 переменных (по Симсу)

Объясняемые переменные	Объясняются шоками в переменных					
	<i>M1</i>		<i>IP</i>		<i>P</i>	
	<i>IW</i>	<i>AW</i>	<i>IW</i>	<i>AW</i>	<i>IW</i>	<i>AW</i>
<i>M1</i>	92	97	4	2	4	1
<i>IP</i>	66	37	28	44	6	18
<i>P</i>	38	14	19	7	43	80

Для обоих периодов дисперсия *M1* почти полностью объясняется инновациями *M1* (табл. 6.15). Инновации *M1* объясняют значительную часть дисперсии *IP* и существенную часть дисперсии *P*. Все это вполне в духе монетаризма, если понимать монетаризм как точку зрения, согласно которой денежная политика играет центральную роль в цикле деловой активности, а динамика денежной массы является хорошим показателем денежной политики.

Между тем ранее в работе (Mehra, 1978) было обнаружено, что приоритет денежной массы как причины по Грейнджеру исчезает, если в систему включается процентная ставка. Для проверки такой возможности Симс добавляет к указанным выше 3 переменным переменную *R* — процентную ставку по 4—6-месячным первоклассным коммерческим бумагам (*prime commercial papers*) и оценивает VAR с 4 переменными. Он приводит декомпозиции дисперсий ошибок прогнозов этих 4 переменных на 48 месяцев вперед, полученные при использовании упорядочения $R \rightarrow M1 \rightarrow P \rightarrow IP$ (табл. 6.16).

Таблица 6.16

Декомпозиции дисперсий ошибок прогнозов 4 переменных (по Симсу)

Объясняемые переменные	Объясняются шоками в переменных							
	<i>R</i>		<i>M1</i>		<i>P</i>		<i>IP</i>	
	<i>IW</i>	<i>AW</i>	<i>IW</i>	<i>AW</i>	<i>IW</i>	<i>AW</i>	<i>IW</i>	<i>AW</i>
<i>R</i>	63	50	28	19	7	4	1	28
<i>M1</i>	39	56	58	42	1	1	1	1
<i>P</i>	1	2	54	32	43	60	3	6
<i>IP</i>	16	30	58	4	7	14	18	52

Как видим, переменная *M1* полностью потеряла в послевоенный период предсказательную роль в отношении *IP*. В то же время дисперсия самой *M1* теперь в значительной мере определяется инновацией процентной ставки.

В работе (Leeper, Sims, Zha, 1996) были рассмотрены модели VAR с 3, 4 и 5 переменными. В простейшей из этих моделей используются 3 переменные: цены, денежная масса и объем производства. Оценивание модели производилось на месячных данных на периоде 1960:1—1996:03 (всего 435 наблюдений). В качестве переменных использовались:

$$LP = 100\ln(CPI_SA), \quad LY = 100\ln(RGDP_MON), \quad LM2 = 100 \ln(M2_SA).$$

(Оценивалась также и модель, в которой вместо *M2* использовался денежный агрегат *M1*.) В уравнения включались константы и по 6 лагов каждой переменной. Анализ откликов системы с указанными 3 переменными производился при упорядочении $P \rightarrow Y \rightarrow M2$. Функции импульсного отклика для оцененной модели VAR показаны на рис. 6.17.

Можно сказать, что такое упорядочение «дает фору» шоку в переменной *P*. И хотя это упорядочение исходит из того, что наиболее эндогенной переменной являются деньги, из приведенных графиков (два первых графика в нижнем ряду) этого не видно. Этот вывод подтверждается поведением декомпозиций дисперсий прогнозов (рис. 6.18).

Цены и объем производства играют весьма ограниченную роль в объяснении дисперсий прогнозов денег, так что эндогенность денежного агрегата не выявляется явным образом. В то же время деньги играют более заметную роль в объяснении прогнозов этих двух макроэкономических переменных.

Между тем возникает желание посмотреть, что будет, если все же применить упорядочение, использованное Симсом. В упорядочении Симса $M2 \rightarrow Y \rightarrow P$ получают декомпозиции дисперсий, представленные на рис. 6.19.

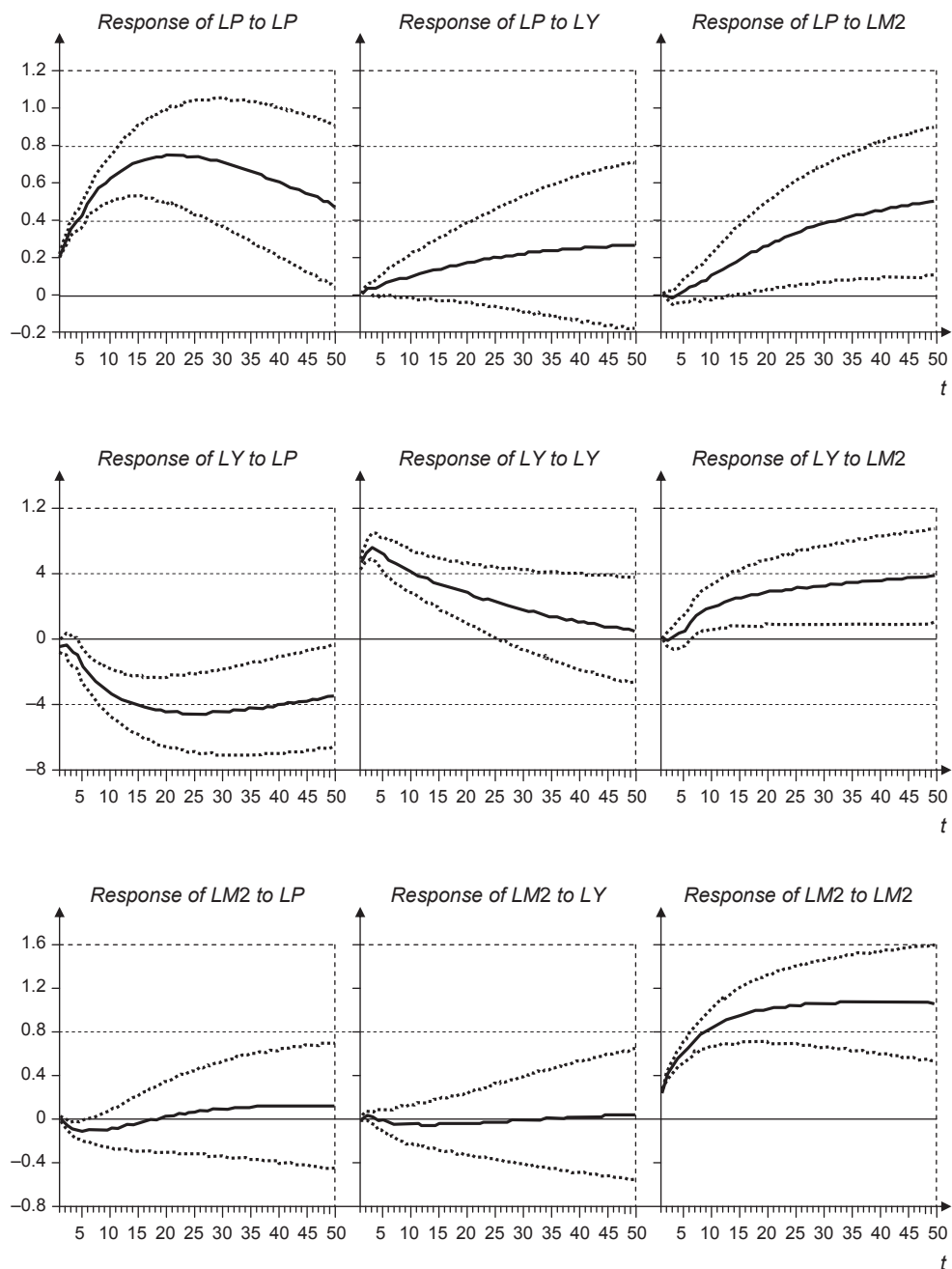
Response to Cholesky One S.D. Innovations ± 2 S.E.

Рис. 6.17

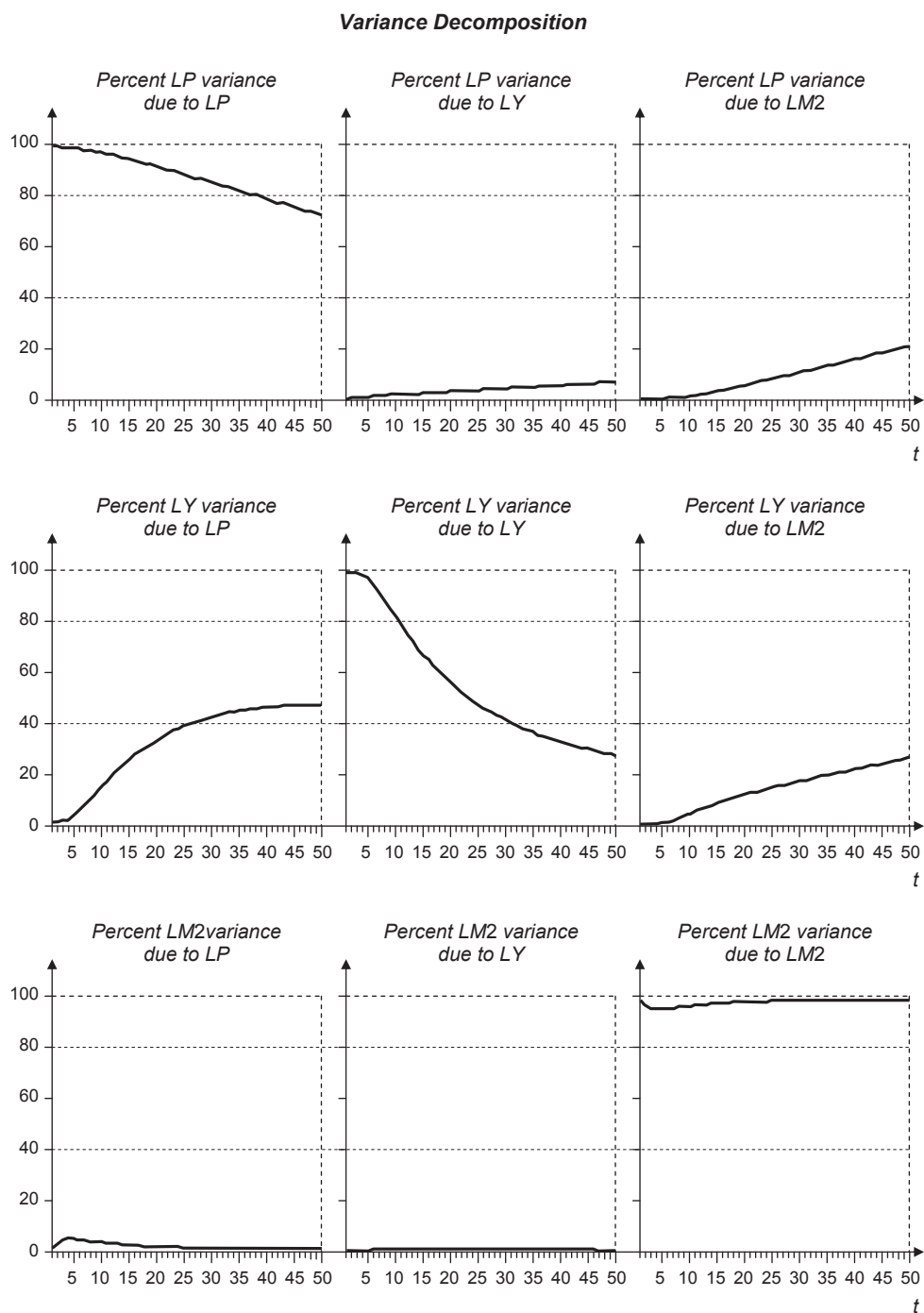


Рис. 6.18

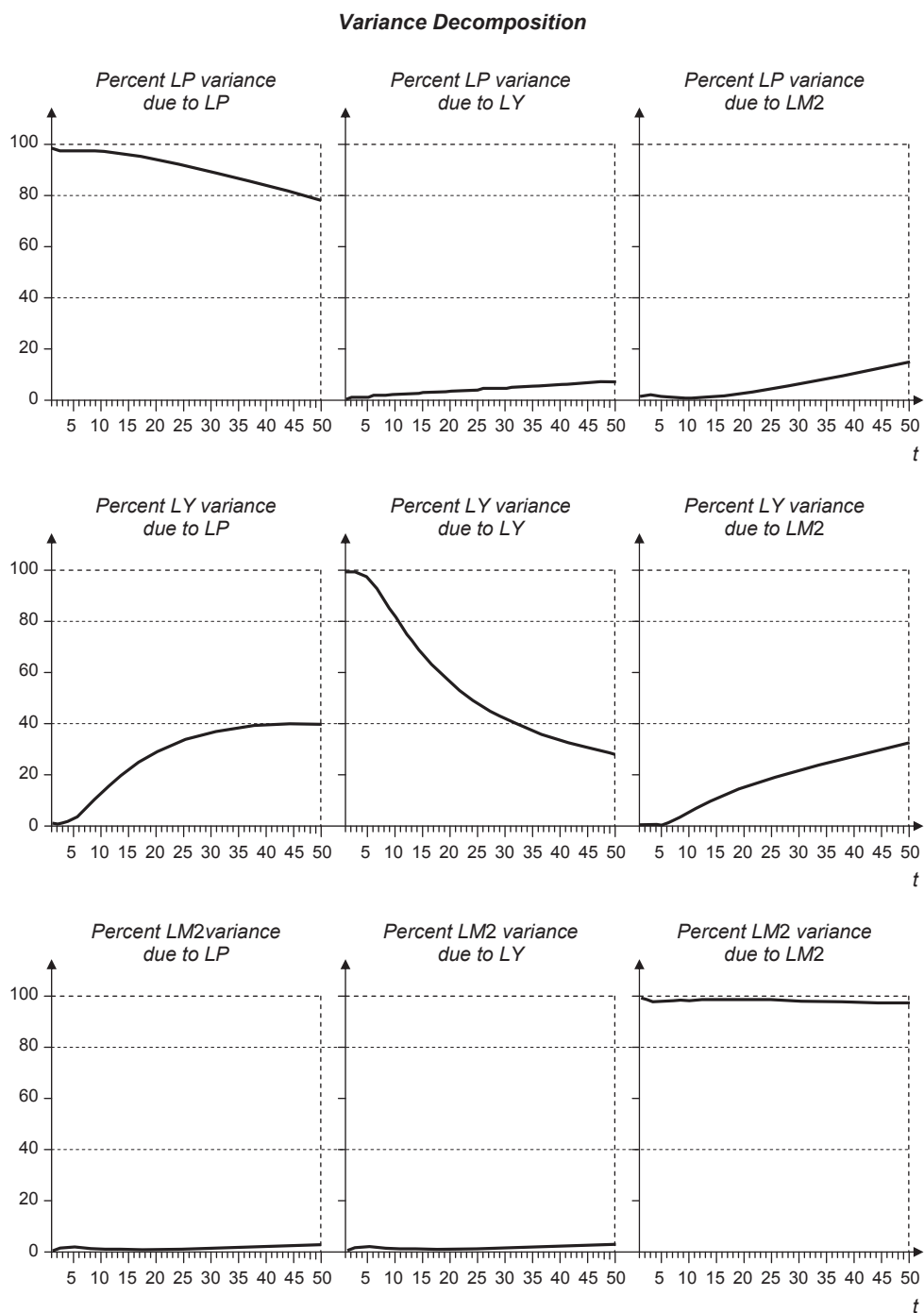


Рис. 6.19

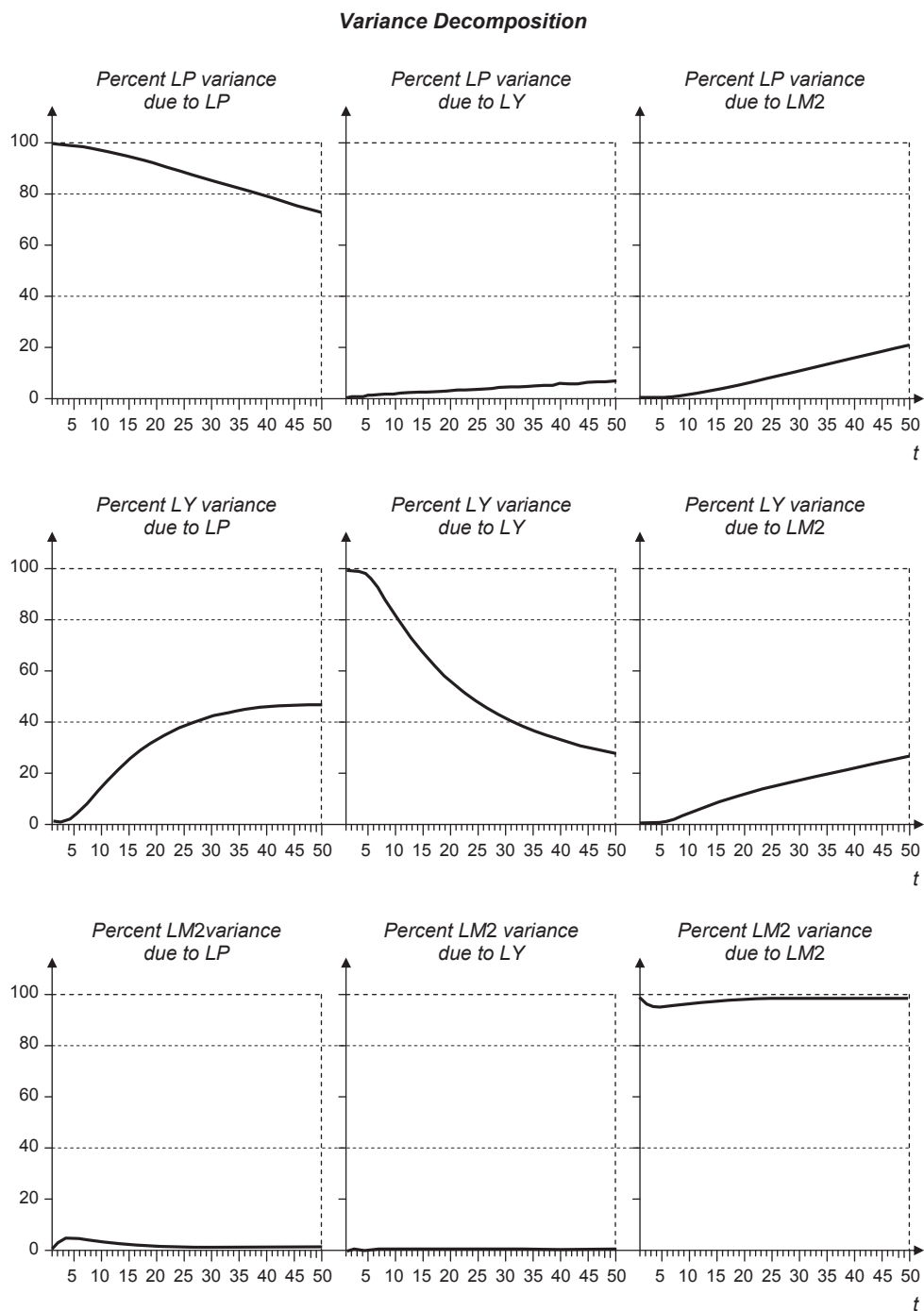


Рис. 6.20

При использовании упорядочения (Leeper, Sims, Zha, 1996) получаем результаты, изображенные на рис. 6.20. Графики очень похожи (хотя и не совпадают).

Сравним теперь функции откликов. Для наглядности в обоих случаях графики принудительно выведены в одном и том же порядке. Рисунок 6.21 соответствует упорядочению $M2 \rightarrow Y \rightarrow P$, рис. 6.22 — упорядочению $P \rightarrow Y \rightarrow M2$.

Почему все так похоже? Получается, что порядок вхождения переменных для этих данных не имеет значения. Но это возможно только в том случае, если инновации в приведенной VAR не коррелированы между собой. На сей раз корреляции между остатками имеют значения, приведенные в табл. 6.17.

Таблица 6.17

Корреляции инноваций между остатками

<i>LP</i>	<i>LY</i>	<i>LM2</i>
1.000000	-0.086962	-0.094147
-0.086962	1.000000	-0.002022
-0.094147	-0.002022	1.000000

Во всяком случае, в обоих вариантах:

- эндогенность $M2$ незаметна;
- $M2$ играет более заметную роль в объяснении дисперсий прогнозов реального GDP .

В работе 1980 г. Симс дополнил модель с 3 переменными (цены, объем производства, деньги) четвертой переменной, которую он поместил в упорядочении непосредственно перед деньгами. Это явилось отражением движения центральных банков к использованию процентных ставок в качестве инструмента монетарной политики. Посмотрим, как введение в модель такой переменной (в работе Leeper, Sims, Zha это переменная $TBILL3$ — процентные ставки по 3-месячным казначейским обязательствам) отразится на получаемых выводах.

Функции импульсного отклика при упорядочении $TBILL3 \rightarrow M2 \rightarrow P \rightarrow Y$ изображены на рис. 6.23, а на рис. 6.24 — декомпозиции дисперсий ошибок прогнозов.

Здесь обнаруживается предположенная эндогенность $M2$ по отношению к $TBILL3$ (третий график в последней строке), но опять значимо не выявляется эндогенность $M2$ по отношению к CPI и GDP . Инновации в процентных ставках имеют большую предсказательную силу для GDP , чем инновации в $M2$.

При анализе графиков обнаруживаются также две «загадки»: «загадка ликвидности» (*liquidity puzzle*) и «загадка цен» (*price puzzle*).

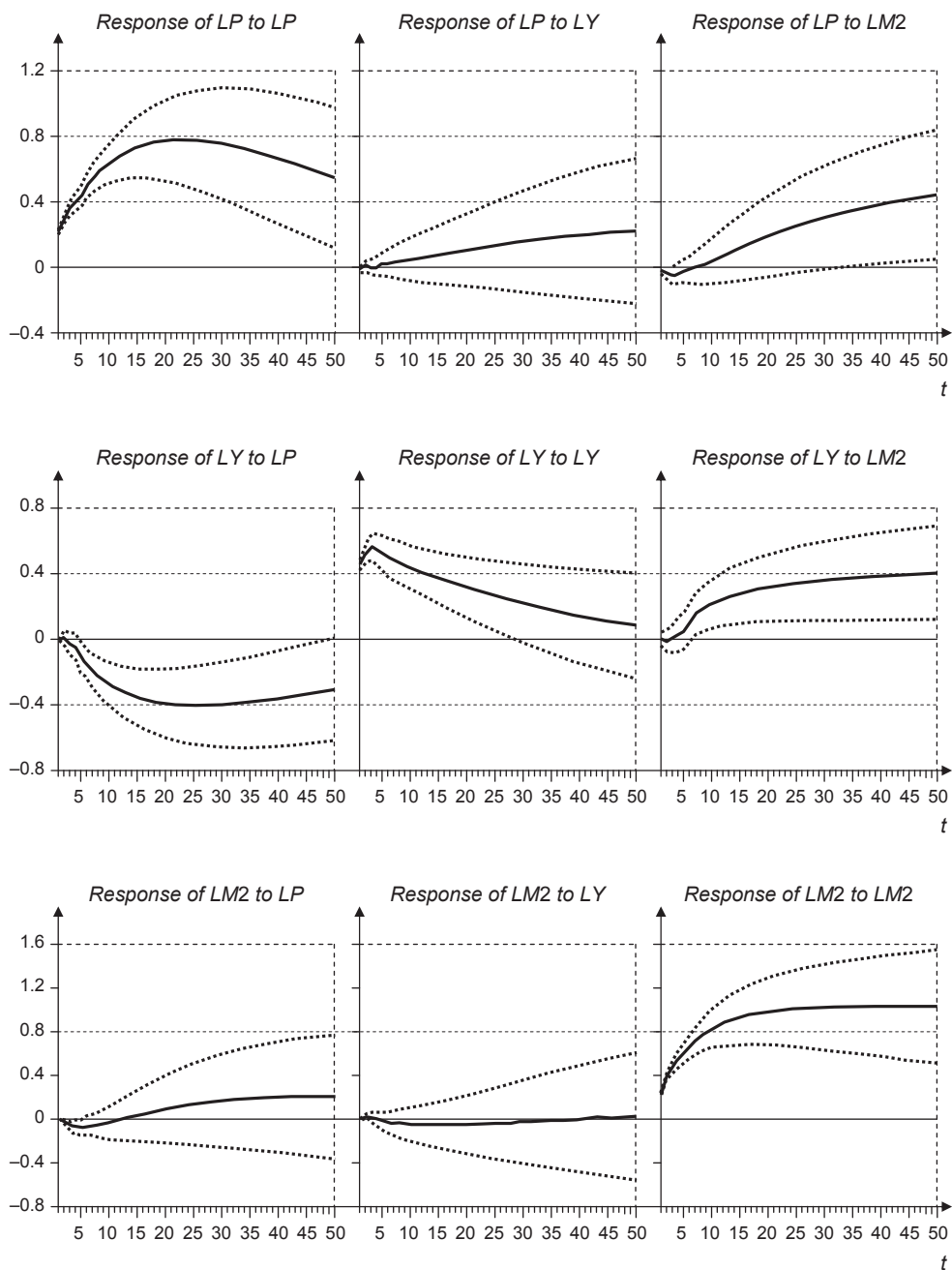
Response to Cholesky One S.D. Innovations ± 2 S.E.

Рис. 6.21

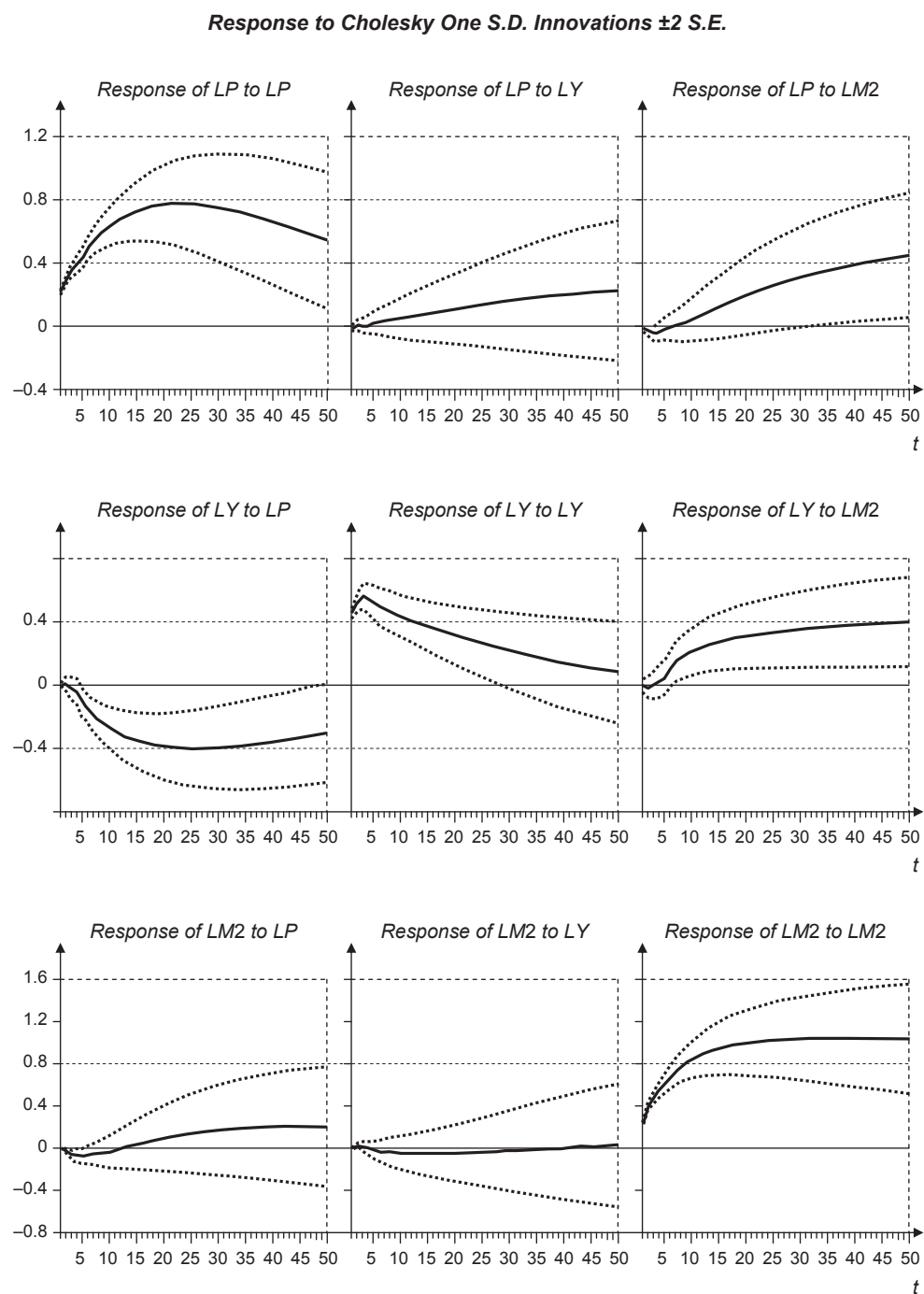


Рис. 6.22

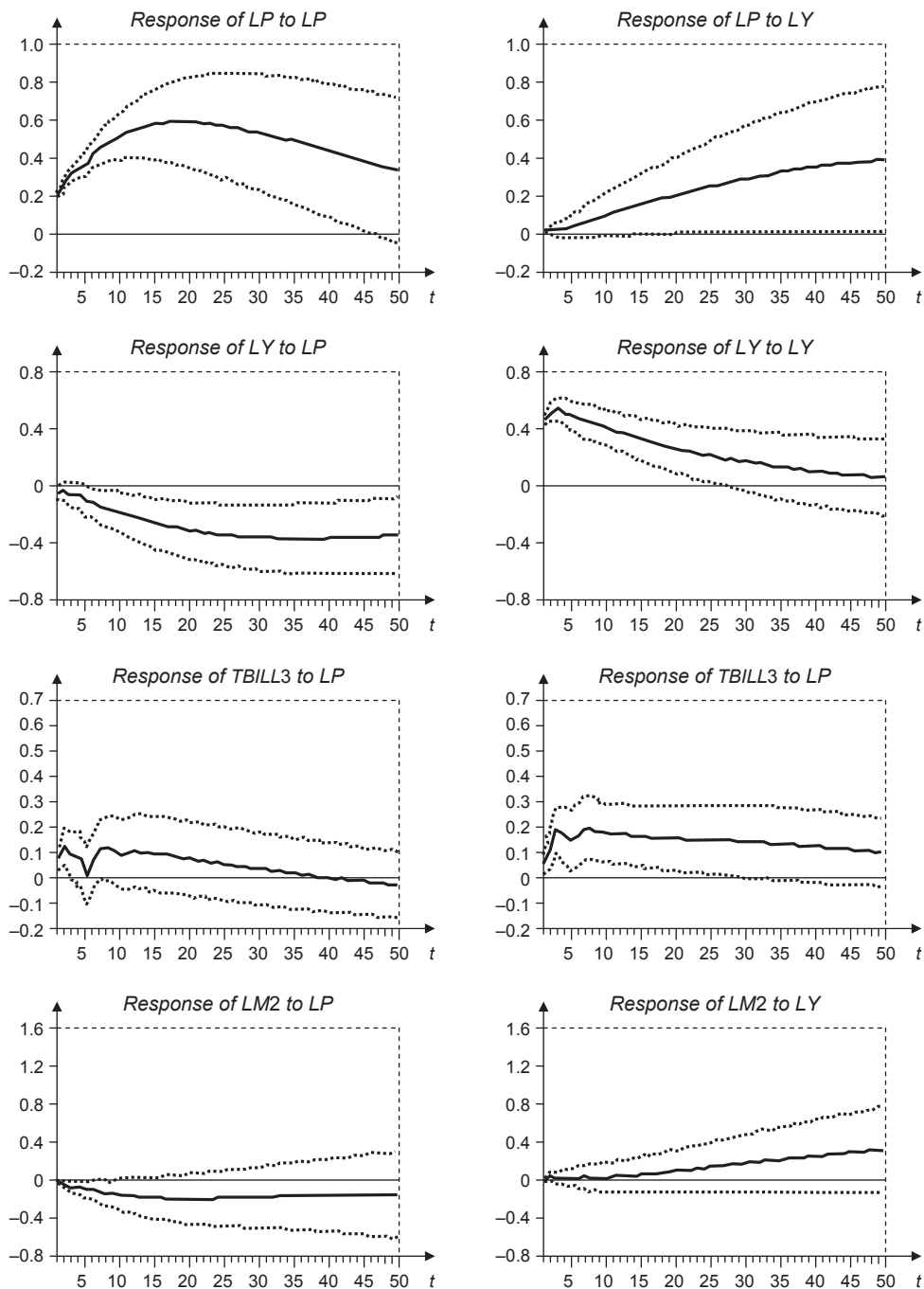
Response to Cholesky One S.D. Innovations ± 2 S.E.

Рис. 6.23 (начало)

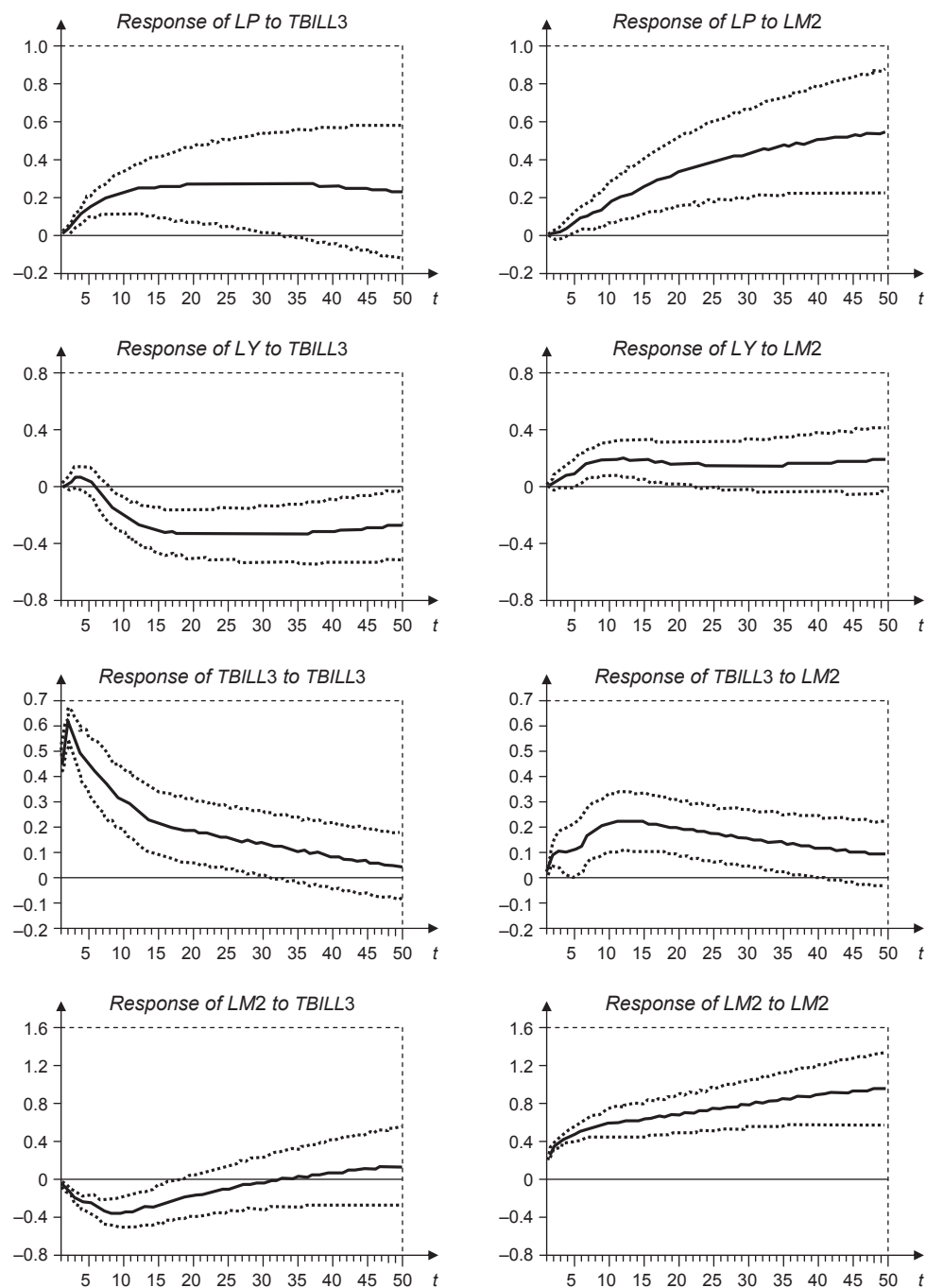


Рис. 6.23 (продолжение)

Variance Decomposition

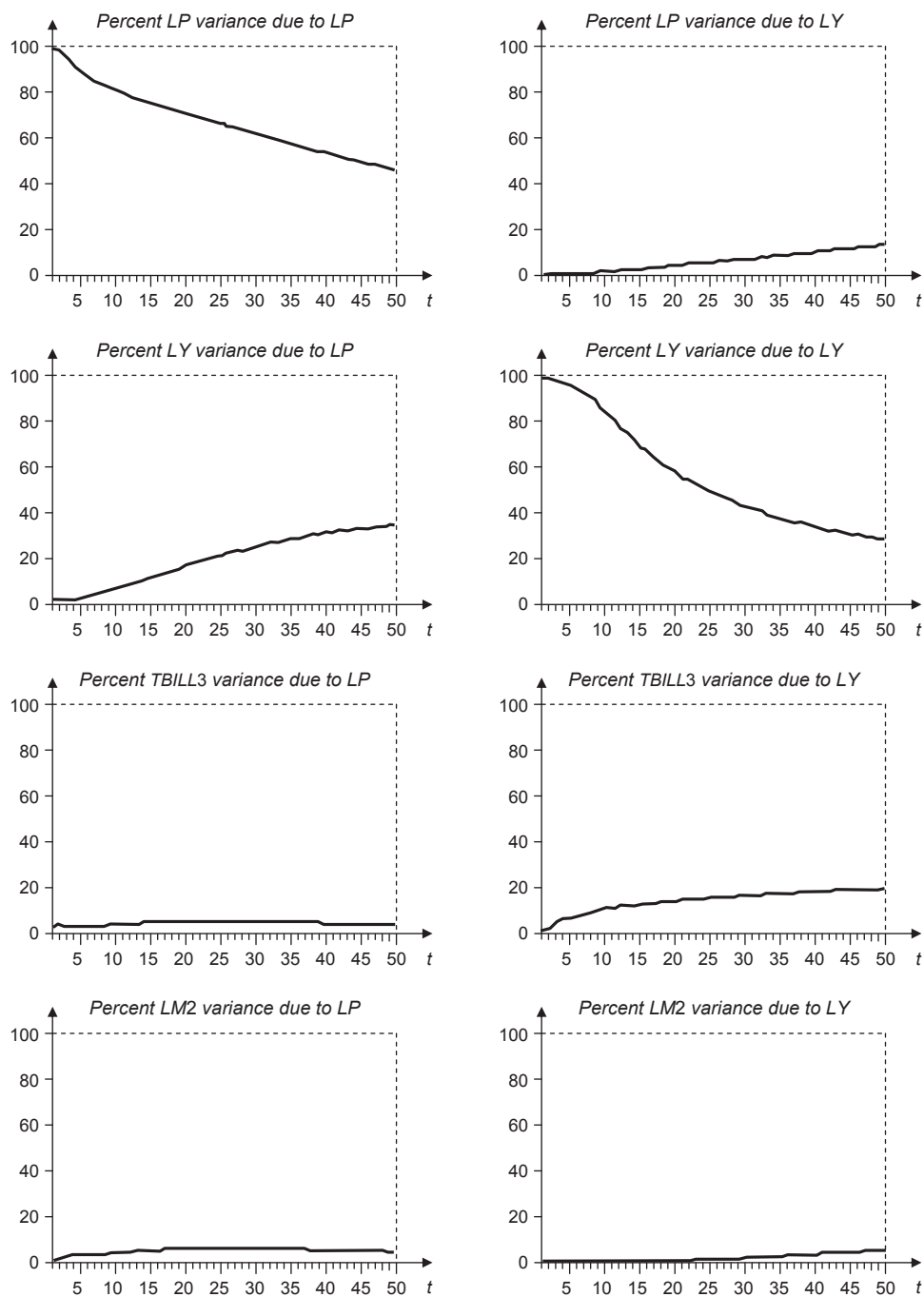


Рис. 6.24 (начало)

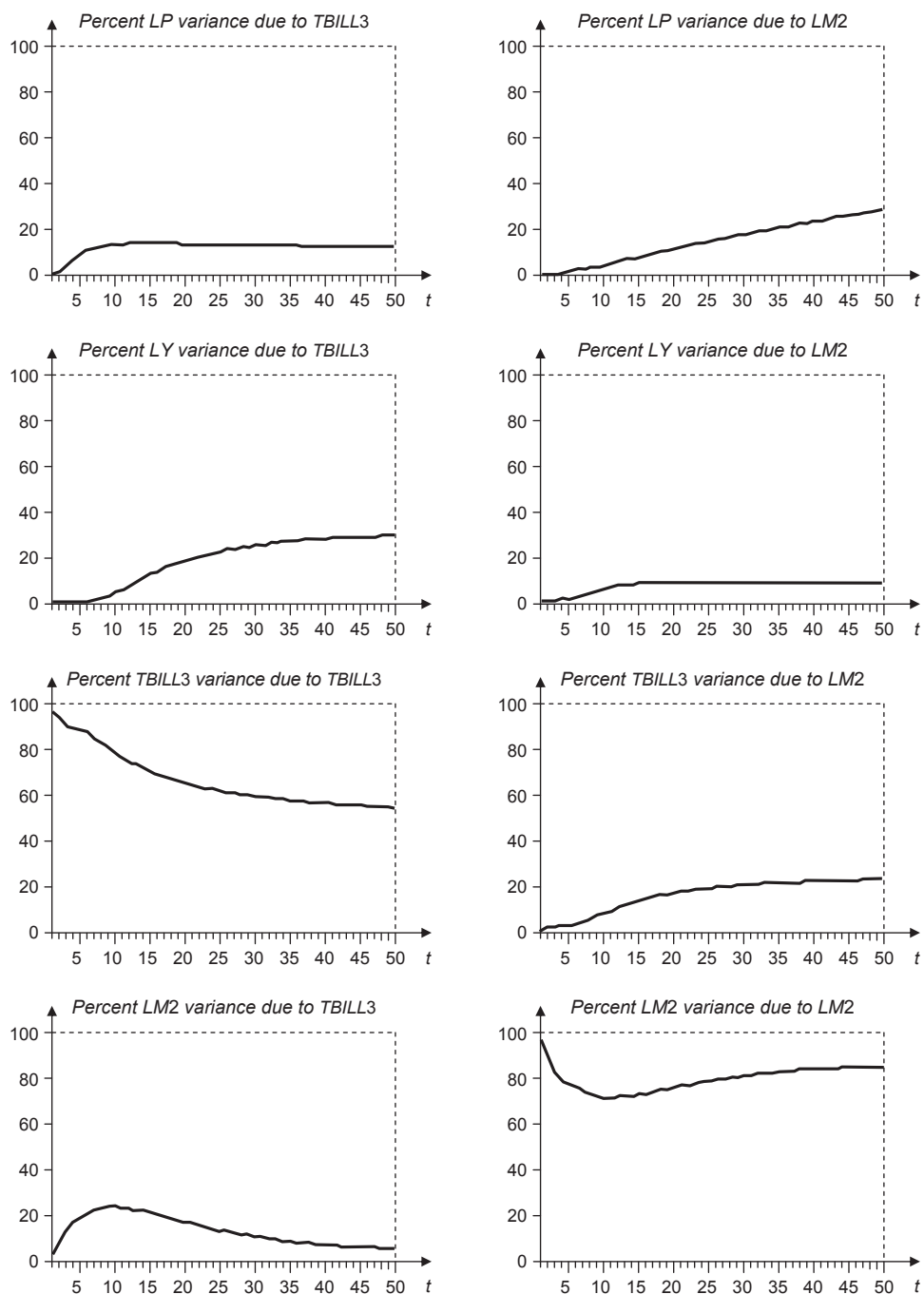


Рис. 6.24 (продолжение)

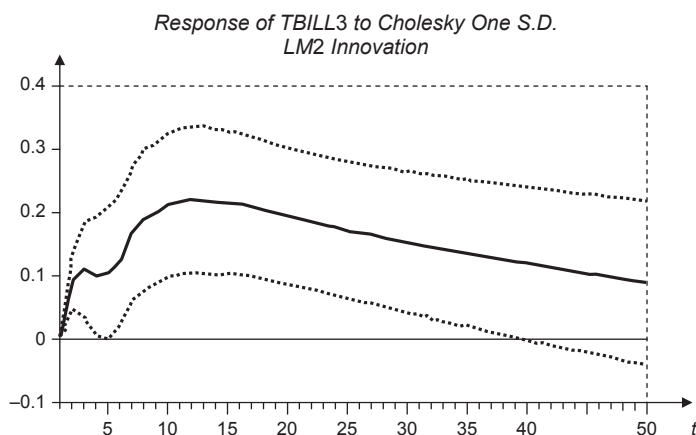


Рис. 6.25

«Загадка ликвидности» связана с наблюдаемым поведением процентных ставок в ответ на шок в $M2$, а именно некоторым их убыванием на начальном отрезке графика (рис. 6.25).

«Загадка цен» связана с их возрастанием после повышения процентных ставок (рис. 6.26).

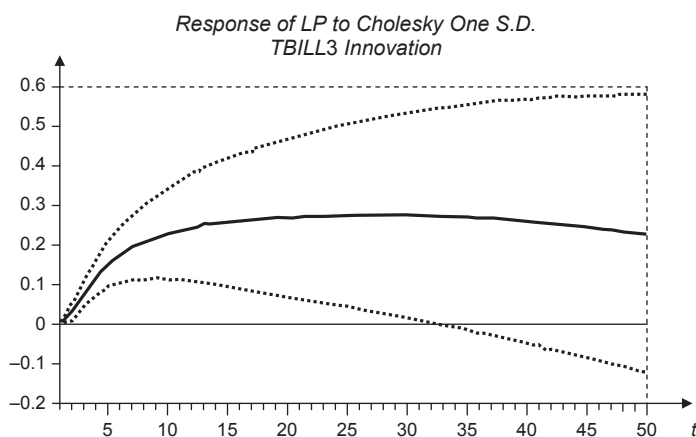


Рис. 6.26

Получается, что повышение процентных ставок, вопреки теоретическим ожиданиям, приводит к инфляции. Такое положение может быть вызвано недостаточным количеством переменных в рассмотренной модели VAR с 4 переменными. В модели может быть пропущена переменная, являющаяся опережающим индикатором инфляции, на которую реагирует ФРС. При этом

положительная корреляция между такой переменной и процентными ставками выразится в положительной корреляции между процентными ставками и ценами. В упоминавшейся выше работе (Sims, 1992) Симс показал важность включения в модель подобных переменных, дополнив модель переменной «индекс цен на товары» (*commodity price index*). В работе (Sims, Zha, 1998) в качестве такой дополнительной переменной был взят индекс цен производителей на промежуточные продукты (*producers' price index for intermediate goods*).

Продолжим наше исследование на базе данных, использовавшихся в работе (Leeper, Sims, Zha, 1996), дополнив только что рассмотренную модель с 4 переменными переменной *PCM* (индекс Международного валютного фонда или *commodity price level*). В работе (Sims, 1992) эта переменная помещается в упорядочении $TBILL3 \rightarrow M2 \rightarrow P \rightarrow Y$ между переменными *TBILL3* и *M2*, что дает в итоге упорядочение

$$TBILL3 \rightarrow PCM \rightarrow M2 \rightarrow P \rightarrow Y.$$

При таком упорядочении получаем функции импульсного отклика, изображенные на рис. 6.27—6.29 (оценивание этой VAR проводилось на периоде 1983:01—1996:03).

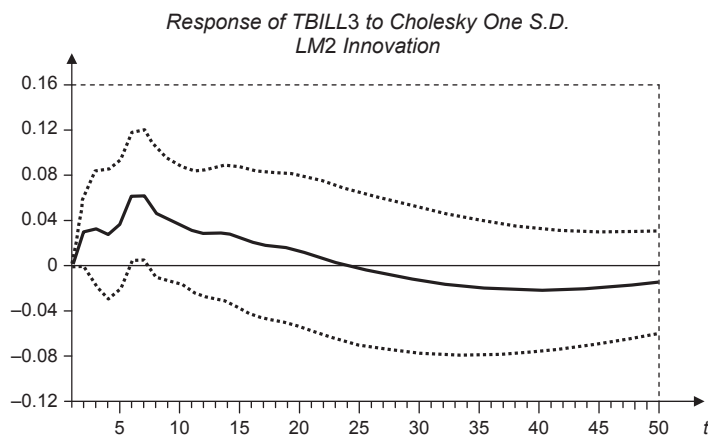


Рис. 6.27

Если изменить порядок вхождения переменной *PCM*, ставя ее в упорядочении на первое место ($PCM \rightarrow TBILL3 \rightarrow M2 \rightarrow P \rightarrow Y$), то получим практически идентичную картину. Один из графиков изображен на рис. 6.30.

Сравним таблицы декомпозиций дисперсий ошибок прогнозов для цен, полученные для моделей с 4 (табл. 6.18) и с 5 переменными (табл. 6.19). Влияние на дисперсию ошибки прогноза цен инноваций *TBILL3* значительно ослабло в модели с 5 переменными вследствие влияния инноваций *PCM*.

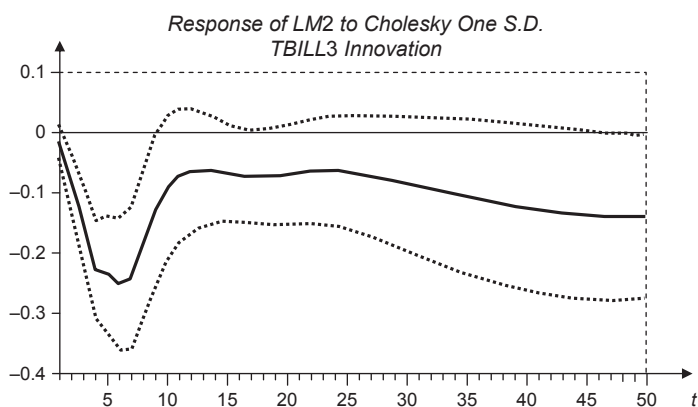


Рис. 6.28

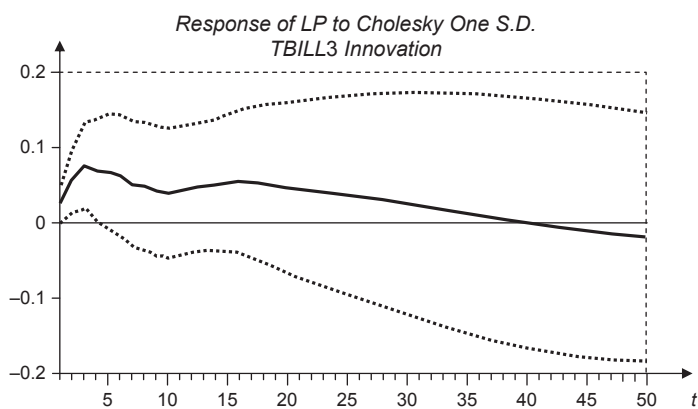


Рис. 6.29

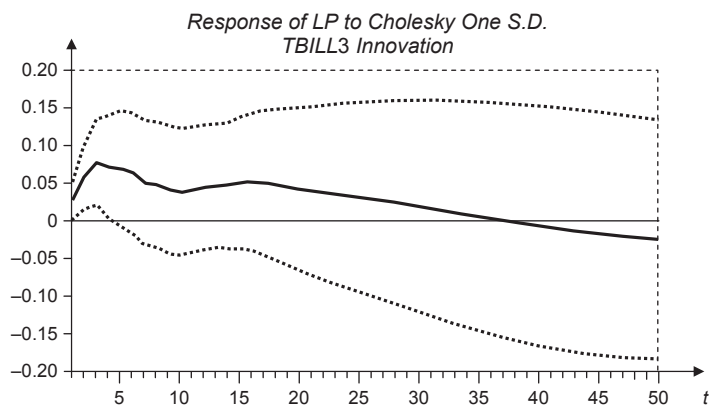


Рис. 6.30

Таблица 6.18

Декомпозиции дисперсий ошибок прогнозов для цен (4 переменные)

<i>Period</i>	<i>S.E.</i>	<i>TBILL3</i>	<i>LM2</i>	<i>LP</i>	<i>LY</i>
1	0.203194	3.413529	0.288594	96.29788	0.000000
2	0.378587	7.398695	0.177139	91.84278	0.581387
3	0.513714	11.734230	0.116070	87.80450	0.345204
4	0.618051	13.843170	0.560451	85.34941	0.246972
5	0.701903	16.271620	1.390146	82.01270	0.325531
6	0.787011	18.798720	2.274997	78.05292	0.873358
7	0.863454	20.380930	3.283058	75.02154	1.314472
8	0.922003	21.481470	4.407218	72.46652	1.644795
9	0.966344	22.256570	5.316719	70.56604	1.860662
10	1.002817	22.867860	5.864264	69.32608	1.941795
<i>Cholesky Ordering: TBILL3 LM2 LP LY</i>					

Таблица 6.19

Декомпозиции дисперсий ошибок прогнозов для цен (5 переменных)

<i>Period</i>	<i>S.E.</i>	<i>LPCM</i>	<i>TBILL3</i>	<i>LM2</i>	<i>LP</i>	<i>LY</i>
1	194.0635	0.589508	2.571722	0.003872	96.83490	0.000000
2	291.7764	1.480334	5.353879	0.079890	92.68270	0.403199
3	376.3085	1.188461	7.979387	0.166852	90.41000	0.255300
4	449.2765	1.089946	8.396279	0.165222	90.04370	0.304855
5	501.9941	1.364401	8.855732	0.297712	89.23682	0.245336
6	559.7134	2.331255	9.184664	0.435117	87.77732	0.271647
7	610.2575	4.147522	8.940558	0.662976	85.99399	0.254951
8	659.1834	6.147062	8.552394	1.059989	84.01543	0.225131
9	706.7642	8.372840	8.041072	1.429213	81.94286	0.214015
10	746.2973	10.53817	7.537850	1.635337	80.01996	0.268676
<i>Cholesky Ordering: LPCM TBILL3 LM2 LP LY</i>						

В работе (Beranke, Blinder, 1992) было выбрано несколько переменных, которые можно рассматривать как инструменты монетарной политики. Это денежные агрегаты *M1*, *M2* и 3 процентные ставки: учетная ставка ФРС США (*FUND*), ставки по 3-месячным казначейским обязательствам (*BILL*) и 10-летним казначейским облигациям (*BOND*). В табл. 6.20 приведены декомпозиции

дисперсий ошибок прогнозов на 24 месяца вперед для 9 макроэкономических показателей (в каждой строке приводится декомпозиция для показателя, указанного в первой графе, в модели VAR, содержащей этот показатель, *CPI*, два денежных агрегата и три процентные ставки). Хотя авторы предполагали выявить существенное влияние переменной *FUND*, они для большей убедительности результатов поставили эту переменную последней в упорядочении, предоставив ей возможность проявиться и в таких невыгодных для нее условиях: $CPI \rightarrow M1 \rightarrow M2 \rightarrow BILL \rightarrow BOND \rightarrow FUNDS$. Результаты для периода 1959:7—1989:12 приведены в табл. 6.20. Использованы месячные данные из *DRI* базы данных; *M1*, *M2*, *industrial production*, *employment*, *housing starts* взяты в логарифмах уровней. *Personal income*, *retail sales*, *consumption* дефлированы и взяты в логарифмах уровней.

Таблица 6.20

Декомпозиции дисперсий ошибок прогнозов (по периоду 1959:7—1989:12)

<i>Forecasted variable</i>	<i>Own Lags</i>	<i>CPI</i>	<i>M1</i>	<i>M2</i>	<i>BILL</i>	<i>BOND</i>	<i>FUNDS</i>
<i>Industrial production</i>	36.6	3.1	15.4	8.7	8.0	0.8	27.4
<i>Capacity utilization</i>	39.7	1.3	21.0	3.5	9.5	1.7	23.3
<i>Employment</i>	38.9	7.0	10.5	0.6	9.8	2.7	30.6
<i>Unemployment rate</i>	31.9	7.2	10.5	0.6	9.9	1.9	37.9
<i>Housing starts</i>	28.8	1.4	3.9	1.8	38.6	14.3	11.2
<i>Personal income</i>	48.2	4.3	20.8	0.1	6.9	3.3	16.3
<i>Retail sales</i>	32.4	15.5	5.1	4.4	27.4	1.1	14.1
<i>Consumption</i>	18.2	13.1	16.0	2.2	28.4	5.3	16.8
<i>Durable goods</i>	41.3	6.8	14.7	5.5	10.3	2.6	18.8

Как видно из табл. 6.20, переменная *FUNDS* вносит наибольший вклад в дисперсию прогнозов для 5 из 9 прогнозируемых макроэкономических показателей.

Для более короткого периода — с 1959:7 по 1979:9 (*pre-Volcker period*) — получены результаты, приведенные в табл. 6.21.

Переменная *FUNDS* и здесь вносит наибольший вклад в дисперсию прогнозов для 5 из 9 прогнозируемых макроэкономических показателей, причем эти 5 показателей те же, что и на полном периоде.

Заметим, что если переменная *FUNDS* ставится в упорядочении первой среди переменных, выражающих денежную политику, то ее вклад в дисперсию прогнозов, как правило, возрастает, но не очень существенно. Это говорит о том, что информация, содержащаяся в этой переменной, почти ортогональна информации, содержащейся в других предикторах.

Таблица 6.21

Декомпозиции дисперсий ошибок прогнозов (по периоду 1959:7—1979:9)

<i>Forecasted variable</i>	<i>Own Lags</i>	<i>CPI</i>	<i>M1</i>	<i>M2</i>	<i>BILL</i>	<i>BOND</i>	<i>FUNDS</i>
<i>Industrial production</i>	36.3	2.7	11.8	6.5	11.5	3.3	27.8
<i>Capacity utilization</i>	39.9	2.4	12.4	4.5	10.8	5.6	24.3
<i>Employment</i>	41.4	1.8	5.8	0.2	10.4	3.2	37.9
<i>Unemployment rate</i>	44.9	1.3	4.9	1.3	11.6	2.2	33.8
<i>Housing starts</i>	45.2	9.9	8.3	6.3	11.8	9.6	9.0
<i>Personal income</i>	34.5	17.7	7.0	0.5	11.9	14.9	13.4
<i>Retail sales</i>	49.2	6.0	9.9	2.7	16.7	4.1	11.2
<i>Consumption</i>	18.9	21.1	13.2	3.3	11.7	16.4	15.5
<i>Durable goods</i>	41.9	1.2	16.9	5.8	7.9	7.4	18.9

Авторы цитируемой работы отмечают, что добавление в модель временного тренда, изменение выборки, переход к *NSA* данным (без сезонной коррекции) сказываются на результатах довольно мало. Дифференцирование нестационарных переменных приводит к значительному уменьшению прогностической силы переменной *FUNDS*, но все же оставляет эту переменную лучшей в большинстве случаев. Таким образом, реальные эффекты денежной политики могут передаваться, скорее, через процентные ставки, чем через денежные агрегаты.

Блочно-треугольная матрица

В связи с произволом в выборе упорядочения переменных при использовании метода Холецкого и зависимостью (в общем случае) результатов от конкретного выбора упорядочения предлагались рецепты преодоления этой проблемы в духе подхода Симса, но с менее строгими ограничениями на матрицу A , которые предполагают блочно-треугольную форму этой матрицы в виде:

$$A = \begin{pmatrix} A_{11} & O & \dots & O \\ A_{21} & A_{22} & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kk} \end{pmatrix}.$$

В соответствии с задачами исследования правильная идентификация может оказаться необходимой только для отдельных блоков такой матрицы. Такой подход использован, например, в работе (*Christiano, Eichenbaum, Evans*, 1998), которую сейчас и рассмотрим.

Пусть S_t — инструмент денежной политики руководящего денежно-кредитного учреждения (ЦБ), например, денежный агрегат или ставка по федеральным ценным бумагам. Предполагается, что

$$S_t = f(\Omega_t) + \sigma_S \varepsilon_t^S,$$

где f — линейная функция (функция отклика), обеспечивающая обратную связь — реакцию на состояние экономики;

Ω_t — информационное множество ЦБ;

$\sigma_S \varepsilon_t^S$ — шок денежной политики, не объясненный реакцией ЦБ на состояние экономики;

ε_t^S — случайная величина, имеющая единичную дисперсию;

σ_S — стандартное отклонение шока денежной политики.

Предположение рекурсивности

Случайная величина ε_t^S не коррелирована со всеми элементами Ω_t (ортогональна всем элементам Ω_t).

Авторы трактуют ε_t^S как результат ошибки измерений предшествующих данных, используемых на момент принятия решения.

Сохраним для удобства обозначения, использованные ранее, т.е. будем записывать структурную VAR в виде

$$A y_t = C(L) y_{t-1} + \zeta_t = C_1 y_{t-1} + \dots + C_p y_{t-p} + \zeta_t, \quad y_t = (y_1, \dots, y_k)^T,$$

где ζ_t — инновационная последовательность *i.i.d.* $(k \times 1)$ -векторов с нулевым математическим ожиданием.

Будем предполагать, что в приведенной форме модель имеет вид

$$y_t = \Pi_1 y_{t-1} + \dots + \Pi_p y_{t-p} + u_t,$$

где $\Pi_j = A^{-1} C_j$, $u_t = A^{-1} \zeta_t$, $\text{Cov}(u_t) = \Sigma$,

и что вектор u_t связан с вектором фундаментальных инноваций ε_t соотношением $u_t = D \varepsilon_t$. Тогда, умножив обе части приведенной формы на матрицу $A_0 = D^{-1}$, получим:

$$A_0 y_t = A_0 \Pi_1 y_{t-1} + \dots + A_0 \Pi_p y_{t-p} + A_0 u_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t.$$

Для восстановления матриц A_0, A_1, \dots, A_p по приведенной форме достаточно знать матрицу A_0 . Для восстановления элементов этой матрицы по элементам матрицы Σ используем соотношение

$$\Sigma = \text{Cov}(u_t) = \text{Cov}(D \varepsilon_t) = \text{Cov}(A_0^{-1} \varepsilon_t) = A_0^{-1} \text{Cov}(\varepsilon_t) (A_0^{-1})^T = A_0^{-1} (A_0^{-1})^T.$$

В матрице A_0 в общем случае надо восстанавливать все k^2 ее элементов. Но для этой цели у нас только $\frac{k(k+1)}{2}$ уравнений. (В случае двумерной VAR имеются 3 уравнения с 4 неизвестными.) Поэтому, чтобы задача была вообще осуществимой, надо накладывать ограничения на структуру матрицы A_0 . В схеме Симса — Холецкого предполагалось, что матрица D треугольная. Это обеспечивало однозначность восстановления матрицы A_0 для каждого выбранного упорядочения переменных. Кристиано, Эйхенбаум, Эванс в своей работе поступают несколько иначе. Они разбивают вектор y_t задействованных в VAR переменных на 3 части:

$$y_t = \begin{pmatrix} X_{1t} \\ S_t \\ X_{2t} \end{pmatrix} \begin{matrix} \} k_1 \\ \} 1 \\ \} k_2 \end{matrix},$$

здесь:

X_{1t} содержит переменные, текущие значения которых входят в Ω_t ;

S_t — упомянутый ранее инструмент денежной политики ЦБ;

X_{2t} содержит переменные, входящие в Ω_t только с запаздывающими значениями;

$k_1 \geq 0, k_2 \geq 0$; если $k_1 = 0$, то $k_2 \geq 1$; если $k_2 = 0$, то $k_1 \geq 1$.

В соответствии с этим разбиением предполагается следующая блочно-нижнетреугольная структура матрицы A_0 :

$$A_0 = \begin{pmatrix} A_{11} & O & O \\ A_{21} & A_{22} & O \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{matrix} k_1 \\ 1 \\ k_2 \end{matrix}, \quad A_{22} = \frac{1}{\sigma_S}, \quad \sigma_S > 0. \quad (6.9)$$

- Нулевой блок в средней строке отражает предположение о том, что ЦБ не наблюдает X_{2t} при установлении S_t .
- Первый нулевой блок в первой строке A_0 исключает прямое влияние S_t на X_{1t} .
- Второй нулевой блок в первой строке A_0 исключает косвенное влияние шоков денежной политики на X_{1t} через посредство переменных, входящих в X_{2t} .

Условие (6.9) не обеспечивает идентификацию всех элементов матрицы A_0 (остаются неразличимыми первые k_1 уравнений и последние k_2 уравнений). Однако оно обеспечивает идентификацию представляющего основной интерес объекта — динамического отклика y_t на шоки денежной политики.

Авторы доказывают, что:

- существует непустое множество матриц A_0 вида (6.9), удовлетворяющих соотношению $A_0^{-1}(A_0^{-1})^T = \Sigma$, одна из которых является нижней треугольной матрицей;
- каждая матрица из этого семейства порождает одну и ту же функцию импульсного отклика элементов y_t на шоки денежной политики ε_t^S ;
- если из указанного семейства всегда выбирается нижняя треугольная матрица, то импульсные отклики переменных в y_t на шоки денежной политики инвариантны относительно упорядочения элементов внутри блоков X_{1t} и X_{2t} (заметим, что это условие не выполняется в отношении других фундаментальных инноваций).

При проведении практических исследований с использованием указанных результатов ключевым становится вопрос о спецификации S_t и Ω_t . Кристиано, Эйхенбаум, Эванс рассматривают две «базовые» (*benchmark*) спецификации, использованные в более ранних работах. Анализируются квартальные данные по США за период с 1965:3 по 1995:2. В VAR(4) включаются следующие переменные:

- Y_t — логарифм реального *GDP*;
 P_t — логарифм дефлятора *GDP*;
 $PCOM_t$ — логарифм индекса цен на промежуточные продукты;
 FF_t — процентная ставка Федерального резерва;
 NBR_t — минус логарифм незаимствованных резервов (*non-borrowed reserves*);
 TR_t — логарифм совокупных резервов (*total reserves*);
 M_t — денежный агрегат *M1* (или *M2*).

Все переменные, за исключением FF_t , скорректированы на сезонность.

Следуя работам (McCallum, 1983), (Bernanke, Blinder, 1992) и (Sims, 1986), авторы в качестве инструмента денежной политики сначала рассматривают:

$$S_t = FF_t,$$

используя в этом случае следующее упорядочение входящих в модель переменных:

$$Y \rightarrow P \rightarrow PCOM \rightarrow FF \rightarrow NBR \rightarrow TR \rightarrow M.$$

Затем, следуя работе (Eichenbaum, Martin, 1992), в качестве инструмента денежной политики они рассматривают:

$$S_t = NBR_t,$$

используя упорядочение:

$$Y \rightarrow P \rightarrow PCOM \rightarrow NBR \rightarrow FF \rightarrow TR \rightarrow M.$$

Графики функций импульсного отклика для *FF*- и *NBR*-моделей во многом похожи, хотя графики для *FF*-модели больше согласуются с теоретическими представлениями. Положительный шок в *FF* ведет к сокращению реального *GDP* и к понижению, в конечном счете, обеих ценовых переменных. Первоначальное возрастание цен является лишь еще одним подтверждением упомянутой ранее «загадки цен». Положительный шок в *FF* ведет также к устойчивому уменьшению незаимствованных резервов (напомним, что при определении переменной *NBR* использовался знак «минус») и к снижению, в конечном счете, общих резервов и *M1*.

Долговременные ограничения

Поскольку экономическая теория часто не в состоянии предложить разумные одновременные ограничения (а чем больше переменных включается в систему, тем большее количество таких ограничений требуется для идентифицируемости), в работах (*Blanchard, Quah, 1989*), (*Shapiro, Watson, 1988*), (*Gali, 1992*) было предложено накладывать ограничения на долговременные свойства системы. Соответствующие **долговременные ограничения** (*long run restrictions*) обычно базируются на некоторых постулатах нейтральности.

Рассмотрим для простоты структурную модель авторегрессии (SVAR) первого порядка

$$Ay_t = Cy_{t-1} + \zeta_t,$$

в которой

$$\zeta_t = B\varepsilon_t,$$

где $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})^T$ — вектор не коррелированных между собой фундаментальных инноваций, имеющих нулевые средние и единичные дисперсии.

Для получения приведенной формы VAR умножим обе части уравнения SVAR на матрицу A^{-1} :

$$y_t = A^{-1}Cy_{t-1} + A^{-1}\zeta_t = \Pi_1 y_{t-1} + A^{-1}B\varepsilon_t = \Pi_1 y_{t-1} + D\varepsilon_t = \Pi_1 y_{t-1} + u_t.$$

Инновации u_{1t}, \dots, u_{kt} в приведенной форме:

$$y_t = \Pi_1 y_{t-1} + u_t,$$

являются линейными комбинациями фундаментальных инноваций,

$$u_t = D\varepsilon_t,$$

где $D = A^{-1}B$.

Мы уже рассматривали задачу идентификации SVAR, полагая матрицу B диагональной, с положительными диагональными элементами. Это означает просто, что инновации ζ_t в SVAR пропорциональны фундаментальным инно-

вациям, а значит, они не коррелированы между собой, но имеют разные дисперсии. Мы обсудили случай, когда матрица A^{-1} предполагается нижней треугольной, что соответствует схеме Холецкого, использованной Симсом, и приводит к рекурсивной SVAR. Затем обсудили методику, использованную в работе Кристиано, Эйхенбаума, Эванса, дающую матрице A больше свободы. Предполагается, что эта матрица блочно-нижнетреугольная, с выделенным элементом на диагонали, соответствующим инструменту денежной политики, представляющему первоочередной интерес.

Подход, о котором пойдет речь, основан на **структурной модели скользящего среднего** (SVMA — *structural vector moving average*), получаемой на основе модели SVAR.

Если VAR(1) удовлетворяет условию стабильности, то из приведенной формы можно получить:

$$(I_k - \Pi_1 L)y_t = u_t \rightarrow \\ \rightarrow y_t = (I_k - \Pi_1 L)^{-1} u_t = (I_k - \Pi_1 L)^{-1} A^{-1} B \varepsilon_t = \Theta(L) \varepsilon_t = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j}.$$

Пусть имеем дело с двумя рядами. Тогда:

$$\Theta_j = \begin{pmatrix} \theta_{11,j} & \theta_{12,j} \\ \theta_{21,j} & \theta_{22,j} \end{pmatrix}, \quad j = 0, 1, 2, \dots, \\ \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \theta_{11,0} & \theta_{12,0} \\ \theta_{21,0} & \theta_{22,0} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} + \begin{pmatrix} \theta_{11,1} & \theta_{12,1} \\ \theta_{21,1} & \theta_{22,1} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix} + \dots$$

Это не что иное, как **разложение Вольда** (*Wold decomposition*) для двумерного ковариационно-стационарного процесса.

Предположим, что шок ε_{2t} не оказывает немедленного влияния на y_{1t} . Это означает, что $\theta_{12,0} = 0$ в SMA, так что:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \theta_{11,0} & 0 \\ \theta_{21,0} & \theta_{22,0} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} + \begin{pmatrix} \theta_{11,1} & \theta_{12,1} \\ \theta_{21,1} & \theta_{22,1} \end{pmatrix} \begin{pmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{pmatrix} + \dots, \\ \Theta_0 = \begin{pmatrix} \theta_{11,0} & 0 \\ \theta_{21,0} & \theta_{22,0} \end{pmatrix}.$$

Но Θ_0 — не что иное, как $A^{-1}B = D$. Если $D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}$, то принятое

предположение записывается в виде $d_{12} = 0$.

Для определения остальных 3 элементов матрицы D и восстановления матриц A и B можно использовать 3 элемента ковариационной матрицы $\text{Cov}(u_t) = \Sigma$, как мы уже делали раньше.

Здесь для идентификации SVAR используются:

- ограничения на матрицу B (диагональная, с положительными диагональными элементами);
- предположение об отсутствии немедленной реакции y_{1t} на шок ε_{2t} .

Долговременные ограничения предполагают отсутствие не немедленной, а накопленной реакции. Соответственно, если опять имеем дело с двумя рядами, то «нейтральность» влияния шока ε_{2t} на y_{1t} понимается как выполнение соотношения $\theta_{12}(1) = \sum_{j=0}^{\infty} \theta_{12,j} = 0$ (сумма импульсных откликов первой переменной на шок ε_{2t} равна нулю). При этом

$$\Theta(1) = \begin{pmatrix} \sum_{j=0}^{\infty} \theta_{11,j} & 0 \\ \sum_{j=0}^{\infty} \theta_{21,j} & \sum_{j=0}^{\infty} \theta_{22,j} \end{pmatrix} = \begin{pmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{pmatrix}.$$

Но поскольку $y_t = (I_k - \Pi_1 L)^{-1} u_t = (I_k - \Pi_1 L)^{-1} A^{-1} B \varepsilon_t = \Theta(L) \varepsilon_t$, то:

$$\Theta(L) = (I_k - \Pi_1 L)^{-1} D,$$

$$\Theta(1) = (I_k - \Pi_1)^{-1} D = (I - A^{-1} C)^{-1} A^{-1} B.$$

Это приводит к нелинейному ограничению на коэффициенты SVAR.

Обозначим для простоты $Q = (I_k - \Pi_1)^{-1}$. Имеем:

$$\Sigma = \text{Cov}(u_t) = \text{Cov}(D \varepsilon_t) = D \text{Cov}(\varepsilon_t) D^T = D D^T.$$

В то же время

$$\Theta(1) = QD,$$

$$\Theta(1)\Theta(1)^T = Q D D^T Q^T = Q \Sigma Q^T.$$

При сделанных предположениях матрица $\Theta(1)$ — нижняя треугольная и однозначно восстанавливается через элементы матрицы Σ .

Бланшар и Квах (*Blanchard, Quah, 1989*) предполагают, что Σ — единичная матрица, считая, что это просто вопрос нормировки шоков. При таком предположении $D = A^{-1}$, и $A^{-1} = (I - \Pi_1)\Theta(1)$, т.е. матрица A может быть восстановлена.

Обратимся теперь собственно к работе (*Blanchard, Quah, 1989*). В ней рассматривается VAR для двух рядов (квартальные данные):

- $\Delta \ln GNP$, где GNP — реальный GNP ;
- U — уровень безработицы.

Если полагать, что это стабильная VAR, то для нее получаем представление Вольда:

$$y_t = \begin{pmatrix} \Delta \ln GNP_t \\ U_t \end{pmatrix}, \quad \Theta(L)\varepsilon_t = \sum_{j=0}^{\infty} \Theta_j \varepsilon_{t-j},$$

где $\varepsilon_t = \begin{pmatrix} \varepsilon_t^{AS} \\ \varepsilon_t^{AD} \end{pmatrix}$,

ε_t^{AD} и ε_t^{AS} — не коррелированные между собой как в совпадающие, так и в несовпадающие моменты времени фундаментальные инновации;

$$E(\varepsilon_t) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad Cov(\varepsilon_t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

При этом ε_t^{AD} авторы интерпретируют как возмущения совокупного спроса (шоки предложения денег или скорости обращения денег, шоки фискальной политики, «спонтанный оптимизм» предпринимателей — то, что Дж.М. Кейнс назвал «жизнерадостностью» (*animal spirits*)), а ε_t^{AS} — как возмущения совокупного предложения (технологический прогресс, предложение трудовых ресурсов, цены на нефть).

Поскольку VAR стационарна, ни одна из этих двух фундаментальных инноваций не может оказывать перманентное воздействие ни на $\Delta \ln GNP_t$, ни на U_t . В то же время обе они могли бы оказывать долговременное воздействие на уровень GNP_t . Авторы исходят из того, что возмущения совокупного спроса не оказывают долговременного воздействия на уровень GNP_t . Это выражается соотношением:

$$\begin{pmatrix} \Delta \ln GNP_t \\ U_t \end{pmatrix} = \begin{pmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{pmatrix} \begin{pmatrix} \varepsilon_t^{AS} \\ \varepsilon_t^{AD} \end{pmatrix}.$$

В статье приведены результаты оценивания VAR(8) по квартальным данным на периоде с 1950:2 по 1987:4. Имея в виду известный мировой нефтяной кризис 1973 г., авторы вычислили отдельно средние значения для темпов прироста GNP на подпериодах 1948:2 — 1973:4 и 1974:1 — 1987:4, которые в пересчете на годовые величины составили 3.62% и 2.43% соответственно, и вычли их из значений этого ряда. Затем был оценен и удален тренд из ряда уровней безработицы. Эти операции проводились с целью «остационарить» VAR, хотя бы приближенно, чтобы можно было воспользоваться изложенной выше методикой.

Заметим, что сделанные ранее идентифицирующие ограничения накладывают единственное ограничение на декомпозиции дисперсий ошибок прогнозов: доля вкладов возмущений предложения в дисперсию ошибок прогнозов валового продукта стремится к 100% с ростом горизонта прогноза.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В чем состоит основной недостаток идентификации структурной формы VAR с использованием упорядочения по Холецкому? Как производится идентификация с использованием блочно-треугольной матрицы?
2. Как можно добиться идентифицируемости структурной формы на основе долгосрочных ограничений?

Тема 6.4

НЕСТАБИЛЬНЫЕ VAR

Можно было бы и далее умножать примеры использования указанной выше методологии, однако при ближайшем их рассмотрении обращает на себя внимание поведение временных рядов, включаемых в модели VAR. Если, например, вернуться к модели $TBILL3 \rightarrow M2 \rightarrow P \rightarrow Y$, то на проанализированном периоде входящие в нее ряды имеют вид, изображенный на рис. 6.31.

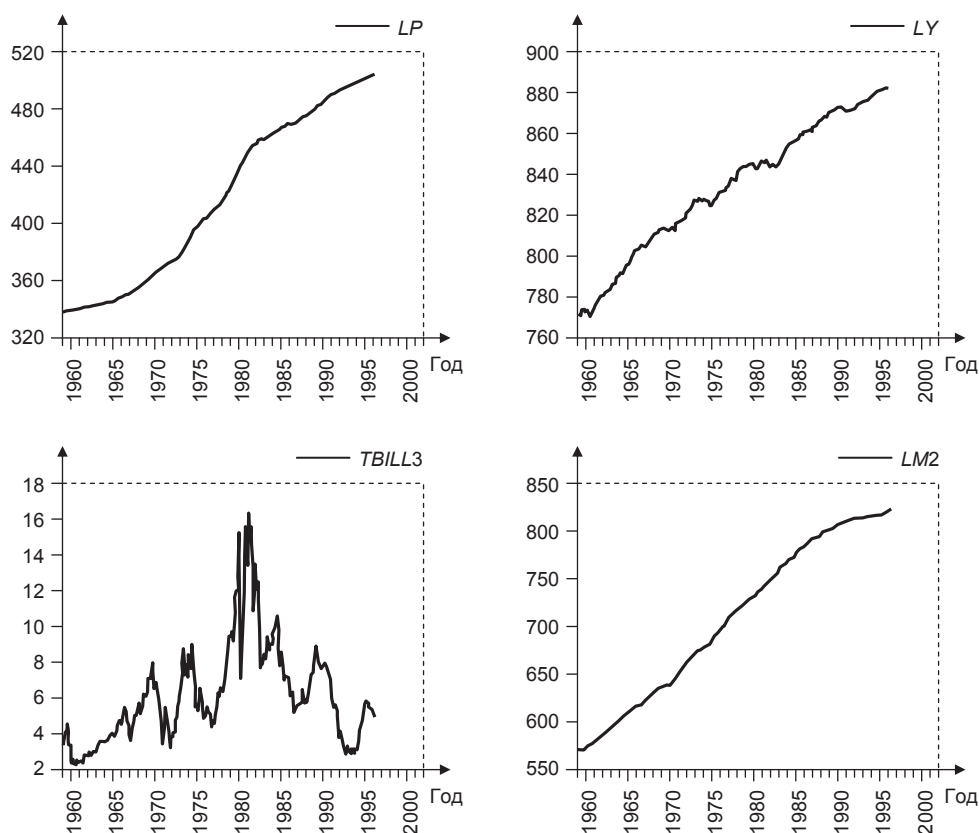


Рис. 6.31

Естественно, возникает вопрос о законности использования инструментария, предназначенного для анализа стабильных моделей VAR, в случаях, когда условие стабильности не выполнено. Этому вопросу посвящена работа Филлипса (*Phillips*, 1995).

Филлипс обнаружил, что при наличии единичных корней или корней, близких к единице, импульсные отклики, полученные на основе оцененной приведенной VAR без ограничений, в долгосрочном плане сходятся в пределе не к истинным откликам, а к некоторым случайным величинам. Предельное распределение не является нормальным и несимметрично, что объясняет несимметричность распределений, получаемых для импульсных откликов при статистическом моделировании. К тому же предсказания на основании нестационарной VAR без ограничений не сходятся к оптимальным прогнозам (по крайней мере, на больших горизонтах прогнозов), а дисперсия оптимального прогноза, как правило, занижается.

Напротив, если имеет место коинтегрированность рядов, образующих VAR, и оценивается модель VAR, учитывающая ограничения, накладываемые коинтеграцией, то построенные по оцененной модели импульсные отклики и декомпозиции дисперсий прогнозов состоятельны, а прогнозы асимптотически оптимальны. Правда, для этого надо или правильно специфицировать ранг коинтеграции, или получить для него состоятельную оценку.

Заметим, что в модели VAR с 5 переменными *TBILL3*, *PCM*, *M2*, *P*, *Y* наличие коинтеграции между переменными на периоде 1983:01 — 1996:12 не отвергается.

Попробуем проверить, что же происходит на практике при построении функций отклика по оцененным моделям.

ПРИМЕР 6.3.1

Смоделирована VAR(1), которая отличается от ранее построенной тем, что в ней $\pi_{22} = 2/3$ (а не $5/8$, как это было ранее). В новой модели один из корней характеристического уравнения равен 1, так что VAR нестабильна. Графики полученных рядов изображены на рис. 6.32.

Проверка на коинтегрированность дает положительные результаты. Протоколы оценивания UVAR(1) (VAR без ограничений) и ECM (модели коррекции ошибок) приведены в табл. 6.22 и 6.23 соответственно. В табл. 6.24 указаны коэффициенты приведенной формы, соответствующие оцененной ECM.

Разница между оценками, полученными по модели UVAR (VAR без ограничений на коэффициенты) и по модели с ограничениями на коэффициенты (*RRR* — *reduced rank regression*), весьма мала.

Сравним последовательности импульсных откликов, полученные по оцененным моделям (на рис. 6.33 — отклики для оцененной UVAR, на рис. 6.34 — отклики для оцененной ECM).

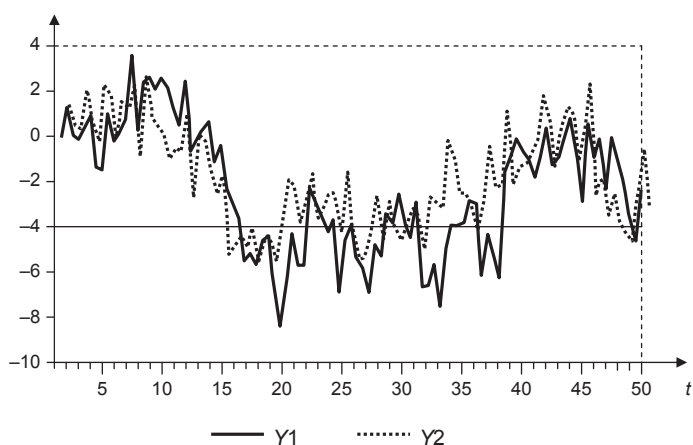


Рис. 6.32

Таблица 6.22

Оценка VAR без ограничений (UVAR(1))

Vector Autoregression Estimates; Standard errors in () & t-statistics in []

Переменная	Y1	Y2
Y1(-1)	0.614554	0.253295
	(0.04487)	(0.05647)
	[13.6956]	[4.48537]
Y2(-1)	0.488234	0.655997
	(0.05317)	(0.06692)
	[9.18179]	[9.80288]

Таблица 6.23

Оценивание модели коррекции ошибок (ECM)

Vector Error Correction Estimates; Standard errors in () & t-statistics in []

Cointegrating Eq:	CointEq1	
Y1(-1)	1.000000	
Y2(-1)	-1.243561	
	(0.08927)	
	[-13.9303]	
Error Correction:	D(Y1)	D(Y2)
CointEq1	-0.389636	0.266943
	(0.04030)	(0.05079)
	[-9.66722]	[5.25541]

Таблица 6.24

Коэффициенты приведенной формы, соответствующие оцененной ЕСМ

Переменная	Y1	Y2
Y1(-1)	0.610	0.270
Y2(-1)	0.484	0.664

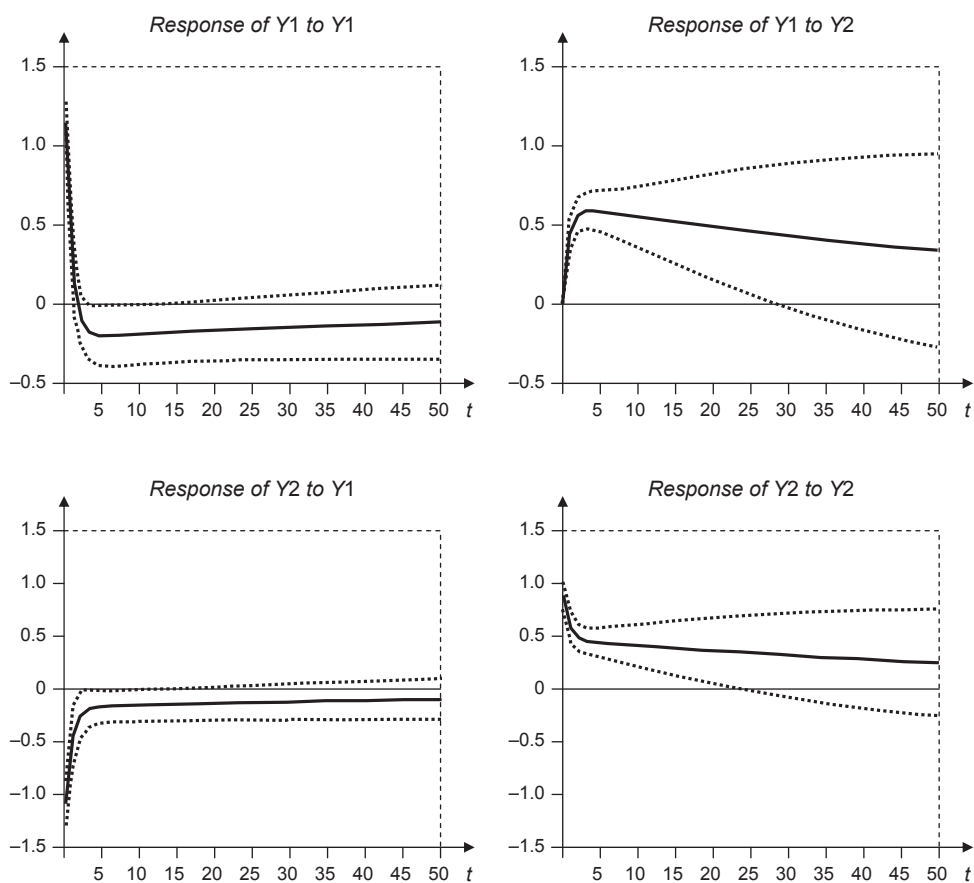
Response to Cholesky One S.D. Innovations ± 2 S.E.

Рис. 6.33

Поведение последовательностей откликов, полученных по оцененной VAR и по оцененной ЕСМ, весьма различно. В связи с этим интересно посмотреть на поведение последовательностей откликов в модели, использованной в DGP.

Response to Cholesky One S.D. Innovations

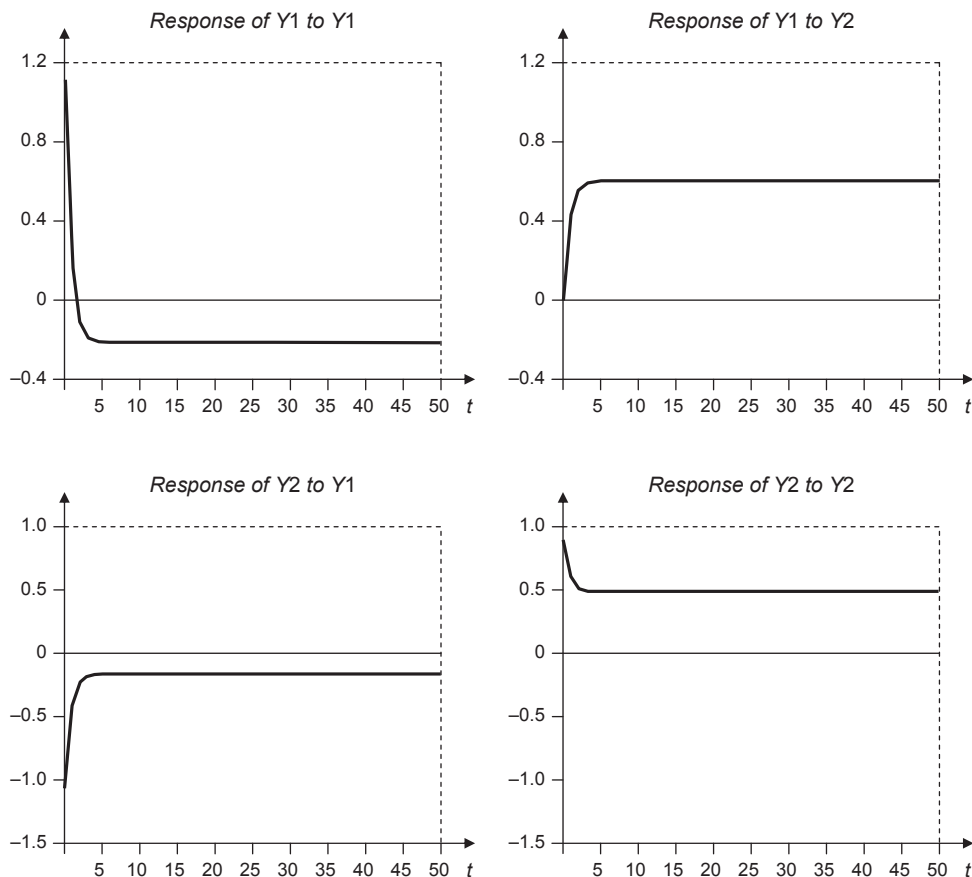


Рис. 6.34

Построим функции откликов на шок в $Y1$ по модели, порождавшей данные, используя те же последовательности фундаментальных инноваций, что и в DGP . Для этого прибавим единицу ко второму значению ряда инноваций для первого уравнения — получим модифицированный ряд инноваций для первого уравнения. Специфицируем новую модель порождения данных, в которой коэффициенты те же, что и в DGP , но ряд инноваций для первого уравнения — модифицированный, как указано выше. Построим реализацию этой модифицированной модели в виде модифицированного ряда $Y1_modif$ и модифицированного ряда $Y2_modif$. Тогда ряд $(Y1_modif - Y1)$ представляет последовательность импульсных откликов переменной $Y1$ на шок этой переменной, а ряд $(Y2_modif - Y2)$ — последовательность импульсных откликов переменной $Y2$ на шок переменной $Y1$. Из графиков на рис. 6.35 видно, что

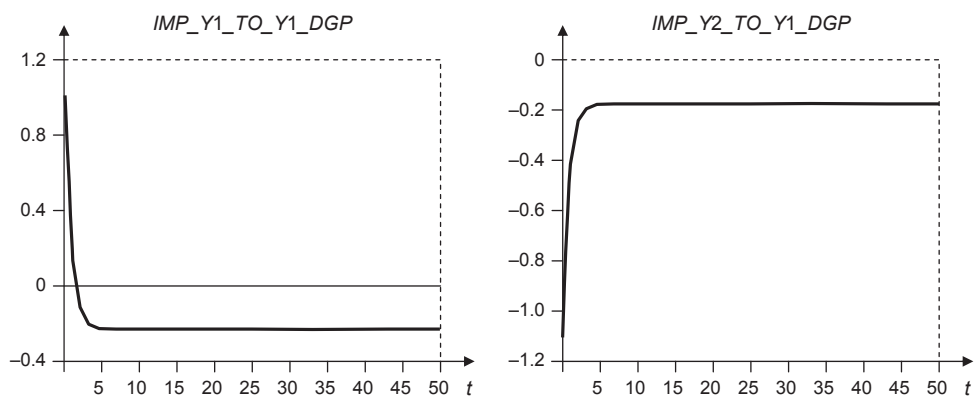


Рис. 6.35

Response to Cholesky One S.D. Innovations ± 2 S.E.

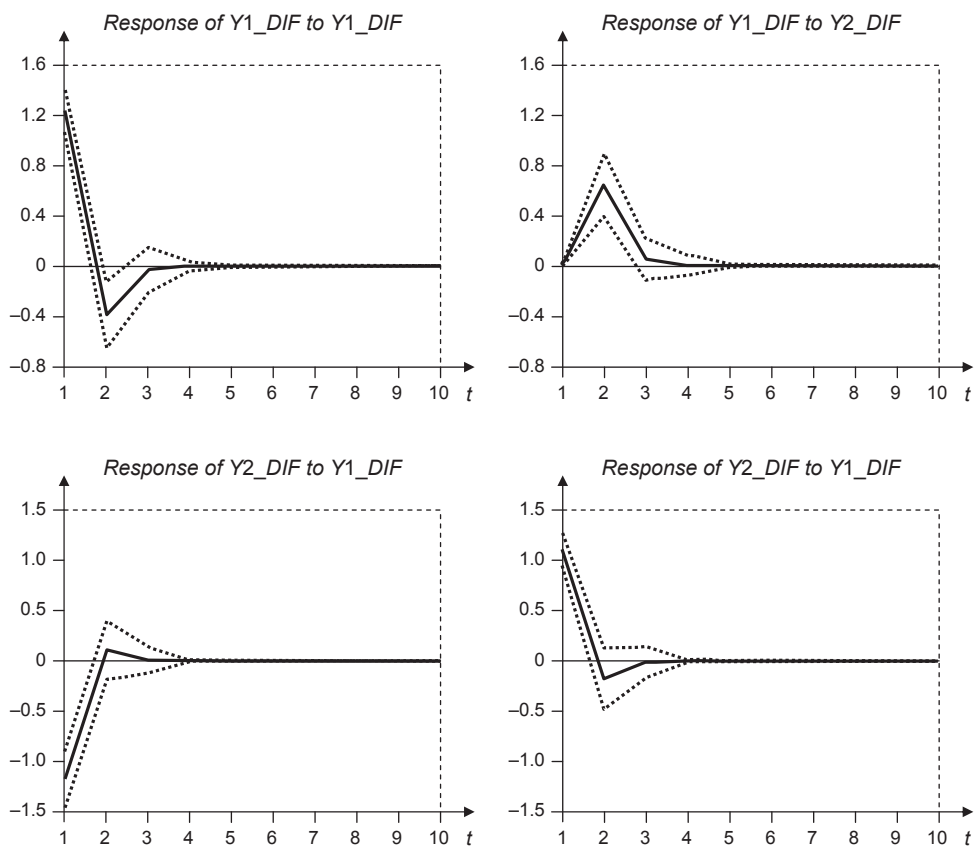


Рис. 6.36

в долгосрочном плане поведение функций отклика при использовании UVAR не соответствует поведению функций отклика в *DGP*, тогда как поведение функций отклика при использовании ЕСМ соответствует поведению функций отклика в *DGP*.

Интересно обратить внимание на поведение функций импульсного отклика в коинтегрированной системе в плане их «выполаживания» и на поведение откликов приращений в оцененной модели VAR с двумя переменными для продифференцированных рядов (рис. 6.36). ■

Рассмотрим теперь пример с 3 рядами.

ПРИМЕР 6.3.2

Смоделируем реализации 3 рядов, следуя *DGP*:

$$\Delta y_{1t} = \varepsilon_{1t}, \quad y_{2t} = y_{1t} + \varepsilon_{2t}, \quad y_{3t} = y_{1t} + \varepsilon_{3t}, \quad t = 1, \dots, 500,$$

где ε_{1t} , ε_{2t} , ε_{3t} — не коррелированные между собой гауссовские инновации с нулевыми средними и единичными дисперсиями, $y_{10} = 0$.

Получаемые $I(1)$ -ряды коинтегрированы, ранг коинтеграции равен 2.

Построим функции импульсного отклика (при упорядочении $Y1 \rightarrow Y2 \rightarrow Y3$), соответствующие оцененной модели UVAR (рис. 6.37), оцененной модели ЕСМ с рангом коинтеграции 1 (рис. 6.38) и оцененной модели коинтеграции с рангом коинтеграции 2 (рис. 6.39). (В обе модели ЕСМ не включаются константа и тренд).

Как видно, использование при оценивании ЕСМ различных предположений о ранге коинтеграции приводит к разным результатам в отношении функций импульсного отклика. Это весьма затрудняет интерпретацию результатов VAR-анализа при обращении к данным экономической статистики. ■

Ранее отмечалось, что если VAR в уровнях нестационарна и ряды коинтегрированы, то VAR в разностях не может иметь конечный порядок — надо специфицировать ЕСМ.

Если же VAR в уровнях нестационарна, ряды имеют порядок $I(1)$ и не коинтегрированы, то можно на законных основаниях строить функции импульсного отклика VAR в разностях, если VAR в разностях стабильна. В этом случае ряд разностей Δy_t представляется в виде $\Delta y_t = (1 - L)y_t = C(L)\varepsilon_t$, где $C(L)$ — функция импульсного отклика VAR в разностях, $C(L) = C_0 + C_1L + C_2L^2 + \dots$. Если $C^*(L)$ — функция импульсного отклика в уровнях (*accumulated responses*), так что $y_t = C^*(L)\varepsilon_t$, $C^*(L) = C_0^* + C_1^*L + C_2^*L^2 + \dots$, то ее можно восстановить последовательным накоплением импульсных откликов в разностях:

$$C_j^* = C_0 + C_1 + \dots + C_j.$$

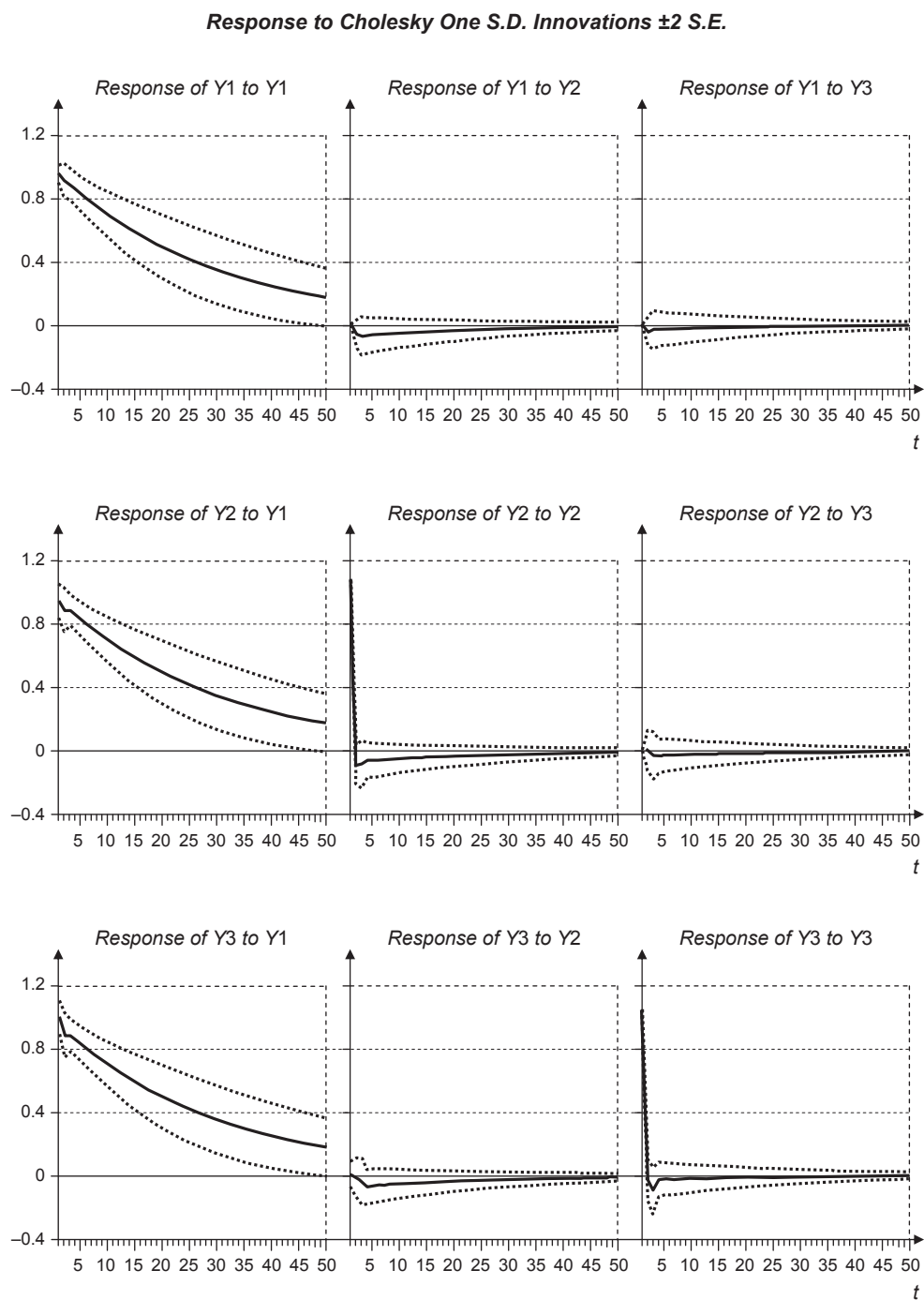


Рис. 6.37

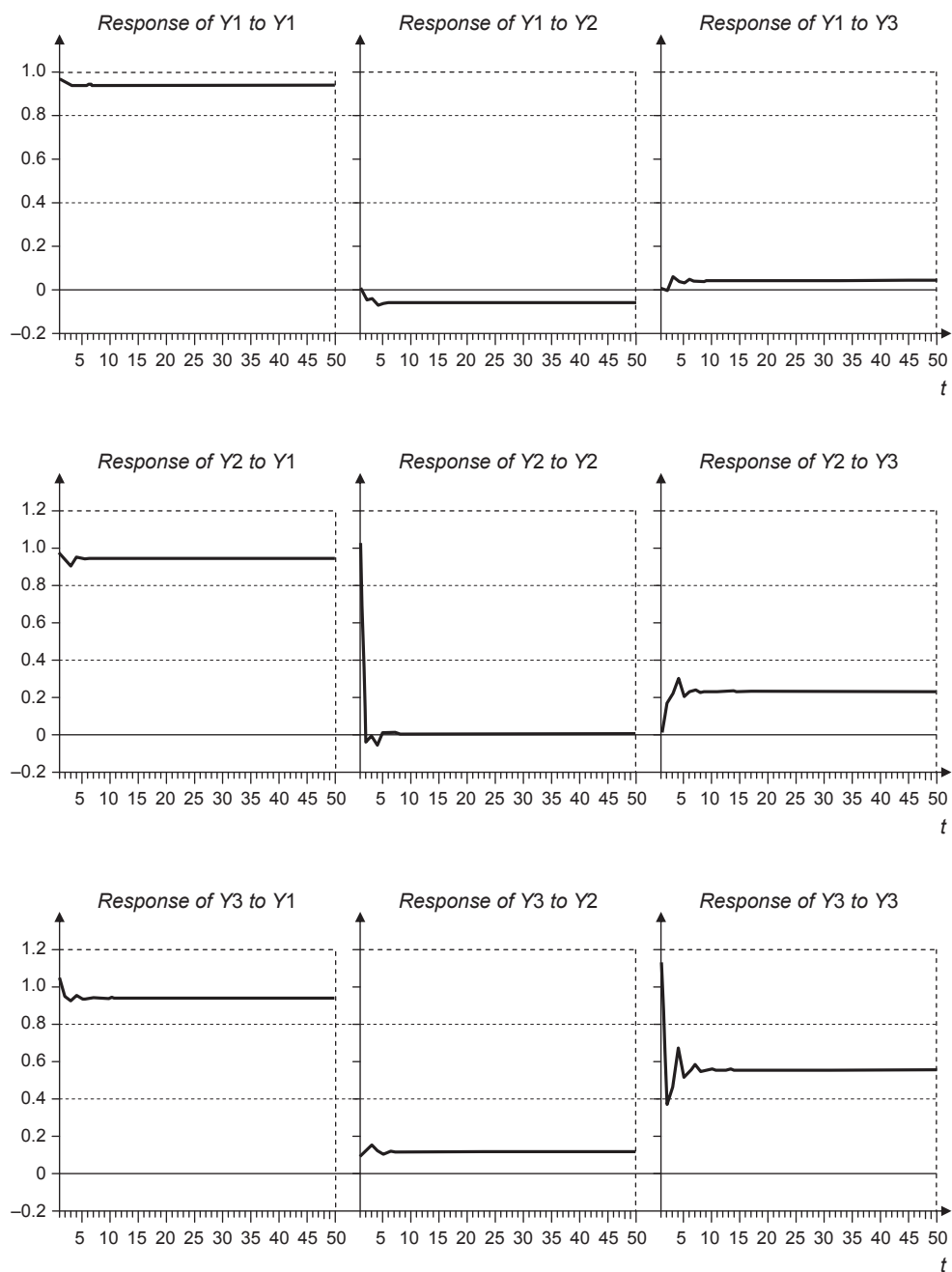
Response to Cholesky One S.D. Innovations

Рис. 6.38

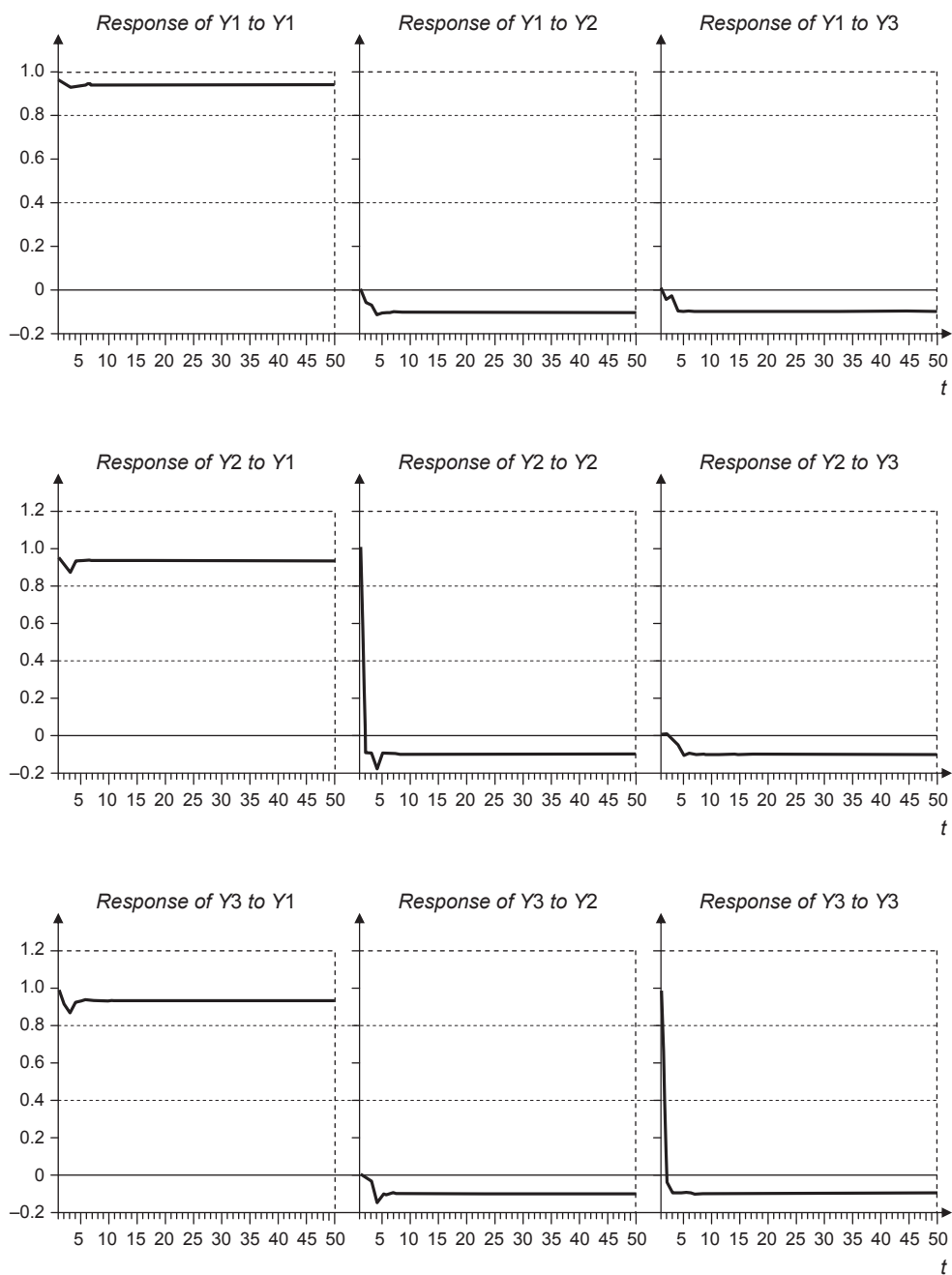
Response to Cholesky One S.D. Innovations

Рис. 6.39

При практическом использовании методологии VAR для построения функции импульсного отклика для ряда Δy_t к имеющимся данным подбирается VAR конечного порядка, и оцененная модель обращается для получения *МА*-коэффициентов, составляющих функцию импульсного отклика. Разумеется, ряд Δy_t может не иметь конечного *AR*-представления, и использование VAR конечного порядка основано на предположении о том, что *AR*-представление существует и может быть приближено VAR моделью конечного (быть может, достаточно высокого) порядка. Однако такое предположение может и не выполняться.

Выше уже говорилось о том, что стационарный процесс, имеющий представление Вольда в виде скользящего среднего (возможно, бесконечного порядка), может не иметь представления в виде процесса авторегрессии (быть может, бесконечного порядка). В качестве примера был приведен процесс *МА*(1):

$$x_t = \varepsilon_t - \varepsilon_{t-1} = (1 - L) \varepsilon_t,$$

где $\varepsilon_t \sim i.i.d.$, с нулевым средним и конечной дисперсией.

Подобно этому векторный стационарный процесс, имеющий представление Вольда в виде векторного скользящего среднего (*MAR* — *MA-representation*), может не иметь авторегрессионного представления (*ARR* — *AR-representation*). В качестве примера рассмотрим модель *VMA*(1) (векторное скользящее среднее первого порядка):

$$\begin{aligned} x_{1t} &= 3 + \varepsilon_{1t} - \varepsilon_{1,t-1}, \\ x_{2t} &= -3 + \varepsilon_{2t} - 0.5 \varepsilon_{2,t-1} + 0.5 \varepsilon_{1,t-1}, \end{aligned} \quad (6.10)$$

где ε_{1t} и ε_{2t} — не коррелированные между собой процессы белого шума.

В этой модели компонента x_{1t} порождается автономно, независимо от x_{2t} , и получается в результате дифференцирования ряда

$$y_{1t} = 3t + \varepsilon_{1t},$$

стационарного относительно линейного тренда. Соответственно ряд x_{1t} является передифференцированным.

В векторно-матричной форме рассматриваемый процесс записывается следующим образом:

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} + \begin{pmatrix} 1-L & 0 \\ 0.5L & 1-0.5L \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Уравнение $\det \begin{pmatrix} 1-z & 0 \\ 0.5z & 1-0.5z \end{pmatrix} = 0$ принимает вид

$$(1-z)(1-0.5z) = 0$$

и имеет два корня: $z = 1$ и $z = 2$. Наличие единичного корня означает необратимость рассматриваемой *VMA*-модели.

Если ряд разностей порождается необратимой VMA-моделью, то для нее не существует авторегрессионного представления, и в такой ситуации кажется бессмысленным использовать VAR-методологию для построения импульсных откликов. Однако в работе (Marcet, 2001) утверждается, что последний вывод является в определенной мере заблуждением.

Пусть ряд разностей порождается необратимой VMA-моделью. Предлагается, несмотря на это, действовать обычным образом: по имеющимся статистическим данным построить VAR(p)-модель для ряда Δy_t и соответствующую ей модель VMA, получить по этому MA-представлению импульсные отклики и путем накопления последних построить накопленные импульсные отклики $C_j^{*,p}$. Утверждается, что если ковариационная матрица инноваций процесса Δy_t не вырождена, то

$$C_j^{*,p} \rightarrow C_j^* \quad \text{при } p \rightarrow \infty,$$

где C_j^* — накопленный отклик, построенный на основе истинной VMA-модели для ряда разностей.

Вернемся к модели (6.10) и заметим, что ее MA-представление есть просто

$$\begin{aligned} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} &= \begin{pmatrix} 3 \\ -3 \end{pmatrix} + \begin{pmatrix} 1-L & 0 \\ 0.5L & 1-0.5L \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \\ &= \begin{pmatrix} 3 \\ -3 \end{pmatrix} + \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0.5 & -0.5 \end{pmatrix} L \right) \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \end{aligned}$$

так что, в использованных выше обозначениях:

$$C_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C_1 = \begin{pmatrix} -1 & 0 \\ 0.5 & -0.5 \end{pmatrix}, \quad C_j = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{для } j = 2, 3, \dots,$$

т.е. импульсные отклики переменных на шоки инноваций имеют значения, приведенные в табл. 6.25.

Таблица 6.25

Импульсные отклики переменных на шоки инноваций

t	1	2	3	4	5	6
$Imp(x_{1t} \varepsilon_{11})$	1	-1	0	0	0	0
$Imp(x_{2t} \varepsilon_{11})$	0	0.5	0	0	0	0
$Imp(x_{1t} \varepsilon_{21})$	0	0	0	0	0	0
$Imp(x_{2t} \varepsilon_{21})$	1	-0.5	0	0	0	0

Соответственно накопленные импульсные отклики, построенные по истинной VMA-модели, имеют вид:

$$C_0^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C_1^* = C_0 + C_1 = \begin{pmatrix} 0 & 0 \\ 0.5 & 0.5 \end{pmatrix}, \quad C_j = \begin{pmatrix} 0 & 0 \\ 0.5 & 0.5 \end{pmatrix} \text{ для } j = 2, 3, \dots$$

Их значения приведены в табл. 6.26.

Таблица 6.26

Импульсные отклики, построенные по истинной VMA-модели

j	1	2	3	4	5
$C_{j,11}^*$	1	0	0	0	0
$C_{j,21}^*$	0	0.5	0.5	0.5	0.5
$C_{j,12}^*$	0	0	0	0	0
$C_{j,22}^*$	1	0.5	0.5	0.5	0.5

Смоделируем реализацию модели (6.10) длины $T = 500$ и оценим модель VAR(5) по первым 50 наблюдениям. На рис. 6.40 показаны графики накопленных импульсных откликов для оцененной модели. Накопленные импульсные отклики, построенные на основании модели VAR(5), оцененной по 50 наблюдениям, существенно отличаются от теоретических.

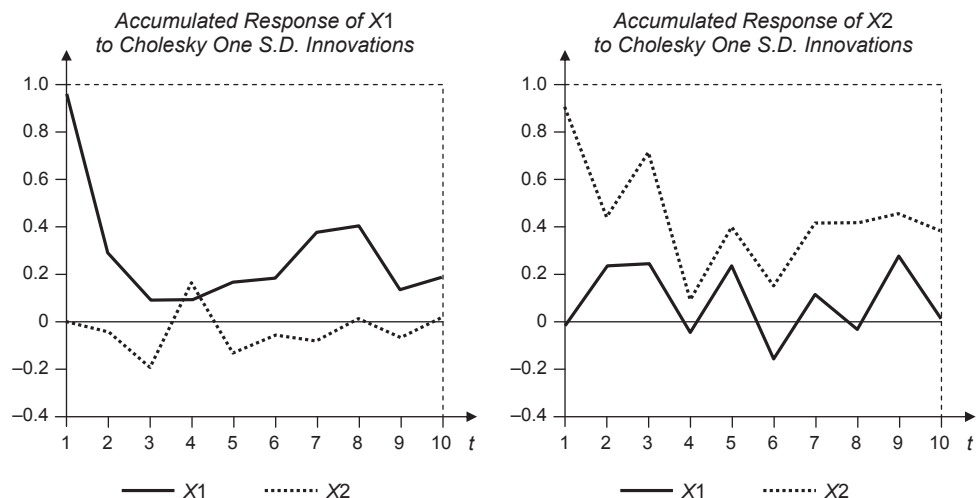


Рис. 6.40

Оценим теперь VAR(10) первым по 100 наблюдениям. На рис. 6.41 показаны графики накопленных импульсных откликов для оцененной модели.

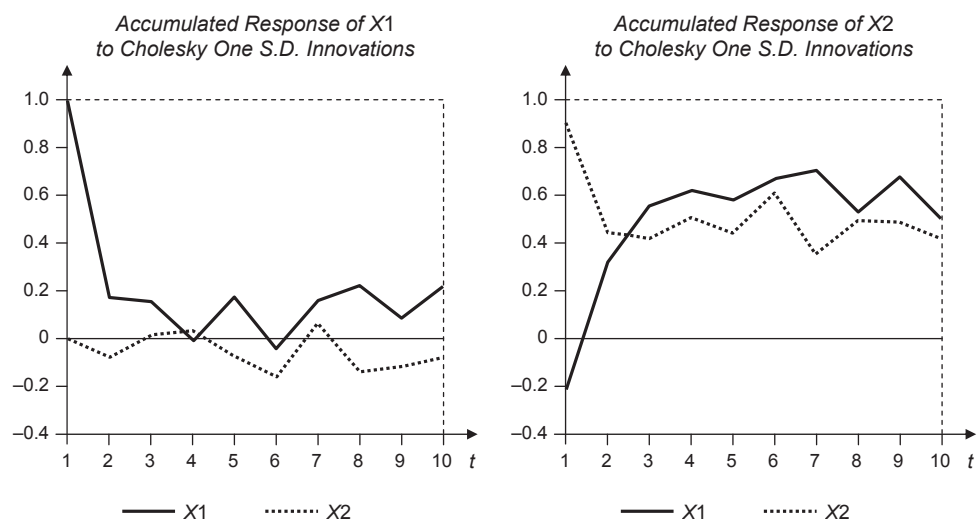


Рис. 6.41

Накопленные импульсные отклики, построенные на основании модели VAR(10), оцененной по 100 наблюдениям, существенно ближе к теоретическим. Еще ближе к теоретическим накопленные импульсные отклики, построенные на основании модели VAR(50), оцененной по 500 наблюдениям (рис. 6.42).

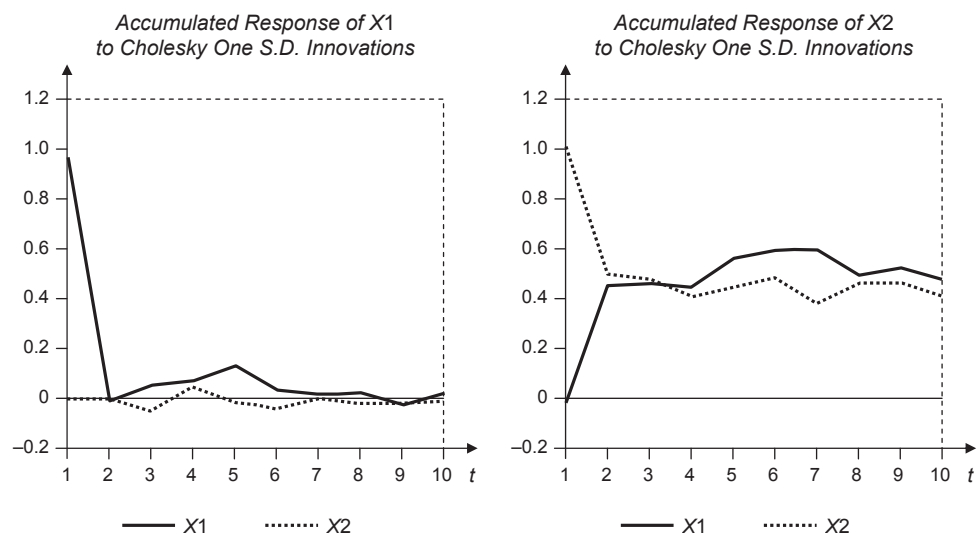


Рис. 6.42

Причинность по Грейнджеру, функции импульсного отклика и декомпозиция дисперсий

Выше было исследовано влияние текущих и прошлых значений изучаемых переменных, а также их инноваций на будущие значения переменных. Тесты на наличие причинности по Грейнджеру/блочную экзогенность фактически только выявляют наличие G -причинности в рамках имеющейся выборки, но не указывают ни на знак и силу выявленной связи, ни на динамику усиления/ослабления влияния одних переменных на другие с течением времени. Функции импульсного отклика и декомпозиции дисперсий позволяют проследить динамику влияния шоковых воздействий на будущие значения переменных. Свойство отсутствия G -причинности между переменными двух групп является исключительным свойством двух подмножеств переменных данного процесса. В то же время декомпозиция дисперсий ошибок прогнозов не единственна: она зависит от матрицы, приводящей к фундаментальным инновациям.

Важно отметить, что если набор переменных в составе y_t разбивается на две части:

$$y_t = \begin{pmatrix} y_t^{(1)} \\ y_t^{(2)} \end{pmatrix},$$

где $y_t^{(1)} = (y_{1,t}, \dots, y_{N_1,t})^T$ — первая группа переменных;

$y_t^{(2)} = (y_{N_1+1,t}, \dots, y_{N,t})^T$ — вторая группа переменных,

и переменные второй группы не являются Грейнджер-причиной для переменных первой группы, то это вовсе не означает, что доли дисперсий ошибок прогнозов переменных первой группы, соответствующие инновациям в уравнениях для переменных второй группы, равны нулю. Последнее может иметь место, только если отсутствует перекрестная коррелированность инноваций в уравнениях для переменных первой группы с инновациями в уравнениях для переменных второй группы. Если же это условие нарушено, то доли дисперсий ошибок прогнозов переменных первой группы, соответствующие инновациям в уравнениях для переменных второй группы, отличны от нуля.

Вернемся к рассмотрению примеров, в которых производилась проверка на наличие причинности по Грейнджеру.

ПРИМЕР 6.3.3

Ранее рассматривалась модель, в которой были использованы переменные $LP = 100\ln(CPI_SA)$, $LY = 100\ln(RGDP_MON)$, $LM2 = 100\ln(M2_SA)$, и оценивание модели производилось на периоде 1974:01—1980:03. Там было выявлено наличие G -причинной связи в направлении от LY к $LM2$. Построим

теперь в оцененной VAR для разностей импульсные функции и декомпозиции дисперсий, используя тот же порядок вхождения переменных, что и в работе (Leeper, Sims, Zha, 1996), т.е. $P \rightarrow Y \rightarrow M2$.

Полученные декомпозиции дисперсий ошибок прогнозов представлены на рис. 6.43. Заметим, что доля инноваций LP_DIF в дисперсии ошибок прогнозов переменной $LM2_DIF$ доходит до 22%, тогда как доля инноваций LY_DIF достигает лишь 6.5%.

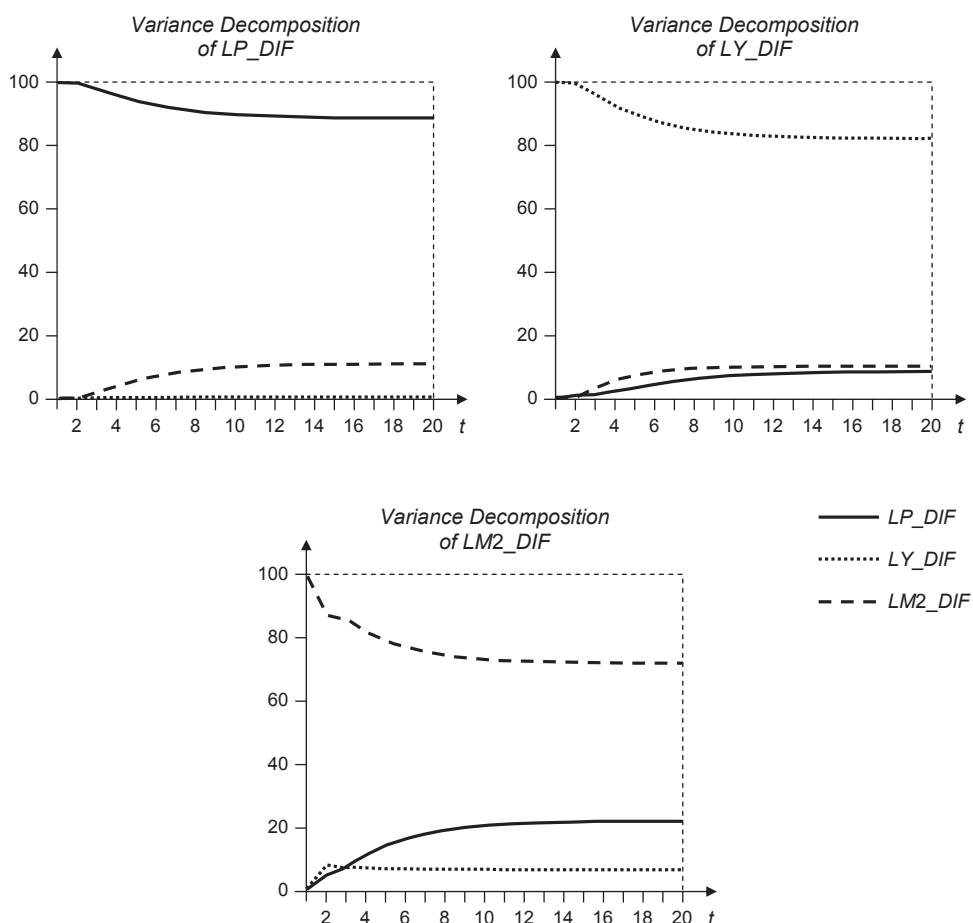


Рис. 6.43

Функции импульсного отклика представлены на рис. 6.44. И здесь влияние инноваций LY_DIF на $LM2_DIF$ значимо только при продвижении на один шаг, тогда как влияние инноваций LP_DIF на $LM2_DIF$ прослеживается в течение более длительного времени. ■

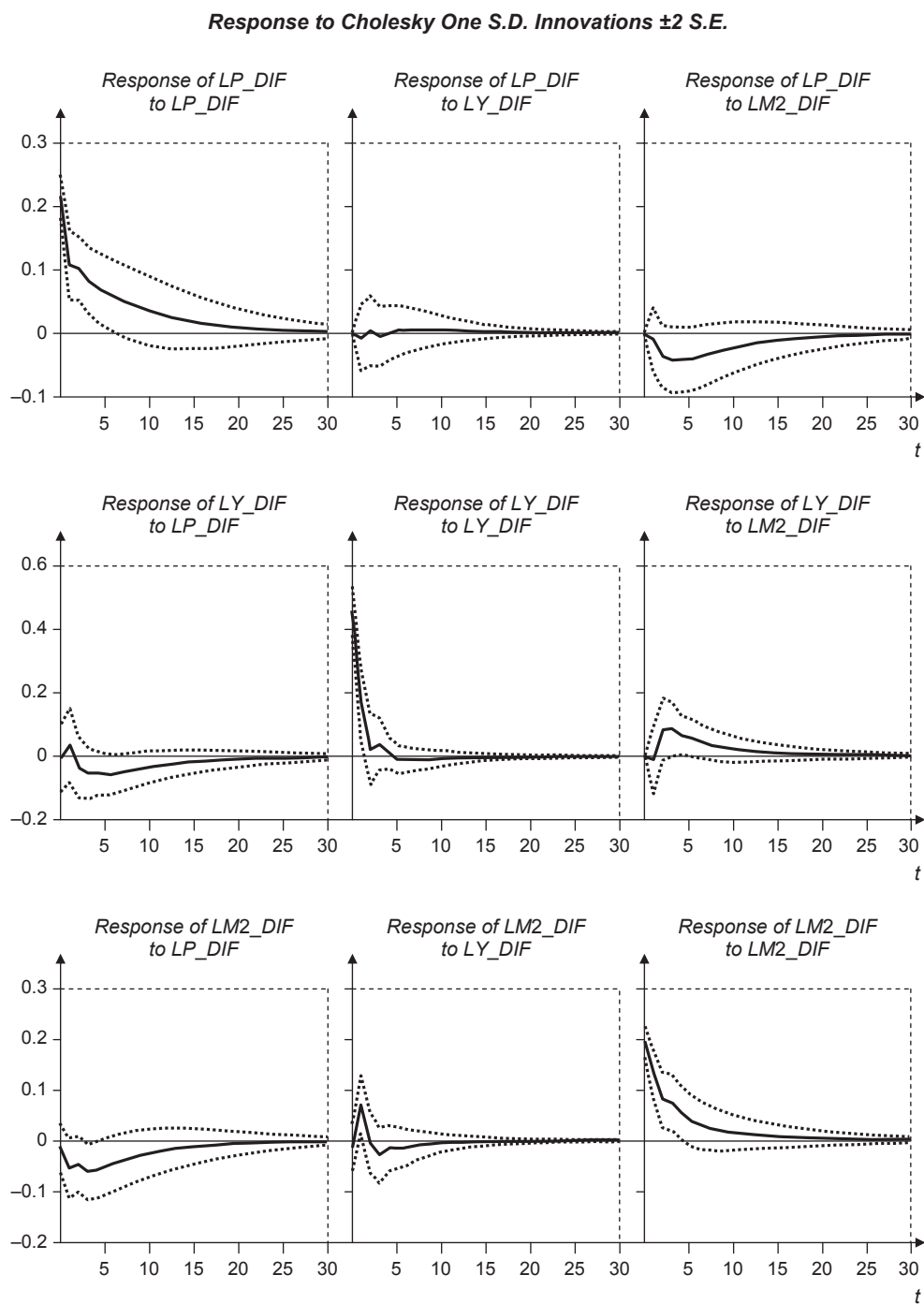


Рис. 6.44

ПРИМЕР 6.3.4

Рассмотрим опять данные по фармацевтической фирме *Lydia Pinkham*. Ранее, привлекая к исследованию связей между переменными *Sales* и *Adver* дополнительную переменную *Income*, был сделан вывод о том, что $Adver \xrightarrow{G} Sales$ и $Sales \xrightarrow{G} Adver$. Построение функций импульсного отклика и декомпозиций дисперсий ошибок прогнозов проведем на основе оцененной модели коррекции ошибок для этих 3 переменных (значения всех переменных взяты в логарифмах уровней). При построении используем упорядочение $Income \rightarrow Adver \rightarrow Sales$.

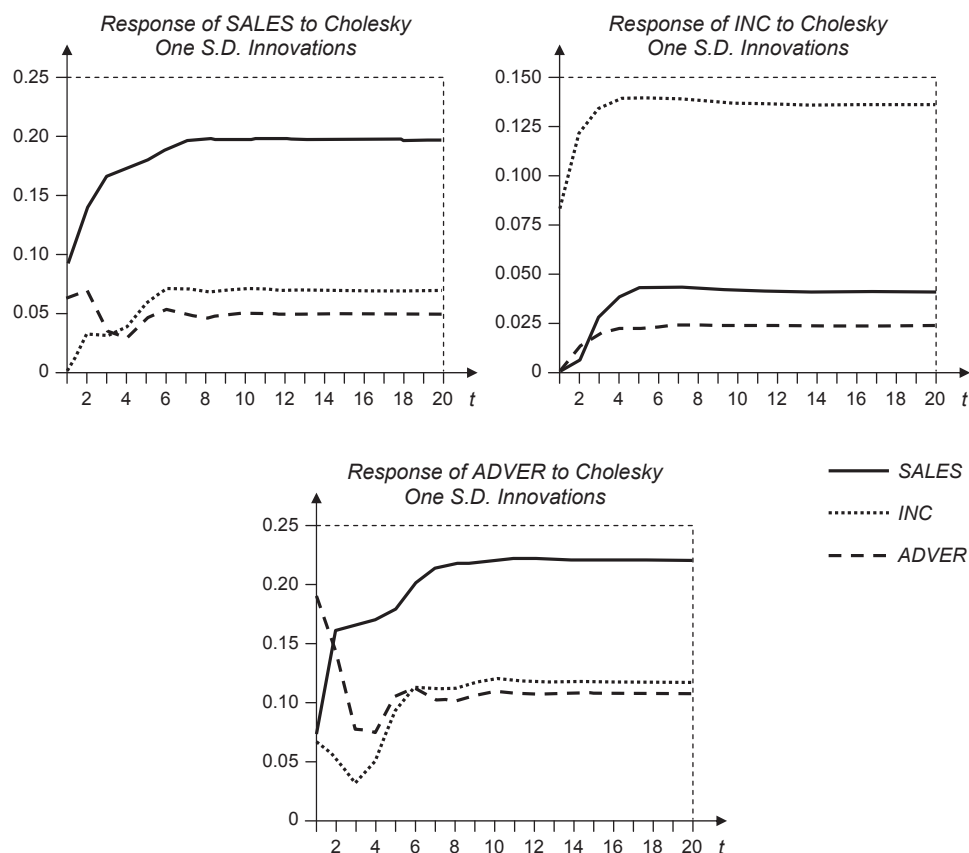


Рис. 6.45

Функции импульсного отклика изображены на рис. 6.45, а декомпозиции дисперсий ошибок прогнозов — на рис. 6.46. Отметим весьма высокий вклад в дисперсию ошибок прогнозов переменной *Adver* переменной *Sales* и достаточно малый вклад переменной *Adver* в дисперсию прогнозов переменной *Sales*. ■

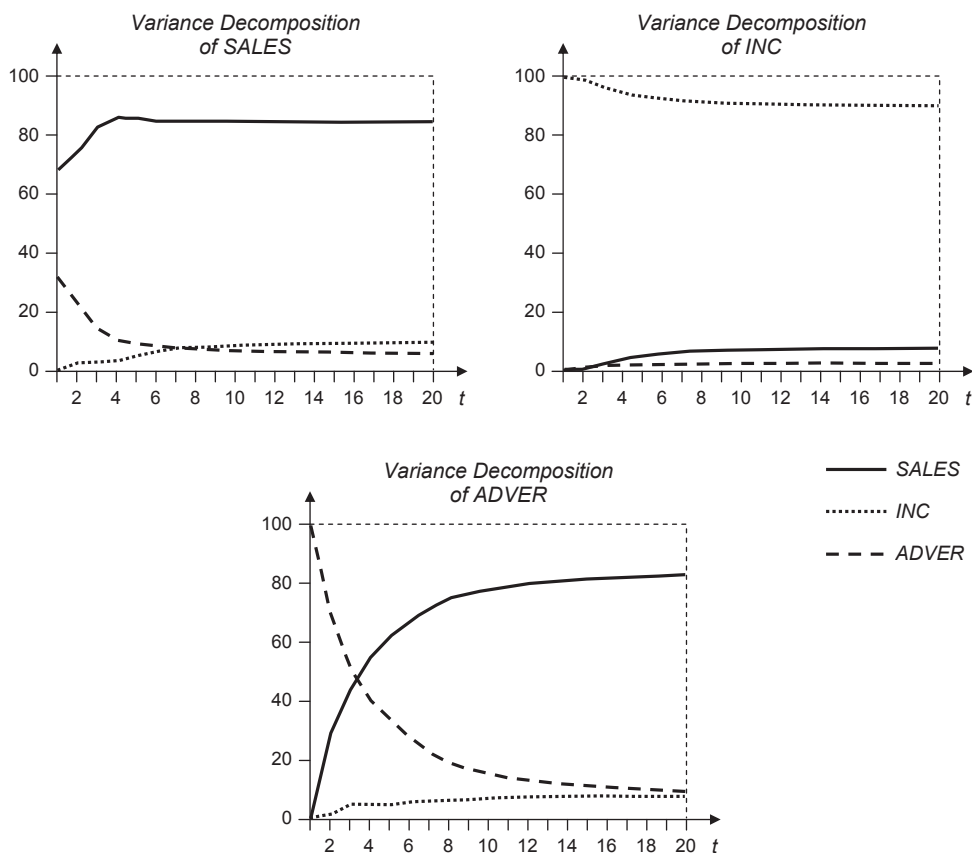


Рис. 6.46

Идентификация в коинтегрированных VAR

Рассмотрим для простоты коинтегрированную VAR первого порядка. Тогда соответствующая ей ЕСМ в приведенной форме имеет простой вид:

$$\Delta y_t = \Pi y_{t-1} + u_t,$$

где $\Pi = \alpha\beta^T$;

α — матрица коэффициентов при коинтегрирующих векторах в различных уравнениях;

β — матрица, образованная коинтегрирующими векторами (если их несколько).

Пусть опять

$$u_t = D\varepsilon_t,$$

где ε_t — вектор фундаментальных инноваций.

Пусть

$$y_t = (z_t, x_t)^T,$$

z_t, x_t — переменные — интегрированные первого порядка;

$(1, -1)^T$ — коинтегрирующий вектор, так что $z_t - x_t$ — стационарный ряд,

$$\alpha = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \Pi = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} (1 \quad -1) = \begin{pmatrix} \alpha_{11} & -\alpha_{11} \\ \alpha_{21} & -\alpha_{21} \end{pmatrix};$$

матрица Π имеет ранг 1. Соответственно ЕСМ тогда принимает вид:

$$\begin{pmatrix} \Delta z_t \\ \Delta x_t \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \end{pmatrix} (1, -1) \begin{pmatrix} z_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Последнее соотношение можно переписать в следующей форме, предложенной в работе (Mellander, Vredin, Warne, 1993):

$$\begin{aligned} \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (1-L) & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (z_t - x_t) \\ \Delta x_t \end{pmatrix} = \\ = \begin{pmatrix} \alpha_{11} & 0 \\ \alpha_{21} & 0 \end{pmatrix} \begin{pmatrix} (z_{t-1} - x_{t-1}) \\ \Delta x_{t-1} \end{pmatrix} + \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \end{aligned}$$

что, в свою очередь, можно записать в виде:

$$\left(\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (1-L) & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \alpha_{11}L & 0 \\ \alpha_{21}L & 0 \end{pmatrix} \right) \begin{pmatrix} (z_t - x_t) \\ \Delta x_t \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Долговременные отклики можно вычислить отсюда, полагая формально в стоящей первой слева матрице $L = 1$ и умножая обе части равенства на обратную к полученной матрице:

$$\begin{pmatrix} (z_t - x_t) \\ \Delta x_t \end{pmatrix} = \begin{pmatrix} -\alpha_{11} & 1 \\ -\alpha_{21} & 1 \end{pmatrix}^{-1} \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

что в ранее использованных обозначениях записывается как

$$\begin{pmatrix} (z_t - x_t) \\ \Delta x_t \end{pmatrix} = \begin{pmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Поскольку система коинтегрирована, имеющиеся фундаментальные шоки естественно идентифицировать, определяя один из них как перманентный, а другой — как преходящий. Первый шок связан с общим стохастическим трендом (*common trend*), а второй — с коинтеграционным соотношением. Если идентифицировать в качестве преходящего фундаментальный шок ε_{2t} , то надо наложить ограничение $\theta_{22}(1) = 0$. Поскольку же

$$\begin{aligned} \begin{pmatrix} -\alpha_{11} & 1 \\ -\alpha_{21} & 1 \end{pmatrix}^{-1} \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} &= \frac{1}{\alpha_{21} - \alpha_{11}} \begin{pmatrix} 1 & -1 \\ \alpha_{21} & -\alpha_{11} \end{pmatrix} \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} = \\ &= \frac{1}{\alpha_{21} - \alpha_{11}} \begin{pmatrix} d_{11} - d_{21} & d_{12} - d_{22} \\ \alpha_{21}d_{11} - \alpha_{11}d_{21} & \alpha_{21}d_{12} - \alpha_{11}d_{22} \end{pmatrix}, \end{aligned}$$

то это сводится к линейному ограничению:

$$\alpha_{21}d_{12} - \alpha_{11}d_{22} = 0.$$

Наличие такого ограничения обеспечивает идентификацию матрицы D , если параметры α_{21} , α_{11} известны (оценены при проведении коинтеграционного анализа). Заметим, что если $\alpha_{11} = 0$, то данный метод идентификации равносильен применению схемы Холецкого (в этом случае указанное линейное ограничение сводится к условию $d_{12} = 0$).

Изложенная здесь схема применялась в работе (Favero, Giavazzi, Spaventa, 1997). В ней рассматривался вопрос о том, чем объясняется наличие в некоторых странах Европы высокого спреда доходности государственных облигаций по отношению к доходности государственных облигаций Германии. При этом сам спред (SP_t) представлялся в виде суммы двух составляющих:

$$SP_t = ER_t + DR_t,$$

где ER_t — составляющая, связанная с риском изменения обменного курса валюты;

DR_t — составляющая, связанная с риском дефолта.

Соответственно здесь ECM имеет вид:

$$\Delta ER_t = \alpha_{11}(SP_{t-1} - ER_{t-1}) + u_{1t} = \alpha_{11}(SP_{t-1} - ER_{t-1}) + d_{11}\varepsilon_{1t} + d_{12}\varepsilon_{2t},$$

$$\Delta SP_t = \alpha_{21}(SP_{t-1} - ER_{t-1}) + u_{2t} = \alpha_{21}(SP_{t-1} - ER_{t-1}) + d_{21}\varepsilon_{1t} + d_{22}\varepsilon_{2t}.$$

Действуя, как было указано выше, получим соотношение для долгосрочного влияния шоков:

$$\begin{pmatrix} DR_t \\ \Delta SP_t \end{pmatrix} = \frac{1}{\alpha_{21} - \alpha_{11}} \begin{pmatrix} d_{11} - d_{21} & d_{12} - d_{22} \\ \alpha_{21}d_{11} - \alpha_{11}d_{21} & \alpha_{21}d_{12} - \alpha_{11}d_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

При проведении коинтеграционного анализа была проверена и не отвергнута гипотеза H_0 : $\alpha_{11} = 0$, так что идентификация перманентного и преходящего шоков обеспечивается наложением ограничения $d_{12} = 0$. При этом шок ε_{2t} является преходящим, а шок ε_{1t} перманентным. Перманентный шок управляет общим стохастическим трендом и интерпретируется как шок, носящий интернациональный характер. Преходящий шок не имеет долгосре-

менного влияния на спред, отражается только в краткосрочных флуктуациях вокруг общего стохастического тренда и интерпретируется как локальный, специфический для данной страны.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Законно ли использование инструментария, предназначенного для анализа стабильных моделей VAR, в случаях, когда условие стабильности не выполнено?
2. Как следует строить функции импульсного отклика в случае неинтегрированной VAR?
3. Как следует строить функции импульсного отклика в случае коинтегрированной VAR? Какие проблемы возникают при практическом построении этих функций по имеющимся статистическим данным?
4. Что можно предпринять в случае, когда векторный временной ряд не может быть представлен в виде векторного процесса авторегрессии (хотя бы и бесконечного порядка)?
5. Как соотносятся между собой причинность по Грейнджеру, функции импульсного отклика и декомпозиция дисперсий?
6. Как можно произвести идентификацию в коинтегрированной VAR?

Р а з д е л 7

ТЕСТЫ НА ЕДИНИЧНЫЕ КОРНИ И НЕЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ. ДИНАМИЧЕСКИЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Тема 7.1

ТЕСТЫ НА ЕДИНИЧНЫЕ КОРНИ И НЕЛИНЕЙНЫЕ ПРЕОБРАЗОВАНИЯ

Проверка на наличие единичных корней стала достаточно общей практикой при работе с временными рядами. В то же время следует уделять внимание вопросу о подходящем преобразовании переменных, которое должно предшествовать такой проверке.

В макроэкономике и эмпирических финансах (цены на акции, обменные курсы) часто имеются убедительные аргументы в пользу использования логарифмического преобразования переменных. Для других рядов (например, процентных ставок) автоматически предполагается спецификация в уровнях. Между тем во многих приложениях априори не ясно, где присутствует предполагаемый единичный корень — в ряде уровней или в ряде логарифмов, так что представляет интерес проверка гипотезы единичного корня в контексте не полностью специфицированного нелинейного преобразования данных.

Начнем рассмотрение с работы Грейнджера и Холлманна (*Granger, Hallmann, 1991*).

Базовым для теории распределений статистик различных критериев, используемых для проверки на наличие единичного корня, является результат Филлипса (*Phillips, 1987*) о предельном (асимптотическом) распределении статистики Дики — Фуллера в ситуации, когда процесс порождения данных имеет вид:

$$DGP: y_t = y_{t-1} + u_t, \quad y_0 = 0,$$

и оценивается статистическая модель:

$$SM: y_t = \rho y_{t-1} + u_t, \quad t = 1, \dots, T.$$

При выводе предельного распределения Филлипс опирается на следующие предположения относительно u_t :

- а) $E(u_t) = 0$;
- б) $\sup_t E|u_t|^\beta < \infty$ для некоторого $\beta > 2$;
- в) предел $\sigma^2 = \lim_{T \rightarrow \infty} E \left\{ \frac{1}{T} \left(\sum_{t=1}^T u_t \right)^2 \right\}$ существует, и его значение больше нуля;
- г) последовательность $\{u_t\}_1^\infty$ обладает свойством сильного перемешивания с коэффициентами α_m , удовлетворяющими соотношению $\sum_{m=1}^\infty \alpha_m^{1-\frac{2}{\beta}} < \infty$.

✓ **Замечание 7.1.1.** Пусть имеется какая-нибудь последовательность случайных величин $\{X_k, k \in Z\}$, $F_{-\infty}^j$ — множество всех событий, связанных со случайными величинами $\{X_k, k \leq j\}$, а F_{j+n}^∞ — множество всех событий, связанных со случайными величинами $\{X_k, k \leq j+n\}$. Для каждого $n \geq 1$ обозначим:

$$\alpha_n = \sup_{j \in Z} \sup_{A \in F_{-\infty}^j, B \in F_{j+n}^\infty} |P(A \cap B) - P(A)P(B)|.$$

Последовательность $\{X_k, k \in Z\}$ обладает свойством **сильного перемешивания** (*strong mixing*), если $\alpha_n \rightarrow 0$ при $n \rightarrow \infty$.

Условие а) — обычное. Условие б) означает равномерную ограниченность моментов $E|u_t|^\beta$ по t . В условии в) имеем:

$$\sigma^2 = \lim_{T \rightarrow \infty} T \left\{ E \left(\frac{1}{T} \sum_{t=1}^T u_t \right)^2 \right\} = \lim_{T \rightarrow \infty} T E(\bar{u}_T^2),$$

где σ^2 — **долговременная дисперсия** (*long-run variance*) ряда u_t .

Если u_t представляет собой процесс белого шума, то

$$T E(\bar{u}_T^2) = T E \left(\frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T u_t u_s \right) = T \frac{1}{T^2} \sum_{t=1}^T E(u_t^2) = \text{Var}(u_t),$$

так что в этом случае $\sigma^2 = \text{Var}(u_t)$ — долговременная дисперсия ряда равна его обычной дисперсии.

Если процесс u_t стационарный (в широком смысле) и последовательность его автоковариаций $\gamma_j = \text{Cov}(u_t, u_{t-j})$ абсолютно суммируема, т.е.

$$\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty,$$

тогда

$$\sigma^2 = \lim_{T \rightarrow \infty} T E(\bar{u}_T^2) = \sum_{j=-\infty}^{\infty} \gamma_j,$$

откуда вытекает, что $E(\bar{u}_T^2) \rightarrow 0$, т.е. $\bar{u}_T \rightarrow 0$ в среднем квадратичном. Последнее есть не что иное, как проявление закона больших чисел (среднее арифметическое сходится к математическому ожиданию).

В качестве примера пусть

$$u_t = \rho u_{t-1} + \varepsilon_t \text{ — стационарный ряд AR}(1), \quad |\rho| < 1, \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 > 0.$$

Математическое ожидание такого процесса равно нулю, его дисперсия равна:

$$\text{Var}(u_t) = \frac{\sigma_\varepsilon^2}{1 - \rho^2},$$

а j -я автоковариация равна:

$$\gamma_j = \sigma_\varepsilon^2 \frac{\rho^{|j|}}{1 - \rho^2},$$

так что

$$\begin{aligned} \sum_{j=-\infty}^{\infty} |\gamma_j| &= \frac{\sigma_\varepsilon^2}{1 - \rho^2} \sum_{j=-\infty}^{\infty} |\rho|^{|j|} = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \left(2 \sum_{j=0}^{\infty} |\rho|^j - 1 \right) = \\ &= \frac{\sigma_\varepsilon^2}{1 - \rho^2} \left(\frac{2}{1 - |\rho|} - 1 \right) = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \frac{1 + |\rho|}{1 - |\rho|} = \frac{\sigma_\varepsilon^2}{(1 - |\rho|)^2} < \infty. \end{aligned}$$

Долговременная дисперсия равна:

$$\sum_{j=-\infty}^{\infty} \gamma_j = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \sum_{j=-\infty}^{\infty} \rho^{|j|} = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \left(2 \sum_{j=0}^{\infty} \rho^j - 1 \right),$$

т.е.

$$\sigma^2 = \frac{\sigma_\varepsilon^2}{(1 - \rho)^2}$$

и отличается от $\sigma_u^2 = \text{Var}(u_t) = \frac{\sigma_\varepsilon^2}{1 - \rho^2}$, если $\rho \neq 0$.

Заметим, что именно вследствие подобного различия обычной и долговременной дисперсий приходится корректировать оценки стандартных ошибок оцененных коэффициентов в моделях регрессии с автокоррелированными ошибками — метод *Newey — West*.

Заметим еще, что в силу условия а) $E(y_t) = 0$ и в условии в) $\left(\sum_{t=1}^T u_t\right)^2 = y_T^2$,

так что

$$\sigma^2 = \lim_{T \rightarrow \infty} E \left\{ \frac{1}{T} y_T^2 \right\} = \lim_{T \rightarrow \infty} \frac{Var(y_T)}{T}.$$

Дополнительное условие г) вместе с остальными условиями обеспечивает выполнение **функциональной предельной теоремы** (*functional limit theorem*), гарантирующей сходимость функционалов от соответствующих случайных последовательностей к функционалам от стандартного винеровского процесса $W(t)$ (процесса с независимыми нормальными приращениями) при $T \rightarrow \infty$. Покажем, как получается при этих условиях предельное распределение статистики Дики — Фуллера в указанной паре DGP и SM .

Прежде всего заметим, что в этой ситуации оценка наименьших квадратов коэффициента ρ , полученная на основании T наблюдений, равна:

$$\hat{\rho}_T = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2},$$

так что

$$\hat{\rho}_T - 1 = \frac{\sum_{t=1}^T y_{t-1} u_t}{\sum_{t=1}^T y_{t-1}^2}.$$

Обозначим частичные суммы ряда u_t как

$$S_t = \sum_{i=1}^t u_i.$$

Для простоты положим $S_0 = 0$. Тогда $y_t = S_t + y_0$,

$$\begin{aligned} T^{-2} \sum_{t=1}^T y_{t-1}^2 &= T^{-2} \sum_{t=1}^T (S_{t-1} + y_0)^2 = \\ &= T^{-2} \sum_{t=1}^T (S_{t-1}^2 + 2y_0 S_{t-1} + y_0^2) = \\ &= \sigma^2 \sum_{t=1}^T \left(\frac{1}{\sigma \sqrt{T}} S_{t-1} \right)^2 \frac{1}{T} + 2y_0 \sigma T^{-\frac{1}{2}} \sum_{t=1}^T \left(\frac{1}{\sigma \sqrt{T}} S_{t-1} \right) \frac{1}{T} + \frac{1}{T} y_0^2. \end{aligned}$$

Последнее выражение можно представить в следующем виде:

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 = \sigma^2 \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} \left(\frac{1}{\sigma \sqrt{T}} S_{[Tr]} \right)^2 dr + 2y_0 \sigma T^{-\frac{1}{2}} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} \left(\frac{1}{\sigma \sqrt{T}} S_{[Tr]} \right) dr + \frac{1}{T} y_0^2,$$

где $[Tr]$ — целая часть числа Tr , $0 \leq r \leq 1$.

Заметим, что:

- при $r = \frac{t}{T}$: $[Tr] = \left[T \frac{t}{T} \right] = t$;
- при $r = \frac{t-1}{T}$: $[Tr] = \left[T \frac{t-1}{T} \right] = t-1$;
- при $\frac{t-1}{T} \leq r < \frac{t}{T}$: $S_{[Tr]} \equiv S_{t-1}$.

При этом:

$$X_T(r) = \frac{1}{\sigma \sqrt{T}} S_{[Tr]} \text{ — ступенчатая функция, непрерывная справа,}$$

и

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 = \sigma^2 \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} X_T^2(r) dr + 2y_0 \sigma T^{-\frac{1}{2}} \sum_{t=1}^T \int_{\frac{t-1}{T}}^{\frac{t}{T}} X_T(r) dr + \frac{1}{T} y_0^2.$$

Функциональная предельная теорема. Если выполнены условия а) — з), то при $T \rightarrow \infty$

$$X_T(r) \Rightarrow W(r),$$

где $W(r)$ — стандартное броуновское движение (винеровский процесс — *Wiener process*), т.е. случайный процесс с непрерывным временем, для которого:

- $W(0) = 0$;
- приращения $(W(r_2) - W(r_1)), \dots, (W(r_k) - W(r_{k-1}))$ независимы в совокупности, если $0 \leq r_1 < r_2 < \dots < r_k$;
- $W(s) - W(r) \sim N(0, s - r)$ при $s > r$;
- реализации $W(r)$ непрерывны с вероятностью 1.

Знак \Rightarrow обозначает здесь слабую сходимость, или сходимость по распределению.

Не будем вдаваться в детали определения слабой сходимости в пространствах реализаций случайных функций — для объяснения этого необходимо

привлечение весьма сложного математического аппарата. Для нас важны последствия наличия указанной сходимости.

Теорема о непрерывном отображении. Если $X_T(r) \Rightarrow W(r)$ и $h(\bullet)$ — непрерывный функционал на пространстве $D[0, 1]$ всех вещественнозначных функций на отрезке $[0, 1]$, которые непрерывны справа в каждой точке этого отрезка и имеют конечные пределы слева, то

$$h(X_T(r)) \xrightarrow{d} h(W(r)) \quad \text{при } T \rightarrow \infty.$$

Здесь \xrightarrow{d} означает сходимость функции распределения случайной величины, стоящей слева от стрелки, к функции распределения случайной величины, стоящей справа от стрелки (**сходимость по распределению** — *convergence in distribution*), во всех точках непрерывности последней.

Из этих двух теорем и полученного ранее выражения для $T^{-2} \sum_{t=1}^T y_{t-1}^2$ получаем: если выполнены условия а) — з), то при $T \rightarrow \infty$

$$T^{-2} \sum_{t=1}^T y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 (W(r))^2 dr.$$

Займемся теперь числителем дроби в правой части выражения

$$\hat{\rho}_T - 1 = \frac{\sum_{t=1}^T y_{t-1} u_t}{\sum_{t=1}^T y_{t-1}^2}.$$

Имеем:

$$\begin{aligned} T^{-1} \sum_{t=1}^T y_{t-1} u_t &= T^{-1} \sum_{t=1}^T (S_{t-1} + y_0) u_t = T^{-1} \sum_{t=1}^T S_{t-1} u_t + y_0 \bar{u} = \\ &= T^{-1} \sum_{t=1}^T \frac{1}{2} (S_t^2 - S_{t-1}^2 - u_t^2) + y_0 \bar{u} = (2T)^{-1} S_T^2 - (2T)^{-1} \sum_{t=1}^T u_t^2 + y_0 \bar{u} = \\ &= \frac{\sigma^2}{2} (X_T(1))^2 - (2T)^{-1} \sum_{t=1}^T u_t^2 + y_0 \bar{u} \xrightarrow{d} \frac{\sigma^2}{2} (W(1))^2 - \frac{\sigma_u^2}{2}, \end{aligned}$$

т.е.

$$T^{-1} \sum_{t=1}^T y_{t-1} u_t \xrightarrow{d} \frac{\sigma^2}{2} \left((W(1))^2 - \frac{\sigma_u^2}{\sigma^2} \right).$$

Отсюда и из соотношения $T^{-2} \sum_{t=1}^T y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 (W(r))^2 dr$, полученного ранее, находим:

$$T(\hat{\rho}_T - 1) = \frac{T^{-1} \sum_{t=1}^T y_{t-1} u_t}{T^{-2} \sum_{t=1}^T y_{t-1}^2} \xrightarrow{d} \frac{\frac{1}{2} \left((W(1))^2 - \frac{\sigma_u^2}{\sigma^2} \right)}{\int_0^1 (W(r))^2 dr}.$$

Заметим, что этот результат был получен ранее в работе (White, 1958), только в предположении, что $u_t \sim i.i.d. N(0, \sigma_u^2)$.

Обратимся теперь к проверке гипотезы $H_0: \rho = 1$ в рамках статистической модели $SM: y_t = \rho y_{t-1} + u_t, t = 1, \dots, T$. Используемая для проверки этой гипотезы t -статистика имеет вид:

$$t_\rho = \frac{\hat{\rho}_T - 1}{S \left(\sum_{t=1}^T y_{t-1}^2 \right)^{-\frac{1}{2}}} = \frac{\sum_{t=1}^T y_{t-1} u_t}{S \sqrt{\sum_{t=1}^T y_{t-1}^2}},$$

где $S^2 = \frac{1}{T-1} \sum_{t=1}^T (y_t - \hat{\rho}_T y_{t-1})^2$.

В соответствии с предыдущим при $T \rightarrow \infty$

$$t_\rho \xrightarrow{d} \frac{\left(\frac{\sigma}{2\sigma_u} \right) \left((W(1))^2 - \frac{\sigma_u^2}{\sigma^2} \right)}{\left(\int_0^1 (W(1))^2 dr \right)^{\frac{1}{2}}}.$$

В предположении, что $u_t \sim i.i.d. N(0, \sigma_u^2)$, для соответствующей t -статистики Дики — Фуллера получаем:

$$t_\rho \xrightarrow{d} \frac{(W(1))^2 - 1}{2 \left(\int_0^1 (W(r))^2 dr \right)^{\frac{1}{2}}}.$$

Такое же предельное распределение имеет статистика ADF -критерия, если в правую часть уравнения добавляется надлежащее количество запаздывающих разностей.

Предположим теперь, что $x_t = x_{t-1} + u_t, x_0 = 0, u_t \sim i.i.d. N(0, \sigma_u^2)$, так что выполнены все условия а) — в), и пусть $y_t = f(x_t)$. Тогда $y_t = y_{t-1} + \eta_t$, и если функция f дифференцируема, то $\eta_t = f(x_t) - f(x_{t-1}) = f'(x_{t-1} + r_t) \varepsilon_t, r_t \in [0, 1]$,

так что в общем случае нет никаких оснований полагать, что для ряда y_t все условия Филлипса выполнены.

Заметим, что в условии в):

$$\left(\sum_{t=1}^T u_t \right)^2 = y_T^2,$$

и рассмотрим некоторые конкретные нелинейные преобразования:

- преобразование $y_t = x_t^2$.

В этом случае $\eta_t = y_t - y_{t-1} = x_t^2 - x_{t-1}^2 = (x_{t-1} + \varepsilon_t)^2 - x_{t-1}^2 = 2x_{t-1}\varepsilon_t + \varepsilon_t^2$, так что нарушаются все 4 условия а) — с);

- преобразование $y_t = x_t^3$.

Здесь $\eta_t = x_t^3 - x_{t-1}^3 = (x_{t-1} + \varepsilon_t)^3 - x_{t-1}^3 = 3x_{t-1}^2\varepsilon_t + 3x_{t-1}\varepsilon_t^2 + \varepsilon_t^3$, и опять нарушены все 4 условия;

- преобразование $y_t = \text{sgn}(x_t)$.

Здесь $\text{sgn}(x_t) = \begin{cases} -1, & x_t < 0 \\ 0, & x_t = 0 \\ 1, & x_t > 0 \end{cases}$. В этом случае $\left(\sum_{t=1}^T \eta_t \right)^2 = y_T^2 = 1$, так что

$$\sigma^2 = 0;$$

- преобразование $y_t = \sin x_t$.

Грейнджер и Холлман (Granger, Hallman, 1988) показали, что в этом случае процесс y_t стационарный и что предел в условии в) равен 0;

- преобразование $y_t = \exp(x_t)$.

В этом случае:

$$\eta_t = e^{x_t} - e^{x_{t-1}} = e^{x_{t-1} + \varepsilon_t} - e^{x_{t-1}} = (e^{\varepsilon_t} - 1)y_{t-1},$$

$$E(\eta_t) = E((e^{\varepsilon_t} - 1)y_{t-1}) = E(e^{\varepsilon_t} - 1)E(y_{t-1}) = 0,$$

$$\begin{aligned} D(\eta_t) &= E(\eta_t^2) = E(e^{\varepsilon_t} - 1)^2 E(y_{t-1}^2) = E(e^{\varepsilon_t} - 1)^2 D(y_{t-1}) = \\ &= E(e^{\varepsilon_t} - 1)^2 (t-1)\sigma_\varepsilon^2, \end{aligned}$$

так что η_t имеет линейно возрастающую дисперсию, что нарушает условия б) и в).

Грейнджер и Холлман смоделировали 2000 случайных блужданий длины 200 и на этих данных получили эмпирические распределения статистик Дики — Фуллера (при оценивании уравнения с константой в правой части) при различных преобразованиях смоделированных рядов. В табл. 7.1 и 7.2 приведены

Таблица 7.1

Эмпирическое распределение статистики DF

Преобразование	Эмпирические квантили, %								
	1	5	10	25	50	75	90	95	99
x	-3.54	-2.90	-2.62	-2.15	-1.59	-1.05	-0.48	-0.06	0.68
x_2	-4.86	-3.74	-3.21	-2.46	-1.84	-1.15	-0.02	0.87	2.61
x_3	-4.78	-3.78	-3.33	-2.65	-1.99	-1.23	0.11	1.58	4.06
$ x $	-4.76	-3.70	-3.24	-2.60	-2.01	-1.40	-0.80	-0.34	0.48
$\text{sgn}(x)$	-14.25	-11.31	-8.37	-6.05	-4.58	-3.58	-2.67	-2.16	-1.45
$\sin x$	-8.50	-8.00	-7.82	-7.46	-7.07	-6.70	-6.34	-6.17	-5.75
$\exp(x)$	-36.06	-10.13	-8.68	-7.22	-6.03	-5.05	-4.04	-2.99	-1.60

Таблица 7.2

Эмпирическое распределение статистики $ADF(4)$

Преобразование	Эмпирические квантили, %								
	1	5	10	25	50	75	90	95	99
x	-3.5	-2.95	-2.64	-2.14	-1.57	-1.03	-0.40	-0.03	0.83
x_2	-4.23	-3.44	-3.06	-2.45	-1.82	-1.07	0.25	1.20	2.63
x_3	-3.95	-3.39	-3.04	-2.48	-1.87	-1.04	0.61	2.03	4.0
$ x $	-4.23	-3.35	-3.05	-2.47	-1.89	-1.29	-0.59	-0.12	0.78
$\text{sgn}(x)$	-10.92	-6.90	-6.08	-4.01	-2.82	-2.08	-1.56	-1.24	-0.53
$\sin x$	-6.70	-6.15	-5.89	-5.49	-5.06	-4.63	-4.27	-4.08	-3.71
$\exp(x)$	-39.10	-7.76	-5.39	-4.72	-4.07	-3.25	-1.88	2.29	10.50

оцененные 5%-е критические значения этих статистик. В качестве нулевой выступает гипотеза о том, что рассматриваемый ряд является интегрированным порядка 1, а в качестве альтернативной выступает гипотеза о том, что этот ряд стационарный. В таблице Фуллера 5%-е критическое значение равно -2.88 . При использовании этого значения гипотеза единичного корня всегда (правильно) отвергается для $\sin x_t$. В то же время эта гипотеза обычно отвергается и для рядов $\text{sgn}(x)$ и $\exp(x)$, которые было бы неправильным рассматривать как $I(0)$ -ряды. Более частое отклонение гипотезы единичного корня наблюдается и для рядов x_t^2 , x_t^3 , $|x|$.

Смоделируем реализацию случайного блуждания со сносом 0.001. График этой реализации показан на рис. 7.1. На рис. 7.2—7.4 приведены графики преобразованных рядов, построенных по этой реализации.



Рис. 7.1

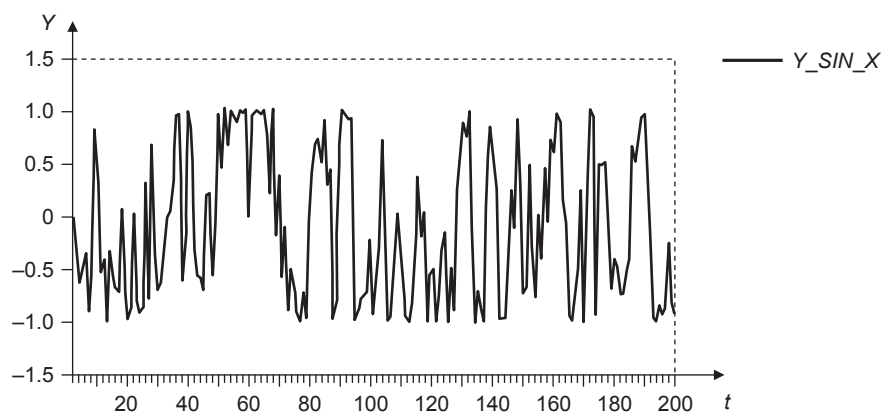


Рис. 7.2

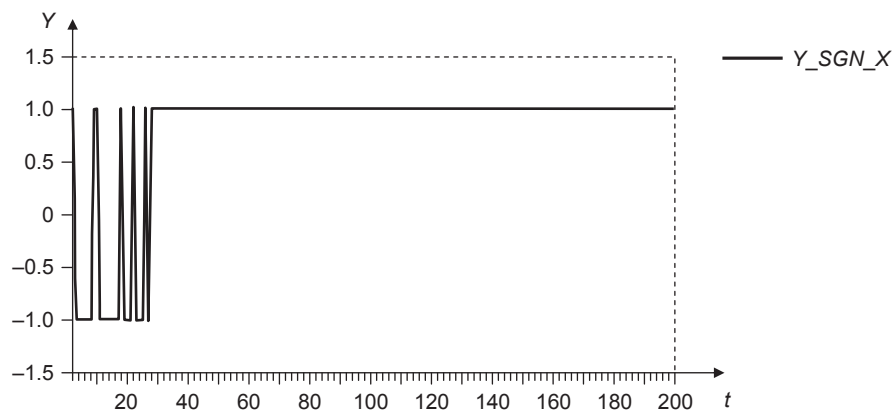


Рис. 7.3

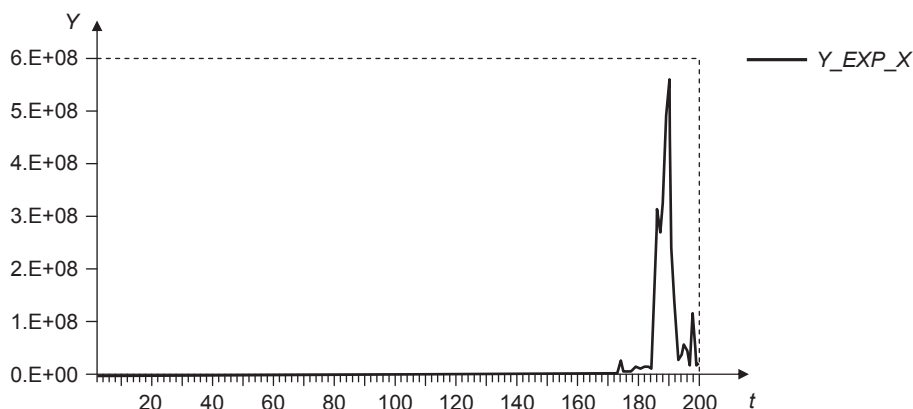


Рис. 7.4

Приведенные на рис. 7.5—7.9 коррелограммы первых 5 рядов напоминают коррелограммы рядов, имеющих единичный корень, а коррелограммы рядов $\sin x$ и $\exp(x)$ (рис. 7.10—7.11) имеют форму, характерную для стационарных рядов.

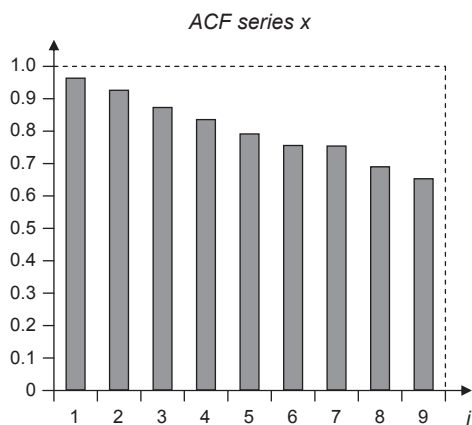


Рис. 7.5

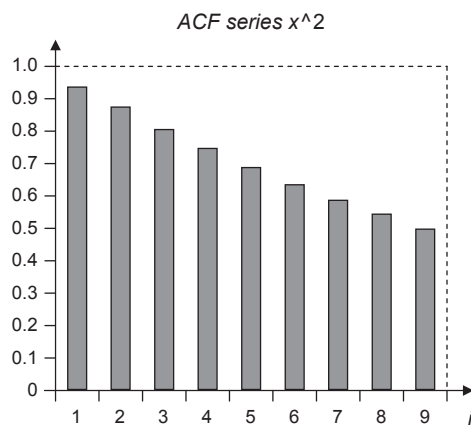


Рис. 7.6

В целом получается, что DF - и ADF -критерии более чувствительны к преобразованиям, чем автокорреляции, они отвергают гипотезу единичного корня в ситуациях, когда автокорреляции преобразованного ряда убывают медленно, по типу ряда с единичным корнем.

В работе (Corradi, 1995) дается объяснение наблюдаемому выше поведению автокорреляций нелинейно преобразованных рядов.

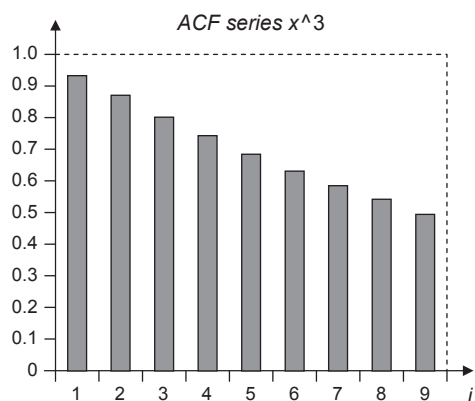


Рис. 7.7

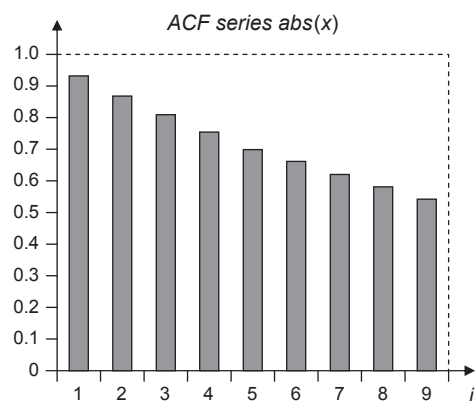


Рис. 7.8

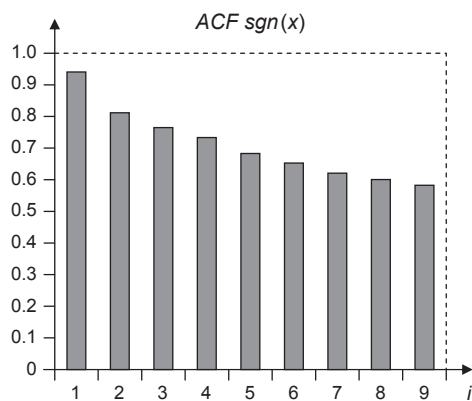


Рис. 7.9

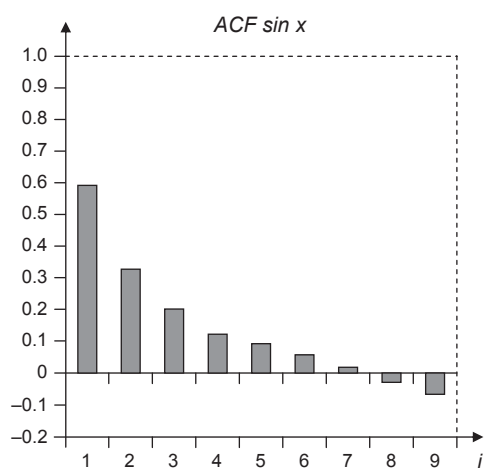


Рис. 7.10

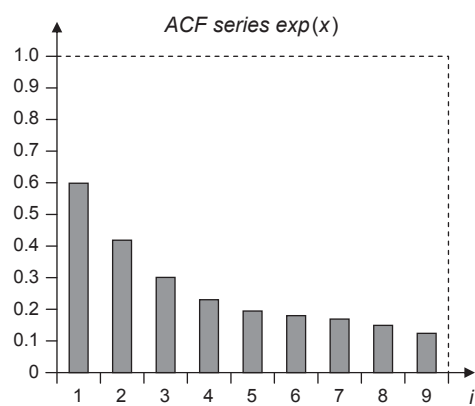


Рис. 7.11

Пусть в качестве *DGP* выступают следующие модели:

$$x_t = \sum_{i=1}^t \varepsilon_i, \quad (7.1)$$

$$x_t = \mu + \sum_{i=1}^t \varepsilon_i, \quad (7.2)$$

$$x_t = ct + \mu + \sum_{i=1}^t \varepsilon_i, \quad (7.3)$$

где $c > 0$ и $t = 1, 2, \dots, T$.

Таким образом,

- модель (7.1) есть случайное блуждание с $x_0 = 0$;
- модель (7.2) есть случайное блуждание с $x_0 = \mu$;
- модель (7.3) есть случайное блуждание с положительным сносом, т.е. с возрастающим детерминированным трендом.

Предполагается, что $\varepsilon_t \sim i.i.d.$ случайные величины с $E(\varepsilon_t) = 0$, $E(\varepsilon_t) = \sigma_\varepsilon^2$, $0 < \sigma_\varepsilon^2 < \infty$, имеющие плотность. Доказаны следующие результаты:

- пусть *DGP* имеет вид модели (7.1) или (7.2). Тогда для любой непрерывной, неограниченной, выпуклой (или вогнутой) функции f процесс $f(x_t)$ не обладает свойством сильного перемешивания;
- пусть *DGP* имеет вид модели (7.3). Тогда для любой непрерывной, строго возрастающей выпуклой функции f процесс $f(x_t)$ не обладает свойством сильного перемешивания.

Таким образом, в указанных ситуациях преобразованный ряд не обладает свойством сильного перемешивания, т.е. является процессом с достаточно долгой памятью, что обычно сопровождается медленным убыванием автокорреляций.

Эти выводы подтверждают большинство наблюдений Грейнджера и Холлмана в отношении коррелограмм. Единственное существенное отличие касается экспоненциального преобразования случайного блуждания. Грейнджер и Холлман, а также Эрмини и Грейнджер (*Ermini, Granger, 1993*) отмечают достаточно быстрое убывание автокорреляций такого ряда. Однако это не противоречит тому факту, что $\exp(x_t)$ не обладает сильным перемешиванием (объяснение — в (*Ermini, Granger, 1993*)).

Обладают ли первые разности преобразованных рядов свойством сильного перемешивания?

Если в качестве *DGP* выступает модель (7.1) или (7.2) и f — непрерывная, строго выпуклая (вогнутая) функция, то

$$f(x_t) = f(x_{t-1}) + \xi_t + (A_t - A_{t-1}),$$

где ξ_t обладает сильным перемешиванием, а разность $(A_t - A_{t-1})$ всегда больше нуля или всегда меньше нуля (в зависимости от того, выпукла или вогнута функция f) и вовсе не обязательно обладает сильным перемешиванием.

Если в качестве DGP выступает модель (7.3) и f — непрерывная, строго возрастающая, выпуклая функция, то

$$f(x_t) = f(x_{t-1}) + \xi_t + (A_t - A_{t-1}),$$

где ξ_t обладает сильным перемешиванием, а разность $(A_t - A_{t-1})$ всегда больше нуля и не обязательно обладает сильным перемешиванием.

Таким образом, первые разности преобразованных рядов не удовлетворяют обычным предположениям функциональной центральной предельной теоремы, поэтому асимптотические распределения статистик критериев Дики — Фуллера для преобразованных рядов перестают быть определенными.

В работе (Kramer, Davies, 2002) авторы, обращаясь к работе Грейнджера и Холлмана, выделяют ситуацию с экспоненциальным преобразованием случайного блуждания z_t . По результатам моделирования (Грейнджера и Холлмана), в этой ситуации DF -критерий, примененный непосредственно к ряду z_t , а к ряду $y_t = \exp(z_t)$, ошибочно отвергал нулевую гипотезу о наличии единичного корня более чем в 75% случаев.

Пусть

$$z_t = \delta + z_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, T,$$

где $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$ и $z_0 = \sigma^2 \cdot \text{constant}$.

Предметом анализа является поведение статистики $T(\hat{\rho}_T - 1)$. Табличные (асимптотические) критические значения этой статистики при гипотезе $H_0: \rho = 1$ равны -13.8 (1%), -8.1 (5%), -5.7 (10%).

Пусть значение T фиксировано. В статье показано следующее:

- если $\sigma^2 \rightarrow 0$, то $\hat{\rho}_T \rightarrow 1$ (по вероятности). Но тогда $T(\hat{\rho}_T - 1) \rightarrow 0$, так что вероятность отвержения нулевой гипотезы стремится к нулю при любом уровне значимости, меньшем 1;
- если $\sigma^2 \rightarrow \infty$ и критическое значение больше, чем $-T$, то вероятность отвержения нулевой гипотезы в этом случае стремится к

$$\begin{aligned} P\left(\max_{t=0, \dots, T-1} y_t^2 > y_T y_{T-1}\right) &= P\left(\max_{t=0, \dots, T-1} \exp(2z_t) > \exp(z_T + z_{T-1})\right) = \\ &= P\left(\max_{t=0, \dots, T-1} (2z_t) > (z_T + z_{T-1})\right). \end{aligned}$$

Последняя вероятность зависит от T и δ . В работе приведены наблюдаемые значения вероятности отвержения нулевой гипотезы при использовании 5%-го табличного критического значения, вычисленные моделированием.

Полученные результаты подтверждают тот факт, что вероятности отвержения нулевой гипотезы возрастают с увеличением дисперсии инноваций и объема выборки, что ведет к слишком частому отвержению нулевой гипотезы при использовании стандартных табличных критических значений. Эти вероятности оказываются меньшими 0.05 только при очень малых δ и при малых объемах выборки.

Авторы рассмотрели также ситуацию, когда единичный корень имеет ряд в уровнях, а проверка производится для ряда логарифмов. Иначе говоря, на этот раз критерий применяется к ряду $y_t = \ln(z_t)$. Разумеется, такая проверка предполагает, что $z_t > 0$ для всех $t = 1, 2, \dots, T$, а это может наблюдаться, только если $z_0 > 0$ достаточно велико и/или имеется значительный положительный снос. В такой ситуации рассматривается предельное поведение оценки:

$$\hat{\rho}_T(\sigma) = \frac{\sum_{t=1}^T \ln(\sigma z_{t-1}) \ln(\sigma z_t)}{\sum_{t=1}^T \ln^2(\sigma z_{t-1})}$$

при $\sigma^2 \rightarrow 0$ и $\sigma^2 \rightarrow \infty$. Показано, что в обоих случаях $\hat{\rho}_T(\sigma) \rightarrow 1$, так что условная вероятность отвержения нулевой гипотезы DF -критерием (при условии отсутствия неотрицательных значений z_t) стремится к нулю при применении DF -критерия к ряду логарифмов случайного блуждания. Иначе говоря, тест в логарифмах является чрезмерно консервативным.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Как определяется долговременная дисперсия временного ряда? Чему она равна в случае стационарного временного ряда?
2. Что представляет собой стандартное броуновское движение (винеровский процесс)?
3. В чем состоит функциональная предельная теорема?
4. В чем состоит теорема о непрерывном отображении?
5. Как используются функциональная предельная теорема и теорема о непрерывном отображении при выводе асимптотического распределения статистик для проверки гипотезы $H_0: \rho = 1$ в рамках статистической модели $SM: y_t = \rho y_{t-1} + u_t$, $t = 1, \dots, T$?
6. Как влияют нелинейные преобразования исходного ряда на выполнение условий Филлипса?
7. Как выглядят автокорреляционные функции нелинейно преобразованных рядов?
8. Что можно сказать о памяти нелинейно преобразованного процесса?
9. К каким результатам приводит ошибочное применение критерия единичного корня не к ряду уровней, а к ряду логарифмов уровней?
10. К каким результатам приводит ошибочное применение критерия единичного корня не к ряду логарифмов уровней, а к ряду уровней?

Тема 7.2

**ДИНАМИЧЕСКИЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ
ДЛЯ ОЦЕНИВАНИЯ КОИНТЕГРИРУЮЩЕГО ВЕКТОРА
СИСТЕМЫ ИНТЕГРИРОВАННЫХ РЯДОВ**

С момента появления исторической работы Энгла и Грейнджера (*Engle, Granger, 1987*) коинтегрирующие регрессии стали одним из стандартных инструментов анализа интегрированных $I(1)$ -переменных.

Если некоторые $I(1)$ -переменные коинтегрированы, то в рассмотренной во второй части учебника двухшаговой процедуре Энгла — Грейнджера на первом шаге производится оценивание коинтеграционного соотношения между этими переменными, дающее состоятельные оценки коэффициентов этого соотношения при наличии и автокоррелированности ошибок в оцениваемом уравнении, и коррелированности ошибок с регрессорами. Наличие такой коррелированности возникает, например, в модели

$$y_{1t} = \alpha + \beta y_{2t} + \varepsilon_{1t}, \quad \Delta y_{2t} = \varepsilon_{2t},$$

где y_{1t} и y_{2t} — $I(1)$ -переменные;

$$\varepsilon_{1t} \sim i.i.d. N(0, \sigma_1^2);$$

$$\varepsilon_{2t} \sim i.i.d. N(0, \sigma_2^2), \text{ если}$$

$$Cov(\varepsilon_{1t}, \varepsilon_{2s}) = \begin{cases} \sigma_{12} \neq 0, & t = s, \\ 0, & t \neq s. \end{cases}$$

В этом случае:

$$Cov(\varepsilon_{1t}, y_{2t}) = Cov(\varepsilon_{1t}, \varepsilon_{21} + \varepsilon_{22} + \dots + \varepsilon_{2t}) = \sigma_{12},$$

$$\begin{aligned} Corr(\varepsilon_{1t}, y_{2t}) &= \frac{\sigma_{12}}{\sqrt{D(\varepsilon_{1t})} \sqrt{D(\varepsilon_{21} + \varepsilon_{22} + \dots + \varepsilon_{2t})}} = \\ &= \frac{\sigma_{12}}{\sqrt{D(\varepsilon_{1t})} \sqrt{t D(\varepsilon_{2t})}} = \frac{\sigma_{12}}{\sigma_1 \sigma_2 \sqrt{t}}. \end{aligned}$$

Если выборка невелика, то оценки могут быть заметно смещены, и тем в большей степени, чем меньше объем выборки. Поскольку переменные нестационарны, оценки коэффициентов не являются асимптотически нормальными, что не позволяет использовать их для статистических выводов о коэффициентах коинтегрирующего вектора. Если система состоит из 3 рядов и более, она может иметь несколько линейно независимых коинтегрирующих векторов, и в этом случае возникают проблемы с экономической интерпретацией.

Указанные трудности привели к развитию альтернативных процедур, среди которых предложенная в работе (*Johansen, 1991*) и рассмотренная

во второй части учебника процедура, улучшающая *OLS* в нескольких направлениях:

- существование единственного коинтеграционного соотношения не предполагается априори, а тестируется;
- все регрессоры рассматриваются как эндогенные;
- можно оценить ранг коинтеграции;
- возможны оценивание коинтегрирующих векторов и проверка различных ограничений на коэффициенты этих векторов;
- возможна проверка выполнения ограничений на коэффициенты коинтегрирующих векторов;
- возможно построение моделей коррекции ошибок с несколькими коинтеграционными соотношениями.

Однако процедура Йохансена имеет и определенные недостатки:

- оценки могут иметь существенное смещение при малых выборках;
- неправильная спецификация даже одного из уравнений системы влияет на оценки коэффициентов во всех остальных уравнениях.

Указанные недостатки преодолеваются в процедуре динамического *OLS* (*DOLS* — *dynamic ordinary least squares*), улучшающей *OLS* и имеющей те же свойства оптимальности, что и процедура Йохансена. Эта процедура была разработана в работах (Phillips, Loretan, 1991), (Saikkonen, 1991), (Stock, Watson, 1993). Статистическое моделирование показывает (Carrion-i-Silvestre, Sansó-i-Rosselló, 2004), что в малых выборках *DOLS* работает лучше, чем еще одна модификация *OLS* (с теми же свойствами оптимальности) — *FM OLS* (*full modified OLS*) (см. (Phillips, Hansen, 1990)).

В чем состоит идея *DOLS*? Рассмотрим ее на нашем примере. В этом примере:

$$\varepsilon_{1t} \sim i.i.d. N(0, \sigma_1^2), \quad \varepsilon_{2t} \sim i.i.d. N(0, \sigma_2^2),$$

$$\text{Cov}(\varepsilon_{1t}, \varepsilon_{2s}) = \begin{cases} \sigma_{12} \neq 0, & t = s, \\ 0, & t \neq s, \end{cases}$$

так что

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim N(0, \Sigma), \quad \text{где } \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

Условное математическое ожидание случайной величины ε_{1t} относительно $\{\varepsilon_{2t}, t = 0, \pm 1, \pm 2, \dots\}$ здесь равно:

$$\begin{aligned} E[\varepsilon_{1t} | \{\varepsilon_{2t}, t = 0, \pm 1, \pm 2, \dots\}] &= E[\varepsilon_{1t} | \varepsilon_{2t}] = \\ &= E(\varepsilon_{1t}) + \frac{\sigma_{12}}{\sigma_2^2} (\varepsilon_{2t} - E(\varepsilon_{2t})) = \frac{\sigma_{12}}{\sigma_2^2} \varepsilon_{2t}. \end{aligned}$$

Если обозначить:

$$v_{1t} = \varepsilon_{1t} - E[\varepsilon_{1t} | \{\varepsilon_{2t}, t = 0, \pm 1, \pm 2, \dots\}],$$

$$\text{то } \text{Cov}(v_{1t}, \varepsilon_{2s}) = \text{Cov}\left(\varepsilon_{1t} - \frac{\sigma_{12}}{\sigma_2^2} \varepsilon_{2t}, \varepsilon_{2s}\right) = \text{Cov}(\varepsilon_{1t}, \varepsilon_{2s}) - \frac{\sigma_{12}}{\sigma_2^2} \text{Cov}(\varepsilon_{1t}, \varepsilon_{2s}) = 0,$$

как при $t \neq s$, так и при $t = s$.

Запишем исходное уравнение в виде:

$$y_{1t} = \alpha + \beta y_{2t} + \varepsilon_{1t} = \alpha + \beta y_{2t} + E[\varepsilon_{1t} | \{\varepsilon_{2t}, t = 0, \pm 1, \pm 2, \dots\}] + v_{1t},$$

$$\text{т.е.} \quad y_{1t} = \alpha + \beta y_{2t} + \frac{\sigma_{12}}{\sigma_2^2} \varepsilon_{2t} + v_{1t},$$

$$\text{или} \quad y_{1t} = \alpha + \beta y_{2t} + \frac{\sigma_{12}}{\sigma_2^2} \Delta y_{2t} + v_{1t},$$

где $\Delta y_{2t} = \varepsilon_{2t}$, и теперь уже

$$\text{Cov}(v_{1t}, y_{2t}) = \text{Cov}(v_{1t}, \varepsilon_{21} + \varepsilon_{22} + \dots + \varepsilon_{2t}) = 0.$$

Таким образом, в этой простейшей ситуации для предупреждения смещения *OLS*-оценки коэффициента β достаточно дополнить правую часть уравнения приращением объясняющей переменной y_{2t} .

В общем случае, когда оценивается коинтеграционное уравнение

$$y_{1t} = \theta_1 + \theta_2 y_{2t} + \dots + \theta_N y_{Nt} + \varepsilon_{1t}, \quad \Delta y_{kt} = \varepsilon_{kt}, \quad k = 2, \dots, N,$$

где $\varepsilon_{kt} \sim i.i.d. N(0, \sigma_k^2)$, в правую часть приходится добавлять не только текущие, но и запаздывающие (*lags*) и опережающие (*leads*) приращения переменных y_{2t}, \dots, y_{Nt} :

$$y_{1t} = \theta_1 + \theta_2 y_{2t} + \dots + \theta_N y_{Nt} + \sum_{j=-\infty}^{\infty} (\gamma_{2j} \Delta y_{2,t-j} + \dots + \gamma_{Nj} \Delta y_{N,t-j}) + v_{1t}.$$

При этом для проверки гипотез о коэффициентах можно использовать стандартные процедуры, основанные на *t*- и *F*-статистиках (*DOLS*-оценки асимптотически нормальны).

Разумеется, реально используется усечение бесконечной суммы в правой части, так что оценивается модель:

$$y_{1t} = \theta_1 + \theta_2 y_{2t} + \dots + \theta_N y_{Nt} + \sum_{j=-K}^K (\gamma_{2j} \Delta y_{2,t-j} + \dots + \gamma_{Nj} \Delta y_{N,t-j}) + v_{1t}.$$

Значение K должно быть достаточно большим для того, чтобы ликвидировать или свести к минимуму коррелированность объясняющих переменных с v_{1t} , но и не излишне большим, поскольку это ведет к ухудшению эффективности оценок.

Нетрудно заметить, что такая процедура уже рассматривалась в разд. 11 ч. 2, где она упоминалась как метод «*leads and lags*». Наименование «*DOLS*» связывают с работой (Stock, Watson, 1993), в которой методика «*leads and lags*» распространялась на коинтегрирующие регрессии с $I(1)$ - и $I(2)$ -переменными. Кратко обсудим такие ситуации чуть ниже.

В примере 11.2.2 разд. 11 ч. 2 была сгенерирована реализация модели:

$$y_t = 5 + z_t + u_t,$$

$$\Delta z_t = \varepsilon_t,$$

$$u_t = \varepsilon_t + 0.65\varepsilon_{t-1} + 0.65\varepsilon_{t+1} + 0.55\varepsilon_{t-2} + 0.55\varepsilon_{t+2} + v_t,$$

где $\varepsilon_t \sim i.i.d. N(0, 1)$;

$v_t \sim i.i.d. N(0, 0.1^2)$,

и оценено уравнение $y_t = \alpha + \beta z_t + u_t$ методом *OLS* по первым 30 наблюдениям.

При этом были получены следующие оценки: $\hat{\alpha} = 4.718$, $\hat{\beta} = 1.235$. Затем правую часть коинтеграционного уравнения мы дополнили текущим, двумя запаздывающими и двумя опережающими приращениями переменной z_t . При этом были получены следующие оценки наименьших квадратов для постоянной и для коэффициента при переменной z_t : $\hat{\alpha} = 4.995$, $\hat{\beta} = 1.0004$. Значения полученных *DOLS*-оценок намного ближе к значениям α и β в модели, порождающей данные.

Сгенерируем теперь реализацию модели $y_t = 5 + z_t + u_t$, где опять $\Delta z_t = \varepsilon_t$,

$$u_t = \varepsilon_t + 0.65\varepsilon_{t-1} + 0.65\varepsilon_{t+1} + 0.55\varepsilon_{t-2} + 0.55\varepsilon_{t+2} + v_t, \quad \varepsilon_t \sim i.i.d. N(0, 1),$$

но на сей раз v_t является стационарным $AR(2)$ -рядом:

$$v_t = 1.3v_{t-1} - 0.4v_{t-2} + \xi_t, \quad \xi_t \sim i.i.d. N(0, 0.1^2).$$

Оценивая уравнение $y_t = \alpha + \beta z_t + u_t$ методом *OLS* по первым 30 наблюдениям, получаем: $\hat{\alpha} = 4.669$, $\hat{\beta} = 1.210$.

Для рядов Δz_t и \hat{u}_t (\hat{u}_t — ряд остатков) часть кросс-коррелограммы имеет вид

<i>RES_OLS, Z_DIF(-i)</i>	<i>RES_OLS, Z_DIF(+i)</i>	<i>i</i>	<i>lag</i>	<i>lead</i>
****	****	0	0.4027	0.4027
****	****	1	0.3853	0.4343
**	***	2	0.2529	0.2802
**	*	3	-0.1517	-0.0670
**	*	4	-0.2302	-0.0530
*		5	-0.1328	0.0020
**	*	10	-0.2370	-0.0514

В соответствии с кросс-коррелограммой опять дополняем правую часть коинтеграционного уравнения текущим, двумя запаздывающими и двумя опережающими приращениями переменной z_t :

$$y_t = \alpha + \beta z_t + \sum_{j=-2}^2 \gamma_j \Delta z_{t-j} + v_t.$$

Оценивая это уравнение методом наименьших квадратов, получаем: $\hat{\alpha} = 4.908$, $\hat{\beta} = 0.958$.

Здесь в результате оценивания расширенного уравнения сохраняется автокоррелированность в остатках, что можно учесть двумя способами.

Во-первых, для оценивания расширенного уравнения можно применить доступный вариант *GLS* или процедуру Кохрейна — Оркатта: этот вариант Сток и Уотсон называют **динамическим обобщенным методом наименьших квадратов** (*DGLS* — *dynamic generalized least squares*). Если, как в нашем примере, v_t имеет структуру $AR(2)$, т.е.

$$v_t = a_1 v_{t-1} + a_2 v_{t-2} + \xi_t,$$

или

$$a(L)v_t = \xi_t, \quad \text{где} \quad a(L) = 1 - a_1 L - a_2 L^2,$$

то надо взять ряд остатков \hat{v}_t , полученных при оценивании расширенного уравнения (*DOLS*), и оценить уравнение

$$\hat{v}_t = a_1 \hat{v}_{t-1} + a_2 \hat{v}_{t-2} + \eta_t.$$

Полученные оценки \hat{a}_1 и \hat{a}_2 используются для оценивания $a(L)$:

$$\hat{a}(L) = 1 - \hat{a}_1 L - \hat{a}_2 L^2.$$

После этого производится авторегрессионное преобразование объясняющей и объясняющих переменных в расширенном уравнении, что приводит к преобразованному уравнению:

$$\hat{a}(L)y_t = \hat{a}(L)\alpha + \beta \hat{a}(L)z_t + \sum_{j=-2}^2 \gamma_j \hat{a}(L)\Delta z_{t-j} + \eta_t.$$

В рамках преобразованного уравнения можно уже на законных основаниях проверять гипотезы о коэффициентах коинтеграционного соотношения, используя для этой цели t - и F -статистики.

Применение указанного подхода в рамках процедуры, реализуемой в *EViews* (с включением в спецификацию уравнения $AR(1)$ и $AR(2)$), приводит к следующим результатам:

$$\hat{\alpha} = 4.955, \quad \hat{\beta} = 0.955.$$

Проверим гипотезу $H_0: \alpha = 5, \beta = 1$. Как следует из результатов проверки (табл. 7.3), эта гипотеза не отвергается.

Таблица 7.3

Проверка гипотезы о коэффициентах коинтеграционного соотношения

Wald Test

<i>Null Hypothesis:</i>	C(1) = 1 C(2) = 5		
<i>F-statistic</i>	1.661326	<i>Probability</i>	0.221017
<i>Chi-square</i>	3.322652	<i>Probability</i>	0.189887

Если динамика ряда v_t определяется моделью AR(p):

$$v_t = a_1 v_{t-1} + a_2 v_{t-2} + \dots + a_p v_{t-p} + \xi_t, \quad (7.4)$$

то можно не изменять спецификацию расширенного уравнения, производя авторегрессионное преобразование переменных, а вместо этого скорректировать надлежащим образом значения t - и F -статистик для проверки гипотез о значениях коэффициентов.

При таком подходе в выражениях для t - и F -статистик несмещенная оценка S^2 дисперсии ошибок, вычисляемая по остаткам \hat{v}_t в расширенном уравнении, заменяется оценкой долговременной дисперсии ряда v_t , вычисляемой по формуле:

$$\lambda^2 = \frac{\sigma_\xi^2}{(1 - a_1 - \dots - a_p)^2}.$$

Для этого оценивается уравнение (7.4), находятся оценки $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_p$ и ряд остатков $\hat{\xi}_t$, и в качестве оценки для λ^2 берется

$$\hat{\lambda}^2 = \frac{\hat{\sigma}_\xi^2}{(1 - \hat{a}_1 - \dots - \hat{a}_p)^2},$$

где $\hat{\sigma}_\xi^2 = \frac{1}{n-p} \sum_{t=p+1}^n \hat{\xi}_t^2$ — оценка дисперсии ряда ξ_t .

При проверке гипотезы $H_0: \beta = \beta^0$ стандартное выражение для t -статистики имеет вид

$$t = \frac{\beta - \beta^0}{S \sqrt{(X^T X)^{-1}_{22}}}.$$

Скорректированное значение равно:

$$t_{adj} = \frac{\beta - \beta^0}{\hat{\lambda} \sqrt{(X^T X)^{-1}_{22}}} = t \frac{S}{\hat{\lambda}}.$$

Аналогично скорректированное значение F -статистики равно

$$F_{adj} = F \frac{S^2}{\hat{\lambda}^2}.$$

В нашем примере оценивание уравнения $\hat{v}_t = a_1 \hat{v}_{t-1} + a_2 \hat{v}_{t-2} + \eta_t$, дает результаты, приведенные в табл. 7.4. Соответственно отсюда находим:

$$\hat{\lambda} = \frac{0.062628}{(1 - 0.993306 + 0.243001)} = 0.2508, \quad \hat{\lambda}^2 = 0.06291.$$

Таблица 7.4

Оценивание динамическим обобщенным методом наименьших квадратов
Dependent Variable: RES_DOLS

Переменная	Коэффициент	Стандартная ошибка	t -статистика	P -значение
AR(1)	0.993306	0.213655	4.649099	0.0001
AR(2)	-0.243001	0.213717	-1.137023	0.2672
<i>R-squared</i>	0.638986	<i>Mean dependent var</i>		0.001909
<i>Adjusted R-squared</i>	0.623289	<i>S.D. dependent var</i>		0.102039
<i>S.E. of regression</i>	0.062628	<i>Akaike info criterion</i>		-2.626581
<i>Sum squared resid</i>	0.090213	<i>Schwarz criterion</i>		-2.529071
<i>Log likelihood</i>	34.832260	<i>Durbin-Watson stat</i>		1.702883
<i>Inverted AR Roots</i>	0.56	0.44		

Скорректированное значение t -статистики для проверки гипотезы $H_0: \beta = 1$ равно

$$t_{adj} = \left(\frac{S}{\hat{\lambda}} \right) \frac{0.957998 - 1}{0.011395} = \frac{0.113386}{0.2508} (-3.686) = -1.666,$$

при этом гипотеза $\beta = 1$ не отвергается.

При проверке в уравнении $DOLS$ гипотезы $H_0: \alpha = 5, \beta = 1$ вычисленное значение F -статистики равно 6.826679, что дает P -значение, равное 0.005496, — эта гипотеза отвергается.

Скорректированное значение F -статистики равно:

$$F_{adj} = 6.826679 \left(\frac{0.113386}{0.2508} \right)^2 = 6.826679 \cdot 0.20439 = 1.3953;$$

ему соответствует P -значение 0.271, при этом гипотеза $H_0: \alpha = 5, \beta = 1$ уже не отвергается.

✓ **Замечание 7.2.1.** Если u_{1t} не является причиной по Грейнджеру для u_{2t} , то в правой части уравнения *DOLS* достаточно оставить только текущее и запаздывающие значения приращений:

$$y_{1t} = \theta_1 + \theta_2 y_{2t} + \dots + \theta_N y_{Nt} + \sum_{j=0}^K (\gamma_{2j} \Delta y_{2,t-j} + \dots + \gamma_{Nj} \Delta y_{N,t-j}) + v_{1t}.$$

В работе (Stock, Watson, 1993) метод *DOLS* применяется для исследования долговременного спроса на деньги в США. Используемые переменные:

m_t — $\log M1$;

gnp_t — $\log GNP$;

r_t — аннуализированная процентная ставка по коммерческим бумагам;

p_t — \log дефлятора *GNP*.

Рассматривались 3 спецификации, в каждой из которых r_t трактовалась как переменная порядка $I(1)$, а разность $(m_t - \theta_p p_t - \theta_{gnp} gnp_t - \theta_r r_t)$ — как $I(0)$ -переменная.

В качестве *первой спецификации* рассматривалась модель, в которой gnp_t , r_t , p_t — некоинтегрированные $I(1)$ -переменные.

Тогда имеем коинтегрированную систему $I(1)$ -рядов (m_t, p_t, gnp_t, r_t) . В соответствии со сказанным выше оценивается модель

$$m_t = \mu_1 + \theta_p p_t + \theta_{gnp} gnp_t + \theta_r r_t + \sum_{j=-K}^K (\gamma_{pj} \Delta p_{t-j} + \gamma_{gnp,j} \Delta gnp_{t-j} + \gamma_{rj} \Delta r_{t-j}) + v_t.$$

Значение K выбиралось равным 2 и 3, использовались как *DOLS*, так и *GLS*. Оценки коэффициентов изменялись в следующих пределах:

для θ_p : 0.997—1.159 (стандартные ошибки в пределах 0.159—0.234);

для θ_{gnp} : 0.685—0.890 (стандартные ошибки в пределах 0.133—0.237);

для θ_r : -0.122—-0.034 (стандартные ошибки в пределах 0.015—0.017).

Предыдущим рассмотрением можно было бы и ограничиться, если бы не неясность с порядками интегрированности переменных m_t и p_t : они имеют то ли порядок 1, то ли порядок 2. В работе (Stock, Watson, 1993) метод «*leads and lags*» был обобщен на ситуации с наличием переменных разных порядков интегрированности.

Пусть максимальный порядок интегрированности компонент N -мерного временного ряда y_t равен d . Предполагается, что ряд имеет представление:

$$\Delta^d y_t = \mu + F(L) \varepsilon_t,$$

где $\Delta^d = (1 - L)^d$, компоненты ε_t не коррелированы между собой;

$$F(L) = \sum_{j=0}^{\infty} F_j L^j, \quad \text{rank } F(1) = k_1, \quad 0 < k_1 < N.$$

(Последнее означает существование по меньшей мере $(N - k_1)$ коинтегрирующих векторов в системе.) Мы рассмотрим здесь только спецификации с порядками интегрированности не более двух.

Общая $I(2)$ модель имеет вид:

$$\Delta^2 y_t^1 = \mu_{1,0} + u_t^1;$$

$$\Delta y_t^2 = \mu_{2,0} + \mu_{2,1}t + \theta_{2,1}^1 \Delta y_t^1 + u_t^2;$$

$$y_t^3 = \mu_{3,0} + \mu_{3,1}t + \mu_{3,2}t^2 + \theta_{3,1}^1 \Delta y_t^1 + \theta_{3,1}^0 y_t^1 + \theta_{3,2}^0 \Delta y_t^2 + u_t^3.$$

Здесь

$$y_t = (y_t^1, y_t^2, y_t^3)^T,$$

вектор y_t^j имеет размер $(k_j \times 1)$, $i = 1, 2, 3$, $u_t = \sum_{j=0}^{\infty} H_j \varepsilon_{t-j}$.

Для простоты изложения коэффициенты при детерминированных составляющих полагаем равными 0, так что рассматривается представление:

$$\text{а) } \Delta^2 y_t^1 = u_t^1;$$

$$\text{б) } \Delta y_t^2 = \theta_{2,1}^1 \Delta y_t^1 + u_t^2;$$

$$\text{в) } y_t^3 = \theta_{3,1}^1 \Delta y_t^1 + \theta_{3,1}^0 y_t^1 + \theta_{3,2}^0 \Delta y_t^2 + u_t^3.$$

Некоторые строки у θ могут быть равными 0 или сами θ равны 0. Второй блок может отсутствовать вовсе. Рассматриваются частные случаи с $k_2 = 0$ или $k_2 = 1$, на базе которых можно анализировать более общие ситуации.

Случай 1. $k_2 = 0$. В этом случае блок б) отсутствует в системе и y_t^2 не входит в в):

$$y_t^3 = \theta_{3,1}^1 \Delta y_t^1 + \theta_{3,1}^0 y_t^1 + u_t^3.$$

Элементы y_t^3 , соответствующие нулевым строкам и в $\theta_{3,1}^0$, и в $\theta_{3,1}^1$, являются $I(0)$ -переменными, и эти переменные не входят ни в какие коинтеграционные соотношения. Элементы y_t^3 , соответствующие нулевым строкам в $\theta_{3,1}^0$, но ненулевым строкам в $\theta_{3,1}^1$, являются $I(1)$ -переменными. Элементы y_t^3 , соответствующие ненулевым строкам в $\theta_{3,1}^0$, являются $I(2)$ -переменными.

В этом случае $DOLS$ и $DGLS$ для $(\theta_{3,1}^1, \theta_{3,1}^0)$ асимптотически эффективны, и для проверки линейных гипотез об их значениях можно использовать хи-квадрат критерий Вальда.

С л у ч а й 2. $k_1 = 1$, $\Delta y_t^2 = \theta_{2,1}^1 \Delta y_t^1 + u_t^2$, $\theta_{2,1}^1$ известно. Тогда оценивается следующее уравнение:

$$y_t^3 = \theta_{3,1}^1 \Delta y_t^1 + \theta_{3,1}^0 y_t^1 + \theta_{3,2}^0 y_t^2 + d_{31}(L) \Delta^2 y_t^1 + d_{32}(L) (\Delta y_t^2 - \theta_{2,1}^1 \Delta y_t^1) + v_t^3.$$

Этому случаю соответствует *вторая спецификация модели*, в которой m_t , $p_t \sim I(2)$ и r_t , Δp_t не коинтегрированы между собой.

Тогда имеем $I(2)$ -систему, в которой

$$y_t^1 = p_t, \quad y_t^2 = \begin{pmatrix} gnp_t \\ r_t \end{pmatrix}, \quad y_t^3 = m_t,$$

$$\theta_{2,1}^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \theta_{3,1}^1 = 0, \quad \theta_{3,1}^0 = \theta_p, \quad \theta_{3,2}^0 = (\theta_{gnp} \quad \theta_r).$$

Иначе говоря:

$$\text{а) } \Delta^2 p_t = u_t^1;$$

$$\text{б) } \Delta \begin{pmatrix} gnp_t \\ r_t \end{pmatrix} = u_t^2;$$

$$\text{в) } m_t = \theta_p p_t + \theta_{gnp} gnp_t + \theta_r r_t + u_t^3.$$

В этом случае оценивается уравнение:

$$m_t = \mu + \theta_p p_t + \theta_{gnp} gnp_t + \theta_r r_t + \\ + \sum_{j=-K}^K (\gamma_{pj} \Delta^2 p_{t-j} + \gamma_{gnp,j} \Delta gnp_{t-j} + \gamma_{rj} \Delta r_{t-j}) + v_t.$$

Значение K выбиралось равным 2 и 3; использовались как *DOLS*, так и *DGLS*. Оценки коэффициентов изменялись в следующих пределах:

для θ_p : 1.022—1.277 (стандартные ошибки в пределах 0.205—0.290);

для θ_{gnp} : 0.723—0.841 (стандартные ошибки в пределах 0.208—0.265);

для θ_r : -0.125—-0.032 (стандартные ошибки в пределах 0.016—0.023).

Третья спецификация:

$$p_t \sim I(2), \quad \text{реальная ставка } (r_t - \Delta p_t) \sim I(0).$$

Тогда имеем $I(2)$ -систему, в которой

$$y_t^1 = p_t, \quad y_t^2 = \begin{pmatrix} gnp_t \\ \Delta^{-1} r_t \end{pmatrix}, \quad y_t^3 = m_t,$$

$$\theta_{2,1}^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \theta_{3,1}^1 = \theta_r, \quad \theta_{3,1}^0 = \theta_p, \quad \theta_{3,2}^0 = (\theta_{gnp} \quad 0).$$

Иначе говоря,

$$\text{а) } \Delta^2 p_t = u_t^1;$$

$$\text{б) } \Delta \begin{pmatrix} gnp_t \\ \Delta^{-1} r_t \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Delta p + u_t^2;$$

$$\text{в) } m_t = \theta_r \Delta p_t + \theta_p p_t + \theta_{gnp} gnp_t + u_t^3.$$

В этом случае оценивается уравнение:

$$m_t = \mu + \theta_p p_t + \theta_{gnp} gnp_t + \theta_r r_t + \\ + \sum_{j=-K}^K (\gamma_{pj} \Delta^2 p_{t-j} + \gamma_{gnp,j} \Delta gnp_{t-j} + \gamma_{rj} (r_t - \Delta p_t)) + v_t.$$

Значение K выбиралось равным 2 и 3; использовались как *DOLS*, так и *GLS*.
Оценки коэффициентов изменялись в следующих пределах:

для θ_p : 0.854—1.087 (стандартные ошибки в пределах 0.141—0.217);

для θ_{gnp} : 0.671—0.922 (стандартные ошибки в пределах 0.141—0.217);

для θ_r : -0.198— -0.002 (стандартные ошибки в пределах 0.013—0.017).

Итоги рассмотрения трех спецификаций:

- во всех спецификациях оценки коэффициента θ_p не отличались от 1 на 10%-м уровне значимости;
- во всех случаях оценки коэффициента θ_{gnp} не отличались от 1 на 10%-м уровне значимости;
- в ряде случаев оценки коэффициента θ_r весьма неточны.

Р а з д е л 8

МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ

Тема 8.1

МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ ДЛЯ ПЕРЕКРЕСТНОЙ ВЫБОРКИ

Понятия стохастической производственной функции и стохастической границы производственных возможностей были предложены в конце 1970-х гг. в работах (*Aigner, Lovell, Schmidt, 1977*) и (*Broeck, Førsund, Hjalmarsson, Meeusen, 1980*).

Теоретически производственная функция выражает максимальный объем выпуска, который можно получить для заданного набора входных переменных при фиксированной технологии производства:

$$q = f(z, \beta),$$

где q — объем выпуска (*output*);

z — вектор значений входных переменных;

β — вектор параметров.

Пусть имеются перекрестные (*cross-section*) статистические данные по N фирмам, так что вектору z_i значений входных переменных для фирмы i соответствует объем выпуска q_i . Для оценивания неизвестного вектора параметров по таким данным в течение довольно многих лет использовался метод наименьших квадратов, согласно которому в качестве оценки вектора β берется вектор $\hat{\beta}$, минимизирующий по всем возможным β^* сумму квадратов

$$\sum_{i=1}^N (q_i - f(z_i, \beta^*))^2.$$

Но такой подход приводит к функции $q = f(z, \hat{\beta})$, которая выражает не максимальный, а «средний» объем выпуска. Началу серьезных усилий в направлении преодоления разрыва между теорией и практикой положила

работа (Farrell, 1957), начиная с которой целью статистического анализа стало оценивание именно **граничной производственной функции** (*frontier production function*). Если $q = f(z, \beta)$ понимается именно таким образом, тогда выпуск фирмы i меньше теоретического (в крайнем случае, равен теоретическому), так что

$$q_i = f_i(z_i, \beta) = f(z_i, \beta) \xi_i,$$

где ξ_i — **уровень эффективности (неэффективности) фирмы i** , $0 < \xi_i \leq 1$.

Если $\xi_i = 1$, то фирма i достигает оптимального выпуска, соответствующего технологии, отраженной в указанной производственной функции, т.е. использует все возможности по увеличению выпуска продукции при заданных значениях входных переменных. Если же $\xi_i < 1$, то фирма не полностью использует эти возможности.

В работе (Aigner, Chu, 1968) было предложено оценивать вектор β методами математического программирования, минимизируя при ограничениях $q_i \leq f(z_i, \beta)$ сумму

$$\sum_{i=1}^N |q_i - f(z_i, \beta)|$$

или сумму

$$\sum_{i=1}^N (q_i - f(z_i, \beta))^2.$$

Если функция $f(z, \beta)$ линейна, то в первом случае это — задача линейного программирования, во втором — квадратичного программирования.

Существенным осложнением при реализации этого подхода явилась крайняя чувствительность получаемых оценок к наличию выбросов в данных. Это привело — начиная с работ (Timmer, 1971) и (Dugger, 1974) — к разработке моделей стохастической границы производственных возможностей. В указанных двух работах использовались те же методы математического программирования, однако допускалось, что из-за возможных ошибок измерений некоторая определенная доля наблюдений может находиться выше границы, описываемой функцией $q = f(z, \beta)$. Выбор этой доли был, по существу, произвольным и не опирался на какое-либо экономическое или статистическое обоснование. Статистические свойства получаемых при этом оценок оставались неизвестными. Чтобы дать этим оценкам статистическое обоснование, Шмидт (Schmidt, 1976) рассматривает модель с аддитивной случайной составляющей:

$$y_i = f(z_i, \beta) + \varepsilon_i, \quad i = 1, \dots, N,$$

где $\varepsilon_i \leq 0$.

Задавая (с точностью до неизвестных параметров) распределение случайной составляющей, можно применить для оценивания модели метод макси-

мального правдоподобия. Предположение о том, что $-\varepsilon_i$ имеет показательное распределение, приводит к линейному программированию. Предположение о том, что $-\varepsilon_i$ имеет полунормальное распределение (усеченное в нуле слева нормальное распределение с нулевым математическим ожиданием), приводит к квадратичному программированию. Однако то, что оба эти метода равносильны методу максимального правдоподобия, не дает большого выигрыша, поскольку здесь не выполнены условия регулярности, при которых гарантируются «хорошие свойства» оценок максимального правдоподобия (состоятельность и асимптотическая нормальность). В частности, требование $y_i \leq f(z_i, \beta)$ приводит к тому, что область значений случайной величины y зависит от значений параметра, который оценивается.

Работа (Aigner, Lovell, Schmidt, 1977) положила начало использованию моделей, в которых ε_i является суммой двух компонент:

$$\varepsilon_i = u_i + v_i.$$

Компонента $u_i \leq 0$ суть отклонение от граничной функции вследствие причин, находящихся под контролем фирмы i , и отражает неэффективность использования этой фирмой имеющихся в ее распоряжении ресурсов (**техническая неэффективность** — *technical inefficiency*). Компонента v_i отражает **внешние шоки** (*external shocks*), не находящиеся под контролем фирмы.

Если производственная функция линейна в логарифмах, то обычно рассматривают модель

$$\ln q_i = \ln f(z_i, \beta) + v_i - u_i, \quad (8.1)$$

где $u_i \geq 0$, т.е. модель

$$q_i = f(z_i, \beta) \exp(v_i) \xi_i,$$

где

$$\xi_i = \exp(-u_i).$$

При этом

$f(z_i, \beta)$ — детерминированная граница производственных возможностей;

$f(z_i, \beta) \exp(v_i)$ — **стохастическая граница производственных возможностей** (*stochastic frontier production function*) фирмы i , учитывающая наличие внешних шоков. Наблюдаемые значения q_i находятся на этой границе или ниже нее.

Обычно предполагается, что u_i и v_i распределены независимо друг от друга и от объясняющих переменных, $u_i \sim i.i.d.$ и имеют строго неотрицательное распределение, $v_i \sim i.i.d.$ и имеют симметричное распределение. Таким образом, выражение (8.1) представляет модель компонент ошибки, в которой ошибка раскладывается на две независимые компоненты.

Оценивание параметров модели производится методом максимального правдоподобия. Для проведения такого оценивания необходимо специфицировать распределения компонент ошибки. Чаще всего предполагается, что $v_i \sim i.i.d. N(0, \sigma_v^2)$, а распределение для u_i выбирается из трех вариантов:

- полунормальное распределение $N^+(0, \sigma_u^2)$;
- экспоненциальное (показательное) распределение;
- усеченное в нуле слева нормальное распределение $N^+(\mu, \sigma_u^2)$ с математическим ожиданием μ .

В пакете *Stata* для первых двух вариантов возможно оценивание модели, допускающей гетероскедастичность условного распределения ошибок при фиксированных значениях объясняющих переменных. Для третьего варианта можно смоделировать условное среднее усеченного распределения $N^+(\mu, \sigma_u^2)$ как линейную функцию от соответствующих объясняющих переменных.

ПРИМЕР 8.1.1

Оценивание в пакете *Stata* модели стохастической границы производственных возможностей с выбором усеченного нормального распределения в качестве распределения для u_i . В табл. 8.1 приведена часть протокола оценивания.

Таблица 8.1

Оценка модели стохастической границы производственных возможностей

<i>Inoutput</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $	[95% Conf. Interval]	
<i>lncapital</i>	0.6029056	0.5703415	1.06	0.290	-0.5149432	1.720754
<i>lnlabor</i>	0.6739399	0.2417019	2.79	0.005	0.2002130	1.147667
<i>_cons</i>	0.1262186	3.2968780	0.04	0.969	-6.3355440	6.587981
<i>mu</i>	-0.7811509	2.3573520	-0.33	0.740	-5.4014760	3.839174

Поскольку оценка для μ статистически незначима, можно использовать полунормальное распределение. В табл. 8.2 приведены результаты оценивания редуцированной модели.

Таблица 8.2

Оценка редуцированной модели границы производственных возможностей

<i>Inoutput</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $	[95% Conf. Interval]	
<i>lncapital</i>	0.6279162	0.5645756	1.11	0.660	-0.4786316	1.734464
<i>lnlabor</i>	0.6741655	0.2413587	2.79	0.005	0.2011112	1.147220
<i>_cons</i>	0.0952752	3.2880660	0.03	0.977	-6.3492160	6.539766

В обоих случаях оценка коэффициента эластичности выпуска по капиталу статистически незначима, так что для объяснения выпуска достаточно использовать лишь количество труда. Результаты оценивания редуцированной модели приведены в табл. 8.3. Статистика критерия отношения правдоподобий для проверки гипотезы $\sigma_u^2 = 0$ принимает здесь значение 64.07. Соответствующее P -значение меньше 0.001, так что эта гипотеза уверенно отвергается.

Таблица 8.3

Оценка редуцированной модели по количеству труда

<i>lnoutput</i>	Коэффициент	Стандартная ошибка	z -статистика	P -значение $P > z $	[95% Conf. Interval]	
<i>lnlabor</i>	0.6716777	0.2422926	2.77	0.006	0.1967929	1.146562
<i>_cons</i>	3.6225380	0.8626296	4.20	0.000	1.9318150	5.313261

После оценивания модели вычисляются прогнозные значения выпуска $lnoutput_f_i$ и остатки $e_i = lnoutput_i - lnoutput_f_i$. Оценки для u_i вычисляются с использованием формулы для условного среднего $E(u_i | \varepsilon_i)$ или формулы для условной медианы $Med(u_i | \varepsilon_i)$. Вычисление оценок технической эффективности опирается на формулу для $E(\exp(u_i) | \varepsilon_i)$. Вместо ε_i в эти формулы подставляются остатки e_i , так что

$$\hat{u}_i = E(u_i | e_i), \quad \hat{\xi}_i = E(\exp(u_i) | e_i).$$

Для последней из 3 оцененных моделей получаем картину, изображенную на рис. 8.1, где наблюдаемые значения логарифма выпуска представлены ромбами, а оцененная детерминированная граница — кружками. Для сравнения на рис. 8.2 проведена прямая, подобранная методом наименьших квадратов. На рис. 8.3 показаны результаты для наблюдений 87 — 100, где имеются наблюдения, в которых выпуск превышает оцененную детерминированную границу. На этом рисунке стохастическая граница представлена квадратами. ■

ПРИМЕР 8.1.2

Для анализа качества оценивания параметров модели и ее технической эффективности смоделируем 500 наблюдений, следующих модели стохастической границы

$$lnoutput_i = lnlabor_i + v_i - u_i, \quad v_i \sim i.i.d. N(0, 0.5^2), \quad u_i \sim i.i.d. N^+(0, 1).$$

Результаты оценивания модели по этим данным приведены в табл. 8.4. Угловой коэффициент и обе дисперсии оценены очень хорошо. Соответственно, ряд остатков e_i (E_F) весьма похож на ряд ε_i (EPS). (На рис. 8.4 показаны первые 20 значений этих рядов).

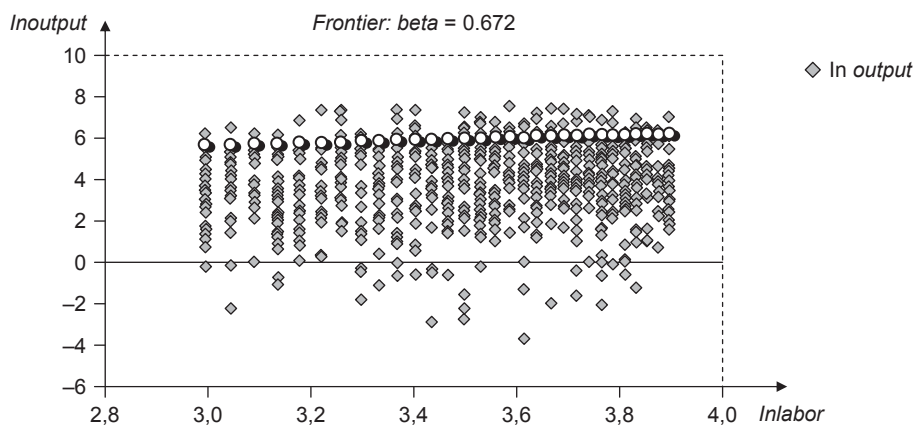


Рис. 8.1

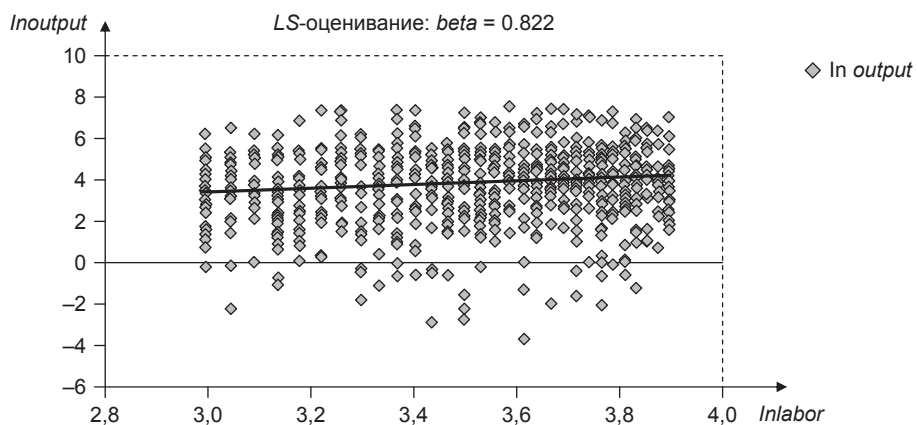


Рис. 8.2

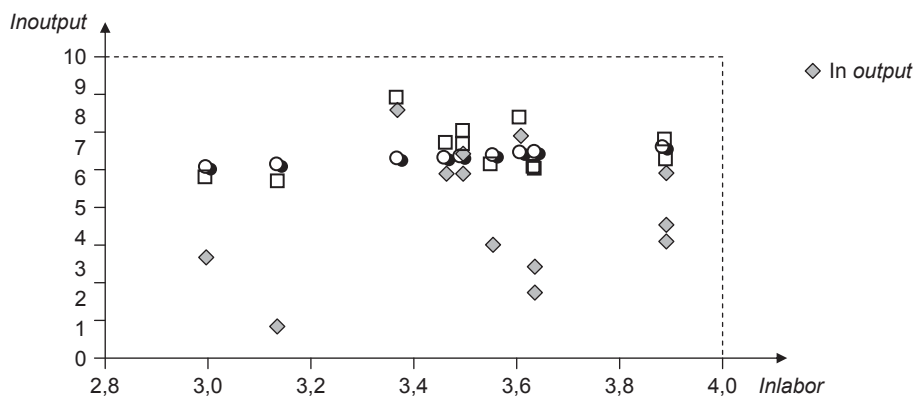


Рис. 8.3

Таблица 8.4

Оценка параметров модели стохастической границы

<i>lnoutput</i>	Коэффициент	Стандартная ошибка	<i>z</i> -статистика	<i>P</i> -значение $P > z $	[95% Conf. Interval]	
<i>lnlabor</i>	1.008306	0.0245941	41.00	0.000	0.9601021	1.056509
<i>sigma_v</i>	0.5037034	0.0542256			0.4078869	0.6220281
<i>sigma_u</i>	1.066262	0.0923971			0.8997109	1.263645
<i>sigma2</i>	1.390632	0.1587568			1.0794750	1.701790
<i>lambda</i>	2.116845	0.1385888			1.8452160	2.388474
<i>Likelihood-ratio test of sigma_u = 0: chibar2(01) = 15.91 Prob>=chibar2 = 0.000</i>						

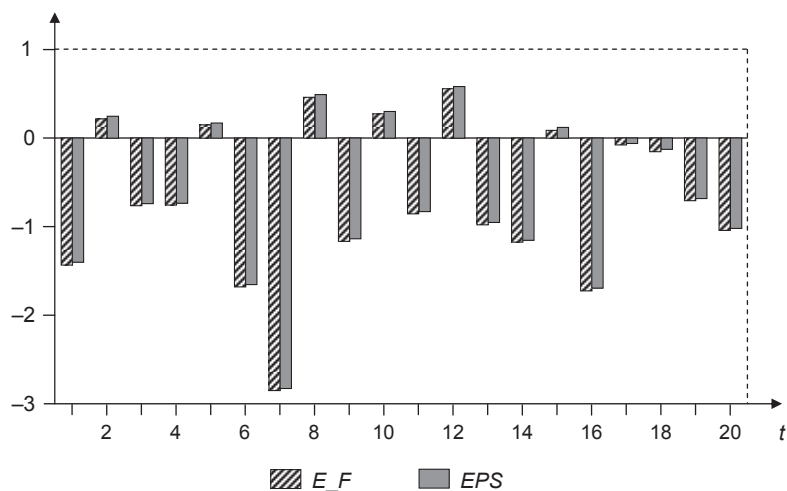


Рис. 8.4

Статистические характеристики этих двух рядов (по всем 500 точкам) представлены на рис. 8.5 и 8.6. Отметим значительное сходство характеристик этих двух рядов.

Статистические характеристики рядов \hat{u}_i (U_F) и u_i (U) представлены на рис. 8.7 и 8.8.

Статистические характеристики ряда технических эффективностей (TE) и ряда оценок технической эффективности (TE_F) приведены соответственно на рис. 8.9 и 8.10. Отметим, что оцененные значения технической эффективности имеют максимальное значение 0.828664, что заметно ниже максимального значения 0.990544 реальной технической эффективности.

Для полноты приведем еще коэффициенты корреляции между истинными и оцененными рядами (табл. 8.5 — 8.8).■

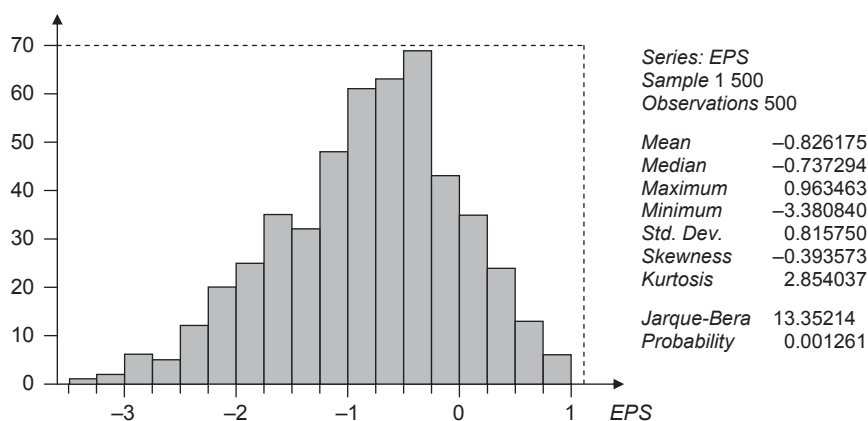


Рис. 8.5

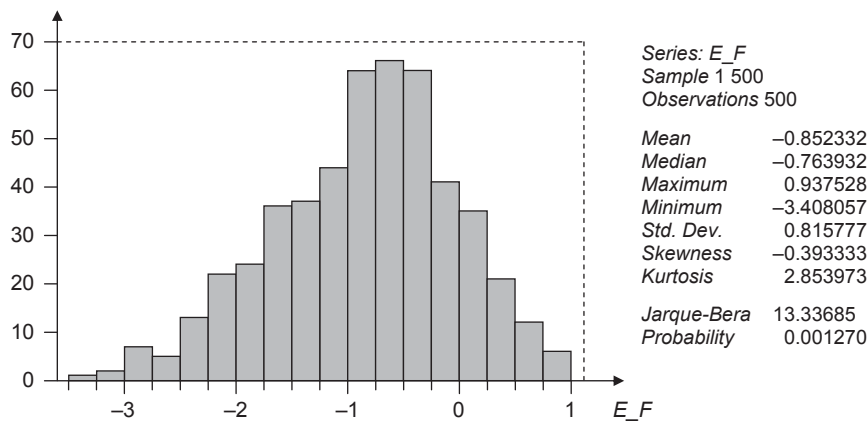


Рис.8.6

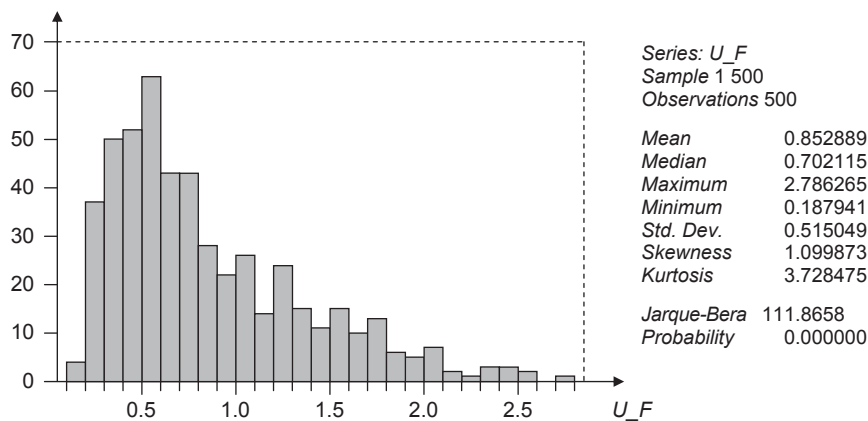


Рис. 8.7

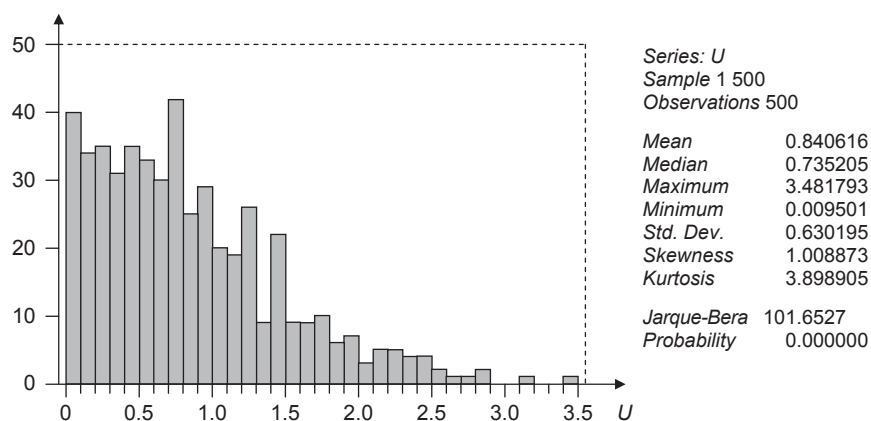


Рис. 8.8

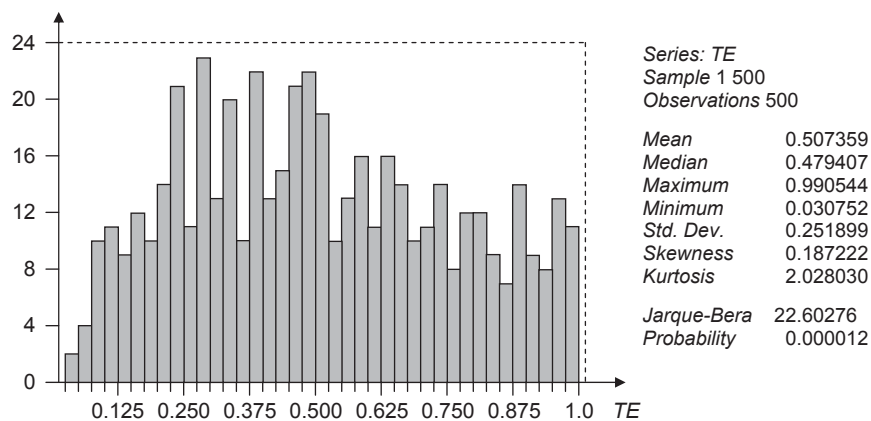


Рис. 8.9

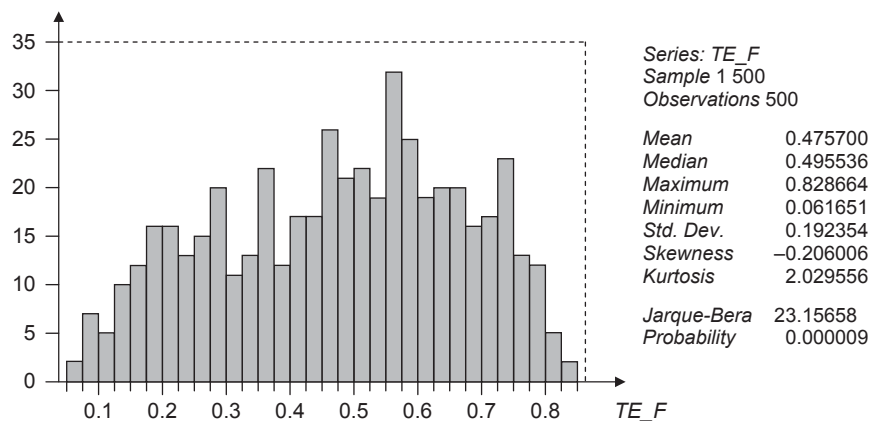


Рис. 8.10

Таблица 8.5

Корреляция между истинным рядом EPS и оцененным

	EPS	E_F
EPS	1	0.9999996
E_F	0.9999996	1

Таблица 8.6

Корреляция между истинным рядом NU и оцененным

	NU	NU_F
NU	1	0.6841711
NU_F	0.6841711	1

Таблица 8.7

Корреляция между истинным рядом U и оцененным

	U	U_F
U	1	0.8237624
U_F	0.8237624	1

Таблица 8.8

Корреляция между истинным рядом TE и оцененным

	TE	TE_F
TE	1	0.7589305
TE_F	0.7589305	1

Проведем теперь небольшое исследование, указывающее на проблемы, возникающие при оценивании параметров детерминированной граничной функции, на которую накладывается случайная составляющая $\varepsilon_i = v_i - u_i$.

Используем в качестве процесса порождения данных модель

$$y_i = 50 + 0.25x_i + v_i - u_i, \quad i = 1, 2, \dots, 200,$$

так что детерминированная граничная функция имеет вид

$$y = 50 + 0.25x.$$

Значения x_i получаются по формуле $x_i = 100 + 300\xi_i$, где $i.i.d.$ случайные величины ξ_i имеют равномерное распределение в интервале $(0, 1)$.

В качестве распределения случайной компоненты u_i будем брать усеченное в нуле нормальное распределение $N^+(\mu, 3^2)$ со значениями $\mu = 3, 6, 9, 12$, а в качестве распределения случайной компоненты v_i возьмем стандартное нормальное распределение $N(0, 1)$. В каждом из 4 случаев сгенерируем по 200 реализаций случайных величин u_i и v_i и вычислим соответствующие им значения y_i . На основе полученной последовательности значений y_i производим оценивание:

- методом наименьших квадратов (*OLS*) параметров α и β модели

$$y_i = \alpha + \beta x_i + \varepsilon_i;$$

- методом максимального правдоподобия (*MLE*) параметров модели

$$y_i = \alpha + \beta x_i + v_i - u_i, \quad u_i \sim N^+(\mu, \sigma_u^2), \quad v_i \sim N(0, \sigma_v^2).$$

В табл. 8.9, 8.10 и на рис. 8.11 приведены результаты *OLS*- и *MLE*-оценивания для случая $\mu = 3$. В табл. 8.11, 8.12 и на рис. 8.12 — результаты *OLS*- и *MLE*-оценивания для случая $\mu = 6$. В табл. 8.13, 8.14 и на рис. 8.13 — результаты *OLS*- и *MLE*-оценивания для случая $\mu = 9$. В табл. 8.15, 8.16 и на рис. 8.14 — результаты *OLS*- и *MLE*-оценивания для случая $\mu = 12$.

Таблица 8.9

Результаты *OLS*-оценивания для $\mu = 3$
Dependent Variable: Y; Method: Least Squares; Sample: 1 200;
Included observations: 200

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
X	0.250406	0.002030	123.3265	0.0000
C	46.16898	0.520325	88.73111	0.0000

Таблица 8.10

Результаты *MLE*-оценивания для $\mu = 3$

y	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $
x	0.2498361	0.0015163	164.76	0.000
$_{cons}$	50.26181	0.5000632	100.51	0.000
$/mu$	3.024968	0.6806073	4.44	0.000
H0: No inefficiency component: $z = -3.569$ Prob $\leq z = 0.000$				

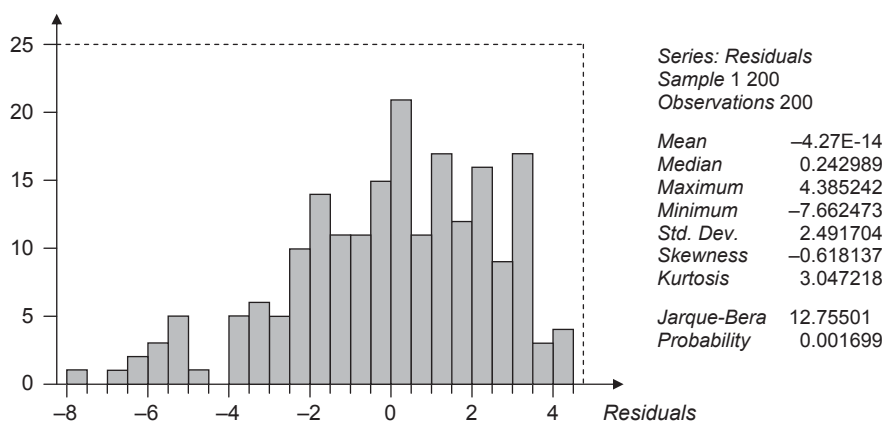


Рис. 8.11

Таблица 8.11

Результаты OLS-оценки для $\mu = 6$

Dependent Variable: Y; Method: Least Squares; Sample: 1 200; Included observations: 200

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
X	0.248631	0.002549	97.53219	0.0000
C	44.41100	0.653272	67.98238	0.0000

Таблица 8.12

Результаты MLE-оценки для $\mu = 6$

y	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $
x	0.2476226	0.0025763	96.12	0.000
_cons	51.6534	1.862753	27.73	0.000
/mu	6.87178	1.828409	3.76	0.000
H0: No inefficiency component: z = -0.603 Prob<=z = 0.273				

Таблица 8.13

Результаты OLS-оценки для $\mu = 9$

Dependent Variable: Y; Method: Least Squares; Sample: 1 200; Included observations: 200

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
X	0.251492	0.002604	96.57229	0.0000
C	40.83894	0.667356	61.19510	0.0000

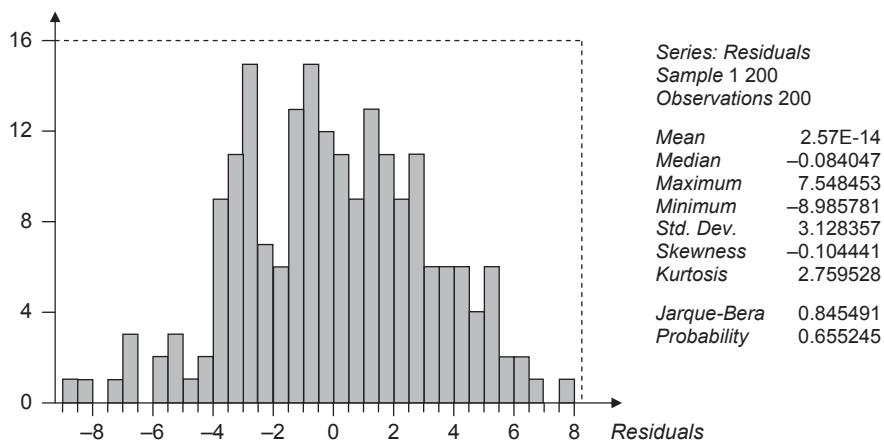


Рис. 8.12

Таблица 8.14

Результаты MLE-оценивания для $\mu = 9$

y	Коэффициент	Стандартная ошибка	z -статистика	P -значение $P > z $
x	0.2513325	0.0026927	93.34	0.000
$_{cons}$	42.41768	—	—	—
$/\mu$	-0.0079909	6.295305	-0.00	0.999
H0: No inefficiency component: $z = -0.280$ Prob $\leq z = 0.390$				

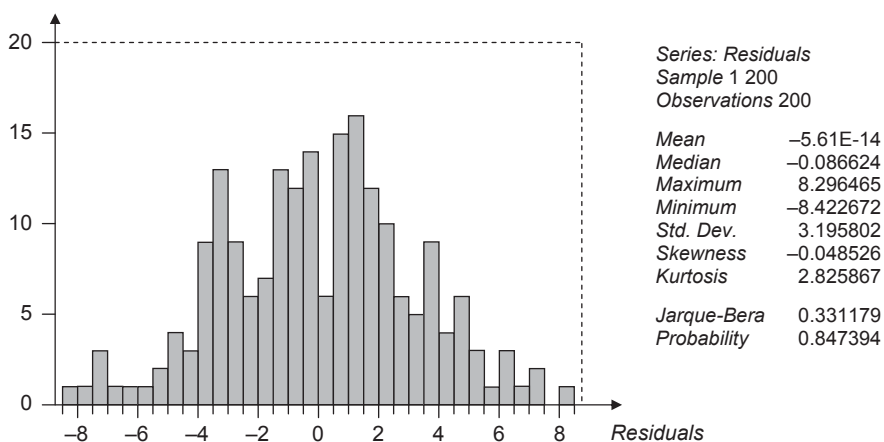


Рис. 8.13

Таблица 8.15

Результаты OLS-оценивания для $\mu = 12$

Dependent Variable: Y; Method: Least Squares; Sample: 1 200; Included observations: 200

Переменная	Коэффициент	Стандартная ошибка	t-статистика	P-значение
X	0.247783	0.002688	92.18696	0.0000
C	38.79056	0.688791	56.31686	0.0000

Таблица 8.16

Результаты MLE-оценивания для $\mu = 12$

y	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $
x	0.2477825	0.0026744	92.65	0.000
_cons	38.85209	11.3368	3.43	0.001
/mu	-0.000011	—	—	—

H_0 : No inefficiency component: $z = 0.200$ Prob $\leq z = 0.579$

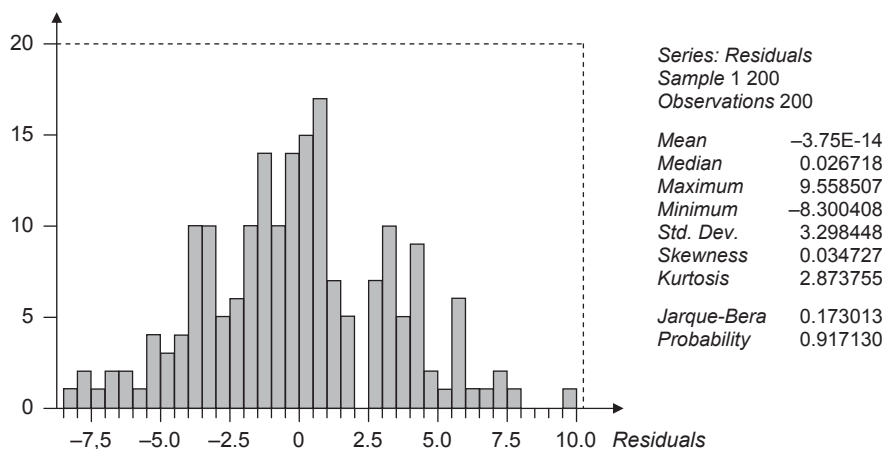


Рис. 8.14

Проанализируем полученные результаты:

- в случае $\mu = 3$ MLE-оценка параметра α (50.26181) близка к использованному в модели значению $\alpha = 50$, тогда как OLS-оценка этого параметра равна 46.16898; MLE-оценка параметра μ (3.024968) близка к использованному в модели значению $\mu = 3$;

- в случае $\mu = 6$ MLE -оценка параметра α (51.6534) также близка к использованному в модели значению $\alpha = 50$, тогда как OLS -оценка этого параметра равна 44.41100; MLE -оценка параметра μ (6.87178) близка к использованному в модели значению $\mu = 6$;
- в случае $\mu = 9$ OLS -оценка параметра α (40.83894) еще дальше отстоит от использованного в модели значения $\alpha = 50$. Но в этом случае смещается и MLE -оценка этого параметра: она равна 42.41768; при этом значение μ оценивается величиной -0.0079909 .
- наконец, в случае $\mu = 12$ OLS и MLE дают близкие друг к другу оценки (38.79056 и 38.85209 соответственно); значение μ оценивается величиной -0.000011 .

В связи с полученными результатами следует обратить внимание на следующее:

- в случае $\mu = 3$ коэффициент асимметрии ряда OLS -остатков равен -0.618137 ; гипотеза равенства нулю компоненты ошибки, отвечающей за наличие неэффективности, отвергается;
- в случае $\mu = 6$ коэффициент асимметрии ряда OLS -остатков равен -0.104441 ; гипотеза равенства нулю компоненты ошибки, отвечающей за наличие неэффективности, не отвергается;
- в случае $\mu = 9$ коэффициент асимметрии ряда OLS -остатков равен -0.048526 ; гипотеза равенства нулю компоненты ошибки, отвечающей за наличие неэффективности, не отвергается;
- в случае $\mu = 12$ коэффициент асимметрии ряда OLS -остатков равен 0.034727 ; гипотеза равенства нулю компоненты ошибки, отвечающей за наличие неэффективности, не отвергается.

Заметим еще, что гипотеза нормальности OLS -остатков отвергается только при $\mu = 3$; при переходе от значения $\mu = 3$ к значениям $\mu = 6$, $\mu = 9$, $\mu = 12$ происходит последовательное возрастание P -значений статистики Харке — Бера от 0.001699 до 0.917130. И это вполне объяснимо: если при $\mu = 3$ распределение суммы $\varepsilon_i = v_i - u_i$ существенно несимметрично (что подтверждается формой гистограммы остатков и достаточно большим по абсолютной величине отрицательным значением коэффициента асимметрии), то с возрастанием μ это распределение становится все более симметричным и приближается к нормальному распределению.

Таким образом, предпосылкой для успешного оценивания параметров модели стохастической границы является достаточно большое по абсолютной величине отрицательное значение коэффициента асимметрии ряда остатков, полученных при OLS -оценивании.

Одним из возможных применений модели стохастической границы является оценивание технической эффективности банков. Это сделано, например,

в работе (Головань, Карминский, Пересецкий, 2007). При этом эффективность банков оценивалась с двух точек зрения:

- с точки зрения выдачи кредитов;
- с точки зрения привлечения депозитов.

В каждом из двух случаев рассматривается производственная функция банка, т.е. банк считается производственным предприятием, преобразующим ресурсы в продукт (соответственно кредиты или привлеченные депозиты).

Для построения моделей использовались данные за период с I квартала 2003 г. по III квартал 2005 г. При оценивании из выборки были исключены Внешэкономбанк и Сбербанк, как работающие в условиях, существенно отличающихся от условий для других коммерческих банков.

При оценивании эффективности банков по выдаче кредитов в качестве модели производственной функции использовалась функция Кобба — Дугласа. В левой части уравнения стоят кредиты нефинансовым организациям KE (рассматривались только долгосрочные кредиты, так как основная задача банков — преобразовывать короткие пассивы в долгосрочные активы, и интересно, насколько эффективно банк это делает). В правой части уравнения стоят депозиты VD , кредиты других банков KDB , административные расходы RSA (отражающие трудовые ресурсы банка).

$$\ln(KE)_i = \beta_0 + \beta_1 \ln(VD)_i + \beta_2 \ln(KDB)_i + \beta_3 \ln(RSA)_i + v_i - u_i.$$

Знаки полученных оценок коэффициентов согласуются с экономической интуицией. Во всех периодах параметр σ_u значимо отличается от нуля, что говорит о наличии эффекта технической эффективности. Оценки коэффициентов модели на различных выборках стабильны; колебания, как правило, лежат в диапазоне, задаваемом стандартными отклонениями. В группах наиболее эффективных и наименее эффективных банков число банков небольшое.

При оценивании эффективности банков по привлечению депозитов рассматривается модель производственной функции с объемом привлеченных депозитов в левой части VD (продукт) и с чистыми активами CA и административными расходами RSA в правой части (ресурсы, с помощью которых банк привлекает депозиты). Модель имеет вид:

$$\ln(VD)_i = \beta_0 + \beta_1 \ln(CA)_i + \beta_2 \ln(RSA)_i + v_i - u_i.$$

Оценки коэффициентов получились еще более устойчивыми, чем в предыдущей модели. Однако распределение эффективности банков с точки зрения привлечения депозитов существенно отличается от распределения эффективности банков с точки зрения выдачи кредитов. Здесь уже гораздо больше как крайне неэффективных банков, так и банков с достаточно высокой эффективностью.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Опишите модель стохастической границы производственных возможностей.
2. Как оценивается модель стохастической границы производственных возможностей?
3. Как влияет на качество оценивания параметров модели стохастической границы коэффициент асимметрии ряда остатков, полученных при оценивании модели методом наименьших квадратов?
4. Как можно использовать модель стохастической границы для оценивания эффективности банков?

Тема 8.2

МОДЕЛИ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ

Выше была рассмотрена модель для перекрестной выборки (*cross-section*). Теперь перейдем к рассмотрению модели для панельных данных.

Предположим, что если не принимать в расчет возможную неэффективность фирм и разного рода неучтенные факторы («ошибки»), то теоретическая производственная функция фирмы, производящей некоторый продукт, имела бы вид:

$$q_{it} = f(z_{it}, \beta),$$

где q_{it} — выпуск фирмы i в период времени t .

Однако вследствие того или иного уровня неэффективности выпуск фирмы i в период времени t оказывается меньше теоретического, так что

$$q_{it} = f(z_{it}, \beta) \xi_{it},$$

где ξ_{it} — уровень эффективности фирмы i в период времени t , $0 < \xi_{it} \leq 1$.

Если $\xi_{it} = 1$, то эта фирма в данный период достигает оптимального выпуска, соответствующего технологии, отраженной в указанной производственной функции. Если же $\xi_{it} < 1$, то эта фирма в данный период не использует все возможности.

Предполагается также, что выпуск подвержен некоторым случайным шокам, так что в итоге получаем:

$$q_{it} = f(z_{it}, \beta) \xi_{it} \exp(v_{it}),$$

или, переходя к логарифму выпуска:

$$\ln q_{it} = \ln f(z_{it}, \beta) + v_{it} - u_{it},$$

где $u_{it} = -\ln \xi_{it}$, $u_{it} \geq 0$.

При использовании панельных данных рассматривают следующие основные модели.

Модель с фиксированными эффектами и постоянной во времени технической эффективностью:

$$y_{it} = \beta_0 + \sum_{j=1}^k \beta_j x_{j,it} + v_{it} - u_i = \alpha + \beta^T X_{it} + v_{it} - u_i,$$

в которой

- $v_{it} \sim i.i.d. N(0, \sigma_v^2)$ и не коррелированы с объясняющими переменными,
- $u_i \geq 0$ имеют произвольное распределение и могут быть коррелированными как с объясняющими переменными, так и с v_{it} .

Оценивая эту модель как обычную модель с фиксированными эффектами, получаем оценки для β_j , а через них — оценки для $\alpha_i = \alpha - u_i$:

$$\hat{\alpha}_i = \bar{y}_{i\bullet} - \hat{\beta}^T \bar{X}_{i\bullet} \quad (\rightarrow \alpha_i \text{ при } T \rightarrow \infty).$$

После этого вычисляем:

$$\max_{i=1, \dots, N} \hat{\alpha}_i = \hat{\alpha}, \quad \hat{u}_i = \hat{\alpha} - \hat{\alpha}_i, \quad \hat{\xi}_{it} = \exp(-\hat{u}_i).$$

Модель со случайными эффектами и постоянной во времени технической эффективностью:

$$y_{it} = \alpha + \beta^T X_{it} + \varepsilon_{it}, \quad \varepsilon_{it} = v_{it} - u_i.$$

В этой модели $u_i \geq 0$ рассматриваются как *i.i.d.* случайные величины с математическим ожиданием μ , не коррелированные с v_{it} и X_{it} . Добавим и вычтем μ в правой части уравнения:

$$y_{it} = (\alpha - \mu) + \beta^T X_{it} + (\varepsilon_{it} + \mu).$$

Если обозначить:

$$\alpha - \mu = \alpha^*, \quad \varepsilon_{it} + \mu = \varepsilon_{it}^*,$$

то тогда

$$y_{it} = \alpha^* + \beta^T X_{it} + \varepsilon_{it}^*, \quad E(\varepsilon_{it}^*) = 0.$$

Используя обычное *RE GLS*-оценивание для этой модели, получаем оценки для α^* и β .

Пусть e_{it}^* — остатки от подобранной модели. Тогда:

$$\hat{u}_i^* = \frac{1}{T} \sum_{t=1}^T e_{it}^* \quad \text{— оценка для } \frac{1}{T} \sum_{t=1}^T v_{it} - u_i + \mu \quad (\rightarrow \mu - u_i \text{ при } T \rightarrow \infty);$$

$\hat{\mu} = \max \hat{u}_i^*$ — оценка для μ ($\rightarrow \mu$ при $N \rightarrow \infty$);
 $\hat{u}_i = \hat{\mu} - \hat{u}_i^*$ — оценка для u_i ($\rightarrow u_i$ при $N, T \rightarrow \infty$);
 $\hat{\xi}_{it} = \exp(-\hat{u}_i)$ — оценка технической эффективности.

Модели со случайными эффектами и постоянной или изменяющейся во времени технической эффективностью — MLE-оценивание. В этих моделях делаются более определенные предположения о распределении случайных эффектов $u_i \geq 0$, что дает возможность использовать метод максимального правдоподобия.

В простейшей спецификации $u_{it} = u_i$ (т.е. u_{it} не зависят от времени), $u_i \sim i.i.d. N^+(\mu, \sigma_u^2)$, $v_{it} \sim i.i.d. N(0, \sigma_v^2)$, причем u_i и v_{it} распределены независимо друг от друга и от объясняющих переменных.

В пакете *Stata* можно производить также оценивание модели, предложенной в работе (Battese, Coelli, 1992) (*time-varying decay model*) с

$$u_{it} = \exp\{-\eta(t - T_i)\} u_i,$$

где T_i — последний период времени для i -й фирмы;

$$u_i \sim i.i.d. N^+(\mu, \sigma_u^2);$$

$$v_{it} \sim i.i.d. N(0, \sigma_v^2), \text{ причем } u_i \text{ и } v_{it} \text{ опять распределены независимо друг от друга и от объясняющих переменных.}$$

Для i -й фирмы базовым является уровень в период T_i . Если $\eta > 0$, то уровень неэффективности фирмы убывает во времени к базовому уровню. Если $\eta < 0$, то уровень неэффективности возрастает до базового уровня.

ПРИМЕР 8.2.1. Модель, инвариантная по времени

Имеются данные по 91 фирме (от 6 до 14 наблюдений на каждую фирму) о количестве произведенной продукции (*widgets*), о количестве затраченных машино-часов (*machines*) и о количестве затраченных человеко-часов (*workers*)¹. Оцениваемая модель имеет вид:

$$\ln widgets_{it} = \beta_0 + \beta_1 \ln machines_{it} + \beta_2 \ln workers_{it} + v_{it} - u_i.$$

Результаты оценивания приведены в табл. 8.17. ■

✓ **Замечание 8.2.1.** Для получения приемлемой точности оценивания параметров μ и σ_u^2 требуются довольно большие выборки.

¹ Статистические данные из пакета *Stata*.

Таблица 8.17

Оценка модели, инвариантной по времени

<i>lnwidgets</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $	[95% Conf. Interval]	
<i>lnmachines</i>	0.2904551	0.0164219	17.69	0.000	0.2582688	0.3226415
<i>lnworkers</i>	0.2943333	0.0154352	19.07	0.000	0.2640808	0.3245858
<i>_cons</i>	3.030983	0.1441022	21.03	0.000	2.748548	3.313418
<i>/mu</i>	1.125667	0.6479217	1.74	0.082	-0.144236	2.39557

ПРИМЕР 8.2.1 (те же данные). Модель с затуханием во времени

Результаты оценивания приведены в табл. 8.18.

Таблица 8.18

Оценка модели с затуханием во времени

<i>lnwidgets</i>	Коэффициент	Стандартная ошибка	z-статистика	P-значение $P > z $	[95% Conf. Interval]	
<i>lnmachines</i>	0.2907555	0.0164376	17.69	0.000	0.2585384	0.3229725
<i>lnworkers</i>	0.2942412	0.0154373	19.06	0.000	0.2639846	0.3244978
<i>_cons</i>	3.028939	0.1436046	21.09	0.000	2.74748	3.310399
<i>/mu</i>	1.110831	0.6452809	1.72	0.085	-0.1538967	2.375558
<i>/eta</i>	0.0016764	0.00425	0.39	0.693	-0.0066535	0.0100064

Оценка параметра $\eta > 0$ близка к нулю и статистически незначима; остальные оценки мало отличаются от предыдущего случая.

Оценивание всех параметров производится методом максимального правдоподобия, после чего вычисляются прогнозные значения u_{it} с использованием формулы для условного среднего $E(u_{it}|\varepsilon_{it})$ или формулы для условной медианы $Med(u_{it}|\varepsilon_{it})$. После этого получаем оценки технической эффективности, опирающиеся на формулу для $E(\exp(u_{it})|\varepsilon_{it})$. Вместо ε_{it} в эти формулы подставляются значения

$$e_{it} = \ln widgets_{it} - \hat{\beta}_0 - \hat{\beta}_1 \ln machines_{it} - \hat{\beta}_2 \ln workers_{it},$$

так что, например,

$$\hat{u}_{it} = E(u_{it}|e_{it}), \quad tech_{it} = E(\exp(u_{it})|e_{it}).$$

В табл. 8.19 приведены оцененные значения для 3 фирм по модели с постоянными эффективностями (ПЭ) и по модели с изменяющимися эффективностями (ИЭ). ■

Таблица 8.19

Оценка эффективности фирм по разным моделям

Фирма 1		Фирма 2		Фирма 3	
ПЭ	ИЭ	ПЭ	ИЭ	ПЭ	ИЭ
0.2735752	0.2727839	0.0537262	0.0527223	0.8113303	0.8113267
0.0735752	0.2723570	0.0537262	0.0529779	0.8113303	0.8115930
0.2735752	0.2729305	0.0537262	0.0532344	0.8113303	0.8118590
0.2735752	0.2735043	0.0537262	0.0534917	0.8113303	0.8121246
0.2735752	0.2740783	0.0537262	0.0537499	0.8113303	0.8123900
0.2735752	0.2746527	0.0537262	0.0540089	0.8113303	0.8126550
0.2735752	0.2752275	0.0537262	0.0542687	0.8113303	0.8129198
0.2735752	0.2758025	0.0537262	0.0545293	0.8113303	0.8131843
0.2735752	0.2763778	0.0537262	0.0547908	0.8113303	0.8134484
0.2735752	0.2769534			0.8113303	0.8137122
0.2735752	0.2775293			0.8113303	0.8139758

Понятие стохастической границы можно использовать для вычисления налогового потенциала регионов. Налоговую нагрузку как следствие проведения политики региональных властей можно представлять как выпуск «предприятия», на входе которого находятся факторы, определяющие налоговые поступления (такие, как налоговая база, ставки налога и т.п.).

Теоретически ввиду сходства проблем методология стохастической границы должна хорошо работать и при оценивании границ налогов. Однако если в случае производственной функции детерминанты объема выпуска довольно ясны, то в случае оценивания границы для налогов положение является менее ясным. Ставки налогов могут быть близкими для разных регионов, но налоговые базы — существенно различаться. При этом, особенно в развивающихся странах, данные о налоговых базах часто недоступны или ненадежны, и для них приходится использовать некоторые прокси (продукт или доход, или такие экономические индикаторы, как уровень образования либо использование электроэнергии). Соответственно построение стохастической границы является успешным только при подборе «правильного» набора детерминантов налоговой нагрузки.

В случае построения границ налоговой нагрузки отклонение налоговой нагрузки региона от налоговой границы следует интерпретировать только как уровень неиспользованного налогового потенциала, но не как меру неэффективности. Недоиспользование налогового потенциала может объясняться по крайней мере двумя факторами:

- региональными предпочтениями (привычка к низкому уровню снабжения общественными благами);
- неэффективностью региональных властей.

Построение границ налоговой нагрузки регионов методом стохастической границы описывается, например, в работе (*Alfirman, 2003*) для Индонезии.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Каковы основные модели стохастической границы для панельных данных?
2. Каким образом оцениваются модели стохастической границы для панельных данных?
3. Как можно использовать понятие стохастической границы для вычисления налогового потенциала регионов?

Задания для семинарских занятий, работы в компьютерном классе и для самостоятельной работы

Методические указания к заданиям, предполагающим использование пакетов компьютерных программ статистического анализа данных, ориентированы на использование пакетов *EViews* 6 и *Stata* 10.

К разделу 1

Задание 1. *Сглаживание временного ряда с использованием простого скользящего среднего*

Повторите исследование, проведенное в примере 5.1.1 для ряда значений индекса ММВБ (ряд *IND*). Данные для этого примера содержатся в табл. П-1 Приложения¹. Их можно также взять на сайте <http://www.micex.ru/marketdata/indices/data/archive>.

Методические указания. Для построения трехпериодного скользящего среднего в пакете *EViews* 6 можно использовать цепочку *Objects* → *New object* → *Model* → $sma_3=(ind(-1)+ind+ind(+1))/3$ или воспользоваться встроенной функцией *@movav*: *Genr* → $sma_3=@movav(ind(1),3)$. Здесь $(ind(1), 3)$ указывает на то, что осредняются текущее и 2 предыдущих значения ряда $ind(1)$, т.е. текущее и по одному соседнему (до и после) значению самого ряда *ind*.

Для построения пятипериодного скользящего среднего можно использовать цепочку *Objects* → *New object* → *Model* → $sma_5=(ind(-2)+ind(-1)+ind+ind(+1)+ind(+2))/5$ или опять воспользоваться встроенной функцией *@movav*: *Genr* → $sma_5=@movav(ind(2),5)$. Здесь $(ind(2), 5)$ указывает на то, что осредняются текущее и 4 предыдущих значения ряда $ind(2)$, т.е. текущее и по два соседних (до и после) значений самого ряда *ind*.

Задание 2. *Сглаживание временного ряда с использованием центрированного скользящего среднего*

1. Повторите исследование, проведенное в примере 5.1.2 для ряда, представляющего квартальные данные об индексе реального объема сельскохозяй-

¹ В заданиях к разделам этой части учебника даются ссылки на Приложение, содержащее необходимые статистические данные и их описание, помещенное после задания 29.

ственного производства в Российской Федерации (ряд *AGRO*). Данные для этого примера содержатся в табл. П-2 Приложения. Их можно также найти на статистическом портале ГУ—ВШЭ <http://stat.hse.ru>.

2. Повторите исследование, проведенное в примере 5.1.3 для ряда, представляющего квартальные данные об импорте товаров и услуг в Российской Федерации (ряд *IMP*). Данные для этого примера содержатся в табл. П-3 Приложения (на сайтах <http://stat.hse.ru> и <http://www.rusimpex.ru> данные за 2004 и 2005 гг. несколько отличаются от использованных нами в примере 5.1.3).

Методические указания (к пунктам задания).

1. Для построения центрированного среднего с периодом 4 можно использовать цепочку $Genr \rightarrow sma_4_centered = (0.5 * agro(-2) + @movsum(agro(1), 3) + 0.5 * agro(2)) / 4$.
2. Аналогичная цепочка используется для ряда *IMP*.

Задание 3. Фильтр Ходрика — Прескотта

1. Используя фильтр Ходрика — Прескотта, выделите долговременную компоненту ряда *AGRO* (см. пример 5.1.4).
2. Используя фильтр Ходрика — Прескотта, выделите долговременную компоненту ряда *IMP* (см. пример 5.1.5).
3. Используя фильтр Ходрика — Прескотта, выделите долговременную компоненту ряда *IND*, рассмотренного в задании 1. Сравните результаты выделения тренда при различных значениях параметра сглаживания (см. пример 5.1.6).

Методические указания (к пунктам задания).

1. Для применения фильтра Ходрика — Прескотта к ряду *AGRO* используйте (в рамках объекта *Series*) цепочку $Agro \rightarrow Procs \rightarrow Hodrick-Prescott\ filter \rightarrow$ название сглаженного ряда (*hptrend*) и значение параметра сглаживания ($lambda=1600$) $\rightarrow OK$.

Аналогично поступите с рядом *IMP* в п. 2.

3. Используйте следующие значения параметра сглаживания: 100 (как для годовых данных), 1600 (как для квартальных данных), 14 400 (как для месячных данных) и 100 000.

Задание 4. Простое экспоненциальное сглаживание

Повторите исследование, проведенное в примере 5.1.7 для ряда *IND*.

Методические указания. В случае, когда программе предоставляется возможность самой выбрать оптимальное значение параметра сглаживания, достаточно в рамках объекта *Series* использовать цепочку $Procs \rightarrow Exponential\ Smoothing \rightarrow Smoothing\ Method: Single \rightarrow$ название сглаженного ряда (*indsm*) и предпрогнозный период $\rightarrow OK$. Если значение параметра сглаживания задается заранее (в примере 5.1.7 это было значение 0.2), то помимо названия сглаженного ряда и предпрогнозного периода следует указать еще значение параметра сглаживания (*Smoothing parameters: Alpha = 0.2*).

Задание 5. Двойное экспоненциальное сглаживание

Повторите исследование, проведенное в примере 5.1.8 для ряда *UNEMP*. Данные, использованные в этом примере, содержатся в табл. П-4 Приложения. Обновленные данные можно найти на сайте <http://www.prime-tass.ru>.

Методические указания. В рамках объекта *Series* используйте цепочку *Procs* → *Exponential Smoothing* → *Smoothing Method: Double* → название сглаженного ряда (*unemp_desm*), предпрогнозный период (1994:1—1996:4), выбранное значение параметра сглаживания $\text{Alpha} = 0.2$ → *OK*. Если программе предоставляется возможность самой выбрать оптимальное значение параметра сглаживания, то задавать значение параметра сглаживания не следует.

Задание 6. Метод Хольта

Повторите исследование, проведенное в примере 5.1.9 для ряда *UNEMP*.

Методические указания. Для применения метода Хольта с предоставлением программе возможности выбора оптимальных значений параметров α и β используйте цепочку *Procs* → *Exponential Smoothing* → *Smoothing Method: Holt-Winters — No seasonal* → название сглаженного ряда (*unemp_hw_est*), предпрогнозный период (1994:1—1996:4) → *OK*.

Задание 7. Метод Хольта — Винтерса

Повторите исследование, проведенное в примере 5.1.10.

Методические указания. Для применения метода Хольта — Винтерса с предоставлением программе возможности выбора оптимальных значений параметров используйте следующие цепочки.

- **Аддитивная сезонность:** *Procs* → *Exponential Smoothing* → *Smoothing Method: Holt-Winters — Additive* → название сглаженного ряда (*AGRO_HW_AD*), предпрогнозный период (1994:1—1996:4) → *OK*.
- **Мультипликативная сезонность:** *Procs* → *Exponential Smoothing* → *Holt-Winters — Multiplicative* → название сглаженного ряда (*AGRO_HW_MU*), предпрогнозный период (1994:1—2003:4) → *OK*.

Задание 8. Прогноз по модели, оцененной методом наименьших квадратов: линейный тренд

Повторите исследование, проведенное в примере 5.1.11.

Методические указания. Используя объект *Equation*, оцените методом наименьших квадратов уравнение линейного тренда на предпрогножном периоде (спецификация *UNEMP C @TREND+1*, *Sample*: 1994:1—1996:4). Для получения ряда прогнозов и характеристик качества прогноза нажмите в меню объекта *Equation* виртуальную кнопку *Forecast* (т.е. прогнозирование) и в открывшемся меню укажите период, для которого строятся прогнозные значения: в нашем случае, это период 1997:1—1997:4. По умолчанию прогнозные значения помещаются во вновь создаваемый ряд (объект *Series*), название которого отличается от названия исходного ряда добавлением в конце буквы *F* (в нашем случае это будет ряд *UNEMPF*).

В дополнение к ряду прогнозных значений можно построить и ряд значений **стандартных ошибок прогнозов**, позволяющий строить доверительные интервалы для будущих значений ряда. Для этого надо в том же меню *Forecast* дополнительно задать в окне *S.E.(optional)* название ряда стандартных ошибок: *UNEMP_SE*. После этого нажимаем виртуальную кнопку *OK* и получаем график, на котором показаны прогнозные значения ряда на 1997 г. и доверительные границы для прогнозов по каждому кварталу, а также таблицу, в которой приведены характеристики качества полученного прогноза на 4 квартала 1997 г.

Задание 9. *Прогноз по модели, оцененной методом наименьших квадратов: полиномиальный тренд*

Повторите исследование, проведенное в примере 5.1.12. Статистические данные для этого задания возьмите из табл. П-5 Приложения.

Методические указания. См. методические указания к заданию 8.

Задание 10. *Прогноз по модели, оцененной методом наименьших квадратов: полиномиальный тренд и сезонность*

Повторите исследование, проведенное в примере 5.1.13.

Методические указания. Для выделения из ряда остатков *RESID_EQ1* циклических составляющих с периодами от 1.5 до 4 лет используйте (в рамках объекта *Series*) цепочку *Proc* → *Frequency Filter* → *Cycle periods* — *Low*: 6.0 *High*: 16.0 → *OK*.

Задание 11. *Прогноз по подобранной стационарной модели AR*

Повторите исследование, проведенное в примере 5.2.1.

Методические указания. В меню объекта *Equation* после оценивания модели на предпрогнозном периоде (1970 — 1990) выбираем *Forecast* и откроется страница для спецификации прогноза:

Forecast

Forecast of
Equation: UNTITLED Series: X

Series names
Forecast name: xf
S.E. (optional):
GAR,CH(optional):

Method
☒ Dynamic forecast
☐ Static forecast
☐ Structural (ignore ARMA)
☒ Coef uncertainty in S.E. calc

Forecast sample
1991|1992

Output
☒ Forecast graph
☒ Forecast evaluation

☒ Insert actuals for out-of-sample observations

OK Cancel

На этой странице по умолчанию:

- дается название для ряда прогнозов: xf ;
- указывается метод получения прогнозных значений:
Method — *Dynamic forecast* (динамический прогноз),
Coef uncertainty in S.E. calc (при вычислении стандартных ошибок прогнозов учитывается неопределенность оценок коэффициентов).

В окне *Forecast sample* надо задать начало и конец прогнозного периода (в нашем примере это 1991, 1992).

После получения прогноза (ряд xf) можно сравнить полученные прогнозные значения с действительными, создав группу из рядов x и xf .

Если на странице в качестве метода выбрать *Static forecast*, то будет получена последовательность одношаговых прогнозов. В нашем примере это будут:

- прогнозное значение на 1991 г., построенное по данным с 1970 по 1990 г.;
- прогнозное значение на 1992 г., построенное по данным с 1970 по 1991 г.

Задание 12. *Прогноз ряда уровней по подобранной стационарной AR модели для ряда разностей*

Повторите исследование, проведенное в примере 5.2.2. Статистические данные приведены в табл. П-6 Приложения.

Методические указания. При построении прогноза для ряда уровней по оцененной стационарной $AR(1)$ модели для ряда разностей удобно поступить следующим образом. В рамках объекта *Equation* вместо спецификации *gnp_dif* с *ar(1)* задайте спецификацию *d(gnp)* с *ar(1)*. Такая спецификация уведомляет программу о том, что в качестве объясняемой переменной используется не ряд уровней, а ряд разностей. Это удобно тем, что прогноз в этом случае выдается не только для ряда разностей, но и для ряда уровней. Чтобы получить прогноз для ряда уровней, в спецификации прогноза надо указать: *Series to forecast* — *gnp*.

Задание 13. *Прогноз по модели авторегрессии с сезонными дамми*

Повторите исследование, проведенное в примере 5.2.3. Обновленные статистические данные об объемах производства тканей в Российской Федерации можно найти на сайте <http://www.prime-tass.ru>.

Задание 14. *Сравнение прогнозов, полученных по модели ARIMA, с адаптивными прогнозами*

Повторите исследование, проведенное в примере 5.2.4, используя обновленные статистические данные об уровнях безработицы в США в период 1987—1990 гг., которые можно найти на сайте <http://research.stlouisfed.org/fred2/data/UNRATE.txt> или на сайте <http://www.davemanuel.com/historical-unemployment-rates-in-the-united-states.php>.

Задание 15. *Прогноз в условиях неопределенности в отношении стационарности/нестационарности ряда наблюдений*

Повторите исследование, проведенное в примере 5.2.5.

К разделу 2

Задание 16. Проверка на отсутствие/наличие причинности по Грейнджеру для двух рядов

1. Рассмотрите модель $VAR(1)$ для двух рядов:

$$Y_{1t} = 0.6 + 0.7Y_{1,t-1} + 0.2Y_{2,t-1} + \varepsilon_{1t},$$

$$Y_{2t} = 0.4 + 0.2Y_{1,t-1} + 0.7Y_{2,t-1} + \varepsilon_{2t}.$$

Выполнено ли условие стабильности для этой VAR ? Если это условие выполнено, то найдите долгосрочное (стабильное) состояние системы.

2. Постройте реализацию указанной модели, полагая $Y_{11} = Y_{21} = 0$ и $D(\varepsilon_{1t}) = D(\varepsilon_{2t}) = 0.01$. Объясните поведение полученных реализаций. Какие значения Y_{11} и Y_{21} следует взять, чтобы система сразу вошла в стабильный режим? Постройте реализацию модели $VAR(1)$ с этими значениями Y_{11} и Y_{21} .
3. По полученной реализации модели (второй вариант) оцените модель $VAR(1)$ для двух рядов. Сравните оценки коэффициентов со значениями коэффициентов, использованными при моделировании. Убедитесь в том, что оцененная $VAR(1)$ является стабильной (проверьте расположение корней характеристического уравнения на комплексной плоскости).
4. Проведите диагностику остатков на предмет обнаружения нарушений стандартных предположений VAR .
5. Подтвердите правильность выбора порядка модели, сравнивая ее с моделями более высокого порядка.
6. Проверьте гипотезу об отсутствии причинности по Грейнджеру для каждого из рядов по отношению к другому ряду.

Методические указания.

1. В компактной форме указанная $VAR(1)$ имеет вид:

$$y_t = \mu + \Pi_1 y_{t-1} + \varepsilon_t,$$

где

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \quad \mu = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}, \quad \Pi_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

или

$$A(L)y_t = \mu + \varepsilon_t,$$

где

$$A(L) = I_2 - \Pi_1 L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.7L & 0.2L \\ 0.2L & 0.7L \end{pmatrix} = \begin{pmatrix} 1-0.7L & -0.2L \\ -0.2L & 1-0.7L \end{pmatrix},$$

так что

$$A(1) = \begin{pmatrix} 0.3 & -0.2 \\ -0.2 & 0.3 \end{pmatrix}, \quad A^{-1}(1) = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix}.$$

Уравнение $\det A(z) = 0$ принимает здесь вид:

$$\det A(z) = \begin{pmatrix} 1 - 0.7z & -0.2z \\ -0.2z & 1 - 0.7z \end{pmatrix} = 0,$$

т.е. $(1 - 0.7z)^2 - (0.2z)^2 = 0$, или $(1 - 0.9z)(1 - 0.5z) = 0$. Оба корня $z = \frac{1}{0.9}$

и $z = \frac{1}{0.5}$ больше 1, т.е. условие стабильности выполняется.

Долгосрочное (стабильное) состояние системы находим по формуле:

$$y_t = A^{-1}(1)\mu = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 5.2 \\ 4.8 \end{pmatrix}.$$

Таким образом, стабильное состояние системы определяется здесь как

$$y_{1t} = 5.2, \quad y_{2t} = 4.8,$$

так что стабильное состояние разности $y_{1t} - y_{2t}$ есть

$$y_1 - y_2 = 0.4.$$

Соответственно с течением времени независимо от начальных условий ряд y_{1t} начинает осциллировать вокруг уровня 5.2, а ряд y_{2t} начинает осциллировать вокруг уровня 4.8; разность $(y_{1t} - y_{2t})$ осциллирует вокруг уровня 0.4.

2. Сгенерируйте ряды $Y1$ и $Y2$, полагая сначала все значения каждого из них равными 0. Специфицируйте объект *Model*:

$$Y1 = c(11) + c(12)*x1(-1) + c(13)*x2(-1) + c(14)*y1(-1) + c(15)*y2(-1)$$

$$Y2 = c(16) + c(17)*x1(-1) + c(18)*x2(-1) + c(19)*y1(-1) + c(20)*y2(-1)$$

и реализуйте эту модель порождения рядов $Y1$ и $Y2$. Объясните совместное поведение полученных рядов.

Измените начальные значения этих рядов на значения, соответствующие стабильному положению системы, т.е., на 5.2 и 4.8, соответственно. Реализуйте ту же модель порождения рядов с новыми начальными значениями. Объясните совместное поведение полученных рядов.

3. Откройте пару рядов $Y1$ и $Y2$ в виде *VAR (Open as VAR...)*. В открывшемся меню укажите следующее (остальное — по умолчанию):

VAR specification: Unrestricted VAR (*VAR* без ограничений на коэффициенты),

Lag intervals: 1 1 (в модели учитываются запаздывания только на один шаг).

→ OK.

При этом автоматически создается объект *VAR* без имени (*UNTITLED*), в котором приведены результаты оценивания модели *VAR*. Сравните полученные оценки коэффициентов со значениями коэффициентов, использованными при моделировании. Убедитесь в том, что оцененная *VAR(1)* яв-

ляется стабильной. Для этого проверьте расположение корней характеристического уравнения (не путать с обратным характеристическим уравнением!) на комплексной плоскости, используя переходы:

View → Lag Structure → AR Roots Graph.

Оцененная система стабильна, если все корни характеристического уравнения расположены внутри единичного круга на комплексной плоскости. По цепочке

View → Lag Structure → AR Roots Table

получаем таблицу значений величин, обратных корням AR -полинома, в нижней части которой указывается, удовлетворяет ли оцененная VAR условию устойчивости.

4. Для диагностики остатков используйте

View → Residual Tests

и процедуры, предлагаемые в выпадающем подменю:

- *Portmanteau Autocorrelation Test...* — вычисляет многомерные статистики Бокса — Пирса и Люнга — Бокса. При нулевой гипотезе об отсутствии автокорреляций в рядах остатков до порядка h включительно обе статистики приблизительно распределены как $\chi^2(k^2(h-p))$, где k — число переменных, включенных в векторную авторегрессию; p — порядок векторной авторегрессии;
 - *Autocorrelation LM Test...* — дает возможность провести многомерный тест для проверки отсутствия автокорреляции определенного порядка.
 - *Normality Test...* — представляет многомерные расширения теста Харке — Бера на нормальность остатков модели;
 - *White Heteroskedasticity (No Cross Terms / Cross Terms)* — данные тесты являются расширениями теста Уайта для случая систем уравнений.
5. Для подтверждения правильности выбора порядка модели можно оценить модель большего порядка, например, $VAR(2)$, и проверить на совместную незначимость оценки коэффициентов при $y_1(-2)$ и $y_2(-2)$. Это можно сделать прямым вычислением статистики критерия отношения правдоподобий (LR -критерий):

$$lrstat = 2(l_2 - l_1) = T(\ln |\hat{\Omega}_1| - \ln |\hat{\Omega}_2|),$$

где l_p — логарифм максимума функции правдоподобия при оценивании $VAR(p)$,

$|\hat{\Omega}|$ — определитель оцененной ковариационной матрицы случайных ошибок модели, вычисляемой двумя разными способами:

$$|\hat{\Omega}| = \det\left(\frac{1}{T} \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_t^T\right) \quad \text{или} \quad |\hat{\Omega}| = \det\left(\frac{1}{T-p} \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_t^T\right).$$

Оба значения $|\hat{\Omega}|$ приводятся в протоколе оценивания, соответственно, как *Determinant resid covariance* и *Determinant resid covariance (dof adj.)*, а значение логарифма максимума функции правдоподобия обозначено как *Log likelihood*.

При вычислении статистики критерия следует брать одинаковое количество наблюдений T при оценивании обеих VAR ! Кроме того, согласно рекомендации Симса, предпочтительнее брать в формуле для статистики критерия не само T , а значение $(T - k)$, где k — количество параметров в одном уравнении модели большего (именно так производится вычисление в *EViews* 6) порядка. Таким образом, вычисленное по указанной выше формуле значение *lrstat* рекомендуется умножать на $\frac{T - k}{T}$.

Вычислите скорректированное значение $lrstat = 2(l_1 - l_2) \frac{T - k}{T}$ и примените LR -критерий с этим значением статистики, ориентируясь на асимптотическое хи-квадрат распределение этой статистики с числом степеней свободы, равным количеству зануляемых коэффициентов.

В рамках объекта *Var* для $VAR(2)$ произвести проверку рассматриваемой гипотезы можно, используя цепочку

View \rightarrow *Lag Structure* \rightarrow *Lag Exclusion Tests*.

В появляющейся таблице указано значение скорректированной статистики LR -критерия и соответствующее ему P -значение, рассчитанное по распределению хи-квадрат.

Для выбора глубины запаздываний (порядка VAR) можно использовать также сопоставление значений информационных критериев для оцененных VAR различных порядков. Для этого следует использовать цепочку

View \rightarrow *Lag Structure* \rightarrow *Lag Length Criteria*;

при этом следует указать максимально возможную глубину запаздываний.

6. Оцените уравнение

$$Y_{1t} = c_1 + a_1 Y_{1,t-1} + b_1 Y_{2,t-1} + \varepsilon_{1t},$$

и проверьте гипотезу $b_1 = 0$.

Оцените уравнение

$$Y_{2t} = c_2 + a_2 Y_{1,t-1} + b_2 Y_{2,t-1} + \varepsilon_{2t}$$

и проверьте гипотезу $a_2 = 0$.

Сравните результаты оценивания этих уравнений с результатом оценивания модели $VAR(1)$. Почему полученные результаты совпадают?

Создайте группу из рядов y_1 и y_2 . Нажмите кнопку *View* и выберите: *Granger Causality*. В открывшемся окне установите количество запаздываний, равное порядку авторегрессии, и нажмите *OK*. Получите протокол, имеющий вид таблицы (табл. 3-1).

Таблица 3-1

Pairwise Granger Causality Tests			
Sample: 1 100			
Lags: 1			
Null Hypothesis:	Obs	F-Statistic	Prob.
Y2 does not Granger Cause Y1	98	14.0588	0.0003
Y1 does not Granger Cause Y2		9.57407	0.0026

Сравните результаты, указанные в табл. 3-1, с результатами, полученными при непосредственном оценивании двух уравнений. Как связаны между собой значения F -статистик, приведенные в таблице, и значения t -статистик, использованные при тестировании отдельных уравнений?

Задание 17. Причинность по Грейнджеру в модели VAR с тремя переменными

1. Рассмотрите следующую модель VAR(1) для трех рядов:

$$\begin{aligned} Y_t &= 0.6Y_{t-1} + 0.5Z_{t-1} + \varepsilon_{1t}, \\ X_t &= 0.6Y_{t-1} + 0.25Z_{t-1} + \varepsilon_{2t}, \\ Z_t &= 0.25X_{t-1} + 0.6Z_{t-1} + \varepsilon_{3t}. \end{aligned}$$

Выполнено ли для этой модели условие стабильности? Является ли переменная X G -причиной для переменной Y ?

2. Постройте реализацию указанной модели, полагая $Y_1 = X_1 = Z_1 = 0$ и генерируя ε_{1t} , ε_{2t} и ε_{3t} как независимые между собой процессы гауссовского белого шума с $D(\varepsilon_{1t}) = \dots = D(\varepsilon_{3t}) = 1$. Для каждой пары переменных проверьте гипотезы о том, что одна из переменных не является G -причиной для второй переменной.

Методические указания.

1. Матричный полином $A(L)$ имеет здесь вид:

$$\begin{aligned} A(L) &= I_3 - \Pi_1 L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.6L & 0 & 0.5L \\ 0 & 0.6L & 0.25L \\ 0 & 0.25L & 0.6L \end{pmatrix} = \\ &= \begin{pmatrix} 1-0.6L & 0 & -0.5L \\ 0 & 1-0.6L & -0.25L \\ 0 & -0.25L & 1-0.6L \end{pmatrix}, \end{aligned}$$

$$\det A(L) = (1 - 0.6L)^3 - (1 - 0.6L)0.25^2 L^2 = (1 - 0.6L)(1 - 0.85L)(1 - 0.35L)$$

→ уравнение $\det A(z) = 0$ имеет корни $\frac{1}{0.6}$, $\frac{1}{0.85}$ и $\frac{1}{0.35}$, соответственно

→ модель стабильна.

Из третьего уравнения получаем: $Z_{t-1} = 0.25X_{t-2} + 0.6Z_{t-2} + \varepsilon_{3,t-1}$, так что

$$Y_t = 0.6Y_{t-1} + 0.5Z_{t-1} + \varepsilon_{1t} = 0.6Y_{t-1} + 0.5(0.25X_{t-2} + 0.6Z_{t-2} + \varepsilon_{3,t-1}) + \varepsilon_{1t},$$

$$Y_t = 0.6Y_{t-1} + 0.125X_{t-2} + 0.3Z_{t-2} + \varepsilon_{1t} + 0.5\varepsilon_{3,t-1}.$$

В рамках последнего уравнения прошлые значения переменной X_t помогают в предсказании значения Y_t , так что $x \xrightarrow{G} y$.

2. Построив реализацию указанной модели с помощью объекта *Model*, образуйте объект *Group*, содержащий три смоделированных ряда Y_t , X_t , Z_t . В меню этого объекта выберите:

$$\text{View} \rightarrow \text{Granger Causality} \rightarrow \text{Lags to Include} = 1.$$

Проинтерпретируйте результаты, представленные в полученной таблице. Почему гипотеза о том, что переменная X не является G -причиной для Y , может быть отвергнутой?

Измените глубину запаздываний, полагая $\text{Lags to Include} = 2$. Изменился ли результат проверки гипотезы $x \xrightarrow{G} y$? Если изменился, то почему? (При увеличении глубины запаздываний до двух выявляется влияние значения X_{t-2} .)

Проверьте ту же гипотезу, не прибегая к встроенной проверке, а оценивая непосредственно уравнение

$$Y_t = c(1)Y_{t-1} + c(2)Y_{t-2} + c(3)X_{t-1} + c(4)X_{t-2} + \varepsilon_t.$$

и проверяя гипотезу $c(3) = c(4) = 0$. Совпало ли значение полученной F -статистики со значением, приведенным в таблице, полученной в рамках объекта *Group*? Если не совпало, то почему? (Это связано с включением или невключением в правую часть уравнения для проверки гипотезы постоянной составляющей.)

Задание 18. Блочная экзогенность в модели VAR с тремя переменными

Рассматривается модель $VAR(2)$ для переменных y_{1t}, y_{2t}, y_{3t} :

$$y_t = \mu + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \varepsilon_t,$$

где $y_t = (y_{1t}, y_{2t}, y_{3t})^T$, $\mu = (\mu_1, \mu_2, \mu_3)^T$, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t})^T$,

$\Pi_r = (\pi_{ij,r})$ — (3×3) -матрица коэффициентов при $y_{1,t-r}, y_{2,t-r}, y_{3,t-r}$ в трех уравнениях.

Матрицы Π_r имеют вид:

$$\Pi_1 = \begin{pmatrix} \pi_{11,1} & \pi_{12,1} & \pi_{13,1} \\ 0 & \pi_{22,1} & 0 \\ 0 & \pi_{32,1} & 0 \end{pmatrix}, \quad \Pi_2 = \begin{pmatrix} 0 & 0 & \pi_{13,2} \\ 0 & \pi_{22,2} & 0 \\ \pi_{31,2} & 0 & \pi_{33,2} \end{pmatrix}.$$

Какие из переменных (группы переменных) являются блочно экзогенными в отношении остальных переменных?

Задание 19. Проверка причинности по Грейнджеру и блочной экзогенности в модели VAR (открытая VAR)

1. Постройте реализацию открытой VAR:

$$Y_{1t} = 0.6 + 0.7Y_{1,t-1} + 0.2Y_{2,t-1} + 0.1X_{1,t-1} + 0.2X_{2t} + \varepsilon_{1t},$$

$$Y_{2t} = 0.4 + 0.2Y_{1,t-1} + 0.7Y_{2,t-1} + 0.2X_{1t} + 0.4X_{2,t-1} + \varepsilon_{2t},$$

где

$$X_{1t} = 0.7X_{1,t-1} + \varepsilon_{3t}, \quad X_{11} = 0,$$

$$X_{2t} = 0.5X_{1,t-1} + \varepsilon_{4t}, \quad X_{21} = 0,$$

и в правых частях уравнений стоят независимые между собой процессы гауссовского белого шума с $D(\varepsilon_{1t}) = \dots = D(\varepsilon_{4t}) = 0.01$.

Положите $Y_{11} = Y_{21} = 0$. Объясните поведение полученных реализаций.

2. Как надо изменить при моделировании начальные значения Y_{11} и Y_{21} , чтобы система быстрее выходила на стабильный режим?
3. Используя смоделированные реализации с начальными значениями, соответствующими стабильному режиму, проверьте гипотезу о том, что переменные Y_{1t} и Y_{2t} не являются причиной по Грейнджеру для переменной X_{1t} . Проверьте гипотезу о том, что переменные Y_{1t} и Y_{2t} не являются причиной по Грейнджеру для переменной X_{2t} .
4. Проверьте гипотезу о блочной экзогенности переменных X_{1t} и X_{2t} .

Методические указания.

1. Проверьте выполнение условия стабильности модели.
2. Найдите стабильное решение и на его основе вычислите необходимые значения Y_{11} и Y_{21} , учитывая, что $X_{11} = 0$, $X_{21} = 0$.

Здесь μ и матричный полином $A(L)$ — те же, что и ранее, а

$$B(L) = B_0 + B_1L = \begin{pmatrix} 0 & 0.2 \\ 0.2 & 0 \end{pmatrix} + \begin{pmatrix} 0.1 & 0 \\ 0 & 0.4 \end{pmatrix} L = \begin{pmatrix} 0.1L & 0.2 \\ 0.2 & 0.4L \end{pmatrix},$$

так что

$$B(1) = B_0 + B_1 = \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \end{pmatrix}.$$

Матрица долгосрочных мультипликаторов равна:

$$C(1) = A^{-1}(1)B(1) = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \end{pmatrix} = \begin{pmatrix} 1.4 & 2.8 \\ 1.6 & 3.2 \end{pmatrix},$$

так что стабильное решение есть

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5.2 \\ 4.8 \end{pmatrix} + \begin{pmatrix} 1.4 & 2.8 \\ 1.6 & 3.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

т.е.

$$y_1 = 5.2 + 1.4x_1 + 2.8x_2,$$

$$y_2 = 4.8 + 1.6x_1 + 3.2x_2.$$

Смоделируйте реализации этой открытой системы в случае, когда x_{1t} и x_{2t} — независимые друг от друга $AR(1)$ -ряды,

$$x_{1,t} = 0.7x_{1,t-1} + v_{1t}, \quad x_{2,t} = 0.5x_{2,t-1} + v_{2t}, \quad v_{1t} \text{ и } v_{2t} \sim i.i.d. N(0, 1).$$

В качестве начальных значений при моделировании возьмите:

$$x_{11} = x_{21} = 0, \quad y_{11} = 5.2, \quad y_{21} = 4.8.$$

3. Создайте объект *System* под именем *sys01* и специфицируйте ее как модель $VAR(1)$ с 4 переменными X_{1t} , X_{2t} , Y_{1t} , Y_{2t} .

$$X1 = C(1) + c(2)*x1(-1) + c(3)*x2(-1) + c(4)*y1(-1) + c(5)*y2(-1)$$

$$X2 = C(6) + c(7)*x1(-1) + c(8)*x2(-1) + c(9)*y1(-1) + c(10)*y2(-1)$$

$$Y1 = C(11) + c(12)*x1(-1) + c(13)*x2(-1) + c(14)*y1(-1) + c(15)*y2(-1)$$

$$Y2 = C(16) + c(17)*x1(-1) + c(18)*x2(-1) + c(19)*y1(-1) + c(20)*y2(-1)$$

Оцените эту систему методом наименьших квадратов по смоделированным ранее данным. Создайте объект *Var* под именем *var1* и специфицируйте ее как модель $VAR(1)$ с теми же переменными, оцените эту модель. Сравните результаты оценивания системы *sys01* и модели *var1*. (Они совпадают.) Проведите диагностику остатков в оцененной модели *var1*. Подтвердите правильность выбора порядка модели, сравнивая ее с моделями более высокого порядка.

Для проверки гипотезы о том, что переменные Y_{1t} и Y_{2t} не являются причиной по Грейнджеру для переменной X_{1t} , оцените уравнение

$$X_{1t} = m1 + a1*Y_{1,t-1} + b1*Y_{2,t-1} + c1*X_{1,t-1} + d1*X_{2,t-1} + \varepsilon_{1t}$$

и проверьте гипотезу $a1 = b1 = 0$.

3. Для проверки гипотезы о том, что переменные Y_{1t} и Y_{2t} не являются причиной по Грейнджеру для переменной X_{2t} , оцените уравнение

$$X_{2t} = m2 + a2*Y_{1,t-1} + b2*Y_{2,t-1} + c2*X_{1,t-1} + d2*X_{2,t-1} + \varepsilon_{2t}$$

и проверьте гипотезу $a2 = b2 = 0$.

4. Для проверки гипотезы о блочной экзогенности переменных X_{1t} и X_{2t} можно воспользоваться встроенной процедурой в рамках объекта *sys01*:

View → *Coefficient tests* → *Wald Coefficient Test*,

задав соответствующее условие на коэффициенты:

$$c(4) = c(5) = c(9) = c(10) = 0.$$

Задание 20. Ложная причинность по Грейнджеру

1. Рассмотрите модель $VAR(1)$ для двух рядов:

$$Y_{1t} = Y_{1,t-1} + \varepsilon_{1t},$$

$$Y_{2t} = Y_{2,t-1} + \varepsilon_{2t}.$$

Выполнено ли условие стабильности для этой VAR ?

2. Постройте реализацию заданной модели, полагая $Y_{11} = Y_{21} = 0$ и $D(\varepsilon_{1t}) = D(\varepsilon_{2t}) = 0.01$, $Cov(\varepsilon_{1t}, \varepsilon_{2t}) = 0$. Объясните поведение полученных реализаций.
3. По полученной реализации модели проверьте гипотезы о том, что одна из переменных не является причиной по Грейнджеру для другой переменной.

Методические указания.

1. В компактной форме указанная $VAR(1)$ имеет вид:

$$y_t = \Pi_1 y_{t-1} + \varepsilon_t,$$

где

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \quad \Pi_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

или

$$A(L)y_t = \varepsilon_t,$$

где

$$A(L) = I_2 - \Pi_1 L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix} = \begin{pmatrix} 1-L & 0 \\ 0 & 1-L \end{pmatrix},$$

так что

$$A(1) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \det A(1) = 0.$$

Уравнение $\det A(z) = 0$ принимает здесь вид:

$$\det A(z) = \begin{vmatrix} 1-z & 0 \\ 0 & 1-z \end{vmatrix} = 0,$$

т.е. $(1 - z^2) = 0$. Оба корня этого уравнения равны 1, условие стабильности не выполняется.

3. Результаты проверки указанных гипотез в смоделированной модели приведены в табл. 3-2.

Таблица 3-2

<i>Pairwise Granger Causality Tests</i>			
<i>Sample:</i> 1 100			
<i>Lags:</i> 1			
<i>Null Hypothesis:</i>	<i>Obs</i>	<i>F-Statistic</i>	<i>Prob.</i>
<i>Y2 does not Granger Cause Y1</i>	98	0.10520	0.7464
<i>Y1 does not Granger Cause Y2</i>		5.93519	0.0167

Переменная Y_1 признается причиной по Грейнджеру для переменной Y_2 , хотя ряды значений этих переменных генерировались независимо друг от друга.

Задание 21. Некоинтегрированная VAR: причинность в краткосрочном плане

1. Рассмотрите модель VAR(1) для рядов разностей:

$$\Delta Y_{1t} = 0.6Y_{1,t-1} + \varepsilon_{1t},$$

$$\Delta Y_{2t} = 0.4Y_{1,t-1} + 0.6\Delta Y_{2,t-1} + \varepsilon_{2t}.$$

Является ли такая VAR в разностях стабильной? Как выглядит VAR для уровней, какой порядок она имеет? Имеется ли коинтеграция между рядами Y_{1t} и Y_{2t} ?

2. В табл. П-7 приведена реализация указанной VAR. Проверьте полученную пару рядов Y_{1t} и Y_{2t} на коинтегрированность.
3. Рассмотрите вопрос о наличии G-причинности одной из переменных в отношении другой переменной в краткосрочном плане.
4. Подтвердите на основании одних только сгенерированных данных правильность выбора порядка VAR при оценивании.

Методические указания.

1. VAR в разностях стабильна (выполнено условие стабильности). VAR в уровнях имеет порядок 2. Коинтеграция между рядами Y_{1t} и Y_{2t} отсутствует (оба уравнения в разностях не имеют в правых частях корректирующей составляющей).
2. Для проверки на коинтегрированность создайте из рядов y_1 и y_2 объект *Group* и используйте встроенную процедуру проверки. Результаты приведены в табл. 3-3. Ряды признаются некоинтегрированными.

Таблица 3-3

Selected (0.05 level*) Number of Cointegrating Relations by Model					
Data Trend:	None	None	Linear	Linear	Quadratic
Test Type	No Intercept	Intercept	Intercept	Intercept	Intercept
	No Trend	No Trend	No Trend	Trend	Trend
Trace	0	0	0	0	0
Max-Eig	0	0	0	0	0
* Critical values based on MacKinnon-Haug-Michelis (1999).					

3. Оцените модель VAR(1) для разностей. Результаты оценивания приведены в табл. 3-4.

Выявляется односторонняя краткосрочная G-причинность в направлении от Y_1 к Y_2 . Проверьте наличие такой G-причинности, используя объект *Group* для разностей. Результаты приведены в табл. 3-5.

Таблица 3-4

	$D(Y1)$	$D(Y2)$
$D(Y1(-1))$	0.623747 (0.081340) [7.668040]	0.310236 (0.091520) [3.389720]
$D(Y2(-1))$	0.083872 (0.074210) [1.130120]	0.486420 (0.083500) [5.825250]
C	-0.026739 (0.102970) [-0.259670]	0.108625 (0.115860) [0.937560]

Таблица 3-5

Pairwise Granger Causality Tests			
Sample: 1 100			
Lags: 1			
Null Hypothesis:	Obs	F-Statistic	Prob.
$D(Y2)$ does not Granger Cause $D(Y1)$	98	1.27717	0.2613
$D(Y1)$ does not Granger Cause $D(Y2)$		11.49020	0.0010

Результат тот же. Сравните значения F -статистик в табл. 3-5 с соответствующими значениями t -статистик в табл. 3-4. Какая между ними связь?

4. Оценив на основе сгенерированных данных $VAR(1)$ для разностей, воспользуйтесь процедурой выбора порядка VAR , следуя цепочке

View → Lag Structure → Lag Length Criteria.

Из 5 критериев выбора 3 указывают на модель первого порядка (табл. 3-6).

Таблица 3-6

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-312.7545	NA	3.462284	6.917681	6.972865	6.939944
1	-266.6040	89.258110*	1.371068	5.991296	6.156847*	6.058086*
2	-262.3325	8.073582	1.363148*	5.985330*	6.261248	6.096646
3	-261.0224	2.418627	1.446726	6.044448	6.430735	6.200291
4	-259.1233	3.422617	1.516135	6.090621	6.587275	6.290990
5	-257.9949	1.983899	1.616650	6.153735	6.760756	6.398630
6	-256.5517	2.474069	1.712766	6.209928	6.927316	6.499349
7	-253.4501	5.180665	1.750655	6.229673	7.057429	6.563621
8	-252.6215	1.347747	1.882307	6.299373	7.237496	6.677847

Задание 22. *Коинтегрированная VAR: причинность в краткосрочном и долгосрочном плане*

В табл. П-8 приведены данные об объемах продаж медицинских препаратов, произведенных фармацевтической фирмой *Lydia Pinkham (Sales)*, и о расходах фирмы на рекламу этих препаратов (*Adver*) в период с 1907 по 1960 г. Учитывая исследования различных авторов, для анализа этих данных перейдите к логарифмам уровней.

1. Определите порядки интегрированности этих рядов, проверьте наличие коинтегрированности между этими рядами.
2. Оцените долговременное соотношение между рядами и модель коррекции ошибок. Проведите анализ на наличие/отсутствие долгосрочной и краткосрочной причинности между рассматриваемыми рядами.

Задание 23. *Коинтегрированная VAR: причинность в краткосрочном и долгосрочном плане*

В задании 22 тестирование было проведено для пары рядов *Adver* и *Sales*, при этом не принималось в расчет возможное влияние других рядов. С целью учета такого влияния протестируйте наличие или отсутствие соотношений $Adver \rightarrow Sales$, $Sales \rightarrow Adver$ (в смысле предшествования по Грейнджеру) с дополнительным включением в тестовые уравнения совокупного располагаемого личного дохода *Income*. Примените к этой тройке рядов подход Тода — Ямамото.

Задание 24. *Структурная VAR, функции импульсного отклика*

1. Докажите стабильность следующей $VAR(1)$:

$$y_{1t} = \frac{5}{8}y_{1,t-1} + \frac{1}{2}y_{2,t-1} + u_{1t},$$

$$y_{2t} = \frac{1}{4}y_{1,t-1} + \frac{5}{8}y_{2,t-1} + u_{2t}.$$

Методические указания. Запишите систему в виде

$$A(L)y_t = u_t$$

и проверьте выполнение условия стабильности в отношении корней уравнения $\det A(z) = 0$.

2. Полагая $\Sigma = Cov(u_t) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, выразите u_{1t} и u_{2t} через фундаментальные

инновации, используя два варианта упорядочения. Для каждого упорядочения составьте модель структурной VAR .

3. В каждом варианте упорядочения фундаментальных инноваций вычислите значения функций импульсного отклика на единичный шок инноваций для моментов времени $t = 1$ и $t = 2$.
4. Используя пакет *EViews*, постройте последовательности импульсных откликов на единичные шоки инноваций в каждом из вариантов упоря-

дочения. Сравните полученные первые два отклика со значениями, вычисленными в п. 3.

Методические указания.

- Сгенерируйте две последовательности $eps1$ и $eps2$, имитирующие гауссовский белый шум. Убедитесь в том, что обе эти последовательности проходят диагностические тесты.
- Создайте объект *Model* под именем m_orig , специфицируйте соответствующую *VAR*, используя сгенерированные последовательности, и получите реализацию этой *VAR* (ряды y_1 и y_2).
- На основе ряда $eps1$ (не изменяя сам ряд $eps1$) образуйте новый ряд $eps1_modif$, увеличив первое значение ряда $eps1$ на 1.
- Создайте новый объект *Model* под именем m_modif , аналогичный предыдущему, но использующий вместо ряда $eps1$ ряд $eps1_modif$. Получите реализацию этой модели (ряды $y1_modif$ и $y2_modif$).
- Постройте последовательности импульсных откликов на единичный шок инноваций ε_{1t} в виде рядов

$$delta1 = y1_modif - y1 \quad \text{и} \quad delta2 = y2_modif - y2.$$

Сравните полученные первые два отклика со значениями, вычисленными в п. 3.

- Постройте графики функций импульсного отклика.
5. Используя смоделированные ряды $y1$ и $y2$, оцените модель *VAR*(1) для этих рядов. Сравните коэффициенты оцененной модели с коэффициентами модели, использованной в модели порождения этих рядов. Обратите внимание на доверительные интервалы для (отдельных!) значений импульсного отклика.

Методические указания. Для получения значений функций импульсных откликов и графиков этих функций в оцененной *VAR* используйте переход: *View/Impulse Response*. Для получения таблицы значений укажите *Display Format: Table*. Для получения графиков функций вместе с доверительными интервалами укажите: *Multiple Graphs*, а также: *Response Standard Errors: Analytic* или *Monte-Carlo*. При использовании первого из двух указанных выше упорядочений во вкладке *Impulse Definition* в окне *Cholesky Ordering* укажите порядок вхождения рядов: $y1 \ y2$. При использовании второго из двух указанных выше упорядочений, во вкладке *Impulse Definition* в окне *Cholesky Ordering* укажите порядок вхождения рядов: $y2 \ y1$.

6. Используя оцененную *VAR*, постройте таблицы и графики декомпозиций дисперсий ошибок прогнозов при каждом из двух упорядочений.

Методические указания. Для получения таблицы и графиков декомпозиций дисперсий ошибок прогнозов в оцененной *VAR* используйте переход: *View/Variance Decomposition*. Далее используйте методические указания к п. 5.

Задание 25. Построение функций импульсного отклика
и декомпозиций дисперсий ошибок прогнозов для модели *VAR*,
оцененной на основании имеющихся статистических данных

В рабочем файле **lszusa_data.wf1** приведены статистические данные о следующих показателях экономики США:

- CPI_SA* — индекс потребительских цен (сезонно-скорректированный);
RGDP_MON — реальный валовой внутренний продукт;
M2_SA — денежный агрегат M2 (сезонно-скорректированный);
TBILL3 — процентные ставки по 3-месячным казначейским обязательствам;
PCM — (*commodity price level*) — индекс цен на инвестиционные товары, не включающий нефть.

Файл **lszusa_data.wf1** можно получить, зайдя на следующий интернет-адрес: <http://didattica.unibocconi.eu/myigier/index.php?IdUte=48917&idr=10020&lingua=eng&comando=Apri>. Далее выбрать: *Data and Exercises for Chapter 4* и скачать предлагаемый ZIP-файл, из которого можно извлечь файл **lszusa_data.wf1**.

При построении моделей *VAR* все переменные, за исключением процентных ставок, возьмите в логарифмах уровней: $LP = 100\ln(CPI_SA)$, $LY = 100\ln(RGDP_MON)$, $LM2 = 100\ln(M2_SA)$, $LPCM = 100\ln(PCM)$.

1. Следуя работе (Leeper, Sims, Zha, 1966), оцените на периоде с 1960:01 по 1996:03 модель *VAR*(6) для переменных *LP*, *LY* и *LM2*. Постройте функции импульсного отклика для оцененной системы при упорядочении $P \rightarrow Y \rightarrow M2$, в котором наиболее эндогенной переменной предполагаются деньги. Подтверждается ли эндогенность денежного агрегата характером этих функций? Постройте декомпозиции дисперсий ошибок прогнозов для оцененной системы. Подтверждается ли эндогенность денежного агрегата поведением этих дисперсий?
2. Примените упорядочение, использованное Симсом: $M2 \rightarrow Y \rightarrow P$. Сравните результаты полученные при использованных двух упорядочениях. Объясните результаты сравнения.
3. Дополните модель *VAR*(6) переменной *TBILL3* и оцените расширенную модель на том же периоде. Постройте функции импульсных откликов и декомпозиции дисперсий ошибок прогнозов для оцененной системы, используя упорядочение $TBILL3 \rightarrow M2 \rightarrow Y \rightarrow P$. Обнаруживается ли здесь эндогенность *M2* по отношению к *CPI* и *GDP*? Обнаруживается ли эндогенность *M2* по отношению к *TBILL3*? Наблюдаются ли в поведении функций импульсного отклика «загадка ликвидности» и «загадка цен»? Чем может быть вызвано такое поведение функций импульсного отклика?
4. Дополните модель *VAR*(6), использованную в п. 3, переменной *LPCM* и оцените расширенную модель на том же периоде. Постройте функции импульсного отклика и декомпозиции дисперсий ошибок прогнозов для оцененной системы, используя упорядочение $TBILL3 \rightarrow PCM \rightarrow M2 \rightarrow P \rightarrow Y$. Сравните результаты, полученные для моделей *VAR*(6) с 4 и 5 переменными.

Задание 26. *Оценивание модели стохастической границы производственных возможностей*

Методические указания. В пакете *Stata* откройте файл **frontier1.dta** и выберите в главном меню: *Statistics* → *Linear regression and related* → *Frontier models*. Укажите в соответствующих окнах объясняемую (*lnoutput*) и объясняющие (*lncapital* и *lnlabor*) переменные, а также выбранное распределение для u_i (*Truncated normal distribution*). Это равносильно выполнению команды *frontier lnoutput lncapital lnlabor, distribution (tnormal)*. Результаты выполнения этой команды отражены в табл. 3-7.

Таблица 3-7

. frontier lncapital lnlabor, distribution (normal)						
Stoc. frontier normal/truncated-normal model				Number of obs = 756		
				Wald chi2(2) = 8.84		
Log likelihood = -1506.9346				Prob > chi2 = 0.0121		
lnoutput	Коэффициент	Стандартная ошибка	z-статистика	P-значение P > z	[95% Conf. Interval]	
lncapital	0.6029056	0.5703415	1.06	0.290	-0.5149432	1.720754
lnlabor	0.6739399	0.2417019	2.79	0.005	0.2002130	1.147667
_cons	0.1262186	3.296878	0.04	0.969	-6.3355440	6.587981
/mu	-0.7811509	2.357352	-0.33	0.740	-5.401476	3.839174
/lnsigma2	2.22012	0.3819963	5.81	0.000	1.471421	2.968819
/ilgtgamma	2.259689	0.292037	7.74	0.000	1.687307	2.832071
sigma2	9.208434	3.517588			4.355419	19.46891
gamma	0.905483	0.0249936			0.8438696	0.9443845
sigma_u2	8.33808	3.319079			1.832805	14.84335
sigma_v2	0.8703535	0.2794375			0.3226661	1.418041
H0: No inefficiency component: z = -8.194 Prob <= z = 0.0000						

Здесь

sigma_v2 — оценка для σ_v^2 ;

sigma_u2 — оценка для σ_u^2 ;

gamma — оценка для $\gamma = \frac{\sigma_u^2}{\sigma_s^2}$;

sigma2 — оценка для $\sigma_s^2 = \sigma_v^2 + \sigma_u^2$.

Поскольку значение γ должно заключаться в пределах от 0 до 1, оптимизация производится для обратного логита от γ , и соответствующая оценка приводится в таблице как *ilgtgamma*. Поскольку значение σ_s^2 должно быть положительным, оптимизация проводится для $\ln(\sigma_s^2)$, и соответствующая оценка приводится в таблице как *lnsigma2*. Наконец, *mu* есть оценка для μ . Последняя строка таблицы указывает на результаты проверки гипотезы $\sigma_u^2 = 0$.

Поскольку оценка для μ статистически незначима, вместо усеченного нормального распределения (с произвольной точкой усечения) можно использовать полунормальное распределение, т.е. использовать выбор: *Half-normal*.

После оценивания модели оценки для u_i вычисляются с использованием команды *predict* и формулы для условного среднего $E(u_i | \varepsilon_i)$ (опция *u*) или формулы для условной медианы $Med(u_i | \varepsilon_i)$ (опция *m*). Применяя опцию *te*, получаем оценки технической эффективности, опирающиеся на формулу для $E(\exp(u_i) | \varepsilon_i)$.

Задание 27. *Анализ качества оценивания параметров модели стохастической границы производственных возможностей и технической эффективности*

Проведите анализ качества оценивания параметров модели и технической эффективности по образцу примера 8.1.2.

Задание 28. *Оценивание модели стохастической границы для панельных данных: модель, инвариантная по времени*

Проведите оценивание модели стохастической границы по образцу примера 8.2.1, используя статистические данные из файла **xtfrontier1.dta** в пакете *Stata 10* (<http://www.stata-press.com/data/r8/xtfrontier1.dta>).

Задание 29. *Оценивание модели стохастической границы для панельных данных: модель с затуханием во времени*

Проведите оценивание модели стохастической границы по образцу примера 8.2.2, используя те же статистические данные, что и в задании 28.

Приложение

ТАБЛИЦЫ СТАТИСТИЧЕСКИХ ДАННЫХ К ЗАДАНИЯМ

Таблица П-1

**Значения индекса ММББ (Московской межбанковской валютной биржи)
в период с 18 сентября 2007 г. по 14 февраля 2008 г. (ряд IND)**

<i>t</i>	IND	<i>t</i>	IND	<i>t</i>	IND	<i>t</i>	IND	<i>t</i>	IND
1	1669.06	21	1821.53	41	1878	61	1969.91	81	1788.34
2	1730.74	22	1827.74	42	1855.98	62	1928.99	82	1654.83
3	1716.60	23	1799.93	43	1833.28	63	1895.25	83	1636.79
4	1725.11	24	1801.42	44	1805.84	64	1887.61	84	1570.90
5	1736.37	25	1772.35	45	1816.00	65	1908.63	85	1661.99
6	1703.72	26	1799.36	46	1785.39	66	1899.42	86	1705.20
7	1744.91	27	1792.05	47	1776.84	67	1926.45	87	1635.33
8	1755.48	28	1816.90	48	1793.03	68	1931.20	88	1678.42
9	1759.44	29	1854.47	49	1816.85	69	1929.31	89	1642.63
10	1733.39	30	1872.00	50	1791.89	70	1918.75	90	1574.33
11	1795.06	31	1855.63	51	1813.09	71	1904.27	91	1640.49
12	1778.66	32	1874.73	52	1822.59	72	1907.39	92	1677.78
13	1777.82	33	1866.49	53	1850.64	73	1888.86	93	1632.37
14	1802.45	34	1879.41	54	1838.78	74	1906.86	94	1631.10
15	1808.42	35	1907.63	55	1834.73	75	1912.29	95	1581.96
16	1827.99	36	1903.55	56	1895.39	76	1914.48	96	1568.88
17	1817.97	37	1917.29	57	1891.16	77	1931.14	97	1605.10
18	1848.82	38	1891.27	58	1913.38	78	1921.38	98	1665.47
19	1835.21	39	1890.01	59	1945.53	79	1839.04	99	1688.30
20	1839.38	40	1868.62	60	1951.38	80	1805.42	100	1687.90

Таблица П-2

**Квартальные данные об индексе реального объема
сельскохозяйственного производства (1993:01 = 100) за 12-летний период
с I квартала 1994 г. по IV квартал 2005 г. (ряд AGRO)**

Год/ квартал	AGRO	Год/ квартал	AGRO	Год/ квартал	AGRO	Год/ квартал	AGRO
1994/1	87.0	1997/1	69.1	2000/1	70.3	2003/1	76.5
1994/2	120.1	1997/2	98.3	2000/2	101.9	2003/2	107.8
1994/3	205.0	1997/3	192.3	2000/3	176.2	2003/3	192.6
1994/4	115.7	1997/4	105.6	2000/4	100.8	2003/4	121.9

Окончание табл. II-2

Год/ квартал	AGRO	Год/ квартал	AGRO	Год/ квартал	AGRO	Год/ квартал	AGRO
1995/1	77.4	1998/1	68.8	2001/1	71.7	2004/1	76.1
1995/2	106.9	1998/2	99.2	2001/2	105.5	2004/2	107.8
1995/3	192.7	1998/3	149.8	2001/3	196.6	2004/3	202
1995/4	106.4	1998/4	90.7	2001/4	105.0	2004/4	125.4
1996/1	73.4	1999/1	66.7	2002/1	75.8	2005/1	76
1996/2	101.7	1999/2	98.9	2002/2	109.1	2005/2	108.6
1996/3	178.2	1999/3	158	2002/3	192.6	2005/3	210.8
1996/4	105.3	1999/4	97	2002/4	113.5	2005/4	126.7

Таблица II-3

**Квартальные данные об импорте товаров и услуг
в Российскую Федерацию за 12-летний период с I квартала 1994 г.
по IV квартал 2005 г., млрд долл. (ряд IMP)**

Год/ квартал	IMP	Год/ квартал	IMP	Год/ квартал	IMP	Год/ квартал	IMP
1994/1	11.3	1997/1	15.2	2000/1	10.000	2003/1	15.982
1994/2	12.1	1997/2	17.2	2000/2	10.400	2003/2	18.200
1994/3	12.5	1997/3	18.4	2000/3	11.100	2003/3	19.573
1994/4	14.5	1997/4	21.2	2000/4	13.400	2003/4	22.314
1995/1	13.4	1998/1	17.8	2001/1	11.300	2004/1	19.632
1995/2	14.6	1998/2	17.4	2001/2	13.600	2004/2	22.725
1995/3	15.8	1998/3	13.4	2001/3	13.238	2004/3	24.678
1995/4	18.8	1998/4	9.4	2001/4	15.619	2004/4	29.272
1996/1	16.4	1999/1	9.1	2002/1	12.347	2005/1	24.781
1996/2	17.2	1999/2	10.1	2002/2	14.768	2005/2	29.310
1996/3	17.0	1999/3	9.5	2002/3	15.725	2005/3	31.507
1996/4	17.5	1999/4	10.8	2002/4	18.125	2005/4	39.500

Таблица II-4

**Квартальные данные о количестве безработных в России
за период с I квартала 1994 г. по IV квартал 1997 г., млн человек (ряд UNEMP)**

Год/ квартал	UNEMP	Год/ квартал	UNEMP	Год/ квартал	UNEMP	Год/ квартал	UNEMP
1994/1	5.0	1995/1	5.8	1996/1	6.7	1997/1	7.5
1994/2	5.5	1995/2	6.0	1996/2	7.0	1997/2	7.9
1994/3	5.7	1995/3	6.4	1996/3	7.1	1997/3	7.9
1994/4	5.7	1995/4	6.6	1996/4	7.2	1997/4	8.1

Таблица П-5

**Квартальные данные о размерах совокупной прибыли
нефинансовых корпораций США (до уплаты налогов)
за период с I квартала 1985 г. по I квартал 1989 г.,
млрд долл. (ряд *PROFITS*)**

Год/ квартал	<i>PROFITS</i>		Год/ квартал	<i>PROFITS</i>		Год/ квартал	<i>PROFITS</i>		Год/ квартал	<i>PROFITS</i>
1985/1	168.7		1986/1	155.9		1987/1	196.6		1988/1	228.4
1985/2	162.1		1986/2	167.2		1987/2	207.9		1988/2	240.5
1985/3	176.0		1986/3	176.2		1987/3	224.6		1988/3	240.4
1985/4	174.2		1986/4	191.0		1987/4	211.6		1988/4	246.6
									1989/1	246.3

Таблица П-6

**Квартальные данные об объеме валового
национального продукта США за период с I квартала 1947 г.
по IV квартал 1961 г. (ряд *GNP*)**

Год/ квартал	<i>GNP</i>		Год/ квартал	<i>GNP</i>		Год/ квартал	<i>GNP</i>		Год/ квартал	<i>GNP</i>
1947/1	224		1951/1	318		1955/1	386		1959/1	474
1947/2	228		1951/2	326		1955/2	394		1959/2	487
1947/3	232		1951/3	333		1955/3	403		1959/3	484
1947/4	242		1951/4	337		1955/4	409		1959/4	491
1948/1	248		1952/1	340		1956/1	411		1960/1	503
1948/2	256		1952/2	339		1956/2	416		1960/2	505
1948/3	263		1952/3	346		1956/3	421		1960/3	504
1948/4	264		1952/4	358		1956/4	430		1960/4	503
1949/1	259		1953/1	364		1957/1	437		1961/1	504
1949/2	255		1953/2	368		1957/2	440		1961/2	515
1949/3	257		1953/3	366		1957/3	446		1961/3	524
1949/4	255		1953/4	361		1957/4	442		1961/4	538
1950/1	266		1954/1	361		1958/1	435			
1950/2	275		1954/2	360		1958/2	438			
1950/3	293		1954/3	365		1958/3	451			
1950/4	305		1954/4	373		1958/4	464			

Таблица П-7

Смоделированные данные к заданию 21 (ряды Y1 и Y2)

t	Y1	Y2	t	Y1	Y2	t	Y1	Y2
1	0	0	34	-31.1633	3.842037	67	-20.46560	21.83319
2	0	0	35	-32.4059	3.804359	68	-19.25990	23.02866
3	-1.15231	1.058442	36	-33.1520	3.878898	69	-17.36810	23.27426
4	-2.97469	1.715719	37	-34.3600	3.214015	70	-16.10600	25.48272
5	-4.64481	2.783577	38	-35.2051	1.961030	71	-16.64920	27.43664
6	-4.45392	3.527210	39	-35.1501	0.524794	72	-15.59800	26.80269
7	-4.12834	4.889087	40	-36.8966	0.638066	73	-14.86250	27.95718
8	-6.26758	4.961772	41	-37.8199	-1.100460	74	-13.28590	27.49484
9	-7.26253	3.295569	42	-38.8979	-2.196450	75	-11.39750	27.77684
10	-9.35684	3.129074	43	-40.9594	-1.942150	76	-8.53231	28.63911
11	-12.41120	2.559852	44	-40.6722	-2.531830	77	-8.31045	30.90254
12	-15.65880	-1.083640	45	-40.4914	-3.010720	78	-7.71818	30.72931
13	-18.85520	-3.059580	46	-39.0452	-3.740140	79	-6.29836	29.64129
14	-20.99170	-4.042540	47	-38.9890	-2.226380	80	-5.72012	28.80901
15	-21.10220	-4.906840	48	-39.4057	-1.989900	81	-5.09867	27.57486
16	-19.36630	-2.321830	49	-38.7585	-1.803700	82	-6.23828	25.64758
17	-19.66180	-0.159230	50	-38.5549	-0.332440	83	-6.01278	23.45752
18	-18.52530	1.425131	51	-38.0779	1.796697	84	-5.25983	22.14006
19	-19.10790	1.405937	52	-37.5275	3.025768	85	-5.05585	21.27268
20	-19.55680	-0.135950	53	-35.5974	3.217876	86	-5.76288	21.45392
21	-20.48270	-1.424450	54	-35.2090	4.604728	87	-6.29981	24.52842
22	-21.09230	-3.198520	55	-35.5308	4.460981	88	-7.45207	26.19114
23	-21.92370	-4.260850	56	-36.1209	3.747267	89	-8.83052	26.27494
24	-22.20870	-3.714910	57	-35.6420	4.291034	90	-7.82135	24.91165
25	-20.91330	-2.077430	58	-34.4174	7.291425	91	-7.95053	23.76549
26	-20.50610	0.839003	59	-33.0686	8.719487	92	-9.48873	23.70169
27	-21.24410	2.773934	60	-31.7113	10.314190	93	-10.33840	23.60316
28	-22.23470	4.220075	61	-29.4035	8.320796	94	-9.55520	22.53140
29	-22.69870	3.671827	62	-28.0767	9.224016	95	-8.23511	20.90438
30	-24.02260	1.207308	63	-26.2872	10.752960	96	-7.28981	19.60656
31	-25.70890	1.276263	64	-25.0938	13.116630	97	-6.52929	18.59868
32	-27.59920	1.841462	65	-23.2852	15.981520	98	-6.21592	18.05265
33	-28.68550	2.622085	66	-23.1385	18.914750	99	-5.21268	18.72333
						100	-4.30148	18.13552

Таблица II-8

**Данные об объемах продаж медицинских препаратов,
произведенных фармацевтической фирмой *Lydia Pinkham* (SALES),
о расходах фирмы на рекламу этих препаратов (ADVER)
и о совокупном располагаемом личном доходе (INC)**

Год	SALES	ADVER	INC	Год	SALES	ADVER	INC	Год	SALES	ADVER	INC
1907	1016	608	—	1925	3438	1800	73.0	1943	2602	1164	133.5
1908	921	451	29.5	1926	2917	1941	77.4	1944	2518	1102	146.8
1909	934	529	30.2	1927	2359	1229	77.4	1945	2637	1145	150.4
1910	976	543	30.5	1928	2240	1373	77.5	1946	2177	1012	160.6
1911	930	525	31.9	1929	2196	1611	83.1	1947	1920	836	170.1
1912	1052	549	33.9	1930	2111	1568	74.4	1948	1910	941	189.3
1913	1184	525	34.8	1931	1806	983	63.8	1949	1984	981	189.7
1914	1089	578	35.8	1932	1644	1046	48.7	1950	1787	974	207.7
1915	1087	609	40.2	1933	1814	1453	45.7	1951	1689	766	227.5
1916	1154	504	47.8	1934	1770	1504	52.0	1952	1866	920	238.7
1917	1330	752	55.2	1935	1518	807	58.3	1953	1896	964	252.5
1918	1980	613	62.3	1936	1103	339	66.2	1954	1684	811	256.9
1919	2223	862	63.3	1937	1266	562	71.0	1955	1633	789	274.4
1920	2203	866	71.5	1938	1473	745	65.7	1956	1657	802	292.9
1921	2514	1016	60.2	1939	1423	749	70.4	1957	1569	770	308.8
1922	2726	1360	60.3	1940	1767	862	76.1	1958	1390	639	317.9
1923	3185	1482	69.7	1941	2161	1034	93.0	1959	1387	644	337.3
1924	3351	1608	71.4	1942	2336	1054	117.5	1960	1289	564	351.8

Литература

Литература к разделу 1

1. Андерсон Т. (1976). Статистический анализ временных рядов. М.: Мир.
2. Бокс Дж., Дженкинс Г. (1974). Анализ временных рядов. Прогноз и управление. М.: Мир. Вып. 1, 2. (Русский перевод издания: Box G.E.P., Jenkins G.M. (1970) Time series analysis: Forecasting and control, San Francisco: Holden-Day.)
3. Носко В.П., Бузаев А.В., Кадочников П.А., Пономаренко С.С. (2003). Анализ прогнозных свойств структурных моделей и моделей с включением результатов опросов предприятий, сравнение с прогнозами по моделям временных рядов. Научные труды № 64. М.: ИЭПП.
4. Энтов Р.М., Носко В.П., Юдин А.Д., Кадочников П.А., Пономаренко С.С. (2002). Проблемы прогнозирования некоторых макроэкономических показателей. М.: ИЭПП.
5. Box G.E.P., Jenkins G.M., Reinsel G.C. (2008). Time Series Analysis: Forecasting and Control. 4th Edition. Wiley.
6. Beveridge S., Nelson C. (1981). A new approach to decomposition of economic times series into permanent and transitory components with particular attention to the measurement of the business cycle // Journal of Monetary Economics. Vol. 7. P. 151—174.
7. Campbell J.J., Perron P. (1991). Pitfalls and Opportunities: What Macroeconomists should know about Unit Roots // Macroeconomics Annual. 1991. NBER. P. 141—201.
8. Clements M.P., Hendry D.F. (1996). Multi-step estimation for forecasting // Oxford Bulletin of Economics and Statistics. Vol. 58. № 4. P. 657—684.
9. Clements Michael P., Hendry D.F. (1998a). Forecasting Economic Processes // International Journal of Forecasting. Vol. 14. № 1. P. 111—131.
10. Clements M.P., Hendry D.F. (1998b). Forecasting Economic Time Series. Cambridge: Cambridge University Press. (The Marshall Lectures on Economic Forecasting.)
11. Clements M.P., Hendry D.F. (2000). Forecasting with difference-stationary and trend-stationary models. Discussion Paper Series, Number 5. Department of Economics, University of Oxford.
12. Clements M.P., Hendry D.F. (2001). Forecasting Non-stationary Economic Time Series. Cambridge, Massachusetts: The MIT Press.
13. Dickey D.A., Bell W.R., Miller R.B. (1986). Unit Roots in Time Series Models: Tests and Implications // American Statistica. Vol. 40. P. 12—26.
14. Diebold F.X. (1998). The Past, Present and Future of Macroeconomic Forecasting // Journal of Economic Perspectives. Vol. 12. № 2. P. 175—192.

15. *Diebold F.X., Kilian L.* (2000). Unit-Root Tests Are Useful for Selecting Forecasting Models // *Journal of Business and Economic Statistics*. Vol. 18. № 3. P. 265—273.
16. *Diebold F.X., Mariano R.S.* (1995). Comparing Predictive Accuracy // *Journal of Business & Economic Statistics*. Vol. 13. № 3. P. 253—63.
17. *Enders W.* (1995). *Applied Econometric Time Series*. New York: Wiley.
18. *Granger C.W.J., Newbold P.* (1986). *Forecasting Economic Time Series*. Second Edition. New York: Academic Press.
19. *Gregory A.W., Smith G.W., Yetman J.* (2001). Testing for Forecast Consensus // *Journal of Business and Economic Statistics*. Vol. 19. № 1. P. 34—43.
20. *Hamilton J.D.* (1994). *Time Series Analysis*. Princeton. Princeton University Press.
21. *Harvey D.I., Leybourne S.J., Newbold P.* (1998). Tests for Forecast Encompassing // *Journal of Business and Economic Statistics*. Vol. 16. № 2. P. 254.
22. *Kumar K., Gill D.S.* (2000) On Forecasting Economic Time Series Data: A Comparative Study // *Indian Journal of Economics*. Vol. 81. P. 265—273.
23. *Snyder R.D., Ord J.K., Koehler A.B.* (2001) Prediction Intervals for ARIMA Models // *Journal of Business and Economic Statistics*. Vol. 19. № 2. P. 217—225.
24. *Stock J.H.* (1996). VAR, Error Correction and Pretest Forecasts at Long Horizons // *Oxford Bulletin of Economics and Statistics*. Vol. 58. № 4. P. 685—701.
25. *Stock J.H., Watson M.W.* (1996). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series // *Journal of Business and Economic Statistics*. Vol. 14. № 1. P. 11—30.
26. *Swanson N.R., White H.* (1997). Forecasting Economic Time Series Using Flexible versus Fixed Specification and Linear versus Nonlinear Econometric Models // *International Journal of Forecasting*. Vol. 13. № 4. P. 439—461.
27. *Wallis K.F., Whitley J.D.* (1991). Sources of Error in Forecasts and Expectations: UK economic models 1984-8 // *Journal of Forecasting*. Vol. 10. P. 231—253.
28. *West K.D.* (2001). Tests for Forecast Encompassing When Forecasts Depend on Estimated Regression // *Journal of Business and Economic Statistics*. Vol. 19. № 1. P. 29—33.

Литература к разделу 2

1. *Bernanke B.S., Blinder A.S.* (1992). The Federal Funds Rate and the Channels of Monetary Transmission // *American Economic Review*. Vol. 82. P. 901—921.
2. *Blanchard O., Quah D.* (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances // *The American Economic Review*. Vol. 79 (4). P. 655—673.
3. *Campbell J.Y., Perron P.* (1991). Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots // *NBER Macroeconomics Annual 1991*. Vol. 6 / eds. O.J. Blanchard, S. Fischer. Cambridge: MIT Press. January 1992.
4. *Christiano L.J., Eichenbaum M., Evans C.* (1998). Monetary Policy Shocks: What Have We Learned and to What End? // *NBER Working Paper*, 6400/1998.
5. *Clements M.P., Hendry D.F.* (1999). *Forecasting Non-stationary Economic Time series* // MIT Press. Cambridge; Massachusetts; London, 1999.
6. *Diebold F.X., Kilian L.* (2000). Unit-root tests are useful for selecting forecasting models // *Journal of Business and Economic Statistics*. Vol. 18. 265—273.
7. *Eichenbaum M.* (1992). Comment on Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy // *European Economic Review*. Vol. 10. P. 1001—1011.
8. *Favero C.A.* (2001). *Applied macroeconometrics* // Oxford University Press. Oxford: xi + 282 pp.

9. Favero C., Giavazzi F., Spaventa L. (1997). High Yields: the Spread on German Interest Rates // *Economic Journal*. Vol. 107. P. 956—986.
10. Gali J. (1992). Variability of Durable and Nondurable Consumption: Evidence for Six O.E.C.D. Countries // *Review of Economics & Statistics*. Vol. 75 (3). P. 418—28.
11. Hamilton J.D. (1994). *Time Series Analysis* // Princeton University Press. Princeton.
12. Lucas R.E. Jr. (1978). Macro-economic Policy Evaluation: A Critique // *Carnegi-Rochester Conference Series on Public Policy*. Vol. 1. P. 19—46.
13. Lee J., Shin B.S., In C. (1996). Causality between Advertising and Sales: New Evidence from Cointegration // *Applied Economic Letters*. Vol. 3. P. 299—301.
14. Leeper E.M., Sims C.A., Zha T. (1996). What Does Monetary Policy Do? // *Brookings Papers on Economic Activity*. 1996. Issue 2. P. 1—63.
15. Marcat A. (2001). Overdifferencing VAR's is Ok // Working Paper. UPF (Universitat Pompeu Fabra, Barcelona).
16. McCallum (1983). A Reconsideration of Sims' Evidence Regarding Monetarism // *Economics Letters*. Vol. 13. 167—171.
17. Mehra Y.P. (1978). Is Money Exogenous in Money-Demand Equations // *Journal of Political Economy*. Vol. 86. № 2. Part 1. P. 211—228.
18. Mellander E., Vredin A., Warne A. (1993). Stochastic Trends and Economic Fluctuations in a Small Open Economy // *Journal of Applied Econometrics*. Vol. 7. Issue 4. P. 369—94.
19. Phillips (1995). Impulse Response and Forecast Error Variance Asymptotics in Nonstationary VAR's // *Cowles Foundation Discussion Paper* 1102.
20. Shapiro M., Watson M. (1988). The sources of business cycle fluctuations: NBER Chapters // *NBER Macroeconomics Annual*. 1988. Vol. 3. P. 111—156. National Bureau of Economic Research, Inc.
21. Sims C. (1972). Money, Income, and Causality // *American Economic Review*. Vol. 62. P. 540—552.
22. Sims C.A. (1980). Comparison on Interwar and Postwar Cycles: Monetarism Revisited // *American Economic Review*. 1980. May. P. 250—257.
23. Sims C.A. (1980). Macroeconomics and reality // *Econometrica*. Vol. 48 (1). P. 1—48.
24. Sims C.A. (1992). Interpreting the Macroeconomic Time Series Facts: the Effects of Monetary Policy // *European Economic Review*. Vol. 36. P. 975—1011.
25. Sims C.A., Zha T.A. (1998). Does monetary policy generate recessions? // Working Paper 98-12. Federal Reserve Bank of Atlanta.
26. Toda H.Y., Phillips P.C.B. (1994). Vector autoregression and causality: A theoretical overview and simulation study // *Econometric Reviews*. Vol. 13. P. 259—285.
27. Toda H.Y., Yamamoto T. (1995). Statistical inferences in vector autoregressions with possibly integrated processes // *Journal of Econometrics*. Vol. 66. 225—50.

Литература к разделу 3

1. Carrion-i-Silvestre J.L., Sansó-i-Rosselló A.S. (2004). Testing the Null Hypothesis of Cointegration with Structural Breaks // Unpublished Manuscript. Departament d'Econometria. Estadística i Economia Espanyola. Universitat de Barcelona.
2. Corradi V. (1995). Nonlinear transformations of integrated time series: a reconsideration // *Journal of Time Series Analysis*. Vol. 16. 1995. P. 539—549.
3. Engle R.F., Granger C.W.J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing // *Econometrica*. Vol. 55. № 2. P. 251—276.
4. Ermini L., Granger C.W.J. (1993). Some Generalizations on the Algebra of I(1) Processes // *Journal of Econometrics*. Vol. 58. P. 369—384.

5. Granger C.W.J., Hallmann J. (1988). The algebra of I(1) // Finance and Economics Discussion Series. Vol. 45. Board of Governors of the Federal Reserve System (U.S.).
6. Granger C.W.J., Hallmann J. (1991). Nonlinear transformations of integrated time series // Journal of Time Series Analysis. Vol. 12. P. 207—224.
7. Kramer W., Davies L. (2002). Testing for unit roots in the context of misspecified logarithmic random walks // Economics Letters. Vol. 74. P. 313—319.
8. Phillips P.C.B. (1987). Time Series Regression with a Unit Root // Econometrica. Vol. 55. P. 277—301.
9. Phillips P.C.B., Hansen B.E. (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes // Review of Economic Studies. Vol. 57(1). P. 99—125.
10. Phillips P.C.B., Loretan M. (1991). Estimating long-run economic equilibria // Review of Economic Studies. Vol. 58. P. 407—436.
11. Saikkonen P. (1991). Asymptotically efficient estimation of cointegration regressions // Econometric Theory. Vol. 7. P. 1—21.
12. Stock J.H., Watson M.W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems // Econometrica. Vol. 61. P. 783—820.

Литература к разделу 4

1. Aigner D.J., Chu S.F. (1968). On estimating the industry production function // American Economic Review. Vol. 58. P. 826—839.
2. Aigner D., Lovell K., Schmidt P. (1977). Formulation and Estimation of Stochastic Frontier Function Models // Journal of Econometrics. Vol. 6. P. 21—37.
3. Alfirman L. (2003). Estimating Stochastic Frontier Tax Potential: Can Indonesian Local Governments Increase Tax Revenues Under Decentralization? // Working Paper No. 03—19. Center for Economic Analysis, Department of Economics, University of Colorado.
4. Battese G.E., Coelli T.J. (1992). Frontier Production Function, technical efficiency and panel data: with application to Paddy Farmers in India // Journal of Productivity Analysis. P. 153—169.
5. Broeck J. van den, Forsund F.R., Hjalmarsson L., Meeusen W. (1980). On the estimation of deterministic and stochastic frontier production functions // Journal of Econometrics. Vol. 13. P. 117—138.
6. Dugger R. (1974). An application of bounded nonparametric estimating functions to the analysis of bank cost and production functions. Unpublished Ph.D. dissertation (University of North Carolina. Chapel Hill, NC).
7. Farrell M.J. (1957). The measurement of productive efficiency // Journal of the Royal Statistical Society (Series A, general). Vol. 120. № 3. P. 253—281.
8. Schmidt P. (1976). On the statistical estimation of parametric frontier production functions // Review of Economics and Statistics. Vol. 58. P. 238—239.
9. Timmer C.P. (1971). Using a probabilistic frontier production function to measure technical efficiency // Journal of Political Economy. Vol. 79. P. 776—794.
10. Головань С.В., Карминский А.М., Пересецкий А.А. (2007). Факторы, влияющие на эффективность российских банков // Модернизация экономики и государство / отв. ред. Е.Г. Ясин; Гос. ун-т — Высшая школа экономики. В 3 кн. Кн. 2. М.: Изд. дом ГУ—ВШЭ. С. 188—206.

Глоссарий

К разделу 1

Аддитивная модель (*additive model*) временного ряда — представление уровня ряда в виде суммы тренда, циклической, сезонной и нерегулярной компонент.

Бевериджа — Нельсона разложение (*Beveridge-Nelson decomposition*) — представление стохастической составляющей ряда, описываемого моделью $ARMA(p, 1, q)$, в виде суммы стохастического тренда и стационарной составляющей.

Взвешенное скользящее среднее (*WMA — weighted moving average*) — метод сглаживания, в котором наблюдения усредняются с различными весами.

Двойное экспоненциальное сглаживание, метод Брауна (*DES — double exponential smoothing*) — метод, в котором процедура экспоненциального сглаживания применяется дважды (с одним и тем же значением параметра сглаживания): сначала к исходному ряду, а затем — к сглаженному ряду.

Диболда — Мариано статистика (*Diebold — Mariano statistics*) — статистика, используемая для проверки гипотезы об отсутствии различий между прогнозными свойствами двух альтернативных моделей.

Локальный прямолинейный тренд (*local linear trend*) временного ряда — понятие, используемое при прогнозировании будущих значений ряда адаптивными методами.

Локальный текущий уровень (*current level*) временного ряда — понятие, используемое при прогнозировании будущих значений ряда адаптивными методами.

Метод скользящего окна (*rolling forecast*) для построения прогнозов — метод построения прогнозов, при котором длина предпрогнозного периода не изменяется со временем: при поступлении новых данных «самые старые» данные отбрасываются.

Метод Хольта (метод линейного тренда) (*Holt's linear trend algorithm*) — метод, используемый для прогнозирования рядов с выраженным линейным трендом без сезонных составляющих. При этом прогноз идет в на-

правлении текущего локального линейного тренда, определяемого текущим уровнем ряда и текущим угловым коэффициентом локального тренда.

Метод Хольта — Винтерса (*Holt — Winter's algorithm*) — обобщение метода Хольта, допускающее наличие сезонного фактора в динамике ряда. При этом сезонный фактор может быть как аддитивным, так и мультипликативным.

Мультипликативная модель (*multiplicative model*) временного ряда — представление уровня ряда в виде произведения тренда, циклической, сезонной и нерегулярной компонент.

Нерегулярная компонента (*irregular component*) временного ряда — компонента временного ряда, остающаяся после выделения из него систематических эффектов.

Передифференцированный временной ряд (*overdifferenced series*) — ряд, полученный в результате дифференцирования стационарного временного ряда.

Перманентная компонента (*permanent component*) временного ряда — составляющая нестационарного временного ряда, включающая детерминированный и стохастический тренды.

Простое скользящее среднее (*SMA — simple moving average*) с периодом усреднения K — формируется путем усреднения K последовательных значений ряда x_1, x_2, \dots, x_T .

Простое экспоненциальное сглаживание (*SES — single exponential smoothing*) — метод, в котором «текущий уровень» ряда в i -м периоде есть взвешенное среднее текущего значения ряда и «текущего уровня» ряда в $(i - 1)$ -м периоде.

Рекурсивный метод прогнозирования (*recursive forecast*) — метод построения прогнозов, при котором предпрогнозный период увеличивается по мере поступления новых данных.

Систематические эффекты (*systematic effects*) — наблюдаемые в динамике уровней ряда основная тенденция (тренд), циклические колебания, сезонная изменчивость.

Стохастическая компонента (*stochastic component*) временного ряда — составляющая нестационарного временного ряда, состоящая из стохастического тренда и стационарной компоненты, имеющей нулевое среднее.

Транзитивная компонента (*transitive component*) временного ряда — стационарная составляющая нестационарного временного ряда.

Фильтр Ходрика — Прескотта (*Hodrick — Prescott filter*) — инструмент, используемый для получения сглаженной оценки долговременного тренда ряда x_t .

Центрированное скользящее среднее (*centered moving average*) — дополнительное усреднение полученного ряда простых скользящих средних, использующее малое четное значение периода усреднения.

Экспоненциальное сглаживание (*exponential smoothing*) — методика, при которой наблюдения усредняются с разными весами, при этом значения весов, приписываемых прошлым наблюдениям, убывают экспоненциальным образом по мере «старения» наблюдений.

Root Mean Squared Error (RMSE) — корень из среднего квадрата ошибки прогноза ряда на прогнозном периоде.

Mean Absolute Error (MAE) — средняя абсолютная ошибка прогноза ряда на прогнозном периоде.

Mean Absolute Percent Error (MAPE) — средняя абсолютная процентная ошибка прогноза ряда на прогнозном периоде.

К разделу 2

Блочная экзогенность (*block exogeneity*) первой группы переменных (или некоторой переменной) в отношении второй группы переменных (или другой переменной) — в модели векторной авторегрессии: запаздывающие значения второй группы переменных не входят в уравнения для первой группы переменных.

Декомпозиция (разложение) дисперсий ошибок прогнозов на h шагов вперед, $h = 1, 2, \dots$ (*variance decomposition*) — разложение дисперсий ошибок прогнозов рассматриваемых рядов на h шагов вперед, сделанных в момент t , на компоненты, обусловленные значениями фундаментальных инноваций в момент $t + h$, $h = 1, 2, \dots$

Долговременные ограничения (*long-run restrictions*) — ограничения на долговременные свойства системы, представляющей структурную модель векторной авторегрессии, обеспечивающие ее идентифицируемость.

Методология Toda — Ямамото (*Toda and Yamamoto approach*) — процедура, позволяющая производить проверку на причинность по Грейнджеру без предварительной проверки гипотезы единичного корня для определения порядков интегрированности рядов и проверки гипотезы об их коинтегрированности/некоинтегрированности.

Причинность по Грейнджеру (*Granger causality*) для двух временных рядов — переменная x не является причиной по Грейнджеру для переменной y (x is *Granger-noncausal* for y), если дополнительная информация о значениях x_T, x_{T-1}, \dots не улучшает оптимальный прогноз для y_{T+h} — для его построения достаточно использовать только значения y_T, y_{T-1}, \dots . Если же такая информация улучшает оптимальный прогноз, то переменная x является причиной по Грейнджеру для переменной y (x *Granger-causes* y , $x \xrightarrow{G} y$).

Причинность по Грейнджеру (*Granger causality*) для N временных рядов — переменные второй группы (включающей N_2 из N переменных) не являются Грейнджер-причиной для переменных первой группы (включающей

$N_1 = N - N_2$ переменных), если дополнительная информация о значениях переменных второй группы не улучшает оптимальный прогноз переменных первой группы — для его построения достаточно использовать только значения $y_{iT}, y_{i,T-1}, \dots$ ($i = 1, \dots, N_1$) первой группы переменных.

Если же такая информация улучшает оптимальный прогноз, то переменные второй группы являются причиной по Грейнджеру для переменных первой группы.

Причинность по Грейнджеру в долгосрочном плане (*long-run causality*) — понятие, связанное с наличием или отсутствием корректирующих составляющих в правых частях уравнений модели коррекции ошибок.

Причинность по Грейнджеру в краткосрочном плане (*long-run causality*) — понятие, связанное с наличием или отсутствием запаздывающих разностей отдельных переменных в правых частях уравнений модели коррекции ошибок.

Фундаментальные инновации (*fundamental innovations*) — инновации, не коррелированные между собой в совпадающие моменты времени. Являются основой для построения структурной *VAR*, а также для построения функций импульсных откликов и декомпозиций дисперсий ошибок прогнозов. Обычно предполагается, что фундаментальные инновации имеют единичные дисперсии.

Функции импульсных откликов (*impulse response function — IRF*) — функции, представляющие реакции значений $y_{1,t+s}, y_{2,t+s}, \dots$ на единичные импульсные изменения отдельных фундаментальных инноваций в момент t при фиксированных значениях всех остальных фундаментальных инноваций во все моменты времени.

К разделу 3

Динамический метод наименьших квадратов (*DOLS — dynamic ordinary least squares*) — процедура, уменьшающая смещение *OLS*-оценок коэффициентов коинтегрирующей регрессии. В случае когда система состоит из $I(1)$ -рядов (имеющих порядок интегрированности 1), сводится к добавлению в правую часть уравнения текущих, запаздывающих и опережающих разностей всех объясняющих переменных (метод «*leads-and-lags*»).

Динамический обобщенный метод наименьших квадратов (*DGLS — dynamic generalized least squares*) — вариант *DOLS*, учитывающий автокоррелированность в остатках расширенного уравнения *DOLS*.

Долговременная дисперсия (*long-run variance*) ряда u_t —

$$\sigma^2 = \lim_{T \rightarrow \infty} T \left\{ E \left(\frac{1}{T} \sum_{t=1}^T u_t \right)^2 \right\} = \lim_{T \rightarrow \infty} T E(\bar{u}_T^2).$$

Если процесс u_t стационарный (в широком смысле) и последовательность его автоковариаций $\gamma_j = \text{Cov}(u_t, u_{t-j})$ абсолютно суммируема, тогда

$$\sigma^2 = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j.$$

Функциональная предельная теорема (*functional limit theorem*) — теорема, гарантирующая сходимость функционалов от соответствующих случайных последовательностей к функционалам от стандартного винеровского процесса $W(t)$.

К разделу 4

Граничная производственная функция (*frontier production function*) — максимальный объем выпуска, который можно получить для заданного набора входных переменных при фиксированной технологии производства: $q = f(z, \beta)$, где q — объем выпуска (*output*); z — вектор значений входных переменных; β — вектор параметров.

Модели стохастической границы со случайными эффектами (*random-effect stochastic frontier*) — модели стохастической границы производственных возможностей для панельных данных, в которых случайные составляющие, соответствующие внешним шокам и технической неэффективности, распределены независимо друг от друга и от объясняющих переменных. Могут иметь постоянную или изменяющуюся во времени техническую эффективность.

Модель стохастической границы с фиксированными эффектами и постоянной во времени технической эффективностью (*fixed-effect stochastic frontier*) — модель стохастической границы производственных возможностей для панельных данных с фиксированными эффектами.

Стохастическая граница производственных возможностей фирмы i (*stochastic frontier production function*) — граница производственных возможностей фирмы i , учитывающая наличие внешних шоков.

Уровень эффективности/неэффективности i -й фирмы (*technical inefficiency*) — коэффициент ξ_i в представлении $q_i = f(z_i, \beta) \xi_i$ объема выпуска i -й фирмы ($0 < \xi_i \leq 1$), где $f(z, \beta)$ — граничная производственная функция.

Предметный указатель

- Бинарного выбора модель** 191
 гомпит-модель 192
 критерий Хосмера-Лемешоу 200
 латентная переменная 207
 логит-модель 192
 метод максимального правдоподобия 192
 показатели качества
 LRI 197
 McFadden's R2 197
 доля неправильных предсказаний 196
 псевдо-*R2* 197
 пробит-модель 192
 проверка гипотезы нормальности 42
 проверка гипотезы одинаковой распределенности ошибок 210
 процесс порождения данных 207
 пороговое значение 207
 сравнение альтернативных моделей 199
 функция правдоподобия 191
- Бинарного выбора модель для панельных данных** 249
 логит-модель с фиксированными эффектами 250
 пробит-модель со случайными эффектами 253
- Блочная экзогенность** 397
 проверка
- Бройша — Пагана критерий**
 для индивидуальных и временных эффектов 160
 для индивидуальных эффектов 141
 для проверки статистической независимости ошибок в разных уравнениях 122
- Взвешенный метод наименьших квадратов**
 107, 116, 190
 доступная версия 116
- Винеровский процесс** 493
- Внутри-оценка** 126—127
- Временной ряд**
 инерционность 379
 передифференцированный 386
 прогнозирование — см. *Прогнозирование временного ряда*
 сглаживание — см. *Сглаживание временного ряда*
 декомпозиция ряда
 аддитивная модель 327
 мультипликативная модель 327
 нерегулярная компонента 327
 разложение Бевериджа — Нельсона 383—385
 компоненты ряда
 перманентная 384
 стационарная 385
 стохастический тренд
 транзитивная 384
 циклическая 384
 систематические эффекты 327
 сезонная изменчивость 327
 тренд (основная тенденция) 327
 циклические колебания 327
 сезонная изменчивость 327
- Временные ряды**
 коинтегрированные
 в узком смысле 86
 коинтеграционное пространство 87
 базис 87
 коинтегрирующий вектор 86
 идентифицирующие ограничения 88
 ранг коинтеграции 86
- Гипотеза единичного корня**
 поведение статистики Дики-Фуллера при нелинейных преобразованиях 489

- Дамми переменные 123
- Долговременная дисперсия 490
- Долговременное положение равновесия системы 86
- отклонение от положения равновесия 86
- Инструмент** 43
- проверка пригодности 58
- J*-статистика 58
- Коинтегрированная VAR** 87
- модель коррекции ошибок (ECM) 87
- коэффициенты адаптации 97
- ранг коинтеграции 87
- Коинтегрирующие векторы
- оценивание
- идентифицирующие ограничения 92
- неявная форма 92
- явная форма 92
- динамический *GLS* 508
- динамический *OLS* 505
- Коэффициент детерминации
- внутри 143
- между 143
- полный 142
- Критерий отношения правдоподобий 209
- Логит** 205
- Между-оценка** 138
- Метод инструментальных переменных 176
- условия ортогональности 177
- Метод максимального правдоподобия 192
- логарифмическая функция правдоподобия 192
- функция правдоподобия 191
- Методология VAR 416
- декомпозиция дисперсий ошибок прогнозов 434
- для нестабильных VAR 467, 485
- структурная VAR 416, 428
- IMA*-представление 433, 464
- восстановление по приведенной форме 429
- блочно-треугольная матрица 459
- долговременные ограничения 463
- в коинтегрированных VAR 485
- Холецкого разложение 430
- фундаментальные инновации 419
- упорядочивание 419
- функция импульсного отклика 419, 433
- Модель коррекции ошибок
- без ограничений 88
- структурная 88
- идентифицируемость 90
- приведенная форма 89
- Модель стохастической границы
- внешние шоки 517
- для перекрестной выборки 515
- для панельных данных 531
- с фиксированными эффектами 532
- со случайными эффектами 532, 533
- производственная функция
- границная 516
- стохастическая граница производственных возможностей 517
- техническая неэффективность 517
- уровень эффективности 516
- Мультиномиальная логит-модель 219
- оценивание 220
- прогнозирование по оцененной модели 223
- Обобщенный метод моментов** 178
- оптимальная взвешивающая матрица 179
- правильность спецификации
- критерий Саргана 183
- Оценивание параметров структурного уравнения
- двухшаговый метод наименьших квадратов 43
- косвенный метод наименьших квадратов 42
- оценка максимального правдоподобия с ограниченной информацией 53
- оценка максимального правдоподобия с полной информацией 50—52
- обобщенный метод наименьших квадратов 49
- трехшаговый метод наименьших квадратов 49
- Панельные данные** 105
- несбалансированная панель 163
- сбалансированная панель 163
- Панельные данные, модели
- OLS* — дамми-модель 130

- двунаправленные 156
 - с фиксированными эффектами 156
 - со случайными эффектами 158
- динамическая модель 173
- индивидуально-специфические переменные 168
- модель кажущихся несвязанными регрессий (*SUR*) 116
- модель ковариационного анализа 115
- модель компонент дисперсии 135
- модель Мундлака 148
- модель несвязанных между собой регрессий 115
- модель пула 105
- модель с фиксированными эффектами, *FE*-модель 130
- модель со случайными эффектами 134
 - стандартная, *RE*-модель 135
- модель Хаусмана — Тейлора 170
- эндогенные объясняющие переменные 164
- Панельные данные, оценивание
 - ковариационная оценка 130
 - FE*-оценка 130
 - RE*-оценка 137
 - SUR*-оценка 117
 - доступная 118
 - внутри-оценка 127
 - двунаправленная 157
 - инструментальная 164
 - двунаправленная *GLS*-оценка 160
 - между-оценка 138
 - инструментальная 164
 - оценка Хаусмана — Тейлора 172
- Переменная
 - бинарная 185
 - дихотомическая 185
 - индикаторная 185
 - инструментальная 42
 - объясняющая 185, 191
 - предопределенная 23
 - экзогенная 23
 - эндогенная 15, 23
- Порядковая пробит-модель 213
 - оценивание 214
 - прогнозирование по оцененной модели, 214
 - стандартная нормализация 214
- Предельный эффект
 - в модели бинарного выбора 203
 - в логит-модели 204
- Причинность по Грейнджеру
 - в долгосрочном плане 406
 - в краткосрочном плане 406
 - в нестационарных VAR 399
 - коинтегрированных 402
 - некоинтегрированных 400
 - в стабильных VAR 398
 - для двух временных рядов 392
 - для нескольких временных рядов 398
- Причинность по Грейнджеру, проверка на отсутствие 394, 398, 401, 404, 407
- методология Тода — Ямамото 405
- Прогнозирование временного ряда
 - выбор модели 360—361, 379
 - для передифференцированного ряда 388
 - информационное множество 362
 - качество прогнозов
 - характеристики
 - bias proportion* 352
 - covariance proportion* 352
 - MAE* 352
 - MAPE* 352
 - RMSE* 351
 - variance proportion* 352
 - Тейла коэффициент 352
 - сравнение 382
 - Диболда — Мариано статистика 383
- по моделям AR, MA, ARMA, ARIMA
 - по модели AR(p) 363
 - по модели ARIMA(p, d, q) 369
 - по модели ARMA(p, q) 366
 - интервальные 367
 - по модели MA(q) 364
 - рекурсивный метод 381
 - скользящего окна метод 381
- функция потерь 362
- экстраполяция с использованием методов сглаживания 336
 - локальный прямолинейный тренд 341
 - текущий угловой коэффициент 344
 - локальный текущий уровень 336
- экстраполяция с использованием оцененного тренда 348
- Распределение
 - стандартное логистическое 192
 - стандартное нормальное 192
 - стандартное экстремальных значений (Гомпертца) 192

- стандартное экстремальных значений (Гумбея) 219
- Сглаживание временного ряда**
 - скользящее среднее
 - взвешенное 337
 - простое 328
 - центрированное 331
 - экспоненциальное сглаживание 337
 - двойное (Брауна метод) 341
 - Хольта метод 344
 - Хольта-Винтерса метод 347
 - простое 337
 - Ходрика-Прескотта фильтр 334
- Система одновременных уравнений**
 - приведенная форма 12
 - проверка правильности спецификации 55
 - критерий Дарбина — Ву — Хаусмана 57
 - критерий Хаусмана 56
 - прогнозирование 80
 - рекурсивная 40
 - структурная форма 11
 - идентифицируемость
 - в целом 36
 - неидентифицируемая 18
 - ограничения
 - исключающего типа 27
 - неисключающего типа 32
 - система с тождествами 34
 - роль константы 33
 - структурное уравнение
 - идентифицируемое точно 28
 - идентифицируемость 26
 - порядковое условие 27
 - ранговое условие 27
 - неидентифицируемое 28
 - нормированное 24
 - сверхидентифицируемое 23
 - частично идентифицируемая 19
- Сходимость по распределению 494**
- ТобитII-модель**
 - двухшаговая процедура Хекмана 244
 - лямбда Хекмана 244
 - стандартная 243
 - функция правдоподобия 245
- ТобитI-модель**
 - стандартная 243
- Тобит-модель**
 - стандартная 232
 - усеченная модель регрессии 233
 - цензурированная модель регрессии 232
- Тобит-модель со случайными эффектами 261**
- Условная логит-модель 219**
- Цензурированные данные 231**
- Шансы 204**
- Эффекты**
 - временные 156
 - дифференциальные 134, 156
 - индивидуальные 156
 - случайные 134
 - фиксированные 130

Учебник

Владимир Петрович НОСКО

ЭКОНОМЕТРИКА

Книга 2

Гл. редактор *Ю.В. Луизо*

Редактор *Н.В. Андрианова*

Художник *В.П. Кориунов*

Компьютерная подготовка оригинал-макета *О.Ю. Гудкова*

Корректоры *М.А. Миловидова, Н.Н. Цыркова*

Подписано в печать 14.04.2011. Формат 70 × 100 ¹/₁₆.

Бумага офсетная. Гарнитура Таймс. Печать офсетная.

Усл. печ. л. 46,8. Тираж 3000 экз. Заказ

Изд. № 219.

Издательский дом «Дело» РАНХиГС

119571, Москва, пр-т Вернадского, 82

Коммерческий отдел — тел.: (495) 433-2510, (495) 433-2502

com@anx.ru

www.delo.anx.ru

Отпечатано в ОАО «Можайский полиграфический комбинат»

143200, г. Можайск, ул. Мира, 93

www.oaompk.ru, www.oaompk.pf тел.: (495) 745-84-28, (49638) 20-685

ISBN 978-5-7749-0655-0



9 785774 906550