

330.43(075.8)

Л-844

Ірина Лук'яненко

Юрій Городніченко

СУЧАСНІ

ЕКОНОМЕТРИЧНІ
МЕТОДИ
У ФІНАНСАХ

330.43(075.8)
1844

Ірина Лук'яненко
Юрій Городніченко



163870

СУЧАСНІ ЕКОНОМЕТРИЧНІ МЕТОДИ У ФІНАНСАХ

Навчальний посібник

Рекомендовано Міністерством
освіти і науки України



Київ  2002

Наказом Міністерства
Університетом
"Київсько-Могилянська
академія"

ББК 65.26я73

Л84

Рекомендовано Міністерством освіти і науки України
(лист № 14/18.2-1020 від 14.05.2002)

Рецензенти:

Ю. М. Бажал, доктор економічних наук, професор,
декан факультету економічних наук НаУКМА
В. В. Вітлінський, доктор економічних наук, професор кафедри
економіко-математичних методів КНЕУ
О. А. Корольов, доктор економічних наук, професор кафедри
статистики та економетрії КНТЕУ

Лук'яненко І. Г., Городніченко Ю. О.

Л84 Сучасні економетричні методи у фінансах. Навчальний посіб-
ник.— К.: Літера ЛТД, 2002.— 352 с.

ISBN 966-7543-29-3

Пропонований посібник є першим вітчизняним путівником у світ практичного застосування сучасних економетричних методів. Базовий теоретичний матеріал викладений доступно та прозоро. Усі розділи ілюструються реальними прикладами з детальними поясненнями та коментарями, що дозволяє на практиці зрозуміти особливості застосування кожного методу. Наведено покроковий алгоритм реалізації розглянутих методів у пакеті E.VIEWS. Посібник узагальнює досвід викладацької та науково-дослідницької роботи авторів протягом останніх років і може бути використаний в курсах «Аналіз часових рядів», «Фінансове прогнозування», «Економетрика», «Макроекономічне прогнозування», а також як допоміжний матеріал для курсів «Бюджетна система», «Податкова система», «Фінансове програмування», «Макроекономічний аналіз» тощо.

Для студентів, аспірантів, докторантів, наукових працівників, фінансових аналітиків, усіх, хто займається моделюванням складних економічних процесів.

ББК 65.26я73

ISBN 966-7543-29-3

© І. Г. Лук'яненко,
Ю. О. Городніченко, 2002
© «Літера ЛТД», 2002

ЗМІСТ

Передмова	7
Розділ 1. Застосування ARIMA-моделей в емпіричних дослідженнях	
1.1. Теоретичні основи моделювання за допомогою ARIMA-моделей	10
1.1.1. Поняття моделювання за допомогою часових рядів. Приклади найпростіших моделей ($AR(p)$, $MA(q)$, $ARMA(p, q)$)	10
1.1.2. Поняття лагового оператора	15
1.1.3. Загальні принципи побудови ARMA-моделей	16
1.1.4. Стаціонарні та нестаціонарні часові ряди. Перевірка часових рядів на стаціонарність	17
1.1.5. Властивості найпростіших авторегресійних (AR), ковзного середнього (MA) та змішаних (ARMA) процесів. Поняття про автокореляційну функцію (ACF)	20
1.1.5.1. Властивості AR(1)-процесу	20
1.1.5.2. Властивості автокореляційного процесу другого порядку AR(2)	25
1.1.5.3. Корені полінома лагового оператора	30
1.1.5.4. Поняття часткової автокореляційної функції	33
1.1.5.5. Властивості процесів ковзного середнього (MA-процесів)	36
1.1.5.6. Властивості ARMA-процесів	38
1.1.6. Ідентифікація, оцінювання та тестування ARIMA(p, d, q)-моделей	40
1.1.6.1. Ідентифікація ARIMA-моделей	40
1.1.6.2. Процедура Хеннона та Пісанена для визначення порядку p та q ARMA(p, q)-процесу. Оцінювання та тестування остаточного вигляду ARIMA-моделі ...	44
1.1.7. Прогнозування за ARIMA-моделями	48
1.1.7.1. Основна ідея прогнозування за ARIMA-моделями	48
1.1.7.2. Властивості прогнозів за найпростішими ARIMA-моделями	50
1.2. Застосування ARIMA(p, d, q)-моделей на практиці: випадок податкових надходжень в Україні	56

Питання для самоперевірки	69
Вправи	70
Список літератури	74
Розділ 2. Застосування VAR-моделей в емпіричних дослідженнях	
2.1. Теоретичні основи моделювання за допомогою VAR-моделей	76
2.1.1. Поняття моделювання за допомогою системи авторегресійних моделей	76
2.2. Проблеми ототожнення та оцінювання VAR-моделей 82 у структурній формі	80
2.3. Оцінка VAR-моделей у приведеній формі	82
2.4. Вибір порядку (p) VAR-моделі	83
2.5. Зв'язок власних значень λ_i та рангу матриці A	84
2.6. Поняття стабільності VAR-моделей	85
2.7. Прогнозування на основі VAR-моделей	87
2.8. Інтерпретація результатів моделювання	89
2.8.1. Аналіз функції імпульсних відгуків	89
2.8.2. Декомпозиція дисперсій помилок прогнозів у VAR-моделюванні	94
2.9. Застосування VAR(p)-моделей на практиці: аналіз зв'язку між споживанням та доходом	97
Питання для самоперевірки	110
Вправи	110
Список літератури	112
Розділ 3. Застосування моделей корегування помилки в емпіричних дослідженнях	
3.1. Поняття моделі корегування помилки та коінтеграції	115
3.2. Коінтеграція нестационарних (інтегрованих) змінних	117
3.3. Поняття спільного стохастичного тренду у випадку коінтеграції	119
3.4. Механізм корегування помилки та коінтеграція	121
3.5. Зв'язок між рангом матриці, коінтеграційним рівнянням та моделлю корегування помилки	124
3.6. Перевірка часових рядів на коінтеграцію	126
3.6.1. Тест перевірки рядів на коінтеграцію Інгла—Гренджера (Engle—Granger)	127

3.6.2. Тест перевірки часових рядів на коінтеграцію Йохансена	131
3.7. Застосування моделей корегування помилки (ЕСМ) на практиці: моделювання податкових надходжень в Україні	140
Питання для самоперевірки	161
Вправи	163
Список літератури	164
Розділ 4. Застосування симультативних систем рівнянь в емпіричних дослідженнях	
4.1. Теоретичні основи побудови економетричних симультативних моделей. Основні поняття в оцінюванні симультативних систем рівнянь. Структурна та приведені форми симультативних моделей	168
4.2. Методи оцінювання невідомих параметрів у моделях симультативних рівнянь	171
4.3. Перевірка системи симультативних рівнянь на ототожненість	172
4.3.1. Умова порядку	173
4.3.2. Рангова умова ототожнення	173
4.4. Оцінка точно ототожненого рівняння: метод непрямих найменших квадратів	175
4.5. Оцінка переототожненого рівняння: двокроковий метод найменших квадратів	175
4.6. Рекурсивні моделі	178
4.7. Застосування моделей симультативних рівнянь на практиці: спрощена макромодель України	181
Питання для самоперевірки	218
Вправи	219
Список літератури	225
Розділ 5. Застосування моделей лонгітюдних даних в емпіричних дослідженнях	
5.1. Особливості моделювання за допомогою моделей лонгітюдних даних. Приклади найпростіших моделей	229
5.2. Приклади застосування моделей лонгітюдних даних в моделюванні бюджетних процесів на регіональному рівні	236
5.3. Теоретичні основи моделювання за допомогою моделей лонгітюдних даних з фіксованими ефектами	243

5.3.1. <i>F</i> -тест одночасної рівності значень фіксованих ефектів (m_i)	248
5.4. Теоретичні основи моделювання за допомогою моделей лонгітюдних даних з випадковими ефектами	249
5.5. Проблеми вибору між моделями з фіксованими та випадковими ефектами	256
5.5.1. Тест Хаусмана для перевірки кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} в моделях лонгітюдних даних	257
5.5.2. Оцінка параметрів лонгітюдних даних методом інструментальних змінних	259
5.6. Критерії якості оцінювання моделей лонгітюдних даних	262
5.7. Особливі випадки в моделях лонгітюдних даних. Тестування наявності автокореляції та гетероскедастичності	264
5.8. Корегування середньоквадратичних відхилень параметрів моделей лонгітюдних даних при наявності автокореляції та (або) гетероскедастичності	266
5.9. Застосування моделей лонгітюдних даних на практиці: моделювання податкових надходжень на регіональному рівні. Прогнозування витрат місцевих органів влади для України	269
Питання для самоперевірки	300
Вправи	301
Список літератури	302
ДОДАТКИ	
Додаток 1	305
Додаток 2	307
Додаток 3	311
Додаток 4	314
Додаток 5	315
Статистичні таблиці	326
Глосарій	333
Предметний покажчик	337
Загальний список літератури	342

ПЕРЕДМОВА

Економетричне моделювання стрімко увійшло в різноманітні сфери економічних досліджень. Важко уявити сучасного економіста без володіння хоча б базовим економетричним інструментарієм. Економетричні методи та моделі все більше використовуються не тільки у прогнозуванні, а й для підтвердження певних гіпотез щодо розвитку економічних процесів, аналізу взаємозв'язків різноманітних факторів, виявлення їх впливу на макро- та мікроекономічні явища, для емпіричного тестування економічної теорії, розробки і аналізу сценаріїв економічного розвитку та прийняття відповідних рішень. Як бачимо, ареал застосування економетричних методів є вельми широким. Звичайно, не слід забувати, що економетричні методи — це лише інструментарій дослідження, який не може замінити глибокого володіння економічною теорією, знання макро- та мікроекономічних процесів. Водночас моделювання таких процесів за допомогою економетричних методів, у свою чергу, дозволяє проникнути в саму їхню сутність, зрозуміти тонку матерію взаємозв'язків. Таким чином, завдяки економетриці теорія та практика стають двома сторонами однієї медалі — наукового розуміння великого світу економіки та фінансів.

Сьогодні економетрика переживає своє друге народження. Вона пропонує досліднику такий різноманітний арсенал методів та моделей, що їх використання робить можливим те, що не так давно ще здавалося неможливим. При цьому неймовірно зросла точність висновків на основі коректного застосування сучасних економетричних методів.

Звісно, щоб використовувати весь арсенал економетричних методів у емпіричних дослідженнях, його необхідно освоїти та вміти адекватно застосовувати на практиці. Для цього й було написано дану книжку, яка є своєрідним путівником у сфері найбільш поширених сьогодні економетричних методів, а саме інтегрованих методів авторегресії та ковзного середнього (ARIMA); векторних авторегресійних моделей (VAR); моделей корегування помилки (ECM); моделей лонгітюдних (панельних) даних та симулятивних систем рівнянь.

Автори свідомо зробили наголос не на теоретичних аспектах кожного методу, а на тонкощах їх практичного застосування для аналізу реальних проблем української економіки. Відзначимо, що у практичній реалізації ми послуговувалися пакетом E.VIEWS. Це, однак, не заважає реалізації всього

арсеналу економетричних методів за допомогою інших поширених економетричних пакетів, таких як STATA, SAS, STATISTICA, SPSS, RATS.

Обрана нами схема побудови посібника є дещо незвичайною. Справді, якщо вивчення економетричних методів, як правило, починають з оволодіння базовою теорією, то ми намагались насамперед наблизити користувача посібника до практики. Його організовано таким чином, що для першого знайомства з кожним методом та розуміння його практичного застосування достатньо розібрати, а якщо можливо, і повторити крок за кроком приклад, наведений наприкінці кожного розділу для ілюстрації методу. Всі приклади побудовано на реальній економічній інформації, тому аналіз їх розв'язків дає відповіді на питання, пов'язані з дійсною економічною проблемою.

Детальний аналіз кожного методу супроводжується ґрунтовними коментарями та мінімальним екскурсом до необхідної теорії, яка докладніше подана у теоретичній частині книги. Після того як практичний приклад розібрано, користувач може послідовно засвоїти необхідні теоретичні поняття, розглянувши основи моделювання за кожним методом. При цьому від студентів, аспірантів, фахівців та практиків вимагається знайомство тільки з базовими поняттями економетрики.

Вказаний підхід до побудови посібника обрано з чисто педагогічних міркувань. Стандартне викладення економетричних методів надто переобтяжене абстрактними поняттями, що відповідно ускладнює сприйняття. Тому ми пропонуємо, на перший погляд, незвичайний шлях — від практики до теорії, на якій ця практика побудована, — і сподіваємось досягти на ньому успіху. Наш досвід переконує, що практичне засвоєння складних економетричних понять спочатку дозволяє потім без великих труднощів перейти до їх теоретичного опанування.

Зазначимо, що пропонований науковий посібник є логічним продовженням базового підручника з економетрики І. Лук'яненко, Л. Краснікової «Економетрика» (Київ: Знання, 1998. — 493 с.). Після видання останнього пройшло достатньо часу. Сьогодні класичні економетричні методи стали стандартним інструментарієм у дослідженнях та практиці. Водночас продовжують з'являтися нові економетричні методи, збільшується статистична база, зростає розуміння необхідності ширшого використання економетричного апарату, застосування лонгітюдних даних тощо. Все це, у свою чергу, підвищує вимоги щодо оволодіння сучасними методами економетрики як до тих, хто має досвід роботи в галузях прикладних економічних та фінансових досліджень, так і до тих, хто вперше опановує її апарат.

У посібнику зроблено одну з перших в Україні спроб об'єднати в єдиний логічний блок сучасні економетричні методи. З класичних методів, частково викладених у вказаній вище «Економетриці», ми залишили тільки симулятивні системи рівнянь, інтерес до яких останнім часом відроджується.

З огляду на те що ця книга є перш за все навчальним посібником, кожний з п'яти розділів супроводжується аналізом результатів розв'язку конкретного питання, пов'язаного з певною економічною проблемою, детальним коментарем та описом покрокової реалізації використаного методу за допомогою пакета E.VIEWS. Наприкінці розділів вміщено контрольні запитання для самоперевірки, різноманітні вправи та список основної літератури з кожного методу.

Посібник є результатом науково-дослідницької роботи авторів за останні роки та багаторічного досвіду викладання курсу «Економетрика» для бакалаврів Національного університету «Києво-Могилянська Академія» і магістрів з економічної теорії (програма EERS при НаУКМА), стажування та навчання в Боулдерському економічному інституті (США), Блумінгтонському університеті (США), Словацькому аграрному університеті (Нітра, Словаччина), Мічіганському університеті (США), а також досвіду співпраці в рамках магістерської програми EERC з професорами Ленсом Бренменом, Чарльзом Стілом, Ізраелем Луцьким, Гаффаром Мугалом, Хартмутом Ліманом та плідної співпраці з Управлінням зведеного аналізу виконання бюджету та фінансово-економічної статистики Департаменту по бюджету Міністерства фінансів України.

РОЗДІЛ 1. ЗАСТОСУВАННЯ ARIMA-МОДЕЛЕЙ В ЕМПІРИЧНИХ ДОСЛІДЖЕННЯХ

1.1. ТЕОРЕТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ ЗА ДОПОМОГОЮ ARIMA-МОДЕЛЕЙ

1.1.1. Поняття моделювання за допомогою часових рядів. Приклади найпростіших моделей (AR(p), MA(q), ARMA(p, q))

Економічні процеси можна моделювати за допомогою різних засобів, серед яких основними є:

- 1) багатофакторні регресійні моделі;
- 2) симульативні системи регресійних рівнянь;
- 3) ARIMA-моделі (інтегровані моделі авторегресії та ковзного середнього);
- 4) VAR- (вектор-авторегресійні — vector autoregressive) та ECM-моделі (моделі корегування помилки);
- 5) моделі лонгітюдних, тобто панельних, даних (panel data models).

Моделювання економічних процесів на основі багатофакторних регресійних моделей набуло поширення в прикладних дослідженнях. Цей підхід вже став класичним. Його використовують як для аналізу впливу різноманітних факторів на досліджуваний показник, так і для прогнозування. Макроекономічні моделі різних країн, що базувались на симульативних системах регресійних рівнянь, домінували в економічному прогнозуванні з 1960 по 1970 р.

Проте цей клас моделей робить низку неявних припущень¹ щодо структури залежності між випадковими змінними. Ця структура визначається наявними інституціями, такими як закони, ділова практика, режим економічної політики (наприклад фіксований або плаваючий обмінний курс). При зміні інституційного оточення моделі часто не здатні адекватно відображати моделювані процеси, що є, по суті, лейтмотивом критики економетричних моделей відомим американським економістом Р. Лукасом (R. Lucas). Ця критика змусила економетристів почати пошуки нових методів економічного прогнозування, зокрема нових ефективних малозатратних методів, одним з яких є ARIMA-моделювання.

¹ Зокрема припускається, що інституції незмінні у часі, і їхні зміни враховуються або контролюються.

Основоположна публікація Дж. Бокса (Box) та Г. Дженкінса (Jenkins) «Аналіз часових рядів: прогнозування та контроль» [3] відкрила новий напрямок в методах прогнозування, стала теоретичною основою ARIMA-аналізу та методології, яка є порівняно новим поколінням засобів прогнозування, заснованих на аналізі ймовірнісних (стохастичних) властивостей часових рядів². При цьому відповідний часовий ряд моделюється лише за допомогою його минулих значень (лагів) та екзогенної випадкової величини.

ARIMA-моделі часто називають атеоретичними моделями, оскільки в переважній більшості випадків моделювання не вимагає економіко-теоретичного підґрунтя специфікації моделі. Необхідною умовою застосування ARIMA-методології є стаціонарність часового ряду, який розглядається. Якщо він є стаціонарним (поняття стаціонарності розглянемо трохи пізніше), то його можна змоделювати різними способами, зокрема за допомогою двох складових — авторегресійної (AR) та ковзного середнього (MA). Відповідно сама модель є комбінацією цих двох складових.

Почнемо аналіз ARIMA-моделей з аналізу їх першої авторегресійної складової. Загальний вигляд авторегресійної моделі:

$$Y_t = f(Y_{t-1}, Y_{t-2}, \dots, u_t), \quad (1.1)$$

де Y_t — значення показника в період часу t , Y_{t-1}, Y_{t-2}, \dots — значення показника в попередні періоди часу, відповідно з лагом 1, 2 і так далі, u_t — випадкова величина (збурення).

Як легко побачити в моделях типу (1.1), значення досліджуваного показника Y_t може бути пояснене, виходячи з його минулих (лагових) значень та випадкової величини (збурення). Оскільки більшість економічних часових рядів має значну автокореляцію, моделі на зразок (1.1) здатні забезпечувати високу прогнозну точність при невисоких затратах.

Для того, щоб записати авторегресійне рівняння в явному вигляді, необхідно насамперед визначити: 1) форму функціональної залежності (наприклад, можна використовувати логарифми змінних); 2) кількість лагів, яку необхідно включати в модель (порядок процесу); 3) структуру випадкової величини (збурення).

Найпростішою ARIMA-моделлю є модель, що складається тільки з авторегресійної складової першого порядку та відповідно називається авторегресійним процесом першого порядку (first-order autoregressive process) AR(1):

$$\text{AR}(1): Y_t = m + \alpha Y_{t-1} + u_t, \quad (1.2)$$

² Часові ряди часто називають динамічними рядами.

де Y_t, Y_{t-1} — поточне та лагове значення досліджуваного показника Y , m — константа, u_t — випадкова величина, яка за припущенням є білим шумом, тобто припускаємо, що $E(u_i) = 0$ для $\forall i$ та $E(u_i^2) = \sigma^2$ для $\forall i$; $E(u_i u_j) = 0$ для всіх $i \neq j$. Крім того, за нашою класифікацією дана модель є лінійною та має один лаг.

Неважко побачити, що в авторегресійній моделі першого порядку AR(1) показник Y_t можна представити комбінацією константи m , коефіцієнта α , початкового значення (Y_0) та комбінацією поточного і всіх попередніх значень випадкової величини. Зауважимо, що багато в чому ARIMA-моделі використовують апарат різницевих рівнянь. Справді, враховуючи рекурентність формули (1.2), перепишемо її для всіх періодів часу t , починаючи з першого:

$$\begin{aligned} Y_1 &= m + \alpha Y_0 + u_1, \\ Y_2 &= m + \alpha Y_1 + u_2 = m + \alpha(m + \alpha Y_0 + u_1) + u_2 = m(1 + \alpha) + \\ &+ \alpha^2 Y_0 + (u_2 + \alpha u_1), \\ &\dots \\ Y_t &= m(1 + \alpha + \alpha^2 + \dots + \alpha^{t-1}) + \alpha^t Y_0 + (u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \dots \\ &\dots + \alpha^{t-1} u_1). \end{aligned} \tag{1.3}$$

Помноживши обидві частини виразу (1.3) почергово на $u_t, u_{t-1}, u_{t-2}, \dots$, знайдемо відповідні математичні сподівання обох частин. При цьому отримаємо:

$$E(Y_t u_t) = \sigma^2, E(Y_t u_{t-1}) = \alpha \sigma^2, E(Y_t u_{t-2}) = \alpha^2 \sigma^2, \dots \tag{1.4}$$

З виразу легко побачити, що Y_t корелює з поточним і всіма попередніми значеннями випадкової величини (збурення), але не корелює з усіма її майбутніми значеннями. Y_{t-1} не корелює з поточними значеннями випадкової величини u_{t-1} та всіма її майбутніми значеннями, але корелює з усіма попередніми і так далі. Таким чином, класичне припущення регресійного аналізу про незалежність факторів та випадкових величин $E(x_i u_j) = 0$ не є вже справедливим для всіх i та j , коли незалежна змінна в регресійному рівнянні є лаговою ендогенною змінною. Звідси випливає, що не всі коефіцієнти коваріації між незалежною змінною та випадковою величиною у випадку авторегресійних моделей дорівнюють нулю, тобто оцінки, отримані за класичним методом найменших квадратів (МНК-оцінки), стають зміщеними. Тому потрібно досить обережно підходити до оцінювання параметрів у авторегресійних моделях. Детально проблеми, пов'язані з оцінюванням ARIMA-моделей, розглянемо пізніше.

Зазначимо, що авторегресійна модель першого порядку є звичайно найпростішою. Загальною авторегресійною моделлю є модель p -го порядку — $AR(p)$.

$$AR(p): Y_t = m + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t, \quad (1.5)$$

де $Y_t, Y_{t-1}, \dots, Y_{t-p}$ — поточне та лагові значення досліджуваного показника, p — кількість лагів у моделі, m — константа, $\alpha_1, \alpha_2, \dots, \alpha_p$ — невідомі параметри, значення яких необхідно оцінити, u_t — випадкова величина (збурення). Якщо вона є білим шумом, тобто якщо $E(u_i) = 0$ для $\forall i$; $E(u_i^2) = \sigma^2$ для $\forall i$; $E(u_i u_j) = 0$ для всіх $i \neq j$, то $AR(p)$ називається чистим авторегресійним процесом p -го порядку.

Авторегресійні моделі різних порядків не є єдиним механізмом, який дозволяє моделювати зміну Y . Іншою поширеною базовою формою АRIМА-моделей є модель ковзного середнього (МА). Найпростішою моделлю ковзного середнього є також модель першого порядку МА(1), яка має такий формалізований вигляд:

$$MA(1): Y_t = \mu + \varepsilon_t - \beta_1 \varepsilon_{t-1}. \quad (1.6)$$

У моделі ковзного середнього досліджуваний показник Y є комбінацією константи, поточного та всіх минулих значень випадкової величини (збурення), яка за припущенням вважається білим шумом: $E(\varepsilon_i) = 0$ для $\forall i$; $E(\varepsilon_i^2) = \sigma^2$ для $\forall i$; $E(\varepsilon_i \varepsilon_j) = 0$ для всіх $i \neq j$.

Загальною моделлю ковзного середнього є модель q -го порядку, тобто МА(q)-процес:

$$MA(q): Y_t = \mu + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q}, \quad (1.7)$$

де ε_t є білим шумом.

Якщо випадкова величина є білим шумом, модель (1.7) вважається чистим МА(q)-процесом.

Комбінація авторегресійної складової p -го порядку та складової ковзного середнього q -го порядку дає загальну АRМА(p, q)-модель, або змішаний авторегресійний процес з ковзною середньою.

$$ARMA(p, q): Y_t = m + \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q}. \quad (1.8)$$

Зазначимо, що ми можемо отримати загальну АRМА(p, q)-модель з АR(p)-

процесу, припускаючи складнішу форму для випадкової величини (збурення) u_t . Якщо допускати, що в моделі (1.5) збурення u_t не є білим шумом, то найбільш поширеною альтернативою буде представлення u_t у вигляді процесу ковзної середньої певного порядку або в загальному випадку MA(q)-процесу.

$$u_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q}, \quad (1.9)$$

де ε_t є білим шумом.

Комбінація (1.9) та (1.5) дає змішану авторегресійну ковзну середню, або ARMA(p, q)-процес.

$$Y_t = m + \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q}. \quad (1.10)$$

Сьогодні ARMA-моделі дуже поширені в практиці прогнозування. Можна виділити декілька основних причин, які сприяють їх практичному застосуванню:

а) не завжди попередня інформація про можливі взаємозв'язки між часовими рядами може бути добре обґрунтована. В цьому випадку чисто статистична модель, що зв'язує поточні та попередні значення досліджуваного показника, може використовуватись для короткострокових прогнозів;

б) інколи з добре відомих структурних моделей економічної теорії можна отримати моделі типу авторегресійних або моделей ковзного середнього, особливо при оцінці приведеної (reduced) форми симультативних систем рівнянь, тобто при виразі ендогенних змінних структурної моделі через попередньо визначені та екзогенні змінні. Проілюструємо це простим прикладом.

Розглянемо як макроекономічну найпростішу модель типу

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + u_t \text{ (споживання)}, \quad (1.11)$$

$$Y_t = C_t + I_t \text{ (тотожність національного доходу)}, \quad (1.12)$$

де C_t — споживання, Y_t — національний дохід, I_t — інвестиції.

Математично ця система двох рівнянь з трьома змінними «пояснює» будь-які дві змінні через третю. В економіці, в рамках даної моделі, ми традиційно вважаємо споживання (C) та національний дохід (Y) величинами, чиї значення залежать від змін в інвестиціях (I) та випадкової величини (u). Тобто C, Y є ендогенними змінними моделі (1.11)—(1.12), а I — екзогенною змінною.

Підставляючи (1.11) в (1.12), отримуємо:

$$C_t - (\alpha_2/1 - \alpha_1) \cdot C_{t-1} = \alpha_0/(1 - \alpha_1) + (\alpha_1/1 - \alpha_1) \cdot I_t + (1/(1 - \alpha_1)) \cdot u_t.$$

Помножимо тотожність доходу, лаг якої дорівнює одному періоду, на $\alpha_2/(1 - \alpha_1)$ і віднімемо отриманий результат від первiсної тотожності (1.12):

$$Y_{t-1} = C_{t-1} + I_{t-1} \mid \cdot \alpha_2 / (1 - \alpha_1) \Rightarrow (\alpha_2 / (1 - \alpha_1)) \cdot Y_{t-1} = (\alpha_2 / (1 - \alpha_1)) \cdot C_{t-1} +$$

$$+ (\alpha_2 / (1 - \alpha_1)) \cdot I_{t-1} \Rightarrow$$

$$Y_t = C_t + I_t$$

$$\frac{(\alpha_2 / (1 - \alpha_1)) \cdot Y_{t-1} = (\alpha_2 / (1 - \alpha_1)) \cdot C_{t-1} + (\alpha_2 / (1 - \alpha_1)) \cdot I_{t-1}}{Y_t - (\alpha_2 / (1 - \alpha_1)) \cdot Y_{t-1} = [C_{t-1} - (\alpha_2 / (1 - \alpha_1)) \cdot C_{t-1}] + [I_{t-1} - (\alpha_2 / (1 - \alpha_1)) \cdot I_{t-1}]} \Rightarrow$$

$$C_t - \frac{\alpha_0}{1 - \alpha_1} C_{t-1} = \frac{\alpha_0}{1 - \alpha_1} + \frac{\alpha_1}{1 - \alpha_1} I_t + \frac{1}{1 - \alpha_1} u_t, \tag{1.13}$$

$$Y_t - \frac{\alpha_0}{1 - \alpha_1} Y_{t-1} = \frac{\alpha_0}{1 - \alpha_1} + \frac{1}{1 - \alpha_1} (I_t - \alpha_2 I_{t-1}) + \frac{1}{1 - \alpha_1} u_t. \tag{1.14}$$

У лiвих частинах регресiйних рiвнянь (1.13) та (1.14) обидвi змiннi C та Y мають авторегресiйну компоненту першого порядку з однаковим коефiцiєнтом при лаговiй змiннiй. Праву частину кожного рiвняння можна розглядати як випадкову величину (збурення), чiї властивостi залежать вiд поведiнки змiнної I (iнвестицiї). Якщо I є бiлим шумом навколо певного середнього, то споживання можна представити чистим AR(1)-процесом, а нацiональний дохід Y_t в (1.14) є чистим ARMA(1,1)-процесом.

Для скороченого запису ARIMA-моделей, як правило, використовується лаговий оператор. Введемо поняття лагового оператора та розглянемо деякi його властивостi.

1.1.2. Поняття лагового оператора

Позначимо лаговий оператор через L . Якщо вiн стоїть перед будь-якою змiнною з часовим iндексом, то дає попереднє значення вiдповiдної змiнної. Наприклад: $L(x_t) = x_{t-1}$. Вiдповiдно $L^2(x_t) = L[L(x_t)] = L(x_{t-1}) = x_{t-2}, \dots, L^s(x_t) = x_{t-s}$.

$$(1 - L)x_t = x_t - x_{t-1} = \Delta x_t; L(1 - L)x_t = x_{t-1} - x_{t-2} = \Delta x_{t-1},$$

де Δ є оператором першої рiзницi.

У багатьох алгебраїчних операцiях лаг-оператор може тлумачитись як скаляр. Важливим є знаходження оберненого значення для виразу лагового

оператора L . Для цього введемо в лаговому записі поняття полінома, або полінома в L . Позначимо його через $A(L) = (1 - \alpha L)$ — багаточлен першого ступеня в L . Розглянемо добуток

$$(1 - \alpha L)(1 + \alpha L + \alpha^2 L^2 + \alpha^3 L^3 + \dots + \alpha^p L^p) = 1 - \alpha^{p+1} L^{p+1}, \quad (1.15)$$

де $p \rightarrow \infty$, $\alpha^{p+1} L^{p+1} \rightarrow 0$, $|\alpha| < 1$.

Виходячи з (1.15), можемо записати обернену величину полінома $A(L)$ як

$$A^{-1}(L) = 1/(1 - \alpha L) = 1 + \alpha L + \alpha^2 L^2 + \alpha^3 L^3 + \dots \quad (1.16)$$

Оберненою величиною до найпростішого лагового полінома першого порядку $A(L) = (1 - \alpha L)$ є поліном вигляду (1.16).

Лагові оператори та лагові поліноми надалі будемо використовувати для спрощення запису моделей часових рядів та деяких положень теорії.

1.1.3. Загальні принципи побудови ARMA-моделей

Побудову ARMA-моделей можна розбити на три основні етапи. На першому необхідно перевірити часові ряди на стаціонарність. Ця перевірка є вельми важливою, оскільки ARMA-моделі застосовуються лише для стаціонарних часових рядів. Якщо ж вони виявились нестаціонарними, то їх необхідно звести до стаціонарних операцією різниць (поняття стаціонарності, перевірки рядів на стаціонарність та способи зведення нестаціонарних рядів до стаціонарного вигляду ми розглянемо в наступному підрозділі). На другому етапі побудови ARMA-моделей, виходячи з аналізу автокореляційних властивостей трансформованого ряду, вибирають декілька ARMA-специфікацій з метою визначення найкращої. Обрана специфікація вважається остаточною, якщо її помилки будуть білим шумом. Якщо ні, то другий етап необхідно повторити. Після того як найкраща ARMA-модель відібрана, можна переходити до третього етапу, тобто використовувати її для прогнозу.

Як можна побачити, для побудови ARMA-моделей слід насамперед перевірити часові ряди на стаціонарність, тому розглянемо поняття стаціонарності ряду детальніше.

1.1.4. Стаціонарні та нестаціонарні часові ряди. Перевірка часових рядів на стаціонарність

В емпіричних дослідженнях, які ґрунтуються на моделюванні за допомогою часових рядів, необхідно, щоб динамічні ряди даних, які лежать в їх основі, були стаціонарними. Стохастичний процес вважається стаціонарним, якщо його середнє та дисперсія не залежать від часу, а значення коваріації між двома періодами часу залежить лише від проміжку між ними (лага) і не залежить від конкретного періоду часу³. Щоб пояснити це визначення, припустимо, що Y_t є стохастичним часовим рядом з такими характеристиками:

середнє: $E(Y_t) = \mu$;

варіація: $var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$;

коваріація: $\gamma_k = [E(Y_t - \mu)(Y_{t+k} - \mu)]$.

Припустимо, ми змінюємо значення Y з Y_t на Y_{t+m} . В такому випадку якщо Y_t є стаціонарним рядом, то середнє, дисперсія і значення коефіцієнтів автоковаріації Y_{t+m} мають бути такими ж, як і для Y_t . Отже, стаціонарним називається часовий ряд, математичне сподівання, дисперсія та автоковаріація якого не залежать від часу⁴.

Якщо математичне сподівання (середнє), дисперсія і автоковаріація (у різних проміжках) змінюються в залежності від часу, то такий часовий ряд є нестаціонарним. Його можна легко звести до стаціонарного операторами різниць.

Перший висновок про стаціонарність або нестаціонарність часового ряду можна зробити, аналізуючи графік часового ряду. Стаціонарність рядів умовно поділяється на три типи: стаціонарність навколо ненульової константи (CS), стаціонарність навколо нуля (ZS) та тренд-стаціонарність, або стаціонарність навколо тренду (TS). Всі три випадки відображено на рисунку 1, відповідно а), б) та в).

Стаціонарні ряди ще називаються динамічно стабільними або такими, що мають нульовий порядок інтеграції, що формалізовано записується: $Y_t \sim I(0)$. Графічне зображення нестаціонарного ряду не дозволяє виявити певну закономірність у його зміні. Нестаціонарні ряди є динамічно нестабільними та можуть мати різний, відмінний від нуля, порядок інтеграції. Порядком інтеграції є число, що показує, скільки разів часовий ряд потребує застосування оператора перших різниць, щоб стати стаціонарним рядом.

³ Стаціонарність часового ряду має й інші визначення. В даному випадку даємо визначення слабкої стаціонарності.

⁴ Далі ми розглянемо детально поняття стаціонарності. На даному етапі достатньо знати, що стаціонарність ряду потрібна для того, щоб висновки щодо вибірки можна було поширювати на генеральну сукупність.



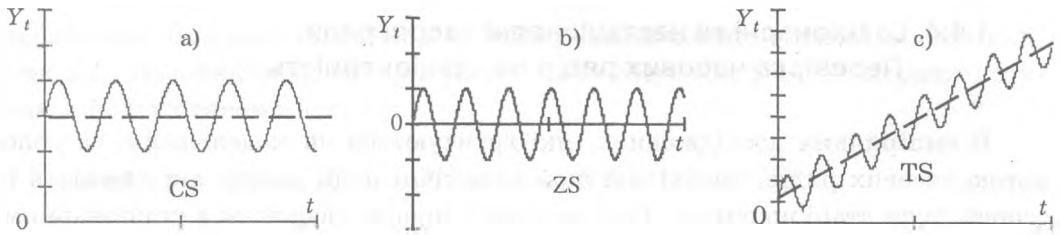


Рис. 1. Стаціонарні ряди. Різні випадки

Позначимо через d порядок інтеграції. Кажуть, що часовий ряд має одиничний корінь, або порядок інтеграції одиниця ($Y_t \sim I(1)$; $d = 1$), якщо $\Delta Y_t = Y_t - Y_{t-1}$ є стаціонарним рядом, тобто має нульовий порядок інтеграції ($\Delta Y_t = (Y_t - Y_{t-1}) \sim I(0)$; $d = 0$).

Часовий ряд має два одиничних корені, або порядок інтеграції 2, якщо його другі різниці є стаціонарним рядом: $Y_t \sim I(2)$; $d = 2$; $\Delta^2 Y_t = (\Delta Y_t - \Delta Y_{t-1}) \sim I(0)$. У загальному випадку часовий ряд має порядок інтеграції d : $Y_t \sim I(d)$, якщо $\Delta^d Y_t = (\Delta^{d-1} Y_t - \Delta^{d-1} Y_{t-1}) \sim I(0)$. Відмітимо, що якщо ряд стаціонарний, то будь-які його різниці залишаються стаціонарним рядом: $Y_t \sim I(0)$; $\Delta Y_t \sim I(0)$ і т. д.

Простий та розширений тести Дікі—Фуллера для перевірки рядів на стаціонарність

Звичайно візуальний контроль не є достатнім для висновку про стаціонарність/нестаціонарність часового ряду. Одним з формальних критеріїв для перевірки рядів на стаціонарність та визначення порядку інтеграції є тест Дікі—Фуллера (Dickey—Fuller test, DF) або розширений тест Дікі—Фуллера (augmented Dickey—Fuller test, ADF). В основі цього тесту є така регресія:

$$\Delta Y_t = a_0 + a_1 t + b Y_{t-1} + \sum_{i=1}^k c_i \Delta Y_{t-i} + \varepsilon_t, \quad (1.17)$$

де $\Delta Y_t = Y_t - Y_{t-1}$, t — часовий тренд, ε_t — випадкова величина, a_0 , a_1 , b , c_i — невідомі коефіцієнти регресії.

Якщо всі $c_i = 0$, $i = 1, \dots, k$, то (1.17) є базовою регресією для DF-тесту; в противному разі — для ADF-тесту. Включення лагів необхідне, щоб уникнути автокореляції залишків, до якої чутливий розподіл статистики МакКіннона (McKinnon) (за нульовою гіпотезою). Варто зазначити, що на практиці, як правило, використовують ADF-тест з кількістю лагів (k) меншою за 10 % від спостережень, щоб уникнути проблеми автокореляції. Крім того, правильна специфікація моделі (1.17) є особливо важливою при застосуванні цього тесту, тому слід визначити, чи потрібно включати вільний член a_0 та тренд (t) в модель.

Можна порадити таке евристичне правило: якщо візуально з графіка ряду не можна зробити висновок про наявність тренду, то необхідно включати лише перетин в (1.17), навіть якщо ряд коливається навколо нуля (в останньому випадку на наступних етапах застосування тесту перетин виключається з базової моделі). Якщо візуально ряд має тренд, то в моделі (1.17) необхідно залишити і перетин, і тренд. Зрозуміло, для визначення правильної специфікації (1.17) можна використати й складніші формалізовані процедури.

На основі базової моделі (1.17) тестуються такі гіпотези:

$H_0: b = 0$, або часовий ряд є нестационарним: $Y_t \sim I(d)$, $d > 0$;

$H_0: b < 0$, або часовий ряд є стаціонарним: $Y_t \sim I(0)$, $d > 0$.

Нульова гіпотеза H_0 відкидається, якщо отриманий коефіцієнт $b < 0$ та розрахована τ -статистика за абсолютною величиною більша за абсолютну величину критичного значення статистики МакКіннона для тестування на наявність одиничного кореня (його поняття розглянемо пізніше) при заданому рівні значимості α , тобто у формалізованому записі, якщо статистика МакКіннона для тестування одиничного кореня

$$|\tau\text{-stat}| = \left| \frac{b}{S_e(b)} \right|$$

більша за критичне значення при заданому рівні значимості α , де $S_e(b)$ — стандартна помилка (середнє квадратичне відхилення) оціненого параметра в моделі (1.17).

Визначення порядку інтеграції нестационарного ряду

Якщо на першому етапі ADF-тесту нульова гіпотеза H_0 не може бути відкинута, тобто ряд є нестационарним: $Y_t \sim I(d)$, $d > 0$, то постає наступне питання про порядок інтегрованості ряду. І знову застосовується ADF-тест, тільки вже для перевірки стаціонарності перших різниць ряду, а базове регресійне рівняння набуває форми других різниць:

$$\Delta^2 Y_t = a_0 + bY_{t-1} + \sum_{i=1}^k c_i \Delta^2 Y_{t-i} + \varepsilon_t. \tag{1.18}$$

Рівняння (1.18) є базовим рівнянням розширеного тесту Дікі—Фуллера для перевірки на стаціонарність перших різниць (ΔY_t) і, як правило, не містить тренду, оскільки на практиці майже виключена ситуація наявності тренду ряду в перших різницях. На основі базової моделі (1.18) тестуються такі гіпотези:

$H_0: b = 0$, або часовий ряд перших різниць є нестационарним: $\Delta Y_t \sim I(d)$, $d > 0$;

$H_0: b < 0$, або часовий ряд перших різниць є стаціонарним: $\Delta Y_t \sim I(0)$, $d > 0$, тобто ряд має порядок інтеграції 1: $Y_t \sim I(1)$, $d = 1$.

Нульова гіпотеза H_0 відкидається, якщо отриманий коефіцієнт $b < 0$ та розрахована τ -статистика за абсолютною величиною більша за абсолютну величину критичного значення статистики МакКіннона для тестування на наявність одиничного кореня при заданому рівні значимості α . Якщо ми не можемо відкинути нульову гіпотезу, то продовжуємо тестування далі для наступних різниць, поки не отримаємо стаціонарний ряд і відповідне значення порядку інтеграції.

Після того як часовий ряд визнано стаціонарним або зведено до стаціонарного необхідною кількістю різниць, слід перейти до другого етапу побудови ARMA-моделей, а саме визначення найкращої ARMA-специфікації.

Власне, поняття «найкраща специфікація» робить очевидним той факт, що остаточну ARMA-модель відбирають з множини можливих специфікацій, що асоціюються як з окремими AR- та MA-процесами, так і з їх численними ARMA-комбінаціями різних порядків. Тому, щоб зрозуміти правило відбору найкращої специфікації ARMA-моделі, розглянемо детально основні властивості AR-, MA-, ARMA-процесів.

1.1.5. Властивості найпростіших авторегресійних (AR), ковзного середнього (MA) та змішаних (ARMA) процесів. Поняття про автокореляційну функцію (ACF)

1.1.5.1. Властивості AR(1)-процесу

Розглянемо властивості AR-процесів, починаючи з найпростішого процесу першого порядку. Загальною специфікацією AR(1)-моделі є така:

$$\text{AR}(1): Y_t = m + \alpha Y_{t-1} + \varepsilon_t, \quad (1.19)$$

де ε_t є білим шумом.

Використовуючи лаговий оператор, можемо переписати модель (1.19) інакше:

$$(1 - \alpha L)Y_t = m + \varepsilon_t, \quad (1.20)$$

де $A(L) = 1 - \alpha L$, $1/(1 - \alpha L) = 1 + \alpha L + \alpha^2 L^2 + \dots \Rightarrow$

$$\text{AR}(1): Y_t = (1 + \alpha L + \alpha^2 L^2 + \dots)(m + \varepsilon_t). \quad (1.21)$$

Вирази (1.20) та (1.21) є еквівалентними представленнями AR(1)-моделі. Розпишемо (1.21) детальніше. Відмітимо, що лаг-оператор, який застосовується до константи будь-яку кількість разів, просто репродукує її значення, тому матимемо:

$$Y_t = m(1 + \alpha + \alpha^2 + \dots) + (\varepsilon_t + \alpha\varepsilon_{t-1} + \alpha^2\varepsilon_{t-2} + \dots). \quad (1.22)$$

Математичне сподівання AR(1)-процесу

Використовуючи вираз (1.22), розглянемо математичне сподівання AR(1)-процесу та доведемо, що воно не залежить від часу.

За умови, що $|\alpha| < 1$, тобто стаціонарності часового ряду, отримаємо:

$$E(Y_t) = \frac{m}{1 - \alpha} = \mu. \quad (1.23)$$

Вираз (1.23) є формулою для середньої величини або, іншими словами, для математичного сподівання AR(1)-процесу. Можна побачити, що в даному випадку змінна Y має постійну, а не умовну середню, тобто середню, незалежну від часу.

Дисперсія AR(1)-процесу

Можна легко показати, що дисперсія змінної Y , як і математичне сподівання, не залежатиме від часу. Інакше кажучи, якщо часовий ряд Y можна представити у вигляді авторегресійного процесу першого порядку AR(1), то його дисперсія також буде постійною безумовною величиною, тобто

$$\sigma_y^2 = E(Y_t - \mu)^2 = \sigma_\varepsilon^2 / (1 - \alpha^2). \quad (1.24)$$

Доведення формули (1.24)

Доведення того факту, що дисперсія часового ряду, який можна представити у вигляді авторегресійного процесу першого порядку (AR(1)), буде постійною безумовною величиною у вигляді (1.24).

Для цього представимо AR(1) у вигляді:

$$Y_t = m + \alpha Y_{t-1} + \varepsilon_t = m(1 + \alpha + \alpha^2 + \dots) + (\varepsilon_t + \alpha\varepsilon_{t-1} + \alpha^2\varepsilon_{t-2} + \dots).$$

Як було показано раніше, математичне сподівання такого процесу

$$E(Y_t) = \frac{m}{1-\alpha} = \mu = m(1 + \alpha + \alpha^2 + \dots). \quad (1.25)$$

Використовуючи формули (1.22), (1.23) та (1.25), отримаємо:

$$[Y_t - E(Y_t)] = [Y_t - \mu] = \varepsilon_t + \alpha\varepsilon_{t-1} + \alpha^2\varepsilon_{t-2} + \dots \quad (1.26)$$

Для отримання дисперсії AR(1)-процесу піднесемо обидві частини (1.26) до квадрата і знайдемо математичне сподівання такого виразу:

$$\text{Var}(Y_t) = E[(Y_t - \mu)^2] = E[\varepsilon_t^2 + \alpha^2\varepsilon_{t-1}^2 + \alpha^4\varepsilon_{t-2}^2 + \dots + 2\alpha\varepsilon_t\varepsilon_{t-1} + 2\alpha^2\varepsilon_t\varepsilon_{t-2} + \dots].$$

За припущенням $E(\varepsilon_t \varepsilon_{t-j}) = 0$, $E(\varepsilon_{t-j})^2 = \sigma^2$, а сума нескінченно спадної геометричної прогресії $(1 + \alpha^2 + \alpha^4 + \dots) = 1/(1 - \alpha^2)$ за умови, що $|\alpha| < 1 \Rightarrow \text{Var}(Y_t) = \sigma^2/(1 - \alpha^2)$. Отже, формулу (1.24) доведено. Як бачимо, дисперсія AR(1) не залежить від часу, якщо $|\alpha| < 1$.

Автокореляційна функція AR(1)-процесу

Для того, щоб знайти автоковаріаційну та автокореляційні функції AR(1)-процесу та визначити їх властивості, представимо авторегресійний процес першого порядку в спрощеному вигляді без константи. Для цього використаємо значення математичного сподівання $E(Y_t) = m/(1 - \alpha) = \mu$, тоді авторегресійну модель першого порядку

$$Y_t = m + \alpha Y_{t-1} + \varepsilon_t \quad (1.27)$$

ми можемо переписати у новому вигляді без константи:

$$X_t = \alpha X_{t-1} + \varepsilon_t, \text{ де } X_t = Y_t - \mu. \quad (1.28)$$

Покажемо, що вирази (1.27) та (1.28) є еквівалентними:

$$\begin{aligned} E(Y_t) = m + \alpha E(Y_{t-1}) &\Rightarrow m/(1 - \alpha) = m + \alpha m/(1 - \alpha) \Rightarrow \\ m/(1 - \alpha) &= [m(1 - \alpha) + \alpha m]/(1 - \alpha) = (m - \alpha m + \alpha m)/(1 - \alpha). \end{aligned} \quad (1.29)$$

Далі віднімемо від обох частин виразу (1.27) відповідні тотожні вирази $m/(1 - \alpha)$ та $m + \alpha m/(1 - \alpha)$ і отримаємо:

$$Y_t = m + \alpha Y_{t-1} + \varepsilon_t, \quad | - m/(1 - \alpha) = - m - \alpha m/(1 - \alpha) \Rightarrow$$

$$\begin{aligned}
 Y_t - m/(1 - \alpha) &= m + \alpha Y_{t-1} + \varepsilon_t - m - \alpha m/(1 - \alpha), \Rightarrow \\
 Y_t - m/(1 - \alpha) &= \alpha(Y_{t-1} - m/(1 - \alpha)) + \varepsilon_t.
 \end{aligned}
 \tag{1.30}$$

Враховуючи введене позначення, а саме: $X_t = Y_t - \mu$, $X_{t-1} = Y_{t-1} - \mu$, остаточно отримаємо:

$$X_t = \alpha X_{t-1} + \varepsilon_t, \tag{1.31}$$

$$X_t = Y_t - \mu, X_{t-1} = Y_{t-1} - \mu, \dots \tag{1.32}$$

Формулу для дисперсії авторегресійного процесу першого порядку $Var(Y_t) = \sigma^2/(1 - \alpha^2)$ можна легко отримати зі спрощеного виразу (1.31). Щоб продемонструвати це, піднесемо обидві частини (1.31) до квадрата і знайдемо математичне сподівання обох частин:

$$E(X_t^2) = \alpha^2 E(X_{t-1}^2) + E(\varepsilon_t^2) + 2\alpha E(X_{t-1}\varepsilon_t), \tag{1.33}$$

де X_{t-1} залежить лише від $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$, тому ε_t не корелює з усіма попередніми значеннями X через припущення білого шуму, тобто $E(X_{t-1}\varepsilon_t) = 0$.

Коли α задовольняє умову $|\alpha| < 1$, то з (1.32) легко побачити, що: $\sigma_y^2 = E(X_t^2) = E(X_{t-1}^2) = \dots$, а рівняння (1.33) відповідно набуває вигляду:

$$\sigma_y^2 = \alpha^2 \sigma_y^2 + \sigma_\varepsilon^2 \text{ або}$$

$$\sigma_y^2 = \sigma_\varepsilon^2 / (1 - \alpha^2).$$

Використовуючи спрощений вигляд AR(1), а саме (1.31), введемо поняття автоковаріаційної та автокореляційної функції. Для цього вираз $X_t = \alpha X_{t-1} + \varepsilon_t$ помножимо послідовно на $X_{t-1}, X_{t-2}, \dots, X_{t-s}$ і знайдемо математичне сподівання обох частин. Відповідно отримаємо:

$$\begin{aligned}
 E(X_t X_{t-1}) &= \alpha E(X_{t-1}^2) + E(X_{t-1}\varepsilon_t), \\
 E(X_t X_{t-2}) &= \alpha E(X_{t-1} X_{t-2}) + E(X_{t-2}\varepsilon_t), \\
 &\dots \\
 E(X_t X_{t-s}) &= \alpha E(X_{t-1} X_{t-s}) + E(X_{t-s}\varepsilon_t).
 \end{aligned}
 \tag{1.34}$$

Зауважимо, що $E(X_{t-1}\varepsilon_t) = E(X_{t-2}\varepsilon_t) = \dots = E(X_{t-s}\varepsilon_t) = 0$. Позначимо математичне сподівання добутку X з нульовим лагом, тобто $E(X_t^2), E(X_{t-1}^2), \dots$

через γ_0 . Легко побачити, що $\gamma_0 = E(X_t^2) = E(X_{t-1}^2) = \dots = \sigma_y^2$, тобто буде дорівнювати дисперсії AR(1)-процесу. Через γ_1 позначимо математичне сподівання добутку X з одиничним лагом X , тобто $E(X_t X_{t-1})$, $E(X_{t-1} X_{t-2})$. Через γ_s позначимо відповідно математичне сподівання добутку X з лагом s , тобто $E(X_t X_{t-s})$, $E(X_{t-1} X_{t-1-s})$, ...

Враховуючи введені позначення, вираз (1.34) легко переписати у рекурентному вигляді:

$$\begin{aligned} \gamma_1 &= \alpha\gamma_0, \gamma_0 = \sigma_y^2 = E(x_{t-1}^2), \\ \gamma_2 &= \alpha\gamma_1, \\ &\dots \\ \gamma_k &= \alpha\gamma_{k-1} = \alpha^k \gamma_0, k = 1, 2, \dots, \end{aligned} \quad (1.35)$$

$$\text{де } \gamma_0 = \sigma_y^2 = E(x_t^2) + E(x_{t-1}^2) = \dots \quad (1.36)$$

Коефіцієнти γ_k називаються коефіцієнтами автоковаріації, для них справедлива рекурентна формула (1.35). Вони відповідають математичному сподіванню добутку двох значень модельованого ряду X з певним лагом та є абсолютною величиною. Для того, щоб перейти до відносних величин, введемо поняття коефіцієнта автокореляції, який позначимо через ρ_k , $k = 1, 2, \dots$

Коефіцієнти автокореляції (теоретичні) часового ряду визначаються так:

$$\rho_k = \frac{\overline{E(X_t X_{t-k})}}{\sqrt{\text{Var}(X_t) \cdot \text{Var}(X_{t-k})}} = \frac{\gamma_k}{\gamma_0}, k = 1, 2, \dots \quad (1.37)$$

Вибіркові коефіцієнти автокореляції розраховуються за формулою:

$$r_k = \frac{\sum_{t=k+1}^n (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}, \quad (1.38)$$

де $\bar{x} = \sum_{t=1}^n \frac{x_t}{n}$, n — кількість спостережень.

Коефіцієнти автокореляції і автоковаріації симетричні відносно нульового лага, тобто $E(X_t X_{t-k}) = E(X_t X_{t+k})$ та $\rho(X_t X_{t-k}) = \rho(X_t X_{t+k})$. Таким чином, можна розглядати лише додатні значення лагів⁵. Відповідно до коефі-

⁵ Антилиаги (або від'ємні лаги) називаються випередженнями.

цієнтів автоковаріації для визначення коефіцієнтів автокореляції різних порядків також справедлива рекурентна формула, а саме коефіцієнти автокореляції для AR(1) визначаються за формулою

$$\rho_k = \alpha \rho_{k-1} = \alpha^k; \quad k = 1, 2, \dots \quad (1.39)$$

Як можна легко побачити з формули (1.39), коефіцієнт автокореляції нульового порядку завжди дорівнює одиниці. Крім того, він є спадною функцією від номера лага, тобто спадає зі збільшенням лага, оскільки ми припускаємо, що $|\alpha| < 1$. Іншими словами, формула (1.39) визначає автокореляційну функцію часового ряду (autocorrelation function of the series — ACF). Графічне зображення цієї функції являє собою корелограму (correlogram). Вигляд корелограми залежить від знака параметра α . На рисунках 2, 3 відображено графіки-корелограми для різних значень α у випадку AR(1)-процесу.

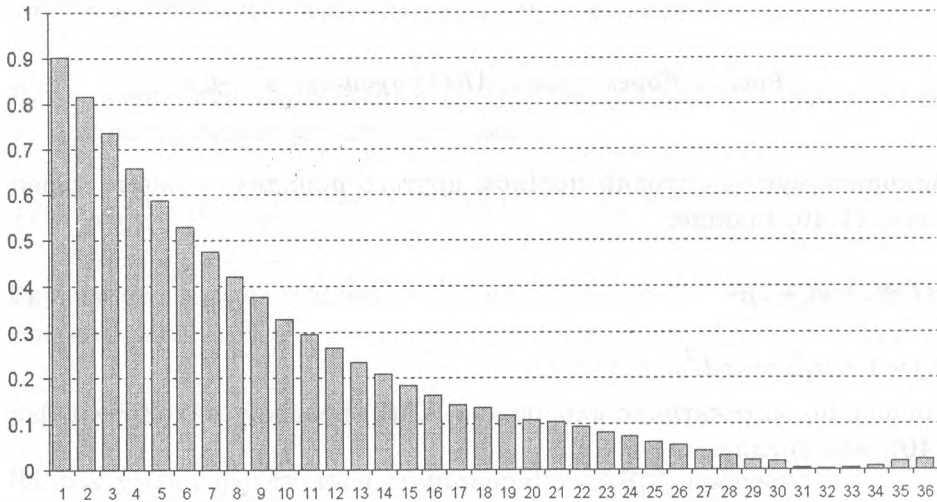


Рис. 2. Корелограма AR(1)-процесу: $\alpha = 0.9$

1.1.5.2. Властивості автокореляційного процесу другого порядку AR(2)

Математичне сподівання AR(2)

Загальною специфікацією AR(2)-моделі є:

$$AR(2): Y_t = m + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t, \quad (1.40)$$

де ε_t за припущенням вважається білим шумом.

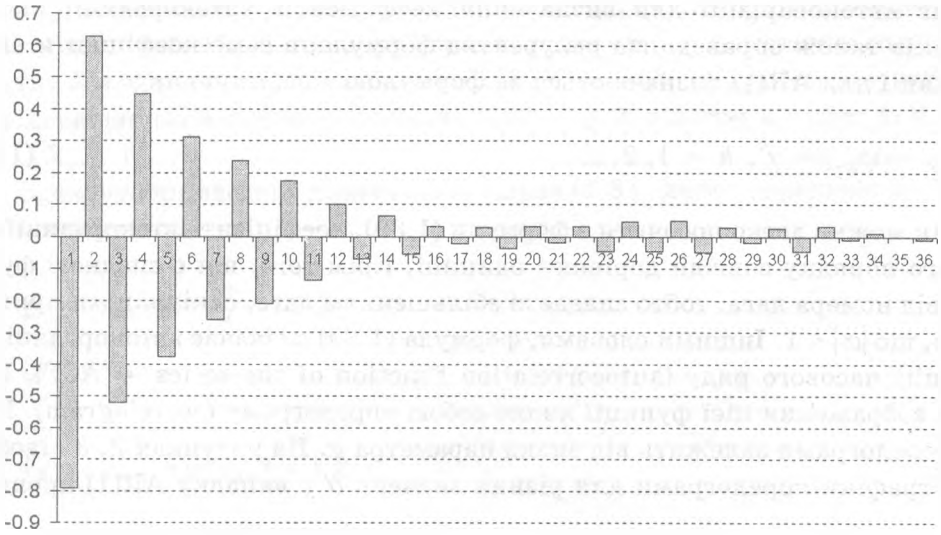


Рис. 3. Корелограма $AR(1)$ -процесу: $\alpha = -0.8$

Використовуючи лаговий поліном другого порядку, можемо переписати модель (1.40) інакше:

$$A(L)Y_t = m + \varepsilon_t, \quad (1.41)$$

де $A(L) = 1 - \alpha_1 L - \alpha_2 L^2$.

Відповідно математичне сподівання $AR(2)$ -процесу, виходячи з формули (1.40), має вигляд:

$$E(Y_t) = m / (1 - \alpha_1 - \alpha_2) = \mu. \quad (1.42)$$

З цієї формули можна побачити, що математичне сподівання $AR(2)$ -процесу не залежить від часу.

Доведення формули (1.42)

Для доведення формули (1.42) скористаємось записом $AR(2)$ -процесу з використанням лагового полінома другого порядку:

$$A(L) Y_t = m + \varepsilon_t,$$

де $A(L) = 1 - \alpha_1 L - \alpha_2 L^2$.

Далі представимо лаговий поліном другого порядку як добуток двох поліномів першого порядку з коренями β_1, β_2 і отримаємо:

$$(1 - \alpha_1 L - \alpha_2 L^2) = (1 - \beta_1 L)(1 - \beta_2 L) = 1 - (\beta_1 + \beta_2)L + \beta_1 \beta_2 L^2,$$

де $\beta_1 + \beta_2 = \alpha_1, \beta_1 \beta_2 = -\alpha_2$.

Позначимо через $A_1(L) = (1 - \beta_1 L)$ лаговий поліном першого порядку в L . Розглянемо добуток:

$$(1 - \beta_1 L)(1 + \beta_1 L + \beta_1^2 L^2 + \dots + \beta_1^p L^p) = 1 - \beta_1^{p+1} L^{p+1} \rightarrow 1,$$

оскільки $\beta_1^{p+1} L^{p+1} \rightarrow 0$, коли $p \rightarrow \infty$ та $|\beta_1| < 1$.

$$\text{Звідси маємо: } 1/(1 - \beta_1 L) = (1 + \beta_1 L + \beta_1^2 L^2 + \dots + \beta_1^p L^p).$$

Якщо позначити через $A_2(L) = (1 - \beta_2 L)$ інший поліном першого ступеня в L , то для нього також виконується умова:

$$1/(1 - \beta_2 L) = (1 + \beta_2 L + \beta_2^2 L^2 + \dots + \beta_2^p L^p), \text{ якщо } |\beta_2| < 1.$$

Зауважимо, що відповідно до наших позначень умови $|\beta_1| < 1$ та $|\beta_2| < 1$ рівнозначні умові:

$$\alpha_1 + \alpha_2 < 1; \alpha_2 - \alpha_1 < 1; |\alpha_2| < 1.$$

За цих умов гарантується стаціонарність AR(2)-процесу.

Представимо лаги полінома другого порядку як добуток двох поліномів першого порядку: $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 = (1 - \beta_1 L)(1 - \beta_2 L)$. Відповідно тепер AR(2) можемо записати у вигляді: $A(L)Y_t = (1 - \beta_1 L)(1 - \beta_2 L)Y_t = m + \varepsilon_t$, звідки маємо:

$$\begin{aligned} (1 - \beta_2 L)Y_t &= (1/(1 - \beta_1 L))(m + \varepsilon_t) = (1 + \beta_1 L + \beta_1^2 L^2 + \dots)(m + \varepsilon_t) = \\ &= m(1 + \beta_1 + \beta_1^2 + \dots) + (\varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \dots). \end{aligned}$$

Виразимо Y_t :

$$\begin{aligned} Y_t &= (1/(1 - \beta_2 L)) \left[m(1/(1 - \beta_1)) + (\varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \dots) \right] = \\ &= (1 + \beta_2 L + \beta_2^2 L^2 + \dots) \left[m(1/(1 - \beta_1)) + (\varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \dots) \right] = \end{aligned}$$

$$= m(1/(1 - \beta_1))(1 + \beta_2 + \beta_1^2 + \dots) + (1 + \beta_2 L + \beta_1^2 L^2 + \dots) + \\ + (\varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \dots).$$

Відповідно математичне сподівання Y_t буде:

$$E(Y_t) = m \frac{1}{(1 - \beta_1)} \frac{1}{(1 - \beta_2)} = \frac{m}{1 - a_1 - a_2}.$$

Формулу (1.42) доведено.

Тепер знайдемо формулу для визначення дисперсії AR(2)-процесу. Для цього перепишемо AR(2) у спрощеному вигляді та введемо такі позначення: $X_t = Y_t - \mu$.

Виходячи із введеного позначення, (1.40) можна переписати у вигляді:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t. \quad (1.43)$$

Дисперсія AR(2)-процесу

Для знаходження дисперсії AR(2)-процесу помножимо обидві частини (1.43) на X_t і знайдемо математичне сподівання обох частин:

$$E(X_t^2) = \alpha_1 E(X_t X_{t-1}) + \alpha_2 E(X_t X_{t-2}) + \alpha_1 E(X_t \varepsilon_t). \quad (1.44)$$

Зазначимо, що $E(X_t \varepsilon_t) = \sigma_\varepsilon^2$, оскільки, як було показано раніше, $E(Y_t \varepsilon_t) = \sigma_\varepsilon^2$ (формула 1.4).

Використовуючи раніше введені позначення для коефіцієнтів автоковаріації, перепишемо (1.44) у вигляді:

$$\gamma_0 = \alpha_1 \gamma_1 + \alpha_2 \gamma_2 + E(X_t \varepsilon_t) = \alpha_1 \gamma_1 + \alpha_2 \gamma_2 + \sigma_\varepsilon^2. \quad (1.45)$$

Далі помножимо поетапно обидві частини (1.43) відповідно на X_{t-1} та X_{t-2} і знайдемо математичне сподівання обох частин:

$$E(X_t X_{t-1}) = \alpha_1 E(X_{t-1}^2) + \alpha_2 E(X_{t-1} X_{t-2}) + E(X_{t-1} \varepsilon_t). \quad (1.46)$$

$$E(X_t X_{t-2}) = \alpha_1 E(X_{t-1} X_{t-2}) + \alpha_2 E(X_{t-2}^2) + E(X_{t-2} \varepsilon_t). \quad (1.47)$$

За визначенням $E(X_{t-1} \varepsilon_t) = 0$ та $E(X_{t-2} \varepsilon_t) = 0$.

Використовуючи раніше введені позначення для коефіцієнтів автоковаріації, перепишемо (1.46) та (1.47) у вигляді:

$$\begin{aligned}\gamma_1 &= \alpha_1\gamma_0 + \alpha_2\gamma_1, \\ \gamma_2 &= \alpha_1\gamma_1 + \alpha_2\gamma_0.\end{aligned}\tag{1.48}$$

Підставляючи (1.48) в (1.45) і спрощуючи, отримаємо:

$$\gamma_0 = (1 - \alpha_2)\sigma_\varepsilon^2 / [(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)].\tag{1.49}$$

Формула (1.49) є формулою дисперсії AR(2)-процесу. Оскільки дисперсія є величиною додатною, то необхідно, щоб кожний елемент в дужках був додатним, що, в свою чергу, дає умови стаціонарності для AR(2)-процесу:

$$\alpha_2 + \alpha_1 < 1; \alpha_2 - \alpha_1 < 1; |\alpha_2| < 1.$$

Зазначимо, що ці ж умови ми також отримали при доведенні формули математичного сподівання AR(2)-процесу.

Нагадаємо, що коефіцієнти автокореляції ρ_k визначаються за формулою:

$$\rho_k = \frac{E(X_t X_{t-k})}{\sqrt{\text{Var}(X_t) \cdot \text{Var}(X_{t-k})}} = \frac{\gamma_k}{\gamma_0}, k = 1, 2, \dots$$

Для того, щоб у системі (1.48) перейти до коефіцієнтів автокореляції, необхідно розділити γ_1 та γ_2 на γ_0 . Відповідно отримаємо систему для коефіцієнтів автокореляції, яку ще називають системою Юла—Уокера (Yule—Walker):

$$\begin{aligned}\rho_1 &= \alpha_1 + \alpha_2\rho_1, \\ \rho_2 &= \alpha_1\rho_1 + \alpha_2.\end{aligned}\tag{1.50}$$

З (1.50) відповідно отримаємо вирази для коефіцієнтів автокореляції першого та другого порядку:

$$\rho_1 = \alpha_1 / (1 - \alpha_2); \rho_2 = \alpha_1^2 / (1 - \alpha_2).\tag{1.51}$$

Для коефіцієнтів автокореляції загального порядку справедливою є така рекурентна формула:

$$\rho_k = \alpha_1\rho_{k-1} + \alpha_2\rho_{k-2}, k = 3, 4, \dots\tag{1.52}$$

з двома початковими умовами, які знаходять за формулою (1.51):

$$\rho_1 = \alpha_1/1 - \alpha_2, \rho_2 = \alpha_1^2/(1 - \alpha_2) + \alpha_2.$$

Формула (1.52) відповідно визначає автокореляційну функцію (ACF) для авторегресійного процесу другого порядку. Умови стаціонарності забезпечують зменшення значення автокореляційної функції при збільшенні лагів. Вигляд автокореляційної функції залежить від коренів лагового полінома. При дійсних коренях автокореляційна функція є спадною показниковою функцією від лага. Якщо корені є комплексними, то вона перетворюється на спадну синусоїду. На рисунках 4 та 5 зображено автокореляційну функцію AR(2) в залежності від різних значень коренів полінома.

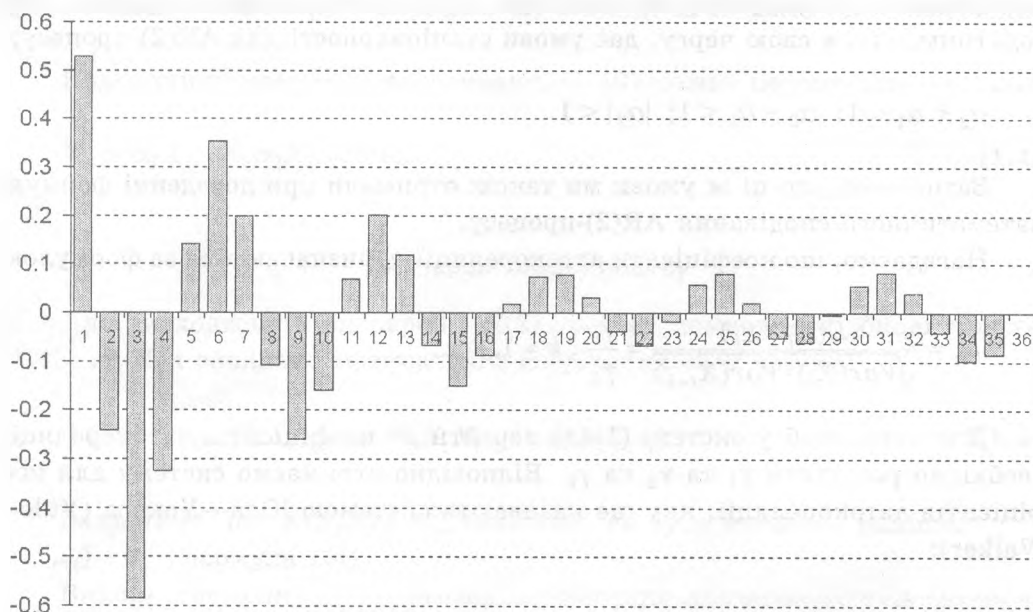


Рис. 4. Корелограма: $\alpha_1 = 0.9$, $\alpha_2 = -0.7$

1.1.5.3. Корені полінома лагового оператора

Проілюструємо поняття коренів полінома лагового оператора для авторегресійних процесів на прикладі авторегресійного процесу другого порядку. Використовуючи поліном лагового оператора, AR(2)-процес запишемо у вигляді:

$$\text{AR}(2): X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t, \text{ або } A(L)X_t = \varepsilon_t, \text{ де } A(L) = 1 - \alpha_1 L - \alpha_2 L^2.$$

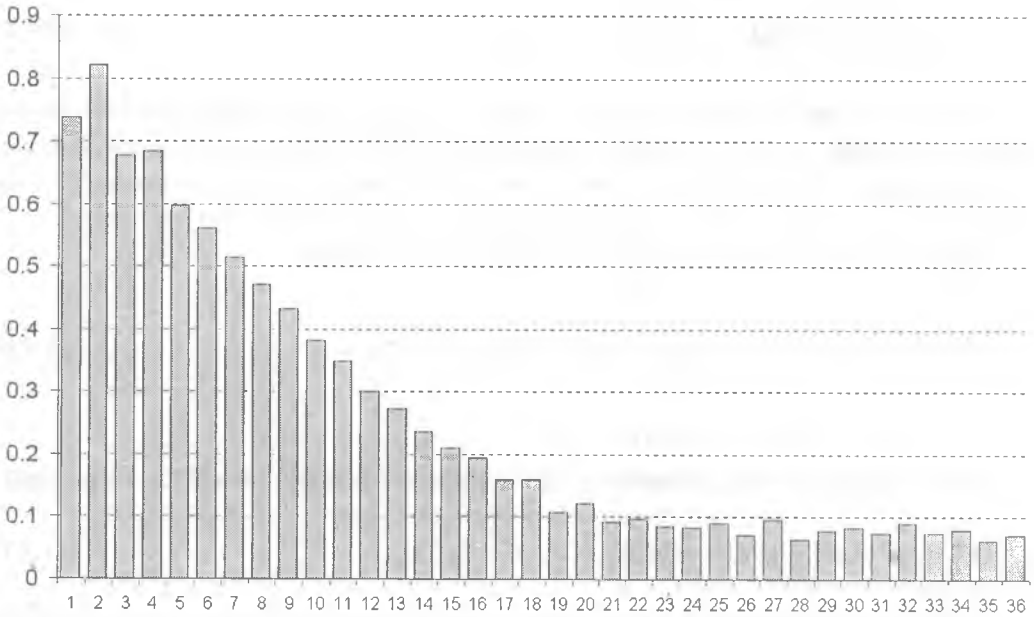


Рис. 5. Корелограма: $\alpha_1 = 0.3$, $\alpha_2 = 0.6$

Виразимо $A(L)$ як добуток двох поліномів першого порядку:

$$A(L) = 1 - \alpha_1 L - \alpha_2 L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L). \tag{1.53}$$

Зв'язок між α та λ можна виразити так: $\lambda_1 + \lambda_2 = \alpha_1$ та $\lambda_1 \lambda_2 = -\alpha_2$.

Змінну λ можна розглядати як корені $\lambda^2 - \alpha_1 \lambda - \alpha_2 = 0$, що є характеристичним рівнянням авторегресійного процесу другого порядку. Його корені відповідно дорівнюють:

$$\lambda_1, \lambda_2 = \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2}. \tag{1.54}$$

Якщо під коренем стоїть від'ємне число, тобто $a_1^2 + 4a_2 < 0$, то ми маємо пару комплексних коренів, які можуть бути записані так:

$$\lambda_1, \lambda_2 = a \pm b_i, \text{ де } a = \frac{a_1}{2}; b = \frac{1}{2} \sqrt{-(a_1^2 + 4a_2)}.$$

Абсолютне значення, або модуль, кожного комплексного кореня можна записати так:

$$|\lambda_j| = \sqrt{a^2 + b^2} = \sqrt{-\alpha_2}; j = 1, 2. \quad (1.55)$$

Для того, щоб розглянути, за яких значень параметрів AR(2)-процесу ми можемо отримати комплексні корені полінома лагового оператора, а отже й синусоїдальну автокореляційну функцію, повернемося до виразу (1.53) $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L)$, де $\lambda_1 + \lambda_2 = \alpha_1$; $\lambda_1 \lambda_2 = -\alpha_2$.

Представимо обернений до $A(L)$ поліном у вигляді:

$$A^{-1}(L) = \frac{1}{(1 - \lambda_1 L)(1 - \lambda_2 L)} = \frac{c}{(1 - \lambda_1 L)} + \frac{d}{(1 - \lambda_2 L)}, \quad (1.56)$$

де $c = -\lambda_1/(\lambda_2 - \lambda_1)$; $d = \lambda_2/(\lambda_2 - \lambda_1)$.

Враховуючи (1.56), запишемо AR(2)-процес як суму двох AR(1)-процесів

$$X_t = A^{-1}(L)\varepsilon_t = c/(1 - \lambda_1 L)\varepsilon_t + d/(1 - \lambda_2 L)\varepsilon_t. \quad (1.57)$$

З умов стаціонарності AR(1)-процесів випливають умови стаціонарності AR(2)-процесу, а саме:

$$|\lambda_1| < 1, |\lambda_2| < 1. \quad (1.58)$$

При цьому коли корені λ_1 та λ_2 набувають дійсних значень, то корелограма має вигляд спадної показникової функції. Якщо корені λ_1 та λ_2 є комплексними числами, то корелограма є спадною синусоїдальною функцією (див. рисунок 4). Формально це відбувається за умови:

$$|\lambda_j| = \sqrt{a^2 + b^2} = \sqrt{-\alpha_2} < 1; j = 1, 2, \quad (1.59)$$

що рівнозначна умові $0 < -\alpha_2 < 1$. (1.60)

Отже, коли другий коефіцієнт AR(2)-процесу знаходиться в межах $0 < -\alpha_2 < 1$, то можна однозначно очікувати спадну синусоїдальну форму автокореляційної функції.

Спадний як показниковий, так і синусоїдальний графік автокореляційної функції є графічним підтвердженням стаціонарності процесу, оскільки при цьому виконується основна умова стаціонарності (1.58): $|\lambda_1| < 1, |\lambda_2| < 1$.

Іншими словами, умова стаціонарності — це умова, за якої корені характеристичного рівняння лежать в межах одиничного кола. Для дійсних коренів це твердження є абсолютно зрозумілим. Проілюструємо його для комплексних коренів.

Кожний комплексний корінь AR(2)-процесу можна представити у вигляді:

$$\lambda_j = a + ib = \rho(\cos f + i \sin f), \text{ де } \operatorname{tg} f = b/a; \rho = \sqrt{a^2 + b^2} = |\lambda_j| < 1; j = 1, 2.$$

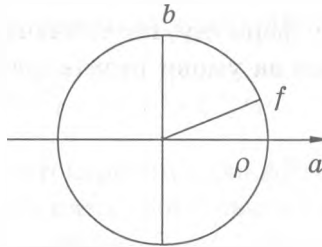


Рис. 6. Геометричне місце точок комплексних коренів характеристичного рівняння

Альтернативним є твердження, що корені полінома лагового оператора $A(Z) = 1 - \alpha_1 Z - \alpha_2 Z^2$ мають лежати поза межами одиничного кола. Покажемо, що ці твердження є альтернативними, тобто що між коренями характеристичного рівняння та коренями полінома лаг-оператора існує зворотний зв'язок. Корені $A(Z)$ ми можемо отримати, розв'язавши рівняння: $A(Z) = 1 - \alpha_1 Z - \alpha_2 Z^2 = (1 - \lambda_1 Z)(1 - \lambda_2 Z) = 0$. Корені: $Z_j = 1/\lambda_j$ ($j = 1, 2$), таким чином, якщо λ лежать в межах одиничного кола, Z лежать поза його межами. Умова стаціонарності, як правило, формулюється в літературі в іншому варіанті, а саме: корені полінома лагового оператора повинні знаходитись поза межами одиничного кола.

1.1.5.4. Поняття часткової автокореляційної функції

Часткова автокореляційна функція PARF (partial autocorrelation function) має велике значення в дослідженні авторегресійних процесів. Саме вигляд часткової автокореляційної функції допомагає у визначенні порядку авторегресійного процесу. Практично всі сучасні пакети прикладних програм видають графік часткової автокореляційної функції. Отже, розглянемо її поняття детальніше. Проілюструємо сутність цієї функції на прикладі AR(2)-та AR(3)-процесів.

Спочатку покажемо, що в AR(2)-процесі коефіцієнт a_2 і є значенням часткового коефіцієнта кореляції між x_t та x_{t-2} за умови, що x_{t-1} залишається незмінним. Для цього введемо певні позначення, а саме позначимо відповідні коефіцієнти кореляції між лаговими змінними через:

$$r_{12} = \operatorname{corr}(x_t, x_{t-1}); r_{23} = \operatorname{corr}(x_{t-1}, x_{t-2}); r_{13} = \operatorname{corr}(x_t, x_{t-2}), \quad (1.61)$$

де індекс 1 відповідає x_t , 2 — x_{t-1} , 3 — x_{t-2} .

Нагадаємо, що для стаціонарних процесів з урахуванням раніше введених позначень справедливо:

$$r_{12} = \text{corr}(x_t, x_{t-1}) = r_{23} = \text{corr}(x_{t-1}, x_{t-2}) = \rho_1, r_{13} = \text{corr}(x_t, x_{t-2}) = \rho_2. \quad (1.62)$$

Крім того, нагадаємо відому формулу, що визначає частковий коефіцієнт кореляції між двома змінними за умови незмінності третьої:

$$r_{13,2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{(1 - r_{12}^2)(1 - r_{23}^2)}}. \quad (1.63)$$

Формула (1.63) визначає частковий коефіцієнт кореляції між x_t та x_{t-3} за умови, що x_{t-2} залишається незмінним.

Підставимо (1.62): $r_{12} = r_{23} = \rho_1$; $r_{13} = \rho_2$ в (1.63) і отримаємо:

$$r_{13,2} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}. \quad (1.64)$$

Якщо повернутися до системи рівнянь Юла—Уокера для AR(2)-процесу (1.49) і розв'язати його для α_2 , після певного ланцюга перетворень отримаємо такий результат:

$$\left. \begin{aligned} \rho_1 &= \alpha_1 + \alpha_2 \rho_1 \\ \rho_2 &= \alpha_1 \rho_1 + \alpha_2 \end{aligned} \right\} \Rightarrow \left. \begin{aligned} \rho_1^2 &= \alpha_1 \rho_1 + \alpha_2 \rho_1^2 \\ \rho_2 &= \alpha_1 \rho_1 + \alpha_2 \end{aligned} \right\} \Rightarrow$$

$$\rho_1^2 - \rho_2 = \alpha_1 \rho_1 + \alpha_2 \rho_1^2 - \alpha_1 \rho_1 - \alpha_2 \Rightarrow \rho_1^2 - \rho_2 = \alpha_2 (\rho_1^2 - 1).$$

Звідки маємо:

$$\alpha_2 = -\left[(\rho_1^2 - \rho_2) / (\rho_1^2 - 1) \right] = (\rho_2 - \rho_1^2) / (1 - \rho_1^2) = r_{13,2}. \quad (1.65)$$

Формула (1.65) показує, що частковий коефіцієнт автокореляції для AR(2)-процесу дорівнює коефіцієнту α_2 цього процесу.

Відповідно, для AR(3)-процесу частковий коефіцієнт автокореляції між X_t та X_{t-3} (за умови, що X_{t-1} та X_{t-2} є незмінними) дорівнює коефіцієнту α_3 цього процесу, для AR(p)-процесу частковий коефіцієнт автокореляції між X_t та X_{t-p} (за умови, що $X_{t-1}, X_{t-2}, \dots, X_{t-p-1}$ є незмінними) дорівнює коефіцієнту α_p цього процесу.

Крім того, коефіцієнти часткової автокореляції мають дуже важливу властивість. Якщо ми маємо авторегресійний процес p -го порядку, то всі част-

кові коефіцієнти автокореляції з лагом більшим за p дорівнюють нулю. Отже, аналіз форми функції часткової автокореляції (PACF) допомагає при визначенні порядку авторегресійного процесу.

Проілюструємо це знову на прикладі процесу AR(2). Покажемо, що для нього частковий коефіцієнт автокореляції з лагом 3 та всі наступні дорівнюють нулю.

Для цього спочатку розглянемо авторегресійний процес третього порядку:

$$AR(3): X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \alpha_3 X_{t-3} + \varepsilon_t.$$

Частковий коефіцієнт автокореляції для AR(3)-процесу між X_t та X_{t-3} (за умови, що X_{t-1} , X_{t-2} є незмінними) дорівнює коефіцієнту α_3 цього процесу. Якщо процес є AR(2), то відповідно коефіцієнт $\alpha_3 = 0$. Для ілюстрації цього положення розглянемо систему Юла—Уокера для AR(3)-процесу, яка має такий вигляд:

$$\begin{aligned} \rho_1 &= \alpha_1 + \alpha_2 \rho_1 + \alpha_3 \rho_2; \\ \rho_2 &= \alpha_1 \rho_1 + \alpha_2 + \alpha_3 \rho_1; \\ \rho_3 &= \alpha_1 \rho_2 + \alpha_2 \rho_1 + \alpha_3. \end{aligned} \tag{1.66}$$

Якщо наш процес не є AR(3), а лише AR(2), то для нього коефіцієнти автокореляції визначаються за рекурентними формулами (1.52) та (1.51):

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}; \quad k = 3, 4, \dots$$

Відповідно коефіцієнт автокореляції третього порядку за цією формулою дорівнює:

$$\rho_3 = \alpha_1 \rho_2 + \alpha_2 \rho_1. \tag{1.67}$$

Підстановка (1.67) в (1.66) дає $\alpha_3 = 0$.

Таким чином, частковий коефіцієнт автокореляції (PACF) для AR(2)-процесу буде обриватись після другого лага, тобто всі часткові коефіцієнти автокореляції з лагом три та вище дорівнюватимуть нулю. Відповідно, для ряду AR(3) частковий коефіцієнт автокореляції буде обриватись після третього лага, а для AR(p)-процесу — після p -го. Отже, властивості ACF та PACF авторегресійних процесів можна представити так:

Процес	Поведінка ACF	Поведінка PACF
AR(p)	Нескінченний процес, що асимптотично наближається до нуля	Скінченний процес, дорівнює нулю після лага p

Як бачимо, завдяки цим властивостям можна легко визначити порядок авторегресійного процесу, а саме: якщо PACF процесу дорівнює нулю після k -го лага, то ми маємо відповідно $AR(k)$ -процес. ACF процесу дає уявлення про його стаціонарність, і якщо він є спадною функцією, то процес візуально можна вважати стаціонарним. Графік PACF авторегресійного процесу $AR(2)$ представлено на рис. 7.

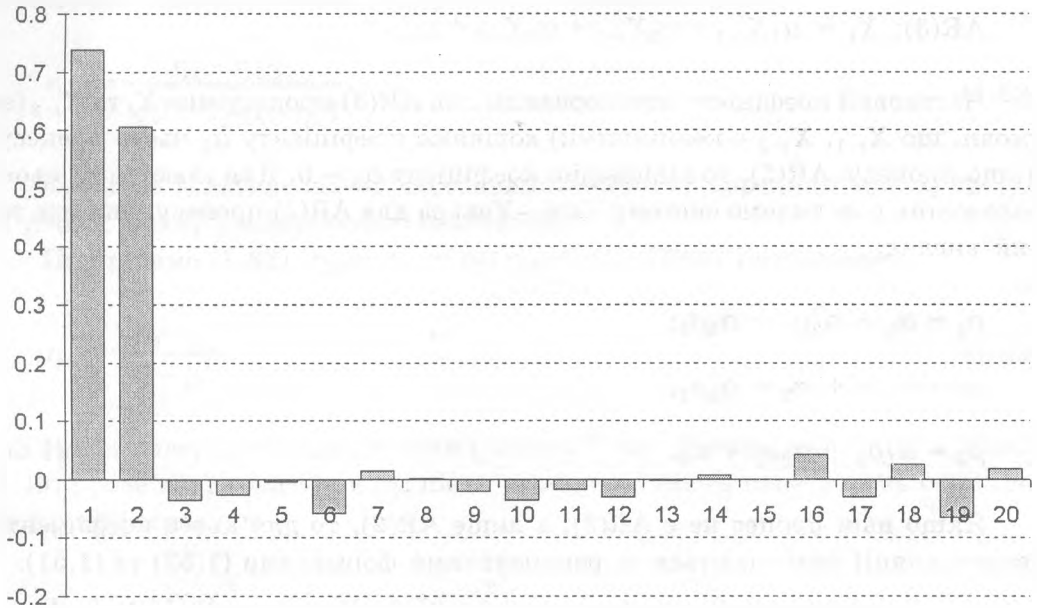


Рис. 7. Часткова корелограма PACF: $AR(2)$ -процес, $\alpha_1 = 0.3$, $\alpha_2 = 0.6$

1.1.5.5. Властивості процесів ковзного середнього (МА-процесів)

Властивості $MA(1)$ -процесу

Почнемо вивчення основних властивостей процесів ковзного середнього з розгляду найпростішого процесу першого порядку $MA(1)$, який має вигляд:

$$MA(1): Y_t = \mu + \varepsilon_t - \beta_1 \varepsilon_{t-1}, \quad (1.68)$$

де збурення ε_t є білим шумом: $E(\varepsilon_i) = 0$ для $\forall i$; $E(\varepsilon_i^2) = \sigma^2$ для $\forall i$; $E(\varepsilon_i \varepsilon_j) = 0$ для всіх $i \neq j$.

Як бачимо, в моделі ковзного середнього досліджуваний показник Y є комбінацією константи, поточного та всіх минулих значень випадкової величини (збурення), яка за припущенням вважається білим шумом.

Запишемо MA(1)-процес у спрощеному вигляді без константи, для чого введемо позначення $X_t = Y_t - \mu$, звідки:

$$\text{MA(1): } X_t = \varepsilon_t - \beta_1 \varepsilon_{t-1}. \quad (1.69)$$

Можна показати, що процеси ковзного середнього пов'язані з авторегресійними процесами. Для цього перепишемо AR(1)-процес таким чином:

$$X_t = \varepsilon_t + \alpha \varepsilon_{t-1} + \alpha^2 \varepsilon_{t-2} + \dots, \quad (1.70)$$

де ε_t — білий шум.

Отже, як бачимо, AR(1) може бути переписаний у вигляді MA-процесу нескінченного порядку MA(∞).

Визначимо коефіцієнти автоковаріації MA(1)-процесу. Для цього обидві частини (1.69) домножимо послідовно на X_t, X_{t-1}, X_{t-2} та знайдемо математичні сподівання обох частин. Використовуючи раніше введені позначення, відповідно отримаємо:

$$\gamma_0 = E(X_t^2) = E(\varepsilon_{t-2}^2 - \beta_1 \varepsilon_t \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-1}^2) = (1 + \beta_1^2) \sigma_\varepsilon^2. \quad (1.71)$$

$$\gamma_1 = E(X_t X_{t-1}) = -\beta_1 \sigma_\varepsilon^2. \quad (1.72)$$

$$\gamma_2 = E(X_t X_{t-2}) = 0; \gamma_3 = \gamma_4 = \dots = 0. \quad (1.73)$$

Зазначимо, що $\sigma_\varepsilon^2 = E(X_t^2) = E(X_{t-1}^2) = \dots$

Відповідно до (1.71), (1.72) та (1.73) коефіцієнти автокореляції MA(1)-процесу дорівнюють нулю після першого лага:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = -\frac{\beta_1}{(1 + \beta_1^2)}; \quad (1.74)$$

$$\rho_2 = \rho_3 = \dots = 0. \quad (1.75)$$

Зазначимо, що в свою чергу MA(1)-процес може бути записаний як авторегресійний процес нескінченного порядку AR(∞): $\varepsilon_t = X_t + \beta_1 X_{t-1} + \beta_1^2 X_{t-2} + \dots$, що можна переписати як

$$X_t = -\beta_1 X_{t-1} - \beta_1^2 X_{t-2} - \dots + \varepsilon_t. \quad (1.76)$$

Отже, MA(1) \sim AR(∞), а AR(1) \sim MA(∞). Тому властивості MA-процесу є

протилежними властивостям AR-процесу, а саме: для MA-процесу часткові коефіцієнти автокореляції не обриваються після лага, який дорівнює порядку процесу, а поступово зменшуються до нуля, а коефіцієнти автокореляції обриваються, тобто дорівнюють нулю, після лага, що є рівним порядку процесу. В загальному випадку поведінку корелограм (ACF) та часткових корелограм (PACF) можна підсумувати у вигляді таблиці.

Властивості ACF та PACF процесів ковзного середнього

Процес	Поведінка ACF	Поведінка PACF
AR(p)	Нескінченно зменшується до нуля	Скінченний: дорівнює нулю після лага p
MA(q)	Скінченний: дорівнює нулю після лага q	Нескінченно зменшується до нуля

Як бачимо, завдяки цим властивостям можна легко визначити порядок процесу ковзного середнього, а саме: якщо ACF процесу дорівнює нулю після k -го лага, то відповідно ми маємо MA(k)-процес. PACF процесу дає уявлення про оберненість (invertibility) процесу: якщо він є спадною функцією, то процес візуально можна вважати таким, який можна обернути.

Розглянемо поняття оберненості MA-процесів. MA-процес у формі (1.76) має сенс, лише коли $|\beta_1| < 1$, тобто коли MA-процес можна обернути, інакше найбільший вплив на поточне X має найвіддаленіше значення показника, що суперечить здоровому глузду. Отже, умова $|\beta_1| < 1$ забезпечує можливість знаходження оберненого процесу.

Умова оберненості для MA-процесів подібна до умови стаціонарності авторегресійних процесів, але на відміну від них не встановлює жодного обмеження для β_1 .

1.1.5.6. Властивості ARMA-процесів

ARMA-процесами є змішані авторегресійні процеси та процеси ковзного середнього. В загальному випадку ARMA(p, q)-процес має такий вигляд:

$$A(L)x_t = B(L)\varepsilon_t, \quad (1.77)$$

де p — порядок авторегресійного процесу; $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p$ — поліном лагового оператора авторегресійного процесу; q — порядок процесу ковзного середнього; $B(L) = 1 - \beta_1 L - \beta_2 L^2 - \dots - \beta_q L^q$ — поліном лагового оператора процесу ковзного середнього.

Для того, щоб ARMA(p, q)-процес був стаціонарним, необхідно, щоб корені полінома $A(L)$ лежали поза межами одиничного кола. Для того, щоб викону-

була умова оберненості (invertibility), необхідно, щоб корені полінома $B(L)$ також лежали поза межами одиничного кола.

Якщо обидві умови виконуються, то ARMA(p, q)-процес може бути виражений як чисті AR- або MA-процеси нескінченного порядку, а саме:

$$B^{-1}(L)A(L)X_t = \varepsilon_t, \text{ або } X_t = A^{-1}(L)B(L)\varepsilon_t, \quad (1.78)$$

де ε_t є білим шумом.

Найпростішим змішаним ARMA-процесом є ARMA(1,1)-процес:

$$\text{ARMA}(1,1): X_t = \alpha X_{t-1} + \varepsilon_t - \beta \varepsilon_{t-1}, \quad (1.79)$$

де ε_t є білим шумом.

Дисперсія ARMA(1,1)

Для знаходження дисперсії ARMA(1,1)-процесу піднесемо обидві частини (1.79) до квадрата, знайдемо математичне сподівання обох частин та, враховуючи раніше введені позначення, отримаємо:

$$\sigma_x^2 = \gamma_0 = \left[(1 - 2\alpha\beta - \beta^2)/(1 - \alpha^2) \right] \sigma_\varepsilon^2. \quad (1.80)$$

Для знаходження коефіцієнтів автоковаріації покроково помножимо на X_{t-1}, X_{t-2}, \dots та знайдемо математичне сподівання обох частин:

$$\gamma_1 = \alpha\gamma_0 - \beta\sigma_\varepsilon^2 = \left[(\alpha - \beta)(1 - \alpha\beta)/(1 - \alpha^2) \right] \sigma_\varepsilon^2. \quad (1.81)$$

Для коефіцієнтів автоковаріації вищого порядку, як легко побачити, справедлива така рекурентна формула:

$$\gamma_k = \alpha\gamma_{k-1}, \quad k = 2, 3, \dots \quad (1.82)$$

При цьому зазначимо, що коефіцієнт автоковаріації першого порядку залежить від параметрів як AR-, так і MA-процесів. Коефіцієнт автоковаріації другого та вищих порядків залежить тільки від параметра AR-процесу, він зменшується експоненційно з коефіцієнтом спадання, заданим параметром AR-процесу.

Коефіцієнти автокореляції ARMA(1,1)-процесу відповідно дорівнюють:

$$\rho_1 = \gamma_1/\gamma_0 = \alpha - \beta\sigma_\varepsilon^2/\gamma_0. \quad (1.83)$$

Для коефіцієнтів автокореляції вищих порядків справедливою є така рекурентна формула:

$$p_k = \alpha p_{k-1}, k = 2, 3, \dots \quad (1.84)$$

ARMA(p, q)-процеси еквівалентні:

— AR-процесам нескінченного порядку AR(∞): $B^{-1}(L)A(L)X_t = \varepsilon_t$,

— процесам ковзного середнього нескінченного порядку MA(∞):

$$X_t = A^{-1}(L) B(L)\varepsilon_t.$$

Таким чином, як ACF, так і PACF ARMA-процесу не будуть обриватися після певного лага, а нескінченно спадатимуть до нуля, що зробить надзвичайно складним визначення порядку ARMA-процесів.

Властивості ACF та PACF процесів авторегресії та ковзного середнього

<i>Поведінка ACF</i>	<i>Поведінка PACF</i>
Нескінченний процес, що асимптотично спадає до нуля	Нескінченний процес, що асимптотично спадає до нуля

1.1.6. Ідентифікація, оцінювання та тестування ARIMA(p, d, q)-моделей

1.1.6.1. Ідентифікація ARIMA-моделей

Як було показано в попередньому розділі, якщо часовий ряд можна описати змішаним авторегресійним процесом та процесом ковзної середньої (ARMA(p, q)), то корелограма та графік часткової автокореляційної функції не допомагають при визначенні порядку процесу. Тому потрібно використовувати складніші, формальні процедури для ідентифікації процесу. Нагадаємо, що ARIMA-методологія є застосовною лише для стаціонарних рядів, тому першим кроком в ідентифікації процесу є перевірка часового ряду на стаціонарність. Детально перевірку рядів на стаціонарність описано в розділі 3, а ми лише нагадаємо деякі основні аспекти цієї перевірки. Чому необхідно, щоб часові ряди були стаціонарними при ARIMA-моделюванні? Перш за все, ми використовуємо ці моделі для прогнозу, а прогнозувати можна поведінку лише тих процесів, основні характеристики яких, середня, дисперсія та коефіцієнти автокореляції не залежать від часу.

Якщо середня, варіація і автоковаріація (у різних проміжках) змінюються в залежності від часу, то часовий ряд є нестаціонарним і ми не можемо передбачити поведінку такого процесу. Проілюструємо це твердження на прикладі авторегресійного процесу першого порядку AR(1).

Нехай ми вважаємо, що досліджуваний часовий ряд може бути описаний моделлю авторегресійного процесу першого порядку:

$$\text{AR(1): } Y_t = m + \alpha Y_{t-1} + u_t, \tag{1.85}$$

де u_t є білим шумом.

Дослідимо, як ряд Y поводить себе протягом певного часу. Для цього перепишемо модель (1.85) для різних проміжків часу і отримаємо:

$$Y_1 = m + \alpha Y_0 + u_1$$

$$Y_2 = m + \alpha(m + \alpha Y_0 + u_1) + u_2$$

.....

$$Y_t = m(1 + \alpha + \alpha^2 + \dots + \alpha^{t-1}) + \alpha^t Y_0 + (u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \dots + \alpha^{t-1} u_1).$$

Припускаючи, що процес почався давно, можна переписати (1.85) як

$$Y_t = m(1 + \alpha + \alpha^2 + \dots + \alpha^{t-1}) + \alpha^t Y_0 + (u_t + \alpha u_{t-1} + \alpha^2 u_{t-2} + \dots + \alpha^{t-1} u_1). \tag{1.86}$$

Звідси, якщо $|\alpha| < 1$, то середнє $E(Y_t) = m(1 + \alpha + \alpha^2 + \dots)$, або $E(Y_t) = \mu = \frac{m}{1 - \alpha}$;

$$\text{дисперсія: } \text{Var}(Y_t) = \sigma_y^2 = \frac{\sigma^2}{1 - \alpha^2} = \gamma_0;$$

коефіцієнти автоковаріації:

$$\gamma_1 = E[(Y_t - \mu)(Y_{t-1} - \mu)] = \alpha \sigma_y^2,$$

$$\gamma_2 = \alpha^2 \sigma_y^2,$$

.....,

$$\gamma_k = \alpha^k \sigma_y^2;$$

$$\text{коефіцієнт автокореляції } \rho_k = \frac{\gamma_k}{\gamma_0}.$$

Якщо $|\alpha| < 1$, середня, варіація і коваріація ряду Y є константами і незалежні від часу, отже, процес AR(1) є стаціонарним.

Якщо $\alpha = 1$, то процес AR(1) є нестаціонарним, або, іншими словами, має одиничний корінь (unit root). У такому випадку (1.85) набуває вигляду:

$$Y_t = m + Y_{t-1} + u_t. \tag{1.87}$$

При $m \neq 0$ процес (1.87) є випадковим нестаціонарним процесом зі зміщенням (random walk with drift). Якщо в (1.87) $m = 0$, ми маємо простий

випадковий процес (simple random work). У нестационарних процесах середня та дисперсія необмежено зростають з часом. Дійсно, підставляючи в формулу (1.86), отримаємо вирази для середнього та дисперсії.

Середнє $E(Y_t|Y_0) = mt + Y_0$ зростає до нескінченності при $t \rightarrow \infty$;

дисперсія $\text{Var}(Y_t|Y_0) = E[(Y_t - E(Y_t|Y_0))^2] = E(u_t + u_{t-1} + u_{t-2} + \dots + u_1)^2 = t\sigma^2$ зростає до нескінченності при $t \rightarrow \infty$.

Якщо ряд є нестационарним, ми не можемо знайти сталих середньої та дисперсії, тому потрібно шукати можливі перетворення ряду, які можуть звести його до стаціонарного. Такими перетвореннями і є операція різниць.

Якщо $|\alpha| > 1$, то ряд Y є вибуховим, що є надзвичайно рідкісним явищем в економічних часових рядах. Зазначимо, що вибуховий ряд не можна звести до стаціонарного ніякою кількістю операцій різниць.

Наприклад, розглянемо модель:

$$(1 - \alpha L) Y_t = m + \varepsilon_t. \quad (1.88)$$

Якщо $|\alpha| = 0.67 < 1$, то процес буде стаціонарним, у даному випадку AR(1)-процесом. При $|\alpha| = 1$ процес буде нестационарним, але його можна операцією різниць звести до стаціонарного. При $|\alpha| = 1.45 > 1$ процес буде вибуховим, його не можна операцією різниць звести до стаціонарного.

Крім того, в загальному випадку модель (1.88) може включати часовий тренд:

$$(1 - \alpha L)(Y_t - \delta_0 - \delta_1 t) = \varepsilon_t. \quad (1.89)$$

Якщо в моделі (1.89) $|\alpha| < 1$, то ряд, який можна представити цією моделлю, називають тренд-стаціонарним часовим рядом (TS, trend stationary), якщо ж $\alpha = 1$, то часовий ряд є нестационарним і називається різницево-стаціонарним (DS, difference stationary).

Тренд-стаціонарні часові ряди коливаються навколо певного тренду (в найпростішому випадку навколо лінійного), при цьому спостерігається постійне збільшення рівня середньої (mean), що і робить ряд формально нестационарним, але якщо виокремити тренд, то ряд стає стаціонарним.

Проілюструємо це твердження. Для цього модель (1.89) перепишемо у більш звичному вигляді, а саме:

$$Y_t = \delta_0 + \delta_1 t + u_t, \quad (1.90)$$

$$u_t = \alpha u_{t-1} + \varepsilon_t, \quad (1.91)$$

де збурення ε_t є білим шумом, при $|\alpha| < 1$ вираз (1.91) є AR(1)-процесом.

Після елементарних перетворень (1.90) представимо у вигляді

$$Y_t = [\delta_0(1 - \alpha) + \alpha\delta_1] + \delta_1(1 - \alpha)t + \alpha Y_{t-1} + \varepsilon_t. \quad (1.92)$$

Введемо заміну змінних і отримаємо:

$$Y_t = \beta_1 + \beta_2 t + \alpha Y_{t-1} + \varepsilon_t, \quad (1.93)$$

де $\beta_1 = [\delta_0(1 - \alpha) + \alpha\delta_1]$; $\beta_2 = \delta_1(1 - \alpha)$.

Оскільки $|\alpha| < 1$, то (1.93) є AR(1)-процесом + часовий тренд (в даному випадку лінійний).

Операція перших різниць дасть чистий стаціонарний процес, а саме:

$$\Delta Y_t = \delta_1 + \Delta u_t, \quad (1.94)$$

де $\Delta Y_t \sim I(0)$, оскільки $\Delta u_t \sim I(0)$. Нагадаємо, що u_t є стаціонарним AR(1)-процесом при $|\alpha| < 1$, а операція різниць стаціонарного процесу знову дає стаціонарний процес.

При $\alpha = 1$ процес (1.90)–(1.91) є нестаціонарним, але його можна звести до стаціонарного операцією перших різниць (в даному найпростішому випадку). Справді, при $\alpha = 1$ з (1.92) отримаємо:

$$\Delta Y_t = \delta_1 + \varepsilon_t. \quad (1.95)$$

Оскільки ε_t є білим шумом, тобто $\varepsilon_t \sim I(0)$, то і $\Delta Y_t \sim I(0)$. Отже, операцією різниць ми можемо нестаціонарний процес звести до стаціонарного.

Продемонструємо це для загального випадку. Загальний ARMA(p, q)-процес запишемо з включенням лінійного тренду:

$$Y_t - \delta_0 - \delta_1 t = u_t;$$

$$A(L)u_t = B(L)\varepsilon_t, \quad (1.96)$$

де $A(L)$ і $B(L)$ є багаточленами порядку p і q в лаговому операторі, збурення ε_t є білим шумом, $\delta_1 \neq 0$. Коли всі корені полінома лагового оператора авторегресійного процесу $A(L)$ знаходяться поза одиничним колом, ми маємо стаціонарний ARMA(p, q)-процес, або, іншими словами, TS-процес. Якщо поліном лагового оператора $A(L)$ має хоча б один одиничний корінь, ми маємо нестаціонарний (DS) процес.

В цьому випадку, якщо вважати, що (1.96) має один одиничний корінь, ми можемо записати лаг-поліном $A(L)$ у вигляді добутку двох поліномів:

$$A(L) = (1 - L)(1 - \lambda_2 L) \dots (1 - \lambda_p L) = (1 - L)A^*(L), \quad (1.97)$$

де всі $(p - 1)$ корені $A^*(L)$ лежать поза межами одиничного кола. Враховуючи (1.97), оператор перших різниць перетворить (1.96) в стаціонарний процес вигляду:

$$A^*(L)(\Delta Y_t - \delta_1) = B(L)\varepsilon_t, \quad (1.98)$$

де $\Delta Y_t \sim I(0)$, оскільки всі корені полінома $A^*(L)$ лежать поза межами одиничного кола. Отже, перші різниці ряду можуть бути змодельовані як стаціонарний ARMA($p - 1, q$)-процес.

Зауважимо, що якщо поліном $A(L)$ в (1.96) матиме не один, а два одиничних корені, то аналогічні перетворення за допомогою оператора других різниць дадуть нам стаціонарний процес, тобто $\Delta^2 Y_t$ буде стаціонарним процесом. Рекурсивно можна показати, що кількість одиничних коренів дорівнює кількості різниць, необхідних для того, щоб звести ряд до стаціонарного.

В сучасній літературі, як ми вже зазначали, нестаціонарні ряди прийнято називати інтегрованими. Ступінь інтегрування — це мінімальне число разів, при якому ряд операцією різниць зводиться до стаціонарного. Отже, першим кроком в ідентифікації ARIMA(p, d, q)-процесів є визначення порядку інтегрування (d). Порядок інтеграції визначається за допомогою тесту Дікі—Фуллера, який детально було розглянуто раніше. Після визначення порядку інтеграції ряд відповідною кількістю разів зводиться різницями до стаціонарного, після чого подальшим кроком є визначення порядку авторегресійної складової (p) та складової ковзного середнього (q) перетвореного ряду.

Якщо первісний ряд є стаціонарним, то ми маємо лише ARMA(p, q)-процес, і нашим наступним кроком є визначення порядку авторегресійної складової (p) та складової ковзного середнього (q).

Повторимо, що визначення порядку ARMA(p, q)-процесу є досить складним. Ми опишемо одну з процедур, процедуру Хеннона та Ріссанена (Hannan and Rissanen), яка певним чином дозволяє формалізувати цей процес.

1.1.6.2. Процедура Хеннона та Ріссанена для визначення порядку p та q ARMA(p, q)-процесу. Оцінювання та тестування остаточного вигляду ARIMA-моделі

Процедура складається з певних послідовних етапів.

Етап 1. Визначення порядку авторегресійної складової процесу (p).

Для цього на першому кроці оцінюється декілька специфікацій AR-

процесу, наприклад AR(1), AR(2), ..., AR(k):

$$AR(k): \Delta Y_t = m + \alpha_1 Y_{t-1} + \dots + \alpha_k Y_{t-k} + u_t, \quad (1.99)$$

де u_t є білим шумом.

Значимо, що на цьому етапі невідомі коефіцієнти авторегресійних процесів оцінюються методом найменших квадратів (МНК).

На другому кроці з множини оцінених AR(1)-, AR(2)-, ..., AR(k)-моделей обирається одна з найменшим значенням інформаційного критерію Акайка (Akaike Information Criterion, далі AIC- або Акайк-критерій):

$$AIC = \frac{2k}{n} + \log \left(\frac{1}{n} u'u \right), \quad (1.100)$$

де n — кількість спостережень в часовому ряді; k — число параметрів авторегресійного процесу, яке потрібно оцінити, u — збурення AR-процесу.

Зауважимо, що чим меншим є значення AIC-критерію, тим кращою вважається оцінена модель. Концептуально AIC-критерій побудований таким чином, що за кожний додатковий лаг, тобто вищий порядок AR-моделі, вводиться штраф. Крім того, зазначимо, що помилки (\hat{u}_t) моделі (1.99) будуть вхідними даними для специфікації MA-складової на другому етапі.

Етап 2. *Визначення порядку складової ковзного середнього q , тобто специфікації ARMA(p, q)-процесу, коли в залежності від обраної на першому кроці специфікації AR-процесу (зазначимо, що їх може бути й декілька) починаємо поступово будувати ARMA, підключаючи різні порядки ковзної середньої, наприклад ARMA($p, 1$), ARMA($p, 2$), ... При цьому в моделях використовуються помилки, які отримані з остаточно обраної на першому етапі AR(p)-моделі. Наприклад, для специфікації ARMA($p, 1$) нам потрібно оцінити невідомі параметри такої моделі:*

$$Y_t = m + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \hat{u}_t - \beta_1 \hat{u}_{t-1} + \varepsilon_t, \quad (1.101)$$

де ε_t є білим шумом.

Після оцінки послідовності ARMA-специфікацій методом найменших квадратів обирається найкраща за найменшим значенням інформаційного критерію Шварца (Schwarz Information Criterion, далі SIC- або Шварц-критерій):

$$SIC = \frac{k \ln n}{n} + \ln \left(\frac{1}{n} \varepsilon' \varepsilon \right), \quad (1.102)$$

де k — кількість невідомих параметрів моделі, ε — збурення оцінюваних моделей типу (1.101); n — кількість спостережень часового ряду. Зауважимо, що SIC-критерій потужніший за AIC-критерій. Штрафи, що вводяться за додаткові коефіцієнти моделі, в цьому критерії є більшими, хоча за логікою побудови обидва критерії є однаковими. Найменше значення Шварц-критерію вважається найкращим.

Обрана ARMA(p, q)-модель (їх може бути й декілька) дає лише ідею про можливий порядок авторегресійної складової та ковзної середньої. Для того, щоб використовувати модель для подальшої роботи, її потрібно переоцінити. Нагадаємо, що на кожному проміжному етапі модель оцінювалась методом найменших квадратів, який в даному випадку може дати зміщені оцінки.

Етап 3. Оцінювання ARMA-моделі. Для цього, як правило, використовується нелінійний метод найменших квадратів, хоча можливо застосування методу максимальної правдоподібності. Після оцінювання ARMA-моделей вказаним вище методом (наприклад, по умовчання в пакеті E.VIEWS для оцінювання ARMA-процесів обирається нелінійний МНК) проводиться діагностичне тестування побудованої моделі.

Етап 4. Тестування побудованої моделі. Її можна вважати остаточною, якщо залишки (помилки) цієї моделі є білим шумом. Якщо ні, то процедуру Хеннона та Ріссанена потрібно повторити спочатку, тому взагалі ця процедура є досить складним ітеративним процесом.

Тестування залишків на білий шум може відбуватись за допомогою аналізу графіка корелограми та часткових коефіцієнтів автокореляції. При цьому вони повинні мати неструктурований вигляд та знаходитись в межах певного інтервалу довіри. В багатьох пакетах прикладних програм таким обирається 95 %-й інтервал. У припущенні, що залишки розподілені нормально, 95 %-й інтервал довіри для будь-якого ρ_k буде дорівнювати $(0 \pm 1,96)$. Тому, якщо обчислене значення ρ_k потрапляє в цей інтервал, гіпотеза, що справжнє $\rho_k = 0$, не відкидається. Але якщо воно потрапить за межі інтервалу довіри, то ми відкидаємо гіпотезу, що справжнє $\rho_k = 0$, з імовірністю помилки 5 % (див. таблиці 1.1—1.2).

Формальним критерієм перевірки залишків на білий шум є Q-статистика Бокса—Пірса—Люнга (Box—Pierce—Ljung), яка перевіряє загальну гіпотезу про одночасну рівність всіх коефіцієнтів автокореляції (ρ_k) нулю. Ljung—Box (LB)-статистика, або Box—Pierce—Ljung Q-статистика, розраховується за формулою:

$$Q = n(n+2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n-k} \right) \sim \chi_{(m-k)}^2, \quad (1.103)$$

Таблиця 1.1. Корелограма білого шуму

ACF	PACF	Лаг	AC	PAC	Q-Stat	Prob
.	.	1	-0.038	-0.038	1.4801	0.224
.	.	2	0.046	0.045	3.6251	0.163
.	.	3	0.052	0.055	6.3025	0.098
.	.	4	0.018	0.020	6.6230	0.157
.	.	5	0.024	0.021	7.2191	0.205
.	.	6	-0.001	-0.004	7.2196	0.301
.	.	7	0.026	0.021	7.8762	0.344
.	*	8	-0.057	-0.058	11.104	0.196
.	.	9	0.055	0.048	14.163	0.117
.	.	10	-0.046	-0.041	16.334	0.090

Таблиця 1.2. Корелограма нестационарного ряду

ACF	PACF	Лаг	AC	PAC	Q-Stat	Prob
*****	*****	1	0.985	0.985	970.88	0.000
*****		2	0.971	0.048	1916.3	0.000
*****		3	0.958	0.010	2837.3	0.000
*****		4	0.945	0.002	3734.7	0.000
*****	*	5	0.931	-0.062	4605.6	0.000
*****		6	0.917	0.003	5451.2	0.000
*****		7	0.903	-0.004	6272.1	0.000
*****		8	0.888	-0.026	7067.7	0.000
*****		9	0.876	0.057	7841.7	0.000
*****		10	0.862	-0.047	8591.9	0.000

У таблицях 1.1 та 1.2 AC — коефіцієнт автокореляції, PACF — коефіцієнт часткової автокореляції, Q-Stat — Q-статистика, Prob — імовірність того, що коефіцієнт автокореляції до k -го лага включно дорівнює нулю.

де n — кількість спостережень; k — кількість оцінених параметрів в ARMA-моделі; m — кількість коефіцієнтів автокореляції, які перевіряються на одночасну рівність нулю. Розрахована Q-статистика має χ^2 (хі-квадрат)-розподіл зі ступенями свободи $(m - k)$.

Нульовою гіпотезою при цьому буде гіпотеза, яка перевіряє одночасну рівність m коефіцієнтів автокореляції нулю, проти альтернативної гіпотези, що хоча б один з коефіцієнтів не дорівнює нулю:

$$\begin{cases} H_0: \hat{\rho}_1 = \hat{\rho}_2 = \dots = \hat{\rho}_m = 0 \\ H_1: \text{хоча б один з них} \neq 0. \end{cases}$$

Правило прийняття рішень: ми приймаємо⁶ нульову гіпотезу (H_0), якщо розрахункове значення Q-статистики менше за критичне значення, знайдене за таблицями χ^2 -розподілу при ступенях свободи ($m - k$) та заданому рівні статистичної значимості (α), тобто $Q < \chi_{\alpha, (m-k)}^2$, або $p\text{-value} > 10\%$. Якщо нульова гіпотеза прийнята, то залишки вважаються білим шумом і оцінена ARMA-модель може використовуватись для прогнозу.

Зазначимо, що ми можемо використати тести і для оцінених коефіцієнтів моделі, наприклад, провести тестування значимості змінних, включених в модель, або певної підмножини цих змінних. Можна також протестувати ефект додавання однієї чи більше змінних до моделі. Крім того, при проведенні процедури Хеннона та Ріссанена можна включати в модель сезонні складові.

1.1.7. Прогнозування за ARIMA-моделями

1.1.7.1. Основна ідея прогнозування за ARIMA-моделями

Основною метою оцінювання ARIMA-моделей є їх подальше використання для прогнозування.

Загальна ARIMA(p, d, q)-модель може бути записана у вигляді, що включає константу

$$A(L)\Delta^d Y_t = m + B(L)\varepsilon_t, \quad (1.104)$$

$$\text{або без константи } A(L)\Delta^d Y_t = B(L)\varepsilon_t, \quad t = 1, 2, \dots, n. \quad (1.105)$$

Відповідно прогноз Y_t в період часу ($n + s$) позначимо через \hat{Y}_{n+s} або $\hat{Y}_n(s)$. При цьому $s \geq 1$ є прогнозним періодом або періодом випередження.

При прогнозуванні можливі два типи помилок, пов'язаних

— з різницею між дійсними та оціненими коефіцієнтами моделі, що використовується для прогнозу;

— з ігноруванням випадкових величин (майбутніх збурень).

Класично береться до уваги лише другий тип помилок, який за введеними нами позначеннями набуває вигляду:

$$e_{n+s} = Y_{n+s} - \hat{Y}_{n+s}, \quad (1.106)$$

де e_{n+s} — помилка прогнозу; Y_{n+s} — невідоме дійсне значення Y в ($n + s$)-й

⁶ Точніше, не відкидаємо.

період часу; \hat{Y}_{n+s} — прогноз значення Y_{n+s} , зроблений на основі оціненої ARIMA-моделі при кількості спостережень n .

При цьому звичайно метою прогнозування за ARIMA-моделями є передбачення майбутніх значень часового ряду з мінімально можливою помилкою. Враховуючи, що значення помилок можуть бути як від'ємними, так і додатними, використовується поняття мінімуму середнього квадрата помилок (MSE). Відповідно оптимальним є прогноз, при якому мінімізується середній квадрат помилок MSE-прогнозу, тобто обирається таке прогнозне значення \hat{Y}_{n+s} , при якому мінімізується $E(e_{n+s}^2) = E\{(Y_{n+s} - \hat{Y}_{n+s})^2\}$. Зауважимо, що оскільки помилка прогнозу є випадковою величиною, ми мінімізуємо математичне сподівання квадрата помилок.

Можна довести (ми наводимо це твердження без доведення), що мінімізація середнього квадрата помилок прогнозного значення \hat{Y}_{n+s} означає, що воно є не що інше, як умовне сподівання Y_{n+s} за всіх спостережень часового ряду до періоду n , тобто $\hat{Y}_{n+s} = E(Y_{n+s} | Y_n, Y_{n-1}, \dots, Y_1)$.

Прогнозне значення \hat{Y}_{n+s} може бути отримане рекурсивно з оціненої ARIMA-моделі. Проілюструємо загальний принцип розрахунку прогнозу. Нехай ми маємо ARIMA(p, d, q)-модель, яка означає, що первісний ряд Y є нестационарним та його зведено до стаціонарного операцією різниць d разів. Позначимо через z_t зведений до стаціонарного часовий ряд, члени якого $z_t = \Delta^d Y_t$. Тоді ARIMA(p, d, q)-модель цього ряду має вигляд:

$$z_t = m + \alpha_1 z_{t-1} + \dots + \alpha_p z_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \dots - \beta_q \varepsilon_{t-q}, \quad t = 1, 2, \dots, n. \quad (1.107)$$

Розрахуємо за моделлю (1.106) прогноз на s періодів (\hat{z}_{n+s}). Почнемо розрахунок з прогнозу на один період вперед, що буде умовним математичним сподіванням значення z_{n+1} в моделі (1.107):

$$\hat{z}_{n+1} = E(z_{n+1} | z_n, \dots, z_1) = \alpha_1 z_n + \dots + \alpha_p z_{n-p+1} - \beta_1 \hat{\varepsilon}_n - \dots - \beta_q \hat{\varepsilon}_{n-q+1} + m, \quad (1.108)$$

де — $E(z_{n+1} | z_n) = 0, E(\varepsilon_n) = \hat{\varepsilon}_n = e_n, E(\varepsilon_{n-1}) = \hat{\varepsilon}_{n-1} = e_{n-1}, \dots$ — помилки, отримані за оціненою моделлю (1.107). Відповідно:

$$\hat{z}_{n+2} = E(z_{n+2} | z_n, \dots, z_1) = \alpha_1 \hat{z}_{n+1} + \alpha_2 z_n + \dots + \alpha_p z_{n-p+2} - \beta_2 \hat{\varepsilon}_n - \dots - \beta_q \hat{\varepsilon}_{n-q+2} + m,$$

.....

$$\hat{z}_{n+s} = E(z_{n+s} | z_n, \dots, z_1) = \alpha_1 \hat{z}_{n+s-1} + \dots + \alpha_p z_{n-p+s} - \beta_s \hat{\varepsilon}_n - \beta_q \hat{\varepsilon}_{n-q+s} + m. \quad (1.109)$$

Зауважимо, що якщо $s > p$ та $s > q$, то (1.109) набуває вигляду:

$$\hat{z}(s) = \hat{z}_{n+s} = \alpha_1 \hat{z}_{n+s-1} + \alpha_2 \hat{z}_{n+s-2} + \dots + \alpha_p \hat{z}_{n+s-p} + m.$$

Після того як отримано прогнозні значення перетвореного часового ряду z_t , можемо знайти прогнозні значення первісного ряду Y шляхом додавання z_t d разів.

Якщо Y має порядок інтеграції один, тобто $d = 1$, тоді прогноз на s періодів вперед для ряду Y можна отримати так:

$$\hat{Y}_{n+s} = Y_n + \hat{z}_{n+1} + \hat{z}_{n+2} + \dots + \hat{z}_{n+s}.$$

Якщо Y має порядок інтеграції два, тобто $d = 2$, тоді прогноз на s періодів вперед для ряду Y можна отримати так:

$$\begin{aligned} \hat{Y}_{n+s} &= Y_n + [\Delta Y_n + \hat{z}_{n+1}] + [\Delta Y_n + \hat{z}_{n+1} + \hat{z}_{n+2}] + \dots + [\Delta Y_n + \hat{z}_{n+1} + \dots + \hat{z}_{n+s}] = \\ &= Y_n + s\Delta Y_n + s\hat{z}_{n+1} + (s-1)\hat{z}_{n+2} + \dots + \hat{z}_{n+1}. \end{aligned} \quad (1.110)$$

Аналогічно (1.110) можна отримати прогноз для ряду Y з порядком інтеграції d .

Розглянемо детальніше деякі властивості прогнозів за найпростішими моделями.

1.1.7.2. Властивості прогнозів за найпростішими ARIMA-моделями

AR-процеси

Розглянемо властивості прогнозів AR(1)-процесу. Запишемо його у вигляді

$$\text{AR}(1): Y_t - \mu = \alpha(Y_{t-1} - \mu) + \varepsilon_t \quad |\alpha| < 1 \quad (1.111)$$

або

$$Y_t = \mu(1 - \alpha) + \alpha Y_{t-1} + \varepsilon_t, \quad (1.112)$$

де μ — середнє процесу; ε_t — білий шум.

Позначимо, як і раніше, Y_{n+s} — невідоме дійсне значення Y в $(n+s)$ -й період часу; \hat{Y}_{n+s} — прогнозне значення Y_{n+s} , $e_{n+s} = Y_{n+s} - \hat{Y}_{n+s}$ — помилка прогнозу. Оптимальним буде прогноз, який мінімізує середній квадрат помилок, тобто MSE-прогноз. Нагадаємо, що такий прогноз дорівнює прогнозному значенню \hat{Y}_{n+s} , отриманому як умовне сподівання Y_{n+s} за всіх спостережень часового ряду до періоду n , тобто $\hat{Y}_{n+s} = E(Y_{n+s} | Y_n, Y_{n-1}, \dots, Y_1)$.

Таким чином, якщо дійсне значення за AR(1)-моделлю в період часу $(n+1)$

$$Y_{n+1} - \mu = \alpha(Y_n - \mu) + \varepsilon_{n+1} \Rightarrow Y_{n+1} = (1 - \alpha)\mu + \alpha Y_n + \varepsilon_{n+1}, \quad (1.113)$$

то відповідно прогнозне значення в період часу $(n + 1)$

$$\hat{Y}_{n+1} = E(Y_{n+1}|Y_n, Y_{n-1}, \dots, Y_1) = (1 - \alpha)\mu + \alpha Y_n, \quad (1.114)$$

оскільки $E(\varepsilon_{n+1}) = 0$.

Перепишемо (1.114) у вигляді:

$$\hat{Y}_{n+1} - \mu = \alpha(Y_n - \mu). \quad (1.115)$$

З (1.115) легко побачити, що відхилення прогнозного значення від середнього процесу в період часу $(n + 1)$ буде дорівнювати частці відхилення дійсного значення в період n від середнього.

Помилка прогнозу в період часу $(n + 1)$ становить $e_{n+1} = Y_{n+1} - \hat{Y}_{n+1} = \varepsilon_{n+1}$, а дисперсія помилки: $\text{Var}(e_{n+1}) = \text{Var}(\varepsilon_{n+1}) = \sigma^2$.

Дійсне значення за AR(1)-моделлю в період часу $(n + 2)$

$$\begin{aligned} Y_{n+2} &= (1 - \alpha)\mu + \alpha Y_{n+1} + \varepsilon_{n+2} = (1 - \alpha)\mu + \alpha [(1 - \alpha)\mu + \alpha Y_n + \varepsilon_{n+1}] + \varepsilon_{n+2} = \\ &= (1 - \alpha^2)\mu + \alpha^2 Y_n + \alpha \varepsilon_{n+1} + \varepsilon_{n+2}. \end{aligned}$$

Прогнозне значення Y в період $(n + 2)$

$$\hat{Y}_{n+2} = E(Y_{n+2}|Y_n, Y_{n-1}, \dots, Y_1) = (1 - \alpha^2)\mu + \alpha^2 Y_n, \text{ або}$$

$$\hat{Y}_{n+2} - \mu = \alpha^2(Y_n - \mu) = \alpha(\hat{Y}_n - \mu). \quad (1.116)$$

Помилка прогнозу

$$e_{n+2} = Y_{n+2} - \hat{Y}_{n+2} = \alpha \varepsilon_{n+1} + \varepsilon_{n+2}.$$

$$\text{Дисперсія помилки } \text{Var}(e_{n+2}) = (1 + \alpha^2)\sigma_\varepsilon^2.$$

Відповідно для періоду $(n + s)$ отримаємо:

$$Y_{n+s} = (1 - \alpha^s)\mu + \alpha^s Y_n + (\varepsilon_{n+s} + \alpha \varepsilon_{n+s-1} + \dots + \alpha^{s-1} \varepsilon_{n+1}).$$

$$\hat{Y}_{n+s} = E(Y_{n+s}|Y_n, Y_{n-1}, \dots, Y_1) = (1 - \alpha^s)\mu + \alpha^s Y_n, \text{ або}$$

$$\hat{Y}_{n+s} - \mu = \alpha^s (Y_n - \mu). \quad (1.117)$$

Помилка прогнозу

$$e_{n+s} = Y_{n+s} - \hat{Y}_{n+s} = \varepsilon_{n+s} + \alpha\varepsilon_{n+s-1} + \dots + \alpha^{s-1}\varepsilon_{n+1}. \quad (1.118)$$

$$\text{Дисперсія помилки } \text{Var}(e_{n+s}) = (1 + \alpha^2 + \alpha^4 + \dots + \alpha^{2(s-1)})\sigma_\varepsilon^2. \quad (1.119)$$

Висновки

Як легко побачити з (1.116) та (1.118), прогнозне значення, отримане за AR(1)-процесами, зі зростанням періоду випередження прямує до середнього, тобто $\hat{Y}_{n+s} \rightarrow \mu$ при $s \rightarrow \infty$, а дисперсія помилки прогнозу прямує до дисперсії процесу:

$$\text{Var}(e_{n+s}) \rightarrow \frac{\sigma^2}{1 - \alpha^2} = \sigma_y^2 \text{ при } s \rightarrow \infty.$$

МА(1)-процес

Розглянемо властивості прогнозів МА(1)-процесу. Запишемо його у вигляді

$$\text{МА(1): } \hat{Y}_t = \mu + \varepsilon_t - \beta\varepsilon_{t-1}, \quad (1.120)$$

де μ — середнє процесу; ε_t — білий шум.

Позначимо, як і раніше, Y_{n+s} — невідоме дійсне значення Y в $(n+s)$ -й період часу; \hat{Y}_{n+s} — прогнозне значення Y_{n+s} , $e_{n+s} = Y_{n+s} - \hat{Y}_{n+s}$ — помилка прогнозу. Як і раніше, оптимальним прогнозом буде MSE-прогноз. Нагадаємо, що такий прогноз дорівнює прогнозованому значенню \hat{Y}_{n+s} , отриманому як умовне сподівання Y_{n+s} в залежності від усіх спостережень часового ряду до періоду n , тобто $\hat{Y}_{n+s} = E(Y_{n+s} | Y_n, Y_{n-1}, \dots, Y_1)$.

Таким чином, якщо дійсне значення за МА(1)-моделлю в період часу $(n+1)$

$$Y_{n+1} = \mu + \varepsilon_{n+1} - \beta\varepsilon_n,$$

то відповідно прогнозне значення в період часу $(n+1)$

$$\hat{Y}_{n+1} = E(Y_{n+1} | Y_n, Y_{n-1}, \dots, Y_1) = \mu - \beta\varepsilon_n, \quad (1.121)$$

оскільки $E(\varepsilon_{n+1}) = 0$. З (1.121) легко побачити, що прогнозне значення в період часу $(n+1)$ дорівнюватиме середньому процесу за виключенням частки помилки в попередній період.

Помилка прогнозу в період часу $(n + 1)$ становить $e_{n+1} = Y_{n+1} - \hat{Y}_{n+1} = \varepsilon_{n+1}$, а дисперсія помилки $\text{Var}(e_{n+1}) = \text{Var}(\varepsilon_{n+1}) = \sigma^2$.

Дійсне значення Y в $(n + 2)$ -й період

$$Y_{n+2} = \mu + \varepsilon_{n+2} - \beta\varepsilon_{n+1}.$$

Прогнозне значення Y в $(n + 2)$ -й період

$$\hat{Y}_{n+2} = E(Y_{n+2}|Y_n, Y_{n-1}, \dots, Y_1) = \mu, \tag{1.122}$$

оскільки $E(\varepsilon_{n+1}) = 0$, $E(\varepsilon_{n+2}) = 0$.

Помилка прогнозу в період $(n + 2)$

$$e_{n+2} = Y_{n+2} - \hat{Y}_{n+2} = \varepsilon_{n+2} - \beta\varepsilon_{n+1}.$$

Дисперсія помилки прогнозу в період $(n + 2)$ відповідно

$$\text{Var}(e_{n+2}) = E[(e_{n+2})^2] = (1 + \beta^2)\sigma_\varepsilon^2 = \sigma_y^2. \tag{1.123}$$

Як можна побачити з (1.123), дисперсія прогнозу помилки в $(n + 2)$ -й період дорівнює дисперсії величини Y .

Для періоду $(n + s)$ отримаємо:

$$Y_{n+s} = \mu + \varepsilon_{n+s} - \beta\varepsilon_{n+s-1},$$

$$\hat{Y}_{n+s} = E(Y_{n+s}|Y_n, Y_{n-1}, \dots, Y_1) = \mu, \text{ для } s \geq 2. \tag{1.124}$$

Помилка прогнозу для періоду $(n + s)$

$$e_{n+s} = Y_{n+s} - \hat{Y}_{n+s} = \varepsilon_{n+s} - \beta\varepsilon_{n+s-1}. \tag{1.125}$$

Дисперсія помилки

$$\text{Var}(e_{n+s}) = E[(e_{n+s})^2] = (1 + \beta^2)\sigma_\varepsilon^2 = \sigma_y^2. \tag{1.126}$$

Висновки

З (1.124) та (1.126) випливає, що прогнозне значення, отримане за МА(1)-процесами, починаючи з другого періоду дорівнює середньому процесу, а дисперсія помилки прогнозу дорівнює дисперсії Y . Отже, прогнозування на основі МА(1)-процесів можливе лише на один період вперед. Після нього прогноз просто дорівнює середньому значенню.

ARMA(1,1)-процес

Після того, як ми розглянули властивості прогнозів найпростіших процесів, розглянемо властивості змішаного процесу, зокрема ARMA(1,1). Запишемо його одразу як відхилення від середнього:

$$\text{ARMA}(1,1): Y_t - \mu = \alpha(Y_{t-1} - \mu) + \varepsilon_t - \beta\varepsilon_{t-1}, \quad (1.127)$$

де μ — середнє процесу; ε_t — білий шум.

Позначимо через Y_{n+s} невідоме дійсне значення Y в $(n+s)$ -й період часу; \hat{Y}_{n+s} — прогнозне значення Y_{n+s} , тоді $e_{n+s} = Y_{n+s} - \hat{Y}_{n+s}$ — помилка прогнозу. Дійсне значення за ARMA(1,1)-моделлю в період часу $(n+1)$

$$Y_{n+1} - \mu = \alpha(Y_{n+1} - \mu) + \varepsilon_{n+1} - \beta\varepsilon_n.$$

Відповідно прогнозне значення в період часу $(n+1)$

$$\hat{Y}_{n+1} - \mu = E[(Y_{n+1} - \mu) | Y_n, Y_{n-1}, \dots, Y_1] = \alpha(Y_n - \mu) - \beta\varepsilon_n.$$

Таким чином, з (1.115) легко побачити, що прогнозне значення за ARMA(1,1)-моделлю в період часу $(n+1)$ відрізняється від прогнозу за AR(1)-моделлю тільки на величину $(\beta\varepsilon_n)$.

Помилка прогнозу в період часу $(n+1)$, як і для всіх прогнозів за попередніми моделями, $e_{n+1} = Y_{n+1} - \hat{Y}_{n+1} = \varepsilon_{n+1}$, а дисперсія помилки відповідно дорівнює дисперсії випадкової величини (збурення): $\text{Var}(e_{n+1}) = \text{Var}(\varepsilon_{n+1}) = \sigma^2$.

Дійсне значення Y в $(n+2)$ -й період

$$Y_{n+2} - \mu = \alpha(Y_{n+1} - \mu) + \varepsilon_{n+2} - \beta\varepsilon_{n+1} = \alpha^2(Y_n - \mu) + \varepsilon_{n+2} - (\alpha - \beta)\varepsilon_{n+1} - \alpha\beta\varepsilon_n.$$

Прогнозне значення Y в період $(n+2)$

$$\hat{Y}_{n+2} - \mu = E[(Y_{n+1} - \mu) | Y_n, Y_{n-1}, \dots, Y_1] = \alpha^2(Y_n - \mu) - \alpha\beta\varepsilon_n, \quad (1.128)$$

оскільки $E(\varepsilon_{n+1}) = 0$, $E(\varepsilon_{n+2}) = 0$.

Помилка прогнозу в період $(n+2)$

$$e_{n+2} = Y_{n+2} - \hat{Y}_{n+2} = \varepsilon_{n+2} - (\alpha - \beta)\varepsilon_{n+1}.$$

Дисперсія помилки прогнозу в період $(n+2)$ відповідно

$$\text{Var}(e_{n+2}) = E[(e_{n+2})^2] = [1 + (\alpha - \beta)^2 \sigma_\varepsilon^2]. \quad (1.129)$$

Рекурсивно для періоду $(n + s)$ отримаємо:

$$\begin{aligned} Y_{n+s} - \mu &= \alpha(Y_{n+s-1} - \mu) + \varepsilon_{n+s} - \beta\varepsilon_{n+s-1} = \alpha(\alpha(Y_{n+s-2} - \mu) + \varepsilon_{n+s-1} - \beta\varepsilon_{n+s-2}) + \\ &+ \varepsilon_{n+s} - \beta\varepsilon_{n+s-1} = \dots = \\ &= \alpha^s(Y_n - \mu) + \varepsilon_{n+s} + (\alpha - \beta)\varepsilon_{n+s-1} + \dots + \alpha^{s-2}(\alpha - \beta)\varepsilon_{n+1} - \alpha^{s-1}\varepsilon_n. \end{aligned}$$

Відповідно прогноз для $(n + s)$ -го періоду дорівнює:

$$\hat{Y}_{n+s} - \mu = E[(Y_{n+s} - \mu) | Y_n, Y_{n-1}, \dots, Y_1] = \alpha^s(Y_n - \mu) - \alpha^{s-1}\beta\varepsilon_n. \quad (1.130)$$

Як легко побачити з (1.130), $\hat{Y}_{n+s} \rightarrow \mu$ при $s \rightarrow \infty$.

Помилка прогнозу для періоду $(n + s)$ за моделлю ARMA(1,1)

$$e_{n+s} = Y_{n+s} - \hat{Y}_{n+s} = \varepsilon_{n+s} + (\alpha - \beta)\varepsilon_{n+s-1} + \dots + \alpha^{s-2}(\alpha - \beta)\varepsilon_{n+1}.$$

Дисперсія помилки

$$\text{Var}(e_{n+s}) = E[(e_{n+s})^2] = \sigma_\varepsilon^2 [1 + (\alpha - \beta)^2 + \dots + \alpha^{2(s-2)}(\alpha - \beta)^2]. \quad (1.131)$$

Висновки

Формули (1.130) та (1.131) показують, що прогнозне значення, отримане за ARMA(1,1)-процесами, аналогічно прогнозам за AR(1)-процесами зі зростанням періоду випередження прямує до середнього, тобто $\hat{Y}_{n+s} \rightarrow \mu$ при $s \rightarrow \infty$, а дисперсія помилки прогнозу прямує до дисперсії (безумовної) процесу $\text{Var}(e_{n+s}) \rightarrow \sigma_y^2$ при $s \rightarrow \infty$.

Ми розглянули сутність прогнозування на основі ARIMA-моделей та прогнозні властивості найпростіших процесів. Зауважимо, що немає процедури, яка б давала однозначну специфікацію моделі ARIMA-процесу, тому на практиці постає питання про вибір найкращої з декількох моделей. Оскільки ARIMA-моделі атеоретичні, вибір специфікації здійснюється на основі статистичних характеристик моделей. Найпростішим критерієм вибору, на наш погляд, є середня абсолютна процентна помилка (MAPE) (mean absolute percentage error). MAPE є одним із найпоширеніших критеріїв, що використовується при порівнянні точності прогнозів різнорідних об'єктів, бо вимірюється в безрозмірних величинах, а саме у відсотках. Вважається, що значення MAPE менше за 10 % дає високу точність прогнозу, а значить, і якості моделі; від 10 до 20 — добру, від 20 до 50 — задовільну, понад 50 % — незадовільну. MAPE розраховується за формулою

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t^f - Y_t^a}{Y_t^a} \right| (100\%), \quad (1.132)$$

де Y_t^f — прогнозне значення показника в період часу t ; Y_t^a — фактичне значення показника в період часу t ; n — кількість спостережень.

Поряд з MAPE можна розглядати медіану абсолютних помилок у відсотках, яка має перевагу у випадку асиметричного розподілу помилок, коли середнє може бути зміщеним внаслідок небагатьох екстремальних значень, а також інші критерії.

1.2. ЗАСТОСУВАННЯ ARIMA(p, d, q)-МОДЕЛЕЙ НА ПРАКТИЦІ: ВИПАДОК ПОДАТКОВИХ НАДХОДЖЕНЬ В УКРАЇНІ

Для того, щоб проілюструвати практичне застосування ARIMA(p, d, q)-моделей, розглянемо прогнозування надходжень у Зведений бюджет України податку на додану вартість (VAT). Щомісячні дані про збір цього податку наведено в таблиці додатку 1.

Завдання 1.1.

Необхідно побудувати ARIMA(p, d, q)-модель та зробити на її основі прогноз на 6 наступних місяців.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. Зауважимо, що в процесі виконання цього завдання ми будемо використовувати пакет прикладних програм E.VIEWS.

Крок 1.

Перевірте, чи є часовий ряд (VAT) стаціонарним. Побудуйте та проаналізуйте корелограму часового ряду. Якщо коефіцієнти автокореляції (ACF) та часткової автокореляції (PACF) підтримують гіпотезу нестаціонарності, поясніть чому. У випадку нестаціонарності ряду знайдіть порядок інтеграції, використовуючи розширений тест Дікі—Фуллера (ADF).

Крок 2.

Застосуйте процедуру Хеннона—Рісанена для ідентифікації ARIMA-процесу. Визначте порядок авторегресійної складової AR-процесу. Коротко поясніть, чому ви впевнені, що знайшли правильний порядок авторегресійного процесу (p).

Крок 3.

Знайдіть порядок MA-процесу. Коротко поясніть, чому ви вважаєте, що знайшли правильний порядок ковзного середнього процесу (q).

Крок 4.

Зробіть загальний висновок щодо побудованої ARIMA(p, d, q)-моделі.

Крок 5.

Використовуючи побудовану ARIMA(p, d, q)-модель, зробіть прогноз на наступні 6 та 12 місяців і поясніть отримані результати.

Розв'язок завдання та розширений коментар моделювання доходу на додану вартість (VAT) за допомогою ARIMA-моделей

Для того, щоб прокоментувати прогнозування на основі ARIMA-моделей на прикладі податку на додану вартість (VAT), введемо такі позначення:

- ARIMA(p, d, q): p — порядок авторегресійної складової; q — порядок складової ковзного середнього (MA-складової); d — порядок інтеграції часового ряду;
- VAT — податок на додану вартість (млрд грн);
- CPI — індекс споживчих цін (дані наведені в таблиці додатку 1).

Зауважимо, що моделювання податку на додану вартість ми проводимо за допомогою економетричного пакета E.VIEWS.

Крок 1. Підготовка вхідних даних.

Припустимо, що необхідні для моделювання дані з таблиці додатку 1 знаходяться в робочому файлі E.VIEWS «ARIMA_data.wf1»⁷. Нагадаємо, що ми моделюємо часовий ряд VAT (податок на додану вартість), дані по якому наведені у номінальних одиницях. Для того, щоб уникнути інфляційного викривлення інформації, необхідно перейти до реальних одиниць виміру. Отже, всі спостереження ряду VAT поділити на відповідний індекс споживчих цін (CPI). В пакеті E.VIEWS для цього нам потрібно згенерувати новий часовий ряд $VATR=VAT/CPI$. У командному рядку наберіть

```
genr VATR=VAT/CPI
```

або використайте меню: **Procs/Generate series** та у вікні, що з'являється, надрукуйте $VATR=VAT/CPI$.

Використання для моделювання показників у реальних одиницях виміру, як правило, зменшує порядок інтеграції на один. Чому? Тому, що зазвичай

⁷ Якщо ні, їх можна легко імпортувати, наприклад з EXCEL, або ввести безпосередньо.

рівень цін є нестационарним рядом, а оскільки ціна входить складовою в показники у номінальному виразі, це збільшує порядок інтеграції часового ряду.

ARIMA-моделювання може застосовуватися лише до **стаціонарних часових рядів**, тому дуже важливо знайти порядок їх інтеграції, щоб коректно звести нестационарний ряд до стаціонарного.

Крок 2. Знаходження порядку інтеграції часового ряду.

A.

Застосуйте розширений критерій Дікі—Фуллера (ADF) до часового ряду:

- виберіть серію VATR;
- виберіть на панелі меню вікна, що відкрилось, опцію **View/Unit root test**;
- переконайтесь, що

1. Обрано розширений тест Дікі—Фуллера (ADF).
2. Задано тренд та перетин (ми рекомендуємо включати як тренд, так і перетин на першому етапі перевірки, щоб запобігти помилці специфікації. Також нагадаємо, що включення до моделі недоречних (тобто іррелевантних) змінних призводить до менш серйозних наслідків, ніж виключення важливих змінних).
3. Специфіковано кількість лагів.
4. Задано потрібний порядок різниць (перший, другий і т. ін.) або рівні (нульовий порядок різниць, тобто значення досліджуваного часового ряду).

В нашому випадку, при заданій на першому кроці кількості лагів $n = 1$ в рівнях (нульовий порядок різниць), отримано такі результати за ADF-тестом (див. таблицю 1.3.).

Як показують отримані результати, розрахована величина МакКіннона τ (ADF Test Statistic) в абсолютному виразі менша за критичну величину навіть за рівня статистичної значимості 10 %. Як правило, в практичних дослідженнях, якщо ADF-тест відхиляє гіпотезу стаціонарності ряду із заданим малим числом лагів, він відкидає її і при більшій кількості заданих лагів. Однак з метою підтвердження цього емпіричного висновку перевірте, чи для $n = 12$ ряд буде також нестационарним. Варто зважити на те, що не існує якогось загального правила відбору потрібної кількості лагів. Традиційно економічна теорія або самі дані містять інформацію про необхідну кількість лагів у базовому регресійному рівнянні ADF-тесту. В нашому випадку вхідними даними є щомісячні дані, тому можна очікувати, що включення 12 лагів⁸ є значимими для ADF-рівняння.

⁸ Сезонність надходжень податку, очевидно, призводить до автокореляції залишків у періодах t та $t - 12$. Отже, включення 12-го лага є виправданим *a priori*.

Таблиця 1.3. Результати ADF-тесту

ADF Test Statistic	-3.019579	1% Critical Value*	-4.0727	
		5% Critical Value	-3.4645	
		10% Critical Value	-3.1585	
*McKinnon critical values for rejection of hypothesis of a unit root.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(VATR)				
Method: Least Squares				
Sample(adjusted): 1993:03 1999:12				
Included observations: 82 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
VATR(-1)	-0.232382	0.076958	-3.019579	0.0034
D(VATR(-1))	0.012173	0.113318	0.107425	0.9147
C	0.001849	0.000802	2.305808	0.0238
@TREND(1993:01)	-1.71E-05	1.03E-05	-1.662316	0.1005
R-squared	0.119954	Mean dependent var	-9.18E-05	
Adjusted R-squared	0.086106	S.D. dependent var	0.001581	
S.E. of regression	0.001511	Akaike info criterion	-10.10456	
Sum squared resid	0.000178	Schwarz criterion	-9.987161	
Log likelihood	418.2871	F-statistic	3.543922	
Durbin-Watson stat	1.931517	Prob(F-statistic)	0.018314	

Тести перевірки часових рядів на стаціонарність вважаються недостатньо потужними, особливо у невеликих вибірках, ось чому дуже корисно додатково аналізувати корелограми, які є менш формальним апаратом перевірки на стаціонарність. Якщо коефіцієнти автокореляції (ACF) та часткової автокореляції (PACF) швидко згасають (тобто прямують до нуля), то це вказує на те, що часовий ряд є стаціонарним. Але у нашому випадку можна спостерігати поступове згасання коефіцієнтів автокореляції для ряду VATR, що свідчить про його нестаціонарність. Можна створити корелограму, вибираючи View/Correlogram і специфікуючи порядок різниць: рівні (нульовий порядок), перший, другий і т. д. (див. таблицю 1.4).

Якщо після перевірки ряду за допомогою корелограм та ADF-тесту можна зробити висновок про його нестаціонарність, переходимо до етапу Б.

Б.

Повторіть кроки, описані в пункті А, але замість використання значень самого ряду (в рівнях) використайте модифікований ряд, значення якого є першими різницями ряду VATR: $DVATR = D(VATR_t) = VATR_t - VATR_{t-1}$ ⁹. Знову

⁹ Для створення нового часового ряду DVATR потрібно набрати в командному рядку `genr DVATR=D(VATR)`.

Таблиця 1.4. Корелограма

Sample: 1993:01 1999:12

Included observations: 84

Autocorrelation	Partial Correlation	Лар	AC	PAC	Q-Stat	Prob
. *****	. *****	1	0.868	0.868	65.529	0.000
. *****	. .	2	0.750	-0.012	115.07	0.000
. *****	. .	3	0.639	-0.038	151.44	0.000
. ****	. .	4	0.532	-0.048	176.96	0.000
. ***	. .	5	0.442	0.001	194.80	0.000
. ***	. .	6	0.364	-0.006	207.09	0.000
. **	. .	7	0.300	0.004	215.55	0.000
. **	. .	8	0.234	-0.056	220.74	0.000
. *	. .	9	0.193	0.056	224.34	0.000
. *	. .	10	0.158	-0.005	226.79	0.000
. *	. **	11	0.185	0.227	230.19	0.000
. **	. **	12	0.264	0.251	237.17	0.000
. **	** .	13	0.253	-0.303	243.66	0.000
. **	. *	14	0.265	0.091	250.92	0.000
. **	. .	15	0.269	0.018	258.52	0.000
. **	. .	16	0.268	0.015	266.17	0.000

застосуйте розширений критерій Дікі—Фуллера (ADF) до часового ряду, тільки вже DVATR:

- виберіть серію DVATR;
- виберіть на панелі меню вікна, що відкрилось, **View/Unit root test**;
- переконайтесь, що
 1. Обрано розширений тест Дікі—Фуллера (ADF).
 2. Задано тренд та перетин.
 3. Специфіковано кількість лагів.
 4. Задано потрібний порядок різниць (перший в даному випадку).

При заданій кількості лагів $n = 1$ та перших різницях отримано такі результати (див. таблицю 1.5).

Як показують розрахунки, величина МакКіннона τ -статистики (ADF Test Statistic) в абсолютному виразі більша за критичну величину при 1 %-му рівні статистичної значимості. Таким чином, можна зробити висновок про стаціонарність ряду в перших різницях, але нам слід бути обережними та перевірити стаціонарність для більшої кількості лагів.

Для $n = 12$ відповідно отримаємо:

ADF Test Statistic	-4.478346	1%	Critical Value*	-4.0928
		5%	Critical Value	-3.4739
		10%	Critical Value	-3.1640

Таблиця 1.5. Результати ADF-тесту

ADF Test Statistic	-8.467939	1% Critical Value*	-4.0742
		5% Critical Value	-3.4652
		10% Critical Value	-3.1589

*McKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(VATR,2)
 Method: Least Squares
 Date: 01/28/01 Time: 14:36
 Sample(adjusted): 1993:04 1999:12
 Included observations: 82 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
VATR(-1)	-1.329333	0.156984	-8.467939	0.0000
D(VATR(-1))	0.177290	0.105344	1.682966	0.0964
C	-0.000634	0.000348	-1.823041	0.0722
@TREND(1993:01)	1.06E-05	7.08E-06	1.493185	0.1395
R-squared	0.608388	Mean dependent var	-3.62E-05	
Adjusted R-squared	0.593130	S.D. dependent var	0.002328	
S.E. of regression	0.001485	Akaike info criterion	-10.13861	
Sum squared resid	0.000170	Schwarz criterion	-10.02037	
Log likelihood	414.6138	F-statistic	39.87440	
Durbin-Watson stat	2.143652	Prob(F-statistic)	0.000000	

Як бачимо, розрахована величина МакКіннона τ -статистики (ADF Test Statistic) в абсолютному виразі більша за критичну величину навіть при 1 %-му рівні статистичної значимості. Таким чином, висновок про стаціонарність ряду в перших різницях підтверджено і при більшій кількості лагів.

На підтримку нашого висновку проаналізуємо корелограму ряду перших різниць D(VATR), яка також свідчить про практичну відсутність систематичності в автокореляції (див. таблицю 1.6).

Зауважимо, що, якщо ряд виявився б нестационарним у перших різницях, нам необхідно було б перейти до розгляду різниць другого порядку (в E.VIEWS вони позначаються: $D(VATR,2)=D(VATR)-D(VATR(-1))$). Якщо ж серія нестационарна і у других різницях, застосовуємо треті різниці і так далі. Eviews дозволяє легко та швидко обчислювати будь-які різниці за допомогою оператора D(Назва ряду, Кількість різниць). У практичній роботі використання вже других різниць достатньо для економічних рядів. Як правило, переважна більшість економічних показників є стаціонарними у других різницях. Таким чином, послідовне застосування ADF-тесту дозволяє знайти порядок інтеграції часового ряду, тобто на цьому кроці ми знаходимо параметр d у скороченій системі позначень ARIMA(p, d, q). В нашому випадку ряд VATR має порядок інтеграції один.

Таблиця 1.6. Корелограма

Sample: 1993:01 1999:12

Included observations: 83

Autocorrelation	Partial Correlation	Lag	AC	PAC	Q-Stat	Prob
. * .	. * .	1	-0.099	-0.099	0.8502	0.356
. * .	. * .	2	-0.153	-0.164	2.8861	0.236
. .	. * .	3	-0.049	-0.086	3.0952	0.377
. .	. * .	4	-0.043	-0.089	3.2622	0.515
. .	. .	5	0.057	0.019	3.5568	0.615
. .	. .	6	-0.026	-0.045	3.6209	0.728
. .	. .	7	-0.048	-0.054	3.8314	0.799
. * .	. * .	8	-0.093	-0.123	4.6394	0.795
. .	. .	9	0.008	-0.041	4.6457	0.864
** .	** .	10	-0.207	-0.283	8.7961	0.552
. * .	** * .	11	-0.185	-0.333	12.163	0.352
. ****	. ***	12	0.466	0.329	33.718	0.001
. * .	. * .	13	-0.106	-0.169	34.857	0.001
. .	. .	14	0.014	0.045	34.877	0.002
. .	. .	15	-0.022	-0.024	34.928	0.003
. * .	. * .	16	0.092	0.147	35.811	0.003

Крок 3. Знаходження порядку AR-складової (p).

Зауважимо, що не існує правила для знаходження ідеального порядку (p) авторегресійного процесу¹⁰. Та навіть за відсутності теоретично обґрунтованого універсального правила для знаходження ідеальної кількості лагів у моделі застосовуються певні процедури, які з достатньою точністю дозволяють знаходити порядок авторегресійного процесу. Ми пропонуємо одну з найпростіших та найпоширеніших — аналіз корелограм. Оскільки в нашому випадку перші різниці є стаціонарними, ми аналізуємо корелограму ряду DVATR (вона є другою з представлених вище корелограм). Можна помітити, що останній значимий елемент ACF та PACF (часткові коефіцієнти автокореляції) на цьому графіку має лаг 12, тому є сенс включити до 12 лагів у авторегресійну складову при моделюванні часового ряду VATR.

Це дає нам змогу оцінити попередню модель авторегресійної складової AR(12). Нагадаємо, що для попереднього оцінювання може використовуватись МНК:

$$DVATR = C(1) + C(2)*DVATR(-1) + C(3)*DVATR(-2) + C(4)*DVATR(-3) +$$

¹⁰ Нагадаємо, що порядок AR-складової дорівнює кількості лагів в AR-моделі.

$$\begin{aligned}
& + C(5)*DVATR(-4) + C(6)*DVATR(-5) + C(7)*DVATR(-6) + C(8)*DVATR(-7) + \\
& + C(9)*DVATR(-8) + C(10)*DVATR(-9) + C(11)*DVATR(-10) + \\
& + C(12)*DVATR(-11) + C(13)*DVATR(-12), \tag{1.1}
\end{aligned}$$

або в командах E.VIEWS:

LS DVATR C DVATR(-1 TO -12).

Зауважимо, що в E.VIEWS не обов'язково генерувати новий ряд, що складається з перших різниць, ми можемо оцінювати рівняння, використовуючи первісно заданий ряд VATR.

Для цього формулу (1.1) потрібно представити у вигляді:

$$\begin{aligned}
D(VATR) = & C(1) + C(2)*D(VATR(-1)) + C(3)*D(VATR(-2)) + C(4)*D(VATR(-3)) + \\
& + C(5)*D(VATR(-4)) + C(6)*D(VATR(-5)) + C(7)*D(VATR(-6)) + \\
& + C(8)*D(VATR(-7)) + C(9)*D(VATR(-8)) + C(10)*D(VATR(-9)) + \\
& + C(11)*D(VATR(-10)) + C(12)*D(VATR(-11)) + C(13)*D(VATR(-12)). \tag{1.2}
\end{aligned}$$

Хоча ми повинні обережно підходити до виключення незначимих факторів, їх зазвичай виключають з рівняння відповідно до статистичної значимості параметрів або різноманітних інформаційних статистик на зразок АІС-критерію. Пізніше ми розглянемо цю проблему, але трохи в іншому аспекті. Після оцінки рівняння (1.2) (див. таблицю 1.7.) нам необхідно підраховувати залишок (помилку), обираючи меню Procs/Make residual series (ми також маємо визначити назву ряду, де будуть зберігатися значення помилок (залишків), по умовчанням E.VIEWS дає їм назву RESID01). Після цього ми можемо переходити до кроку 4.

Крок 4. Знаходження порядку складової ковзного середнього $MA(q)$ в моделі $ARIMA(p, d, q)$.

Зауважимо, що на попередніх кроках ми знайшли порядок інтеграції (d) та порядок авторегресійної складової (p). Для остаточної ідентифікації ARIMA-моделі залишилось знайти порядок ковзної середньої (q). Ковзне середнє обчислюється на основі неспостережуваних випадкових величин. Для його оцінки ми використовуємо помилки (залишки), які отримані на кроці 3. Далі ми додаємо підраховані залишки до рівняння третього кроку (1.2). Знову постає питання, скільки лагів для залишків обрати, тобто якому порядку MA -

Таблиця 1.7. Результати оцінки моделі AR(12)

Dependent Variable: DVATR

Method: Least Squares

Sample(adjusted): 1994:02 1999:12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000188	0.000149	-1.268665	0.2096
DVATR(-1)	-0.391511	0.114125	-3.430542	0.0011
DVATR(-2)	-0.217590	0.115855	-1.878127	0.0654
DVATR(-3)	-0.083497	0.105447	-0.791837	0.4317
DVATR(-4)	-0.074356	0.100048	-0.743201	0.4604
DVATR(-5)	-0.098207	0.097426	-1.008014	0.3176
DVATR(-6)	-0.191804	0.097768	-1.961820	0.0546
DVATR(-7)	-0.213345	0.097073	-2.197774	0.0320
DVATR(-8)	-0.133016	0.092882	-1.432102	0.1575
DVATR(-9)	-0.037546	0.091788	-0.409050	0.6840
DVATR(-10)	-0.174493	0.090490	-1.928311	0.0587
DVATR(-11)	-0.264599	0.091735	-2.884380	0.0055
DVATR(-12)	0.357113	0.102582	3.481245	0.0010
R-squared	0.524402	Mean dependent var	1.41E-06	
Adjusted R-squared	0.426002	S.D. dependent var	0.001443	
S.E. of regression	0.001094	Akaike info criterion	-10.63471	
Sum squared resid	6.94E-05	Schwarz criterion	-10.22041	
Log likelihood	390.5321	F-statistic	5.329301	
Durbin-Watson stat	1.959028	Prob(F-statistic)	0.000006	

складової віддати перевагу. Як і раніше, можна відмітити, що не існує процедури для встановлення ідеального порядку MA, тому повертаємось до аналізу другої корелограми. Її аналіз показує, що необхідно включити 12 лагів для ідентифікації MA-складової (ACF має статистично значимі коефіцієнти для лага 12). Таким чином, методом МНК ми попередньо оцінюємо наступне ARIMA(12, 1, 12)-рівняння:

$$\begin{aligned}
 DVATR = & C(1) + C(2)*DVATR(-1) + C(3)*DVATR(-2) + C(4)*DVATR(-3) + \\
 & + C(5)*DVATR(-4) + C(6)*DVATR(-5) + C(7)*DVATR(-6) + C(8)*DVATR(-7) + \\
 & + C(9)*DVATR(-8) + C(10)*DVATR(-9) + C(11)*DVATR(-10) + \\
 & + C(12)*DVATR(-11) + C(13)*DVATR(-12) + C(14)*RESID01(-1) + \\
 & + C(15)*RESID01(-2) + C(16)*RESID01(-3) + C(17)*RESID01(-4) + \\
 & + C(18)*RESID01(-5) + C(19)*RESID01(-6) + C(20)*RESID01(-7) + \\
 & + C(21)*RESID01(-8) + C(22)*RESID01(-9) + C(23)*RESID01(-10) + \\
 & + C(24)*RESID01(-11) + C(25)*RESID01(-12),
 \end{aligned}
 \tag{1.3}$$

або в командах E.VIEWS:

LS DVATR C DVATR(-1 TO -12) RESID01(-1 TO -12).

Таблиця 1.8. Результати оцінки моделі AR(12, 1, 12)

Dependent Variable: DVATR
 Method: Least Squares
 Sample(adjusted): 1995:02 1999:12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.005775	0.011692	0.493940	0.6245
DVATR(-1)	9.857377	20.73947	0.475296	0.6376
DVATR(-2)	11.04669	21.81177	0.506455	0.6158
DVATR(-3)	9.364731	18.96973	0.493667	0.6247
DVATR(-4)	7.395617	14.37778	0.514378	0.6103
DVATR(-5)	5.265482	11.16224	0.471723	0.6401
DVATR(-6)	4.527325	9.482976	0.477416	0.6361
DVATR(-7)	4.929830	10.22957	0.481920	0.6330
DVATR(-8)	5.573729	11.20858	0.497274	0.6222
DVATR(-9)	5.213252	10.21027	0.510589	0.6129
DVATR(-10)	3.643312	7.645796	0.476512	0.6368
DVATR(-11)	4.196044	8.742020	0.479986	0.6343
DVATR(-12)	5.937596	11.36148	0.522608	0.6046
RESID01(-1)	-10.40299	20.76468	-0.500995	0.6196
RESID01(-2)	-7.758026	13.73593	-0.564798	0.5759
RESID01(-3)	-4.386170	9.163657	-0.478648	0.6353
RESID01(-4)	-3.644705	6.166026	-0.591095	0.5584
RESID01(-5)	-1.584439	4.197105	-0.377508	0.7081
RESID01(-6)	-1.311141	2.822066	-0.464603	0.6452
RESID01(-7)	-0.784660	1.836754	-0.427199	0.6719
RESID01(-8)	-0.452408	1.262256	-0.358412	0.7223
RESID01(-9)	-0.503305	0.859159	-0.585811	0.5619
RESID01(-10)	-0.091420	0.687955	-0.132886	0.8951
RESID01(-11)	-0.385569	0.461710	-0.835090	0.4095
RESID01(-12)	0.257299	0.422527	0.608953	0.5466
R-squared	0.724169	Mean dependent var	-1.59E-05	
Adjusted R-squared	0.529465	S.D. dependent var	0.001487	
S.E. of regression	0.001020	Akaike info criterion	-10.64213	
Sum squared resid	3.54E-05	Schwarz criterion	-9.761813	
Log likelihood	338.9427	F-statistic	3.719327	
Durbin-Watson stat	1.894913	Prob(F-statistic)	0.000246	

Щоб перевірити, чи можна вважати попередньо оцінену ARIMA(12, 1, 12)-модель першим наближенням до придатної для прогнозу ARIMA-моделі, необхідно перевірити, чи є залишки такої моделі білим шумом. Перевірку на білий шум (незалежність залишків) ми робимо на 5-му кроці. Відмітимо, що отримана попередня ARIMA(12, 1, 12)-модель є досить громіздкою з незначними факторами, тому пізніше, на 6-му кроці ми спробуємо її покращити.

Крок 5. Аналіз ARIMA-моделей.

Найголовніша вимога ARIMA-моделей є надзвичайно простою: залишки мають бути незалежними один від одного. Зокрема, має бути відсутня серійна кореляція залишків. Оскільки в моделі наявні лагові залежні змінні, ми можемо використовувати *h*-тест Дарбіна—Уотсона (Durbin—Watson) для перевірки залишків на серійну автокореляцію, але цей тест вимагає додаткових розрахунків. Альтернативним тестом у цьому випадку є LM-тест Бройша—Годфрея на серійну кореляцію (Breusch—Godfrey (BG) Serial Correlation LM). Щоб застосувати цей тест, виберіть в меню **View/Residual tests/Serial Correlation LM test** (треба також визначити порядок серійної кореляції — останній лаг випадкової величини (залишку), який може корелювати з поточним значенням залишку). Для нашої моделі:

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.838946	Probability	0.613136
Obs*R-squared	18.52270	Probability	0.100717

Ми приймаємо гіпотезу про відсутність серійної кореляції, якщо значення *p*-value статистики (Obs*R-squared) > 10 %. В даному випадку *p*-value = 10.0717 > 10 %. Зауважимо, що Obs означає кількість спостережень *n*, а R-squared — коефіцієнт детермінації базової моделі тесту, *p*-value (probability) — імовірність того, що ми приймаємо хибну нульову гіпотезу (помилка першого типу).

Перевірити, чи є залишки нормально розподіленими, можна за допомогою тесту Жарку—Бера (Jarque—Bera) (**View/Residual tests/Histogram-Normality**), (див. наведену нижче схему).

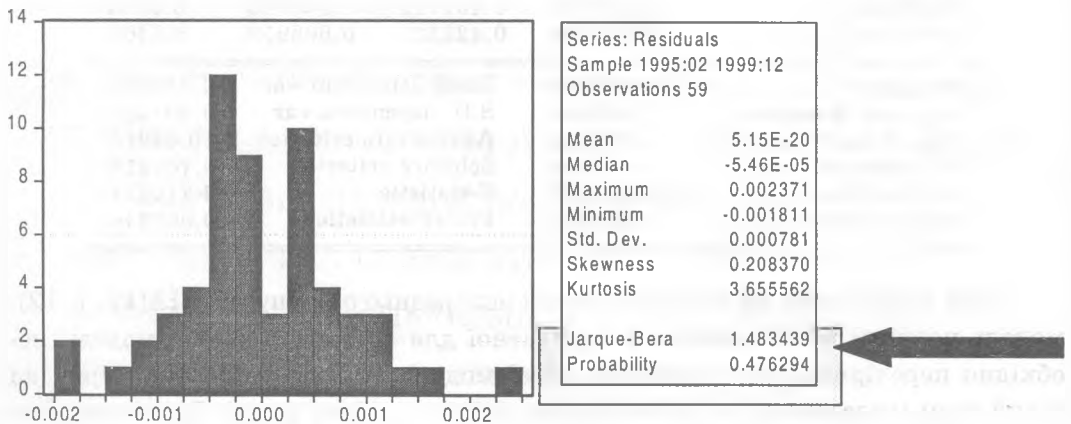


Рис. 8. Результати тесту Жарку—Бера

Цей тест випробовує нульову гіпотезу (H_0): залишки розподілені нормально проти альтернативної, коли вони не розподілені нормально. Ми приймаємо нульову гіпотезу, якщо p -value Жарку—Бера-статистики $> 10\%$. У нашому конкретному випадку $p = 48\%$, що свідчить про нормально розподілені залишки.

Крок 6. Вибір моделі.

Оцінена ARIMA(12, 1, 12)-модель є досить громіздкою, тому спробуємо її покращити.

Відмітимо, що коефіцієнт детермінації залежить тільки від AR-умов, тому цей критерій адекватності моделі є не зовсім придатним для відбору найкращої з декількох моделей. Для того, щоб вибрати придатну й негроміздку модель, пропонуються дві статистики: інформаційний критерій Акайка (AIC) та інформаційний критерій Шварца (SIC). Ці критерії можуть бути застосовані до обох моделей, які оцінюються та аналізуються на третьому та четвертому кроках. Якщо включення додаткового лага не змінює AIC та SIC, ми можемо прийняти рішення не додавати до моделі цей лаг. Крім того, ми можемо виключити незначимі лаги з рівняння. У випадку з VATR зазначимо, що лише лаги 1, 2, 6, 7, 10, 11, 12 є значимими на третьому кроці перевірки адекватності моделі. За цим же критерієм для MA-складової залишимо тільки лаги 2 та 4. Аналогічного висновку можна дійти, використовуючи інформаційні критерії. Обрана нами модель має вигляд:

Таблиця 1.9. Результати переоцінки ARIMA-моделі

Method: Least Squares
 Sample(adjusted): 1994:06 1999:12
 Included observations: 67 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.13E-05	0.000128	-0.323162	0.7478
DVATR(-1)	-0.348521	0.096790	-3.600800	0.0007
DVATR(-2)	0.111197	0.145780	0.762778	0.4487
DVATR(-6)	-0.143786	0.088057	-1.632866	0.1080
DVATR(-7)	-0.183983	0.086963	-2.115646	0.0388
DVATR(-10)	-0.176345	0.088548	-1.991512	0.0512
DVATR(-11)	-0.214760	0.089938	-2.387849	0.0203
DVATR(-12)	0.474850	0.094517	5.023988	0.0000
RESID01(-2)	-0.580792	0.210175	-2.763376	0.0077
RESID01(-4)	-0.296118	0.126356	-2.343533	0.0226
R-squared	0.585800	Mean dependent var	-3.58E-05	
Adjusted R-squared	0.520400	S.D. dependent var	0.001465	
S.E. of regression	0.001015	Akaike info criterion	-10.81167	
Sum squared resid	5.87E-05	Schwarz criterion	-10.48261	
Log likelihood	372.1909	F-statistic	8.957176	
Durbin-Watson stat	2.042114	Prob(F-statistic)	0.000000	
Breusch-Godfrey Serial Correlation LM Test:				

Закінчення таблиці 1.9.

F-statistic	0.982628	Probability	0.479437
Obs*R-squared	13.91110	Probability	0.306421

Як показують результати розрахунків, критерії BG та JB свідчать про адекватну ідентифікацію ARIMA-моделі. Проте це не єдино можливий варіант ARIMA-моделі для VAT. Достатня доля суб'єктивізму при виборі остаточної ідентифікації сприяє величезній різноманітності моделей. Крім того, відмітимо, що процедура ідентифікації є багатокроковою, тому ми втрачаємо точність упродовж її виконання. Однак коли модель вже остаточно обрана, необхідно її переоцінити коректнішими методами. В пакеті E.VIEWS для оцінки остаточної специфікації моделі використовується вбудована процедура оцінки ARIMA-моделі (нелінійний метод найменших квадратів). Для її застосування необхідно записати в командах E.VIEWS:

LS DVATR C AR(1) AR(2) AR(6) AR(7) AR(10) AR(11) AR(12) MA(2) MA(4)

Крок 7. Прогнозування за ARIMA-моделями.

Прогнозування — досить легка процедура, бо нам потрібно знати лише початкове значення часового ряду та значення залежних змінних з відповідними лагами. Це вельми спрощує завдання — нам не треба робити припущення щодо майбутніх значень незалежних факторів, як в інших методах, наприклад у лінійній регресії, хоча ми традиційно припускаємо, що минулі тенденції зберігатимуться у майбутньому. Розрахунок прогнозів за ARIMA-моделями та помилок прогнозу вручну — дуже виснажлива процедура, але прикладні пакети, зокрема E.VIEWS, забезпечують швидку оцінку точкових та інтервальних прогнозів. Щоб зробити прогноз, виберіть з головного меню E.VIEWS Procs/Forecast або Forecast у меню діалогового вікна оцінки регресії. Потім у відкритому вікні виберіть:

- початкове значення;
- період випередження прогнозу;
- результат: стандартні помилки (On/Off), графік (On/Off).

Якщо ви створюєте регресію у різницях (як D(VATR)), можете обрати кінцеві прогнозні значення як у різницях, так і у рівнях. Якщо ж ви створюєте регресію у DVATR (на основі безпосереднього ряду в різницях), отримати прогнозні значення можна тільки у різницях. *Зауважте це для себе!*

Ще одна важлива опція — **Static/Dynamic forecasting**. Якщо обрано опцію **Static** (статичне прогнозування), то прогноз розраховується на основі фактичних значень лагових змінних, тоді як при динамічному прогнозуванні (**Dynamic forecasting**) використовуються розраховані (fitted) значення лаго-

вих змінних. В обох підходах помилки також обраховуються по-різному, але процедура їх розрахунку досить складна, тому ми її не наводимо (для отримання додаткової інформації дивіться E.VIEWS Help). Статичне прогнозування ефективне, якщо ми хочемо отримати прогноз на один крок уперед. Різниця між цими двома підходами незначна для короткострокового прогнозу.

Крім того, зауважимо, що не слід вибирати для ARIMA-моделей опцію **Structural (ignore ARMA)** тому, що тоді E.VIEWS вважатиме всі помилки рівними нулю, що не має сенсу в умовах ARIMA-моделювання.

Завдання 1.2.

Використовуючи дані таблиці додатку 1, побудуйте ARIMA-моделі для інших податкових надходжень.

Питання для самоперевірки

1. Поясніть, чому стаціонарність часових рядів є необхідною умовою розробки моделей авторегресії та ковзного середнього.
2. В чому полягає відмінність між інтегрованими моделями авторегресії, ковзного середнього (ARIMA) та моделями авторегресії та ковзного середнього (ARMA)?
3. Наведіть визначення нестационарного процесу та поясніть, що означає порядок інтеграції.
4. Запишіть загальний вигляд моделі авторегресії p -го порядку. Детально прокоментуйте модель.
5. Дайте визначення лагового оператора та лагового полінома. Запишіть поліном лага для AR(1)-процесу.
6. Опишіть загальну схему побудови ARMA-моделей.
7. Як перевірити часові ряди на стаціонарність? Чи завжди нестационарний ряд можна звести до стаціонарного?
8. Якщо часовий ряд є нестационарним з порядком інтеграції 3 [1(3)], то скільки разів вам необхідно його перетворювати операцією різниць, щоб він став стаціонарним?
9. Чим відрізняється тест перевірки часового ряду на стаціонарність Дікі—Фуллера (DF) від розширеного тесту Дікі—Фуллера (ADF)?
10. Запишіть лаг-поліном для AR(2)-процесу.
11. Що таке одиничний корінь? Поясніть детально, в чому полягає його важливість для ARIMA-моделювання.
12. Що таке автокореляційна (ACF) та часткова автокореляційна (PACF) функції?

13. Як за допомогою корелограми та часткової корелограми можна визначити порядок процесу авторегресії?
14. Запишіть та проінтерпретуйте модель ковзного середнього першого порядку $MA(1)$. Прокоментуйте основні властивості $MA(1)$ -процесу.
15. Як можна визначити порядок MA -процесу?
16. Прокоментуйте основні властивості $ARMA(1,1)$ -процесу.
17. Детально опишіть алгоритм ідентифікації та оцінки $ARIMA(p, d, q)$ -процесу.
18. Детально опишіть процедуру Хеннона та Ріссанена для визначення порядку p та q $ARIMA(p, d, q)$ -процесів.
19. Яким чином можна визначити порядок інтеграції (d) часового ряду?
20. В чому полягає основна ідея прогнозу за $ARIMA$ -моделями?
21. Чим відрізняються прогнози за $AR(1)$ - та $MA(1)$ -моделями?
22. Назвіть основні властивості прогнозів за $AR(1)$ -моделями.

Вправи

Вправа 1.1.

Використовуючи умовні дані наведеної нижче таблиці 1.10, побудуйте вибірккові корелограми до 25 лага включно для часових рядів Y_2 , Y_3 , Y_4 та Y_5 . Яку загальну картину ви спостерігаєте? Оцініть візуально, який з цих часових рядів може бути стаціонарним.

Зауваження. Для виконання вправ 1.1—1.4 використайте пакет E.VIEWS або інший доступний вам статистичний пакет.

Таблиця 1.10. Вхідні дані (макропоказники за 1980.01—2001.04)

Період	Y_1	Y_2	Y_3	Y_4	Y_5
1980.01	2872.8	1990.6	1800.5	44.7	24.5
1980.02	2860.3	2020.1	1807.5	44.4	23.9
1980.03	2896.6	2045.3	1824.7	44.9	23.3
1980.04	2873.7	2045.2	1821.2	42.1	23.1
1981.01	2942.9	2073.9	1849.9	48.8	23.8
1981.02	2947.4	2098.0	1863.5	50.7	23.7
1981.03	2966.0	2106.6	1876.9	54.2	23.8
1981.04	2980.8	2121.1	1904.6	55.7	23.7
1982.01	3037.3	2129.7	1929.3	59.4	25.0
1982.02	3089.7	2149.1	1963.3	60.1	25.5
1982.03	3125.8	2193.9	1989.1	62.8	26.1
1982.04	3175.5	2272.0	2032.1	68.3	26.5
1983.01	3253.3	2300.7	2063.9	79.1	27.0

Продовження таблиці 1.10.

Період	Y_1	Y_2	Y_3	Y_4	Y_5
1983.02	3267.6	2315.2	2062.0	81.2	27.8
1983.03	3264.3	2337.9	2073.7	81.3	28.3
1983.04	3289.1	2382.7	2067.4	85.0	29.4
1984.01	3259.4	2334.7	2050.8	89.0	29.8
1984.02	3267.6	2304.5	2059.0	91.2	30.4
1984.03	3239.1	2315.0	2065.5	97.1	30.9
1984.04	3226.4	2313.7	2039.9	86.8	30.5
1985.01	3154.0	2282.5	2051.8	75.8	30.0
1985.02	3190.4	2390.3	2086.9	81.0	29.7
1985.03	3249.9	2354.4	2114.4	97.8	30.1
1985.04	3292.5	2389.4	2137.0	103.4	30.6
1986.01	3356.7	2424.5	2179.3	108.4	32.6
1986.02	3369.2	2434.9	2194.7	109.2	35.0
1986.03	3381.0	2444.7	2213.0	110.0	36.6
1986.04	3416.3	2459.5	2242.0	110.3	38.3
1987.01	3466.4	2463.0	2271.3	121.5	39.2
1987.02	3525.0	2490.3	2280.8	129.7	40.0
1987.03	3574.4	2541.0	2302.6	135.1	41.4
1987.04	3567.2	2556.2	2331.6	134.8	42.4
1988.01	3591.8	2587.3	2347.1	137.5	43.5
1988.02	3707.0	2631.9	2394.0	154.0	44.5
1988.03	3735.6	2653.2	2404.5	158.0	46.6
1988.04	3779.6	2680.9	2421.6	167.8	48.9
1989.01	3780.8	2699.2	2437.9	168.2	50.5
1989.02	3784.3	2697.6	2435.4	174.1	51.8
1989.03	3807.5	2715.3	2454.7	178.1	52.7
1989.04	3814.6	2728.1	2465.4	173.4	54.5
1990.01	3830.8	2742.9	2464.6	174.3	57.6
1990.02	3732.6	2692.0	2414.2	144.5	58.7
1990.03	3733.5	2722.5	2440.3	151.0	59.3
1990.04	3808.5	2777.0	2469.2	154.6	60.5
1991.01	3860.5	2783.7	2475.5	159.5	64.4
1991.02	3844.4	2776.7	2476.1	143.7	68.4
1991.03	3864.5	2814.1	2487.4	147.6	71.9
1991.04	3803.1	2808.8	2468.6	140.3	72.4
1992.01	3756.1	2795.0	2484.0	114.4	70.0
1992.02	3771.1	2824.8	2488.9	114.0	68.4
1992.03	3754.4	2829.0	2502.5	114.6	69.2
1992.04	3759.6	2832.6	2539.3	109.9	72.5
1993.01	3783.5	2843.6	5556.5	113.6	77.0
1993.02	3886.5	2867.0	2604.0	133.0	70.5
1993.03	3944.4	2903.0	2639.0	145.7	83.1
1993.04	4012.1	2960.6	2678.2	141.6	84.2
1994.01	4089.5	3033.2	2703.8	155.1	83.3
1994.02	4144.0	3065.9	2741.1	152.6	82.2
1994.03	4166.4	3102.7	2754.6	141.8	81.7
1994.04	4194.2	3118.5	2784.8	136.3	83.4

Закінчення таблиці 1.10.

Період	Y_1	Y_2	Y_3	Y_4	Y_5
1995.01	4221.8	3123.6	2824.9	125.2	87.2
1995.02	4254.8	3189.6	2849.7	124.8	90.8
1995.03	4309.0	3156.5	2893.3	129.8	94.1
1995.04	4333.5	3178.7	2895.3	134.2	97.4
1996.01	4390.5	3227.5	2922.4	109.2	105.1
1996.02	4387.7	3281.4	2947.9	106.0	110.7
1996.03	4412.6	3272.6	2993.7	111.0	112.3
1996.04	4427.1	3266.2	3012.5	119.2	111.0
1997.01	4460.0	3295.2	3011.5	140.2	108.0
1997.02	4515.3	3241.7	3046.8	157.9	105.5
1997.03	4559.3	3285.7	3075.8	169.1	105.1
1997.04	4625.5	3335.8	3074.6	176.0	106.3
1998.01	4655.3	3380.1	3128.2	195.5	109.6
1998.02	4704.8	3386.3	3147.8	207.2	113.3
1998.03	4734.5	3407.5	3170.6	213.4	117.5
1998.04	4779.7	3443.1	3202.9	226.0	121.0
1999.01	4809.8	3473.9	3200.9	221.3	124.6
1999.02	4832.4	3450.9	3208.6	206.2	127.1
1999.03	4845.6	3466.9	3241.1	195.7	129.1
1999.04	4859.7	3493.0	3241.6	203.0	130.7
2000.01	4880.8	3531.4	3258.8	199.1	132.3
2000.02	4900.3	3545.3	3258.6	193.7	132.5
2000.03	4903.3	3547.0	3281.2	196.3	133.8
2000.04	4855.1	3529.5	3251.8	199.0	136.2
2001.01	4824.0	3514.8	3241.1	189.7	137.8
2001.02	4840.7	3537.4	3252.4	182.7	136.7
2001.03	4862.7	3539.9	3271.2	189.6	138.1
2001.04	4868.0	3547.5	3271.1	190.3	138.5

Примітка. Дані наведені в цінах 1997 р., млрд у. о.: Y_1 — валовий внутрішній продукт; Y_2 — прибуток, що залишається в розпорядженні домогосподарств; Y_3 — витрати домогосподарств на споживання; Y_4 — прибуток корпорацій після сплати податків; Y_5 — чисті дивіденди корпорацій.

Вправа 1.2.

Кожний з часових рядів вправи 1.1. перевірте на наявність одиничного кореня за допомогою простого тесту Дікі—Фуллера (DF). Якщо він є присутнім, то як ви охарактеризуєте такий часовий ряд?

Вправа 1.3.

Продовження вправи 1.2. Як ви можете довести, що розширений тест Дікі—Фуллера (ADF) є кращим за простий тест Дікі—Фуллера (DF)? Обґрунтуйте детально ваші висновки.

Вправа 1.4.

Утворіть перші різниці з часових рядів, наведених в таблиці 1.10. Відобразіть їх графічно. Крім того, побудуйте вибіркові корелограми кожного з рядів до 25 лага включно. Які висновки ви можете зробити стосовно часових рядів у перших різницях?

Вправа 1.5.

В роботі [24, с.127] наведена така модель, розрахована на період 1948—1984 рр.:

$$\Delta X_t = 31.03 - 0.188X_{t-1}$$

$$se = (12.50) \quad (0.088)$$

$$(t = \tau) \quad (-2.35),$$

де X_t — ресурси приватного сектора Великобританії в період t ; se — вибіркове середньоквадратичне відхилення; τ — статистика.

Зауваження. 5%-не критичне значення τ -статистики дорівнює -2.95 , а 10 %-не — -2.60 .

(а) Чи можна на основі результатів даної регресії зробити висновок про стаціонарність або нестаціонарність часового ряду X ? Чи є одиничний корінь у цього часового ряду? Яким чином ви можете про це дізнатися?

(б) Якби вам потрібно було використати звичайний τ -тест для перевірки на значимість нахилу, який висновок ви зробили б? Чи зробили б ви на основі вашої попередньої перевірки висновок про стаціонарність часового ряду?

(в) Розгляньте такі результати регресійної моделі:

$$\Delta^2 X_t = 4.76 - 1.39X_{t-1} + 0.313\Delta^2 X_{t-1}$$

$$se = (5.06) \quad (0.236) \quad (0.163)$$

$$(t = \tau) \quad (-5.89),$$

де Δ^2 є оператором других різниць. Отримане тепер значення τ -статистики свідчить про статистичну значимість коефіцієнта при змінній з лагом один. Що ви можете тепер сказати про стаціонарність часового ряду?

Зауваження. Метою попередньої регресії є визначення наявності в часовому ряді другого одиничного кореня.

Список літератури

1. *Anderson T. W.* The Statistical Analysis of Time Series.— New York : Wiley, 1983.
2. *Beguin J.-M., Gourieroux C., Montfort A.* Identification of a Mixed Autoregressive-Moving Average Process: The Corner Method. Time Series.— Amsterdam : North Holland, 1980.— P. 423—436.
3. *Box G. E. P., Jenkins G. M.* Time Series Analysis: Forecasting and Control / Revised Edition.— San Francisco : Holden Day, 1976.
4. *Campbell John Y., Perron Pierre.* Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Working Papers Series, 1991.
5. *Charemza W. W., Deadman D. F.* New Direction in Econometric Practice.— Brookfield, VT : Edward Eglar, 1992.
6. *Davidson J. E. H.* Problems with the Estimation of Moving Average Processes // Journal of Econometrics, 1981, № 16.— P. 295—310.
7. *Dickey D. A., Fuller W. A.* Distribution of the Estimators for the Autoregressive Time Series with a Unit Root // Journal of the American Statistical Association, 1979, № 74.— P. 427—431.
8. *Dickey D. A., Bell W. R., Miller R. B.* Unit Roots in Time Series Models: Tests and Implications // American Statistician, 1986, № 40.— P. 12—26.
9. *Dickey D. A., Fuller W. A.* Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root // Econometrica, 1981, № 49.— P. 1057—1072.
10. *Frances Philip Hans.* Time Series Models for Business and Economic Forecasting.— Cambridge : Cambridge University Press, 1999.— 280 p.
11. *Fuller W. A.* Introduction to Statistical Time Series.— New York : Wiley, 1976.
12. *Granger C. W.* Some Recent Developments in a Concept of Causality // Journal of Econometrics, 1988, № 39.— P. 199—211.
13. *Greene William H.* Econometric Analysis.— 4th ed.— New Jersey : Prentice Hall, 2000.— 1004 p.
14. *Gujarati Damodar N.* Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, 1995.— 838 p.
15. *Hamilton James D.* Time Series Analysis.— Princeton, New Jersey : Princeton University Press, 1994.
16. *Hannan E. J.* Multiple Time Series.— New York : Wiley, 1970.
17. *Hannan E. J.* The Estimation of the Order of an ARMA Process // Annals of Statistics, 1980, № 8.— P. 1071—1081.
18. *Hatanaka M.* Time-Series-Based Econometrics. Unit Roots and Cointegration.— Oxford : Oxford University Press, 1996.
19. *Harvey A. C.* Time Series Models.— Oxford : Philip Allan, 1981.
20. *Johnston J., Dinardo J.* Econometric Methods.— New York : McGraw-Hill, 1997.

21. *Kennedy P.* A Guide to Econometrics.— Cambridge, Massachusetts : The MIT Press, 1998.— 468 p.
22. *Lucas R. E.* Econometric Policy Evaluation: A Critique // *K. Brunner, A. H. Meltzer.* The Philips Curve and Labor Markets. Carnegie-Rochester Conference. Series on Public Policy, 1976, № 1.— P. 19—46.
23. *Makridakis S., Wheelwright S. C., McGee V. E.* Forecasting: Method and Application.— 2nd ed.— New York : Wiley, 1983.
24. *Mills Terence C.* Time Series Techniques for Economists.— Cambridge : Cambridge University Press, 1998.
25. *Mitchell W. F.* Testing for Unit Roots and Persistence in OECD Unemployment Rates // *Applied Econometrics*, 1993, № 25.— P. 1489—1501.
26. *Pindyck R. S., Rubinfeld D. L.* Econometric Models and Economic Forecasts.— 3rd ed.— New York : McGraw-Hill, 1991.
27. *Буряковський В. В. та ін.* Податки. Навчальний посібник.— Дніпропетровськ : Пороги, 1998.— 646 с.
28. Закон України «Про Державний бюджет України на 2000 рік» // Відомості Верховної Ради, 2000, № 8.
29. Закон України «Про бюджетну систему України» // ВВР, 1991, № 1.
30. Закон України «Про оподаткування прибутку підприємств» // ВВР, 1995, № 4.
31. Закон України «Про податок на додану вартість» від 3 квітня 1997 р. // ВВР, 1997, № 21.
32. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика. Практикум з використанням комп'ютера.— К. : Знання, 1998.— 217 с.
33. *Лук'яненко І. Г., Городніченко Ю. О.* Прогнозування податкових надходжень за допомогою моделей корегування помилки // *Фінанси України*, 2001, № 7.— С. 89—99.
34. *Холден К., Піл Д. А., Томпсон Дж. Л.* Економічне прогнозування. Вступ.— К. : Інформтехніка-ЕМЦ, 1996.— 215 с.
35. *Черняк О. І., Ставицький А. В.* Динамічна економетрика.— К. : КВІЦ, 2000.— 120 с.

РОЗДІЛ 2. ЗАСТОСУВАННЯ VAR-МОДЕЛЕЙ В ЕМПІРИЧНИХ ДОСЛІДЖЕННЯХ

2.1. ТЕОРЕТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ ЗА ДОПОМОГОЮ VAR-МОДЕЛЕЙ

2.1.1. Поняття моделювання за допомогою системи авторегресійних моделей

Моделювання економічних процесів за допомогою ARIMA-моделей дозволяє виявити динамічний зв'язок між поточними та лаговими значеннями досліджуваного показника. ARIMA-моделі є зручним інструментом коротко- та середньострокового прогнозування окремих часових рядів. Однак сучасні дослідження концентрують все більшу увагу на розробці апарату одночасного моделювання декількох часових рядів за допомогою системи динамічних рівнянь ARIMA-процесів, що дозволяє включати та досліджувати взаємозворотні зв'язки між показниками та їх лаговими значеннями. Таким чином, VAR-моделі є розширенням концепції ARIMA-моделювання окремого часового ряду. Термін «вектор» в цьому випадку показує, що моделюються одночасно два або більше часових рядів. Термін «авторегресійні» означає включення лагових значень залежних змінних у праву частину кожного окремого рівняння системи.

Проілюструємо сутність VAR-моделювання простим прикладом. Припустимо, що маємо тільки два часових ряди: $\{Y_{1t}\}$ та $\{Y_{2t}\}$. Найпростішу VAR-модель в даному випадку можна записати у вигляді:

$$Y_{1t} = \gamma_{10} - \gamma_{12}Y_{2t} + \beta_{11}Y_{1,t-1} + \beta_{12}Y_{2,t-1} + u_{1t}, \quad (2.1)$$

$$Y_{2t} = \gamma_{20} - \gamma_{21}Y_{1t} + \beta_{21}Y_{1,t-1} + \beta_{22}Y_{2,t-1} + u_{2t}, \quad (2.2)$$

де γ_{10} , γ_{20} , γ_{12} , γ_{21} , β_{11} , β_{12} , β_{21} , β_{22} — невідомі коефіцієнти, які пов'язують поточні та минулі значення показників Y_{1t} та Y_{2t} . При цьому припускається, що Y_{1t} та Y_{2t} є стаціонарними процесами; випадкові величини (збурення)

u_{1t} та u_{2t} є білим шумом. Крім того, припускається, що $\{u_{1t}\}$ та $\{u_{2t}\}$ не корелюють між собою.

Рівняння (1.2)—(2.2) є (VAR)-моделлю першого порядку, яка є зручною формою ілюстрації основних властивостей складніших вектор-авторегресійних моделей вищих порядків¹. Структура такої моделі дозволяє промодельовувати зворотні зв'язки між змінними. Дійсно, якщо $\gamma_{12} \neq 0$ та $\beta_{12} \neq 0$, то система (2.1)—(2.2) моделює зворотні зв'язки між минулим і поточним значеннями Y_{2t} та поточним значенням Y_{1t} . При цьому γ_{12} відображає зміну Y_{1t} при зміні Y_{2t} на одиницю, а β_{12} — зміну Y_{1t} при зміні $Y_{2,t-1}$ на одиницю. Аналогічна ситуація справедлива й для другого рівняння системи. Зрозуміло, що шоки, або збурення, u_{1t} та u_{2t} впливають і на Y_{1t} , і на Y_{2t} . Очевидно, якщо $\gamma_{21} \neq 0$, то u_{1t} має непрямий вплив на зміну Y_{2t} ; відповідно, якщо $\gamma_{12} \neq 0$, то u_{2t} має непрямий вплив на зміну Y_{1t} .

Основна відмінність VAR-моделей від симультативних систем рівнянь полягає в тому, що ми *a priori* не проводимо розподілу змінних на ендогенні та екзогенні. Всі змінні припускаються ендогенними².

Система (2.1)—(2.2) є структурною формою VAR-моделей, тому що Y_{1t} має вплив на Y_{2t} , а Y_{2t} , в свою чергу, впливає на зміну Y_{1t} . Як і у випадку симультативних систем рівнянь, структурну форму VAR-моделі завжди можна записати в приведеній формі, тобто виразити всі ендогенні змінні тільки через предетерміновані змінні. Для цього спочатку систему (2.1)—(2.2) запишемо у матричному вигляді:

$$\begin{bmatrix} 1 & \gamma_{21} \\ \gamma_{21} & 1 \end{bmatrix} \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \quad (2.3)$$

$$\text{або } LY_t = B_0 + B_1 Y_{t-1} + U_t, \quad (2.4)$$

$$\text{де } L = \begin{bmatrix} 1 & \gamma_{21} \\ \gamma_{21} & 1 \end{bmatrix}, Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix}, B_0 = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix}, B_1 = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}, U_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \quad (2.5)$$

Якщо помножити обидві частини (2.5) на L^{-1} , отримаємо VAR у приведеному, або, як прийнято називати у VAR-методології, стандартному, вигляді:

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t,$$

¹ Якщо $\gamma_{12} = \gamma_{21} = \beta_{11} = \beta_{21} = 0$, то система (2.1)—(2.2) розпадається на два окремих рівняння AR(1) відповідно для $\{Y_{1t}\}$ та $\{Y_{2t}\}$.

² Це властиво й іншим варіантам VAR, наприклад VARMA (включно зі складовою ковзної середньої).

$$\text{де } A_0 = L^{-1}B_0, A_1 = L^{-1}B_1, \varepsilon_t = L^{-1}U_t. \quad (2.6)$$

В розгорнутому вигляді (2.4) запишеться так:

$$Y_{1t} = a_{10} + a_{11}Y_{1,t-1} + a_{12}Y_{2,t-1} + \varepsilon_{1t}, \quad (2.7)$$

$$Y_{2t} = a_{20} + a_{21}Y_{1,t-1} + a_{22}Y_{2,t-1} + \varepsilon_{2t}, \quad (2.8)$$

де a_{10}, a_{20} — невідомі коефіцієнти, які є елементами матриці A_0 ; $a_{11}, a_{12}, a_{21}, a_{22}$ — невідомі коефіцієнти, які пов'язують поточні та минулі значення показників Y_{1t} та Y_{2t} і є елементами матриці A_1 ; ε_{1t} та ε_{2t} є елементами вектора шоків (збурень) U .

Важливо підкреслити, що випадкові величини ε_{1t} та ε_{2t} є складовими двох шоків (збурень): u_{1t} та u_{2t} , оскільки: $\varepsilon_{1t} = L^{-1}U_t$. Відповідно

$$\varepsilon_{2t} = \begin{pmatrix} \frac{u_{1t} - \gamma_{12}u_{2t}}{1 - \gamma_{12}\gamma_{21}} \\ \frac{u_{2t} - \gamma_{21}u_{1t}}{1 - \gamma_{12}\gamma_{21}} \end{pmatrix}.$$

Враховуючи те, що u_{1t} та u_{2t} , за припущенням, є білим шумом, тобто відповідають усім класичним припущенням: мають нульове математичне сподівання, постійну дисперсію та нульову автокореляцію, їх комбінація ε_{1t} та ε_{2t} також є білим шумом.

Продемонструємо це на прикладі ε_{1t} . Спочатку знайдемо математичне сподівання ε_{1t} .

$$E(\varepsilon_{1t}) = E\left(\frac{u_{1t} - \gamma_{12}u_{2t}}{1 - \gamma_{12}\gamma_{21}}\right) = 0. \quad (2.9)$$

Відповідно дисперсія ε_{1t} дорівнює:

$$E(\varepsilon_{1t}^2) = E\left[\left(\frac{u_{1t} - \gamma_{12}u_{2t}}{1 - \gamma_{12}\gamma_{21}}\right)^2\right] = \frac{\sigma_{u_1}^2 - \gamma_{12}^2\sigma_{u_2}^2}{(1 - \gamma_{12}\gamma_{21})^2} = \text{const}. \quad (2.10)$$

Як легко побачити з (2.10), дисперсія є постійною та незалежною від часу. Розглянемо тепер автокореляцію між ε_{1t} та $\varepsilon_{1,t-i}$:

$$E(\varepsilon_{1t}\varepsilon_{1,t-i}) = E\left(\frac{(u_{1t} - \gamma_{12}u_{2t})(u_{1,t-i} - \gamma_{12}u_{2,t-i})}{(1 - \gamma_{12}\gamma_{21})^2}\right) = 0 \text{ для } i \neq 0, \quad (2.11)$$

оскільки, за припущенням, білі шуми u_{1t} та u_{2t} не корелюють між собою.

Аналогічно можна показати, що помилка також є білим шумом, тобто має нульове математичне сподівання, постійну дисперсію та нульову автокореляцію. Слід, однак, зауважити, що ε_{1t} та ε_{2t} на відміну від u_{1t} та u_{2t} корелюють між собою і коефіцієнт кореляції між ними не дорівнює нулю.

$$E(\varepsilon_{1t}\varepsilon_{2t}) = E\left(\frac{(u_{1t} - \gamma_{12}u_{2t})(u_{2t} - \gamma_{21}u_{1t})}{(1 - \gamma_{12}\gamma_{21})^2}\right) = -\frac{\gamma_{21}\sigma_{u_1}^2 + \gamma_{12}\sigma_{u_2}^2}{(1 - \gamma_{12}\gamma_{21})^2} \neq 0. \quad (2.12)$$

Тільки в особливих випадках ($\gamma_{12} = \gamma_{21} = 0$) ε_{1t} та ε_{2t} не корелюють між собою.

Дисперсійно-коваріаційну матрицю VAR-моделі (2.7)—(2.8) можна записати у вигляді:

$$\text{VAR-COV} = \begin{bmatrix} \text{var}(\varepsilon_{1t}) & \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) \\ \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) & \text{var}(\varepsilon_{2t}) \end{bmatrix}. \quad (2.13)$$

Позначимо $\text{var}(\varepsilon_{it}) = \sigma_i^2$, $i = 1, 2$; $\sigma_{12} = \sigma_{21} = \text{cov}(\varepsilon_{1t}, \varepsilon_{2t})$, тоді (2.13) можна записати у вигляді:

$$\text{VAR-COV} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}. \quad (2.14)$$

Зазначимо, що якщо $a_{11} = a_{12} = 0$, то система (2.7)—(2.8) розпадається на два окремих рівняння AR(1) відповідно для $\{Y_{1t}\}$ та $\{Y_{2t}\}$.

Ще раз підкреслимо, що система (1.2)—(2.2) називається структурною VAR-моделлю, а (2.7)—(2.8) — VAR-моделлю в приведеній (reduced) формі. На практиці більш поширеним є останній вираз VAR-моделі, тому VAR-модель у приведеній формі частіше називається стандартною VAR-моделлю.

Для того, щоб виявити, чи існує взаємозалежність між показниками, які включаються у VAR-модель, використовують різні тести причинності, крім того, при інтерпретації вектор-авторегресійних моделей широко застосовують імпульсний аналіз (аналіз реагування на шоки) та декомпозицію дисперсій. В економетричному аналізі нова технологія такого роду дозволяє досконало описати та проінтерпретувати взаємозв'язки між економічними змінними. Імпульсний аналіз та декомпозицію дисперсій розглянемо пізніше. Спочатку з'ясуємо методи оцінювання невідомих параметрів VAR-моделей та проблеми, пов'язані з їх застосуванням.

2.2. ПРОБЛЕМИ ОТОТОЖНЕННЯ ТА ОЦІНЮВАННЯ VAR-МОДЕЛЕЙ У СТРУКТУРНІЙ ФОРМІ

Проблеми оцінки невідомих параметрів VAR-моделей легко пояснюють, чому саме VAR-моделі у приведеній формі широко використовуються на практиці. Параметри такої системи можна без проблем оцінити методом найменших квадратів (МНК). VAR-модель у структурній формі (1.2)—(2.2) не може бути оцінена звичайним МНК, тому що Y_{1t} корелює з випадковою величиною u_{2t} , а Y_{2t} корелює з u_{1t} . Але, як і у випадку симультативних систем рівнянь, ми можемо оцінити VAR у приведеній формі, а потім на основі оцінених коефіцієнтів такої системи знайти коефіцієнти структурної системи. При цьому, як завжди, виникає питання, чи можливо співставити результати оцінки системи (2.7)—(2.8) та системи (2.1)—(2.2), тобто питання ототожнення. На жаль, в загальному випадку зіставити результати оцінювання неможливо. Ми можемо зрозуміти чому, якщо порівняємо кількість параметрів структурної форми VAR з кількістю параметрів приведеної форми. Для нашого ілюстративного прикладу необхідно оцінити 6 параметрів моделі в приведеній формі: a_{10} , a_{20} , a_{11} , a_{12} , a_{21} , a_{22} , а також оцінити $\text{var}(\varepsilon_{it}) = \sigma_i^2$, $i = 1, 2$; $\sigma_{12} = \sigma_{21} = \text{cov}(\varepsilon_{1t}, \varepsilon_{2t})$, тобто всього 9 параметрів. З іншого боку, для VAR-моделі в структурній формі ми маємо оцінити 8 параметрів: γ_{10} , γ_{20} , γ_{12} , γ_{21} , β_{11} , β_{12} , β_{21} , β_{22} , а також $\sigma_{u_1} = \sigma_{Y_1}$ та $\sigma_{u_2} = \sigma_{Y_2}$, тобто 10 параметрів. Отже, якщо ми не накладемо обмеження на один з параметрів структурної форми моделі, то не зможемо її оцінити, виходячи з параметрів приведеної форми. Таким чином, система (2.1)—(2.2) буде в цьому випадку неототожненою. Якщо на один з параметрів буде накладено обмеження, система в цілому буде точно ототожненою; якщо обмеження накладено більше ніж на один параметр, система буде переототожненою.

Проілюструємо детальніше проблему точного ототожнення структурної VAR-моделі. Припустимо, що накладено наступне обмеження на параметр $\gamma_{21} = 0$. Перепишемо систему (2.1)—(2.2), враховуючи введене обмеження:

$$Y_{1t} = \gamma_{10} - \gamma_{12}Y_{2t} + \beta_{11}Y_{1,t-1} + \beta_{12}Y_{2,t-1} + u_{1t}, \quad (2.15)$$

$$Y_{2t} = \gamma_{20} + \beta_{21}Y_{1,t-1} + \beta_{22}Y_{2,t-1} + u_{2t}. \quad (2.16)$$

Як видно з (2.15)—(2.16), обмеження $\gamma_{21} = 0$ є випадком, при якому зміна впливає на величину Y_{1t} , але зміна Y_{1t} не впливає на зміну Y_{2t} . Матриця коефіцієнтів L^{-1} при цьому має вигляд:

$$L^{-1} = \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix}.$$

Помножимо (2.15)—(2.16) на L^{-1} :

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \quad (2.17)$$

З (2.17) після спрощення отримаємо:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} - \gamma_{12}\gamma_{20} \\ \gamma_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} - \gamma_{12}\beta_{21} & \beta_{12} - \gamma_{12}\beta_{22} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} - \gamma_{12}u_{2t} \\ u_{2t} \end{bmatrix}. \quad (2.18)$$

Оцінка (2.18) методом найменших квадратів аналогічна оцінці наступної системи в приведеній формі:

$$Y_{1t} = a_{10} + a_{11}Y_{1,t-1} + a_{12}Y_{2,t-1} + \varepsilon_{1t}, \quad (2.19)$$

$$Y_{2t} = a_{20} + a_{21}Y_{1,t-1} + a_{22}Y_{2,t-1} + \varepsilon_{2t}, \quad (2.20)$$

де $\varepsilon_{1t} = u_{1t} - \gamma_{12}u_{2t}$, $\varepsilon_{2t} = u_{2t}$; $a_{10} = \gamma_{10} - \gamma_{12}\gamma_{20}$; $a_{11} = \beta_{11} - \gamma_{12}\beta_{21}$;
 $a_{12} = \beta_{12} - \gamma_{12}\beta_{22}$; $a_{20} = \gamma_{20}$; $a_{21} = \beta_{21}$; $a_{22} = \beta_{22}$.

Відповідно елементи VAR-COV-матриці дорівнюють:

$$\text{Var}(\varepsilon_1) = \sigma_{Y_1}^2 + \gamma_{12}^2\sigma_{Y_2}^2; \text{Var}(\varepsilon_2) = \sigma_{Y_2}^2, \text{Cov}(\varepsilon_1, \varepsilon_2) = \gamma_{12}\sigma_{Y_2}^2.$$

Оцінивши параметри a_{10} , a_{20} , a_{11} , a_{12} , a_{21} , a_{22} , $\text{Var}(\varepsilon_1)$, $\text{Var}(\varepsilon_2)$, $\text{Cov}(\varepsilon_1, \varepsilon_2)$, ми відповідно можемо знайти параметри: γ_{10} , γ_{20} , γ_{12} , γ_{21} , β_{11} , β_{12} , β_{21} , β_{22} , $\sigma_{Y_1}^2$ та $\sigma_{Y_2}^2$, які є параметрами VAR-моделі в структурній формі.

Цей простий приклад дає нам змогу ввести важливе поняття у VAR-моделюванні, а саме поняття декомпозиції Чолеського (Choleski decomposition). Проаналізуємо детально систему (2.18). Обмеження, накладене на параметр $\gamma_{21} = 0$, призводить до того, що обидва шоки u_{1t} та u_{2t} впливають на поточне значення Y_{1t} , але тільки один шок u_{2t} впливає на Y_{2t} . Таким чином, значення ε_{2t} повністю визначається збуреннями (шоками) послідовності $\{Y_{2t}\}$. Декомпозиція помилок в такий спосіб (або трикутна декомпозиція) називається декомпозицією Чолеського: $\varepsilon_{1t} = u_{1t} - \gamma_{12}u_{2t}$, $\varepsilon_{2t} = u_{2t}$.

Аналогічно випадку точної ототожненості можна розглянути випадок переототожненості, коли на два або більше параметрів VAR у структурній формі накладені певні обмеження. Як правило, переототожненість³ призво-

³ У випадку переототожненості ми маємо більше, ніж потрібно, рівнянь для визначення параметрів структурної моделі з оцінених параметрів приведеної (скороченої) форми.

дить до нелінійних обмежень на параметри, але сучасні економетричні пакети легко розв'язують цю проблему. Отже, певні обмеження на параметри системи VAR у структурній формі дають можливість її оцінки на основі приведеної форми.

2.3. ОЦІНКА VAR-МОДЕЛЕЙ У ПРИВЕДЕНІЙ ФОРМІ

Як вже відмічалось вище, параметри VAR-моделі в приведеній (скороченій) формі без проблем можна оцінити методом найменших квадратів, якщо кількість лагових змінних в кожному рівнянні системи однакова. Якщо кількість лагів у різних рівняннях системи різна, то таку систему не можна оцінити методом найменших квадратів. Як правило, в таких випадках VAR-моделі оцінюються методом уявно⁴ непов'язаної регресії (seemingly unrelated regressions, SUR), який є доступним у багатьох економетричних пакетах прикладних програм.

Розглянемо проблему оцінювання VAR у приведеній формі детальніше для загального випадку. Досі ми лише ілюстрували основні поняття VAR-моделювання на прикладі найпростішої VAR(1)-моделі, моделі першого порядку. На практиці можливо побудувати VAR-модель, що складається з n рівнянь, кожне з яких, у свою чергу, включає p лагів для кожної з n змінних системи.

В загальному вигляді VAR-модель p -го порядку, що складається з n рівнянь, можна представити таким чином:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{nt} \end{bmatrix} = \begin{bmatrix} A_{10} \\ A_{20} \\ \vdots \\ A_{n0} \end{bmatrix} + \begin{bmatrix} A_{11}(L) & A_{12}(L) & \cdots & A_{1n}(L) \\ A_{21}(L) & A_{22}(L) & \cdots & A_{2n}(L) \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}(L) & A_{n2}(L) & \cdots & A_{nn}(L) \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-2} \\ \vdots \\ Y_{n,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}, \quad (2.21)$$

де A_{i0} — параметри, що є перетинами в кожному рівнянні VAR-системи, $A_{ij}(L)$ — поліном лагового оператора, $\{Y_{jt}\}$, $j = 1, 2, \dots, n$; $t = 1, 2, \dots, T$ — вектори ендогенних змінних. Відмітимо, що в систему (2.21) може бути включений також вектор екзогенних змінних, наприклад $\{X_{it}\}$, $i = 1, 2, \dots, k$. Індивідуальні коефіцієнти, що входять в поліном лагового оператора $A_{ij}(L)$, часто позначаються $a_{ij}(1)$, $a_{ij}(2)$, ... Як правило, всі рівняння системи мають однакову кількість лагів. Ми також будемо вважати, що всі поліноми лагового

⁴ В даному контексті слово «уявно» потрібно розуміти як «за припущенням», оскільки регресії напевно будуть пов'язаними в більшості випадків, тобто кореляція залишків регресій буде напевно відмінною від нуля.

оператора $A_{ij}(L)$ мають однаковий порядок, наприклад p . $\{\varepsilon_{jt}\}$, $j = 1, 2, \dots, n$ — вектор випадкових величин (innovations), які можуть корелювати між собою, але не корелюють зі своїми лаговими значеннями та відповідно лаговими ендогенними змінними. Дисперсійно-коваріаційна матриця, як правило, позначається $\Sigma_{(n \times n)}$ і має відповідно порядок $(n \times n)$.

Найкраща оцінка кожного рівняння системи (2.21) може бути отримана за допомогою МНК. Можна показати, що в цьому випадку (при однаковій кількості лагів у кожному рівнянні системи) оцінювання МНК дає консистентні⁵ (consistent) та асимптотично ефективні (asymptotically effective) оцінки⁶. Відмітимо, що припущення про відсутність серійної автокореляції не є вирішальним в даному випадку, оскільки вона може бути абсорбована збільшенням кількості лагів у системі.

Якщо деякі рівняння VAR-системи вміщують змінні, які не включені в інші рівняння, то оцінка системи на основі МНК неможлива. В такому випадку VAR-модель необхідно оцінювати методом уявно непов'язаних регресій SUR, що забезпечує ефективні оцінки при різній кількості лагів у рівняннях системи.

Відмітимо також, що кількість лагів, яку необхідно включати у VAR-модель, є певним компромісом між можливою помилкою специфікації та різким зменшенням ступенів свободи. Дійсно, якщо в модель включено p лагів, то кожне з n рівнянь системи вміщує $n \cdot p$ коефіцієнтів плюс перетин. Це звичайно вимагає достатньої довжини часових рядів. Отже, оптимальний вибір порядку VAR-моделі є серйозною проблемою в аналізі часових рядів. Зазначимо, що, як і у випадку ARIMA-моделі, універсальної методології не існує, але можна використовувати певні процедури, які досить ефективні у практичному застосуванні. Розглянемо можливі підходи до вибору порядку VAR-моделей.

2.4. ВИБІР ПОРЯДКУ (p) VAR-МОДЕЛІ

Найпростішим підходом до вибору порядку VAR-моделі є аналіз довжини часового ряду. Оцінюючи порядок VAR-моделі, можна спиратись на попередній досвід, аналіз динаміки процесів, але звичайно ці підходи є суб'єктивними і можуть використовуватись лише як перше наближення для визначення порядку. Крім того, незначна довжина часового ряду є досить суворим обмеженням на порядок VAR-моделі.

⁵ Тобто межа по ймовірності оцінки параметра існує і збігається до значення параметра для генеральної сукупності.

⁶ Тобто асимптотично мають найменшу дисперсію серед класу консистентних оцінок.

Існують і більш формальні критерії, одним з яких є АІС-критерій (інформаційний критерій Акайка). Ми розглядали застосування інформаційного АІС-критерію у випадку АRІМА-моделей, тепер розглянемо його для вибору порядку VAR-моделей.

На першому етапі обирається максимально можливе значення порядку VAR-моделі, позначимо його p^* . Далі оцінюється методом найменших квадратів певна множина VAR-моделей з різною кількістю лагів $p = 1, 2, 3, \dots, p^*$. Для кожної з оцінених моделей розраховується АІС-статистика за формулою:

$$AIC(p) = \ln|\hat{V}| + \frac{2pn^2}{T}, \quad p = 1, 2, 3, \dots, p^*, \quad (2.22)$$

де n — кількість часових рядів у VAR-моделі; T — кількість спостережень (нагадаємо, що вона має бути однаковою для кожного часового ряду); p — порядок оціненої VAR-моделі; $|\hat{V}|$ — детермінант коваріаційної матриці помилок моделі, оціненої методом найменших квадратів.

Серед оцінених моделей обирається модель порядку p_{\max} ($0 \leq p_{\max} \leq p^*$) з найменшим значенням АІС.

2.5. ЗВ'ЯЗОК ВЛАСНИХ ЗНАЧЕНЬ λ_i ТА РАНГУ МАТРИЦІ A

У VAR-моделюванні корисним є зв'язок власних значень матриці A та її рангу. Нагадаємо, що рангом матриці $A_{(n \times n)}$ є кількість незалежних рядків (стовпців) матриці. Як правило, ранг матриці позначається через r : $\text{rang}(A) = r$. Матриця має повний ранг, якщо $\text{rang}(A) = n$. Для нашого аналізу важливо, що ранг матриці дорівнює також кількості ненульових її власних значень λ_i . Розглянемо спочатку особливі випадки. Нехай $\text{rang}(A) = n$, це означає, що матриця має повний ранг, а відповідно всі її характеристичні значення відрізняються від нуля. Дійсно, якщо всі рядки матриці A лінійно незалежні, то детермінант такої матриці не дорівнює нулю. Отже, виходячи з широко відомої властивості власних векторів, а саме:

$$|A| = \prod_{i=1}^n \lambda_i,$$

для того, щоб детермінант не дорівнював нулю, необхідно, щоб всі власні значення матриці відрізнялись від нуля.

В іншому особливому випадку, коли $\text{rang}(A) = 0$, кожний елемент матриці A повинен дорівнювати нулю. В такому випадку характеристичне рівняння $|(A - \lambda I)| = 0$ перетворюється на рівняння $\lambda^n = 0$ з єдиною можливим розв'язком $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

Розглянемо тепер загальний випадок: $0 < \text{ранг}(A) = r < n$. Оскільки перестановка рядків не змінює значення детермінанта, ми завжди можемо переписати $\left| (A - \lambda I) \right| = 0$ так, що першими r -рядками будуть лінійно незалежні рядки матриці A . Таким чином, детермінант перших r -рядків включатиме r характеристичних коренів (власних значень), а інші $(n - r)$ корені будуть дорівнювати нулю.

2.6. ПОНЯТТЯ СТАБІЛЬНОСТІ VAR-МОДЕЛЕЙ

Стабільність VAR-моделей є необхідною умовою їх практичного застосування. Стабільність означає, що послідовність зовнішніх шоків для VAR-системи має кінцевий спадний ефект, тобто якщо шоки затухають з часом, VAR-модель є стаціонарною. Крім того, можна обчислити довгострокові мультиплікатори [13, 17, 20]. Детальніше пояснення шоків та довгострокових мультиплікаторів наводиться в розділі 2.7.

Поняття та умови стабільності VAR-моделей проілюструємо на прикладі VAR-моделі першого порядку:

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t, \quad (2.23)$$

де $\{Y_t\} = \{Y_{1t}, Y_{2t}, \dots, Y_{nt}\}$ є вектор порядку $(n \times 1)$; A_0 — вектор порядку $(n \times 1)$ з елементами; A_1 — квадратна матриця $(n \times n)$ з елементами a_{ij} ; $\{\varepsilon_t\} = \{\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{nt}\}$ — вектор збурень порядку $(n \times 1)$.

Для того, щоб перевірити стабільність системи (2.23), ми повинні брати до уваги тільки однорідне рівняння:

$$Y_t = A_1 Y_{t-1}. \quad (2.24)$$

Ми можемо використати метод невизначених коефіцієнтів та представити рішення (2.24) у вигляді: $Y_t = c \lambda^t$ [11, 20]. Якщо підставити його в (2.24), отримаємо:

$$c \lambda^t = A_1 c \lambda^{t-1}, \text{ звідки } \lambda c = A c, \text{ або } (A - \lambda I) c = 0, \quad (2.25)$$

де I є одиничною матрицею $(n \times n)$, λ — невідомий скаляр⁷; c — невідомий вектор порядку $(n \times 1)$.

⁷ Для тих, хто знайомий з матричною алгеброю, зауважимо, що λ є не що інше, як власне (або характеристичне) значення матриці (eigenvalue), а вектор c — характеристичний вектор (eigenvector), відповідний знайденому значенню λ .

Якщо матриця $(A - \lambda I)$ є несингулярною (невиродженою), то очевидно ми можемо отримати з (2.25) тільки тривіальне рішення $Y = 0$. Якщо матриця є сингулярною (виродженою), то ми отримуємо ненульове рішення. Сингулярність матриці $(A - \lambda I)$ означає, що її детермінант має дорівнювати нулю:

$$|(A - \lambda I)| = 0. \quad (2.26)$$

Умова (2.26) є характеристичним рівнянням матриці A , яке є поліномом n -го степеня відносно невідомих значень λ , а отже, повинно мати n коренів. Дійсно, детермінант $|(A - \lambda I)| = 0$ обов'язково вміщує значення λ^n , яке впливає з виразу:

$$(a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda) \dots (a_{nn} - \lambda).$$

Таким чином, характеристичне рівняння (2.26) є поліномом n -го порядку:

$$\lambda^n + z_1 \lambda^{n-1} + z_2 \lambda^{n-2} + \dots + z_{n-1} \lambda + z_n = 0. \quad (2.27)$$

З (2.27) випливає, що матриця A порядку $(n \times n)$ повинна мати n характеристичних коренів λ_i , $i = 1, 2, \dots, n$, які ще називаються власними значеннями (eigenvalues) матриці A . Якщо кожне значення λ_i підставити в (2.25), то ми отримуємо відповідні їм вектори c , які називаються характеристичними, або власними, векторами матриці A .

Необхідною та достатньою умовою стабільності (стаціонарності) VAR-моделі є умова, щоб усі власні значення λ_i матриці A знаходились в межах одиничного кола, тобто їх значення за абсолютною величиною повинні бути менші за одиницю $|\lambda_i| < 1$.

Якщо розглянути узагальнену VAR-модель p -го порядку

$$Y_{1t} = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t, \quad (2.28)$$

то стабільність системи залежить від власних векторів матриці: $\pi = I - A_1 - A_2 - \dots - A_p$, які є коренями рівняння

$$|\lambda^p I - \lambda^{p-1} A_1 - \dots - \lambda A_{p-1} - A_p| = 0. \quad (2.29)$$

2.7. ПРОГНОЗУВАННЯ НА ОСНОВІ VAR-МОДЕЛЕЙ

Основною метою розробки VAR-моделей, як і у випадку ARIMA-моделей, є їх широке використання для прогнозів⁸. При цьому на основі VAR-моделей ми можемо отримати одночасно прогнози багатьох взаємопов'язаних економічних показників. Звичайно, передбачення майбутніх значень часових рядів має проводитись, як і у випадку ARIMA-моделей, з мінімально можливою помилкою.

Нагадаємо, що при прогнозуванні можливі два типи помилок: пов'язані з різницею між дійсними та оціненими коефіцієнтами моделі, що використовується для прогнозу, та пов'язані з ігноруванням випадкових величин (майбутніх збурень). Класично беруть до уваги лише другий тип помилок, тому при прогнозуванні намагаються мінімізувати саме його.

Як і раніше, позначимо прогноз Y_t в період часу $(t + s)$ через \hat{Y}_{t+s} , або $\hat{Y}_t(s)$. При цьому $s \geq 1$ є прогнозним періодом, або періодом випередження. Відповідно помилки прогнозу дорівнюють: $e_{t+s} = Y_{t+s} - \hat{Y}_{t+s}$. З огляду на те що значення помилок можуть бути як від'ємними, так і додатними, використовується поняття мінімуму середнього квадрата помилок (MSE). Відповідно, оптимальним є прогноз, при якому мінімізується середній квадрат помилок прогнозу. Тобто обирається таке прогнозне значення \hat{Y}_{t+s} , при якому мінімізується $E(e_{t+s}^2) = E\{(Y_{t+s} - \hat{Y}_{t+s})^2\}$. Зауважимо, що, оскільки помилка прогнозу є випадковою величиною, ми мінімізуємо математичне сподівання квадрата помилок.

Мінімізація середнього квадрата помилок аналогічна прогнозному значенню \hat{Y}_{t+s} , отриманому як умовне сподівання Y_{t+s} при даних всіх спостереженнях часового ряду до періоду t , тобто $\hat{Y}_{t+s} = E(Y_{t+s} | Y_t, Y_{t-1}, \dots, Y_1)$.⁹

Проілюструємо загальний принцип розрахунку прогнозу для випадку VAR-моделювання на прикладі найпростішої VAR-моделі в стандартному вигляді:

$$Y_{1t} = a_{10} + a_{11}Y_{1,t-1} + a_{12}Y_{2,t-1} + \varepsilon_{1t}, \quad (2.29)$$

$$Y_{2t} = a_{20} + a_{21}Y_{1,t-1} + a_{22}Y_{2,t-1} + \varepsilon_{2t}. \quad (2.30)$$

⁸ Надзвичайно важливим є також застосування VAR-моделей для визначення причинності у зв'язках між змінними та аналізу функцій імпульсного відгуку (розглянемо пізніше), що дають можливість знайти не лише величину остаточного (довгострокового) впливу, а й короткострокову динаміку взаємодії змінних. Крім того, VAR-моделі дають можливість ідентифікувати джерело шоків за допомогою так званих структурних VAR-моделей.

⁹ Наводимо це твердження без доведення.

Перепишемо (2.29)—(2.30) в матричному вигляді:

$$Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t. \quad (2.31)$$

Для спрощення виключимо з (2.31) вектор перетинів і отримаємо

$$Y_t = AY_{t-1} + \varepsilon_t. \quad (2.32)$$

Таким чином, якщо дійсні значення за VAR(1)-моделлю (2.32) в період часу $(t + 1)$

$$Y_{t+1} = AY_t + \varepsilon_{t+1},$$

то відповідно прогнозні значення в період часу $(t + 1)$

$$\hat{Y}_{t+1} = E(Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1) = AY_t,$$

оскільки $E(\varepsilon_{t+1}) = 0$.

Вектор помилок прогнозу в період часу $(t + 1)$

$$e_{t+1} = Y_{t+1} - \hat{Y}_{t+1} = \varepsilon_{t+1}.$$

Вектор дійсних значень за VAR(1)-моделлю в період часу $(t + 2)$

$$Y_{t+2} = AY_{t+1} + \varepsilon_{t+2} = A^2 Y_t + A\varepsilon_{t+1} + \varepsilon_{t+2}.$$

Прогнозні значення Y в період $(t + 2)$

$$\hat{Y}_{t+2} = E(Y_{t+2} | Y_t, Y_{t-1}, \dots, Y_1) = A^2 Y_t.$$

Помилки прогнозу

$$e_{t+2} = Y_{t+2} - \hat{Y}_{t+2} = A\varepsilon_{t+1} + \varepsilon_{t+2}.$$

Відповідно для періоду $(t + s)$ отримаємо:

$$Y_{t+s} = A^s Y_t + A^{s-1} \varepsilon_{t+1} + \dots + A\varepsilon_{t+s-1} + \varepsilon_{t+s}. \quad (2.33)$$

$$\hat{Y}_{t+s} = E(Y_{t+s} | Y_t, Y_{t-1}, \dots, Y_1) = A^s Y_t. \quad (2.34)$$

Помилки прогнозу

$$e_{t+s} = Y_{t+s} - \hat{Y}_{t+s} = \varepsilon_{t+s} + A\varepsilon_{t+s-1} + A^{s-1}\varepsilon_{t+1}. \quad (2.35)$$

Якщо позначити VAR-COV-матрицю векторів випадкових величин через $\Sigma = E(\varepsilon_t \varepsilon_t')$, то з виразу (2.35) легко побачити, що VAR-COV-матриця для помилок в період $(t + s)$ відповідно дорівнює:

$$\Sigma(s) = \Sigma + A\Sigma A' + A^2\Sigma(A')^2 + \dots + A^{s-1}\Sigma(A')^{s-1}. \quad (2.36)$$

Ілюстрація принципу прогнозування на основі найпростішої VAR-моделі демонструє надзвичайну простоту отримання прогнозів, коли вся необхідна для цього інформація міститься лише в часових рядах досліджуваних змінних. Сучасні пакети прикладних програм дозволяють отримати прогнози для VAR-моделей вищих порядків з великою кількістю досліджуваних показників.

На відміну від ARIMA- VAR-моделі дозволяють проводити також економічний аналіз результатів. Звичайно, самі коефіцієнти VAR-моделей важко проінтерпретувати, але ми можемо проінтерпретувати результати функції імпульсних відгуків (impulse response function) та декомпозицію дисперсій (variance decomposition). Розглянемо ідею аналізу функції імпульсних відгуків та декомпозиції дисперсій детальніше.

2.8. ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ МОДЕЛЮВАННЯ

2.8.1. Аналіз функції імпульсних відгуків

Функція імпульсних відгуків показує зміну ендогенних показників у відповідь на шок (зміну одного зі збурень системи).

Розглянемо VAR у приведеній формі:

$$Y_{1t} = a_{10} + a_{11}Y_{1,t-1} + a_{12}Y_{2,t-1} + \varepsilon_{1t}, \quad (2.37)$$

$$Y_{2t} = a_{20} + a_{21}Y_{1,t-1} + a_{22}Y_{2,t-1} + \varepsilon_{2t}.$$

Зміна в ε_{1t} миттєво викликає зміну в Y_{1t} , а також зміну в усіх майбутніх значеннях Y_1 та Y_2 , через динамічну структуру системи (2.37). Функція імпульсних відгуків якраз і описує зміну поточного та майбутніх значень

ендогенних змінних як реакцію на шок, що дорівнює одному середньоквадратичному відхиленню (standard deviation).

Проілюструємо цю ідею простим прикладом. Для спрощення припускаємо, що перетини в (2.37) відсутні, тобто в матричному вигляді маємо:

$$Y_t = AY_{t-1} + \varepsilon_t.$$

Нехай нам відома матриця коефіцієнтів:

$$A = \begin{pmatrix} 0.3 & 0.1 \\ 0.5 & 0.2 \end{pmatrix}$$

та варіаційно-коваріаційна матриця збурень:

$$\text{VAR-COV} = \Sigma = \begin{pmatrix} 9 & 7 \\ 7 & 16 \end{pmatrix}.$$

Перший елемент варіаційно-коваріаційної матриці є дисперсією першого збурення, тобто $\sigma_{\varepsilon_1}^2 = 9$. Побудуємо імпульсну функцію в припущенні, що перше збурення змінюється на одне середньоквадратичне відхилення, тобто на 3. Крім того, припустимо, що початкові значення ендogenous змінних дорівнюють нулю:

$$Y_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \varepsilon_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

За нашим припущенням, в перший період часу ε_{11} зростає на 4, а в інші проміжки часу знову спадає до нуля; друге збурення залишається без змін. Проаналізуємо ланцюгову зміну Y як реакцію на одноразовий шок, викликаний зміною першого збурення.

$$Y_1 = \begin{pmatrix} 0.3 & 0.1 \\ 0.5 & 0.2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

$$Y_2 = \begin{pmatrix} 0.3 & 0.1 \\ 0.5 & 0.2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 1.5 \end{pmatrix}.$$

$$Y_3 = \begin{pmatrix} 0.3 & 0.1 \\ 0.5 & 0.2 \end{pmatrix} \begin{pmatrix} 0.9 \\ 1.5 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.42 \\ 0.85 \end{pmatrix}.$$

$$Y_4 = \begin{pmatrix} 0.3 & 0.1 \\ 0.5 & 0.2 \end{pmatrix} \begin{pmatrix} 0.42 \\ 0.85 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.211 \\ 0.38 \end{pmatrix}.$$

.....

Цей процес можна продовжити далі, відповідно імпульсна функція має вигляд:

Імпульсна функція (IRF) від $\varepsilon_1 = (30)$		
Період	y_{1t}	y_{2t}
1	3	0
2	0.9	1.5
3	0.42	0.85
4	0.211	0.38
...

Сучасні пакети видають графічне зображення імпульсної функції. Аналогічно можна побудувати імпульсну функцію для другого збурення. Таким чином, вона вимірює ефект на значення ендогенних змінних системи в поточний та майбутні періоди часу, викликаний зміною одного з показників у поточний період часу на одне середньоквадратичне відхилення в поточний період часу.

Зауважимо, що у VAR-моделях в приведеній формі інтерпретація імпульсної функції ускладнюється, оскільки збурення в такій системі, на відміну від структурної форми, корелюють між собою. Якщо кореляція існує, то відповідно є загальний компонент, який не може бути пов'язаний лише з однією змінною. Одне з компромісних рішень цієї проблеми полягає в штучному присвоєнні ефекту від загального компоненту першій змінній VAR-моделі. Звичайно, це не вирішує проблеми в цілому, крім того, зі зміною порядку рівнянь змінюється й сама імпульсна функція, що ускладнює її інтерпретацію.

Іншим виходом є представлення збурень приведенної форми VAR-моделей через збурення структурної форми, які за припущенням не корелюють між собою. Розглянемо таку можливість детальніше. Це дозволить більш формально описати ідею імпульсної функції.

Нагадаємо, що авторегресійні (AR) процеси завжди можна представити у вигляді процесів ковзної середньої. По аналогії вектор-авторегресійні моделі (VAR) можна представити у вигляді вектор-моделей ковзної середньої (VMA). Дійсно, повернемося до найпростішої VAR-моделі (2.31) та розпишемо її як VMA:

$$\begin{aligned}
 Y_t &= A_0 + A_1 Y_{t-1} + \varepsilon_t = A_0 + A_1 (A_0 + A_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \\
 &= (I + A_1) A_0 + A_1^2 Y_{t-2} + A_1 \varepsilon_{t-1} + \varepsilon_t,
 \end{aligned}
 \tag{2.38}$$

де I — одинична матриця другого порядку.

Після n ітерацій відповідно отримаємо:

$$\begin{aligned} Y_t &= (A_0 + A_1 + \dots + A_1^n) A_0 + \sum_{i=0}^n A_1^i \varepsilon_{t-1} + A_1^{n+1} Y_{t-n-1} = \\ &= (I + A_1)^{-1} A_0 + \sum_{i=0}^n A_1^i \varepsilon_{t-i} + A_1^{n+1} Y_{t-n-1}. \end{aligned} \quad (2.39)$$

Звичайно, якщо кількість ітерацій нескінченно збільшувати, останній член зникає і з (2.39) ми отримаємо граничне значення, яке має такий вираз:

$$Y_t = (I - A_1)^{-1} A_0 + \sum_{i=0}^{\infty} A_1^i \varepsilon_{t-i} = \mu + \sum_{i=0}^{\infty} A_1^i \varepsilon_{t-i}, \quad (2.40)$$

де $\mu = (\bar{Y}_1 \bar{Y}_2)$; $\bar{Y}_1 = [a_{10}(1 - a_{22}) + a_{12}a_{20}]/D$; $\bar{Y}_2 = [a_{20}(1 - a_{11}) + a_{21}a_{10}]/D$;

$$D = (1 - a_{11})(1 - a_{22}) - a_{12}a_{21}.$$

Таким чином, (2.40) є представленням VAR-моделі у вигляді ковзної середньої, тобто представленням кожної ендогенної змінної у вигляді залежності виключно від збурень, або, іншими словами, від двох типів шоків (ε_{1t} та ε_{2t}). Отже, саме VMA-представлення VAR-моделей дає змогу відобразити ланцюговий імпульс різноманітних шоків на змінні VAR-моделі.

Проілюструємо імпульсну функцію в більш формалізованому вигляді на прикладі все тієї ж найпростішої VAR-моделі, розписавши (2.31) детальніше:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \quad (2.41)$$

або розписавши детально (2.40), отримаємо VAR в іншому вигляді:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \quad (2.42)$$

Саме (2.42) виражає Y в залежності від значень збурень ε_{1t} та ε_{2t} . Нагадаємо, що збурення ε_{1t} та ε_{2t} корелюють між собою, в той час як збурення структурної форми цієї VAR-моделі (2.3) u_{1t} та u_{2t} між собою не корелюють. Тому в (2.42) ε_{1t} та ε_{2t} краще виразити через u_{1t} та u_{2t} . Нагадаємо формулу (2.6), яка пов'язує ці збурення у випадку найпростішої вектор-авторегресійної моделі:

$$\varepsilon_t = L^{-1}U_t, \text{ де } L^{-1} = \begin{bmatrix} 1 & \gamma_{12} \\ \gamma_{21} & 1 \end{bmatrix}^{-1}; U_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \quad (2.43)$$

Відповідно (2.43) можна розписати у вигляді:

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{bmatrix} 1 / (1 - \gamma_{12}\gamma_{21}) \\ -\gamma_{21} \quad 1 \end{bmatrix} \begin{bmatrix} 1 & -\gamma_{12} \\ -\gamma_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \quad (2.44)$$

Об'єднавши (2.44) та (2.42), отримаємо:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} + \begin{bmatrix} 1 / (1 - \gamma_{12}\gamma_{21}) \\ -\gamma_{21} \quad 1 \end{bmatrix} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -\gamma_{12} \\ -\gamma_{21} & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \quad (2.45)$$

Для спрощення вигляду (2.44) введемо нові позначення, а саме матрицю порядку (2×2) H_i з елементами $h_{jk}(i)$:

$$H_i = \begin{bmatrix} A_1^i / (1 - \gamma_{12}\gamma_{21}) \\ -\gamma_{21} \quad 1 \end{bmatrix} \begin{bmatrix} 1 & -\gamma_{12} \\ -\gamma_{21} & 1 \end{bmatrix}. \quad (2.46)$$

Відповідно (2.44) можна записати в нових позначеннях:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} h_{11}(i) & h_{12}(i) \\ h_{21}(i) & h_{22}(i) \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \quad (2.47)$$

або

$$Y_t = \mu + \sum_{i=0}^{\infty} H_i u_{t-i}. \quad (2.48)$$

Коефіцієнти матриці H_i можуть використовуватись для аналізу ефектів на повну зміну досліджуваних показників Y , що викликана шоками в u_{1t} або в u_{2t} . Наприклад, чотири елементи матриці в початковий (нульовий) період часу $h_{jk}(0)$ є миттєвими мультиплікаторами (impact multipliers). При цьому: $h_{11}(0)$ визначає миттєву зміну Y_{1t} у відповідь на одиничну зміну в u_{1t} ; $h_{12}(0)$ визначає миттєву зміну Y_{1t} у відповідь на одиничну зміну в u_{2t} ; $h_{21}(0)$ визначає миттєву зміну Y_{2t} у відповідь на одиничну зміну в u_{1t} ; $h_{22}(0)$ визначає миттєву зміну Y_{2t} у відповідь на одиничну зміну в u_{2t} .

Відповідно елементи матриці $h_{jk}(1)$ відображають зміну показників через один період у відповідь на зміну в u_{1t} та в u_{2t} : $h_{11}(1)$ відображає зміну в $Y_{1,t+1}$ у відповідь на одиничну зміну в u_{1t} ; $h_{12}(1)$ відображає зміну в $Y_{1,t+1}$ у відповідь на одиничну зміну в u_{2t} ; $h_{21}(1)$ відображає зміну в $Y_{2,t+1}$ у відповідь на одиничну зміну в u_{1t} ; $h_{22}(1)$ відображає зміну в $Y_{2,t+1}$ у відповідь на одиничну зміну в u_{2t} . Аналогічно інтерпретуються елементи матриці для різних періодів випередження.

Чотири множини коефіцієнтів: $h_{11}(i)$, $h_{12}(i)$, $h_{21}(i)$, $h_{22}(i)$ називаються імпульсними функціями відгуків. Графічне відображення цих коефіцієнтів

залежно від періоду часу (i) дає можливість візуального аналізу поведінки досліджуваних показників у залежності від різних шоків.

Сумарний (акумуляований) ефект одиничного імпульсу в u_{1t} (або в u_{2t}) може бути отриманий шляхом додавання відповідних коефіцієнтів імпульсної функції відгуку. Наприклад, через n періодів часу ефект одиничної зміни в u_{2t} на значення $Y_{1, t-1}$ є $h_{12}(n)$. Таким чином, через n періодів сумарний ефект від зміни в u_{2t} на послідовність значень $\{Y_{1t}\}$ дорівнює сумі

$$\sum_{i=0}^{\infty} h_{12}(i).$$

Якщо $n \rightarrow \infty$, то ми отримаємо відповідно довгострокові мультиплікатори (long-run multiplier).

На жаль, ідея виразу випадкових величин приведеної форми ε_{1t} та ε_{2t} через збурення структурної форми u_{1t} та u_{2t} має істотний недолік. Як вже демонструвалось вище, часто VAR-модель в структурній формі є недоототожненою, тому ми не можемо оцінити її параметри та вимушені вводити певні обмеження для ідентифікації моделі, найбільш поширеним з яких є декомпозиція Чолеського.

Іншим підходом є ортогоналізація ε_{1t} та ε_{2t} . Ідея ортогоналізації полягає в перетворенні випадкових величин ε_{1t} та ε_{2t} , які корелюють між собою, на такі, що не корелюють, тобто ортогональні. Пояснення ідеї ортогоналізації викликає необхідність введення додаткових понять з теорії матриць, тому в рамках даної роботи ми її детально не розглядатимемо.

2.8.2. Декомпозиція дисперсій помилок прогнозів у VAR-моделюванні

Для того, щоб глибше зрозуміти взаємозв'язки між показниками у VAR-системі, корисно розглянути властивості помилок прогнозу, а також декомпозицію їх дисперсій. Декомпозиція дисперсій помилок прогнозів дозволяє проаналізувати вплив різноманітних шоків на дисперсію помилки прогнозів для різних періодів випередження. Проілюструємо ідею декомпозиції дисперсії помилки прогнозу на прикладі найпростішої VAR (2.31): $Y_t = A_0 + A_1 Y_{t-1} + \varepsilon_t$. Припустимо, що ми оцінили параметри цієї моделі, тобто знаємо коефіцієнти матриць A_0 та A_1 і хочемо отримати прогнози на декілька періодів уперед: \hat{Y}_{t+i} , $i = 1, 2, 3, \dots$ Для кожного періоду ми відповідно отримуватимемо помилки прогнозу, різницю між дійсними значеннями та прогнозами. Нагадаємо, як їх отримують. Дійсні значення за VAR(1)-моделлю (2.31) в період часу $(t + 1)$ дорівнюють:

$$Y_{t+1} = A_0 + A_1 Y_t + \varepsilon_{t+1}.$$

Відповідно прогнознi значення в період часу $(t + 1)$

$$\hat{Y}_{t+1} = E(Y_{t+1} | Y_t, Y_{t-1}, \dots, Y_1) = A_0 + A_1 Y_t.$$

Звідси помилки прогнозу в період часу $(t + 1)$ $e_{t+1} = Y_{t+1} - \hat{Y}_{t+1} = \varepsilon_{t+1}$.

Аналогічно для другого періоду випередження $(t + 2)$ дійсні значення

$$Y_{t+2} = A_0 + A_1 Y_{t+1} + \varepsilon_{t+2} = A_0 + A_1 (A_0 + A_1 Y_t + \varepsilon_{t+1}) + \varepsilon_{t+2}.$$

Прогнознi значення Y в період $(t + 2)$:

$$\hat{Y}_{t+2} = E(Y_{t+2} | Y_t, Y_{t-1}, \dots, Y_1) = (I + A_1)A_0 + A_1^2 Y_t.$$

Помилки прогнозу

$$e_{t+2} = Y_{t+2} - \hat{Y}_{t+2} = A_1 \varepsilon_{t+1} + \varepsilon_{t+2}.$$

Відповідно для періоду $(t + s)$ отримаємо:

$$\hat{Y}_{t+s} = E(Y_{t+s} | Y_t, Y_{t-1}, \dots, Y_1) = (I + A_1 + A_1^2 + \dots + A_1^{t-1})A_0 + A_1^t Y_t.$$

Помилки прогнозу для періоду $(t + s)$

$$e_{t+s} = Y_{t+s} - \hat{Y}_{t+s} = \varepsilon_{t+s} + A_1 \varepsilon_{t+s-1} + \dots + A_1^{s-1} \varepsilon_{t+1}. \quad (2.49)$$

Таким чином, ми отримали послідовність помилок прогнозу для найпростішої VAR-моделі. Тепер розглянемо інше представлення цієї моделі, а саме у формі ковзного середнього (див. (2.48)):

$$Y_t = \mu + \sum_{i=0}^{\infty} H_i u_{t-i}.$$

Звичайно, як VAR-модель, так і її вираз у формі VMA, описують абсолютно однакові процеси, тому ми можемо представити помилки прогнозу як функцію від збурень u , використовуючи останню форму представлення. Відповідно помилка прогнозу на один період вперед дорівнюватиме $e_{t+1} = Y_{t+1} - \hat{Y}_{t+1} = H_0 u_{t+1}$. В загальному виразі на s періодів вперед помилка прогнозу

$$e_{t+s} = Y_{t+s} - \hat{Y}_{t+s} = \left(\mu + \sum_{i=0}^{\infty} H_i u_{t+s-i} \right) - \hat{Y}_{t+s} = \sum_{i=0}^{s-1} H_i u_{t+s-i}. \quad (2.50)$$

З виразу (2.50) випишемо помилки прогнозу для першого показника Y_1 в період $(t + s)$:

$$e_{t+s} = Y_{t+s} - \hat{Y}_{t+s} = h_{11}(0)u_{1,t+s} + h_{11}(1)u_{1,t+s-1} + \dots + h_{11}(s-1)u_{1,t+1} + \\ + h_{12}(0)u_{2,t+s} + h_{12}(1)u_{2,t+s-1} + \dots + h_{12}(s-1)u_{2,t+1}. \quad (2.51)$$

Позначимо дисперсію помилки прогнозу в період $(t + s)$ через $\sigma_{Y_1}^2(S)$, тоді

$$\sigma_{Y_1}^2(S) = \sigma_{Y_1}^2 [h_{11}^2(0) + h_{11}^2(1) + \dots + h_{11}^2(s-1)] + \sigma_{Y_2}^2 [h_{12}^2(0) + h_{12}^2(1) + \dots + \\ + h_{12}^2(s-1)], \quad (2.52)$$

де $\sigma_{Y_1}^2 = \sigma_{u_1}^2$; $\sigma_{Y_2}^2 = \sigma_{u_2}^2$.

З формули (2.52) легко побачити, що дисперсія помилки прогнозу для s -го періоду випередження є функцією відповідно двох шоків u_1 та u_2 , більш того, ми можемо знайти частки дисперсії, пов'язані з кожним окремим шоком. Так, частка дисперсії помилки прогнозу першого показника, яка викликана шоками цього показника, в загальній дисперсії помилки прогнозу є:

$$\frac{\sigma_{Y_1}^2 [h_{11}^2(0) + h_{11}^2(1) + \dots + h_{11}^2(s-1)]}{\sigma_{Y_1}^2(s)}, \quad (2.53)$$

а частка дисперсії помилки прогнозу першого показника, яка викликана шоками другого показника, в загальній дисперсії помилки прогнозу дорівнює

$$\frac{\sigma_{Y_1}^2 [h_{12}^2(0) + h_{12}^2(1) + \dots + h_{12}^2(s-1)]}{\sigma_{Y_1}^2(s)}. \quad (2.54)$$

Отже, декомпозиція дисперсій дає змогу оцінити пропорції дисперсії, викликані шоками різних змінних, та, відповідно, оцінити ступінь впливу взаємозв'язків між показниками. Наприклад, якщо шоки в u_2 жодним чином не пояснюють дисперсію помилки прогнозу показника Y_1 (відповідна доля дисперсії дорівнює нулю) для всіх періодів випередження, ми можемо вважати, що Y_1 є екзогенним показником. Якщо зобразити пропорції дисперсій відповідно періоду випередження на числовій осі, ми отримаємо графіки для декомпозиції дисперсій, які легко інтерпретувати. Загалом імпульсна функ-

ція відгуків та декомпозиція дисперсії становлять інструментарій аналізу та декомпозицій шоків (innovation accounting).

Відмітимо, що проблеми з декомпозицією дисперсій є аналогічними проблемам з імпульсною функцією відгуків та пов'язані насамперед з оцінкою послідовностей збурень $\{Y_{1t}\}$ та $\{Y_{2t}\}$. Для їх оцінок інколи можна використати декомпозицію Чолеського, частіше ми повинні використовувати складніші методи, які в рамках цієї роботи ми наводити не будемо. Зазначимо лише, що сучасні пакети прикладних економетричних програм без проблем дають змогу проаналізувати аналітично та графічно декомпозицію дисперсій помилок прогнозу.

2.9. ЗАСТОСУВАННЯ VAR(p)-МОДЕЛЕЙ НА ПРАКТИЦІ: АНАЛІЗ ЗВ'ЯЗКУ МІЖ СПОЖИВАННЯМ ТА ДОХОДОМ

Для того, щоб проілюструвати практичне застосування VAR(p)-моделей, розглянемо модель взаємозв'язку між споживанням та доходом. Зазначимо, що дану модель ми розглянемо на умовних даних, але ускладнену версію VAR з механізмом корегування помилки в наступному розділі ми прокоментуємо для моделювання податкових надходжень на реальних даних для України.

Більшість економетричних моделей побудовані на певних припущеннях щодо реальних економічних процесів. Економетристи накладають обмеження на функціональну форму, набір змінних у рівняннях, число лагів тощо. Такі припущення рідко можна заперечити на практиці, оскільки часто вони мають неявний характер. При цьому, якщо припущення не виконуються, отримані оцінки можуть бути зміщеними та неефективними. Для того, щоб уникнути проблем, пов'язаних з формулюванням припущень на основі певного знання економічної теорії, розроблено новий економетричний апарат, зокрема вектор-авторегресійні моделі (VAR). Часто такий тип моделей називають атеоретичними, оскільки при їх побудові економічна теорія майже не використовується. Принцип «нехай дані пояснюють самі себе» є основою VAR-моделювання. На жаль, такий підхід не залишає багато можливостей для економічного аналізу, тому VAR-моделі зазвичай застосовують для прогнозування, хоча результати прогнозу доповнюються аналізом імпульсних функцій відгуку та декомпозиції дисперсій.

Перед розглядом конкретного прикладу з VAR-моделювання ще раз підкреслимо, що стаціонарність часових рядів є одним з фундаментальних припущень в традиційних економетричних дослідженнях. Інакше кажучи, необхідно, щоб ряди мали постійну дисперсію та математичне сподівання, значення яких не залежать від періоду часу. У випадку нестаціонарних часових рядів оцінені коефіцієнти моделей можуть бути хибними, а помилка

регресії (estimated error) — вельми викривленою. Щоб уникнути такої пастки, нестационарні часові ряди необхідно перетворити на стаціонарні операцією різниць.

Просту VAR-модель p -го порядку традиційно представляють у вигляді:

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{i=1}^p \beta_i x_{t-i},$$

$$x_t = \sum_{i=1}^p \gamma_i y_{t-i} + \sum_{i=1}^p \mu_i x_{t-i}.$$

Неважко помітити, що порядок моделі (p) треба якимось чином визначити. На жаль, немає жодних формальних тестів чи методологій, щоб точно визначити правильний порядок VAR-моделі. Як правило, для цього можуть бути застосовані лише евристичні методи. Нагадаємо, що зазвичай використовують Акайк- або Шварц-критерії, які пропонують «емпірично ефективну» кількість лагів у VAR-моделях. З цього моменту (цієї точки зору) моделювання стає мистецтвом. Незважаючи на певну невизначеність стосовно параметра p , треба пам'ятати, що краще додати до моделі зайві лаги, аніж зменшити їх необхідну кількість. В останньому випадку в моделі можливе зміщення оцінок, викликане помилкою специфікації, в той час як у першому випадку можлива лише втрата деякої ефективності оцінених коефіцієнтів.

Таким чином, першим практичним етапом побудови VAR-моделі є визначення порядку p . Далі необхідно оцінити VAR-модель, зробити прогноз та провести аналіз функції імпульсних відгуків та декомпозиції дисперсій.

Для розрахунку конкретного прикладу ми використаємо квартальні дані, наведені в книзі Р. А. Ruud [34].

Дані доступні також на web-сторінці: <http://www.econ-line.com/data/NQ.zip>. Крім того, вони наведені в таблиці додатку 2.

Завдання 2.1.

Необхідно побудувати VAR(p)-модель залежності між споживанням та доходом, що залишається у розпорядженні домогосподарств, провести аналіз та зробити прогноз на наступний квартал.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. Зауважимо, що в процесі виконання цього завдання ми використовували пакет E.VIEWS.

Для виконання завдання введемо такі позначення:

VAR(p) — вектор-авторегресійна модель p -го порядку; p — порядок вектор-авторегресійної моделі; LNC — ряд логарифмів реальних значень споживання короткострокових товарів (квартальна інформація); DLNC — ряд пер-

ших різниць логарифмів реальних значень споживання короткострокових товарів (квартальна інформація); LNY — ряд реальних значень доходу, що залишається у розпорядженні домогосподарств (квартальна інформація); DLNY — ряд перших різниць логарифмів реальних значень доходу, що залишається у розпорядженні домогосподарств (квартальна інформація). Дані наведені в таблиці додатку 2.

Зауважимо, що на web-сторінці (<http://www.econ-line.com/data/NQ.zip>) наведені лише перші різниці даних часових рядів, тому необхідно самостійно зробити відтворення первісних рядів. Результат цього перетворення також наведено в таблиці додатку 2.

Послідовність виконання завдання 2.1.

Крок 1. Вибір порядку (p) та оцінка VAR-моделі.

Перед визначенням порядку VAR-моделі необхідно перевірити часові ряди логарифмів споживання та доходу на стаціонарність, використовуючи розширений тест Дікі—Фуллера (ADF). Чи є вони стаціонарними?

Побудуйте та проаналізуйте корелограми відповідних часових рядів. Якщо коефіцієнти автокореляції (ACF) та часткової автокореляції (PACF) підтримують гіпотезу нестационарності, поясніть, на чому базується ваш висновок.

У випадку нестационарності рядів перевірте, чи будуть перші різниці стаціонарними рядами. Поясніть ваші висновки.

Використовуючи критерії Акайка та Шварца, визначте порядок VAR-моделі. Прокоментуйте результати.

Крок 2. Імпульсний аналіз та аналіз декомпозиції дисперсій на основі оціненої VAR-моделі.

На основі оціненої на першому кроці моделі проведіть аналіз імпульсної функції відгуків та декомпозиції дисперсій. Поясніть отримані результати.

Крок 3. Прогнозування на основі оціненої VAR-моделі.

На основі оціненої на першому кроці VAR(11)-моделі знайдіть прогноз на один квартал вперед. Поясніть отримані результати.

Крок 4. Висновки.

Зробіть загальний висновок щодо побудованої VAR(p)-моделі (зробіть ваші висновки самостійно).

Розв'язок завдання 2.1 та розширений коментар моделювання за допомогою VAR-моделей

Зауважимо, що побудова та аналіз VAR-моделей проводиться за допомогою економетричного пакета E.VIEWS.

Крок 1. Вибір порядку (p) та оцінювання VAR-моделі.

У більшості економетричних підходів основним є припущення про стаціонарність часових рядів. У нашому прикладі натуральний логарифм від функції споживання не є стаціонарним рядом в рівнях, про що свідчать результати розширеного тесту Дікі—Фуллера (ADF).

Таблиця 2.1. Результати перевірки тесту Дікі—Фуллера (ADF) на стаціонарність ряду LNC

ADF Test Statistic	-1.037008	1%	Critical Value*	-4.0068
		5%	Critical Value	-3.4332
		10%	Critical Value	-3.1402

*McKinnon critical values for rejection of hypothesis of a unit root.

Але він є стаціонарним у перших різницях.

Таблиця 2.2. Результати перевірки тесту Дікі—Фуллера (ADF) на стаціонарність ряду DLNC

ADF Test Statistic	-5.901050	1%	Critical Value*	-4.0070
		5%	Critical Value	-3.4333
		10%	Critical Value	-3.1402

*McKinnon critical values for rejection of hypothesis of a unit root.

Такий самий висновок отримуємо після аналізу корелограми часового ряду LNC.

Таблиця 2.3. Автокореляційна та часткова автокореляційна функції часового ряду LNC

Sample: 1947:2 2000:4

Included observations: 205

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
*****	*****	1	0.986	0.986	202.32	0.000
*****	.	2	0.972	-0.018	399.86	0.000
*****	.	3	0.958	-0.006	592.69	0.000
*****	.	4	0.944	-0.003	780.89	0.000
*****	.	5	0.930	-0.013	964.42	0.000
*****	.	6	0.916	-0.002	1143.4	0.000
*****	.	7	0.902	-0.011	1317.8	0.000

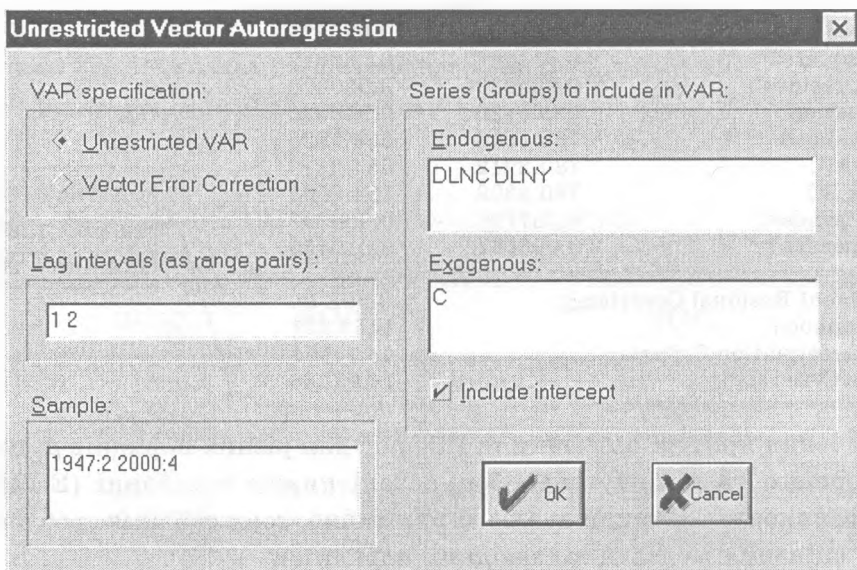
.*****	..	8	0.888	-0.010	1487.6	0.000
.*****	..	9	0.874	-0.017	1652.8	0.000
.*****	..	10	0.859	-0.008	1813.4	0.000
.*****	..	11	0.845	-0.004	1969.6	0.000
.*****	..	12	0.831	-0.001	2121.4	0.000
.*****	..	13	0.817	-0.004	2268.9	0.000
.*****	..	14	0.803	-0.020	2412.0	0.000
.*****	..	15	0.789	0.001	2550.9	0.000
.*****	..	16	0.774	-0.017	2685.5	0.000
.*****	..	17	0.760	-0.003	2815.9	0.000
.*****	..	18	0.746	-0.009	2942.2	0.000
.*****	..	19	0.732	-0.012	3064.3	0.000
.*****	..	20	0.718	0.001	3182.4	0.000

Подібні висновки можна зробити і стосовно ряду LNY (натурального логарифма від доходу).

Оскільки часові ряди LNC та LNY є нестационарними, але їх перші різниці утворюють стаціонарні ряди, то для побудови VAR-моделей використовуються відповідно тільки перші різниці споживання та доходу, тобто ряди DLNC та DLNY.

Для знаходження порядку VAR-моделей спочатку припустимо, що найвищий порядок VAR може дорівнювати 16 (приблизно 10 % кількості спостережень). Відповідно, можливі VAR-моделі першого, другого, третього і т. д. до шістнадцятого порядку. Звичайно аналіз ми починаємо з VAR-моделі першого порядку.

Для того, щоб оцінити VAR(p), оберіть змінні DLNC та DLNY як групу. Потім у вікні, що з'явилося, Procs/Make Vector autoregressive...



У новому вікні необхідно задати порядок VAR. Наприклад, для $p = 1$ в полі *Lag intervals (as range pairs)* напишіть пару 1 1. Для $p = 2$ пишемо 1 2. Для $p = k$, пишемо 1 k .

VAR-модель можна оцінити за допомогою команди E.VIEWS, наприклад, для оцінки VAR-моделі першого порядку з ім'ям VAR1.LS використовується така команда:

VAR1.LS 1 1 DLNC DLNY @ C

(формат команди: ім'я VAR-моделі. 1 р Y1 Y2 @ екзогенна змінна 1 екзогенна змінна 2, де p — порядок VAR).

Таблиця 2.4. Результати оцінки VAR(1)-моделі

Sample(adjusted): 1947:3 1998:3

Included observations: 205 after adjusting

Endpoints

Standard errors & t-statistics in parentheses

	DLNC	DLNY
DLNC(-1)	0.108491 (0.07426) (1.46096)	0.612448 (0.13736) (4.45878)
DLNY(-1)	0.094712 (0.03909) (2.42301)	-0.136060 (0.07230) (-1.88183)
C	0.006166 (0.00066) (9.37191)	0.004729 (0.00122) (3.88652)
R-squared	0.060604	0.089625
Adj. R-squared	0.051303	0.080612
Sum sq. residues	0.005933	0.020300
S.E. equation	0.005420	0.010025
Log likelihood	780.2619	654.1826
Akaike AIC	780.2912	654.2119
Schwarz SC	780.3398	654.2605
Mean dependent	0.007798	0.008397
S.D. dependent	0.005564	0.010455
Determinant Residual Covariance		2.39E-09
Log Likelihood		1453.271
Akaike Information Criteria		-14.11972
Schwarz Criteria		-14,02246

Рівняння
доходу

Статистика AIC
та SIC для вибору
порядку моделі

Значення критеріїв Акайка та Шварца для різних порядків p , починаючи з першого та закінчуючи 16-м, подані нижче у таблиці (E.VIEWS не дозволяє виконувати такий аналіз автоматично, тому таблиця зроблена після окремо оцінених моделей відповідних порядків).

Порядок моделі (P)	Акайк-критерій	Шварц-критерій	Порядок моделі (P)	Акайк-критерій	Шварц-критерій
1	-14.1197	-14.0225	9	-14.1467	-13.5134
2	-14.111	-13.9483	10	-14.1654	-13.463
3	-14.1333	-13.9048	11	-14.1065	-13.3344
4	-14.1584	-13.8636	12	-14.2736	-13.4314
5	-14.1923	-13.8307	13	-14.2485	-13.3356
6	-14.2093	-13.7805	14	-14.2086	-13.2246
7	-14.1853	-13.6888	15	-14.2917	-13.236
8	-14.2016	-13.6369	16	-14.2406	-13.1127

Як можна побачити з таблиці, інформаційні критерії Акайка та Шварца не дають однозначної відповіді щодо вибору специфікації моделі. В загальному випадку одностайність досягається рідко, і доводиться робити певні компромісні рішення, використовуючи інші критерії. Варто також зауважити, що ми аналізуємо попарне значення критеріїв при однакових порядках p , інакше треба застосовувати інші критерії для його визначення.

В даному випадку значення критерію Шварца постійно зростає, значення критерію Акайка, коливаючись, поступово спадає. Оскільки дані критерії не дають переваги жодному з порядків, використаємо той факт, що остання статистично значима кроскореляція цих змінних спостерігається на 15-му лазі. З таблиці ми можемо помітити, що значення критерію не дуже відрізняються від оптимальних значень. Крім того, аналіз коефіцієнтів детермінації, хоча це і надзвичайно недосконалий критерій, також віддає перевагу моделі з порядком 15. Отже, VAR(15) є кінцевою моделлю, оцінку якої можна отримати, використовуючи команди меню, наведені вище, або прямою командою E.VIEWS, яка для нашого прикладу, як було показано, має вигляд:

VAR11.LS 1 15 DLNC DLNY @ C.

Таблиця 2.4. Результат оцінки моделі VAR (15)

Vector Autoregression Estimates

Date: 11/23/01 Time: 11:34

Sample(adjusted): 1951:1 1998:3

Included observations: 191 after adjusting endpoints

Standard errors in () & t-statistics in []

	DLNC	DLNY		DLNC	DLNY
DLNC(-1)	0.180748 (0.08422) [2.14608]	0.676620 (0.14705) [4.60140]	DLNC(-1)	0.100276 (0.05043) [1.98845]	0.178799 (0.08805) [-2.03073]
DLNC(-2)	0.157853 (0.08908) [1.77213]	0.107461 (0.15552) [0.69098]	DLNC(-2)	-0.022655 (0.05199) [-0.43579]	-0.093591 (0.09076) [-1.03114]

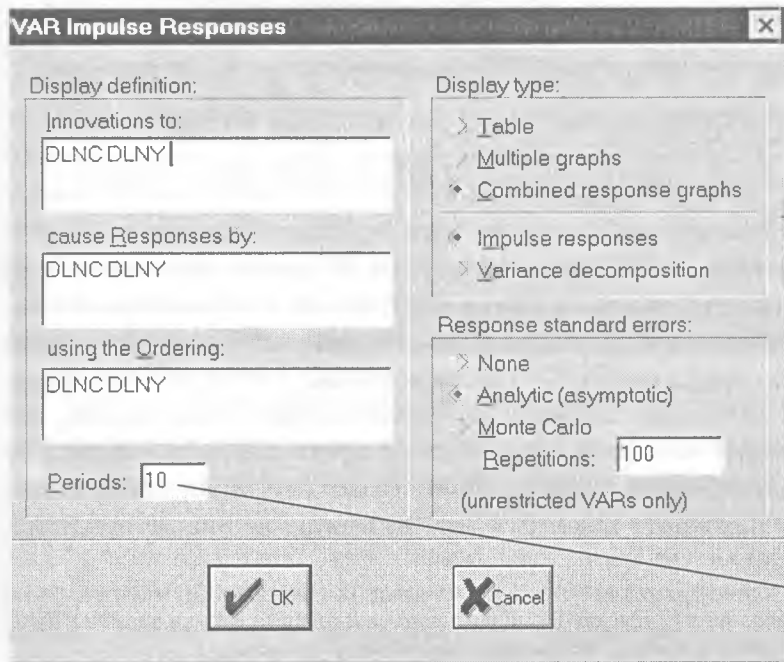
	DLNC	DLNY		DLNC	DLNY
DLNC(-3)	0.186318 (0.08922) [2.08838]	0.392900 (0.15577) [2.52237]	DLNY(-3)	-0.022451 (0.05183) [-0.43317]	-0.128622 (0.09049) [-1.42140]
DLNC(-4)	-0.118598 (0.09085) [-1.30546]	0.428479 (0.15861) [2.70139]	DLNY(-4)	0.043629 (0.04773) [0.91404]	-0.143853 (0.08334) [-1.72618]
DLNC(-5)	-0.159072 (0.09170) [-1.73476]	-0.040927 (0.16010) [-0.25564]	DLNY(-5)	-0.038753 (0.04761) [-0.81400]	-0.228076 (0.08312) [-2.74395]
DLNC(-6)	0.044545 (0.09019) [0.49391]	0.149561 (0.15746) [0.94982]	DLNY(-6)	0.104935 (0.04774) [2.19794]	0.071432 (0.08335) [0.85697]
DLNC(-7)	0.057238 (0.08999) [0.63604]	0.040939 (0.15712) [0.26056]	DLNY(-7)	0.022039 (0.04788) [0.46028]	-0.099384 (0.08360) [-1.18883]
DLNC(-8)	-0.186033 (0.08878) [-2.09554]	-0.096846 (0.15500) [-0.62483]	DLNY(-8)	-0.021982 (0.04805) [-0.45752]	-0.136743 (0.08389) [-1.63013]
DLNC(-9)	-0.027484 (0.08891) [-0.30912]	0.040153 (0.15523) [0.25866]	DLNY(-9)	0.081157 (0.04814) [1.68574]	0.051486 (0.08406) [0.61253]
DLNC(-10)	-0.064276 (0.08530) [-0.75351]	-0.011900 (0.14893) [-0.07990]	DLNY(-10)	0.087532 (0.04726) [1.85196]	0.114193 (0.08252) [1.38381]
DLNC(-11)	0.074663 (0.08461) [0.88245]	0.086528 (0.14772) [0.58575]	DLNY(-11)	0.033679 (0.04635) [0.72665]	0.021236 (0.08092) [0.26243]
DLNC(-12)	-0.023078 (0.08393) [-0.27495]	0.336836 (0.14654) [2.29858]	DLNY(-12)	-0.003599 (0.04384) [-0.08208]	-0.079146 (0.07655) [-1.03395]
DLNC(-13)	-0.221041 (0.08308) [-2.66047]	-0.188053 (0.14506) [-1.29639]	DLNY(-13)	-0.003628 (0.04397) [-0.08251]	-0.013707 (0.07677) [-0.17855]
DLNC(-14)	0.067216 (0.08029) [0.83719]	-0.149673 (0.14018) [-1.06773]	DLNY(-14)	0.013790 (0.04299) [0.32078]	0.021285 (0.07506) [0.28359]
DLNC(-15)	0.080325 (0.07540) [1.06529]	-0.184980 (0.13165) [-1.40512]	DLNY(-15)	-0.029358 (0.04090) [-0.71778]	0.079393 (0.07141) [1.11177]

	DLNC	DLNY
C	0.004738 (0.00144) [3.29229]	0.001628 (0.00251) [0.64787]
R-squared	0.323273	0.330931
Adj. R-squared	0.196387	0.205481
Sum sq. resid	0.003505	0.010684
S.E. equation	0.004680	0.008172
F-statistic	2.547743	2.637944
Log likelihood	770.4932	664.0507
Akaike AIC	-7.743384	-6.628803
Schwarz SC	-7.215528	-6.100947
Mean dependent	0.007989	0.008099
S.D. dependent	0.005221	0.009168
Determinant Residual Covariance	1.11E-09	
Log Likelihood (d.f. adjusted)	1426.857	
Akaike Information Criteria	-14.29170	
Schwarz Criteria	-13.23599	

Крок 2. Імпульсний аналіз та аналіз декомпозиції дисперсій на основі оціненої VAR-моделі.

Аналіз оцінених коефіцієнтів у VAR-моделях, особливо для виявлення короткострокових ефектів, часто не має великого значення, оскільки дуже важко відокремити частковий миттєвий ефект від зміни певної змінної, особливо коли кількість лагів є значною. Ось чому для кращого розуміння дина-

мічних властивостей моделі кориснішим є аналіз імпульсних функцій відгуків (impulse response functions — IRF). Його можна зробити, вибравши **Impulse** у вікні VAR-об'єкта. При цьому відкривається нове вікно:



Визначте горизонт для функції IRF

Головне, що нам треба зробити, це визначити прогнозний горизонт (Periods); змінні, які впливають на досліджувані значення ('affecting' variables) (Innovations to:), та змінні, що піддаються впливу ('affected' variables) (Cause Responses by:). Варто зауважити, що число періодів треба обрати достатньо великим, щоб можливо було побачити, чи є система стабільною.

Ми також можемо визначити тип вихідного аналізу: функції імпульсних відгуків або декомпозицію дисперсій (Impulse responses або Variance decomposition). Імпульсна функція відгуків (IRF) показує явну динаміку зміни всіх змінних всередині системи у відповідь на зміну на одне середньоквадратичне відхилення однієї з них. З іншого боку, декомпозиція дисперсії показує відносну важливість факторів у впливі на динаміку зміни (дисперсію) конкретної змінної системи.

Для нашого прикладу функції імпульсних відгуків представлено на рисунку 2.1.

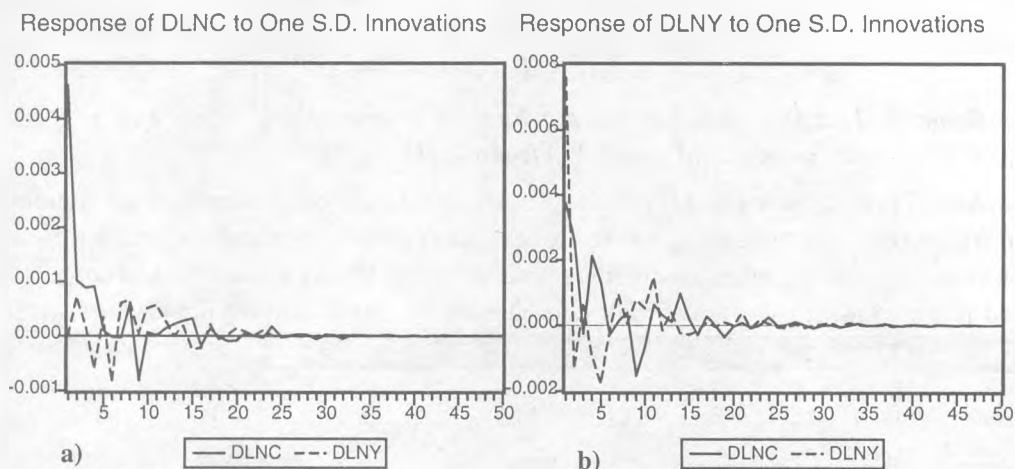


Рис. 2.1. Функції імпульсних відгуків DLNC та DLNY для VAR(15)

Рисунок 2.1.а) показує, як зміни на одне середньоквадратичне відхилення (шок одного стандартного відхилення) в DLNY (пунктирна лінія) та в DLNC (суцільна лінія) впливають на зміну натурального логарифма споживання (DLNC). З зовнішнього вигляду імпульсних функцій відгуків можна помітити, наприклад, що зміна DLNY (зміна в доході) в одне середньоквадратичне відхилення викликає спочатку позитивні зміни у споживанні, але згодом матиме негативні наслідки. Тому чистий ефект від змін у доході не є ясним. Загалом це відповідає теорії постійного доходу, коли споживання слабо залежить від поточного доходу¹⁰. Тим не менше не важко помітити,

¹⁰ Наголосимо, що цей висновок залежить від природи процесу, який лежить в основі часового ряду доходу. Детальнішу інформацію щодо теорії споживання можна знайти в будь-якому підручнику з макроекономіки, наприклад підручнику Менк'ю.

що флуктуації зменшуються і наближаються до нуля зі зростанням часу. З іншого боку, шок в DLNC справляє відносно сильний і тривалий позитивний ефект на динаміку DLNY, що також відповідає теорії постійного доходу. В підсумку, система є стабільною, оскільки відгук згасає та асимптотично наближається до нуля, і система врешті-решт досягає певного стійкого стану.

Рисунок 2.1.b) показує відповідь DLNY на шоки в DLNY (пунктирна лінія) та в DLNC (суцільна лінія). Прокоментуйте його самостійно.

Тепер розглянемо детально декомпозицію дисперсій, яка є відносним індикатором, тому часто виражається у процентах. Для нашого прикладу функції декомпозиції дисперсії представлені на рисунку 2.2.

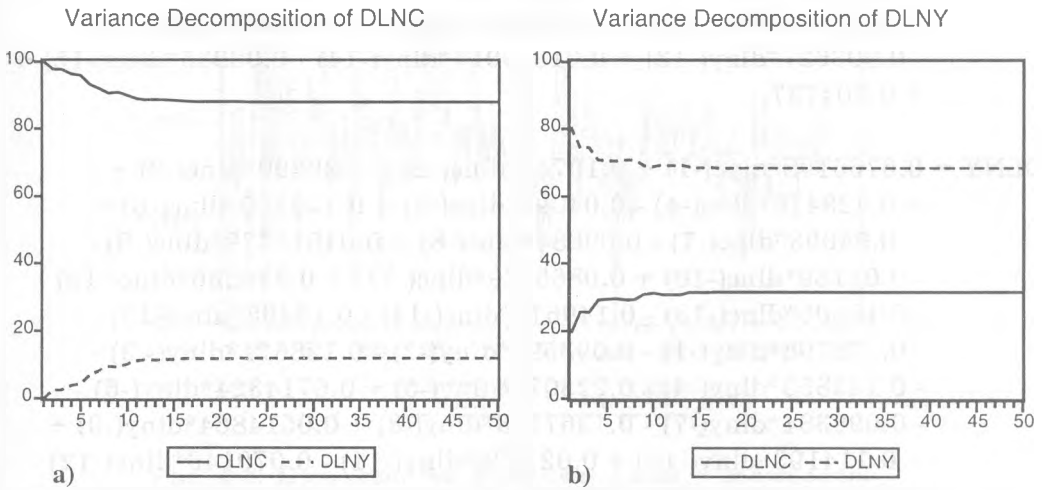


Рис. 2.2. Графіки декомпозиції дисперсій DLNC та DLNY для VAR(11)

Декомпозиція дисперсії для DLNC (рисунок 2.2.a)) показує, що спочатку зміни у споживанні практично пояснюють самі себе, точніше, **віртуально пояснюють самі себе** (virtually self-explanatory). З часом зміни в доході починають відігравати дедалі важливішу роль. Проте ця роль відносно незначна: не більше 20 % від дисперсії DLNC пояснюється змінами в доході.

Як можна побачити з графіка 2.2.b), зміни у доході більш чутливі до змін у споживанні. Наведемо приклад.

Крок 3. Прогнозування на основі оціненої VAR-моделі.

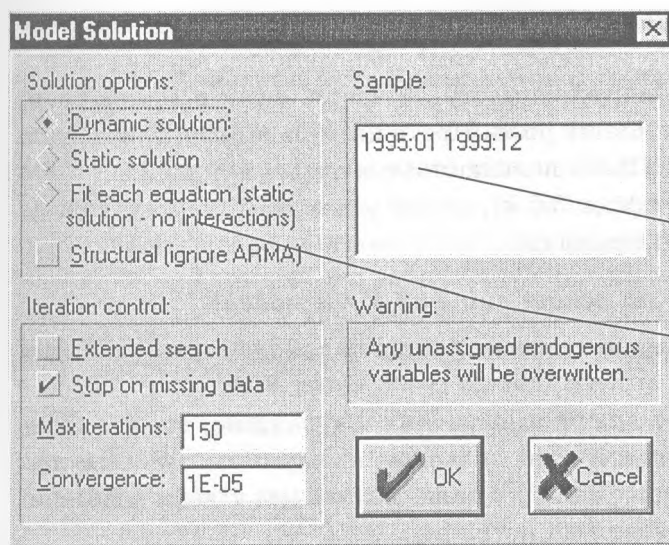
Оцінену VAR(15)-модель ми можемо використати для прогнозування змін у доході та споживанні. Для цього оберіть в меню вікна VAR-об'єкта Procs/Make Model — і модель готова для прогнозування. Зауважимо, що для отримання прогнозу необхідно надати явні значення екзогенним змінним для періоду прогнозування. В даному випадку наша модель має тільки ендогенні змінні, тому їх значення задавати не потрібно.

В умовах нашого прикладу модель для прогнозу матиме вигляд:

ASSIGN @ALL F

$$\begin{aligned}
 DLNC = & 0.180747*dlnc(-1) + 0.15785*dlnc(-2) + 0.186318*dlnc(-3) - \\
 & - 0.118597*dlnc(-4) - 0.159071*dlnc(-5) + 0.044545*dlnc(-6) + \\
 & + 0.057238*dlnc(-7) - 0.18603*dlnc(-8) - 0.027484*dlnc(-9) - \\
 & - 0.064276*dlnc(-10) + 0.074662*dlnc(-11) - 0.023077*dlnc(-12) - \\
 & - 0.2210412*dlnc(-13) + 0.067216*dlnc(-14) + 0.080324*dlnc(-15) + \\
 & + 0.1002762*dlny(-1) - 0.0226551*dlny(-2) - 0.02245*dlny(-3) + \\
 & + 0.04362*dlny(-4) - 0.038752*dlny(-5) + 0.10493*dlny(-6) + \\
 & + 0.02203*dlny(-7) - 0.0219819*dlny(-8) + 0.08115*dlny(-9) + \\
 & + 0.087531*dlny(-10) + 0.033678*dlny(-11) - 0.00359*dlny(-12) - \\
 & - 0.003627*dlny(-13) + 0.01379017*dlny(-14) - 0.02935*dlny(-15) + \\
 & + 0.004737,
 \end{aligned}$$

$$\begin{aligned}
 DLNY = & 0.6766198*dlnc(-1) + 0.10746*dlnc(-2) + 0.39290*dlnc(-3) + \\
 & + 0.428478*dlnc(-4) - 0.04092*dlnc(-5) + 0.149560*dlnc(-6) + \\
 & + 0.04093*dlnc(-7) - 0.09684*dlnc(-8) + 0.04015279*dlnc(-9) - \\
 & - 0.01189*dlnc(-10) + 0.0865279*dlnc(-11) + 0.336836*dlnc(-12) - \\
 & - 0.18805*dlnc(-13) - 0.149673*dlnc(-14) - 0.18498*dlnc(-15) - \\
 & - 0.178798*dlny(-1) - 0.093591*dlny(-2) - 0.128621*dlny(-3) - \\
 & - 0.143853*dlny(-4) - 0.228075*dlny(-5) + 0.0714324*dlny(-6) - \\
 & - 0.099384*dlny(-7) - 0.1367433*dlny(-8) + 0.0514864*dlny(-9) + \\
 & + 0.114193*dlny(-10) + 0.021236*dlny(-11) - 0.079146*dlny(-12) - \\
 & - 0.013706*dlny(-13) + 0.02128*dlny(-14) + 0.079392*dlny(-15) + \\
 & + 0.0016
 \end{aligned}$$



Щоб розв'язати модель, виберіть **Solve** в меню вікна **Model-object Window** та визначте параметри прогнозування.

Визначте прогнозний горизонт: початкову та кінцеву дату

Solution options встановлює, чи лагові ендегенні змінні треба брати з емпіричних рядів (*Static solution*), чи з теоретичних (*Dynamic solution*)

Зауважимо, що якщо командою *ASSIGN @ALL F* визначено розширення *F*, то прогнозні значення змінних автоматично зберігаються пакетом *E.VIEWS* з іменами рядів фактичних значень з розширенням *F* (в нашому прикладі — це часові ряди прогнозних значень *DLNCF* та *DLNYF*). Командою *ASSIGN @ALL* можна надавати й інші розширення для рядів прогнозних даних.

Щоб перевірити точність прогнозу, зробленого на основі оціненої *VAR(15)*-моделі, порівняємо візуально фактичні та розраховані значення часових рядів, побудувавши відповідні графіки.

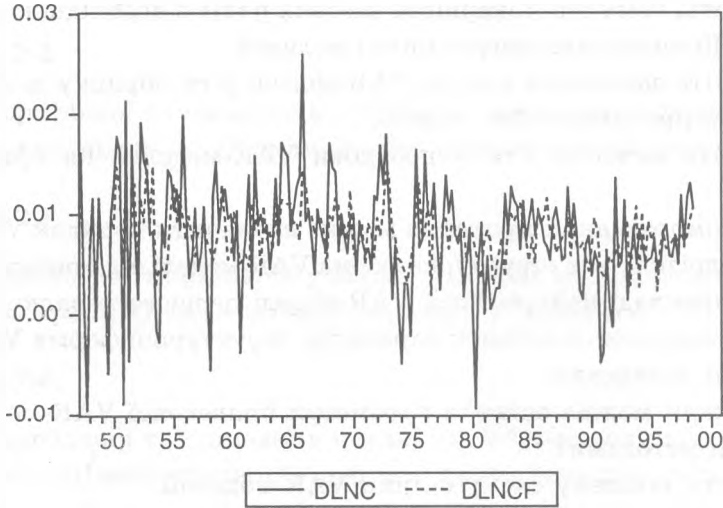


Рис. 2.3. Прогнозні та фактичні зміни у споживанні

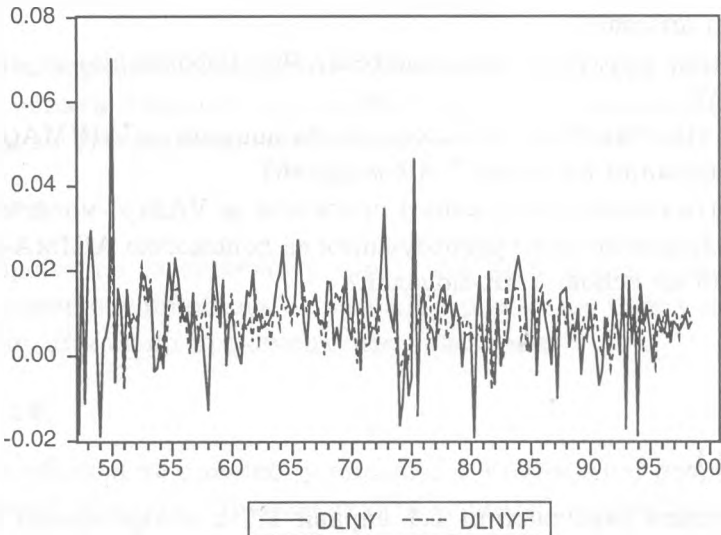


Рис. 2.4. Прогнозні та фактичні зміни у доході

Як можна побачити з наведених графіків, прогнози (fitted) значення обох змінних, обчислені за VAR(15)-моделлю, досить точно відображають фактичні (actual) значення. Отже, прогнозування на основі VAR-моделей є достатньо ефективною процедурою. Звичайно, прогнозувати якість моделей можна перевірити і на основі формальних критеріїв якості. Зробіть це самостійно.

Питання для самоперевірки

1. Поясніть, чому стаціонарність часових рядів є необхідною умовою для розробки VAR(вектор-авторегресійних)-моделей.
2. Запишіть загальний вигляд VAR-моделі p -го порядку для двох змінних. Детально прокоментуйте модель.
3. Опишіть загальну схему побудови VAR-моделей на прикладі двох змінних.
4. За допомогою яких критеріїв можна визначити порядок VAR-моделі?
5. Чим відрізняється структурна форма VAR-моделі від приведеної форми? Поясніть на прикладі найпростішої VAR-моделі першого порядку.
6. Чи завжди можна оцінити параметри структурної форми VAR-моделі? Поясніть ваші доведення.
7. Чи завжди можна оцінити параметри приведеної VAR-моделі? Якщо так, то якими методами?
8. Опишіть загальну схему оцінки VAR-моделей.
9. Прокоментуйте детально особливості прогнозування за допомогою VAR-моделей.
10. Що таке функція імпульсних відгуків? Яку інформацію можна отримати при її аналізі?
11. Що таке декомпозиція дисперсій? Яку інформацію можна отримати при її аналізі?
12. Чим відрізняється моделювання за допомогою ARIMA(p, d, q)-моделей від моделювання на основі VAR-моделей?
13. Назвіть основні властивості прогнозів за VAR(p)-моделями.
14. Чим відрізняється прогнозування за допомогою ARIMA-моделей від прогнозування на основі VAR-моделей?

Вправи

Вправа 2.1.

Використайте дані таблиці 1.1 вправи 1.1 з попереднього розділу про прибуток, що залишається в розпорядженні домогосподарств (Y_2), та витрати домогосподарств на споживання (Y_3) для розробки VAR-моделі двох змінних

на період з 1980.01 по 2000.04. Використайте цю модель для прогнозування значення прибутку, що залишається в розпорядженні домогосподарств (Y_2), та витрат домогосподарств на споживання (Y_3) для чотирьох кварталів 2001 р. та порівняйте їх з фактичними даними, наведеними в таблиці 1.1. Які висновки щодо якості прогнозу ви можете зробити? Які критерії прогнозовної якості моделі можна використати?

Зауваження. Для розв'язання вправ 2.1—2.7 використовуйте пакет E.VIEWS або інший доступний статистичний пакет.

Вправа 2.2.

Для розробленої та оціненої у вправі 2.1 VAR-моделі розрахуйте функцію імпульсних відгуків. Прокоментуйте отримані результати.

Вправа 2.3.

Для розробленої та оціненої у вправі 2.1 VAR-моделі розрахуйте декомпозицію дисперсій. Прокоментуйте отримані результати.

Вправа 2.4.

Для розробленої та оціненої у вправі 2.1 VAR-моделі поясніть, як можна застосувати Шварц-критерій, щоб визначити довжину лага в моделі. Як ви застосуєте цей критерій, щоб визначити відповідну довжину лага у загальному випадку?

Вправа 2.4.

Повторіть вправу 2.1, використовуючи тепер дані про прибуток корпорацій після сплати податків (Y_4) та чисті дивіденди корпорацій (Y_5) з таблиці 1.1 вправі 1.1 попереднього розділу.

Вправа 2.5.

Для розробленої та оціненої у вправі 2.3 VAR-моделі поясніть, на якій формі ви остаточно зупинились — на VAR у рівнях чи VAR у перших різницях — і чому. Детально прокоментуйте пояснення.

Вправа 2.6.

Для розробленої та оціненої у вправі 2.3 VAR-моделі розрахуйте функцію імпульсних відгуків. Прокоментуйте отримані результати.

Вправа 2.7.

Для розробленої та оціненої у вправі 2.3 VAR- моделі розрахуйте декомпозицію дисперсій. Прокоментуйте отримані результати.

Список літератури

1. *Anderson T. W.* The Statistical Analysis of Time Series.— New York : Wiley, 1983.
2. *Anderson T. W., Vahid F.* Testing Multiple Equation System for Common Nonlinear Components // *Journal of Econometrics*, 1998, № 84.— P. 1—36.
3. *Charemza W. W., Deadman D. F.* New Direction in Econometric Practice.— Brookfield : Edward Eglar, 1992.
4. *Enders W.* Applied Econometric Time Series.— New York : Wiley, 1995.
5. *Engle R. F., Granger C. W.* Co-integration and Error Correction: Representation, Estimation and Testing // *Econometrica*, 1987, № 55.— P. 251—276.
6. *Frances Philip Hans.* Time Series Models for Business and Economic Forecasting.— Cambridge : Cambridge University Press, 1999.— 280 p.
7. *Granger C. W.* Some Recent Developments in a Concept of Causality // *Journal of Econometrics*, 1988, № 39.— P. 199—211.
8. *Granger C. W., Andersen A. P.* Introduction to Bilinear Time Series Models.— Goettingen : Vandenhoeck and Ruprecht, 1978.
9. *Granger C. W. J., Newbold P.* Spurious Regressions in Econometrics // *Journal of Econometrics*, 1974, № 2.— P. 111—120.
10. *Granger C. W. J., Terasvirta T.* Modelling Nonlinear Economic Relationships.— Oxford : Oxford University Press, 1993.
11. *Greene William H.* Econometric Analysis.— 4th ed.— New Jersey : Prentice Hall, 2000.— 1004 p.
12. *Gujarati Damodar N.* Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, 1995.— 838 p.
13. *Hamilton James D.* Time Series Analysis.— Princeton, New Jersey : Princeton University Press, 1994.
14. *Hannan E. J.* Multiple Time Series.— New York : Wiley, 1970.
15. *Harvey A. C.* Time Series Models.— Oxford : Philip Allan, 1981.
16. *Hatanaka M.* Time-Series-Based Econometrics. Unit Roots and Cointegration.— Oxford : Oxford University Press, 1996.
17. *Hendry D. F.* Dynamic Econometrics.— Oxford : Oxford University Press, 1995.
18. *Johansen S., Juselius K.* Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money // *Oxford Bulletin of Economics and Statistics*, 1990, № 52.— P. 169—210.

19. *Johansen S.* Statistical Analysis of Cointegration Vector // Journal of Economics Dynamics and Control, 1988, № 12.— P. 231—254.
20. *Johnston J., Dinardo J.* Econometric Methods.— New York : McGraw-Hill, 1997.
21. *Kennedy P. A.* Guide to Econometrics.— Cambridge, Massachusetts : The MIT Press, 1998.— 468 p.
22. *Lukepohl H.* Introduction to Multiple Time Series Analysis.— Berlin : Springer-Verlag, 1991.
23. *Lukepohl H., Reimers H.-E.* Impulse Response Analysis of Cointegrated Systems // Journal of Economic Dynamics and Control, 1992, № 16.— P. 53—78.
24. *Makridakis S., Wheelwright S. C., McGee V. E.* Forecasting: Method and Application.— 2nd ed.— New York : Wiley, 1983.
25. *Meade N., Islam T.* Prediction Intervals for Growth Curve Forecasts // Journal of Forecasting, 1995, № 14.— P. 413—430.
26. *Mills Terence C.* Time Series Techniques for Economists.— Cambridge : Cambridge University Press, 1998.
27. *Ooms M.* Empirical Vector Autoregressive Modeling.— Berlin : Springer-Verlag, 1994.
28. *Osborn D. R., Smith J. P.* The performance of Periodic Autoregressive Models in Forecasting Seasonal UK Consumption // Journal of Business and Economic Statistics, 1989, № 7.— P. 117—127.
29. *Pagan A. R.* The Econometrics of Financial Markets // Journal of Empirical Finance, 1996, № 3.— P. 15—102.
30. *Paulsen J.* Order Determination of Multivariate Autoregressive Time Series with Unit roots // Journal of Time Series Analysis, 1984, № 5.— P. 115—127.
31. *Pindyck R. S., Rubinfeld D. L.* Econometric Models and Economic Forecasts.— 3rd ed.— New York : McGraw-Hill, 1991.
32. *Poirier D. J.* Intermediate Statistics and Econometrics: A Comparative Approach.— Cambridge, Massachusetts : MIT Press, 1995.
33. *Poskitt D. S., Tremayne A. R.* Diagnostic Tests for Multiple Time Series Models // Annals of Statistics, 1982, № 10.— P. 114—120.
34. *Ruud P. A.* An Introduction to Classical Econometric Theory.— Oxford : Oxford University Press, 2000.
35. *Taylor S. J.* Modelling Financial Time Series.— Chichester : Wiley, 1986.
36. *Tiao G. C., Box G. E. P.* Modelling Multiple Time Series with Applications // Journal of the American Statistical Association, 1981, № 76.— P. 802—816.
37. *Tong H.* Non-linear Time Series. A Dynamical System Approach.— Oxford : Oxford University Press, 1990.
38. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика. Практикум з використанням комп'ютера.— К. : Знання, 1998.— 217 с.
39. *Лук'яненко І. Г., Городніченко Ю. О.* Прогнозування податкових

надходжень за допомогою моделей корегування помилки // *Фінанси України*, 2001, № 7.— С. 89—99.

40. *Султан К., Лук'яненко І. Г., Городніченко Ю. О.* Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К. : Видавничий дім «КМ Academia», 2000.— 204 с.

41. *Холден К., Піл Д. А., Томпсон Дж. Л.* Економічне прогнозування. Вступ.— К. : Інформтехніка-ЕМЦ, 1996.

42. *Черняк О. І., Ставицький А. В.* Динамічна економетрика.— К. : КВІЦ, 2000.— 120 с.

РОЗДІЛ 3. ЗАСТОСУВАННЯ МОДЕЛЕЙ КОРЕГУВАННЯ ПОМИЛКИ В ЕМПІРИЧНИХ ДОСЛІДЖЕННЯХ

3.1. ПОНЯТТЯ МОДЕЛІ КОРЕГУВАННЯ ПОМИЛКИ ТА КОІНТЕГРАЦІЇ

Останнім часом в економетричних дослідженнях активно розвивається новий напрям моделювання за допомогою моделей корегування помилки. Фактично ці моделі є структурною формою вектор-авторегресійних (VAR) моделей, або VAR, що включає нестационарні змінні. В попередньому розділі ми зазначали, що оцінювання таких систем є досить складним завданням та вимагає певних додаткових знань, зокрема поняття коінтеграції часових рядів. У цьому розділі ми розглянемо ці поняття детальніше, крім того, розглянемо умови, за якими можна оцінити VAR в структурній формі (з нестационарними змінними), а також умови побудови та оцінки моделей корегування помилки.

Нагадаємо, що коли ми розглядали проблеми моделювання окремих часових рядів за допомогою ARIMA-моделей, то зауважували, що побудова такого роду моделей можлива лише для стаціонарного ряду. Але нестационарний ряд можна звести до стаціонарного шляхом операції різниць! Отже, якщо часовий ряд є нестационарним, то шляхом операції різниць ми можемо звести його до стаціонарного та моделювати перетворений ряд.

Здавалось би, цю ідею можна використати і при побудові систем авторегресійних процесів. Так, якщо ми маємо декілька нестационарних (нагадаємо, що нестационарні ряди називаються також інтегрованими) рядів, то можна утворити з них стаціонарні та далі будувати VAR-моделі в різницях. Але це не зовсім правильно в контексті моделювання взаємозв'язків між коінтегрованими змінними, оскільки коінтеграція змінних дозволяє будувати коректні моделі навіть у випадку їх нестационарності, не перетворюючи часові ряди оператором різниць на стаціонарні. Це дуже важливо для прикладних досліджень, бо, використовуючи оператор різниць, ми втрачаємо цінну «довгострокову» інформацію про динаміку поведінки часового ряду. Тому перетворювати ряди слід лише в разі абсолютної необхідності.

Дійсно, на практиці часто зустрічається ситуація, коли лінійна комбінація нестационарних змінних є стационарною змінною. В такому випадку кажуть, що змінні коінтегровані, або, іншими словами: між ними існує довгостроковий рівноважний зв'язок. Існування рівноважного зв'язку між нестационарними змінними означає, що їхні стохастичні тренди пов'язані, тобто змінні розвиваються в одному напрямку, вони не можуть рухатись незалежно одна від одної. Зв'язок між стохастичними трендами, в свою чергу, свідчить про коінтеграцію змінних. Динамічний шлях коінтегрованих змінних (рис. 3.1) можна уявити як поточне відхилення від довгострокової рівноваги. Отже, в разі наявності коінтеграції між показниками виникає додаткова інформація про відхилення від стану рівноваги, тому просто побудова VAR-моделі в різницях з опущенням складової, що описує таке відхилення, призводить до помилки специфікації. Модель з урахуванням зв'язку між зміною показників та відхиленням від рівноважного стану має спеціальну назву — модель корегування помилки.

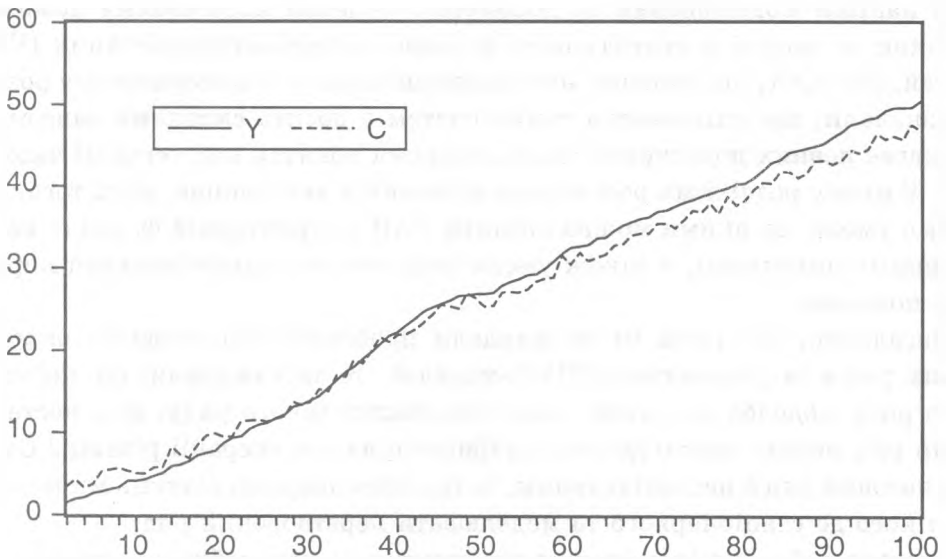


Рис. 3.1. Коінтеграція двох рядів¹: $Y_t = Y_{t-1} + \varepsilon_t$ та $C_t = 1 + 0.9Y_t + u_t$, де ε_t, u_t — білий шум

Таким чином, в основу моделі корегування помилки (error correction model — ECM), або, іншими словами, в основу VAR-моделей з нестационарними змінними, покладена концепція коінтеграції змінних: існування довгострокового зв'язку між рівнями досліджуваних економічних змінних. Тобто

¹ Y можна тлумачити як дохід, а C — як споживання.

припускається існування спільної рівноважної траєкторії руху цих змінних, від якої вони можуть відхилитись в короткострокові періоди, але економічні механізми постійно корегують відповідні відхилення в напрямку відтворення рівноваги, як це можна побачити з рисунку 3.1.

Формалізованою мовою йдеться про те, що навіть якщо ряди нестационарні, але коінтегровані, то лінійна комбінація цих рядів є стаціонарним рядом.

У випадку наявності коінтеграції завжди існує відповідна модель корегування помилки, яка:

- одночасно відображає короткострокові та довгострокові аспекти динаміки досліджуваних показників;
- уникає знаходження хибного (spurious) зв'язку;
- не вимагає попереднього розподілу досліджуваних змінних на екзогенні та ендогенні;
- відповідає основним класичним припущенням економетрики.

Побудова та коректне застосування ЕСМ вимагає попереднього проведення певних логічно пов'язаних між собою етапів. Перш за все, необхідно перевірити ряди на стаціонарність. Якщо вони нестационарні, то треба визначити порядок інтеграції. При однаковому порядку інтеграції можна переходити до перевірки рядів на коінтеграцію. І лише тоді, коли ряди коінтегрують, можна будувати ЕСМ, яка є не чим іншим, як VAR у структурній формі, та оцінювати її невідомі параметри.

Розглянемо кожний етап окремо. Після цього проілюструємо практичне застосування ЕСМ для прогнозування основних статей податків в Україні.

3.2. КОІНТЕГРАЦІЯ НЕСТАЦІОНАРНИХ (ІНТЕГРОВАНИХ) ЗМІННИХ

Коінтеграція є статистичним виразом концепції довгострокового зв'язку між нестационарними (інтегрованими) змінними. Необхідною умовою коінтеграції є однаковий порядок інтегрованості нестационарних змінних. Нехай ми маємо k нестационарних змінних: $\{Y_{1t}\}, \{Y_{2t}\}, \dots, \{Y_{kt}\}$ однакового порядку d ($d > 0$). Ми можемо зробити висновок, що ці змінні коінтегрують, якщо їхня лінійна комбінація є стаціонарним рядом, тобто

$$\gamma_1 Y_{1t} + \gamma_2 Y_{2t} + \dots + \gamma_k Y_{kt} \sim I(0), \tag{3.1}$$

Лінійна комбінація (3.1) відображає довгостроковий рівноважний зв'язок між k змінними: $Y_{1t}, Y_{2t}, \dots, Y_{kt}$.

Наведене нами визначення коінтеграції є найбільш поширеним на практиці, але тлумачить поняття коінтеграції у вузькому розумінні. Інгл (Engle) та Гренджером (Granger) введено загальне поняття коінтеграції [12].

Розглянемо детальніше їхній підхід. Якщо між змінними, що розглядаються, існує довгостроковий зв'язок, то очевидно довгострокова рівновага досягається, коли

$$\gamma_1 Y_{1t} + \gamma_2 Y_{2t} + \dots + \gamma_k Y_{kt} = 0, \quad (3.2)$$

або у матричному вигляді:

$$\gamma Y_t = 0, \quad (3.3)$$

де $\gamma_t = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$, $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$.

Відхилення від довгострокової рівноваги називається помилкою рівноваги, що відповідно дорівнює:

$$e_t = \gamma Y_t. \quad (3.4)$$

Якщо рівновага існує, то необхідно, щоб помилка рівноваги була стаціонарним процесом.

Виходячи з наведених вище понять, Інгл та Гренджер ввели таке загальне поняття коінтеграції. Компоненти вектора $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$ є коінтегрованими порядку d , b : $\sim CI(d, b)$, якщо:

- всі компоненти Y_t мають однаковий порядок інтеграції d ;
- існує вектор коефіцієнтів $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$ такий, що лінійна комбінація $\gamma_1 Y_{1t} + \gamma_2 Y_{2t} + \dots + \gamma_k Y_{kt}$ є інтегрованою величиною порядку $(d - b)$, $b > 0$.

Вектор $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$ називається коінтеграційним вектором (cointegrating vector). Очевидно, що якщо $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$ є коінтеграційним вектором, то для будь-якого додатного значення θ $\theta\gamma = \{\theta\gamma_1, \theta\gamma_2, \dots, \theta\gamma_k\}$ є також коінтеграційним вектором, тому на практиці одна зі змінних використовується для нормалізації коінтеграційного вектора, тобто відповідний їй коефіцієнт повинен дорівнювати одиниці². Для цього коінтеграційний вектор необхідно домножити на $\theta = 1/\gamma_i$, де i відповідає номеру змінної Y_i , за якою ми хочемо нормалізувати коінтеграційний вектор. Наприклад, для того, щоб нормалізувати коінтеграційний вектор відносно змінної Y_1 , необхідно обрати $\theta = 1/\gamma_1$.

² Якщо дослідник має апіорну інформацію щодо структури коінтеграційного рівняння (наприклад, відомо, що певний коефіцієнт дорівнює константі a), тоді нормалізація проводиться відповідно до такої апіорної інформації.

Якщо ми перевіряємо на коінтеграцію k змінних: $Y_t = \{Y_1, Y_2, \dots, Y_k\}^T$, то максимально можна отримати $(k - 1)$ коінтеграційних векторів. Наприклад, якщо ми маємо дві змінні, то повинні отримати один коінтеграційний вектор (звичайно, якщо змінні коінтегрують). Кількість коінтеграційних векторів називається коінтеграційним рангом Y .³

Важливою передумовою коінтеграції нестационарних змінних є однако-вий порядок інтеграції. Звичайно, це не означає, що змінні з однако-вим порядком інтеграції повинні обов'язково коінтегрувати. Якщо змінні мають різний порядок інтеграції, вони не можуть коінтегрувати в принципі. Припустимо, що $Y_1 \sim I(d_1)$, а $Y_2 \sim I(d_2)$, $d_2 > d_1$, тоді можна показати, що їхня лінійна комбінація матиме порядок d_2 , тобто не може бути стаціонарним процесом.

Зауважимо, що у визначенні коінтеграції використовується *лінійна* комбінація нестационарних змінних. Теоретично можна припустити, що існує й нелінійна довгострокова залежність між нестационарними змінними, але сучасний економетричний інструментарій ще не в змозі тестувати нелінійні коінтеграційні залежності.

Надалі ми будемо використовувати поняття коінтеграції у вузькому розумінні, тобто коли лінійна комбінація нестационарних змінних є стаціонарним процесом. Це викликано тим, що на практиці незначна кількість економічних показників має порядок інтеграції більший за одиницю. Звичайно, може виникнути ситуація, коли множина нестационарних змінних другого порядку $I(2)$ може коінтегрувати з порядком $CI(2,1)$, тобто їх лінійна комбінація є нестационарним процесом порядку один $I(1)$, але такі ситуації рідко трапляються у практичних дослідженнях, тому в контексті даної роботи ми їх розглядати не будемо.

3.3. ПОНЯТТЯ СПІЛЬНОГО СТОХАСТИЧНОГО ТРЕНДУ У ВИПАДКУ КОІНТЕГРАЦІЇ

Можливість коінтеграційного зв'язку між нестационарними змінними можна глибше зрозуміти, розглянувши поняття спільного стохастичного тренду в досліджуваних змінних. Поняття спільного стохастичного тренду в коінтегрованих змінних вперше було розглянуто Стоком (Stock) та Вотсоном (Watson) в 1989 р. [42]. Вони показали, що коінтеграція наявна в тому випадку, коли стохастичний тренд однієї змінної можна виразити як лінійну комбінацію трендів інших змінних. Розглянемо їхній підхід детальніше.

³ Далі ми пояснимо, чому можна отримати максимум $(k - 1)$ лінійно незалежних коінтеграційних векторів.

Нехай ми маємо k нестационарних змінних, Y_1, Y_2, \dots, Y_k з однаковим порядком інтеграції. Якщо не враховувати сезонні та циклічні коливання, кожен нестационарну змінну ми можемо представити у такому вигляді:

$$\begin{aligned} Y_{1t} &= \mu_{1t} + \varepsilon_{1t}, \\ &\dots\dots\dots \\ Y_{kt} &= \mu_{kt} + \varepsilon_{kt}, \end{aligned} \tag{3.5}$$

де μ_{it} — випадковий процес (random walk), що є трендом у змінній i в період t , ε_{it} — стаціонарна (нерегулярна) компонента змінної i в період t .

Розглянемо матричний запис формули (3.5):

$$Y_t = \mu_t + \varepsilon_t, \tag{3.6}$$

де вектор $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$; вектор стохастичних трендів $\mu_t = \{\mu_{1t}, \mu_{2t}, \dots, \mu_{kt}\}^T$; вектор нерегулярних компонент $\varepsilon_t = \{\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt}\}^T$.

Якщо один з трендів можна виразити як лінійну комбінацію інших трендів, то знайдеться відмінний від нуля вектор θ такий, що: $\theta_1\mu_{1t} + \theta_2\mu_{2t} + \dots + \theta_k\mu_{kt} = 0$. Помноживши обидві частини (3.6) на вектор θ , отримаємо:

$$\theta Y_t = \theta \mu_t + \theta \varepsilon_t. \tag{3.7}$$

Оскільки $\theta \mu_t = 0$, лінійна комбінація Y є стаціонарним процесом: $\theta Y_t = \theta \varepsilon_t$, тобто змінні Y коінтегрують, а вектор θ є коінтеграційним вектором. Цю логіку легко узагальнити на випадок, коли існує множина лінійних зв'язків між трендами нестационарних змінних Y . В такому випадку існує й відповідна множина коінтеграційних векторів. Дійсно, нагадаємо, що кількість коінтеграційних векторів називається коінтеграційним рангом Y , який класично позначається через r , $0 < r < k$, але не більше ніж $(k - 1)$.⁴ Якщо коінтеграційний ранг дорівнює r , то відповідно матимемо r лінійно незалежних комбінацій трендів, які в матричному вигляді можна записати:

$$\theta \mu_t = 0,$$

⁴ Якщо $r = k$, то кожен з елементів Y є стаціонарним. Можна застосовувати VAR у рівнях. Якщо $r = 0$, ряди є некоінтегрованими.

де, $\theta_t = \begin{bmatrix} \theta_{11} & \dots & \theta_{1k} \\ \dots & \dots & \dots \\ \theta_{r1} & \dots & \theta_{rk} \end{bmatrix}$, матриця порядку $(r \times k)$, $\mu_t = \{\mu_{1t}, \mu_{2t}, \dots, \mu_{kt}\}$.

Отже, один з можливих варіантів розуміння наявності коінтеграції є можливість виразу певних стохастичних трендів через лінійну комбінацію інших трендів системи. Зазначимо, що у випадку двох змінних стохастичні тренди є ідентичними з точністю до скаляра. Наприклад, якщо Y_1 та Y_2 коінтегрують, то відповідно $\theta_1\mu_{1t} + \theta_2\mu_{2t} = 0$. Звідси, тренди відрізняються тільки на скаляр:

$$\mu_{1t} = \frac{\theta_2\mu_{2t}}{\theta_1},$$

тобто випадок коінтеграції можна пояснити наявністю спільних стохастичних трендів у нестационарних змінних.

3.4. МЕХАНІЗМ КОРЕГУВАННЯ ПОМИЛКИ ТА КОІНТЕГРАЦІЯ

Однією з основних властивостей коінтегрованих змінних є те, що їхній розвиток фактично відображає відхилення від довгострокової рівноваги. Зрозуміло, якщо система повертається до рівноважного стану в довгостроковій перспективі, то всі змінні, або принаймні деякі з них, відповідають за повернення до рівноваги. Припустимо, що в певний період часу відхилення від рівноваги збільшується. Відповідно в наступний період часу певні змінні системи повинні змінитись таким чином, щоб зменшити це відхилення. Отже, короткострокова динаміка, або поведінка змінних системи в короткостроковий період часу, залежить від величини відхилення від довгострокової рівноваги. Щоб правильно специфікувати систему з такою поведінкою, ми повинні при моделюванні одночасно розглядати довгострокове відхилення та короткострокову динаміку. Це є основним принципом побудови моделі корегування помилки, в якій зв'язки між змінними системи повинні описуватися з урахуванням механізму корегування відхилення від довгострокової рівноваги. Такий процес називається механізмом корегування помилки. Пояснимо суть цього механізму на прикладі найпростішої моделі корегування помилки.

Припустимо, що змінні Y_1 та Y_2 коінтегрують. Тоді зв'язки між цими двома змінними можна промодельовати за допомогою моделі корегування помилки, яка поєднує короткострокову динаміку з довгостроковим рівноважним зв'язком та у випадку двох змінних має такий формалізований вигляд:

$$\Delta Y_{1t}^* = \alpha_{10} + \sum_{i=1}^k \alpha_{11}(i)\Delta Y_{1,t-i} + \sum_{i=0}^k \alpha_{12}(i)\Delta Y_{2,t-i} + \lambda_1 \hat{u}_{1,t-1} + \varepsilon_{1t}, \quad (3.8)$$

$$\Delta Y_{2t} = \alpha_{20} + \sum_{i=0}^k \alpha_{21}(i) \Delta Y_{1,t-i} + \sum_{i=1}^k \alpha_{22}(i) \Delta Y_{2,t-i} + \lambda_2 \hat{u}_{2,t-1} + \varepsilon_{2t}, \quad (3.9)$$

де $\hat{u}_{1,t-1} = Y_{1,t-1} - \gamma_0 - \gamma_1 Y_{2,t-1} - I(0)$ — рівняння довгострокової рівноваги (коінтеграційне рівняння), нормоване по першій змінній; $\hat{u}_{2,t-1} = Y_{2,t-1} - \gamma'_0 - \gamma'_1 Y_{1,t-1} - I(0)$ — рівняння довгострокової рівноваги (коінтеграційне рівняння), нормоване по другій змінній. Відмітимо, що $\hat{u}_{1,t-1}$ та $\hat{u}_{2,t-1}$ є відхиленням від довгострокової рівноваги. Довгострокова рівновага досягається, якщо $Y_{1,t-1} = \gamma_0 + \gamma_1 Y_{2,t-1}$. Збурення ε_{1t} та ε_{2t} є білим шумом, при цьому вони можуть корелювати між собою.

Для того, щоб модель (3.8)—(3.9) була динамічно стабільною, необхідно, щоб $0 \leq \lambda_1 < 1$, $0 \leq \lambda_2 < 1$.⁵ Зауважимо, що якщо коінтеграційні рівняння моделі нормовані по різних змінних, то знаки цих коефіцієнтів повинні бути від'ємними. Коефіцієнти λ_1 та λ_2 в моделі (3.8)—(3.9) називаються швидкістю пристосування. При цьому коефіцієнт λ_1 є чутливістю (відгуком, responsiveness) ΔY_1 на відхилення від рівноваги, а коефіцієнт λ_2 є чутливістю ΔY_2 на відхилення від рівноваги. Вони показують, на скільки відсотків відхилення від рівноваги миттєво корегуються кожною змінною. Відповідно решта відхилення корегується протягом наступних періодів. Наприклад, якщо $\lambda_1 = -0.3$, це означає, що 30 % відхилення від рівноваги корегується миттєво змінною Y_1 . Чим більшим є значення λ_1 , тим більшою є реакція Y_1 на попереднє відхилення від рівноваги. Якщо значення параметра λ_1 є незначним (тобто близьким до нуля), то змінна Y_1 фактично не відповідає за повернення до рівноваги. Можливі чотири випадки:

Випадок 1	Випадок 2	Випадок 3	Випадок 4
$0 \leq \lambda_1 < 1$, $0 \leq \lambda_2 < 1$. Обидві змінні моделі (3.8)—(3.9) відповідають за повернення до рівноваги. Y_1 та Y_2 є ендогенними змінними відхилення.	$0 \leq \lambda_1 < 1$, $\lambda_2 = 0$. Тільки Y_1 відповідає за повернення до рівноваги. Y_1 є ендогенною змінною, а Y_2 є слабо екзогенною.	$\lambda_1 = 0$, $0 \leq \lambda_2 < 1$. Тільки Y_2 відповідає за відхилення від рівноваги. Y_2 є ендогенною змінною, а Y_1 є слабо екзогенною.	$\lambda_1 = 0$, $\lambda_2 = 0$. Y_1 та Y_2 не коінтегрують, немає довгострокового зв'язку між цими змінними.

Зауважимо, що принаймні один з коефіцієнтів λ повинен відрізнятись від нуля, бо інакше довгострокового рівноважного зв'язку між двома змінними не існує, а отже, й модель не може бути коінтеграційною або моделлю корегування помилки.

⁵ Якщо абсолютна величина коефіцієнтів більша за одиницю, то, як відомо з теорії різницевого рівняння, система «вибухає». Якщо коефіцієнт від'ємний та більший за мінус одиницю, система просто не повертається у рівноважний стан. У принципі, такий випадок вказує на існування багатьох рівноважних станів (multiple equilibria).

Відмітимо також, що модель корегування помилки завжди є статистично коректною, тому що виконуються основні класичні припущення, тобто всі змінні моделі завжди є стаціонарними. Наприклад в моделі (3.8)—(3.9): $\Delta Y_{1t} \sim I(0)$, $\Delta Y_{2t} \sim I(0)$, $\hat{u}_{1,t-1} \sim I(0)$, $\hat{u}_{2,t-1} \sim I(0)$, ε_{1t} та $\varepsilon_{2t} \sim I(0)$ за припущенням.

Аналіз найпростішої моделі корегування помилки показує, що вона відрізняється від звичайної VAR-моделі в перших різницях елементом, що відображає механізм корегування помилки, тобто наявністю коінтеграційного рівняння. Якщо одна зі змінних системи не впливає на зміну іншої, то швидкість пристосування й інші коефіцієнти при ній мають дорівнювати нулю. Це означає відсутність причинного зв'язку в цьому напрямку.

Модель корегування помилки (3.8)—(3.9) легко узагальнити для випадку k змінних. У формалізованому матричному вигляді модель корегування помилки можна записати так:

$$\Delta Y_t = \pi_0 + \pi Y_{t-1} + \pi_1 \Delta Y_{t-1} + \pi_2 \Delta Y_{t-2} + \dots + \pi_p \Delta Y_{t-p} + \varepsilon_t, \quad (3.10)$$

де $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$ — вектор змінних порядку $(k \times 1)$; π_0 — вектор $(k \times 1)$ перетинів з елементами π_{i0} ; π_i , $i = 1, 2, \dots, k$ — матриці $(k \times k)$ коефіцієнтів з елементами $\pi_{jk}(i)$; π — матриця з елементами π_{jk} , серед яких принаймні один чи більше $\pi_{jk} \neq 0$; вектор $(k \times 1)$ нерегулярних компонентів $\varepsilon_t = \{\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt}\}$, які є білим шумом, але такі, що ε_{it} може корелювати з ε_{jt} .

Для спрощення припустимо, що всі змінні $Y_{1t}, Y_{2t}, \dots, Y_{kt}$ мають порядок інтеграції одиниця $(I(1))$. Крім того, в коінтеграційних рівняннях припустимо, що всі перетини дорівнюють нулю. Якщо зв'язки між змінними можна промодельювати у вигляді моделі корегування помилки (3.10), то лінійна комбінація цих нестаціонарних змінних повинна бути стаціонарним процесом. Розв'яжемо (3.10) відносно πY_{t-1} і отримаємо:

$$\pi Y_{t-1} = \Delta Y_t - \pi_0 + \pi_1 Y_{t-1} - \pi_2 \Delta Y_{t-2} - \dots - \pi_p \Delta Y_{t-p} - \varepsilon_t. \quad (3.11)$$

Оскільки кожна складова лівої частини (3.11) є стаціонарним процесом, то πY_{t-1} також має бути стаціонарним процесом. Кожний рядок матриці π є коінтеграційним вектором Y_t . Наприклад, перший рядок матриці π може бути записано у вигляді:

$$\pi_{11} Y_{1,t-1} + \pi_{12} Y_{2,t-2} + \dots + \pi_{1k} Y_{k,t-1}. \quad (3.12)$$

Відповідно вектор $(\pi_{11}, \pi_{12}, \dots, \pi_{1k})$ є коінтеграційним вектором для Y_t . Отже, матриця π відіграє значну роль в моделі корегування помилки (3.11), тому наведемо її основні властивості:

— якщо всі елементи матриці π дорівнюють нулю, тобто механізм коре-

гування помилки відсутній, (3.11) перетворюється на звичайну VAR-модель в перших різницях;

— якщо хоча б один елемент матриці π , $\pi_{jk} \neq 0$, то ΔY_t реагує на відхилення від довгострокової рівноваги в попередньому періоді. В такому випадку моделювання зв'язків між змінними Y_{1t} , Y_{2t} , ..., Y_{kt} у вигляді VAR-моделі в перших різницях є некоректним, оскільки модель матиме помилку специфікації, зокрема невключення статистично важливої змінної⁶.

3.5. ЗВ'ЯЗОК МІЖ РАНГОМ МАТРИЦІ, КОІНТЕГРАЦІЙНИМ РІВНЯнням ТА МОДЕЛЛЮ КОРЕГУВАННЯ ПОМИЛКИ

Ранг матриці π в моделі (3.11) широко використовується на практиці для тестування коінтеграції між змінними. Якщо ранг матриці π менший за k (тобто ранг є неповним), існує коінтеграційне рівняння для векторів Y_1 , Y_2 , ..., Y_k , а відповідно й модель корегування помилки.

Якщо ранг матриці дорівнює нулю, то коінтеграція між змінними відсутня, ми не можемо використовувати модель корегування помилки для опису зв'язку між змінними. Потрібно використовувати VAR у різницях. Чому в різницях? Нагадаємо, що в такому випадку змінні будуть нестационарними і для зведення їх до стаціонарних використовуються операції різниць, а отже, й VAR-модель будується на основі перетворених змінних.

Якщо матриця π має повний ранг⁷, це свідчить про те, що змінні є стаціонарними, а відповідно також, що між ними не може бути коінтеграційного зв'язку. В такому випадку також не існує моделі корегування помилки, а VAR, відповідно, будується у рівнях, оскільки змінні стаціонарні.

Проілюструємо наведені вище твердження для випадку k змінних⁸. Дослідимо такий можливий зв'язок між змінними:

$$Y_t = A_1 Y_{t-1} + \varepsilon_t, \quad (3.13)$$

де $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$ — вектор змінних порядку $(k \times 1)$; A_1 — матриця $(k \times k)$ коефіцієнтів моделі; $\varepsilon_t = \{\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt}\}^T$ — вектор $(k \times 1)$ збурень, які є білим шумом, але такі, що ε_{it} може корелювати з ε_{jt} .

⁶ Нагадаємо, що невключення статистично та/або теоретично важливої змінної призводить до зміщення оцінок.

⁷ Відповідно детермінант матриці не дорівнює нулю.

⁸ Нагадаємо, що в такому випадку може існувати більше ніж одне коінтеграційне рівняння.

Відніmemo Y_{t-1} від обох частин (3.13) і отримаємо:

$$\Delta Y_t = -(I - A_1)Y_{t-1} + \varepsilon_t = \pi Y_{t-1} + \varepsilon_t, \quad (3.14)$$

де $\pi = -(I - A_1)$ є $(k \times k)$ матрицею, відповідно I є одиничною матрицею також порядку k .

Легко побачити, що модель (3.14) є спрощеним випадком (3.11) в припущенні, що перетини відсутні, а всі матриці $\pi_i = 0$.

Тепер легко прослідкувати зв'язок між рангом матриці π та існуванням коінтеграційного рівняння, а отже, й моделі корегування помилки. Якщо ранг матриці π дорівнює нулю, кожний елемент цієї матриці повинен дорівнювати нулю, отже, (3.14) перетворюється на

$$\Delta Y_t = \varepsilon_t. \quad (3.15)$$

Формула (3.15) є не чим іншим, як VAR у перших різницях, при цьому $\Delta Y_t = \varepsilon_t \sim I(0)$, а отже, всі змінні $Y_{it} = Y_{i,t-1} + \varepsilon_{it}$ (3.13) є нестационарними процесами з одиничним коренем. Якщо всі змінні є нестационарними з одиничним коренем, то жодна з їх лінійних комбінацій не може бути стаціонарним процесом, а, відповідно, моделі корегування помилки не існує.

Тепер припустимо, що матриця π має повний ранг. Розв'язок (3.14) в довгостроковій перспективі (при досягненні довгострокової рівноваги) знаходиться на основі k незалежних рівнянь:

$$\begin{aligned} \pi_{11}Y_{1t} + \pi_{12}Y_{2t} + \dots + \pi_{1k}Y_{kt} &= 0 \\ \pi_{21}Y_{1t} + \pi_{22}Y_{2t} + \dots + \pi_{2k}Y_{kt} &= 0 \\ \dots & \dots \dots \dots \\ \pi_{k1}Y_{1t} + \pi_{k2}Y_{2t} + \dots + \pi_{kk}Y_{kt} &= 0. \end{aligned} \quad (3.16)$$

(3.16) є системою k довгострокових рівнянь з k змінними. Система дає розв'язок для довгострокової складової (3.14). Оскільки матриця коефіцієнтів у (3.16) має повний ранг, система (3.16) має єдиний розв'язок: рівноважне значення для будь-якого Y_{it} дорівнює нулю! Отже, всі змінні повинні бути стаціонарними. У такому випадку не можна говорити про коінтеграційний зв'язок, а отже, правильно специфікованою моделлю для опису їх зв'язку є відповідно VAR у рівнях.

Тепер розглянемо випадок, коли ранг матриці π не є повним і дорівнює r , $r > 0$. Відповідно існує r коінтеграційних рівнянь (r коінтеграційних векторів). Якщо $r = 1$, то ми маємо єдиний коінтеграційний вектор, який задається будь-яким рядком матриці π . Кожна послідовність $\{Y_{it}\}$ може бути

записана з урахуванням механізму корегування помилки, наприклад для першої змінної маємо:

$$\Delta Y_{1t} = \pi_{11}Y_{1,t-1} + \pi_{12}Y_{2,t-1} + \dots + \pi_{1k}Y_{k,t-1} + \varepsilon_{1t}. \quad (3.17)$$

Якщо ми нормалізуємо коінтеграційний вектор в (3.17) по першій змінній $Y_{1,t-1}$, позначивши $\pi_{11} = \lambda_1$ та $\gamma_{ij} = \pi_{ij}/\pi_{11}$, то відповідно отримаємо:

$$\Delta Y_{1t} = \lambda_1 (Y_{1,t-1} + \gamma_{12}Y_{2,t-1} + \dots + \gamma_{1k}Y_{k,t-1}) + \varepsilon_{1t}. \quad (3.18)$$

В довгостроковій рівновазі $\{Y_{it}\}$ повинні задовольняти рівнянню:

$$Y_{1t} + \gamma_{12}Y_{2t} + \dots + \gamma_{1k}Y_{kt} = 0. \quad (3.19)$$

В даному випадку нормалізованим коінтеграційним вектором є вектор $\{1, \gamma_{12}, \gamma_{13}, \dots, \gamma_{1k}\}$, а швидкість пристосування дорівнює λ_1 .

Аналогічно у випадку r коінтеграційних векторів довгострокові значення змінних повинні задовольняти системі r таких рівнянь:

$$\begin{aligned} \pi_{11}Y_{1t} + \pi_{12}Y_{2t} + \dots + \pi_{1k}Y_{kt} &= 0 \\ \pi_{21}Y_{1t} + \pi_{22}Y_{2t} + \dots + \pi_{2k}Y_{kt} &= 0 \\ \dots & \dots \dots \dots \\ \pi_{r1}Y_{1t} + \pi_{r2}Y_{2t} + \dots + \pi_{rk}Y_{kt} &= 0, \end{aligned} \quad (3.20)$$

які можуть бути нормалізовані.

Отже, як видно з наведених вище прикладів, аналіз рангу матриці дозволяє зробити висновок не тільки про наявність коінтеграційних рівнянь та їх кількість, але дає змогу коректно специфікувати відповідні моделі та форму змінних (різниця або рівні), що в них входять. Тепер розглянемо детальніше формальні тести, що дозволяють зробити висновок про наявність коінтеграційного зв'язку між змінними.

3.6. ПЕРЕВІРКА ЧАСОВИХ РЯДІВ НА КОІНТЕГРАЦІЮ

Коінтеграція є статистичним виразом концепції довгострокового зв'язку між змінними. У випадку коінтеграції завжди існує відповідна модель корегування помилки, що відображає довгострокові та короткострокові аспекти динаміки досліджуваних змінних [8, 12, 16]. Отже, наступним важливим етапом на шляху побудови ЕСМ є перевірка часових рядів на коінтеграцію. Існує декілька принципово різних підходів до тестування часових рядів

на коінтеграцію, а саме підхід Інгла—Гренджера (Engle—Granger) [12,16] та підхід Йохансена (Johansen) [22, 29, 31], або Стока—Уотсона (Stock—Watson) [41]. Підхід Інгла—Гренджера базується на перевірці стаціонарності помилок рівняння довгострокової рівноваги, підхід Йохансена та Стока—Уотсона — на визначенні рангу матриці π . Розглянемо особливості кожного.

3.6.1. Тест перевірки рядів на коінтеграцію Інгла—Гренджера (Engle—Granger)

Нагадаємо, що необхідною передумовою коінтеграції є однаковий порядок інтеграції нестационарних часових рядів. Іншими словами, часові ряди повинні бути, по-перше, нестационарними, а по-друге — мати однаковий порядок інтеграції. Лише при наявності даної передумови можна ставити питання про можливу їх коінтеграцію. Для ілюстрації тесту Інгла—Гренджера припустимо, що необхідна передумова коінтеграції виконується. Крім того, для спрощення припустимо, що ми маємо лише дві змінні Y_1 та Y_2 . У випадку двох змінних ми можемо зробити висновок, що $Y_{1t} \sim I(d)$ та $Y_{2t} \sim I(d)$, $d > 0$, коінтегрують, якщо їх лінійна комбінація є стаціонарним рядом:

$$Y_{1t} - \gamma_0 - \gamma_1 Y_{2t} \sim I(0), \quad (3.21)$$

$$Y_{2t} - \gamma_0^* - \gamma_1^* Y_{1t} \sim I(0). \quad (3.22)$$

Комбінації (3.21) та (3.22) відображають довгостроковий рівноважний зв'язок між Y_1 та Y_2 .

Перевірка рядів на коінтеграцію за методологією Інгла—Гренджера практично ідентична тесту Дікі—Фуллера або розширеному тесту Дікі—Фуллера для перевірки рядів на стаціонарність (наявність одиничного кореня). Єдина відмінність полягає в тому, що ми застосовуємо тест для перевірки стаціонарності залишків рівнянь довгострокової рівноваги (3.21) або (3.22).

Тому першим етапом перевірки рядів на коінтеграцію за тестом Інгла—Гренджера є оцінювання рівняння довгострокової рівноваги, наприклад нормалізованого по першій змінній (3.21):

$$Y_{1t} = \gamma_0 + \gamma_1 Y_{2t} + \varepsilon_t. \quad (3.23)$$

Якщо ці змінні коінтегровані, то оцінювання (3.23) методом найменших квадратів дає надконсистентні (super-consistent) оцінки параметрів γ_0 та γ_1 . Сток довів [41], що в даному випадку МНК-оцінки γ_0 та γ_1 збігаються до дійсних значень швидше, ніж МНК-оцінки γ_0 та γ_1 у випадку стаціонарності.

Після оцінювання рівняння (3.23) розрахуємо залишки, тобто отримаємо відповідний часовий ряд залишків: $\{\hat{\varepsilon}_t\}$. Якщо залишки є стаціонарним рядом, то можна зробити висновок про коінтеграцію Y_1 та Y_2 . Для перевірки залишків на стаціонарність застосуємо тест Дікі—Фуллера, який в такому контексті називається тестом Інгла—Гренджера⁹.

Базовим рівнянням для тесту Інгла—Гренджера є регресійне рівняння:

$$\Delta \hat{\varepsilon}_t = b \hat{\varepsilon}_{t-1} + \sum_{i=1}^k c_i \Delta \hat{\varepsilon}_{t-1} + u_t, \quad (3.24)$$

де $\Delta \hat{\varepsilon}_t = \hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}$, u_t — випадкова величина, b , c_i — невідомі коефіцієнти регресії, $\hat{\varepsilon}_t$ — залишки рівняння довгострокової рівноваги (3.23). На основі базової моделі (3.24) тестуються такі гіпотези:

H_0 : $b = 0$, або часові ряди Y_1 та Y_2 не коінтегрують: залишки $\hat{\varepsilon}_t$ є нестаціонарними,

H_1 : $b < 0$, або часові ряди коінтегрують: залишки є стаціонарними: $\hat{\varepsilon}_t \sim I(0)$.

Нульова гіпотеза H_0 відкидається, якщо отриманий коефіцієнт $b < 0$ та розрахована τ -статистика за абсолютною величиною більша за абсолютну величину критичного значення τ -статистики МакКіннона для тестування на коінтеграцію при заданому рівні значимості α . Якщо нульова гіпотеза відкидається, то ми вважаємо, що Y_1 та Y_2 коінтегрують, тобто $\hat{\varepsilon}_t \sim I(0)$. Аналогічний тест можна провести і для залишків рівняння довгострокової рівноваги, нормалізованого по другій змінній (3.22). Якщо і для них нульова гіпотеза відкидається, то ми робимо висновок про те, що Y_1 та Y_2 коінтегрують в обох випадках.

Після того як зроблено висновок про коінтеграцію Y_1 та Y_2 , ми можемо оцінити модель корегування помилки, використовуючи залишки з рівняння довгострокової рівноваги:

$$\Delta Y_{1t} = a_{10} - \lambda_1(Y_{1,t-1} - \gamma_1 Y_{2,t-1}) + \sum_{i=1}^k a_{11}(i) \Delta Y_{1,t-i} + \sum_{i=0}^k a_{12}(i) \Delta Y_{2,t-i} + u_{1t}, \quad (3.25)$$

$$\Delta Y_{2t} = a_{20} - \lambda_2(Y_{1,t-1} - \gamma_1 Y_{2,t-1}) + \sum_{i=0}^k a_{21}(i) \Delta Y_{1,t-i} + \sum_{i=1}^k a_{22}(i) \Delta Y_{2,t-i} + u_{2t}, \quad (3.26)$$

де γ_1 — параметр вектора коінтеграції, визначеного (3.23); u_{1t} та u_{2t} — збурення, що є білим шумом, але які можуть корелювати між собою; a_{10} , a_{20} , λ_1 , λ_2 , $a_{10}(i)$, $a_{12}(i)$, $a_{21}(i)$, $a_{22}(i)$ — параметри системи.

⁹ Отже, тест Інгла—Гренджера для перевірки рядів на коінтеграцію є фактично тестом Дікі—Фуллера для залишків.

Інгл та Гренджер запропонували використовувати залишки $\bar{\varepsilon}_{t-1}$, отримані з рівняння довгострокової рівноваги (3.23) в період $(t-1)$, як інструмент для виразу $(Y_{1,t-1} - \gamma_1 Y_{1,t-1})$ в моделях (3.25)—(3.26). Таким чином, використовуючи збережені залишки, ми фактично можемо оцінити таку модель корегування помилки:

$$\Delta Y_{1t} = a_{10} - \lambda_1 \bar{\varepsilon}_{t-1} + \sum_{i=1}^k a_{11}(i) \Delta Y_{1,t-i} + \sum_{i=0}^k a_{12}(i) \Delta Y_{2,t-i} + u_{1t}, \quad (3.27)$$

$$\Delta Y_{2t} = a_{20} - \lambda_2 \bar{\varepsilon}_{t-1} + \sum_{i=0}^k a_{21}(i) \Delta Y_{1,t-i} + \sum_{i=1}^k a_{22}(i) \Delta Y_{2,t-i} + u_{2t}. \quad (3.28)$$

Модель (3.27)—(3.28) є фактично VAR-моделлю в перших різницях, якщо не брати до уваги складову корегування помилки ($\bar{\varepsilon}_{t-1}$). Оцінювання такої моделі аналогічне оцінюванню VAR-моделей, яке ми розглядали у попередньому розділі. Нагадаємо його особливості:

— оцінювання моделі (3.27)—(3.28) методом найменших квадратів дає ефективні оцінки, якщо в кожне рівняння входять однакові змінні з однаковою кількістю лагів;

— оскільки всі складові моделі (3.27)—(3.28) є стаціонарними величинами, застосування традиційних статистичних тестів (t -статистики; F -статистики та інших) залишається коректним.

Після того як модель (3.27)—(3.28) оцінено, необхідним є тест її на адекватність. Існує декілька специфічних процедур, які допомагають зробити висновок про коректність оціненої моделі корегування помилки. На особливу увагу заслуговують коефіцієнти швидкості пристосування λ_1 та λ_2 . Наприклад, якщо λ_1 дорівнює нулю, це означає, що зміни в Y_1 не є чутливими до відхилення від довгострокової рівноваги. Якщо, крім того, і всі коефіцієнти $a_{12}(i)$ дорівнюють нулю, це означає, що $\{\Delta Y_{2t}\}$ не є причиною зміни $\{\Delta Y_{1t}\}$ по Гренджеру (Granger cause).

Якщо змінні коінтегрують, то хоча б один з коефіцієнтів λ_1 та λ_2 повинен статистично значимо відрізнитись від нуля. Якщо обидва значення λ_1 та λ_2 дорівнюють нулю, то модель (3.27)—(3.28) не є моделлю корегування помилки, а є просто VAR у перших різницях. Значення λ_1 та λ_2 за абсолютною величиною повинні бути меншими за одиницю, саме ця умова гарантує, що $\{\Delta Y_{1t}\}$ та $\{\Delta Y_{2t}\}$ збігаються до довгострокового рівноважного зв'язку.

Як і у випадку VAR-моделей, для моделей корегування помилки можна застосовувати імпульсний аналіз та аналіз декомпозиції дисперсій.

Звернемо увагу, що у випадку моделей корегування помилки потрібно бути дуже обережним з висновками щодо серійної кореляції. Як правило, її можна уникнути, збільшивши кількість лагів у моделі.

У спрощеному вигляді тестування на коінтеграцію за методологією Інгла—Гренджера можна представити у вигляді блок-схеми:

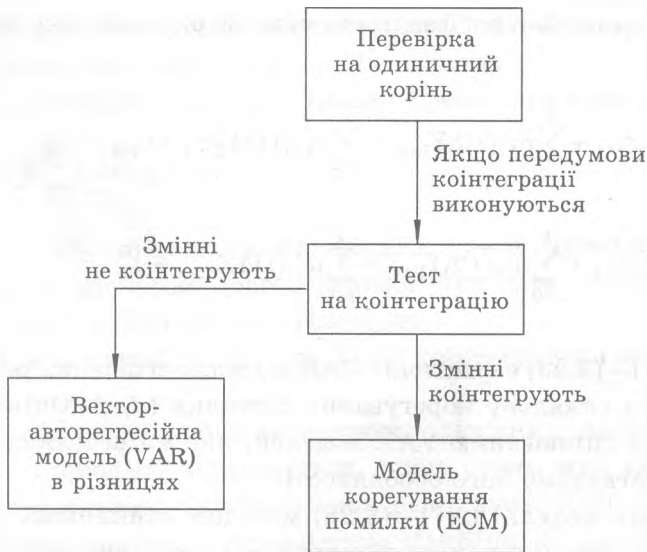


Рис. 3.2. Узагальнена блок-схема перевірки змінних на коінтеграцію за методологією Інгла—Гренджера

Методологія Інгла—Гренджера є простою, але має істотні недоліки. Насамперед це стосується можливого обрахунку залишків рівняння довгострокової рівноваги. Наприклад, у випадку двох змінних для тестування рядів на коінтеграцію ми можемо використовувати залишки від двох коінтеграційних рівнянь, нормованих відповідно по кожній зі змінних:

$$Y_{1t} = \gamma_0 - \gamma_1 Y_{2t} + \varepsilon_{1t}, \quad (3.29)$$

$$Y_{2t} = \gamma_0^* - \gamma_1^* Y_{1t} + \varepsilon_{2t}. \quad (3.30)$$

В теорії доведено, що якщо вибірка нескінченно зростає, то тест на стаціонарність послідовності $\{\varepsilon_{1t}\}$ стає еквівалентним тесту на стаціонарність послідовності $\{\varepsilon_{2t}\}$. На жаль, на практиці ми не часто оперуємо з невеликими вибірками, тому можлива ситуація, коли висновок про наявність коінтеграції залежить від вибору рівняння (тобто від вибору між (3.29) та (3.30)). Це, звичайно, небажана ситуація, тому що тест повинен бути інваріантним до того, яка зі змінних обрана для нормалізації довгострокового рівняння. Проблема ще більш ускладнюється у випадку перевірки на коінтеграцію трьох та більше змінних. У таких випадках ми повинні мати більше ніж одне коінтеграційне рівняння,

а методологія Інгла—Гренджера не включає систематичної процедури для окремого оцінювання багатьох коінтеграційних рівнянь.

Інший недолік методології Інгла—Гренджера полягає в тому, що вона базується на покроковій процедурі. На першому кроці ми оцінюємо залишки з рівняння довгострокової рівноваги $\{\varepsilon_t\}$, а на другому — саму модель корегування помилки. Таким чином, будь-яка помилка, зроблена на першому етапі, автоматично поширюється на другий етап. Вільше того, внаслідок застосування покрокової процедури зменшується ефективність оцінок параметрів.

Методологія Йохансена, Стока та Уотсона багато в чому дозволяє уникнути перелічених вище недоліків. Вона принаймні є однокроковою та дозволяє тестувати наявність багатьох коінтеграційних рівнянь, проводити їх оцінювання. Розглянемо цю методологію детальніше.

3.6.2. Тест перевірки часових рядів на коінтеграцію Йохансена

Методологія Йохансена (Johansen) так само, як Стока і Уотсона, базується на зв'язку між рангом матриці π та її характеристичними коренями [28-32].

Нагадаємо, що в найпростішому випадку однієї змінної $\{Y_t\}$ можна зробити висновок про її стаціонарність на основі тесту Дікі—Фуллера відповідно до значення параметра $(\alpha - 1)$, визначеного з рівняння:

$$Y_t = \alpha_1 Y_{t-1} + \varepsilon_t, \text{ або } \Delta Y_t = (\alpha_1 - 1)Y_{t-1} + \varepsilon_t. \quad (3.31)$$

Якщо параметр $(\alpha - 1) = 0$, ми можемо зробити висновок, що процес $\{Y_t\}$ має одиничний корінь, тобто є нестаціонарним, якщо $(\alpha - 1) \neq 0$, то ми можемо зробити висновок про стаціонарність $\{Y_t\}$ (ми виключаємо з розгляду можливість вибухового процесу, оскільки емпірично такі процеси трапляються дуже рідко).

Концептуально методологія Йохансена фактично узагальнює (3.31) на випадок k змінних. Дослідимо такий найпростіший можливий зв'язок між змінними:

$$Y_t = A_1 Y_{t-1} + \varepsilon_t, \quad (3.32)$$

де $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$ — вектор змінних порядку $(k \times 1)$; A_1 — матриця $(k \times k)$ коефіцієнтів моделі; $\varepsilon_t = \{\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{kt}\}^T$ — вектор $(k \times 1)$ збурень, які є білим шумом, але такі, що ε_{it} може корелювати з ε_{jt} .

Віднявши від обох частин (3.32), отримаємо:

$$\Delta Y_t = -(I - A_1) Y_{t-1} + \varepsilon_t = \pi Y_{t-1} + \varepsilon_t, \quad (3.33)$$

де $\pi = -(I - A_1) = A_1 - I$ є $(k \times k)$ матрицею, відповідно I є одиничною матрицею також розмірності $(k \times k)$.

Як було зазначено в попередньому підрозділі, ранг матриці $\pi = -(I - A_1)$ дорівнює кількості коінтеграційних рівнянь. По аналогії з (3.31), якщо всі значення матриці $\pi = -(I - A_1)$ дорівнюють нулю, то і ранг матриці π дорівнює нулю. В цьому випадку всі змінні $\{Y_{it}\}$ є процесами з одиничним коренем. Це означає, що немає такої лінійної комбінації процесів $\{Y_{it}\}$, яка була б стаціонарним процесом (ми наводимо це твердження без доведення).

Якщо відкинути характеристичні корені, значення яких більше за одиницю, то, якщо ранг матриці π дорівнює k , система (3.33) перетворюється на систему, що збігається до рівнянь різниць, тобто в цьому випадку всі змінні є стаціонарними.

Звичайно, (3.33) представляє собою спрощений випадок, цю систему можна ускладнити, додавши перетини:

$$\Delta Y_t = A_0 + \pi Y_{t-1} + \varepsilon_t, \quad (3.34)$$

де A_0 — вектор констант $\{a_{01}, a_{02}, \dots, a_{0k}\}^T$ розмірності $(k \times 1)$.

Включення константи дозволяє описати наявність лінійного часового тренду в даних. Константи (зсув) можна включати, якщо змінні мають тенденцію до спаду або зростання. В такому випадку ранг матриці π можна розглядати як кількість коінтеграційних рівнянь в очищених від тренду (detrended) даних. В довгостроковій рівновазі $\pi Y_{t-1} = 0$, тобто кожна послідовність $\{\Delta Y_{it}\}$ має очікуване значення a_{i0} . Агрегування всіх цих змін протягом t дає детермінований вираз $a_{i0} \cdot t$. Включення констант в моделі (3.34) означає, що змінні мають не лише стохастичний тренд, а й однаковий детермінований тренд¹⁰. Проблема констант в моделі (3.34) є важливою, тому що відповідна маніпуляція елементами матриці дозволяє включити константу в коінтеграційне рівняння, не включаючи детермінований тренд в систему.

Одним з можливих шляхів включення константи в коінтеграційні рівняння є певні обмеження на значення a_{i0} . Наприклад, якщо ранг матриці π дорівнює одиниці, рядки матриці мають відрізнятись один від одного лише на скаляр, тому (3.34) можна переписати для кожної змінної $\{\Delta Y_{it}\}$ як:

$$\Delta Y_{1t} = \pi_{11} Y_{1,t-1} + \pi_{12} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1} + a_{10} + \varepsilon_{1t},$$

$$\Delta Y_{2t} = \theta_2 (\pi_{11} Y_{1,t-1} + \pi_{12} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1}) + a_{20} + \varepsilon_{2t},$$

.....

¹⁰ Однаковий детермінований тренд не означає, що константи є рівними.

$$\Delta Y_{kt} = \theta_k (\pi_{11} Y_{1,t-1} + \pi_{12} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1}) + a_{k0} + \varepsilon_{kt}. \quad (3.35)$$

де скаляр θ_i є таким, що $\theta_i \times \pi_{1j} = \pi_{ij}$.

Якщо кожна константа може бути представлена у вигляді $a_{i0} = \theta_i a_{10}$, то це означає, що всі змінні $\{\Delta Y_{it}\}$ можуть бути представлені з константою в коінтеграційному рівнянні

$$\begin{aligned} \Delta Y_{1t} &= (\pi_{11} Y_{1,t-1} + \pi_{21} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1} + a_{10}) + \varepsilon_{1t}, \\ \Delta Y_{2t} &= \theta_2 (\pi_{11} Y_{1,t-1} + \pi_{21} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1} + a_{10}) + \varepsilon_{2t}, \\ &\dots\dots\dots \\ \Delta Y_{kt} &= \theta_k (\pi_{11} Y_{1,t-1} + \pi_{21} Y_{2,t-1} + \dots + \pi_{1k} Y_{k,t-1} + a_{10}) + \varepsilon_{kt}. \end{aligned} \quad (3.36)$$

Систему (3.36) можна переписати у матричному вигляді:

$$\Delta Y_t = \pi^* Y_{t-1}^* + \varepsilon_t, \quad (3.37)$$

де $Y_t = \{Y_{1t}, Y_{2t}, \dots, Y_{kt}\}^T$ — вектор змінних порядку $(k \times 1)$; $Y_{t-1}^* = \{Y_{1,t-1}, Y_{2,t-1}, \dots, Y_{k,t-1}, 1\}$ — вектор змінних розмірності $((k + 1) \times 1)$;

$$\pi^* = \begin{bmatrix} \pi_{11} & \dots & \pi_{1k} & a_{10} \\ \pi_{21} & \dots & \pi_{2k} & a_{20} \\ \dots & \dots & \dots & \dots \\ \pi_{k1} & \dots & \pi_{kk} & a_{k0} \end{bmatrix}.$$

Модель (3.33) може бути узагальнена, як і розширений тест Дікі—Фуллера, на випадок авторегресійних процесів вищих порядків:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t, \quad (3.38)$$

$$\text{або } \Delta Y_t = (A_1 - I) Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t. \quad (3.39)$$

Тепер додамо та віднімемо в правій частині (3.39) вираз $(A_1 - I) Y_{t-2}$. Отримаємо:

$$\Delta Y_t = (A_1 - I) \Delta Y_{t-1} + (A_2 + A_1 - I) Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t. \quad (3.40)$$

Далі додамо та віднімемо від обох частин (3.40) вираз $(A_2 + A_1 - I) Y_{t-3}$.

Отримаємо:

$$\Delta Y_t = (A_1 - I)\Delta Y_{t-1} + (A_2 + A_1 - I)Y_{t-2} + (A_3 + A_2 + A_1 - I)Y_{t-3} + \dots \\ \dots + A_p Y_{t-p} + \varepsilon_t.$$

Продовжуючи далі, отримаємо:

$$\Delta Y_t = \sum_{i=1}^{p-1} \pi_i \Delta Y_{t-i} + \pi Y_{t-p} + \varepsilon_t, \quad (3.41)$$

$$\text{де } \pi = -\left(I - \sum_{i=1}^p A_i\right), \pi_i = -\left(I - \sum_{j=1}^i A_j\right).$$

В моделі (3.41), як і раніше, ключове значення має ранг матриці π . Якщо ранг $\pi = 0$, всі елементи матриці дорівнюють нулю, а (3.41) є звичайною VAR-моделлю в перших різницях. Якщо матриця має повний ранг, то всі змінні є стаціонарними і ми не можемо мати коінтеграційних рівнянь. В інших випадках ранг матриці π дорівнює кількості коінтеграційних векторів. В особливому випадку, якщо ранг $\pi = 1$, має місце одне коінтеграційне рівняння, тому в (3.41) вираз πY_{t-p} є механізмом корегування помилки.

Якщо ранг матриці π знаходиться в межах $1 < \text{ранг}(\pi) < k$, то ми маємо декілька коінтеграційних рівнянь. Кількість різних коінтеграційних рівнянь може бути знайдена через аналіз характеристичних коренів матриці π . Ми знаємо, що ранг матриці дорівнює кількості характеристичних коренів, що відрізняються від нуля. Припустимо, що ми упорядкували характеристичні корені матриці π таким чином: $\lambda_1 > \lambda_2 > \dots > \lambda_k$. Якщо змінні Y_i не коінтегрують, ранг матриці π дорівнює нулю. Оскільки $\ln(1) = 0$, кожний з виразів $(\ln(1 - \lambda_1))$ повинен дорівнювати нулю у випадку, коли змінні не коінтегрують. Відповідно, якщо ранг матриці дорівнює одиниці, $0 < \lambda_1 < 1$, тобто перший вираз $(\ln(1 - \lambda_1))$ повинен бути від'ємним, а всі інші $\lambda_i = 0$, таким чином, $(\ln(1 - \lambda_2)) = (\ln(1 - \lambda_3)) = \dots = (\ln(1 - \lambda_k)) = 0$.

На практиці ми можемо отримати лише оцінку матриці π та її характеристичні корені. Тест на кількість характеристичних коренів, що статистично незначимо відрізняються від одиниці, можна провести на основі двох статистик:

$$\lambda_{\text{trace}}(r) = -T \sum_{i=p+1}^n \ln(1 - \hat{\lambda}_i), \quad (3.42)$$

$$\lambda_{\text{max}}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1}), \quad (3.43)$$

де $\hat{\lambda}_i$ — оцінені значення характеристичних коренів (eigenvalues), отримані з оціненої матриці π , T — кількість спостережень. Якщо значення r є визначеним, то статистики просто стають рівними λ_{trace} та λ_{max} .

Перша статистика тестує нульову гіпотезу, що кількість коінтеграційних рівнянь є меншою або рівною r , проти альтернативної, що ранг більший за r . Друга статистика тестує нульову гіпотезу, яка полягає в тому, що кількість коінтеграційних векторів дорівнює r , проти альтернативної, що вона дорівнює $(r + 1)$.

Йохансеном були отримані критичні значення для обох статистик. Розподіл цих статистик залежить від кількості нестационарних компонент за нуль-гіпотезою $(k - r)$, а також від форми вектора A_0 , тобто від того, чи включено зсув у систему (константи в коінтеграційне рівняння).

Для того, щоб протестувати гіпотезу про включення константи в коінтеграційні рівняння проти альтернативної гіпотези про включення зсуву в систему, потрібно оцінити обидва варіанти моделі та відповідно знайти характеристичні корені матриці π в другому випадку, випадку системи без обмежень на параметри (unrestricted model): $\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_k$, та характеристичні корені системи з перетинами в коінтеграційному рівнянні: $\hat{\lambda}_1^*, \hat{\lambda}_2^*, \dots, \hat{\lambda}_k^*$. Далі припустимо, що кількість ненульових характеристичних коренів у системі без обмежень дорівнює r , тоді асимптотично статистика

$$-T \sum_{i=r+1}^k \left[\ln(1 - \hat{\lambda}_i^*) - \ln(1 - \hat{\lambda}_i) \right] \text{ має } \chi^2\text{-розподіл зі ступенями свободи } (k - r).$$

Інтуїтивно зрозуміло, що обидва значення є рівними, якщо обмеження не є дійсними, тобто перший тип моделі не є коректним. Йохансен довів, що якщо значення цієї статистики достатньо велике, то ми можемо відкинути нульову гіпотезу про наявність перетину в коінтеграційних рівняннях та зробити висновок про наявність лінійного тренду в змінних.

Для того, щоб тестувати обмеження в коінтеграційному рівнянні, Йохансен розрізняє дві матриці α та β , обидві розміру $(k \times r)$, такі, що $\pi = \alpha\beta^T$. Матриця β є матрицею значень параметрів у коінтеграційних рівняннях, а матриця α є матрицею вагових коефіцієнтів (коефіцієнтів швидкості пристосування), з якими кожний коінтеграційний вектор входить в k рівнянь VAR-моделі. Через можливу залежність між збуреннями в різних рівняннях неможливо ефективно оцінити матриці α та β МНК, але методом максимальної правдоподібності (maximum likelihood estimation) можна не тільки оцінити (3.38), але й визначити ранг матриці π , використати r найбільш статистично значимих коінтеграційних векторів, щоб сформуванати матрицю β^T та знайти α , такі, що $\pi = \alpha\beta^T$.

Методологію Йохансена легко зрозуміти у випадку одного коінтеграційного рівняння. Припустимо, що ранг матриці π дорівнює одиниці, тоді всі

рядки матриці є лінійними комбінаціями одного коінтеграційного вектора, тобто (3.42) може бути записана як

$$\begin{aligned}\Delta Y_{1t} &= \dots + \pi_{11}Y_{1,t-p} + \pi_{12}Y_{2,t-p} + \dots + \pi_{1k}Y_{k,t-p} + \varepsilon_{1t}, \\ \Delta Y_{2t} &= \dots + \theta_2(\pi_{11}Y_{1,t-1} + \pi_{12}Y_{2,t-p} + \dots + \pi_{1k}Y_{k,t-p}) + \varepsilon_{2t}, \\ &\dots\dots\dots \\ \Delta Y_{kt} &= \dots + \theta_k(\pi_{11}Y_{1,t-p} + \pi_{12}Y_{2,t-p} + \dots + \pi_{1k}Y_{k,t-p}) + \varepsilon_{kt},\end{aligned}\quad (3.44)$$

де θ_i є скалярами та для спрощення вигляду системи $\pi_i \Delta Y_{t-i}$ опущені.

Тепер визначимо $\alpha_i = (\theta_i \pi_{ij})$, тоді кожне рівняння (3.44) можна записати у вигляді

$$\Delta Y_{it} = \dots + \alpha_i(Y_{1,t-p} + \beta_2 Y_{2,t-p} + \dots + \beta_k Y_{k,t-p}) + \varepsilon_{it}, \quad i = 1, 2, \dots, k \quad (3.45)$$

або відповідно у матричному вигляді

$$\Delta Y_{it} = \sum_{i=1}^{p-1} \pi_i \Delta Y_{t-i} + \alpha \beta^T Y_{t-p} + \varepsilon_t, \quad (3.46)$$

де $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_k)^T$ і є простим коінтеграційним вектором, а коефіцієнти швидкості пристосування відповідно дорівнюють: $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)^T$. Після того як вектори α та β^T знайдені, можливе тестування різних обмежень на ці вектори. Звичайно, потрібно пам'ятати, що якщо ми маємо r коінтеграційних векторів, то лише r лінійних комбінацій досліджуваних показників є стаціонарним рядом.

Які ж основні практичні заходи щодо застосування тесту Йохансена для перевірки наявності коінтеграційного зв'язку між двома або більше змінними? Вони складаються з ряду певних кроків, тому розглянемо кожний детально.

Крок 1.

Оцінити модель та визначити ранг матриці π . Практично всі економетричні пакети вміщують різноманітні методи оцінювання. Нагадаємо, що в даному випадку, при наявності можливої кореляції між залишками різних рівнянь, МНК не є найефективнішим методом оцінювання, краще використовувати або метод максимальної правдоподібності, або SUR-метод.

Оцінювати модель можна в трьох різних формах: в припущенні, що всі елементи матриці A_0 дорівнюють нулю; у формі зі зсувом; у формі, коли коінтеграційний вектор має перетин.

Крок 2.

Для того, щоб знайти ранг матриці π , використовуються два різних підходи, що базуються на різних статистиках: λ_{trace} - та λ_{max} -тести. За логікою вони однакові, але по-різному формують альтернативні гіпотези. Розглянемо послідовність формування гіпотез для кожного тесту окремо.

λ_{max} — тест визначення рангу матриці
(кількості коінтеграційних рівнянь)

Знаходження рангу матриці починається з нульової гіпотези (припущення), що він дорівнює нулю (змінні не коінтегрують), проти альтернативної гіпотези, що він дорівнює одиниці:

$$H_0: r = 0,$$

$$H_1: r = 1.$$

Якщо ми відкидаємо нульову гіпотезу, то далі тестуються такі послідовні можливі випадки:

$$H_0: r = 1, H_0: r = 2 \dots H_0: r = k - 1,$$

$$H_1: r = 2, H_1: r = 3 \dots H_1: r = k,$$

аж поки ми не зможемо відкинути нульову гіпотезу. Таким чином, ітеративно знаходимо ранг матриці.

Аналогічні дії використовуються і для λ_{trace} -тесту, але нульові та альтернативні гіпотези будуть трохи іншими.

λ_{trace} — тест визначення рангу матриці
(кількості коінтеграційних рівнянь)

Знаходження рангу матриці починається з нульової гіпотези (припущення), що він дорівнює нулю (змінні не коінтегрують), проти альтернативної гіпотези, що він дорівнює одиниці:

$$H_0: r = 0,$$

$$H_1: r > 0.$$

Якщо ми відкидаємо нульову гіпотезу, то далі тестуються такі послідовні випадки:

$$H_0: r \leq 1, H_0: r \leq 2 \dots H_0: r \leq k - 1,$$

$$H_1: r \geq 2, H_1: r \geq 3 \dots H_1: r \geq k,$$

аж поки ми не зможемо відкинути нульову гіпотезу. Таким чином, ітеративно знаходимо ранг матриці.

Крок 3.

Після того як знайдено ранг матриці та відповідно оцінена модель корегування помилки, аналізуються нормалізовані коінтеграційні рівняння та коефіцієнти швидкості пристосування.

Узагальнено тестування часових рядів на коінтеграцію за методологією Йохансена можна представити у вигляді блок-схеми.

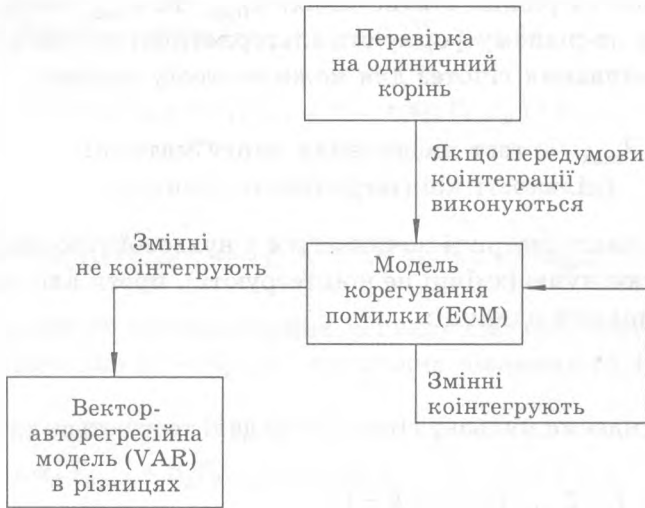


Рис. 3.3. Узагальнена блок-схема перевірки змінних на коінтеграцію за методологією Інгла—Гренджера

Нагадаємо: якщо ми перевіряємо на коінтеграцію дві змінні, то краще використовувати тест Інгла—Гренджера, а якщо перевіряються більше як дві змінні, краще використовувати техніку Йохансена.

Сучасні комп'ютерні програми автоматично видають всі описані вище результати, мета дослідника коректно їх проаналізувати та, в разі необхідності, трансформувати модель.

Таким чином, незважаючи на те, що багато економічних змінних є не-стаціонарними, їхні лінійні комбінації можуть бути стаціонарними. В такому випадку кажуть, що змінні коінтегрують. Коінтегровані змінні дозволяють використовувати моделі корегування помилки для моделювання такого зв'язку між ними, що кожна зі змінних відповідає за відхилення від довгострокової рівноваги. Одним із способів перевірити змінні на коінтеграцію є перевірка на стаціонарність залишків рівняння довгострокової рівноваги. Іншим методом є застосування λ_{trace} або λ_{max} -тесту, які дозволяють визначити ранг матриці, а отже, наявність та кількість коінтеграційних рівнянь. Якщо коінтеграційні рівняння існують, то ми можемо оцінювати та аналізувати модель корегування помилки, яка дозволяє одночасно оцінити зміни в короткостроковому та довгостроковому періодах.

Для випадку двох змінних покажемо за допомогою узагальненої блок-схеми (рисунок 3.4) процедуру вибору коректного типу моделі: вектор-авто-

регресійної або моделі корегування помилки. Для спрощення припустимо, що змінні є нестационарними з порядком інтеграції один. Блок-схема вибору типу моделі: VAR або ECM.

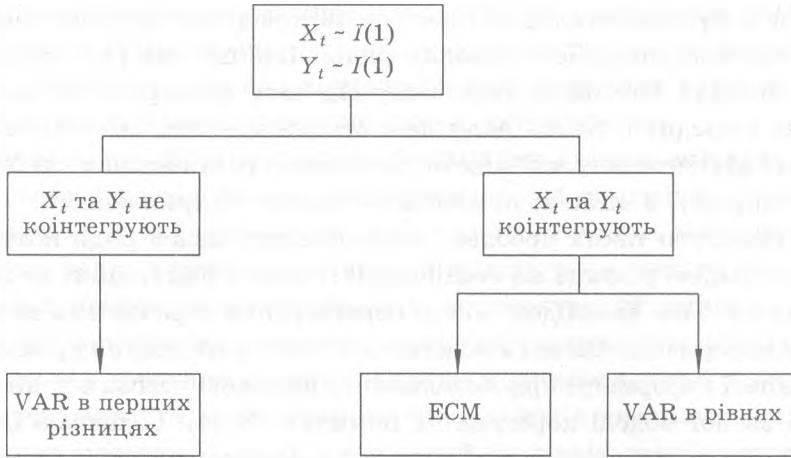


Рис. 3.4. Узагальнена блок-схема процедури вибору коректного типу моделі у випадку двох змінних

Як і для VAR-моделей, важливим інструментом аналізу якості ECM-моделі є імпульсний аналіз та аналіз декомпозиції дисперсій, який також допомагає зробити висновки щодо адекватності моделі. Методологія прогнозування на основі моделей корегування помилки аналогічна методології прогнозування на основі вектор-авторегресійних моделей, тож аналіз якості моделей і прогнозування на основі моделей корегування помилки в теоретичній частині ми розглядати не будемо. Ці питання розглянуто в практичній частині цього розділу при аналізі застосування на практиці моделей корегування помилки.

3.7. ЗАСТОСУВАННЯ МОДЕЛЕЙ КОРЕГУВАННЯ ПОМИЛКИ (ЕСМ) НА ПРАКТИЦІ: МОДЕЛЮВАННЯ ПОДАТКОВИХ НАДХОДЖЕНЬ В УКРАЇНІ

Одним з фундаментальних припущень класичної економетрики є припущення про стаціонарність часових рядів. Інакше кажучи, потрібно, щоб часові ряди мали постійну, незалежну від часу дисперсію та математичне сподівання (середню). Якщо ряди не є стаціонарними, оцінені на їх основі часткові коефіцієнти регресій цих моделей можуть виявитися «хибними» (неінформативними), а оцінені помилки — значно викривленими.

Щоб уникнути таких проблем, нестационарні часові ряди потрібно перетворити операцією різниць на стаціонарні (стільки разів, скільки це необхідно). Зауважимо, що внаслідок такого перетворення втрачається важлива довгострокова інформація. Одним з можливих шляхів розв'язання проблеми втрати довгострокової інформації при моделюванні на основі часових рядів є застосування так званої моделі корегування помилки (ЕСМ). Головною ідеєю цього підходу є оцінка довгострокового рівноважного взаємозв'язку (на основі значень відповідних часових рядів у рівнях) між досліджуваними показниками та його комбінація з короткостроковими зв'язками (оціненими на основі перетворених часових рядів у різницях). Побудова моделі корегування помилки є коректною тільки у випадку коінтеграції часових рядів.

Нагадаємо дуже стисло методологію побудови моделі корегування помилки та прокоментуємо її побудову на основі реальної інформації для прогнозування податкових надходжень у Зведений бюджет України.

Припустимо, що між досліджуваними показниками існує певний довгостроковий зв'язок та лише одна із досліджуваних змінних є ендогенною, тобто такою, що визначається всередині моделі, а решта змінних є екзогенними. Згодом припущенням про екзогенність решти змінних можна буде знехтувати.

Для спрощення будемо також вважати, що всі змінні є нестационарними з порядком інтеграції один ($I(1)$). Це означає, що всі вони є стаціонарними у перших різницях. Як нам вже відомо, ми не можемо застосовувати звичайний метод найменших квадратів для нестационарних рядів, оскільки він дає хибні результати. Хоча можна показати, що якщо змінні коінтегрують між собою, то метод найменших квадратів забезпечує суперконсистентні оцінки для рівняння довгострокової рівноваги (коінтеграційного рівняння).

Нагадаємо також, що коінтеграція означає, що лінійна комбінація змінних (або помилка, що вимірює відхилення від довгострокової рівноваги) є стаціонарним часовим рядом. Вектор коефіцієнтів коінтеграційного рівняння можна ідентифікувати з точністю до скаляра, бо, перемноживши всі коефіцієнти коінтеграційного рівняння на одну й ту ж ненульову константу, ми також отримуємо стаціонарний ряд.

Відповідно першим кроком перевірки існування коінтеграції між досліджуваними змінними є оцінка регресійного рівняння довгострокової рівноваги (їх лінійної комбінації) звичайним методом найменших квадратів (МНК), знаходження помилки даної моделі та перевірка часового ряду помилок на стаціонарність (за розширеним тестом Дікі—Фуллера [ADF] або Філіпса—Перрона [PP]). Якщо помилки є стаціонарними, то можна зробити висновок, що оцінена лінійна комбінація досліджуваних змінних насправді є рівнянням коінтеграції, тобто що змінні коінтегрують.

Зазначимо, що лінійна комбінація змінних є лише оцінкою довгострокового зв'язку і не відображає короткострокової динаміки. Для того, щоб поглибити економетричний аналіз та поліпшити точність моделювання, Інгл та Гренджер запропонували розглядати модель корегування помилки, яка об'єднувала б довгостроковий зв'язок для досліджуваних змінних з лагом одиниця та короткострокову динаміку, виражену залежністю різниць ендогенних змінних від поточних та лагових (затриманих на деякий проміжок часу) різниць екзогенних змінних. Найпростіша модель корегування помилки для випадку двох змінних Y_{1t} та Y_{2t} має такий формалізований вигляд:

$$\Delta Y_{1t} = a_{10} + \sum_{i=1}^k a_{11}(i)\Delta Y_{1,t-i} + \sum_{i=0}^k a_{12}(i)\Delta Y_{2,t-i} - \lambda_1 \hat{u}_{1,t-1} + \varepsilon_{1t},$$

$$\Delta Y_{2t} = a_{20} + \sum_{i=0}^k a_{21}(i)\Delta Y_{1,t-i} + \sum_{i=1}^k a_{22}(i)\Delta Y_{2,t-i} - \lambda_2 \hat{u}_{2,t-1} + \varepsilon_{2t},$$

де $\hat{u}_{1,t-1} = Y_{1,t-1} - \gamma_0 - \gamma_1 Y_{2,t-1} \sim I(0)$ — рівняння довгострокової рівноваги (коінтеграційне рівняння), нормоване по першій змінній; $\hat{u}_{2,t-1} = Y_{2,t-1} - \gamma'_0 - \gamma'_1 Y_{1,t-1} \sim I(0)$ — рівняння довгострокової рівноваги (коінтеграційне рівняння), нормоване по другій змінній. Відмітимо, що $\hat{u}_{1,t-1}$ та $\hat{u}_{2,t-1}$ є відхиленням від довгострокової рівноваги. Довгострокова рівновага досягається, якщо $Y_{1,t-1} = \gamma_0 + \gamma_1 Y_{2,t-1}$. Збурення ε_{1t} та ε_{2t} є білим шумом, при цьому вони можуть корелювати між собою.

Для того, щоб модель корегування помилки була динамічно стабільною, необхідно, щоб $0 \leq \lambda_1 < 1$, $0 \leq \lambda_2 < 1$. Зауважимо, що якщо коінтеграційні рівняння моделі нормовані по різних змінних, то знаки цих коефіцієнтів повинні бути від'ємними, тобто знаходитись на проміжку $[0, 1]$. Коефіцієнти λ_1 та λ_2 називаються швидкістю пристосування. При цьому коефіцієнт λ_1 є чутливістю (responsiveness) ΔY_1 на відхилення від рівноваги, а коефіцієнт λ_2 є чутливістю ΔY_2 на відхилення від рівноваги. Вони показують, на скільки відсотків відхилення від рівноваги миттєво корегуються кожною змінною. Відповідно решта відхилення корегується протягом наступних періодів.

Якщо коефіцієнти λ_1 та λ_2 виходять за межі проміжку $[-1, 0]$, то система є нестабільною, тобто не буде повертатись до стану рівноваги після дії певних шоків.

Для ілюстрації побудови моделі корегування помилки розглянемо прогнозування надходження податку на додану вартість (ПДВ) до Зведеного бюджету України. Необхідні дані наведені в додатку 3.¹¹

Очевидно, що на надходження ПДВ впливає багато факторів, хоча можна припустити, що найголовніший з них є валовий внутрішній продукт (ВВП) — агрегований показник доданої вартості, створеної в економіці, який відображає податкову базу цього податку. Звичайно, можна включити й інші фактори, але вони не є визначальними для ілюстрації нашого завдання, тому для спрощення інші фактори не розглядатимуться. Зазначимо лише, що ми введемо в модель корегування помилки лише фіктивні (dummy) змінні, необхідні для відображення певних змін у податковому законодавстві.

Отже, ми розглянемо побудову моделі корегування помилки для двох змінних: ПДВ та ВВП.

Завдання 3.1.

Перевірити, чи виконуються передумови коінтеграції для часових рядів ПДВ та ВВП, тобто чи є вони нестационарними однакового порядку. Якщо передумови коінтеграції виконуються, перевірити часові ряди на коінтеграцію. У випадку коінтеграції побудувати модель корегування помилки залежності між ПДВ та ВВП (дані наведені в додатку 3). Побудову моделі корегування помилки провести окремо за двома методологіями — Інгла—Гренджера та Йохансена. Провести аналіз оціненої моделі корегування помилки та зробити прогноз.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. В процесі виконання цього завдання ми використовували пакет E.VIEWS.

Послідовність виконання завдання 3.1.

Крок 1. Перевірка виконання передумов коінтеграції.

Перед побудовою моделі корегування помилки перш за все необхідно перевірити, чи є часові ряди, що розглядаються, нестационарними однакового порядку, за допомогою розширеного тесту Дікі—Фуллера. Якщо так, то перейти до кроку 2. Якщо ні, то ми не можемо будувати модель корегу-

¹¹ Зауважимо, що в процесі виконання цього завдання ми використовували пакет E.VIEWS.

вання помилки для змінних (зробіть цей крок самостійно по аналогії до завдання 1.1. побудови ARIMA-моделей, розділ 1).

Застосування методології Інгла—Гренджера для перевірки часових рядів на коінтеграцію та побудови моделі корегування помилки

Крок 2. Перевірка часових рядів на коінтеграцію за методологією Інгла—Гренджера.

Перевірте часові ряди на коінтеграцію за допомогою тесту Інгла—Гренджера. Які висновки ви можете зробити? Прокоментуйте ваші результати.

Крок 3. Побудова моделі корегування помилки за методологією Інгла—Гренджера.

Побудуйте модель корегування помилки для змінних, що розглядаються. Прокоментуйте детально результати розрахунків.

Застосування методології Йохансена для перевірки часових рядів на коінтеграцію та побудови моделі корегування помилки.

Крок 4. Перевірка часових рядів на коінтеграцію та побудова моделі корегування помилки за методологією Йохансена.

Перевірте часові ряди на коінтеграцію за допомогою тесту Йохансена. Які висновки ви можете зробити? Побудуйте модель корегування помилки для змінних, що розглядаються. Прокоментуйте детально результати розрахунків.

Крок 5. Порівняння методології Інгла—Гренджера та Йохансена.

Порівняйте підходи Інгла—Гренджера та Йохансена до побудови моделей корегування помилки. Який з підходів, на вашу думку, є найефективнішим та чому? Обґрунтуйте ваші висновки.

Крок 6. Імпульсний аналіз та аналіз декомпозиції дисперсій на основі оціненої моделі корегування помилки.

На основі оціненої на першому кроці моделі проведіть аналіз імпульсної функції відгуків та декомпозиції дисперсій. Поясніть отримані результати.

Крок 7. Прогнозування на основі оціненої моделі корегування помилки.

На основі оціненої моделі корегування помилки знайдіть прогноз на один квартал. Поясніть отримані результати.

Крок 8. Загальні висновки щодо побудованої моделі корегування помилки. Порівняння результатів моделювання за VAR-та ЕСМ-моделями.

На основі проведених розрахунків, аналізу функцій імпульсного відгуку та результатів прогнозу зробіть загальні висновки щодо ефективності моделювання за допомогою моделей корегування помилки. Крім того, порівняйте результати моделювання за VAR-моделями та ЕСМ-моделями в разі наявності коінтеграції між нестационарними змінними. В чому полягає принципова різниця? Прокоментуйте ваші результати.

Розв'язок завдання 3.1. та розширений коментар моделювання за допомогою моделей корегування помилки за двома методологіями

Для того, щоб прокоментувати прогнозування на основі моделі корегування помилки податку на додану вартість, введемо такі позначення: VAT — податок на додану вартість в номінальному вимірі (млрд грн); CPI — індекс споживчих цін (дані наведені в додатку 3); VATR — податок на додану вартість в реальному вимірі ($VATR = VAT/CPI$); VATRSA — податок на додану вартість, очищений від сезонності¹²; GDP — валовий внутрішній продукт в номінальному вимірі (млрд грн); GDPR — валовий внутрішній продукт в реальному вимірі ($GDPR = GDP/CPI$); GDPRSA — валовий внутрішній продукт, очищений від сезонності.

Для побудови моделі корегування помилки для двох змінних: ПДВ та ВВП використовуватимемо реальні щомісячні дані, наведені в додатку 3.

Крок 1. Перевірка виконання передумов коінтеграції часових рядів.

У завданні 1.1. до першого розділу, а саме для побудови ARIMA-моделей, ми вже показали, що реальні надходження ПДВ є нестационарним рядом, порядок інтеграції якого дорівнює одиниці. Легко можна також показати, що реальні податкові надходження, скореговані на сезонність (VATRSA), та реальний ВВП, скорегований на сезонність (GDPRSA)¹³, є також нестационарними часовими рядами з порядком інтеграції один ($I(1)$). Зробіть таку перевірку самостійно по аналогії до кроку 1 завдання 1.1. розділу 1. Ви повинні отримати результат, який свідчить, що досліджувані показники є нестационарними та мають однаковий порядок інтеграції. Отже, передумови для

¹² Нагадаємо, що десезонулювати часовий ряд в пакеті E.VIEWS можливо за допомогою команди у вікні часового ряду. Просто виберіть в меню вікна Procs/Seasonal Adjustment.

¹³ В загальному випадку потрібно було б працювати з рядами, не очищеними від сезонності, оскільки очищення сезонності може призводити до втрати цінної інформації щодо коінтеграції змінних. Можливі також сезонні одиничні корені. Для ясності застосування ЕСМ-моделей ми вирішили опустити ці тонкощі.

коінтеграції виконано. Тепер ви повинні перевірити ряди на наявність коінтеграції.

Крок 2. Перевірка часових рядів на коінтеграцію за методологією Інгла—Гренджера.

Для того, щоб перевірити часові ряди на коінтеграцію, ми маємо спочатку оцінити довгострокову залежність між надходженнями від ПДВ (VATRSA) та ВВП (GDPRSA) звичайним методом найменших квадратів (ще раз нагадаємо, що ми використовуємо реальні та сезонно скореговані змінні), а потім перевірити помилки оціненої моделі на стаціонарність.

Крок 2. 1. Оцінка регресійного рівняння залежності між надходженнями від ПДВ (VATRSA) та ВВП (GDPRSA) звичайним методом найменших квадратів.

В E.VIEWS це можна зробити за допомогою такої простої команди:

LS VATRSA C GDPRSA.

Або виберіть **Quick/Estimate Equation.**

Dependent Variable: VATRSA

Method: Least Squares

Sample(adjusted): 1995:02 2000:08

Included observations: 67 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002132	0.000718	-2.969788	0.0042
GDPRSA	0.126167	0.015905	7.932574	0.0000
R-squared	0.4911893	Mean dependent var		0.003533
Adjusted R-squared	0.484076	S.D. dependent var		0.000825
S.E. of regression	0.000592	Akaike info criterion		-11.99588
Sum squared resid	2.28E-05	Schwarz criterion		-11.93007
Log likelihood	403.8620	F-statistic		62.92573
Durbin-Watson stat	0.834327	Prob(F-statistic)		0.000000

Як можна побачити з результатів оцінки залежності ПДВ від ВВП, зростання на 1 млрд грн реального ВВП приводить до зростання надходжень ПДВ приблизно на 120 млн грн.

Уважний аналіз оціненої моделі довгострокового зв'язку показує велику ймовірність наявності автокореляції помилок (значення статистики Дарбіна—Уотсона (DW) дорівнює 0.83). Низьке значення статистики DW може слугувати сигналом того, що регресійне рівняння неправильно специфіковане, але не обов'язково свідчить про «хибну регресію». Дійсно, за правилом великого пальця, запропонованим Гренджером та Ньюболдом (Granger and Newbold), існує велика ймовірність того, що ми маємо «хибну регресію», якщо

коефіцієнт детермінації оціненої моделі (R^2) є більшим, аніж значення статистики DW. Проте це правило треба застосовувати дуже обережно і тільки як наближений результат, оскільки статистика DW побудована на припущенні про стаціонарність часових рядів.

Для того, щоб зробити висновок про коінтеграцію часових рядів, нам необхідно перевірити на стаціонарність помилки, розраховані на основі оціненого вище рівняння.

Крок 2.2. Тестування помилок оціненого рівняння на стаціонарність.

Проведемо тест помилок на стаціонарність за допомогою формальних тестів ADF або PP. Спершу згенеруємо помилки, вибравши **Procs/Make Residual Series** в меню вікна об'єкта рівняння. Для застосування розширеного тесту Дікі—Фуллера (ADF) треба вибрати **View/Unit root test** та встановити потрібне число лагів, порядок різниць та специфікацію рівняння (перетин, тренд чи обидва). Ми можемо отримати ті ж самі результати, написавши в командному вікні команду E.VIEWS:

UROOT(t,2) RESID_VAT,

де t означає «включити тренд і перетин у рівняння»; 2 означає, що кількість лагів у моделі дорівнює двом (ви можете обрати будь-яке інше позитивне число).

Зауважте, що ми перевіряємо, чи справджується гіпотеза про те, що $\alpha = 0$, в такому рівнянні:

$$\Delta RESID_t = m + \alpha \cdot RESID_{t-1} + \sum_{i=1}^k c_i \cdot \Delta RESID_{t-i}.$$

Тобто ми тестуємо гіпотезу про наявність одиничного кореня.

При цьому непараметричний тест Філіпса—Перрона (PP) є ефективнішим за ADF-тест у відхиленні нульової гіпотези про наявність одиничного кореня. Наприклад, ADF-тест може дати викривлені результати при невеликій кількості лагів (до трьох) в моделі, в той час як PP¹⁴ є дуже стійким до кількості лагів, в тому числі й до невеликих значень. Якщо ми відкидаємо гіпотезу H_0 , то можемо зробити висновок про стаціонарність помилок, або, іншими словами, про існування коінтеграції (довгострокової залежності) між змінними. Таким чином, можна переходити до наступного кроку, а саме до побудови моделі корегування помилки.

¹⁴ E.VIEWS використовує аналогічну процедуру для того, щоб застосувати Philips-Perron test. Виберіть часовий ряд. У вікні, що відкриється, виберіть View/Unit root test. Виберіть опцію: Philips-Perron test та визначте опції: перетин/тренд (intercept/trend); рівні/різниця (level/differences) та задайте кількість лагів. Нульова гіпотеза є ідентичною тій, яку використовує ADF-тест.

Крок 3. Побудова моделі корегування помилки за методологією Інгла—Гренджера.

Оскільки ми виявили наявність коінтеграції між ПДВ та ВВП, можемо оцінити модель корегування помилки для цих змінних, тобто побудувати залежність перших різниць реальних очищених від сезонності значень ПДВ (VATRSA) від перших різниць реальних очищених від сезонності значень ВВП (GDPRSA) та відхилення від довгострокової рівноваги з одиничним лагом.

Оцінити описану вище модель корегування помилки можна або за допомогою команди

```
LS D(VATRSA) C (VATRSA(-1)-0.1261*GDPRSA(-1)+0.002132) D(GDPRSA)
D(GDPRSA(-2)) D(GDPRSA(-3)) D(VATRSA(-1)) D(VATRSA(-2))
D(VATRSA(-3)) DUMMY9712 DUMMY9801 DUMMY97,
```

або працюючи з меню: **Quick/Estimate equation**, та надрукувати у вікні, що з'являється,

```
D(VATRSA) C (VATRSA(-1)-0.1261*GDPRSA(-1)+0.002132) D(GDPRSA)
D(GDPRSA(-2)) D(GDPRSA(-3)) D(VATRSA(-1)) D(VATRSA(-2))
D(VATRSA(-3)) DUMMY9712 DUMMY9801 DUMMY97.
```

Зазначимо, що ми включили фіктивні змінні для 1997¹⁵ фінансового року, для грудня 1997 і січня 1998 р.

Треба зауважити, що модель є стійкою, тому що коефіцієнт при значенні відхилення від довгострокової рівноваги з лагом один дорівнює -0.30 , тобто знаходиться на проміжку $[-1, 0]$. Значення коефіцієнта в даному випадку вказує на те, що приблизно 30 % відхилення від довгострокового тренду елімінується протягом одного часового періоду (в даному випадку місяця). Також варто звернути увагу на значення DW-статистики. Вона дорівнює близько 2; порівняйте це з першим варіантом МНК-регресії (див. таблицю 3.1).

Крок 4. Перевірка часових рядів на коінтеграцію та побудова моделі корегування помилки за методологією Йохансена.

На жаль, підхід Інгла—Гренджера до оцінювання вектора коінтеграції є іноді неприйнятним, особливо у випадку багатьох змінних, оскільки ми можемо не знати наперед кількості коінтеграційних векторів, тобто рангу матриці Π , де Π отримано з загального вигляду моделі корегування помилки:

$$\Delta Y_t = a + \Pi Y_{t-1} + \sum_{i=1}^k \Gamma_i \cdot \Delta Y_{t-i} + \varepsilon_t \text{ та}$$

¹⁵ Фіктивні змінні дорівнюють одиниці для всіх місяців 1997 р. та нулю для решти місяців.

Таблиця 3.1. Результати оцінки ECM-моделі

Dependent Variable: D(VATRSA)

Method: Least Squares

Sample(adjusted): 1995:01 2000:08

Included observations: 68 after adjusting endpoints

Newey-West HAC Standard Errors & Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.91E-05	6.91E-05	-1.289883	0.2023
VATRSA(-1)-(-0.0021+0.1261*GDPRSA(-1))	-0.301871	0.092518	-3.262836	0.0019
D(GDPRSA)	0.107389	0.017702	6.066570	0.0000
D(GDPRSA(-2))	0.052321	0.019803	2.642075	0.0106
D(GDPRSA(-3))	0.079131	0.020108	3.935330	0.0002
D(VATRSA(-1))	-0.242072	0.123427	-1.961258	0.0547
D(VATRSA(-2))	-0.334155	0.083132	-4.019550	0.0002
D(VATRSA(-3))	-0.407489	0.097451	-4.181489	0.0001
DUMMY9712	0.000813	0.000155	5.250942	0.0000
DUMMY9801	-0.000807	0.000118	-6.842813	0.0000
DUMMY97	0.000305	0.000111	2.739428	0.0082
R-squared	0.817763	Mean dependent var	-1.70E-05	
Adjusted R-squared	0.785792	S.D. dependent var	0.000823	
S.E. of regression	0.000381	Akaike info criterion	-12.75990	
Sum squared resid	8.28E-06	Schwarz criterion	-12.40086	
Log likelihood	444.8367	F-statistic	25.57798	
Durbin-Watson stat	1.976827	Prob(F-statistic)	0.000000	

$$Y_t = \begin{pmatrix} Y_{1t} & Y_{12} & \dots & Y_{1t} \\ Y_{11} & Y_{21} & \dots & Y_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{m1} & Y_{m1} & \dots & Y_{mt} \end{pmatrix}$$

Більш того, хоча методологію Інгла—Гренджера можна розширити на випадок множини коінтеграційних векторів, вона може бути нецприйнятна тому, що її двокрокова природа призводить до втрати ефективності, оскільки корисна інформація, отримана на першому кроці, не враховується на другому.

Методологія Йохансена вирішує обидві проблеми: ефективності (всі тести застосовуються на першому кроці) і множини коінтеграційних рівнянь.

Для застосування тесту Йохансена треба зробити такі кроки.

Виберіть набір змінних і відкрийте їх як групу. Ви можете зробити це як на основі виділення групи змінних безпосередньо в робочому файлі, так і за допомогою команди:

`show name_of_series.`

Потім виберіть **View/Cointegration test.**

У вікні, що відкриється, треба визначити:

- кількість лагів для різниць змінних, що використовуються для тесту на коінтеграцію;
- список екзогенних змінних (вони не входять в коінтеграційне рівняння, але вони наявні у різницях, так само, як різниці змінних, які входять в коінтеграційне рівняння);
- наявність тренду і/або перетину у коінтеграційному рівнянні.

Зауважимо, що ми можемо розглянути всі припущення про перетин і тренд одночасно, вибравши опцію **Summarize all 5 sets of assumptions**.

Johansen Cointegration Test

Cointegrating Equation (CE) and VAR specification:

Test assumes no deterministic trend in data:

- No intercept or trend in CE or test VAR
- Intercept (no trend) in CE - no intercept in VAR

Test allows for linear deterministic trend in data:

- Intercept (no trend) in CE and test VAR
- Intercept and trend in CE - no trend in VAR

Test allows for quadratic deterministic trend in data:

- Intercept and trend in CE - linear trend in VAR

Summary:

- Summarize all 5 sets of assumptions

Information:

The test VAR is estimated in differenced form.
CE and data trend assumptions apply to levels.

***** Warning *****
Test critical values were derived assuming no exogenous series.

Exogenous series in VAR: (don't include C or trend)

Lag intervals (pairs) in VAR:

OK Cancel

Виберіть припущення для коінтеграційного рівняння

Впишіть назви екзогенних змінних, в тому числі фіктивних

Виберіть кількість лагів, які використовуються для різниць у правій частині рівняння

У випадку VATRSA та GDPRSA рядів ми отримуємо (вибравши **Summarize all 5 sets of assumptions**) результат тесту Йохансена (див. таблицю 3.2).

Оскільки ми маємо тільки одну екзогенну змінну, можемо отримати щонайбільше один коінтеграційний вектор. Проте якщо ми включимо тренд, то можемо отримати два коінтеграційних вектори. Перший з них включає всі змінні, в той час як другий містить лише GDPRSA та часовий тренд.

Результат з таблиці 3.2 можна отримати, застосовуючи команду

COINT(S, 1) VATRSA GDPRSA.

Таблиця 3.2. Результат тесту Йохансена

Sample: 1995:01 1999:12
 Included observations: 60
 Series: VATRSA GDPRSA
 Lags interval: 1 to 1

Data Trend:	None	None	Linear	Linear	Quadratic
Rank or No. Of Ces	No Intercept No Trend	Intercept No Trend	Intercept No Trend	Intercept Trend	Intercept Trend
Log Likelihood by Model and Rank					
0	621.3671	621.3671	621.4162	621.4162	622.3830
1	630.1279	631.1579	631.1691	635.0828	636.0097
2	630.1771	633.7596	633.7596	637.7363	637.7363
Akaike Information Criteria by Model and Rank					
0	-20.57890	-20.57890	-20.51387	-20.51387	-20.47943
1	-20.73760	-20.73860	-20.70564	-20.80276	-20.80032
2	-20.60590	-20.65865	-20.65865	-20.72454	-20.72454
Schwarz Criteria by Model and Rank					
0	-20.43928	-20.43928	-20.30444	-20.30444	-20.20019
1	-20.45835	-20.42445	-20.35658	-20.41880	-20.38146
2	-20.18704	-20.16997	-20.16997	-20.16605	-20.16605
L.R. Test:	Rank = 1	Rank = 1	Rank = 2	Rank = 1	Rank = 1

Виберемо другу опцію для вигляду коінтеграційного рівняння: перетин та відсутність тренду (виберіть у вікні **Johansen Cointegration Test** другий набір припущень: **Test assumes no deterministic trend in data: intercept (no trend) in CE — no intercept in VAR** або задайте команду

COINT(B, 1) VATRSA GDPRSA)¹⁶.

Результат тесту Йохансена наводимо в таблиці 3.3.

Нульова гіпотеза тесту Йохансена полягає в тому, що ранг матриці P вважається рівним певному попередньо визначеному числу r (може бути меншим). Якщо нульова гіпотеза відкидається, мінімальний ранг матриці збільшують і нова нульова гіпотеза знову тестується. Наприклад, тест Йохансена починають з нульового рангу (**None**). В нашому випадку бачимо, що тест Йохансена відкидає нульову гіпотезу про рівність рангу нулю. Тому тестується

¹⁶ Інші набори припущень також можна тестувати. Ми вибрали даний набір припущень для отримання результатів, які можна порівняти з результатами, одержаними за технікою Інгла—Гренджера.

Таблиця 3.3. Результат тесту Йохансена

Sample: 1995:01 1999:12

Included observations: 60

Test assumption: No deterministic trend in the data

Series: VATRSA GDPRSA

Lags interval: 1 to 1

Eigenvalue	Likelihood Ratio	5 Percent Critical Value	1 Percent Critical Value	Hypothesized No. of CE(s)
0.278457	24.78509	19.96	24.60	None **
0.083068	5.203300	9.24	12.97	At most 1

() denotes rejection of the hypothesis at 5% (1%) significance level

L.R. test indicates 1 cointegrating equation(s) at 5% significance level

Unnormalized Cointegrating Coefficients:

VATRSA	GDPRSA	C
269.0000	-32.42233	0.481491
-22.80476	29.39616	-1.242374

Normalized Cointegrating Coefficients: 1 Cointegrating Equation(s)

VATRSA	GDPRSA	C
1.000000	-0.120529 (0.02063)	0.001790 (0.00094)

Log likelihood 631.1579

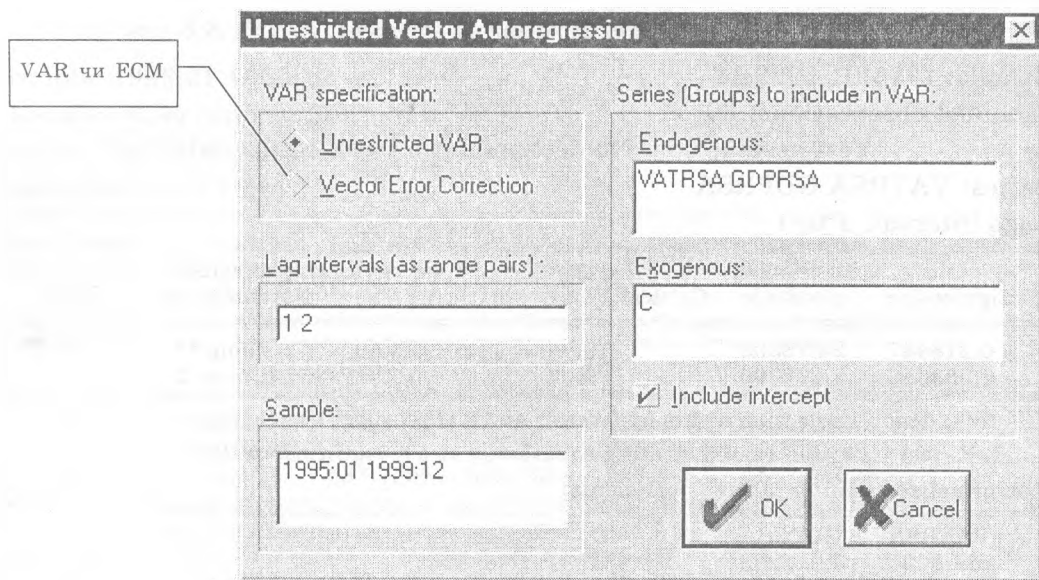
наступна гіпотеза, що ранг матриці дорівнює одиниці, і т. д. В нашому випадку тест не відкидає гіпотези про те, що ранг дорівнює одиниці (менший або дорівнює одиниці).

Ми знаємо, що коінтеграційний вектор можна визначити з точністю до скаляра. Щоб полегшити застосування векторів на практиці, коефіцієнти коінтеграційного вектора нормалізують таким чином, щоб перший ненульовий коефіцієнт у векторі дорівнював одиниці. Ви можете побачити нормалізований вектор в останньому рядку наведеної вище таблиці (стандартні помилки представлені в дужках). Варто зауважити, що ми отримуємо значення параметрів коінтеграційного вектора близьким до значення параметрів, отриманих за методологією Інгла—Гренджера.

Хоча це лише перший крок у побудові ECM- чи VAR-моделі. E.VIEWS пропонує легкий шлях переходу від тестування на коінтеграцію до оцінювання ECM(VAR)-моделі: для побудови моделі оберіть **Procs/Make Vector Autoregression**. Відкривається нове вікно (див. с. 152).

В цьому вікні ви маєте вибрати, з якою моделлю працювати: VAR чи ECM.

Ми знаємо, що $\Pi = \alpha \cdot \beta'$, де β — коінтеграційний вектор. Теоретично можливі три випадки:



1. Ранг (Π) дорівнює нулю. Коінтеграції немає, може застосовуватись лише необмежена (unrestricted) VAR-модель.
2. Ранг (Π) більший за нуль, але менший за k (кількість ендогенних змінних). Є один або більше коінтеграційних векторів. Застосовується ECM-модель.
3. Ранг (Π) дорівнює k . Всі ряди є стаціонарними (trend stationary), і треба застосовувати VAR в рівнях (in levels).

Для нашого прикладу другий випадок має місце, тому що тест Йохансена не відкидає гіпотези про те, що ранг Π дорівнює одиниці.

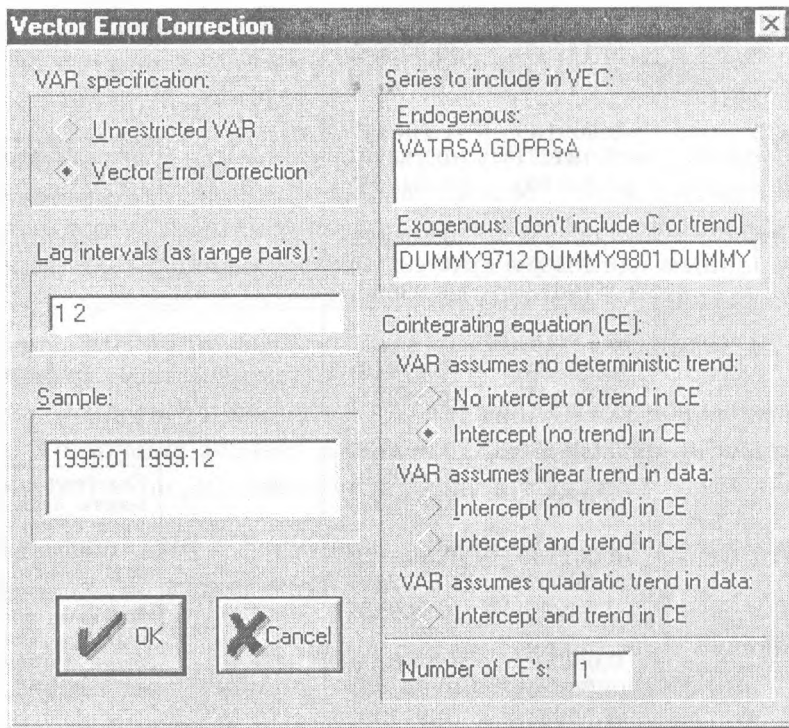
Якщо ми обираємо ECM в підменю **VAR specification**, вікно пропонує нам додаткові опції. В даному випадку ми можемо вибрати специфікацію коінтеграційного рівняння (тренд, перетин і т. д.).

Тепер оцінимо модель, представлену у вікні, що знаходиться вище. Ми вже знаємо з проведених тестів, що існує, можливо, лише одне коінтеграційне рівняння. Тому ми вибираємо **Number of CEs** (кількість коінтеграційних рівнянь), яке дорівнює 1. Зауважимо, що для інших прикладів звичайно можна отримати більше ніж одне коінтеграційне рівняння.

Для порівняння результатів з уже отриманими раніше за методологією Інгла—Гренджера виберіть **VAR assumes no deterministic trend: Intercept (no trend) in CE**. (VAR припускає відсутність детермінованого тренду: перетин (немає тренду) в коінтеграційному рівнянні).

Результат має такий вигляд (див. таблицю 3.4).

Використовуючи ці результати, ми легко можемо визначити довгострокову залежність між змінними (коефіцієнти нормалізовані таким чином, щоб коефіцієнти, які стоять при залежних змінних, дорівнювали одиниці).



Таблиця 3.4. Оцінка ECM-моделі

Sample: 1995:01 1999:12
 Included observations: 60
 Standard errors & t-statistics in parentheses

Cointegrating Eq:	CointEq1	
VATRSA(-1)	1.000000	
GDPRSA(-1)	-0.102309 (0.03964) (-2.58083)	
C	0.001225 (0.00178) (0.68801)	
Error Correction:	D(VATRSA)	D(GDPRSA)
CointEq1	-0.450869 (0.17841) (-2.52720)	-1.635571 (0.95235) (-1.71740)

Нормалізовані коефіцієнти

Зауважте, що залежна змінна у перших різницях

Коефіцієнт, що «вимірює» швидкість збігання до довгострокової рівноваги. Насправді змінна є відхиленням від довгострокової рівноваги

D(VATRSA(-1))	0.046581 (0.16299) (0.28578)	-0.062750 (0.87008) (-0.07212)
D(VATRSA(-2))	-0.569136 (0.13721) (-4.14799)	-1.206497 (0.73243) (-1.64725)
D(GDPRSA(-1))	-0.097801 (0.03505) (-2.79013)	-0.568173 (0.18711) (-3.03650)
D(GDPRSA(-2))	0.067469 (0.03325) (2.02924)	-0.144437 (0.17748) (-0.81381)
DUMMY9712	0.001350 (0.00060) (2.25177)	0.002865 (0.00320) (0.89538)
DUMMY9801	-0.001199 (0.00062) (-1.94290)	0.000129 (0.00329) (0.03918)
DUMMY97	0.000317 (0.00020) (1.55031)	0.001000 (0.00109) (0.91730)
R-squared	0.634985	0.452719
Adj. R-squared	0.585849	0.379046
Sum sq. residues	1.63E-05	0.000465
S.E. equation	0.000560	0.002990
F-statistic	12.92286	6.145022
Log likelihood	368.3890	267.8967
Akaike AIC	-12.01297	-8.663223
Schwarz SC	-11.73372	-8.383977
Mean dependent	-1.63E-05	-3.65E-05
S.D. dependent	0.000871	0.003795

Короткострокові ефекти

Екзогенні змінні, які мають бути стаціонарними. Якщо вони не є стаціонарними в рівнях (in levels), то до них треба застосувати процедуру різниць доти, доки вони не стануть стаціонарними

Критерії якості моделі

Determinant Residual Covariance	1.19E-12
Log Likelihood	653.3367
Akaike Information Criteria	-21.14456
Schwarz Criteria	-20.48135

$$\text{VATRSA} - 0.1023 * \text{GDPRSA} + 0.001225.$$

Всю модель можна представити так:

$$\begin{aligned} D(\text{VATRSA}) = & -0.4508 * (\text{VATRSA}(-1) - 0.1023 * \text{GDPRSA}(-1) + 0.001225) + \\ & + 0.0465 * D(\text{VATRSA}(-1)) - 0.5691 * D(\text{VATRSA}(-2)) - \\ & - 0.0978 * D(\text{GDPRSA}(-1)) + 0.0674 * D(\text{GDPRSA}(-2)) + \\ & + 0.00134 * \text{DUMMY9712} - 0.0011 * \text{DUMMY9801} + \\ & + 0.000316 * \text{DUMMY97} \end{aligned}$$

$$\begin{aligned} D(\text{GDPRSA}) = & -1.635 * (\text{VATRSA}(-1) - 0.1023 * \text{GDPRSA}(-1) + 0.001225) - \\ & - 0.0627 * D(\text{VATRSA}(-1)) - 1.2064 * D(\text{VATRSA}(-2)) - \end{aligned}$$

$$-0.5681 * D(\text{GDPRSA}(-1)) - 0.1444 * D(\text{GDPRSA}(-2)) + \\ + 0.0028 * \text{DUMMY9712} + 0.000129 * \text{DUMMY9801} + \\ + 0.0010 * \text{DUMMY97}.$$

Прокоментуємо отримані результати та звернемо увагу на деякі моменти.

1. Ми бачимо, що система стабільна відносно змінної VATRSA (коефіцієнт відхилення від довгострокової рівноваги [коінтеграційне рівняння] дорівнює -0.45), але нестабільна з точки зору іншої змінної — GDPRSA (-1.63). Хоча потрібно відмітити, що GDPRSA є слабо екзогенною, оскільки оцінений коефіцієнт є статистично незначимим при відносно низькому рівні довіри.
2. Кількість лагів у прикладних дослідженнях часто обирається виходячи з максимальних значень Акайк- та Шварц-критеріїв, якщо немає попередньої апріорної інформації щодо прийнятної їх кількості.

Крок 5. Порівняння методології Інгла—Гренджера та Йохансена.

Зробіть висновки самостійно.

Крок 6. Імпульсний аналіз та аналіз декомпозиції дисперсій на основі оціненої моделі корегування помилки.

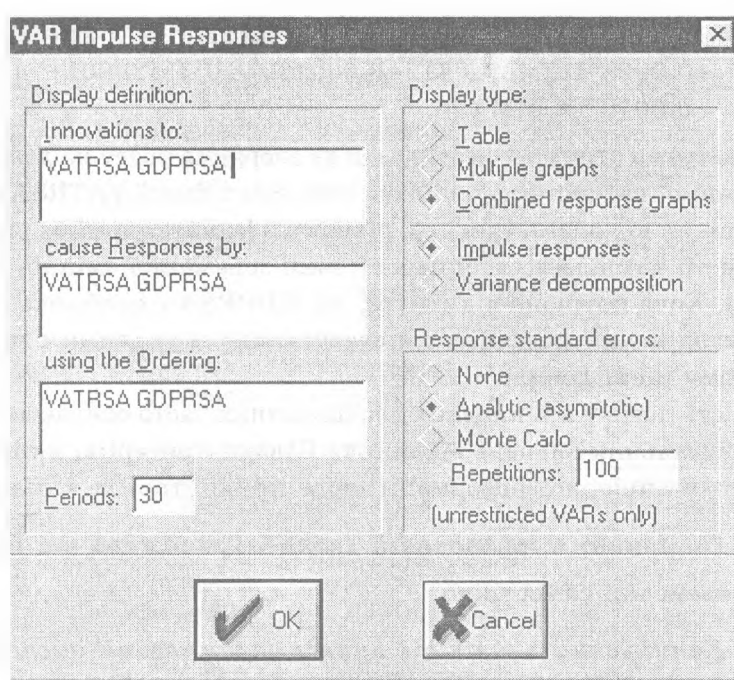
Аналіз оцінених коефіцієнтів, особливо тих, які ілюструють короткострокові ефекти, часто не має великого значення, оскільки дуже важко відокремити часткові моментні ефекти змінної, що має багато лагів. Іншими словами, коли пояснююча змінна має багато лагів у правій частині, то її вплив на поточне значення залежної змінної визначається не лише поточним значенням цієї пояснюючої змінної (або її окремим лагом), а повним набором лагів цієї пояснюючої змінної. Ось чому для кращого розуміння динамічних якостей моделі в багатьох випадках корисним є аналіз імпульсних функцій відгуків (IRF).

Це можна зробити, обравши **Impulse** у вікні VAR-об'єкта. При цьому відкривається нове вікно (див. с. 156).

Ми повинні насамперед визначити прогностичний горизонт (**Periods**), змінні, що впливають (**Innovations to:**), та змінні, що зазнають впливу (**Cause Responses by:**). Варто також зауважити, що кількість періодів потрібно вибирати достатньо великою, аби дослідник міг бачити, збігається система чи ні¹⁷.

Ми також повинні обрати тип оціненого результату (імпульсний відгук чи декомпозиція дисперсії). Функція імпульсного відгуку відображає реакцію змінних всередині системи на шок (зміну) в одне стандартне відхилення

¹⁷ Деякі пакети дають можливість обрахувати характеристичні корені динамічної матриці (companion matrix) і таким чином прямо визначити стабільність системи.



однієї зі змінних системи — з одного боку, з іншого — декомпозиція дисперсії показує відносну важливість факторів у впливі на динаміку (дисперсію) конкретної змінної.

Функції імпульсних відгуків (IRF) для нашого прикладу мають такий вигляд:

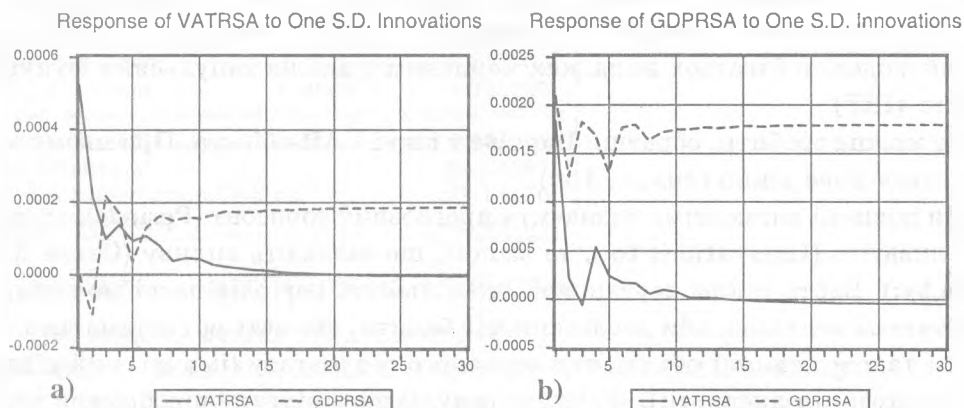


Рис. 3.5. Функції імпульсних відгуків для змінних VATRSA та GDPRSA

На рисунку 3.5.а) показано, яким чином шок (зміна) на одне стандартне відхилення у ВВП (пунктирна лінія) та ПДВ (суцільна лінія) впливає на

надходження від ПДВ. З цієї конкретної імпульсної функції відгуку ми можемо бачити, наприклад, що шок в одне стандартне відхилення у реальному ВВП призводить до росту надходжень від ПДВ, які потім стабілізуються на певному постійному позитивному рівні. З іншого боку, шок у ПДВ зменшується до нуля. До того ж система наближається до певного стабільного стану.

Рисунок 3.5.b) показує, як реагує ВВП на шок у ВВП (пунктирна лінія) та ПДВ (суцільна лінія). Ефект впливу ПДВ на ВВП швидко спадає. Тому наші висновки про слабку екзогенність ВВП виправдані.

Якщо ми оберемо для аналізу опцію «Декомпозиція дисперсій», то отримаємо такі результати:

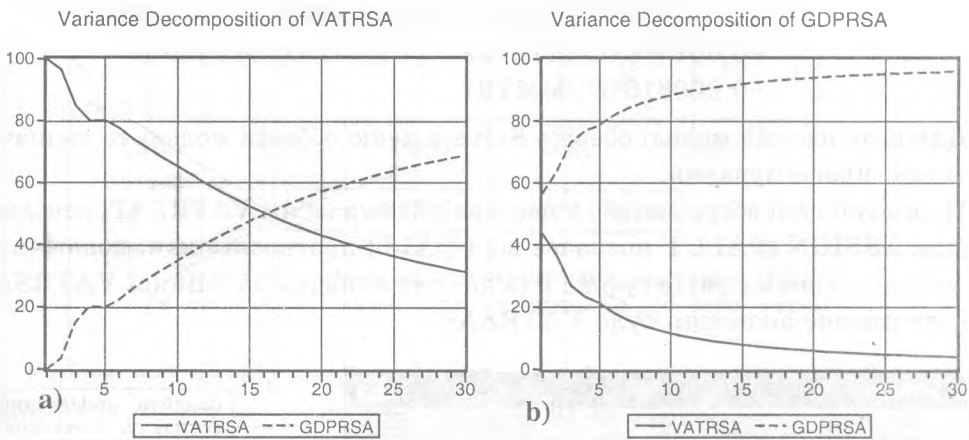


Рис. 3.6. Декомпозиція дисперсій для змінних VATRSA та GDPRSA

Зауважимо, що декомпозиція дисперсії є відносним показником і виражена в процентах.

Декомпозиція дисперсії для ПДВ (VATRSA) (рисунок 3.6.a)) показує, що динаміка лагів ПДВ (VATRSA) є важливішою за дисперсію ВВП в короткостроковій перспективі. Проте важливість динаміки ВВП постійно зростає і на кінець 30-го періоду пояснює більше ніж 60 % варіацій в податкових надходженнях. Якщо збільшити кількість періодів, частка ВВП у поясненні змін ПДВ наближається до 100 %. Стосовно ВВП (рисунок 3.6.b)) не важко помітити, що лаги цього показника пояснюють чималу частку власної дисперсії, тобто цей показник непогано пояснює себе сам.

Крок 7. Прогнозування на основі оціненої моделі корегування помилки.

Раз оцінивши модель, ми можемо використати її для прогнозування податкових надходжень. VAR-об'єкт має бути перетворений на об'єкт моделі. Це

можна зробити, вибравши в меню VAR-об'єкта **Procs/Make Model**. Після цього рівняння для слабко екзогенних змінних вилучаються з системи і модель готова до прогнозування. Для того, щоб зробити прогноз, ми розв'язуємо модель, надаючи конкретні значення екзогенним змінним у прогнозованому горизонті.

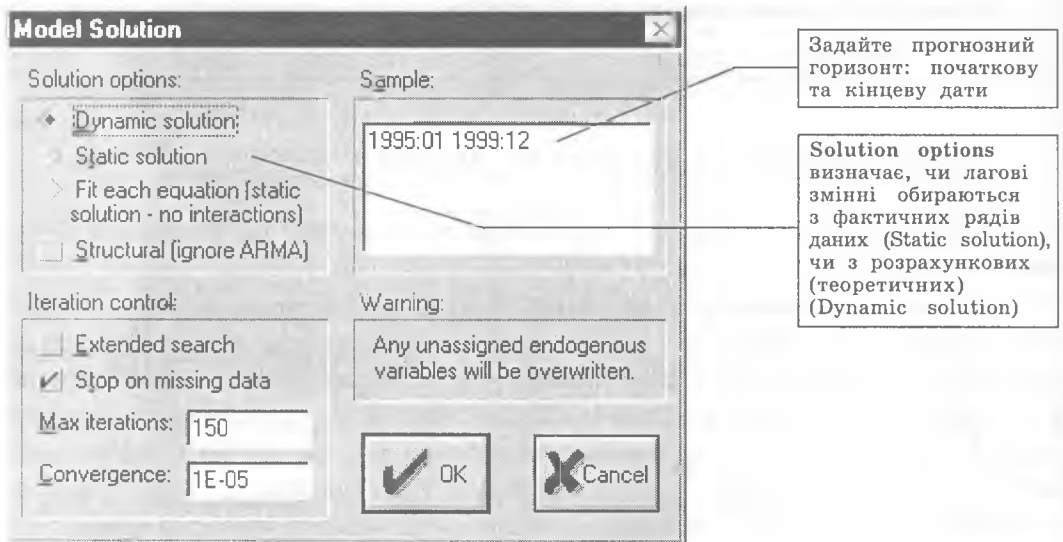
У даному випадку модель складається лише з одного рівняння, тому має дуже простий вигляд:

ASSIGN @ALL F

$$D(VATRSA) = -0.4508 * (VATRSA(-1)) - 0.1023 * GDPRSA(-1) + 0.00122 + 0.0465 * D(VATRSA(-1)) - 0.5691 * D(VATRSA(-2)) - 0.0978 * D(GDPRSA(-1)) + 0.0674 * D(GDPRSA(-2)) + 0.0013 * DUMMY9712 - 0.00119 * DUMMY9801 + 0.000316 * DUMMY97.$$

Для розв'язання моделі оберіть **Solve** в меню об'єкта моделі та визначте параметри прогнозування.

Прогнозні дані зберігаються у часових рядах з ім'ям **VATRSAF**, оскільки команда **ASSIGN @ALL F** позначає всі (@ALL) прогнозні ряди, додаючи до назв усіх ендогенних рядів суфікс **F** (в даному випадку до змінної **VATRSA**), отже, прогнозне значення було **VATRSAF**.



Звичайна найпростіша перевірка якості прогнозу полягає у візуальному порівнянні фактичних дійсних та теоретично розрахованих значень часових рядів. На рисунку 3.7 відображені фактичні та прогнозні значення ПДВ (в реальному вимірі та очищеному від сезонності).

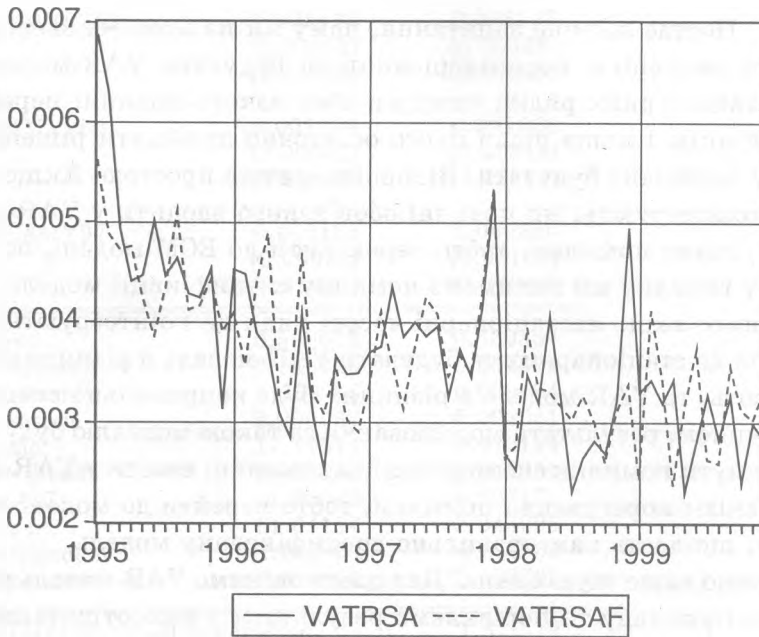


Рис. 3.7. Фактичні та прогнози значення ПДВ (в реальному вимірі та очищеному від сезонності)

Візуальний аналіз показує, що прогнози значення (теоретично розрахований ряд) досить непогано «відображають» фактичні.

Звичайно, висновки, зроблені на основі візуального аналізу, необхідно підкріпити формальними критеріями якості прогнозу, які наведені в розділі 1 для ARIMA-моделей та в розділі 4.

Розрахуйте критерії якості, наприклад MAPE, самостійно.

Зауважимо, що для отримання номінальних обсягів надходжень від ПДВ перемножте ряд VATRSA на сезонний коефіцієнт та ІСЦ. Коефіцієнти сезонності можна отримати на початковому етапі аналізу за допомогою команди `genr VATR_SCOEF=VATRSA/VATR`.

Крок 8. Загальні висновки щодо побудованої моделі корегування помилки. Порівняння результатів моделювання за VAR- та ECM-моделями.

Спробуйте зробити загальні висновки самостійно.

Ми звернемо вашу увагу лише на один, але дуже важливий момент при побудові моделей корегування помилки, а саме на порівняння моделювання за VAR- та ECM-моделями.

Попередній аналіз часових рядів ПДВ та ВВП (в реальному вимірі та очищених від сезонності) показав, що вони є нестационарними з порядком

інтеграції один. Постає логічне запитання, чому ми не можемо звести ряди до стаціонарних операцією перших різниць та будувати VAR-модель для перетворених стаціонарних рядів; чому ми обов'язково повинні перевірити ряди на коінтеграцію і лише після цього остаточно приймати рішення про тип моделі, яку необхідно будувати. Відповідь є дуже простою. Якщо нестационарні ряди коінтегрують, ми повинні обов'язково вводити у VAR-модель механізм корегування помилки, тобто переходити до ECM-моделі, оскільки в протилежному випадку ми матимемо помилку специфікації моделі. Отже, ще раз наголосимо: якщо нестационарні часові ряди не коінтегрують, то ми повинні звести їх до стаціонарних та будувати VAR-модель в різницях. Якщо ряди коінтегрують, то VAR-модель в різницях буде неправильно специфікованою і, таким чином, результати моделювання за такою моделлю будуть хибними. Щоб уникнути помилки специфікації, ми повинні ввести у VAR-модель в різницях механізм корегування помилки, тобто перейти до моделі корегування помилки, що дасть нам правильно специфіковану модель.

Проілюструємо наше зауваження. Для цього оцінимо VAR-модель в різницях для нашого прикладу та порівняємо результати з уже отриманими результатами за моделлю корегування помилки (ECM).

Для побудови VAR оберіть опцію **Unrestricted VAR** у вікні VAR-об'єкта та оцініть модель (див. таблицю 3.5).

Таблиця 3.5. Результати розрахунків за VAR-моделлю

Unrestricted VAR		
Sample: 1995:01 1999:12		
Included observations: 60		
Standard errors & t-statistics in parentheses		
	D(VATRSA)	D(GDPRSA)
D(VATRSA(-1))	-0.194949 (0.14050) (-1.38757)	-0.942780 (0.73058) (-1.29045)
D(VATRSA(-2))	-0.757554 (0.12469) (-6.07574)	-1.897398 (0.64836) (-2.92645)
D(GDPRSA(-1))	-0.062638 (0.03339) (-1.87623)	-0.441664 (0.17360) (-2.54410)
D(GDPRSA(-2))	0.103760 (0.03130) (3.31463)	-0.013079 (0.16278) (-0.08035)
C	-9.63E-05 (8.7E-05) (-1.10869)	-0.000402 (0.00045) (-0.89000)

DUMMY9712	0.001422 (0.00063) (2.26786)	0.003127 (0.00326) (0.95906)
DUMMY9801	-0.001634 (0.00061) (-2.66390)	-0.001385 (0.00319) (-0.43418)
DUMMY97	0.000126 (0.00020) (0.63437)	0.000360 (0.00103) (0.34967)
R-squared	0.599618	0.430354
Adj. R-squared	0.545720	0.353671
Sum sq. Resids	1.79E-05	0.000484
S.E. equation	0.000587	0.003051
F-statistic	11.12514	5.612114
Log likelihood	365.6145	266.6951
Akaike AIC	-11.92048	-8.623171
Schwarz SC	-11.64124	-8.343925
Mean dependent	-1.63E-05	-3.65E-05
S.D. dependent	0.000871	0.003795
Determinant Residual Covariance		1.31E-12
Log Likelihood		650.5803
Akaike Information Criteria		-21.15268
Schwarz Criteria		-20.59419

Неважко помітити, що характеристики рівнянь для показника GDPRSA, а особливо для VATRSA, значно змінилися. Тому ми можемо зробити висновок, що виключення з моделі важливої довгострокової інформації призведе до зміщення.

Завдання 3.2.

Використовуючи дані додатку 3, побудуйте моделі корегування помилки для інших податкових надходжень.

Питання для самоперевірки

1. Поясніть поняття коінтеграції часових рядів. Дайте визначення коінтеграції на прикладі двох часових рядів.
2. Наведіть основні передумови можливої наявності коінтеграції часових рядів.
3. В чому полягає поняття «хибної» регресії? До яких наслідків вона приводить?
4. Якщо часові ряди коінтегрують, то чи можна побудувати модель корегування помилки (ЕСМ)? Які її основні властивості?
5. Запишіть загальний вигляд моделі корегування помилки (ЕСМ) з p -лагами для двох змінних. Детально прокоментуйте модель.

6. Опишіть загальну схему побудови моделі корегування помилки (ЕСМ) на прикладі двох змінних.

7. Чи завжди модель корегування помилки (ЕСМ) є статистично коректною? Якщо так, то поясніть чому.

8. Що таке загальне поняття коінтеграції за Інглом та Гренджером?

9. Чи завжди можна оцінити параметри моделі корегування помилки (ЕСМ)? Поясніть ваші докази.

10. На прикладі двох змінних поясніть поняття коінтеграційного рівняння та коінтеграційного вектора.

11. Поясніть, чому, як правило, робиться нормалізація коінтеграційного рівняння.

12. Якщо на коінтеграцію перевіряється k змінних, яку максимальну кількість коінтеграційних векторів можна очікувати? Прокоментуйте.

13. Що означає поняття довгострокової рівноваги? Запишіть рівняння довгострокової рівноваги для двох змінних. В якому випадку його можна побудувати? Як оцінити рівняння довгострокової рівноваги?

14. Нехай два часових ряди коінтегрують. За допомогою якого тесту це можна перевірити?

15. Опишіть загальну схему механізму корегування помилки.

16. Поясніть поняття спільного стохастичного тренду у випадку коінтеграції.

17. Запишіть загальний вигляд моделі корегування помилки (ЕСМ) для двох змінних. Що означає швидкість пристосування? Поясніть, які коефіцієнти моделі її вимірюють.

18. Запишіть загальний вигляд моделі корегування помилки (ЕСМ) для двох змінних. Що означає стабільність моделі? Поясніть, які значення і яких коефіцієнтів вказують на наявність стабільності моделі.

19. Чим відрізняється VAR в перших різницях від моделі корегування помилки (ЕСМ) для двох змінних з порядком інтеграції один? При наявності коінтеграції, яка з моделей є вірно специфікованою, а яка ні? Поясніть чому.

20. Запишіть загальний вигляд моделі корегування помилки (ЕСМ) для k змінних в матричному вигляді. Прокоментуйте модель.

21. Як пов'язані ранг матриці та кількість коінтеграційних рівнянь для моделі корегування помилки (ЕСМ) у випадку k змінних?

22. Як перевірити часові ряди на стаціонарність? Чи завжди нестационарний ряд можна звести до стаціонарного?

23. Чим відрізняється тест Дікі—Фуллера (DF) перевірки часового ряду на стаціонарність від розширеного тесту Дікі—Фуллера (ADF)?

24. Запишіть базове рівняння тесту Інгла—Гренджера перевірки двох часових рядів на коінтеграцію. Чим відрізняється розширений тест Дікі—

Фуллера (ADF) перевірки часового ряду на стаціонарність від тесту Інгла—Гренджера перевірки двох часових рядів на коінтеграцію?

25. В чому полягають основні недоліки тесту Інгла—Гренджера перевірки часових рядів на коінтеграцію?

26. Прокоментуйте методологію Йохансена перевірки часових рядів на коінтеграцію. Як в методології Йохансена пов'язані між собою ранг матриці та кількість коінтеграційних рівнянь?

27. Детально поясніть методологію Йохансена у випадку наявності одного коінтеграційного рівняння.

28. Поясніть значення статистик $\lambda_{trace}(r)$ та $\lambda_{max}(r, r + 1)$ для тесту Йохансена.

29. В чому, на ваш погляд, полягає перевага методології Йохансена над методологією Інгла—Гренджера?

30. Опишіть схему тестування наявності перетину в коінтеграційному рівнянні.

31. Якщо ми включаємо перетин в коінтеграційне рівняння, чому це свідчить про наявність лінійного часового тренду в даних?

32. Прокоментуйте детально особливості прогнозування за допомогою ЕСМ-моделей.

33. Що таке імпульсна функція відгуків? Яку інформацію можна отримати при її аналізі?

34. Що таке декомпозиція дисперсій? Яку інформацію можна отримати при її аналізі?

35. Чим відрізняється моделювання за допомогою VAR-моделей від моделювання на основі ЕСМ-моделей?

36. Назовіть основні властивості моделей корегування помилки.

Вправи

Вправа 3.1.

Розгляньте часові ряди Y_4 та Y_5 (прибуток та дивіденди відповідно) з вправи 1.1. розділу 1. Оскільки дивіденди залежать від прибутків, розгляньте таку просту модель:

$$Y_5 = \beta_1 + \beta_1 Y_{4t} + u_t.$$

(а) Чи матиме дана регресія ваду від феномена «хибної» регресії? Чому?

(б) Чи коінтегровані часові ряди Y_4 та Y_5 ? Як ви це протестуєте? Якщо після перевірки ви зробите висновок, що вони коінтегрують, то чи зміниться ваша відповідь на запитання (а)?

- (в) Застосуйте механізм корегування помилки (ЕСМ), щоб вивчити коротко- і довгострокову поведінку дивідендів від прибутків.
- (г) Розглянемо часові ряди Y_4 та Y_5 окремо. Які тренди вони утворюють: стохастичні чи детерміновані? Які тести ви використовуєте?
- (д) Припустимо, що дивіденди і прибутки коінтегровані. Тоді замість того, щоб будувати регресійну модель залежності дивідендів від прибутків, ви будете регресійну модель залежності прибутків від дивідендів. Чи буде така регресія правильною?

Зауваження. Для розв'язання вправ 3.1—3.3 використайте пакет E.VIEWS або інший доступний вам статистичний пакет.

Вправа 3.2.

Замість того, щоб оцінювати регресійну модель залежності дивідендів від прибутків у рівнях, припустіть, що ви оцінюєте регресійну модель в перших різницях. Чи включатимете ви перетин у цю регресію? Чому? Наведіть та прокоментуйте ваші розрахунки.

Вправа 3.3.

Продовження вправи 3.2. Як ви перевірите регресійну модель в перших різницях вправи 3.2 на стаціонарність? Що ви можете очікувати апріорі від регресійної моделі й чому? Наведіть та прокоментуйте ваші розрахунки.

Список літератури

1. *Abraham B., Ledolter J.* Statistical Methods for Forecasting.— New York : Wiley, 1989.
2. *Anderson T. W., Vahid F.* Testing Multiple Equation System for Common Nonlinear Components // *Journal of Econometrics*, 1998, № 84.— P. 1—6.
3. *Baillie R. T.* Long Memory Processes and Fractional Integration in Econometrics // *Journal of Econometrics*, 1996, № 73.— P. 5—60.
4. *Balke N. S., Fomby T.* Threshold Cointegration // *International Economic Review*, 1997, № 38.— P. 627—646.
5. *Banerjee A., Dolado J., Galbraith J. W., Hendry D. F.* Cointegration, Error Correction, and the Econometric Analysis of Nonstationary Data.— Oxford : Oxford University Press, 1993.
6. *Boswijk H. P.* Asymptotic Theory for Integrated Processes.— Oxford : Oxford University Press, 1998.
7. *Campbell John Y., Perron Pierre.* Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Working Papers Series, 1991 (Available at www.nber.org).

8. *Charemza W. W., Deadman D. F.* New Direction in Econometric Practice.— Brookfield : Edward Eglar, 1992.
9. *Dickey D. A., Fuller W. A.* Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root // *Econometrica*, 1981, № 49.— P. 1057—1072.
10. *Clements M. P., Hendry D. F.* Forecasting in Cointegrated Systems // *Journal of Forecasting*, 1995, № 10.— P. 127—146.
11. *Enders W.* Applied Econometric Time Series.— New York : Wiley, 1995.
12. *Engle R. F., Granger C. W.* Cointegration and Error Correction: Representation. Estimation and Testing // *Econometrica*, 1987, № 55.— P. 251—276.
13. *Engle R. F., Yoo B. S.* Forecasting and Testing in Cointegrated System // *Journal of Econometrics*, 1987, № 35.— P. 143—159.
14. *Frances Philip Hans.* Time Series Models for Business and Economic Forecasting.— Cambridge : Cambridge University Press, 1999.— 280 p.
15. *Frances P. H., Hylleberg S., Lee H. S.* Spurious Deterministic Seasonality // *Economics Letters*, 1995, № 48.— P. 249—256.
16. *Gonzalo J., Granger C. W.* Estimation of Common Long — Memory Components in Cointegrated System // *Journal of Business and Economic Statistics*, 1995, № 13.— P. 27—36.
17. *Granger C. W.* Some Recent Developments in a Concept of Causality // *Journal of Econometrics*, 1988, № 39.— P. 199—211.
18. *Granger C. W. J., Newbold P.* Spurious Regressions in Econometrics // *Journal of Econometrics*, 1974, № 2.— P. 111—120.
19. *Granger C. W. J., Terasvirta T.* Modelling Nonlinear Economic Relationships.— Oxford : Oxford University Press, 1993.
20. *Greene William H.* Econometric Analysis.— 4th ed.— New Jersey : Prentice Hall, 2000.— 1004 p.
21. *Gujarati Damodar N.* Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, 1995.— 838 p.
22. *Hamilton James D.* Time Series Analysis.— Princeton, New Jersey : Princeton University Press, 1994.
23. *Hansen L. P., Sargent T. J.* Seasonality and Approximation Errors in Rational Expectations Models // *Journal of Econometrics*, 1993, № 55.— P. 21—56.
24. *Hatanaka M.* Time-Series-Based Econometrics. Unit Roots and Cointegration.— Oxford : Oxford University Press, 1996.
25. *Hylleberg S.* Modelling Seasonality.— Oxford : Oxford University Press, 1992.
26. *Hylleberg S., Engle R. F., Granger C. W. J., Yoo B. S.* Seasonal Integration and Cointegration // *Journal of Econometrics*, 1990, № 44.— P. 215—238.
27. *Hendry D. F.* Dynamic Econometrics.— Oxford : Oxford University Press, 1995.

28. *Johansen S., Juselius K.* Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money // Oxford Bulletin of Economics and Statistics, 1990, № 52.— P. 169—210.
29. *Johansen S.* Statistical Analysis of Cointegration Vector // Journal of Economics Dynamics and Control, 1988, № 12.— P. 231—254.
30. *Johansen S.* Cointegration in Partial Systems and the Efficiency of Single — Equation Analysis // Journal of Econometrics, 1992, № 52.— P. 389—402.
31. *Johansen S.* Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.— Oxford : Oxford University Press, 1995.
32. *Johnston J., Dinardo J.* Econometric Methods.— New York : McGraw-Hill, 1997.
33. *Kennedy P.* A Guide to Econometrics.— Cambridge, Massachusetts : The MIT Press, 1998.— 468 p.
34. *Lee T.-H., Tse Y.* Cointegration Tests with Conditional Heteroskedasticity // Journal of Econometrics, 1996, № 73.— P. 401—410.
35. *Lukepohl H.* Introduction to Multiple Time Series Analysis.— Berlin : Springer-Verlag, 1991.
36. *Lukepohl H., Reimers H.-E.* Impulse Response Analysis of Cointegrated Systems // Journal of Economic Dynamics and Control, 1992, № 16.— P. 53—78.
37. *Lin J.-L., Tsay R. S.* Cointegration Constraints and Forecasting: An Empirical Examination // Journal of Applied Econometrics, 1996, № 11.— P. 519—538.
38. *Mills Terence C.* Time Series Techniques for Economists.— Cambridge : Cambridge University Press, 1998.
39. *Pagan A. R.* The Econometrics of Financial Markets // Journal of Empirical Finance, 1996, № 3.— P. 15—102.
40. *Patterson Kerry.* An Introduction to Applied Econometrics: A Time Series Approach.— London : St. Martin Press, 2000.
41. *Paulsen J.* Order Determination of Multivariate Autoregressive Time Series with Unit roots // Journal of Time Series Analysis, 1984, № 5.— P. 115—127.
42. *Stock J., Watson M. W.* Testing for Common Trend // Journal of the American Statistical Association, 1989, № 83.— P. 1097—1107.
43. *Tiao G. C., Box G. E. P.* Modelling Multiple Time Series with Applications // Journal of the American Statistical Association, 1981, № 76.— P. 802—816.
44. *Tong H.* Non-linear Time Series. A Dynamical System Approach.— Oxford : Oxford University Press, 1990.
45. *Veenstra A. W., Frances P. H.* A Cointegration Approach to Forecasting Freight Rates in the Dry Bulk Shipping Sector // Transportation Research A, 1997, № 31.— P. 447—458.

46. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика. Практикум з використанням комп'ютера.— К. : Знання, 1998.— 217 с.
47. *Лук'яненко І. Г., Городніченко Ю. О.* Прогнозування податкових надходжень за допомогою моделей корегування помилки // *Фінанси України*, 2001, № 7.— С. 89 — 99.
48. *Султан К., Лук'яненко І. Г., Городніченко Ю. О.* Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К.: Видавничий дім «КМ Academia», 2000.— 204 с.
49. *Холден К., Піл Д. А., Томпсон Дж. Л.* Економічне прогнозування. Вступ.— К. : Інформтехніка-ЕМІЦ, 1996.— 215 с.
50. *Черняк О. І., Ставицький А. В.* Динамічна економетрика.— К., 2000.— 120 с.

РОЗДІЛ 4. ЗАСТОСУВАННЯ СИМУЛЬТАТИВНИХ СИСТЕМ РІВНЯНЬ В ЕМПІРИЧНИХ ДОСЛІДЖЕННЯХ

4.1. ТЕОРЕТИЧНІ ОСНОВИ ПОБУДОВИ ЕКОНОМЕТРИЧНИХ СИМУЛЬТАТИВНИХ МОДЕЛЕЙ. ОСНОВНІ ПОНЯТТЯ В ОЦІНЮВАННІ СИМУЛЬТАТИВНИХ СИСТЕМ РІВНЯНЬ. СТРУКТУРНА ТА ПРИВЕДЕНІ ФОРМИ СИМУЛЬТАТИВНИХ МОДЕЛЕЙ

Застосування методу найменших квадратів (МНК) для знаходження невідомих параметрів регресійної моделі передбачає, що існує тільки односторонній зв'язок між залежною змінною Y (регресант; ендогенна змінна) та незалежними змінними X (регресори; екзогенні; пояснюючі змінні), тобто припускають, що змінні в правій частині регресійного рівняння (фактори) не корелюють із випадковими величинами. Якщо ці умови не виконуються, тобто якщо деякі фактори можуть, наприклад, в свою чергу, визначатись через Y , то порушується одне з припущень класичного регресійного аналізу про незалежність між собою випадкових величин та факторів. У такому випадку $\text{COV}(X, \varepsilon) \neq 0$ і використання методу найменших квадратів для знаходження невідомих параметрів моделі даватиме зміщені неконсистентні оцінки.

Якщо Y є функцією від X ($Y = f(X)$), а X , в свою чергу, є функцією від Y ($X = g(Y)$), взаємозв'язок між Y та X неможливо описати за допомогою лише одного регресійного рівняння. В такому випадку переходять від регресійної моделі з одним рівнянням до регресійної моделі з багатьма рівняннями, серед яких можуть бути рівняння, які включають X та Y як ендогенні і пояснюючі змінні. Система, що описує таку взаємну залежність між змінними, називається *системою одночасних або симультативних рівнянь*.

Симультативна система рівнянь називається ще структурною моделлю.

Структурна модель — це повна система рівнянь, що описує структуру взаємозв'язків між змінними. Структурні рівняння виражають ендогенні змінні як функції інших ендогенних змінних, предетермінованих змінних (екзогенних та лагових ендогенних і екзогенних змінних) та випадкових

величин. Ендогенні змінні вважаються стохастичними, тоді як предетерміновані змінні тлумачаться як нестохастичні.

Детальніше розглянемо *предетерміновані (або попередньо визначені) змінні*. Їх умовно можна поділити на дві категорії: *поточні екзогенні та лагові екзогенні й ендогенні змінні*. Наприклад, якщо X_{1t} є поточною екзогенною змінною, то X_{1t-1} вважається лаговою екзогенною змінною з одиничним лагом. І X_{1t} , і X_{1t-1} є предетермінованими змінними. Якщо Y_t є ендогенною стохастичною змінною, то Y_{t-1} є лаговою ендогенною змінною, значення якої відоме в поточний період часу t , отже, Y_{t-1} вважається нестохастичною змінною. Звідси лагові ендогенні змінні належать до предетермінованих змінних¹. Право визначати, які змінні в моделі є ендогенними, а які є попередньо визначеними, належить, за певних обмежень, досліднику, який розробляє модель. Виходячи з наведених пояснень, запишемо симультазивну (структурну) модель у загальному випадку:

$$\begin{aligned} Y_{1t} &= \beta_{12}Y_{2t} + \dots + \beta_{1M}Y_{Mt} + \gamma_{11}X_{1t} + \dots + \gamma_{1K}X_{Kt} + \varepsilon_{1t}; \\ Y_{2t} &= \beta_{21}Y_{1t} + \dots + \beta_{2M}Y_{Mt} + \gamma_{21}X_{1t} + \dots + \gamma_{2K}X_{Kt} + \varepsilon_{2t}; \\ &\dots\dots\dots \\ Y_{Mt} &= \beta_{M1}Y_{1t} + \dots + \beta_{MM-1}Y_{M-1,t} + \gamma_{M1}X_{1t} + \dots + \gamma_{MK}X_{Kt} + \varepsilon_{Mt}, \end{aligned} \quad (4.1)$$

де Y_1, Y_2, \dots, Y_M — ендогенні, або залежні, змінні; X_1, X_2, \dots, X_K — предетерміновані, або попередньо визначені, змінні²; $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_M$ — випадкові величини; $t = 1, 2, \dots, N$ — загальна кількість спостережень; β — невідомі параметри при ендогенних змінних; γ — невідомі параметри при попередньо визначених змінних.

Параметри β та γ ще відомі як структурні параметри, або структурні коефіцієнти, моделі.

В наведеній моделі структурні параметри представлені коефіцієнтами β , коли вони належать до ендогенних змінних, та коефіцієнтами γ , коли вони стоять біля попередньо визначених змінних. Так само ендогенні змінні позначено через Y , а екзогенні (попередньо визначені) — через X . Слід відмітити, що не всі змінні мають обов'язково з'являтися у кожному рівнянні. Іншими словами, дослідник може *априорі* накласти обмеження на параметри моделі і припустити, що деякі параметри дорівнюють нулю (наприклад $\gamma_{11} = 0$),

¹ По умовчанню вважається, що випадкові величини ε_t 's не корельовані між собою. Якщо це не виконується, Y_{t-1} корелюватиме з відповідною випадковою величиною ε_t . Таким чином, ми не можемо розглядати її як попередньо визначену.

² Одна з цих X змінних може набувати значення одиниці для отримання перетину в кожному рівнянні.

i , таким чином, змінні, відповідні цим параметрам (в даному випадку X_1), не включаються до складу регресії (в даному випадку першої).

Параметри структурної моделі відображають *прямий вплив* кожного фактора на залежну змінну. За економічним змістом вони можуть бути коефіцієнтами еластичності, граничними нормами і т. п.

Відмітимо, що кожна структурна модель має свою приведену форму, в якій ендогенні змінні виражаються лише через предетерміновані змінні системи та випадкові величини.

Приведена форма структурної моделі

Приведеною (скороченою) формою структурної моделі є модель, в якій ендогенні змінні виражені як функції лише предетермінованих змінних. Приведена форма структурної моделі (4.1) у векторному записі має відповідно вигляд:

$$Y_i = \pi_{i1}X_1 + \pi_{i2}X_2 + \dots + \pi_{ik}X_k + v_i, \quad i = 1, 2, \dots, M, \quad (4.2)$$

де π_{ij} , $i = 1, 2, \dots, M$; $j = 1, 2, \dots, K$ — невідомі параметри (коефіцієнти) приведеної моделі; v — випадкова величина.

Параметри приведеної моделі вимірюють *загальний вплив* (прямий та опосередкований) предетермінованих змінних на ендогенні змінні, в той час як структурні параметри вимірюють *прямий ефект* на ендогенну змінну від зміни фактора.

Оцінювання параметрів приведеної форми на відміну від оцінювання структурних параметрів не є складним завданням. Дійсно, як можна побачити з (4.2), ендогенні змінні в приведеній моделі залежать лише від предетермінованих змінних, які не є стохастичними. Отже, припущення про незалежність між собою випадкових величин та факторів не порушується. Відповідно, $\text{COV}(X, v) = 0$ і можна застосовувати метод найменших квадратів для знаходження невідомих параметрів приведеної форми структурної моделі³.

Можна показати, що між параметрами приведеної моделі та структурними параметрами існує чіткий взаємозв'язок, тобто *значення π є функціями структурних параметрів β та γ* . Звичайно, не завжди оцінювання приведеної моделі дозволяє отримати однозначні величини параметрів симулятивної системи рівнянь в загальному випадку (структурної моделі), тому розроблені спеціальні методи оцінювання невідомих параметрів структурних

³ Тут ми використовуємо припущення, що автокореляція залишків відсутня. В протилежному разі, оскільки в правій частині можуть з'являтися лаги ендогенних змінних, ми знову можемо отримати зміщені неконсистентні оцінки.

моделей. Проте в певних випадках (розглянемо їх нижче) з параметрів приведеної форми можна отримати значення структурних параметрів.

4.2. МЕТОДИ ОЦІНЮВАННЯ НЕВІДОМИХ ПАРАМЕТРІВ У МОДЕЛЯХ СИМУЛЬТАТИВНИХ РІВНЯНЬ

Як ми вже відмічали вище, застосування методу найменших квадратів при оцінці невідомих параметрів у системах симультаивних рівнянь призводить до зміщених та неконсистентних оцінок. Щоб уникнути цього небажаного результату, необхідно застосовувати спеціальні методи оцінювання, які давали б незміщені та консистентні оцінки параметрів. Такі методи розроблені, і їх можна звести до двох великих категорій: методи, що базуються на оцінюванні кожного окремого рівняння системи, та методи одночасного оцінювання невідомих параметрів усієї системи. Їх ще називають відповідно методами з обмеженою інформацією та методами з повною інформацією, а також *методами одного рівняння та системними методами*. Найпопулярнішими серед них є:

1. Метод отримання параметрів з приведеної форми, або метод непрямих найменших квадратів (МННК).
2. Двокроковий МНК (2МНК).
3. Метод максимальної правдоподібності з обмеженою інформацією (ММПОІ).
4. Метод змішаного оцінювання.
5. Трикроковий МНК (3МНК).
6. Метод максимальної правдоподібності з повною інформацією (ММППІ).

Перші чотири методи входять до категорії методів одного рівняння, оскільки вони можуть застосовуватись до кожного окремого рівняння системи. П'ятий та шостий методи (3МНК та ММППІ) називають системними, оскільки вони застосовуються одночасно до всіх рівнянь системи.

Як правило, застосування того або іншого методу оцінювання системи щільно пов'язане з проблемою ототожнення структурної моделі. Окреме рівняння симультаивної моделі може бути недоототожненим, точно ототожненим або переототожненим. Якщо хоча б одне рівняння системи є недоототожненим, то і вся система вважається недоототожненою і, відповідно, така система не має розв'язку або не має однозначного розв'язку. Тобто жодним методом ми не зможемо отримати оцінки невідомих параметрів моделі. Крім того, легко показати, що в даному випадку параметри структурної моделі не можна знайти на основі оцінених параметрів приведеної форми.

Ототожнена модель може бути як точно ототожненою, так і переототожненою.

Якщо всі рівняння системи є точно ототожненими, то і вся система вважається точно ототожненою, і найбільш простим методом оцінювання в даному випадку є метод непрямих найменших квадратів, хоча можна застосовувати і всі інші наведені вище методи. Зазначимо, що точно ототожнену модель ми маємо у випадку, коли можна отримати однозначну оцінку її параметрів, виходячи з оцінених параметрів приведеної форми.

Якщо хоча б одне рівняння симультативної системи є переототожненим, то і вся система є переототожненою. В такому випадку ми не можемо отримати однозначних значень структурних параметрів з параметрів приведеної форми. Іншими словами, деяким параметрам структурної моделі може відповідати множина значень. Очевидно, що отримана приведена форма може бути сумісною з різними структурними моделями, і тоді важко сказати, яку саме модель ми досліджуємо. В даному випадку для оцінювання невідомих параметрів не можна застосовувати метод непрямих найменших квадратів, але існують інші методи, які дозволяють зняти таку неоднозначність. Найбільш поширеними серед таких методів є 2МНК — двокроковий метод найменших квадратів (Two-stage Least squares, TSLS), 3МНК — трикроковий МНК та метод максимальної правдоподібності повної інформації (ММПІІ). Нагадаємо, що 3МНК та ММПІІ належать до системних методів оцінки параметрів системи.

Існують певні критерії, які допомагають визначити доцільність того чи іншого методу. Такі критерії дають, зокрема, правила ототожнення симультативних моделей, які базуються на порядку матриці факторів⁴.

4.3. ПЕРЕВІРКА СИСТЕМИ СИМУЛЬТАТИВНИХ РІВНЯНЬ НА ОТОТОЖНЕНІСТЬ

Сформулюємо правило перевірки на ототожненість у загальному випадку.

Нехай нам необхідно перевірити ототожненість системи симультативних рівнянь

$$Y_{1t} = \beta_{12}Y_{2t} + \dots + \beta_{1M}Y_{Mt} + \gamma_{11}X_{1t} + \dots + \gamma_{1K}X_{Kt} + \varepsilon_{1t};$$

$$Y_{2t} = \beta_{21}Y_{1t} + \dots + \beta_{2M}Y_{Mt} + \gamma_{21}X_{1t} + \dots + \gamma_{2K}X_{Kt} + \varepsilon_{2t};$$

.....

$$Y_{Mt} = \beta_{M1}Y_{1t} + \dots + \beta_{MM-1}Y_{M-1,t} + \gamma_{M1}X_{1t} + \dots + \gamma_{MK}X_{Kt} + \varepsilon_{Mt}, \quad (4.3)$$

⁴ Фактором може бути будь-яка предетермінована та ендогенна змінна.

де Y_1, Y_2, \dots, Y_M — ендogenous, або залежні, змінні; X_1, X_2, \dots, X_K — предетерміновані, або попередньо визначені, змінні; M — кількість ендogenous змінних в симультативній моделі; K — кількість попередньо визначених змінних в моделі. Крім того, введемо додаткові позначення: k та m — відповідно кількість ендogenous та попередньо визначених змінних в окремому рівнянні системи.

4.3.1. Умова порядку

З урахуванням введених позначень сформулюємо необхідну (але не достатню) умову ототожнення, яка має назву «Умова порядку» і може бути визначена двома різними, але еквівалентними способами:

Визначення 1. Для того, щоб окреме рівняння симультативної моделі було ототожненим, з нього має бути виключено щонайменше $(M - 1)$ змінних (ендogenous та предетермінованих), які з'являються в загальній моделі. Якщо виключено рівно $(M - 1)$ змінних, то рівняння, що розглядається, буде точно ототожненим. Якщо виключено більше ніж $(M - 1)$, то воно буде переототожненим.

Визначення 2. Для ототожнення рівняння симультативної моделі кількість предетермінованих змінних, виключених з цього рівняння, має бути не меншою за кількість ендogenous змінних мінус одиниця, включених в рівняння, тобто

$$K - k \geq m - 1. \quad (4.4)$$

Якщо $K - k = m - 1$, рівняння є точно ототожненим, але якщо $K - k > m - 1$, воно є переототожненим.

Як показують наведені визначення, ототожненість рівнянь симультативних моделей можлива у випадку, коли відповідні рівняння приведеної форми виключають одну чи більше змінних, які наявні в якому-небудь іншому рівнянні моделі, але відсутні в досліджуваному рівнянні.

4.3.2. Рангова умова ототожнення⁵

Умова порядку, яка обговорювалась раніше, є необхідною, але не достатньою умовою ототожнення. Тобто може статися так, що навіть при її виконанні,

⁵ Термін «ранг» стосується рангу матриці і визначається найбільшим порядком квадратної матриці (що міститься у даній матриці), детермінант якої ненульовий. З іншого боку, ранг матриці — це найбільше число лінійно незалежних рядків чи колонок цієї матриці.

рівняння все ж таки залишиться неототожненим. Наприклад, навіть якщо умова порядку $K - k \geq t - 1$ виконана, рівняння може бути неототожненим, тому що предетерміновані змінні, які виключені з нього, але присутні в моделі, можуть бути залежними між собою. Через це відповідність між структурними параметрами (β та γ) і параметрами приведеної форми (π) не зберігається. Тобто ми не можемо знайти структурні параметри з оцінених параметрів приведеної форми. Звичайно, потрібно мати як достатню, так і необхідну умову ототожнення. Необхідність та достатність ототожнення гарантується умовою рангу. Рангова умова ототожнення формулюється таким чином:

В симультивній моделі, яка містить M рівнянь з M ендогенними змінними, рівняння буде ототожненим тоді і лише тоді, коли щонайменше один ненульовий визначник порядку $(M - 1)(M - 1)$ може бути створений з коефіцієнтів змінних (як ендогенних, так і предетермінованих), виключених з рівняння, що розглядається, але включених в інші рівняння моделі.

Для спрощення перевірку за ранговою умовою можна розбити на такі кроки:

1. Записати систему симультивних рівнянь в табличній формі.
2. Викреслити коефіцієнти рядка, в якому з'являється рівняння, що розглядається.
3. Викреслити стовпці, що відповідають ненульовим коефіцієнтам.
4. Вхід ліворуч таблиці дасть тільки коефіцієнти змінних, включених в систему, але не в рівняння, що розглядається. Завдяки цьому отримаємо всі можливі форми матриць A -порядку $(M - 1)$ і відповідні визначники. Якщо можна знайти щонайменше один ненульовий визначник, рівняння ототожене. В даному випадку ранг матриці, скажімо A , точно дорівнює $(M - 1)$. Якщо всі можливі $(M - 1)(M - 1)$ -визначники дорівнюють нулю, ранг матриці A менший $(M - 1)$ свідчить про те, що відповідне рівняння неототожене.

На базі умов порядку та умов рангу можна сформулювати загальні принципи ототожнення окремого рівняння структурної моделі, яка складається з M симультивних рівнянь:

1. Якщо $K - k > t - 1$ і ранг матриці A буде дорівнювати $(M - 1)$, то відповідне рівняння переототожене.
2. Якщо $K - k = t - 1$ і ранг матриці A дорівнює $(M - 1)$, рівняння точно ототожене.
3. Якщо $K - k \geq t - 1$ і ранг матриці A менший ніж $(M - 1)$, рівняння неототожене.
4. Якщо $K - k < t - 1$, рівняння неототожене. Ранг матриці A в даному випадку менший за $(M - 1)$.

Коли йдеться про ототожнення, ми маємо на увазі точне ототожнення або переототожнення. Для неототожнених рівнянь, на жаль, невідомо, наскільки вичерпно має бути додаткова інформація для оцінки структурних параметрів.

4.4. ОЦІНКА ТОЧНО ОТОТОЖНЕНОГО РІВНЯННЯ: МЕТОД НЕПРЯМИХ НАЙМЕНШИХ КВАДРАТІВ

Для точно ототоженого структурного рівняння можна отримати структурні параметри з МНК-оцінок параметрів приведеної форми методом непрямих найменших квадратів (МННК). Оцінка параметрів за цим методом умовно розбивається на три етапи.

1. Спочатку отримуємо рівняння приведеної форми, тобто виразимо кожну залежну змінну як функцію виключно предетермінованих змінних (екзогенних та лагових ендогенних) та випадкових величин.
2. Оцінимо параметри кожного рівняння приведеної форми методом найменших квадратів. Це можливо, оскільки пояснюючі змінні в даних рівняннях є предетермінованими (нестохастичними), а отже, не корелюють з випадковими величинами.
3. Знайдемо оцінки структурних параметрів симультазивної моделі з оцінених на другому етапі параметрів приведеної форми. Якщо рівняння точно ототожене, існує однозначна взаємна відповідність між структурними параметрами та параметрами приведеної форми.

Назва методу «непрямі найменші квадрати» (ННК) пов'язана з тим фактом, що структурні параметри отримуються не прямо, а з МНК-оцінок параметрів приведеної форми.

Властивості МННК-оцінок

Ми знаємо, що оцінки параметрів приведеної форми є, згідно з відповідними припущеннями, найкращими незміщеними лінійними оцінками, тобто ефективними. Чи переносяться ці властивості на МННК-оцінки? Можна показати, що МННК-оцінки також успадковують всі властивості МНК-оцінок, але лише асимптотично, тобто вони є консистентними та асимптотично ефективними.

4.5. ОЦІНКА ПЕРЕОТОТОЖНЕНОГО РІВНЯННЯ: ДВОКРОКОВИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ

Для ілюстрації цього методу розглянемо приклад. Нехай ми маємо модель:

$$\text{функція доходу: } Y_{1t} = \beta_{10} + \beta_{11}Y_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + \varepsilon_{1t}; \quad (4.5)$$

$$\text{функція пропозиції грошей: } Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} + \varepsilon_{2t}, \quad (4.6)$$

де Y_1 — дохід, Y_2 — запас грошей, X_1 — інвестиційні витрати, X_2 — витрати уряду на товари та послуги. Змінні X_1 та X_2 є екзогенними, а Y_1 та Y_2 — ендогенними.

Рівняння доходу, яке ми розглядаємо, показує, що дохід визначається пропозицією грошей, інвестиційними витратами та витратами уряду. Рівняння пропозиції грошей показує, що запас грошей визначається Національним банком на базі рівня доходів. Очевидно, що ми маємо симультативну модель.

Застосовуючи умову порядку для її ототожнення, можна побачити, що рівняння доходу є неототожненим, тоді як рівняння пропозиції грошей є переототожненим. Переототожнена функція пропозиції грошей не може бути оцінена за допомогою МНК, тому що існують два значення для оцінки β_{21} (перевірте це самостійно). Потрібно застосовувати інші методи.

Якщо застосувати МНК для оцінки невідомих параметрів рівняння пропозиції грошей, отримані оцінки будуть неконсистентними через кореляцію між змінною Y_1 та випадковою величиною ε_2 . Припустимо, що ми знайшли певний показник, який сильно корелює з Y_1 , але не корелює з ε_2 . Такий показник називається *інструментальною змінною*, або *інструментом*. Якщо його можна знайти, то МНК можна прямо застосовувати для оцінки функції грошової пропозиції. Але як можна отримати дану інструментальну змінну?

Г. Тейл та Р. Басман запропонували спеціальний двокроковий метод найменших квадратів, який використовує інструментальні змінні та дозволяє однозначно оцінити структурні параметри переототожнених рівнянь [10, 11]. З назви видно, що метод складається з двокрокового застосування методу найменших квадратів, тобто з декількох етапів:

1. Щоб позбавитись кореляції між Y_1 і ε_2 , побудуємо спочатку регресійне рівняння залежності Y_1 від всіх предетермінованих змінних системи:

$$Y_{1t} = \hat{\pi}_0 + \hat{\pi}_1 X_{1t} + \hat{\pi}_2 X_{2t} + e_t, \quad (4.7)$$

де e_t є помилками.

Невідомі параметри рівняння (4.7) можна оцінити за допомогою МНК, отримаємо:

$$\hat{Y}_{1t} = \hat{\pi}_0 + \hat{\pi}_1 X_{1t} + \hat{\pi}_2 X_{2t}. \quad (4.8)$$

Рівняння (4.7) є не чим іншим, як регресійною моделлю приведеної форми, тому що в правій частині з'являються лише предетерміновані (екзогенні) змінні. Його ще можна записати у вигляді:

$$Y_{1t} = \hat{Y}_{1t} + e_t, \quad (4.9)$$

отже, змінна Y_1 складається з двох частин: \hat{Y}_{1t} — прогнозової величини та випадкової компоненти (помилки) e_t . Виходячи з класичних припущень методу найменших квадратів, \hat{Y}_{1t} та e_t не корелюють між собою.

2. Переототожене рівняння пропозиції грошей можна записати таким чином:

$$Y_{2t} = \beta_{20} + \beta_{21}(\hat{Y}_{1t} + e_t) + \varepsilon_{2t} = \beta_{20} + \beta_{21}\hat{Y}_{1t} + \varepsilon_t^*, \quad (4.10)$$

де $\varepsilon_t^* = \varepsilon_{2t} + \beta_{21}e_t$.

(4.10) оцінюються методом найменших квадратів, і отримані оцінки будуть консистентними та асимптотично ефективними.

Порівнюючи (4.10) з (4.7), бачимо, що зовні ці рівняння дуже схожі, єдина відмінність полягає в тому, що Y_1 замінено на його оцінку \hat{Y}_{1t} . В чому полягає перевага рівняння у вигляді (4.10)? Хоча Y_1 в початковому рівнянні грошової пропозиції корелює із випадковими величинами $\varepsilon_2, \hat{Y}_{1t}$ в (4.10) не корелює з ε_t^* (у випадку, коли розмір моделі зростає пропорційно). В результаті МНК можна застосувати до (4.10) та отримати незміщені консистентні оцінки параметрів функції пропозиції грошей.

Основна ідея двокрокового методу найменших квадратів полягає в очищенні змінної Y_1 від впливу випадкової величини ε_2 . Це виконується за допомогою регресії приведеної форми Y_1 від всіх предетермінованих змінних в системі (крок 1).

Отримавши оцінки \hat{Y}_{1t} , заміщуємо в початковому рівнянні Y_1 оцінками \hat{Y}_{1t} , очищеними від ендогенності, та переходимо до модифікованого рівняння, в якому знаходимо невідомі параметри методом найменших квадратів (крок 2). Отримані оцінки є консистентними, тобто прямують до справжніх значень при зростанні кількості спостережень.

Характерні особливості двокрокового методу найменших квадратів

Можна відмітити такі особливості методу 2МНК:

1. Метод можна застосувати для оцінки кожного окремого рівняння системи без прямого врахування інших. Отже, для економетричних моделей, що складаються з великої кількості рівнянь, метод 2МНК є дуже економним, тому він широко використовується на практиці.
2. На відміну від МНК, який дає множину значень оцінок параметрів у переототожнених рівняннях, 2МНК забезпечує однозначну оцінку структурних параметрів моделі.
3. Хоча метод 2МНК був спеціально розроблений для переототожнених рівнянь, його також можна застосовувати до точно ототожнених рівнянь. В цьому випадку МНК та 2МНК дадуть ідентичні оцінки.

Зауважимо, що якщо 2МНК застосовувати до регресійних рівнянь, в яких ендогенність/симультаивність відсутня, тобто серед пояснюючих змінних відсутні ендогенні змінні, то ми отримуємо консистентні, але неефективні оцінки.

Якщо в регресіях приведеної форми (крок 1) значення R^2 достатньо великі, скажімо, перевищують 0.8, то значення оцінок, отримані методами МНК та 2МНК, будуть дуже близькими. Однак якщо значення коефіцієнтів детермінації (R^2) в регресійних рівняннях, що оцінюються на першому кроці, є низькими, то 2МНК будуть відрізнятися від МНК-оцінок. При цьому зрозуміло, що при низькому значенні коефіцієнтів детермінації на першому кроці 2МНК інструментальна змінна \hat{Y} , знайдена з приведеної форми, не буде сильно корелювати зі значенням Y , тобто не є адекватним інструментом⁶. Це може призводити до значних проблем в малих вибірках.

4.6. РЕКУРСИВНІ МОДЕЛІ

Особливим випадком симультаивних моделей є рекурсивні моделі. Модель називається рекурсивною, якщо її структурні рівняння можна впорядкувати таким чином, що перше містить у правій стороні лише предетерміновані змінні; друге містить у правій стороні предетерміновані змінні та першу ендогенну змінну і так далі. Наприклад:

$$Y_1 = f(X_1, X_2, \dots, X_k; \varepsilon_1);$$

$$Y_2 = f(X_1, X_2, \dots, X_k; Y_1; \varepsilon_2);$$

$$Y_3 = f(X_1, X_2, \dots, X_k; Y_1, Y_2; \varepsilon_3);$$

.....

$$Y_M = f(X_1, X_2, \dots, X_k; Y_1, Y_2, \dots, Y_{M-1}; \varepsilon_M). \quad (4.11)$$

Однією з особливостей рекурсивної моделі є можливість оцінити її рівняння по одному на кожному кроці за допомогою МНК. Щоб проілюструвати це, перепишемо рекурсивну модель (4.11) як систему лінійних регресійних рівнянь:

⁶ Іншими словами, в даному випадку знайдені інструментальні змінні \hat{Y} будуть невідповідними фактичним значенням Y .

$$Y_1 = \gamma_{10} + \gamma_{11}X_1 + \gamma_{12}X_2 + \dots + \gamma_{1K}X_K + \varepsilon_1;$$

$$Y_2 = \gamma_{20} + \gamma_{21}X_1 + \gamma_{22}X_2 + \dots + \gamma_{2K}X_K + \beta_{21}Y_1 + \varepsilon_2;$$

$$Y_3 = \gamma_{30} + \gamma_{31}X_1 + \gamma_{32}X_2 + \dots + \gamma_{3K}X_K + \beta_{32}Y_1 + \beta_{31}Y_2 + \varepsilon_3;$$

.....

$$Y_M = \gamma_{M0} + \gamma_{M1}X_1 + \gamma_{M2}X_2 + \dots + \gamma_{MK}X_K + \beta_{M1}Y_1 + \beta_{M2}Y_2 + \dots + \varepsilon_M. \quad (4.12)$$

Маючи значення екзогенних змінних, ми можемо застосувати МНК до першого рівняння системи (4.12) і отримати оцінене значення Y для першої ендогенної змінної. Потім застосовуємо обчислене Y_1 (оцінка \hat{Y}_1) як пояснюючу змінну в другому рівнянні. Знову можна застосувати МНК, оскільки всі змінні незалежні від випадкової величини ε_2 . Зокрема, X та ε_2 незалежні за означенням. \hat{Y}_1 і ε_2 також незалежні, оскільки \hat{Y}_1 ортогональна до ε_2 ⁷ та за класичними припущеннями змінні ε_1 та ε_2 є незалежними. Таким чином, \hat{Y}_1 та ε_2 є також незалежними.

Рекурсивні системи також називають *трикутними системами*, оскільки коефіцієнти при ендогенних змінних (β) утворюють трикутну матрицю: головна діагональ цієї матриці містить одиниці, а над головною діагоналлю елементи матриці дорівнюють нулю.

Для того, щоб визначити, чи є модель рекурсивною, достатньо дослідити форму матриці коефіцієнтів β . Якщо вона трикутна, то система рекурсивна. Рекурсивність системи дозволяє нам застосувати для оцінки невідомих параметрів кожного окремого рівняння метод МНК.

Основні висновки про симультативні моделі

Головною особливістю системи симультативних рівнянь є використання залежних (ендогенних) змінних у ролі пояснюючих змінних системи. Така змінна є стохастичною і, отже, корелює з випадковою величиною того рівняння, у якому з'являється як пояснююча змінна. В такому випадку для оцінки невідомих параметрів не можна застосувати класичний метод найменших квадратів, бо отримані на основі цього методу оцінки будуть зміщеними незалежно від кількості спостережень. Тому для оцінювання симультативних рівнянь розроблені спеціальні методи. Їх застосування пов'язане з проблемою ототожнення.

Під проблемою ототожнення мається на увазі можливість отримання кількісних оцінок структурних параметрів рівнянь загальної симультативної

⁷ Тобто коваріація між цими випадковими величинами дорівнює нулю.

моделі з оцінених параметрів приведеної форми. Якщо це можна зробити, то рівняння в симультаивній системі є ототожненим, якщо ні — то недоототожненим.

Ототожене рівняння може бути як точно ототожненим, так і переототожненим. В першому випадку існують однозначні значення структурних параметрів, тоді як в другому структурні параметри можуть приймати більше ніж одне значення.

Проблема ототожнення виникає тому, що одна й та ж інформація може бути сумісною з різними структурними параметрами симультаивних рівнянь, а отже, з різними моделями.

Для перевірки ототожнення рівняння застосовують умови рангу та порядку. Умова рангу є необхідною та достатньою умовою. Вона дає відповідь на запитання, чи є рівняння системи, що розглядається, ототожненим, в той час як умова порядку, що є лише достатньою, а не необхідною умовою ототожнення, дає відповідь на запитання, чи буде ототожене рівняння системи точно ототожненим, чи переототожненим. Якщо умова рангу виконана, умова порядку також виконана, що невірно навпаки.

Припустимо, що рівняння в симультаивній моделі ототожене. Існує декілька методів для оцінки його параметрів. Вони розділяються на дві категорії: методи одного рівняння та системні методи. Найпопулярнішими методами оцінювання невідомих параметрів системи є методи одного рівняння. Серед найчастіше вживаних методів оцінки невідомих параметрів окремого рівняння є МНК, МННК та 2МНК. Хоча МНК в загальному випадку не підходить для оцінювання параметрів симультаивних моделей, його можна застосовувати в особливих випадках, а саме у випадку рекурсивних моделей. Метод ННК підходить для точно ототожнених рівнянь. В цьому методі МНК застосовується до рівнянь приведеної форми, а значення структурних параметрів отримуються з параметрів приведеної форми. Метод 2МНК був спеціально розроблений для переототожнених рівнянь, хоча його також можна застосувати і у випадку точної ототожненості. При цьому оцінки, отримані за допомогою МННК та 2МНК, будуть однакові. Основна ідея методу 2МНК полягає в заміщенні стохастичної ендогенної пояснюючої змінної такою змінною (інструментом), що сильно корелює з ендогенною змінною та не корелює з випадковою величиною. Як інструмент використовується зазвичай оцінена змінна, що є функцією предетермінованих (нестохастичних) змінних у моделі. Двокроковий метод найменших квадратів ще називають методом інструментальних змінних.

Характерною рисою обох методів є те, що отримані за ними оцінки є консистентними та асимптотично ефективними. В умовах невеликої вибірки оцінки можуть не задовольняти властивостям незміщеності та мінімальної дисперсії. Тоді отримані результати повинні інтерпретуватись обережно.

4.7. ЗАСТОСУВАННЯ МОДЕЛЕЙ СИМУЛЬТАТИВНИХ РІВНЯНЬ НА ПРАКТИЦІ: СПРОЩЕНА МАКРОМОДЕЛЬ УКРАЇНИ

Для того, щоб проілюструвати практичне застосування симультативних систем рівнянь, спробуємо розробити та оцінити макроеконометричну модель, яка є спрощеним аналогом IS-LM-моделі для України.

Відмітимо, що в розвинутих країнах, особливо в США, макроеконометричні моделі часто використовуються для визначення політики і відтворення її можливих сценаріїв. Наприклад, економетричні моделі DRI, BEA, FRB/MPS, Taylor, Fair часто застосовуються для передбачення і оцінки певних заходів політики. Як правило, вони ґрунтуються на різних припущеннях. Наприклад, модель Тейлора застосовує теорію раціональних очікувань, тоді як модель DRI спирається на кейнсіанську теорію. В Україні макроеконометричні моделі розробляються в Інституті прогнозування НАН України, в CASE, а також розроблялись у Гарвардському інституті міжнародного розвитку.

Нагадаємо, що оцінка таких моделей вимагає спеціальних методів. Економетрична теорія МНК базується на припущенні, що всі змінні з правого боку рівняння не корелюють з випадковими величинами. Ця умова автоматично задовольняється, коли змінні є екзогенними. Проте в системі рівнянь деякі змінні визначаються всередині системи. Такі змінні називаються ендогенними.

Можна показати, що ендогенні змінні часто корелюють з випадковими величинами. В такому випадку оцінки МНК є зміщеними. Щоб вирішити цю проблему, для оцінювання систем симультативних рівнянь слід використовувати двокроковий (2МНК) або трикроковий метод найменших квадратів (3МНК), або якийсь інший відповідний метод. Сутність цих підходів — знайти інструменти (замінники) для ендогенних змінних. Ці інструменти мають корелювати з ендогенними змінними, але в той же час не корелювати з випадковими величинами.

Метою даного завдання є отримання практичних навичок в розробці найпростіших макроеконометричних моделей та оволодіння методами оцінювання системи симультативних рівнянь.

Завдання 4.1.

Необхідно оцінити спрощену макромодель України та розробити на її основі декілька можливих сценаріїв економічної політики. Вхідні дані для розрахунків наведені в таблиці 4.15.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. Зауважимо, що в процесі виконання цього завдання ми будемо використовувати пакет прикладних програм E.VIEWS.

Крок 1. Визначення ендогенних та екзогенних змінних моделі.

Детально прокоментуйте запропоновану модель. Визначте ендогенні та екзогенні змінні моделі.

Крок 2. Перевірка системи на ототожненість.

За умовою рангу та порядку перевірте систему на ототожненість. Які висновки ви можете зробити? Поясніть їх детально.

Крок 3. Оцінка, тестування та визначення специфікації кожного окремого рівняння системи.

Оцініть кожне окреме рівняння системи. Протестуйте на відповідність класичним припущенням. Детально опишіть ваші результати по кожному окремому рівнянню системи. Коротко поясніть ваші висновки.

Крок 4. Оцінка параметрів системи методом 2МНК (TSLS) для усунення ендогенності.

Розрахуйте параметри системи двокроковим методом найменших квадратів та проінтерпретуйте детально отримані результати.

Крок 5. Тест Хаусмана для оцінки ефективності застосування 2МНК (TSLS).

Проведіть тест Хаусмана для підтвердження гіпотези, що у вашому випадку застосування двокрокового методу є дійсно ефективним для зняття проблеми ендогенності. Прокоментуйте ваші результати та зробіть остаточний висновок щодо ефективності застосування 2МНК.

Зауважимо, що в пакеті E.VIEWS немає прямого тесту Хаусмана, тому потрібно написати цей тест власноруч, використовуючи команди E.VIEWS. В коментарі до цього завдання ми наводимо тест Хаусмана в командах E.VIEWS, написаний нами.

Крок 6. Оцінка параметрів системи методом 3МНК (3SLS).

Розрахуйте параметри системи трикроковим методом найменших квадратів та проінтерпретуйте детально отримані результати. Порівняйте результати оцінювання системи двокроковим та трикроковим методами найменших квадратів. Чи є між ними різниця? Прокоментуйте ваші висновки.

Крок 7. Особливості прогнозування на основі симультазивних моделей.

Проведіть прогнозування на основі розробленої та оціненої симультазивної моделі. Проаналізуйте та детально прокоментуйте отримані вами результати.

Крок 8. Оцінка прогнозової якості моделі.

Перевірте прогнозну якість моделі на основі відомих вам критеріїв якості. Які висновки щодо якості моделі ви можете зробити? Прокоментуйте їх.

Крок 9. Розробка та аналіз сценаріїв на основі симультативних моделей.

На основі розробленої моделі проаналізуємо, що трапиться, якщо уряд України:

- 1) підвищить державні витрати та доходи на 5 % у кожному кварталі 1997—1999 рр.;
- 2) підвищить монетарну базу на 5 % у кожному кварталі 1997—1999 рр.;
- 3) застосує обидва заходи.

Детально прокоментуйте отримані результати. Які висновки ви можете зробити? Обґрунтуйте їх.

Розв'язок завдання та розширений коментар оцінювання за допомогою системи симультативних рівнянь спрощеної макромоделі України

Для того, щоб прокоментувати оцінювання макромоделі, наведемо її специфікацію та введемо певні позначення.

Специфікація моделі

Наведемо загальний вигляд розробленої нами спрощеної макроеконометричної моделі, яка ґрунтується на традиційній IS-LM-моделі:

1. $RCON_t = f_1(RDI_t, R_t, RCON_{t-1}, RCON_{t-4});$
2. $RINV_t = f_2(RGDP_t, RGDP_{t-1}, R_{t-2}, RINV_{t-1}, RINV_{t-4}, t);$
3. $R_t = f_3(RM_2_t, RGDP_t, DLOG\{CPI_t\}, R_{t-1});$
4. $M2_t = f_4(MB_t, [R_t - RNBU_t]);$
5. $RGDP_t = RCON_t + RINV_t + RG_t + NX_t;$
6. $RDI_t = RGDP_t - RTAX_t;$
7. $RM_2 = M2_t / CPI_t,$ (4.3)

де t — часовий період; $RCON$ — реальне споживання, млрд грн. 1996 (джерело: CASE); RDI — реальний дохід у розпорядженні, млрд грн. 1996 (CASE); R — комерційна (ринкова) процентна ставка, щомісячний % (НБУ); $RINV$ — реальні інвестиції, млрд грн. 1996 (CASE); $RGDP$ — реальний ВВП, млрд грн.

1996 (CASE); $RM2$ — реальні грошові запаси M2, млрд грн. 1996 (власні розрахунки); CPI — індекс споживчих цін, середина 1996=100 (CASE/UEPLAC); $M2$ — грошовий агрегат M2, млрд грн (НБУ); MB — грошова база, млрд грн. (НБУ); $RNBU$ — ставка рефінансування НБУ, місячний % (НБУ); RG — реальне державне споживання, млрд грн. 1996 (CASE); NX — реальний чистий експорт, млрд грн. 1996 (CASE); $RTAX$ — реальні державні доходи, млрд грн. 1996 (CASE).

Вхідні дані наведені в додатку 4.

Крок 1. Визначення ендогенних та екзогенних змінних моделі.

В цій моделі ендогенними змінними є

- реальне споживання ($RCON$),
- реальні інвестиції ($RINV$),
- ринкова процентна ставка (R),
- грошовий агрегат M2 ($M2$),
- реальний ВВП ($RGDP$),
- реальний грошовий агрегат M2 ($RM2$),
- реальний дохід в розпорядженні (RDI).

За припущенням екзогенними змінними є

- реальне державне споживання (RG),
- реальні державні доходи ($RTAX$),
- рівень цін (індекс споживчих цін, CPI),
- грошова база (MB),
- ставка рефінансування НБУ ($RNBU$),
- реальний чистий експорт (NX).

Деякі змінні є предетерміновані. Наприклад, в рівнянні (4.3) одна із змінних є лаговою, а саме реальне споживання з лагом один. Реальне споживання є ендогенною змінною, проте реальне споживання з лагом є фіксованою (іншими словами, попередньо визначеною або предетермінованою) змінною в період t . Загальна кількість екзогенних та предетермінованих змінних дорівнює 13.

Крок 2. Перевірка системи на ототожненість.

З'ясуємо, чи є система симульативних рівнянь (4.1) ототожненою (ідентифікованою), чи ні. Ми повинні, зокрема, перевірити умови порядку і рангу.

Перевіримо спочатку умову порядку. Згідно з [10, с. 665], в моделі з M симульативних рівнянь для того, щоб рівняння було ототожненим, кількість попередньо визначених змінних, виключених з рівняння, повинна бути не меншою, ніж кількість ендогенних змінних, включених до рівняння, мінус одиниця, тобто $K - k \geq m - 1$.

Якщо $K - k = t - 1$, рівняння є точно ототожненим, якщо $K - k > t - 1$ — переототожненим, а якщо $K - k < t - 1$ — недоототожненим.

Розглянемо перше рівняння, яке описує реальне споживання. В цьому рівнянні є три ендогенних змінних: $RCON$, RDI та R і три предетерміновані: $DUMMY^{98:02}$, $RCON_{t-1}$ та $RCON_{t-4}$. Отже, t та k дорівнюють 3. Кількість попередньо визначених змінних рівна 13. Таким чином, $K = 13$. Очевидно, що $K - k = 10$ більше, ніж $2 (= t - 1 = 3 - 1)$. Отже, рівняння переототожене.

Умова порядку як необхідна умова ототожнення на практиці майже завжди є достатньою для визначення ототожненості рівняння в тому розумінні, що в практичних випадках умови порядку та рангу дають один і той же результат. В даному випадку ми також перевіримо умову рангу, яка є достатньою для ототожнення. Умова рангу не вказує, чи є рівняння переототожненим, чи точно ототожненим, тож в будь-якому разі ми повинні використати умову порядку, яка дає можливість визначити, чи є система переототоженою, чи точно ототоженою.

Щоб перевірити умову рангу, спочатку побудуємо таблицю (4.1) для всіх регресійних (behavioral) рівнянь моделі (4.3). Зауважимо, що ми маємо 4 регресійних рівняння, інші три є тотожностями, які не потребують перевірки на ототожненість.

Таблиця 4.1.

Рівняння	$RCON_t$	$RCON_{t-1}$	$RCON_{t-4}$	RDI	R_t	R_{t-1}	R_{t-2}	$RINV_t$	$RINV_{t-1}$	$RINV_{t-4}$	$RGDP_t$	$RGDP_{t-1}$	$RM2$	CPI	$M2$	MB	$RNBUS$	RG	NX	$FTUX$	t	$D98:03$
1	+	+	+	+	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+
2	0	0	0	0	0	+	0	0	+	+	+	+	0	0	0	0	0	0	0	0	+	0
3	0	0	0	0	+	+	0	0	0	0	+	0	+	+	0	0	0	0	0	0	+	+
4	0	0	0	0	+	0	0	0	0	0	0	0	0	0	+	+	+	0	0	0	0	0

Перше рівняння системи є ототожненим, тому що відповідна матриця (таблиця 4.2) має ранг 3.

Таблиця 4.2.

$$P1 = \begin{vmatrix} + & + & + & + & + & 0 & 0 & 0 & 0 & 0 & + \\ + & 0 & 0 & + & 0 & + & + & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & + & + & + & 0 \end{vmatrix}$$

Ранг (P1) = 3.

Те ж саме вірно для *другого рівняння* системи, яке також є ототожненим.

Таблиця 4.3.

$$P2 = \begin{vmatrix} + & + & + & + & + & 0 & 0 & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & + & + & + & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & + & 0 & 0 & + & + & + & 0 \end{vmatrix}$$

Ранг (P2) = 3.

Третє рівняння

Таблиця 4.4.

$$P3 = \begin{vmatrix} + & + & + & + & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & + & + & + & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & + & + & + \end{vmatrix}$$

Ранг (P3) = 3.

Четверте рівняння

Таблиця 4.5.

$$P4 = \begin{vmatrix} + & + & + & + & 0 & 0 & 0 & 0 & 0 & 0 & 0 & + \\ 0 & 0 & 0 & 0 & + & + & + & + & + & 0 & 0 & + \\ 0 & 0 & 0 & 0 & + & 0 & 0 & + & 0 & + & + & + \end{vmatrix}$$

Ранг (P4) = 3.

Оскільки кожне рівняння системи є ототожненим, то і вся система є ототожненою. Далі за умовою порядку необхідно визначити, чи буде кожне окреме рівняння точно ототожненим, чи переототожненим. Відповідно до результатів нашої попередньої перевірки за умовою порядку перше рівняння є переототожненим. Самостійно зробіть висновок щодо інших рівнянь системи.

Крок 3. Оцінка, тестування та визначення специфікації кожного окремого рівняння системи.

Як правило, симульативна модель оцінюється в декілька кроків.

По-перше, кожне окреме регресійне рівняння оцінюється за допомогою МНК, щоб перевірити його здатність пояснювати зв'язки між змінними. Якщо регресія їх не пояснює (зокрема має низький коефіцієнт детермінації R^2), немає сенсу використовувати її в моделі. Треба бути особливо уважним при розробці специфікації окремих рівнянь, оскільки неправильна специ-

фікація одного з них може призвести до великих помилок в моделі в цілому. Коротко суть цього зауваження полягає в тому, що визначати специфікацію кожного рівняння треба дуже обережно, щоб отримати значимі результати і дієздатну модель. Проте на цьому етапі не обов'язково, щоб усі змінні були статистично значимими, оскільки ендогенність змінних у правій частині може викривити результати.

По-друге, попередні рівняння¹ об'єднуються в систему рівнянь і оцінюються одночасно. Тотожності включати в систему не слід, оскільки вони оцінки не потребують. Обраний метод (наприклад 2МНК або 3МНК) повинен вирішити проблему ендогенності.

По-третьє, коли модель оцінена, її тестують на те, як вона в цілому відтворює справжні часові ряди. Якщо оцінені значення є досить близькі до реальних, вважають, що модель задовільна. На практиці модель рідко починає працювати належним чином з першої спроби. Зазвичай вона переглядається багато разів. Наприклад, може бути змінена специфікація регресійного рівняння. Часто трапляється, що окреме рівняння добре описує часовий ряд, коли використовується окремо, але в контексті моделі дає погані результати (наприклад зміщення). На жаль, не існує формальних критеріїв щодо вибору специфікації, розробка її є швидше мистецтвом, ніж наукою, і майстерність приходить з досвідом.

Прокоментуємо кожне окреме рівняння нашої моделі.

Крок 3.1. Споживання.

Кейнсіанська економічна теорія визначає споживання як функцію від доходу, що залишається в розпорядженні домогосподарств. У кейнсіанській функції споживання лінійно залежить від автономного споживання (C_0) та доходу в розпорядженні (YD): $C=C_0+cYD$. Проте дослідження в теорії споживання, такі як теорія постійного доходу й гіпотеза життєвого циклу (PI-LCH), припускають, що поточний дохід в розпорядженні мало впливає на споживання.

Головна ідея цих теорій полягає в тому, що домогосподарства вирівнюють своє споживання в часі, тобто формують свої очікування щодо майбутнього доходу і потім споживають (витрачають дохід) на основі цих очікувань. З одного боку, якщо домогосподарства передбачають стабільний (постійний) дохід, вони витратять велику частку його, оскільки впевнені, що матимуть цей дохід у наступному періоді. З другого — мінливий дохід змушує домогосподарства заощаджувати, щоб компенсувати можливі майбутні втрати доходу і споживати приблизно таку ж кількість товарів і послуг. Згідно з PI-

¹ В даному контексті «попередні рівняння» потрібно розуміти як рівняння з попередньою специфікацією.

ЛСН, споживання з лагом, яке об'єднує всю доступну інформацію (історичну і припущення), є найкращим прогнозом поточного споживання [11].

Емпіричні дослідження підтверджують теоретичне припущення про витрачання великої частки постійного доходу і відносно малої частки тимчасового (непостійного) доходу. Р. Холл і Ф. Мішкін показують у своїх роботах [12, 13]: «Споживання не пристосовується у той самий механічний спосіб до кожної зміни доходів. Замість цього споживачі оцінюють джерело зміни доходу та діють лише по тих сигналах, що означають кардинальні зміни в їхньому економічному становищі». Домогосподарства окремо оцінюють частку постійного доходу в їхньому загальному доході, і така оцінка зазвичай робиться на основі всієї доступної інформації. PI-LCH — мікроекономічне обґрунтування теорії споживання на макрорівні. Р. Холл [12] показав, що *найкращим прогнозом майбутнього споживання є споживання попередніх періодів*².

Отже, PI-LCH вважає, що гранична схильність до споживання (MPC) є вищою для постійного доходу й нижчою для тимчасового. На жаль, PI-LCH-теорія має деякі обмеження. Найважливішим з них, пов'язаним з домогосподарствами, є обмеження у вирівнюванні споживання через нездатність капіталізувати майбутні доходи в поточні періоди. За оцінками, близько 20 % домогосподарств США мають це ліквідне обмеження [27, с. 104]. Інші дослідження для США показують, що ця цифра близька до 50 % [21]. Для України ця цифра має бути навіть вищою, оскільки українські домогосподарства не можуть позичати, щоб фінансувати своє споживання. Це означає, що MPC мінливого доходу буде відносно високим, можливо, близьким до MPC постійного доходу. Отже, поточний дохід має стати важливішим для рішень щодо споживання.

На основі попередніх зауважень ми включаємо і дохід, і споживання з лагом, щоб пояснити зміни в поточному споживанні. Коефіцієнти мають бути додатними.

Вважається, що процентна ставка також відіграє роль у прийнятті споживацьких рішень. Якщо ми припускаємо, що багатство є функцією процентної ставки, то можемо розробити кілька ідей щодо того, як домогосподарства реагують на зміни в багатстві, зумовлені, в свою чергу, змінами процентної ставки. При її зростанні ефект заміни завжди зменшує споживання, оскільки воно стає дорожчим, ніж заощадження. Згідно з теорією, зростання процентної ставки стимулює заощадження за припущення, що ефект доходу боржників урівноважує ефект доходу позичальників. Тому збільшення реаль-

² Потрібно зауважити, що теорія споживання залишається *гарячою* темою в макроекономіці. На сьогодні не існує одностайної згоди щодо чинників споживання та каналів їхньої дії.

них процентних ставок зменшує споживання. На жаль, емпірично зв'язок між сукупними заощадженнями та процентними ставками не є очевидним. Деякі дослідження виявили очевидний вплив процентних ставок на заощадження в розвинутих країнах і в країнах, що розвиваються, але більшість інших досліджень показує незначний або протилежний вплив [27, с. 109].

Хоча зміни в процентних ставках мають ефект заміни та ефект доходу і тому остаточний вплив є неоднозначним, ми припускаємо, що зростання процентних ставок призводить до зростання споживання. Іншими словами, градієнт має бути позитивним.

В контексті включення процентної ставки до нашого рівняння споживання ми стикаємося з ситуацією, яка часто трапляється в практичних економетричних дослідженнях. Нам потрібно включити процентну ставку, але яку? На жаль, немає єдиної міри процентної ставки. На ринку існує їх декілька: по депозитах, по кредитах, ломбардна тощо. Щоб узяти обґрунтовану змінну для процентної ставки, ми пропонуємо використати зважену процентну ставку по комерційних кредитах, визначену українськими банками у державній грошовій одиниці.

Ми також включаємо фіктивну змінну для 3-го кварталу 1998 р., щоб подолати вплив інфляційних очікувань напередодні фінансової кризи в Україні.

Отже, оцінимо таку функцію споживання:

$$RCON_t = a_0 + a_1 RDI_t + a_2 R_t + a_3 RCON_{t-1} + a_4 RCON_{t-4} + a_5 DUMMY^{98:03}.$$

Кількість невиключених попередньо визначених змінних дорівнює 10, кількість включених ендогенних змінних мінус 1 дорівнює 2. Оскільки 10 явно більше, ніж 2, ми робимо висновок, що рівняння є переототожненим згідно з умовою порядку. Умова рангу задоволена, оскільки немає невідродженої матриці коефіцієнтів невиключених змінних. E.VIEWS не дає можливості автоматично ототожнити рівняння. Це має зробити сам дослідник.

Оцінка функції споживання за МНК дає такі результати (див. таблицю 4.6).

Загалом коефіцієнти мають очікувані згідно з теорією знаки і є економічно значимими. Наприклад, гранична схильність до споживання від доходу у розпорядженні дорівнює приблизно 20 %. Схильність до споживання у довгостроковому періоді дорівнює приблизно 26 % (= 20 % / [1 - 0.23]), що є все таки низьким значенням.

Як виявилось, номінальна процентна ставка не відіграє значної ролі в поясненні змін у споживанні. Вона статистично незначимо відрізняється від нуля (t -statistic = -0.27, p -value = 0.8)³. Причиною незначимості процентної

³ Висновок є попереднім, оскільки ми ще не враховували можливі ефекти ендогенності. Тим не менше застосування відповідних методів оцінки підтверджує цей висновок.

Таблиця 4.6. Функція споживання

Dependent Variable: RCON

Method: Least Squares

Sample(adjusted): 1996:01 1999:3

Included observations: 15 after adjusting endpoints

Newey-West HAC Standart Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9.888171	2.299068	4.300948	0.0020
RDI	0.195181	0.049872	3.913652	0.0035
R	-0.020336	0.073700	-0.275933	0.7888
RCON(-1)	-0.236457	0.148392	-1.593457	0.1455
RCON(-4)	0.233587	0.114231	2.044867	0.0712
DUMMY9803a	1.828440	0.156958	11.64927	0.0000
R-squared	0.802813	Mean dependent var		13.67295
Adjusted R-squared	0.693265	S.D. dependent var		0.974716
S.E. of regression	0.539833	Akaike info criterion		1.894061
Sum squared resid	2.622780	Schwarz criterion		2.177282
Log likelihood	-8.205461	F-statistic		7.328412
Durbin-Watson stat	2.022826	Prob(F-statistic)		0.005338
Ramsey RESET Test:				
F-statistic	0.020174	Probability		0.890564
Log likelihood ratio	0.037779	Probability		0.845888
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	2.176183	Probability		0.184094
Obs*R-squared	5.750827	Probability		0.056393
White Heteroskedasticity Test:				
F-statistic	0.647723	Probability		0.731102
Obs*R-squared	8.074477	Probability		0.526658
Jarque-Bera	3.331778	Probability		0.189023

ставки є обмежена ліквідність домогосподарств, пов'язана з нерозвинутим в Україні ринком споживчих кредитів. Проте ми сподіваємось, що в майбутньому ця змінна стане важливим фактором.

Як ми і очікували, фіктивна змінна для передкризового кварталу є дуже важливою. Як показує таблиця 4.6, коефіцієнт цієї фіктивної змінної є статистично значимим за будь-якого практичного рівня значимості. Це, очевидно, вказує на бум споживання перед прискороною інфляцією.

Загальна якість регресії досить висока: коефіцієнт детермінації показує, що фактори регресії пояснюють близько 80 % варіації залежної змінної. F-тест свідчить, що регресія є загалом адекватною. Більше того, тест RESET

і тест гетероскедастичності Уайта (White heteroscedasticity test) не вказують на помилки специфікації. Автокореляція помилок також не виявлена ($p(\text{Breusch-Godfrey}) > 10\%$)⁴. Взагалі ця специфікація здається цілком придатною для функції споживання нашої моделі.

Кілька технічних деталей.

Щоб виконати RESET-тест (на вірну специфікацію), вам потрібно

- у вікні об'єкта «рівняння» виділити **View/Stability tests.../Ramsey RESET test**;
- визначити кількість fitted terms (порядок поліноміальності відтворюваних значень), який по умовчання дорівнює 2.

Щоб здійснити тест Бройша—Годфрея (перевірка на автокореляцію помилок),

- у вікні об'єкта «рівняння» виділіть **View/Residual tests.../Serial correlation LM test**;
- визначіть кількість лагів, які треба включити.

Щоб здійснити White heteroscedasticity-тест (перевірка на гетероскедастичність дисперсії),

- у вікні об'єкта «рівняння» виділіть **View/Residual tests.../White heteroscedasticity**.

Слід зазначити, що тест гетероскедастичності Уайта має нульову гіпотезу, яка означає відсутність гетероскедастичності і правильну специфікацію. Таким чином, якщо тест відкидає нульову гіпотезу, то можливою є неправильна специфікація або/та гетероскедастичність.

Щоб здійснити тест Жарку—Бера (нульова гіпотеза: помилки є нормально розподіленими),

- у вікні об'єкта «рівняння» виділіть **View/Residual tests.../Histogram Normality test**.

Крок 3.2. Інвестиції.

З макроекономічної теорії ми знаємо, що інвестиції набагато мінливіші за споживання. Беручи до уваги проциклічну природу інвестицій, ми можемо очікувати, що зростання ВВП веде до зростання інвестиційних витрат. Існує кілька теорій, що пояснюють, чому інвестиції є проциклічними. Наприклад, теорія раціонування кредитів наголошує на тому, що фірми, особливо малі та середні, є об'єктами кредитного нормування. Як результат застосування цього обмеження, фірми не можуть фінансувати свої інвестиційні проекти із зовнішніх джерел, наприклад за рахунок банківських кредитів. Вони можуть покладатися лише на свої внутрішні джерела, такі як прибутки. Оскільки прибутковість фірм⁵ корелює з реальним ВВП, існує також кореляція інвестицій і реального ВВП.

⁴ Ми можемо не застосовувати d-статистику Дарбіна—Уотсона, оскільки маємо залежні змінні з лагами в правій частині рівняння.

⁵ Відповідно до багатьох емпіричних оцінок, ця змінна є проциклічною.

Інша важлива концепція зв'язку між інвестиціями та випуском — концепція акселератора. Хоча акселератор найкраще описує зміни у запасах, його можна застосовувати до всіх видів інвестицій. Ідея акселерації полягає в різниці між автономними та індукованими інвестиціями. Приклади автономних інвестицій — це державні інвестиції або інвестиції у відновлення виробничих потужностей, рівні постійній екзогенній амортизації. Індуковані інвестиції — це пристосування запасів капіталу до бажаного співвідношення капіталу і ВВП. Ідея цієї концепції полягає в тому, що інвестиції залежать від змін у випуску (більш детально в [4, с. 319—347], [27, с. 133—135]).

Мікроекономічна основа акселератора з'являється в контексті кредитних обмежень. Як вже було зазначено, банки турбуються про якість своїх кредитних портфелів. Взагалі малі та середні підприємства є об'єктами раціонування кредитів. Іншими словами, їх доступ до кредитів від фінансових посередників обмежений. Через неспроможність більшості підприємств збільшувати фонди випуском акцій кредитне нормування стає ще впливовішим. Таким чином, підприємства змушені фінансувати свої інвестиції за рахунок поточного доходу та прибутку. Іншим поясненням кредитного нормування є нездатність банків оцінювати ризик. Схоже, що цей аспект є більш впливовим у перехідних економіках, ніж у розвинутих. Внаслідок цього підприємства схильні інвестувати не тоді, коли це потрібно, а тоді, коли це можливо, тобто внутрішнє фінансування є одним з критично важливих факторів поступового пристосування капіталу до бажаного рівня. Припускаючи, що підприємницькі доходи та, можливо, прибутки залежать від випуску й кредитного нормування, ми можемо очікувати, що будь-яке збільшення обсягів випуску стимулює інвестиції, тобто акселератор має врахувати, бодай частково, ефект кредитного обмеження.

Щоб включити до розгляду механізм акселератора, ми пропонуємо додати випуск з лагом до поточного випуску. Проте ми не накладаємо обмеження, що випуск з лагом і поточний випуск повинні мати однакові коефіцієнти. Хоча модель акселератора передбачає це, інші теорії інвестицій — ні. Коротко кажучи, ми очікуємо, що зростання у випуску, вимірюване як реальний ВВП, веде до зростання інвестицій. Інвестиції з лагом повинні відображати поступове пристосування капіталу до бажаного рівня.

Ще одна важлива змінна, що впливає на інвестиційні рішення, — це процентна ставка. Загальне правило: чим вища процентна ставка, тим нижчий рівень інвестицій. Наприклад, житлове будівництво як частина валових інвестицій дуже чутливе до коливань процентної ставки, бо більшість будинків купується за позики. Ринок позик дуже нерозвинутий в Україні. Проте процентна ставка все ж таки впливає на рішення, будувати нові житлові та промислові споруди чи ні, оскільки вони ґрунтуються на теперішній вартості

рентних виплат. Більше того, вартість пристосування до бажаного рівня капіталу також обернено залежна від процентної ставки.

Незважаючи на важливість процентної ставки, вона має обмежене застосування в Україні та багатьох інших пострадянських економіках через м'які бюджетні обмеження підприємств. Багато великих державних підприємств явно чи неявно використовують державні гарантії на їхні борги. В середовищі незахищених прав власності процентна ставка має тенденцію до незначимості. Як результат, банківська система стикається з несприятливим відбором та обмеженням кредитування [2], [28]. Інші причини незначної ролі процентної ставки описані у [7]. Однак останні дані та аналіз економіки підприємств показують, що нині банки та підприємства, надаючи кредити, можуть досить ефективно відрізнити потенційно «поганих» боржників від «хороших» [44]. Таким чином, нам необхідно включити комерційну процентну ставку в регресію, оскільки вона має тенденцію відігравати дедалі важливішу роль у визначенні інвестицій.

Щоб врахувати сезонність інвестицій, ми також включаємо інвестиції з лагом в один і чотири періоди. Інституційні зміни апроксимуються логарифмом часового тренду. В підсумку специфікація виглядає так:

$$RINV_t = a_0 + a_1RGDP_t + a_2RGDP_{t-1} + a_3R_{t-1} + a_4RINV_{t-1} + a_5RINV_{t-4} + a_6Log(Trend_t).$$

Відповідно до умови порядку, це рівняння є переототожненим. Кількість невиключених попередньо визначених змінних, що дорівнює 8, є більшою за кількість включених ендогенних змінних мінус 1, що дорівнює 1. Умова рангу також задоволена: існує щонайменше одна невикористана матриця коефіцієнтів змінних, не включених до рівняння.

Результати оцінки за допомогою МНК представлені в таблиці 4.7.

Отже, результати відповідають нашим очікуванням. Зростання ВВП на 1 млрд миттєво проявляється в зростанні інвестицій на 0,19 млрд. Після початкового шоку інвестиції частково зростають за рахунок механізму пристосування та акселератора. Виявляється, що процентна ставка не відіграє ролі у визначенні динаміки інвестицій: коефіцієнт є статистично незначимим.

В цілому регресія здається правильно специфікованою: відсутня кореляція залишків ($p(\text{Breusch-Godfrey}) \gg 0.1$), немає помилок специфікації ($p(\text{White}) \gg 0.1$, $p(\text{RESET}) \gg 0.1$, відсутня гетероскедастичність ($p(\text{White}) > 0.1$).

Крок 3.3. Процентна ставка.

Процентна ставка, здається, є найбільш нестійкою змінною в нашій моделі. Як фактори реальної економіки, так і монетарні фактори викликають

Таблиця 4.7. Інвестиції

Dependent Variable: RINV

Method: Least Squares

Sample(adjusted): 1996:1 1999:3

Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.646460	2.325515	-3.718085	0.0059
RGDP	0.198813	0.046346	4.289764	0.0027
RGDP(-1)	0.082663	0.045448	1.818835	0.1064
R(-1)	0.155913	0.129180	1.206942	0.2619
RINV(-1)	0.105916	0.137863	0.768269	0.4644
RINV(-4)	0.535761	0.119261	4.492334	0.0020
LOG(@TREND(1995:4))	0.604391	0.274229	2.203964	0.0586
R-squared	0.950415	Mean dependent var	4.764416	
Adjusted R-squared	0.913226	S.D. dependent var	1.239041	
S.E. of regression	0.364990	Akaike info criterion	1.126829	
Sum squared resid	1.065740	Schwarz criterion	1.457253	
Log likelihood	-1.451220	F-statistic	25.55644	
Durbin-Watson stat	2.366468	Prob(F-statistic)	0.000084	
Ramsey RESET Test:				
F-statistic	0.417302	Probability	0.538879	
Log likelihood ratio	0.868579	Probability	0.351349	
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.308211	Probability	0.745734	
Obs*R-squared	1.397484	Probability	0.497210	
White Heteroskedasticity Test:				
F-statistic	1.279408	Probability	0.520366	
Obs*R-squared	13.27118	Probability	0.349642	
Jarque-Bera	0.822788	Probability	0.662726	

коливання в процентній ставці. В нашому конкретному випадку ми застосуємо спрощену модель для процентної ставки, припускаючи, що піднесення економічної активності сприяє зростанню процентної ставки як наслідок вищого попиту на обмежені ресурси, доступні для інвестицій. Найкращим показником для такої змінної економічної активності є реальний ВВП. Ми очікуємо додатний знак для цієї змінної.

Зростання пропозиції грошей за інших рівних умов зменшує процентну ставку. Ми також включаємо поточну інфляцію як наближену оцінку очікуваної інфляції. Зафіксувавши інші фактори на незмінному рівні, ми можемо очікувати, що зростання поточної інфляції відіб'ється на вищих інфляцій-

них очікуваннях i , в свою чергу, призведе до вищих процентних ставок. Логіка цієї тенденції полягає в тому, що кредитори зацікавлені в реальній процентній ставці, а не в номінальній. Хоча короткострокові відхилення можливі, в довгостроковому періоді процентна ставка додатна і практично незмінна. Щоб обґрунтувати це твердження, ми звертаємося до відомого рівняння Фішера: $i = r + \pi$, де i — номінальна процентна ставка, r — реальна процентна ставка, π — рівень інфляції. Згідно з цим рівнянням, інфляція збільшує номінальну процентну ставку один до одного (1:1).

Розвиток фінансового ринку апроксимовано логарифмом часового тренду. Ми також додаємо фіктивну змінну для передкризового 3-го кварталу 1998 р., коли НБУ намагався скоротити вплив капіталу з України ринковими та адміністративними заходами. Часткове пристосування ринку грошей описане процентною ставкою з лагом. Зазначена специфікація підсумована в такому рівнянні:

$$R_t = a_0 + a_1 RM2_t + a_2 RGDP_t + a_3 DLog(CPI_t) + a_4 R_{t-1} + a_5 Log(Trend_t) + a_6 DUMMY^{98:03}.$$

Відповідно до умови порядку, рівняння переототожене (перевірте це). Умова рангу задоволена. Результати оцінки за допомогою МНК представлені в таблиці 4.8.

Як ми очікували, зростання пропозиції грошей зменшує процентну ставку: зростання грошового агрегату М2 на 1 млрд зменшує щомісячну процентну ставку на 0.31 процентного пункту, або приблизно на 4 процентних пункти для річних даних. Вища інфляція, крім того, підвищує процентну ставку. Коефіцієнт реального ВВП є статистично незначимим. Таким чином, ми можемо узагальнити, що реальний ВВП виявився не впливовим на грошовий ринок. Механізм часткового пристосування є досить вагомим в поясненні коливань процентної ставки. Сезонність є також статистично значимим фактором.

В цілому відповідно до формальних критеріїв рівняння правильно специфіковане. P -value для RESET-тесту набагато перевищує критичні 10%. Інші тести (Breusch—Godfrey, White Heteroscedasticity) не виявляють ні автокореляції, ні гетероскедастичності.

Крок 3.4. Пропозиція грошей.

За припущенням, запас грошей є функцією монетарної бази і різниці між ринковою процентною ставкою та ставкою рефінансування, встановленою НБУ. За інших рівних умов зростання монетарної бази призводить до створення «нових» грошей через банківську систему. Грошовий мультиплікатор є відображенням цього процесу створення грошей. Таким чином, ми

Таблиця 4.8. Процентна ставка

Dependent Variable: RINV

Method: Least Squares

Sample(adjusted): 1996:1 1999:3

Included observations: 15 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.078188	1.748271	1.188711	0.2596
RM2	-0.318431	0.081351	-3.914279	0.0024
RGDP	-0.020084	0.026787	-0.749751	0.4691
DLOG(CPI)	17.69707	1.314993	13.45793	0.0000
DUMMY9803	1.414015	0.334399	4.228522	0.0014
R(-1)	0.251284	0.057534	4.367568	0.0011
R(-4)	0.123215	0.047433	2.597651	0.0248
LOG(@TREND(1994:4))	1.469824	0.526635	2.790972	0.0176
R-squared	0.994508	Mean dependent var		6.180526
Adjusted R-squared	0.991013	S.D. dependent var		3.283593
S.E. of regression	0.311276	Akaike info criterion		0.799288
Sum squared resid	1.065821	Schwarz criterion		1.196947
Log likelihood	0.406760	F-statistic		284.5706
Durbin-Watson stat	1.925330	Prob(F-statistic)		0.000000
Ramsey RESET Test:				
F-statistic	1.618260	Probability		0.232120
Log likelihood ratio	2.849865	Probability		0.091382
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.728127	Probability		0.509212
Obs*R-squared	2.646151	Probability		0.266315
White Heteroskedasticity Test:				
F-statistic	1.207679	Probability		0.448080
Obs*R-squared	14.41059	Probability		0.345577
Jarque-Bera	0.721212	Probability		0.697254

очікуємо, що коефіцієнт грошової бази має бути більшим за 1: зростання грошової бази на одну гривню веде до зростання грошового агрегату M2 більше ніж на одну гривню.

Як і монетарна база, розрив між ринковою процентною ставкою і ставкою рефінансування має позитивний вплив на гроші. Якщо розрив великий, уряд проводить експансійну монетарну політику, пропонуючи дешеві фінансові ресурси банківській системі. Проте позичати у Нацбанку є все ж таки скоріше привілей, ніж право. Таким чином, немає автоматичного зв'язку між розривом та монетарною експансією/звуженням.

Ми розглядаємо всі змінні в рівнях. Оскільки інші дослідження показують ймовірну нестационарність змінних, ми маємо перевірити стаціонарність помилок. Якщо помилки стаціонарні, ми маємо коінтеграцію між змінними. В такому випадку МНК дають консистентні оцінки (так само, як інші методи найменших квадратів, в тому числі 2МНК).

Коротко кажучи, ми отримуємо таке регресійне рівняння:

$$M2_t = a_0 + a_1 MB_t + a_1 [R_t - RNBU_t].$$

Це рівняння переототожене відповідно до умови порядку. Кількість невиключених попередньо визначених змінних дорівнює 11, тоді як кількість ендогенних змінних в цьому рівнянні лише 2. Так, одинадцять менше, ніж два мінус одиниця, рівняння є переототожненим. Умова рангу для цього рівняння також задовольняється.

Оцінки МНК для рівняння пропозиції грошей наведені в таблиці 4.9.

Результати оцінки регресії показують, що грошовий мультиплікатор дорівнює 1.7, що збігається з оцінками інших досліджень та нашими власними спостереженнями. За інших рівних умов зростання монетарної бази на 1 млрд грн веде до зростання агрегату M2 на 1.7 млрд грн.

Як не дивно, розрив між ринковою процентною ставкою та ставкою рефінансування має *від'ємний значимий* коефіцієнт. За незмінності інших факторів зменшення розриву на один процентний пункт *збільшує* попит на гроші на 95 млн грн. Коефіцієнт статистично значимий.

RESET-тест свідчить про малу ймовірність помилки специфікації. Автокореляція відхилень також малоімовірна ($p\text{-value}(\text{Breusch—Godfrey}) \gg 0.1$). Гетероскедастичність теж не спостерігається ($p\text{-value}(\text{White}) \gg 0.1$). Розширений тест Дікі—Фуллера відхиляє гіпотезу про нестационарність помилок. Таким чином, отримані оцінки є консистентними. Проте зміщення оцінок через ендогенність змінних в правій частині регресійних рівнянь ми ще не вивчали.

Тепер розглянемо детальніше проблему ендогенності, що виникає в симулятивних моделях.

Крок 4. Оцінка параметрів системи методом 2МНК (TSLS) для усунення ендогенності.

E.VIEWS має широкий набір методів оцінювання для вирішення проблеми ендогенності. Ми зосередимося лише на двокроковому методі найменших квадратів (2МНК) та трикроковому методі найменших квадратів (3МНК). 2МНК (TSLS) застосовується до кожного рівняння окремо. Він має свої переваги і слабкі сторони. З одного боку, нам не потрібно оцінювати всі рівняння одночасно, що може бути незручно. Крім того, менше проблем, які виникають, якщо одне з рівнянь неправильно специфіковане. З іншого боку, ми втра-

Таблиця 4.9. Пропозиція грошей

Dependent Variable: M2

Method: Least Squares

Sample(adjusted): 1995:1 1999:3

Included observations: 19 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.675091	0.121919	5.537191	0.0000
MB	1.723843	0.018203	94.70016	0.0000
R-RNBU	-0.095506	0.033781	-2.827228	0.0121
R-squared	0.998235	Mean dependent var		10.96995
Adjusted R-squared	0.998014	S.D. dependent var		4.490690
S.E. of regression	0.200127	Akaike info criterion		-0.235793
Sum squared resid	0.640811	Schwarz criterion		-0.086671
Log likelihood	5.240036	F-statistic		4523.675
Durbin-Watson stat	2.109587	Prob(F-statistic)		0.000000
Ramsey RESET Test:				
F-statistic	0.017723	Probability		0.895861
Log likelihood ratio	0.022436	Probability		0.880933
Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.193918	Probability		0.825901
Obs*R-squared	0.512161	Probability		0.774080
White Heteroskedasticity Test:				
F-statistic	0.930086	Probability		0.474486
Obs*R-squared	3.989006	Probability		0.407496
Jarque-Bera	0.548383	Probability		0.760186

чаємо ефективність, бо не враховуємо кореляцію між помилками (відхиленнями) окремих рівнянь. Як бачимо, треба вибирати між ефективністю та консистентністю. ЗМНК є ефективнішим за 2МНК, але ЗМНК є вразливішим до помилок специфікації.

Щоб оцінити рівняння одночасно, ми повинні створити новий об'єкт System (виділити Objects/New Object..., обрати System і назвати її); у відкритому вікні виділити Spec та ввести специфікацію системи. Ми повинні точно написати взаємозв'язки між змінними, використовуючи спеціальні позначення для коефіцієнтів C(1), C(2) тощо.

Наприклад, для нашої моделі специфікація має такий вигляд:

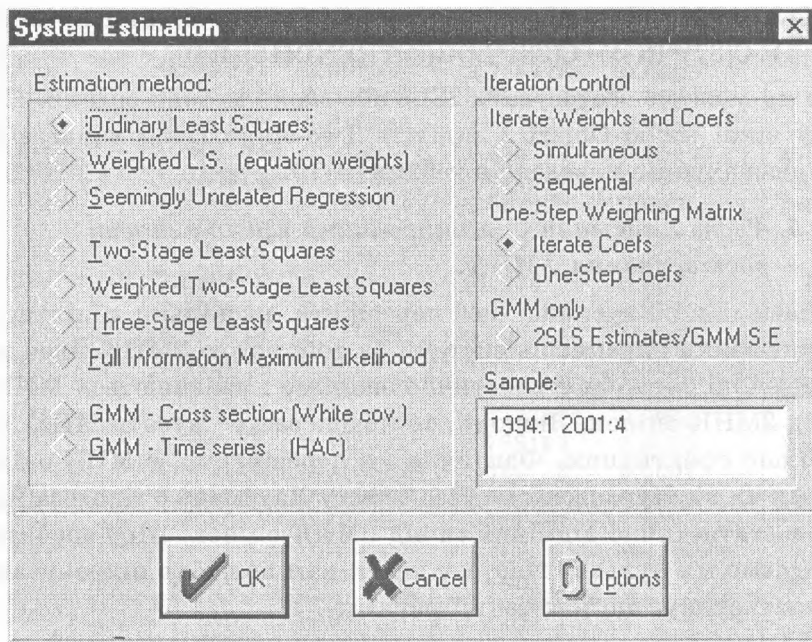
$$RCON=C(1)+C(2)*Rdi+C(3)*R+C(4)*RCON(-1)+C(5)*RCON(-4)+C(6)*DUMMY9803$$

$$RINV=C(11)+C(12)*RGDP+C(13)*RGDP(-1)+C(14)*R(-1)+C(15)*RINV(-1)+C(16)*RINV(-4)+C(17)*LOG(@TREND(1995:4))$$

$$M2=C(21)+C(22)*MB+C(23)*(R-RNBU)$$

$$R=C(31)+C(32)*RM2+C(33)*RGDP+C(34)*DLOG(CPI)+C(35)*DUMMY9803+C(36)*R(-1)+C(37)*R(-4)+C(38)*LOG(@TREND(1994:4)).$$

Кожен коефіцієнт, який слід оцінити, має свій окремий індекс. Якщо індекси збігаються, E.VIEWS сприймає це як обмеження, накладене на коефіцієнти. Пам'ятайте, що ми не розглядаємо тотожності на цьому етапі. Щоб оцінити модель, натисніть **Estimate** в меню **Object** системи. У відкритому вікні оберіть метод оцінки.



В даному ілюстративному випадку ми можемо вибрати лише три перші опції: OLS, WLS та SUR. Натисніть **OK** і отримайте статистичні оцінки. Щоб використати інші методи, що знімають проблему ендогенності, наприклад 2МНК або 3МНК, ми повинні зазначити інструменти. Це можна зробити, додавши список інструментів у вікно специфікації. Натисніть **Spec** у меню **Object** системи. Команда **INST** слугує для специфікації списку інструментів. У нашому інструментами можуть бути всі екзогенні та попередньо визначені змінні системи, зокрема: RGDP(-1), RCON(-1), MB, RNBU, CPI(-1), RG, RTAX, DUMMY9803, @TREND(1994:1), DUMMY03, DUMMY9803, DUMMY9804. Таким чином, ми пишемо в будь-якому місці вікна **Spec** такий рядок:

INST RGDP(-1) RCON(-1) MB RNBU CPI(-1) RG RTAX DUMMY9803
@TREND(1994:1) DUMMY03 DUMMY9804.

У вікні специфікації ми маємо:

INST RGDP(-1) RCON(-1) MB RNBU CPI(-1) RG RTAX DUMMY9803
@TREND(1994:1) DUMMY03 DUMMY9804

RCON=C(1)+C(2)*RDI+C(3)*R+C(4)*RCON(-1)+C(5)*RCON(-4)+
+C(6)*DUMMY9803

RINV=C(11)+C(12)*RGDP+C(13)*RGDP(-1)+C(14)*R(-1)+C(15)*RINV(-1)+
+C(16)*RINV(-4)+C(17)*LOG(@TREND(1995:4))

M2=C(21)+C(22)*MB+C(23)*(R-RNBU)

R=C(31)+C(32)*RM2+C(33)*RGDP+C(34)*DLOG(CPI)+C(35)*DUMMY9803+
+C(36)*R(-1)+C(37)*R(-4)+C(38)*LOG(@TREND(1994:4)).

Тепер ми можемо застосувати 2SLS або схожі методи оцінки. Натисніть Estimate у вікні меню Object і виділіть Two-stage Least squares (2МНК). Результат розрахунків наведено в таблиці 4.10.

Крок 5. Тест Хаусмана для оцінювання ефективності застосування 2МНК.

Виконаємо тест Хаусмана, щоб перевірити, наскільки доцільною є процедура оцінювання параметрів системи за допомогою 2МНК. Ідея тесту полягає в перевірці того, чи статистично значимо відрізняються МНК-оцінки системи від 2МНК-оцінок. Якщо ні, то краще застосовувати МНК, оскільки він буде більш ефективним. Фактично ми перевіряємо, є МНК-оцінки консистентними чи ні. Нульовою гіпотезою тесту Хаусмана є гіпотеза, що МНК-оцінки не є статистично відмінними від 2МНК-оцінок. Щоб зробити висновок, відкидаємо ми нульову гіпотезу чи ні, нам потрібно підрахувати спеціальну χ^2 -статистику, що має такий вигляд:

$$\chi^2 = (b^{2SLS} - b^{OLS})^T \cdot (V^{2SLS} - V^{OLS})^{-1} \cdot (b^{2SLS} - b^{OLS}),$$

де b^{2SLS} — вектор оцінених параметрів системи методом 2МНК; b^{OLS} — вектор оцінених параметрів системи методом МНК; V^{2SLS} — дисперсійно-коваріаційна матриця оцінених параметрів системи методом 2МНК; V^{OLS} — дисперсійно-коваріаційна матриця оцінених параметрів системи методом МНК. Зазначимо, що дана статистика розподілена за χ^2 -розподілом з n ступенями свободи (де n — кількість параметрів системи, які потрібно оцінити). Якщо обрахована статистика перевищує критичне значення, що знаходиться за таблицями χ^2 -розподілу при заданих рівні значимості та ступенях свободи, ми відхиляємо нульову гіпотезу (МНК-оцінки не є статистично відмінними від

Таблиця 4.10. Система рівнянь 2МНК

Estimation Method: Two-Stage Least Squares

Sample: 1995:2 1999:3

Included observations: 18

Total system (unbalanced) observations 66

Instruments: RGDP(-1) RCON(-1) MB RNBUCPI(-1) RG RTAX

DUMMY9803 @TREND(1994:1) DUMMY03 DUMMY9804 C

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	9.471847	2.624375	3.609182	0.0008
C(2)	0.220782	0.132338	1.668325	0.1027
C(3)	0.003678	0.140284	0.026220	0.9792
C(4)	-0.252914	0.166485	-1.519140	0.1362
C(5)	0.236140	0.208137	1.134543	0.2630
C(6)	1.809811	0.585345	3.091870	0.0035
C(11)	-9.625371	2.562820	-3.755773	0.0005
C(12)	0.197141	0.049753	3.962403	0.0003
C(13)	0.100479	0.048268	2.081704	0.0435
C(14)	0.210233	0.143434	1.465712	0.1502
C(15)	0.058240	0.148809	0.391372	0.6975
C(16)	0.574654	0.126081	4.557820	0.0000
C(17)	0.735881	0.300892	2.445668	0.0187
C(21)	0.899219	0.138981	6.470078	0.0000
C(22)	1.699562	0.018619	91.27910	0.0000
C(23)	-0.162924	0.039958	-4.077326	0.0002
C(31)	-0.951648	2.419920	-0.393256	0.6961
C(32)	-0.402061	0.099231	-4.051777	0.0002
C(33)	0.016141	0.035290	0.457382	0.6498
C(34)	14.40937	2.630526	5.477752	0.0000
C(35)	1.388687	0.368532	3.768159	0.0005
C(36)	0.305916	0.070828	4.319122	0.0001
C(37)	0.210235	0.076633	2.743421	0.0089
C(38)	2.358304	0.709955	3.321763	0.0019

Оцінки параметрів для першого рівняння (споживання)

Оцінки параметрів для другого рівняння (інвестиції)

Оцінки для рівняння попиту на

Оцінки параметрів четвертого рівняння

Determinant residual covariance 1.68E-05

Equation: RCON=C(1)+C(2)*RDI+C(3)*R+C(4)*RCON(-1)+C(5)*RCON(-4)+C(6)*DUMMY9803

Observations: 15

R-squared	0.799737	Mean dependent var	13.67295
Adjusted R-squared	0.688479	S.D. dependent var	0.974716
S.E. of regression	0.544029	Sum squared resid	2.663706
Durbin-Watson stat	1.969764		

Міра того, як залишки (помилки) корелюють між рівняннями системи

Статистика для рівняння споживання

Equation: RINV=C(11)+C(12)*RGDP+C(13)*RGDP(-1)+C(14)*R(-1)+C(15)*RINV(-1)+C(16)*RINV(-4)+C(17)*LOG(@TREND(1995:4))

Observations: 15

R-squared	0.946820	Mean dependent var	4.764416
Adjusted R-squared	0.906935	S.D. dependent var	1.239041
S.E. of regression	0.377988	Sum squared resid	1.143000
Durbin-Watson stat	2.420381		

Equation: $M2=C(21)+C(22)*MB+C(23)*(R-RNBU)$
 Observations: 18

R-squared	0.998525	Mean dependent var	11.36078
Adjusted R-squared	0.998329	S.D. dependent var	4.275466
S.E. of regression	0.174784	Sum squared resid	0.458244
Durbin-Watson stat	2.434240		

Equation: $R=C(31)+C(32)*RM2+C(33)*RGDP+C(34)*DLOG(CPI)+C(35)*DUMMY9803+C(36)*R(-1)+C(37)*R(-4)+C(38)*LOG(@TREND(1994:4))$
 Observations: 18

R-squared	0.981771	Mean dependent var	5.562778
Adjusted R-squared	0.969010	S.D. dependent var	1.933664
S.E. of regression	0.340399	Sum squared resid	1.158716
Durbin-Watson stat	2.024850		

2МНК-оцінок), і слід застосовувати 2МНК, оскільки в даному випадку МНК-оцінки будуть зміщеними та неконсистентними. В іншому разі ми повинні використовувати МНК-метод, тому що він є ефективніший за 2МНК.

Тест Хаусмана в командах E.VIEWS

`sys01.ls` 'оцінити систему за OLS (МНК); система вже специфікована
`sys02.tsls` 'оцінити систему за TSLS (2МНК)

'обрахувати різницю між матрицями дисперсії-коваріації
`matrix var_dif=sys02.@coefcov-sys01.@coefcov`

Var-Cov-матриця
коєфіцієнтів

'обрахувати обернену матрицю `var_dif`
`matrix var_dif_inv=@inverse(var_dif)`

Визначити нову
матрицю об'єкта

'обрахувати різницю між коефіцієнтами матриці
`matrix coef_dif=sys02.@coefs-sys01.@coefs`

'транспонувати матрицю `coef_dif`
`matrix coef_dif_trans=@transpose(coef_dif)`

Перемножити
матриці

'обрахувати χ^2 -статистику
`matrix st=coef_dif_trans*var_dif_inv*coef_dif`

Кількість
коєфіцієнтів
у системі

'обрахувати p -значення для статистики
`matrix result=@chisq(st(1,1), sys02.@ncoef)`

`st(1,1)` — значення
статистики
`sys02.ncoef` —
ступені свободи

(or scalar result=@chisq(st(1,1), sys02.@ncoef)

Новий об'єкт
«скаляр»

(Ви можете запусити програмний файл з цими командами. Порада, як запусити програму (припустимо, що ви її вже набрали та записали у файл з ім'ям **hausman.prg**): відкрийте **workfile**, відкрийте програмний файл **hausman.prg**, натисніть **Run**, натисніть **OK**, відкрите вікно показує *p*-значення.)

Якщо запусити цей перелік команд (дивіться також доданий програмний файл: відкрийте його і натисніть **RUN**) для нашого завдання, *p*-значення для розрахованої статистики буде дорівнювати приблизно 0.001. Отже, ми відхиляємо нульову гіпотезу, що МНК-оцінки є консистентними, тому слід використовувати 2МНК.

Завдання

*Виконайте тест Хаусмана окремо для функцій споживання та інвестицій (порада: використайте коментарі в специфікації моделі для інших функцій). Що таке значення χ^2 -статистики? Що таке *p*-значення? Прокоментуйте самостійно.*

Крок 6. Оцінка параметрів системи методом 3МНК (3SLS).

Тепер перейдемо до оцінювання параметрів системи трикроковим методом найменших квадратів. Сутність процедури 3МНК полягає у використанні інформації про кореляцію помилок між рівняннями системи. Це дає змогу знайти точні оцінки параметрів регресійних рівнянь. Ми можемо оцінити, як корелюють помилки, за допомогою матриці кореляцій помилок (залишків). Натисніть **View** у вікні об'єкта **Model** і виділіть **Residuals/Correlation matrix**. Для нашої системи матриця кореляцій має такий вигляд.

Таблиця 4.11. Перехресна кореляція між рівняннями

	RCON	RINV	M2	R
RCON	1.00000	-0.11873	0.38247	0.16695
RINV	-0.11873	1.00000	-0.02699	-0.24739
M2	0.38247	-0.02699	1.00000	0.19421
R	0.16695	-0.24739	0.19421	1.00000

Підсумковий показник цих перехресних кореляцій є детермінантом матриці коваріацій помилок (дивіться результат оцінювання: виділено жирним). На жаль, важко визначити потенційні переваги використання 3МНК *a priori*. Точність оцінок може бути переважена збільшеною вразливістю моделі до помилок специфікації.

Аналогічно тому, як ми оцінювали параметри системи 2МНК, можемо оцінити їх за допомогою 3МНК. Натисніть Estimate в меню об'єкта System і виділіть Three-stage Least squares. Результати наведені в таблиці 4.12.

Таблиця 4.12. Система рівнянь, 3МНК

Estimation Method: Three-Stage Least Squares

Sample: 1995:2 1999:3

Included observations: 18

Total system (unbalanced) observations 66

Instruments: RGDP(-1) RCON(-1) MB RNBU CPI(-1) RG RTAX

DUMMY9803 @TREND(1994:1) DUMMY03 DUMMY9804 C

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	7.857561	1.875251	4.190138	0.0001
C(2)	0.217584	0.094226	2.309182	0.0259
C(3)	-0.013710	0.100951	-0.135805	0.8926
C(4)	-0.150966	0.119237	-1.266102	0.2125
C(5)	0.264438	0.147705	1.790312	0.0806
C(6)	1.549541	0.429018	3.611835	0.0008
C(11)	-9.358521	1.844266	-5.074387	0.0000
C(12)	0.192873	0.036006	5.356627	0.0000
C(13)	0.096827	0.034523	2.804699	0.0076
C(14)	0.223960	0.103392	2.166117	0.0360
C(15)	0.033367	0.106565	0.313117	0.7557
C(16)	0.572333	0.090575	6.318857	0.0000
C(17)	0.745750	0.216237	3.448763	0.0013
C(21)	0.917057	0.120486	7.611311	0.0000
C(22)	1.695975	0.016169	104.8919	0.0000
C(23)	-0.154724	0.035351	-4.376834	0.0001
C(31)	-0.982247	1.754677	-0.559788	0.5786
C(32)	-0.430347	0.071875	-5.987427	0.0000
C(33)	0.020921	0.025736	0.812924	0.4208
C(34)	15.12821	1.890878	8.000626	0.0000
C(35)	1.378607	0.266931	5.164660	0.0000
C(36)	0.309323	0.051214	6.039851	0.0000
C(37)	0.203346	0.055059	3.693280	0.0006
C(38)	2.427796	0.513091	4.731705	0.0000

Determinant residual covariance 1.36E-05

Equation: RCON=C(1)+C(2)*RDI+C(3)*R+C(4)*RCON(-1)+C(5)*RCON(-4)+C(6)*DUMMY9803

Observations: 15

R-squared	0.783861	Mean dependent var	13.67295
Adjusted R-squared	0.663783	S.D. dependent var	0.974716
S.E. of regression	0.565181	Sum squared resid	2.874870
Durbin-Watson stat	1.875793		

Equation: RINV=C(11)+C(12)*RGDP+C(13)*RGDP(-1)+C(14)*R(-1)+C(15)*RINV(-1)+C(16)*RINV(-4)+C(17)*LOG(@TREND(1995:4))

Observations: 15

R-squared	0.944358	Mean dependent var	4.764416
Adjusted R-squared	0.902627	S.D. dependent var	1.239041
S.E. of regression	0.386639	Sum squared resid	1.195915
Durbin-Watson stat	2.417031		

Equation: $M2=C(21)+C(22)*MB+C(23)*(R-RNBU)$

Observations: 18

R-squared	0.998517	11.36078
Adjusted R-squared	0.998319	4.275466
S.E. of regression	0.175272	0.460806
Durbin-Watson stat	2.408633	

Equation: $R=C(31)+C(32)*RM2+C(33)*RGDP+C(34)*DLOG(CPI)+C(35)*DUMMY9803+C(36)*R(-1)+C(37)*R(-4)+C(38)*LOG(@TREND(1994:4))$

Observations: 18

R-squared	0.982450	5.562778
Adjusted R-squared	0.970166	1.933664
S.E. of regression	0.333995	1.115528
Durbin-Watson stat	2.108160	

Зазначте, що метод ЗМНК дає більш точні оцінки для параметрів, ніж 2МНК. Детермінант матриці коваріацій зменшився з $1.68E-05$ в 2МНК до $1.36E-05$ в ЗМНК, або приблизно на 20 %. Таблиця 4.13 також демонструє на прикладі окремої функції споживання, що оцінки параметрів є взагалі більш точними.

Таблиця 4.13. Оцінки параметрів функції споживання

Variables	3SLS			2SLS		
	Coefficient	SE	p-value	Coefficient	SE	p-value
Const	7.858	1.875	0.0001	9.472	2.624	0.0008
RDI	0.218	0.094	0.0259	0.221	0.132	0.1027
R	-0.014	0.101	0.8926	0.004	0.140	0.9792
RCON(-1)	-0.151	0.119	0.2125	-0.253	0.166	0.1362
RCON(-4)	0.264	0.148	0.0806	0.236	0.208	0.2630
DUMMY9803	1.550	0.429	0.0008	1.810	0.585	0.0035

Самостійно перевірте це твердження для останніх трьох рівнянь.

Крок 7. Особливості прогнозування на основі симультативних моделей.

На жаль, E.VIEWS відокремлює оцінювання систем і симуляцію (отримання прогнозних даних для минулих або майбутніх значень для ендогенних змінних системи) з оціненими системами. Опишемо коротко процес симуляції в пакеті E.VIEWS.

Для проведення прогнозу ми повинні створити ще один об'єкт MODEL. Це можна зробити, натиснувши Procs в меню Object і виділивши Make model.

В заново створеному об'єкті ми бачимо, що перший рядок це — **ASSIGN @ALL F**. Ця команда наказує **E.VIEWS** додати **F** до всіх імен розрахованих (прогнозних) значень ендогенних змінних. Наприклад, розраховані (fitted) значення споживання **RCON** будуть збережені під іменем **RCONF**. Ми можемо замінити його на будь-який інший символ чи слово, наприклад **FOR**. **E.VIEWS** дозволяє використання не тільки суфіксів, а й префіксів (**ASSIGN @PREFIX F**).

На цьому етапі ми повинні додати тотожності до специфікації моделі. Зауважимо, що ми можемо використовувати коментарі в специфікації моделі, щоб легко розрізнити рівняння або тотожності між собою. Коментарі можуть бути де завгодно після символу **'**. Наприклад, для нашого завдання ми можемо ввести тотожності в систему для утворення моделі з такими коментарями:

'реальний ВВП: наступний рядок — тотожність ВВП
RGDP=RCON+RINV+RG+NX

'реальний дохід у використанні
RDI=RGDP-RTAX

'реальні грошові залишки
RM2=M2/CPI*100

Після того як ми додали всі тотожності, у вікні специфікації моделі маємо:

ASSIGN @ALL _F

'real GDP
RGDP=RCON+RINV+RG+NX

'disposable income
RDI=RGDP-RTAX

'real money balances
RM2=M2/CPI*100

'real consumption
RCON=10.54203219+0.2876768382*RDI+0.01521275802*R-
-0.2653303418*RCON(-1)+0.07822234519*RCON(-4) +
+1.766620998*DUMMY9803

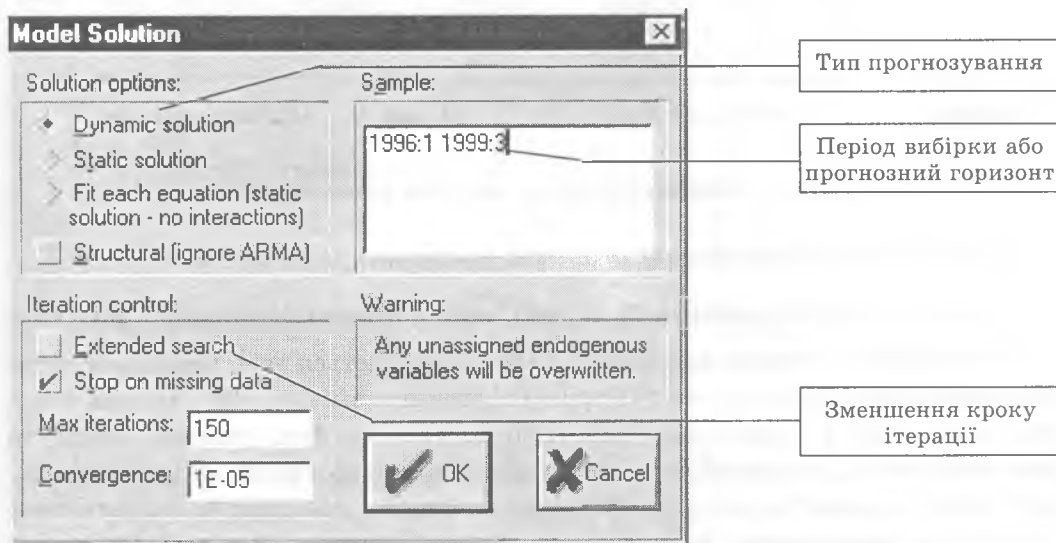
'real investment
RINV=-10.05593616+0.2013080839*RGDP+0.09684253696*RGDP(-1)+
+0.237424563*R(-1)+0.08645828423*RINV(-1)+0.5802259394*RINV(-4)+
+0.7943412564*LOG(@TREND(1995:4))

'money supply
M2=0.9006022025+1.699434837*MB-0.1636280921*(R-RNBU)

'interest rate

$$R = -0.7870717232 - 0.4015562553 * RM2 + 0.01516727059 * RGDP + 14.67748385 * DLOG(CPI) + 1.391526878 * DUMMY9803 + 0.3062138166 * R(-1) + 0.202111356 * R(-4) + 2.314974518 * LOG(@TREND(1994:4))$$

Таким чином, ми підготували модель для розв’язку. Щоб здійснити його, у вікні об’єкта Model натисніть Solve. Відкриється нове діалогове вікно:



Опція **Sample** визначає період, для якого ми робимо прогноз (або моделювання). Впевніться, що ви маєте значення екзогенних змінних для виділеного періоду. В іншому разі неможливо розв’язати модель. Ми обмежимо наше моделювання 3-м кварталом 1999 р.

Опція **Extended search** застосовується, коли значення змінних малі, тому потрібні маленькі кроки в ітераціях під час розв’язання моделі.

Зазначимо, що динамічний та статичний розв’язки (опції: **Dynamic solution** та **Static solution**) можуть давати, і як правило дають, різні результати. Різниця між динамічним та статичним розв’язками полягає в тому, що статичний розв’язок є насправді прогнозуванням на один крок вперед на заданому горизонті прогнозування. Іншими словами, для лагових залежних змінних використовуються фактичні дані. На відміну від статичного розв’язку, динамічний розв’язок використовує розраховані (fitted) значення для лагових залежних змінних. Динамічний розв’язок є точнішою перевіркою моделі, оскільки в статичному розв’язку ми не можемо перевірити, чи є модель динамічно стабільною (тобто прогнозовані динамічні ряди є збіжними). При статичному розв’язку ряди повинні збігатися, бо використовуються фактичні ряди, що збігаються. Таким чином, динамічне рішення є кращим показником якості моделі.

Схематично процес прогнозування зображено на рисунку 4.1.

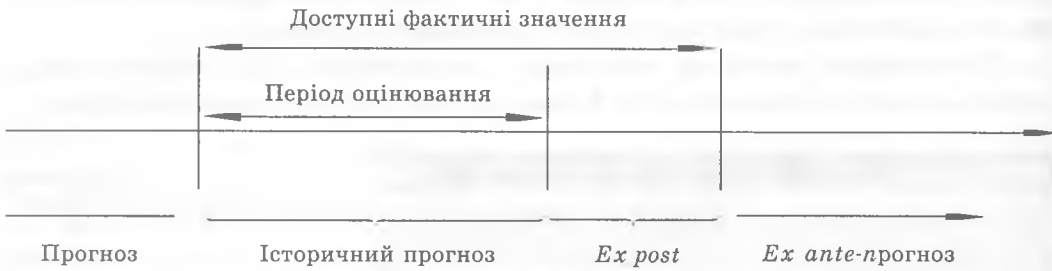


Рис. 4.1. Спрощена схема періодів прогнозування

Самостійно проаналізуйте отримані прогнозні ряди даних.

Крок 8. Оцінка прогнозної якості моделі симультативних рівнянь.

Якість моделі симультативних рівнянь визначається тим, наскільки добре вона відтворює фактичні часові ряди ендогенних змінних. Як можна побачити з рисунка 4.1, тип симуляції (прогнозу) може бути різним. Класично рекомендується для перевірки якості моделі використовувати *ex post*-прогноз, щоб потім порівняти доступні фактичні значення показників з прогнозними даними, отриманими на основі розробленої та оціненої моделі. Ми не виконуємо цю рекомендацію через дуже короткі часові ряди.

Точніше якість прогнозу можна перевірити на основі різноманітних критеріїв якості.

Існує певний набір критеріїв оцінки якості моделі, серед яких найпоширенішими є:

- формальні статистичні критерії (див. нижче),
- критичні точки,
- чутливість початкового періоду,
- чутливість до змін коефіцієнтів.

До формальних статистичних критеріїв належать:

- корінь середньої квадратної помилки $RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$,

- середня помилка $ME = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)$,

- середня процентна помилка $MPE = \frac{1}{N} \sum_{i=1}^N \frac{(y_i - \hat{y}_i)}{y_i}$,

- середня процентна абсолютна помилка $MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{|y_i|}$,

- коефіцієнт нерівності Тейла $U = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^T (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^T y_i^2} + \sqrt{\frac{1}{N} \sum_{i=1}^T \hat{y}_i^2}}$,

- пропорція зміщення $U^M = \frac{(y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N \frac{(y_i - \hat{y}_i)}{y_i}}$,

- пропорція дисперсії $U^S = \frac{(\sigma_{actual} - \sigma_{fitted})^2}{\frac{1}{N} \sum_{i=1}^N \frac{(y_i - \hat{y}_i)}{y_i}}$,

- пропорція коваріації $U^C = \frac{2(1-\rho)(\sigma_{actual} \cdot \sigma_{fitted})}{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$.

Корінь середньої квадратної помилки (RMSE) є достатньо зручним показником якості, бо вимірює середню стандартну помилку прогнозу. **Середня помилка (ME)** дає змогу виміряти зміщення в оцінках. За припущенням, середня помилка повинна дорівнювати нулю, інакше ми матимемо відхилення в оцінках. **Середня процентна помилка (MPE)** забезпечує відносну оцінку зміщення (bias). **Середня процентна абсолютна помилка (MAPE)** схожа з критерієм **RMSE**, але є відносним показником точності моделі.

Коефіцієнт нерівності Тейла (U) є дуже важливим індикатором точності моделі і консистентності отриманих оцінок. За побудовою цей коефіцієнт варіює між нулем і одиницею. Якщо $U = 1$, модель зовсім не може використовуватися для прогнозу. Розраховані (fitted) і фактичні ряди не корелюють. Але якщо $U = 0$, розраховані ряди повністю збігаються з фактичними рядами, тобто прогноз є досконалим.

Коефіцієнт нерівності Тейла (U) може бути розкладеним на три частини: зміщену пропорцію (пропорцію зміщеності) U^M , пропорцію дисперсії U^S і пропорцію коваріації U^C . Зазначимо, що $U^M + U^C + U^S = 1$. Критерій зміщеності пропорції (U^M) використовується, щоб перевірити, чи є систематичне відхилення середніх розрахованих та фактичних рядів, тобто чи дає модель систематично завищені або занижені прогнози. Чим менше значення U^M ,

тим краще. Якщо U^M дорівнює нулю, у розрахованих (прогнозних) значеннях немає зміщень, тобто з моделлю все гаразд. Пропорція дисперсії використовується, щоб переконатися, що модель має достатні динамічні властивості для відтворення дисперсії фактичних рядів. Наприклад, модель може відтворювати систематично менші коливання, ніж фактичні. Як і у випадку критерію U^M , менше значення U^S вказує на менше зміщення (bias). Нарешті, пропорція коваріації вказує, як корелюють фактичні та розраховані ряди. U^C рівне одиниці показує, що фактичні та розраховані ряди корелюють ідеально.

Критичні точки важливі як критерій якості, оскільки деякі моделі можуть бути дуже точними, але погано передбачати зміни в тренді (наприклад поворотні точки в циклах), тобто погано відтворювати критичні точки. Інші моделі можуть бути неточними, але мати багатий динамічний характер. Загалом може бути певний компроміс між точністю та динамічними властивостями моделі. На жаль, немає формального тесту для оцінки цієї властивості. Проте візуальний огляд розрахованих та фактичних рядів звичайно одразу виявляє, добре модель відтворює критичні точки чи ні.

Ще один важливий тест якості моделі — аналіз чутливості до початкового (стартового) періоду симуляції. Якщо модель є нечутливою до змін початкового періоду, то вона є досить хорошою за інших рівних умов. Залежність моделі від початкового періоду є індикатором того, що специфікація є недосконалою (наприклад, можливо в моделі наявні деякі нестационарні часові ряди).

Основні характеристики якості нашої моделі підсумовані в таблиці.

Таблиця 4.14. Оцінка якості моделі

	RGDP	RINV	RCON	R	M2
<i>RMSE</i>	1.0407	0.4835	0.6936	0.4868	0.1790
<i>ME</i>	-0.5377	-0.4172	-0.1205	-0.2612	0.0256
<i>MPE</i>	-0.0178	-0.0868	-0.0093	-0.0554	0.0022
<i>MAPE</i>	0.0336	0.1180	0.0399	0.1024	0.0135
U^M	0.2107	0.4268	0.0293	0.2235	0.0200
U^S	0.1249	0.1107	0.0414	0.0364	0.0147
U^C	0.6643	0.4626	0.9293	0.7401	0.9653
<i>U</i>	0.1473	0.1855	0.3422	0.1703	0.0237

Як можна побачити з таблиці 4.14, загалом результати є задовільними. Наприклад, *MAPE* тільки для деяких рівнянь трохи перевищує заданий 10 %-й рівень точності. Для інших рівнянь системи цей показник набагато нижчий. Коефіцієнт нерівності Тейла *U* досить низький. Крім того, для функції інвестицій ми спостерігаємо постійне переоцінення. В середньому розраховані значення на 8.68 % вищі за фактичні, про що свідчить візуальний аналіз графіка отриманих прогнозних та фактичних даних по інвестиціях.

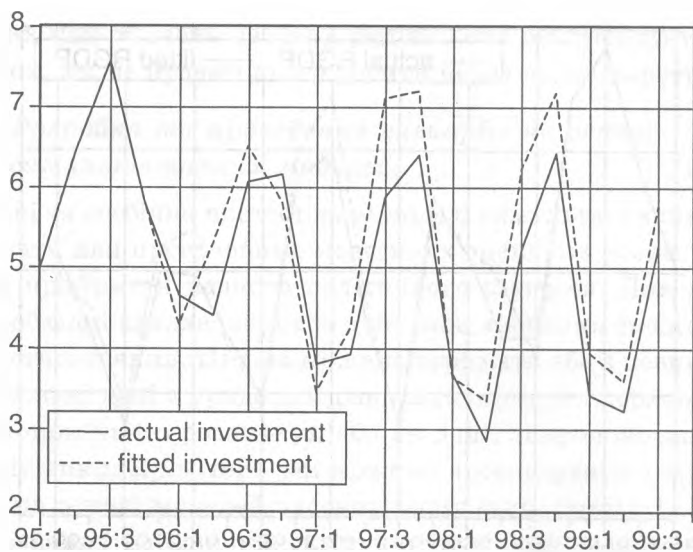


Рис. 4.3. Прогнозні та реальні інвестиції, млрд грн

Графічний аналіз ми можемо зробити також і для інших ендогенних змінних системи, наприклад для процентної ставки та ВВП.

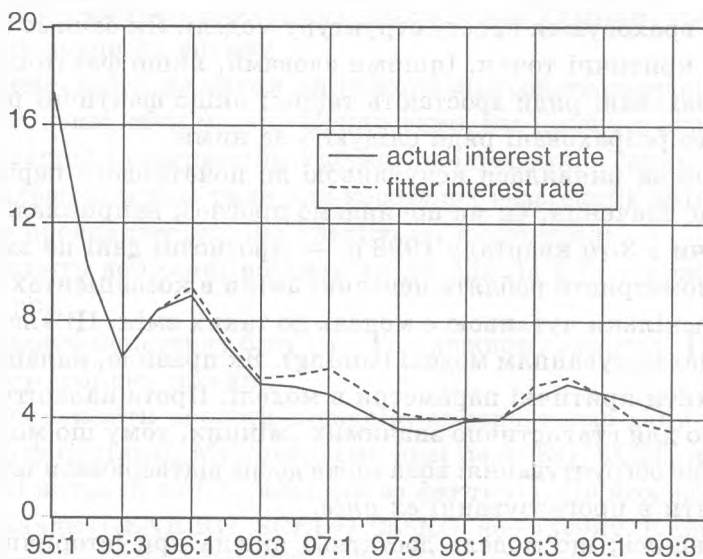


Рис. 4.4. Щомісячна процентна ставка.
Прогнозні та фактичні значення

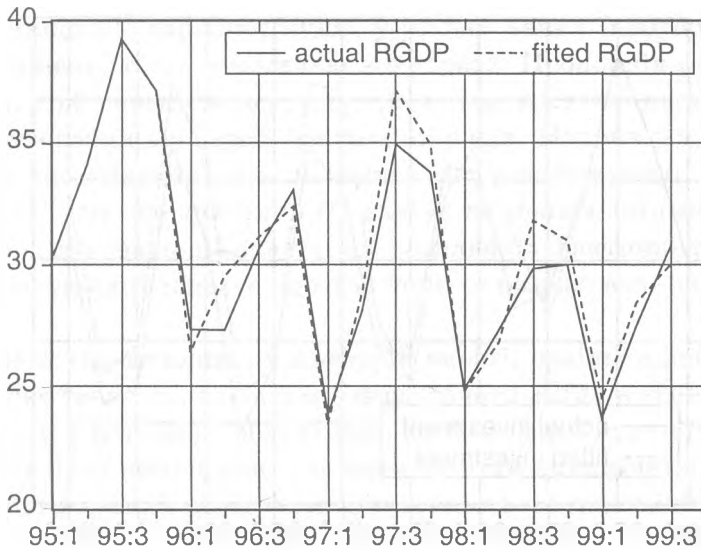


Рис. 4.5. Прогнозний та реальний ВВП (GDP), млрд грн, середина 1996=1

Побудуйте графіки для інших змінних самостійно.

Візуальний огляд графіків переконує, що розраховані значення досить близькі до фактичних. Середня абсолютна процентна помилка (MAPE) не перевищує 10 %, таким чином, прогнозні можливості моделі достатньо високі, особливо враховуючи просту структуру моделі. Як бачимо на графіках, відтворено всі критичні точки. Іншими словами, якщо фактичні ряди зростають, то розраховані ряди зростають також; якщо фактичні ряди починають спадати, то розраховані ряди слідує за ними.

Модель також виявилася нечутливою до початкового періоду. Інакше кажучи, не має значення, чи ми починаємо прогноз, наприклад, з 1-го кварталу 1996 р., чи з 3-го кварталу 1998 р. — прогнозні дані не змінюються.

Іноді економетристи роблять невеликі зміни в коефіцієнтах моделі, щоб перевірити, наскільки чутливою є модель до таких змін. Ці «штучні» зміни називаються налаштуванням моделі (tuning). Як правило, налаштування дає змогу простежити критичні параметри в моделі. Проте налаштування небажане, особливо для статистично значимих змінних, тому що модель втрачає своє статистичне обґрунтування: вона може добре відтворювати історичні дані, але погано діяти в прогнозуванні *ex ante*.

Переконавшись, що модель достатньо якісна при історичному прогнозуванні (ми оцінили якість моделі для цього прогнозного горизонту), слід перевірити якість моделі (тобто провести *ex post*-прогноз), порівнявши фактичні (не включені раніше у вибірку для оцінювання моделі) дані та розраховані значення (див. рисунок 4.1). Якщо розраховані значення добре відтворюють фактичні, модель є задовільною і ми знову оцінюємо систему на основі

всіх доступних нам значень часових рядів. Хоча рекомендується проводити *ex post*-прогноз, ми не будемо цього робити через надто короткі часові ряди.

Крок 8. Розробка та проведення сценаріїв на основі симультативних моделей.

Розроблені та всебічно протестовані моделі симультативних рівнянь можна застосовувати для програвання можливих сценаріїв розвитку або аналізу ефективності прийняття певного політичного рішення. Для реалізації сценаріїв нам необхідні заплановані або очікувані значення екзогенних змінних на момент прогнозування. Ці дані можемо отримати або з певних припущень, або як заплановані дані з урядових джерел (наприклад державного бюджету або планів економічного розвитку), або як дані, запропоновані особами, що приймають рішення. Крім того, ми можемо прогнозувати значення екзогенних змінних на основі моделей часових рядів (див. розділ 1).

Певний заданий набір рядів екзогенних змінних називається сценарієм. Неформально екзогенні змінні можна класифікувати на керовані (*policy variables*) та чисті екзогенні змінні. Керовані екзогенні змінні є змінними, які можуть задаватись та змінюватись особами, що приймають рішення (*policy maker*). Прикладами таких екзогенних змінних можуть бути монетарна база або дисконтна процентна ставка (рефінансування). Інші екзогенні змінні є, по суті, «чистими». Політики можуть не впливати на них. Наприклад, світові ціни на нафту є «чистою» екзогенною змінною для України, оскільки Україна не має на них значного впливу.

Розглянемо, яким же чином ми можемо програти сценарії на основі розробленої та оціненої нами моделі. Припустимо, нас цікавить аналіз результатів вузької монетарної та експансійної фіскальної політики. Таким чином, ми розглядаємо відповідні часові ряди для керованих екзогенних змінних.

Давайте проаналізуємо, що трапиться, якщо уряд:

1) підвищить державні витрати та доходи на 5 % у кожному кварталі 1997—1999 рр.;

2) підвищить монетарну базу на 5 % у кожному кварталі 1997—1999 рр.;

3) застосує обидва заходи.

Щоб зсимулювати перший сценарій, використаємо дану модель з невеликими модифікаціями. Ми генеруємо нові ряди RG_SCE1 , що представляють державні витрати на 5 % вищими за фактичні. Це можна зробити за допомогою команди ($GENR\ RG_SCE1=1.05*RG$) або вручну. В специфікації моделі всюди замініть RG на RG_SCE1 . Таку ж операцію слід зробити для $RTAX$ рядів. Використовуючи команду *Assign*, збережіть результати симуляції в рядах із суфіксами $_SCE1$: $ASSIGN\ @ALL\ _SCE1$. Загалом модель для програвання та аналізу першого сценарію має вигляд:

ASSIGN @ALL _SCE1

Збережіть результати симуляції в рядях із суфіксом SCE1

RGDP=RCON+RINV+RG_SCE1+NX

RDI=RGDP-RTAX_SCE1

Щойно згенеровані екзогенні змінні використовуйте для розвитку моделі

RM2=M2/CPI*100

RCON=10.54203219+0.2876768382*RDI+0.01521275802*R-
-0.2653303418*RCON(-1)+0.07822234519*RCON(-4)+
+1.766620998*DUMMY9803

RINV=-10.05593616+0.2013080839*RGDP+0.09684253696*RGDP(-1)+
+0.237424563*R(-1)+0.08645828423*RINV(-1)+0.5802259394*RINV(-4)+
+0.7943412564*LOG(@TREND(1995:4))

M2=0.9006022025+1.699434837*MB-0.1636280921*(R-RNBU)

R=-0.7870717232-0.4015562553*RM2+0.01516727059*RGDP+
+14.67748385*DLOG(CPI)+1.391526878*DUMMY9803+0.3062138166*R(-1)+
+0.202111356*R(-4) + 2.314974518*LOG(@TREND(1994:4))

Розв'яжіть модель самостійно та отримайте прогнознi ряди для всіх ендогенних змінних.

Щоб просимулювати другий сценарій, повторіть кроки, описані в попередньому параграфі, але замість RG_SCE1 використовуйте MB_SCE2. Третій сценарій — це просто комбінація першого і другого: підставте RG_SCE1 замість RG та MB_SCE2 замість MB. Збережіть результати в ряду *під назвою* _SCE3. Загалом модель для третього сценарію має вигляд:

ASSIGN @ALL _SCE3

RGDP=RCON+RINV+RG_SCE1+NX

RDI=RGDP-RTAX_SCE1

RM2=M2/CPI*100

RCON=10.54203219+0.2876768382*RDI+0.01521275802*R-
-0.2653303418*RCON(-1)+0.07822234519*RCON(-4)+
+1.766620998*DUMMY9803

$$RINV = -10.05593616 + 0.2013080839 * RGDP + 0.09684253696 * RGDP(-1) + 0.237424563 * R(-1) + 0.08645828423 * RINV(-1) + 0.5802259394 * RINV(-4) + 0.7943412564 * LOG(@TREND(1995:4))$$

$$M2 = 0.9006022025 + 1.699434837 * MB_SCE2 - 0.1636280921 * (R - RNBU)$$

$$R = -0.7870717232 - 0.4015562553 * RM2 + 0.01516727059 * RGDP + 14.67748385 * DLOG(CPI) + 1.391526878 * DUMMY9803 + 0.3062138166 * R(-1) + 0.202111356 * R(-4) + 2.314974518 * LOG(@TREND(1994:4))$$

Тепер порівняймо результати розрахунків за різними сценаріями. Ми використовуємо розрахунки за нашою моделлю як базовий сценарій для порівняння. Графічно зобразити ряди можна, використовуючи такі команди (можна зробити те саме, виділивши ряди, відкрити їх як групу, клацнувши правою кнопкою мишки на **Open/As a group...**, потім обрати **View/Graph/Line...**).

GRAPH RGDP_SIM.LINE RGDP_SCE0 RGDP_SCE1 RGDP_SCE2 RGDP_SCE3

SHOW RGDP_SIM

Створити новий об'єкт-графік з назвою RGDP_SIM. Тип: лінійний. Потім іде список рядів

Показати графік

Результат:

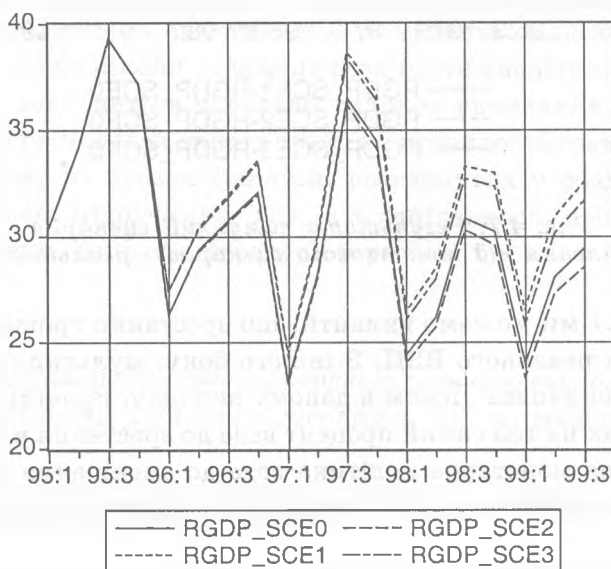


Рис. 4.6. Результати симуляції сценаріїв (рівні): реальний ВВП

Оскільки часто сценарії дуже близькі до результатів базового сценарію або відмінності є дуже малими відносно рівнів, буде доречно представити результати у вигляді відхилень від базового сценарію. Отже, спробуйте запустити

```
GRAPH RGDP_SIM_DEV.LINE RGDP_SCE1-RGDP_SCE0 RGDP_SCE2-
RGDP_SCE0 RGDP_SCE3-RGDP_SCE0
```

```
SHOW RGDP_SIM_DEV.
```

Графічний результат:

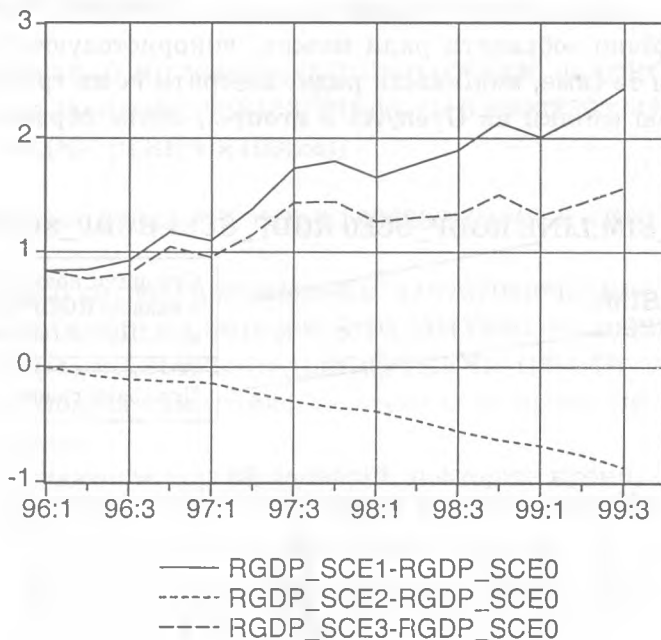


Рис. 4.7. Результати симуляції сценаріїв (відхилення від початкового сценарію): реальний ВВП

З рисунка 4.7 ми можемо виявити, що зростання грошової бази не стимулює зростання реального ВВП. З іншого боку, мультиплікатор збалансованого бюджету виявився дієвим в даному випадку. Зростання в державних витратах і доходах на той самий процент веде до зростання в реальному ВВП. Проте експансійна монетарна політика веде до зменшення процентних ставок, тоді як експансійна фіскальна політика підвищує процентні ставки (див. рисунок 4.8).

Процедура, використана в цій дуже простій симуляції, часто застосовується, щоб визначити різні мультиплікатори, наприклад, мультиплікатор

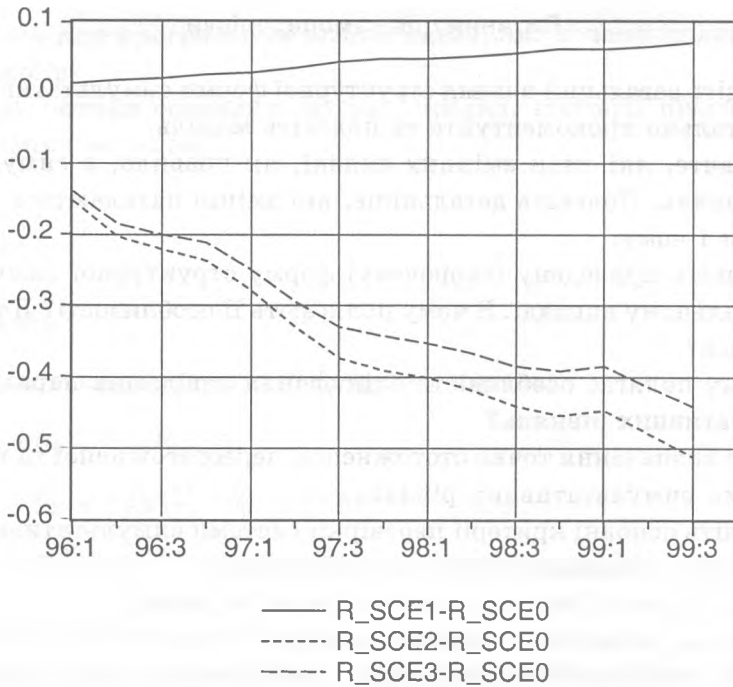


Рис. 4.8. Результати симуляції сценаріїв (відхилення від базового сценарію): процентна ставка

державних витрат. В простій моделі може бути знайдено аналітичний розв’язок. Типовий розв’язок для такого (скорочена форма) взаємозв’язку — це рівняння різниці. Проте на практиці навіть маленькі моделі, що мають нелінійні зв’язки, майже ніколи не можна розв’язати аналітично. Як наслідок, мультиплікатори знаходяться чисельно, шляхом виконання різноматнітних симуляцій. Взагалі в моделях такого типу, як правило, розраховують функції імпульсного відгуку на зразок функцій, розглянутих у розділах 2—3. Але це тема для окремого підручника. Тим, хто зацікавився, пропонуємо прочитати книгу [15] (див. список літератури в розділі 1).

Завдання

Здійснити симуляцію для вашого власного сценарію економічної політики і порівняти результати з базовим сценарієм. Які ви можете зробити висновки?

Для розрахунків використовуйте дані, наведені в додатку 4. Цю таблицю рекомендується поступово оновлювати.

Питання для самоперевірки

1. Наведіть загальний вигляд структурної форми симультазивної системи рівнянь. Детально прокоментуйте та поясніть модель.
2. Визначте, які типи змінних наявні, як правило, в симультазивних системах рівнянь. Поясніть детальніше, які змінні називаються попередньо визначеними і чому.
3. Запишіть приведену (скорочену) форму структурної симультазивної моделі в загальному вигляді. В чому полягають її особливості? Для чого вона застосовується?
4. В чому полягає особливість оцінювання невідомих параметрів моделей симультазивних рівнянь?
5. Дайте визначення точно ототожненої, переототожненої та недоототожненої системи симультазивних рівнянь.
6. Наведіть основні критерії перевірки системи симультазивних рівнянь на ототожненість. Прокоментуйте їх особливості.
7. Наведіть два різних визначення умови порядку.
8. Поясніть, в чому полягає суть рангової умови ототожнення.
9. Яка з умов рангу або порядку є необхідною та достатньою умовою ототожнення?
10. В чому полягає проблема ендогенності в симультазивних системах рівнянь?
11. Що таке інструментальні змінні? Детально поясніть особливості їх застосування.
12. Наведіть алгоритм оцінювання невідомих параметрів регресійного рівняння двокроковим методом найменших квадратів. У яких випадках він застосовується?
13. Наведіть схему перевірки системи симультазивних рівнянь на ототожненість за ранговою умовою.
14. Поясніть суть методу непрямих найменших квадратів (МНК). В яких випадках він може бути застосований?
15. Наведіть основні особливості двокрокового методу найменших квадратів.
16. Поясніть суть трикрокового методу найменших квадратів (ЗМК). В яких випадках він може бути застосований?
17. Наведіть приклад рекурсивної моделі симультазивних рівнянь. Поясніть основні її особливості. Чи можуть бути оцінені невідомі параметри рекурсивної системи методом найменших квадратів (МК)? Поясніть чому.
18. Коротко поясніть різницю між поняттями «історична симуляція», «ex post-прогноз» та «ex ante-прогноз».
19. Коротко поясніть, яким чином симультазивні моделі можуть вико-

ривовуватись для програвання різних сценаріїв. У чому полягає суть розробки сценаріїв?

Наведіть чотири основних, на ваш погляд, критерії прогновної якості симультаивних моделей.

Вправи

Вправа 4.1.

Розгляньте таку макроеконометричну модель:

$$C_t = \alpha_1 + \alpha_2 Y_t + \alpha_3 R_t + \varepsilon_{1t}$$

$$I_t = \beta_1 + \beta_2 R_t + \beta_3 (Y_t - Y_{t-1}) + \varepsilon_{2t}$$

$$R_t = c_1 + c_2 I_t + c_3 M_t + \varepsilon_{3t}$$

$$Y_t = C_t + I_t + G_t,$$

де C_t — витрати на споживання в період t ; I_t — інвестиції в період t ; Y_t — національний дохід в період t ; R_t — процентна ставка в період t ; G_t — витрати уряду в період t ; M_t — пропозиція грошей в період t .

а) Що означає перетотоженість рівняння? Концептуально (не за допомогою математичних викладок) поясніть, чому перетотоженість є проблемою в прикладній економіці.

б) Чи вирішує двокроковий метод найменших квадратів проблему перетотоженості? Поясніть.

в) Використовуючи умову рангу та умову порядку, покажіть, яке з рівнянь (якщо є) є недоототожненим. Точно ототожненим? Переототожненим?

г) Показати, що застосування методу найменших квадратів (МНК) у рівнянні процентної ставки дає зміщені та неконсистентні оцінки параметрів.

д) Уявіть, що ви маєте тільки одне рівняння процентної ставки. Поясніть, яким чином перевірити можливу проблему симультаивності, пов'язану з наявністю в рівнянні витрат на інвестиції. Як розв'язати цю проблему?

Вправа 4.2.

Наведено два визначення умови порядку.

Визначення 1. В моделі, що складається з M симультаивних рівнянь, для того, щоб рівняння було ототожненим (identified), воно повинно виключати не менше ніж $M - 1$ змінних (як ендогенних, так і предетермінованих (predetermined)), що наявні в цілому в системі.

Визначення 2. В моделі, що складається з M симультаивних рівнянь,

для того, щоб рівняння було ототожненим, кількість предетермінованих змінних, виключених з рівняння, повинна бути не меншою за кількість ендогенних змінних, включених в рівняння, мінус одиниця, тобто:

$$K - k \geq m - 1,$$

де m — кількість ендогенних змінних у рівнянні, що розглядається; K — кількість предетермінованих змінних моделі; k — кількість предетермінованих змінних у рівнянні, що розглядається.

Покажіть, що обидві умови порядку є еквівалентними.

Вправа 4.3.

Нехай ми маємо певну структурну симульативну модель:

$$Y_{1t} = \beta_{10} + \beta_{12}Y_{2t} + \gamma_{11}X_{1t} + u_{1t}$$

$$Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} + \gamma_{22}X_{2t} + u_{2t},$$

з якої було отримано таку оцінену приведену форму:

$$\hat{Y}_{1t} = 4 + 3X_{1t} + 8X_{2t}$$

$$\hat{Y}_{2t} = 2 + 6X_{1t} + 10X_{2t}.$$

Виходячи зі значень параметрів приведеної форми, розрахуйте параметри структурної форми.

Вправа 4.4.

Визначте, чи є кожне з наведених нижче тверджень правильним.

- Метод найменших квадратів не можна застосовувати для оцінки невідомих параметрів рівняння симульативної моделі в структурній формі.
- У випадку, коли рівняння є недоототожненим, ми не можемо застосовувати двокроковий метод найменших квадратів.
- Проблема симульативності не виникає в рекурсивній симульативній моделі.
- Двокроковий метод найменших квадратів та інші методи оцінювання структурних рівнянь симульативної системи мають бажані статистичні властивості тільки для великих вибірок.
- Якщо рівняння точно ототожнено, тоді непрямий метод найменших квадратів та двокроковий метод найменших квадратів дають однакові результати при оцінюванні невідомих параметрів.

- е) Якщо рівняння переототожено, тоді непрямий метод найменших квадратів та двокроковий метод найменших квадратів дають однакові результати при оцінюванні невідомих параметрів.
- є) Якщо рівняння не має проблеми симультативності, то МНК та 2МНК дають однакові результати.

Вправа 4.5.

Чому не обов'язково застосовувати двокроковий метод найменших квадратів до точно ототожнених (ідентифікованих) рівнянь? Поясніть це на умовному прикладі.

Вправа 4.6.

Розгляньте таку модифіковану кейнсіанську модель визначення доходу:

$$C_t = \beta_{10} + \beta_{11}Y_t + u_{1t}$$

$$I_t = \beta_{20} + \beta_{21}Y_t + \beta_{22}Y_{t-1} + u_{2t}$$

$$Y_t = C_t + I_t + G_t,$$

де C — витрати на споживання; I — витрати на інвестиції; Y — доход; G — витрати уряду. Зазначимо, що G_t та Y_{t-1} є попередньо визначеними змінними моделі.

- а) Запишіть рівняння у приведеній формі та визначте, які з них є точно ототожненими, а які переототожненими;
- б) який з методів ви використаєте для оцінки точно ототожненого та переототожненого рівняння? Обґрунтуйте свою відповідь.

Вправа 4.7.

Розгляньте результати оцінювання двох рівнянь різними методами, зокрема МНК та 2МНК:

МНК: $W_t = 0.276 + 0.258P_t + 0.046P_{t-1} + 4.959V_t, R^2 = 0.924.$

МНК: $P_t = 2.693 + 0.232W_t - 0.544X_t + 0.247M_t + 0.064M_{t-1}, R^2 = 0.982.$

2МНК: $W_t = 0.272 + 0.257P_t + 0.046P_{t-1} + 4.966V_t, R^2 = 0.920.$

2МНК: $P_t = 2.686 + 0.233W_t - 0.544X_t + 0.246M_t + 0.046M_{t-1}, R^2 = 0.981,$

де W_t, P_t, M_t та X_t — це процентні зміни заробітної плати, цін, цін на імпорт та продуктивності праці (всі процентні зміни дані відносно попереднього періоду), і де V_t відображає незаповнені робочі вакансії (в процентах до загальної кількості робітників).

«Оскільки, як можна побачити, результати оцінювання за МНК та 2МНК практично ідентичні, то немає сенсу застосовувати двокроковий метод найменших квадратів у даному випадку». Прокоментуйте дане твердження.

Вправа 4.8.

Припустіть, що виробництво характеризується виробничою функцією Кобба—Дугласа:

$$Q_i = AK_i^\alpha L_i^\beta,$$

де Q — виробництво; K — затрати капіталу; L — затрати робочої сили; A , α , β — параметри; i — i -та фірма.

При ціні кінцевого виробництва P , ціні робочої сили W і ціні капіталу R та за умов максимізації прибутку ми отримуємо таку емпіричну виробничу модель:

функція виробництва:

$$\ln Q_i = \ln A + \alpha \ln K_i + \beta \ln L_i + \ln u_{1i}, \quad (4.2)$$

граничний продукт функції робочої сили:

$$\ln Q_i = -\ln \beta + \ln L_i + \ln \frac{W}{P} + \ln u_{2i}, \quad (4.3)$$

граничний продукт функції капіталу:

$$\ln Q_i = -\ln \alpha + \ln K_i + \ln \frac{R}{P} + \ln u_{3i}, \quad (4.4)$$

де u_1 , u_2 та u_3 — випадкові величини (disturbances).

В даному разі ми маємо симультативну модель з трьох рівнянь та з трьома ендогенними змінними Q , L і K . Змінні P , R та W є екзогенними.

а) З якими проблемами ви зустрічаєтесь, оцінюючи модель, якщо $\alpha + \beta = 1$, тобто коли ми маємо постійну віддачу від масштабу?

б) Чи можете ви оцінити рівняння, навіть якщо $\alpha + \beta \neq 1$? Дайте відповідь, використовуючи поняття ототожненості системи.

с) Якщо система є неототоженою, що можна зробити для її ототожнення?

Зауваження. Рівняння (4.3) і (4.4) отримано в результаті диференціювання Q по змінних праці та капіталу (відповідно прирівнювали їх до W/P та R/P) і подальшої трансформації результуючих виразів у логарифми з додаванням логарифмів випадкових величин (disturbance terms).

Вправа 4.9.

Розгляньте модель попиту і пропозиції грошей:

$$\text{попит на гроші: } M_t^D = \beta_0 + \beta_1 Y_t + \beta_2 R_t + \beta_3 P_t + u_{1t},$$

$$\text{пропозиція грошей: } M_t^S = \alpha_0 + \alpha_1 Y_t + u_{2t},$$

де M — гроші; Y — доход; R — процентна ставка; P — ціна.

Припустіть, що R і P є попередньо визначеними змінними.

- a) Чи є ототожненою функція попиту?
- b) Чи є ототожненою функція пропозиції?
- c) Який метод ви використаєте, щоб оцінити параметри ототожненого рівняння? Чому?
- d) Припустимо, що ми модифікуємо функцію пропозиції, додаючи пояснюючі змінні Y_{t-1} та M_{t-1} . Як це вплине на проблему ототожненості рівнянь? Чи будете ви тепер використовувати метод, який ви використовували в пункті c)? Чому?

Вправа 4.10.

Розгляньте систему симультативних рівнянь:

$$R_t = \beta_0 - \beta_1 M_t + \beta_2 Y_t + u_{1t}$$

$$Y_t = \alpha_0 - \alpha_1 R_t + u_{2t},$$

де M_t (пропозиція грошей) є екзогенною змінною, R_t — процентна ставка, а Y_t — ВВП.

- a) Поясніть, що визначає дана модель.
- b) Чи ототожені рівняння даної системи?
- c) Оцініть параметри ототожнених рівнянь, використовуючи наведені нижче дані з таблиці 4.15 (дані умовні).
- d) Поясніть метод, який ви використовували для оцінювання рівняння.
- e) Припустимо, що ми змінюємо модель попередньої вправи таким чином:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + \beta_3 Y_{t-1} + u_{1t}$$

$$Y_t = \alpha_0 - \alpha_1 R_t + u_{2t}.$$

1. Визначте, чи є система ототожненою.
2. Оцініть параметри ототожнених рівнянь, використовуючи дані таблиці 4.15.

Таблиця 4.15. Макроекономічні показники 1980—2001 рр., млрд у. о.

Рік	Валовий внутрішній продукт (Y)	Пропозиція грошей (M)	Валові приватні внутрішні інвестиції (I)	Державні федеральні витрати (x_2)	Процентна ставка (R)
1980	1010,7	628,1	150,3	208,5	6.562
1981	1097,2	717,2	175,5	224,3	4.511
1982	1207,0	805,2	205,6	249,3	4.466
1983	1349,6	861,0	243,1	270,3	7.178
1984	1458,6	908,6	245,8	305,6	7.926
1985	1585,9	1023,3	226,0	364,2	6.122
1986	1768,4	1163,7	286,4	392,7	5.266
1987	1974,1	1286,6	358,3	426,4	5.510
1988	2232,7	1388,7	434,0	469,3	7.572
1989	2488,6	1496,7	480,2	520,3	10.017
1990	2708,0	1629,5	467,6	613,1	11.374
1991	3030,6	1792,9	558,0	697,8	13.776
1992	3149,6	1951,9	503,4	770,9	11.084
1993	3405,0	2186,1	546,7	840,0	8.750
1994	3777,2	2374,3	718,9	892,7	9.800
1995	4038,7	2569,4	714,5	969,9	7.660
1996	4268,6	2811,1	717,6	1028,2	6.030
1997	4539,9	2910,8	749,3	1065,6	6.050
1998	4900,0	3071,1	793,6	1109,0	6.920
1999	5250,8	3227,3	832,3	1181,6	8.040
2000	5522,2	3339,0	799,5	1273,6	7.470
2001	5677,5	3439,8	721,1	1332,7	5.490

Вправа 4.11.

Розгляньте наступну модель симультативних рівнянь:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + u_{1t}$$

$$Y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + u_{2t},$$

де змінні визначені у вправі 4.10. Якщо вважати, що I (внутрішні інвестиції) та M є екзогенними змінними, визначте, чи є система ототожненою. Оцініть параметри ототожнених рівнянь, використовуючи дані з таблиці 4.15.

Вправа 4.12.

Припустимо, що ми змінюємо модель у вправі 4.10 так:

$$R_t = \beta_0 + \beta_1 M_t + \beta_2 Y_t + u_{1t}$$

$$Y_t = \alpha_0 + \alpha_1 R_t + \alpha_2 I_t + u_{2t}$$

$$I_t = \gamma_0 + \gamma_1 R_t + u_{3t}.$$

Припустимо також, що M є екзогенною змінною.

а) З'ясуйте, які з рівнянь даної системи є ототожненими.

б) Оцініть параметри ототожнених (ідентифікованих) рівнянь, використовуючи дані з таблиці 4.15. Поясніть, які методи для оцінювання невідомих параметрів ви використовували.

Вправа 4.13.

Розглянемо модель попиту-пропозиції у вигляді:

$$\text{функція попиту: } Q_t = \beta_0 + \beta_1 P_t + \beta_2 X_t + u_{1t},$$

$$\text{функція пропозиції: } Q_t = \beta_0 + \beta_1 P_t + u_{2t},$$

де Q_t — обсяг, P_t — ціна, X_t — доход або витрати.

а) Визначте, яке з рівнянь буде ототожненим. Поясніть чому.

б) Припустіть, що функція пропозиції змінюється таким чином:

$$Q_t = \beta_0 + \beta_1 P_{t-1} + u_{2t},$$

де P_{t-1} — ціна у попередньому періоді.

с) Якщо X (витрати) та P_{t-1} є попередньо визначеними, чи буде мати місце проблема симультативності?

д) Якщо так, то чи будуть функції попиту та пропозиції ототожненими? Якщо вони будуть ототожненими, поясніть, якими методами ви можете їх оцінити. Чи зможете ви з коефіцієнтів приведеної форми рівнянь даної системи отримати структурні коефіцієнти? Якщо так, то покажіть, яким чином.

Список літератури

1. A Guide to FRB/US // A Macroeconomic Model of the United States — Federal Reserve Board.— Washington, D. C., 1996.— 45 p.
2. Begg D. Monetary Policy During Transition: Progress and Pitfalls in Central and Eastern Europe // Oxford Review of Economic Policy, 1990, № 6, Vol. 13 (2).— P. 33—46.
3. Blanchard Oliver J. What is Left of the Multiplier Accelerator // The American Economic Review, 1981, May, Vol. 71, Issue 2.— P. 150—154
4. Dornbush R., Fisher S., Startz R. Macroeconomics.— 7th ed.— Boston : Irwin McGraw-Hill, 1998.— 550 p.
5. Frumkin N. Guide to Economic Indicators.— 2nd ed.— M. E. Sharpe, 1994.— 430 p.

6. *Geiger Linwood T.* Macroeconomic Analysis and Transitional Economy.— Pennsylvania : St. Davids, 1992.
7. *Gronicki Mirek, Piętko Katarzyna.* Interest Rates and Real Economy in Ukraine // Ukrainian Economic Trends, 1998, № 1.— P. 4—5
8. *Gronicki Miroslaw, Piętko Katarzyna.* Macroeconomic Model for Ukraine // CASE Working Paper, 1999.— 56 p.
9. *Gros D., Steinherr A.* Winds of Change: Economic Reforms in Central and Eastern Europe.— London : Longman, 1995.
10. *Gujarati Damodar N.* Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, 1995.— 838 p.
11. *Greene William H.* Econometric Analysis.— 4th ed.— New Jersey : Prentice Hall, 2000.— 1004 p.
12. *Hall Robert E.* Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence // The Journal of Political Economy, 1978, Vol. 86(6).— P. 971—987.
13. *Hall Robert E., Mishkin Frederic S.* The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households // Econometrica, 1982, Vol. 50(2).— P. 461—482
14. *Kaufmann D., Kaliberda A.* Integrating the Unofficial Economy into Dynamics of Post-Socialist Economies: a Framework of Analysis and Evidence // Economic Transition in Russia and the New States of Eurasia, 1996, August.
15. *Kennedy P.* A Guide to Econometrics.— Cambridge, Massachusetts : The MIT Press, 1998.— 468 p.
16. *Klein L.* (editor) Comparative Performance of U. S. Econometric Models.— Oxford : Oxford University Press, 1991.
17. *Krugman Paul R.* Obstfeld Maurice International Economics: Theory and policy.— 3rd ed.— Harper Collins College Publishers, 1994.
18. *Lucas R. E.* Econometric Policy Evaluation: A Critique // *K. Brunner, A. H. Meltzer.* The Philips Curve and Labor Markets. Carnegie-Rochester Conference Series on Public Policy, 1976, № 1. — P. 19—46.
19. Macroeconometric Modeling and Its Use for the Conduct of Policy in Ukraine // TACIS, Ukrainian Economic Trends, 1998, № 9.— P. 82—86.
20. *Mankiw N. Gregory.* Macroeconomics.— 2nd ed.— New York : Worth Publishers, Inc., 1992.— 514 p.
21. *Mankiw Gregory, Campbell John.* Consumption, Income, and Interest rates. Reinterpreting the Time Series Evidence // NBER Working Papers, 1989, № 2924.
22. *McPherson Malcolm F., Rakovski Tzvetana S.* A Small Econometric Model of the Zambian Economy // HIID Development Discussion Papers, 1999, January, № 672.— 30 p.
23. *Pindyck R. S., Rubinfeld D. L.* Econometric Models and Economic Forecasts.— 3rd ed.— New York : McGraw-Hill, 1991.
24. *Qaunes A., Thakur S.* Macroeconomic Accounting and Analysis in

Transition Economies.— Washington, D.C. : International Monetary Fund, 1997.— 183 p.

25. *Ray C.* Fair Testing Macroeconometric Models.— Harvard University Press, 1994.— 421 p.

26. *Ross S. A., Westerfield R. W., Jaffe J.* Corporate Finance.— 4th ed.— Chicago : Irwin, 1996.— 899 p.

27. *Sachs Jeffrey D., Larrain Felipe B.* Macroeconomics in the global economy.— New Jersey : Prentice-Hall, 1993.— 3778 p.

28. *Steinherr A.* Banking Reforms in Eastern European Countries // Oxford Review of Economic Policy, Vol. 13, № 2.— P. 106—125.

29. *Taylor L.* Income Distribution, Inflation and Growth: Lectures on Structuralist Macroeconomic Theory.— Cambridge, Massachusetts : MIT, 1991.— 543 p.

30. Ukraine: Recent Economic Developments. IMF Staff Country Paper, 1999, № 99/42.

31. TACIS. Ukrainian Economic Trends. Monthly Update, 1999, № 5.

32. *Varian Hal R.* Intermediate Microeconomics: A Modern Approach.— 5th ed.— New York : W. W. Norton & Company, 1999.— 711 p.

33. *Аллен Р.* Математическая экономия / Пер. с англ; Под ред. А. Л. Вайнштейна.— М. : Изд-во иностранной литературы, 1963.— 599 с.

34. Бураковський І. Теорія міжнародної торгівлі.— К. : Основи, 1996.— 241 с.

35. Бюджетний і податковий огляд. Третій квартал 1998 р.: Аналітичні статті, коментарі, таблиці: Група фіскального аналізу при Комітеті Верховної Ради України з питань бюджету.

36. Бюджетний і податковий огляд. Четвертий квартал 1998 р.: Аналітичні статті, коментарі, таблиці: Група фіскального аналізу при Комітеті Верховної Ради України з питань бюджету.

37. *Геєць В., Скрипниченко М., Соколик М., Шумська С.* Секторальні макромоделі прогнозування економіки України // Економіст, 1998, № 5.— С. 58—67.

38. *Гончар В.* Моделювання інфляції в Україні // Вісник НБУ, 1998, № 2.— С. 52—56.

39. *Гроніцкі М., Ширмер Я., П'єнтка К.* Альтернативні сценарії розвитку економіки України у 1998—1999 роках // Економіст, 1998, № 5.— С. 72—76.

40. Декрет Кабінету Міністрів України «Про прибутковий податок з громадян» // ВВР, 1993, № 10.

41. *Джонстон Дж.* Эконометрические методы / Пер. с англ. и предисловие Рыбкина А. А.— М. : Статистика, 1980.— 444 с.

42. Закон України «Про оподаткування прибутку підприємств» // ВВР, 1995, № 4.

43. Закон України «Про податок на додану вартість» // ВВР, 1997, № 21.

44. *Ільчук С.* Погані борги в Україні створює держава // Перспективні дослідження. I квартал 1999 р., № 1.— С. 17—33.

45. *Кваджа С., Лук'яненко І. Г., Городніченко Ю. О.* Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К. : Видавничий дім «КМ Academia», 2000.— 204 с.
46. *Ковальчук Т., Коваль Н.* Основні чинники інфляції в Україні // *Економіка України*, 1996, № 3, 4.— С. 31—40, 22—32.
47. *Корольов О. А.* Економетрія. Навчальний посібник: КНТУ.— К., 2000.— 660 с.
48. *Корнаи Я.* Дефіцит.— М. : Наука, 1992.— 608 с.
49. *Крючкова И. П.* Инфляция и процентные ставки // *Деньги и кредит*, 1995, № 3.— С. 20—25
50. *Крючкова І. В.* Прогноз основних макроекономічних показників України до 2000 року: Інститут економічного прогнозування НАН України.— К., 1999.— 66 с.
51. *Лук'яненко І. Г.* Попит на гроші. Ілюстрація класичного підходу // *Наукові записки НаУКМА. Економіка*, т. 16.— К. : Видавничий дім «Педагогіка», 2000.— С. 38—42.
52. *Лук'яненко І. Г., Городніченко Ю. О.* Економетричне моделювання процесів у грошово-кредитній сфері // *Наукові записки НаУКМА. Економіка*, т. 15.— К. : Видавничий дім «Педагогіка», 1999.
53. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика. Теорія та практика.— К.: Знання, 1998.— 493 с.
54. *Лук'яненко І. Г., Краснікова Л. І.* Економетрика. Практикум з використанням комп'ютера.— К. : Знання, 1998.— 217 с.
55. *В. Ф. Бесєдін, І. В. Богдан, Л. Л. Варениченко та ін.* Основні прогностичні показники економічного і соціального розвитку України на 2000 рік // *За ред. В. Ф. Бесєдіна*.— К. : Інтелект, 1999.— 159 с.
56. *Наконечний С. І., Терещенко Т. О., Романюк Т. П.* Економетрія. Підручник. КНЕУ.— К., 2000.— 296 с.
57. *Райская Н.* Временные лаги в динамике инфляции // *Вопросы экономики*, 1996, № 8.— С. 34—41
58. *Рывок в рыночную экономику*.— К. : Феникс, 1997.— 304 с.
59. Указ Президента України № 857/98 «Про деякі зміни в оподаткуванні» від 7 серпня 1998 р.
60. *Холден К., Піл Д. А., Томпсон Дж. Л.* Економічне прогнозування. Вступ.— К. : Інформтехніка-ЕМЦ, 1996.
61. *Черваньов Д. М., Комашко О. В.* Економетрика. Курс лекцій.— РВЦ КІЕМБСС, 1998.— 61 с.
62. *Экономическая статистика* / Под ред. Иванова Ю. Н.— М. : Инфра-М, 1998.— 480 с.
63. *Єлейко В.* Основи економетрії.— Львів : Марка, 1995.

РОЗДІЛ 5. ЗАСТОСУВАННЯ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ В ЕМПІРИЧНИХ ДОСЛІДЖЕННЯХ

5.1. ОСОБЛИВОСТІ МОДЕЛЮВАННЯ ЗА ДОПОМОГОЮ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ. ПРИКЛАДИ НАЙПРОСТІШИХ МОДЕЛЕЙ

Моделі, які базуються на використанні лонгитюдних (панельних) даних (panel data model), досить широко застосовуються при моделюванні процесів на мікрорівні, особливо при моделюванні доходів та споживання домогосподарств, аналізі ринку праці тощо. Сьогодні, у зв'язку з надзвичайною актуальністю підвищення ефективності регіональної політики у всіх колишніх країнах Радянського Союзу, значення сучасного апарату моделювання на основі панельних даних стрімко зростає. Дійсно, регіональні соціально-економічні та політичні процеси легше моделювати та аналізувати, використовуючи саме цей апарат. В чому ж полягає особливість моделювання економічних процесів на основі моделей лонгитюдних даних, в чому їх перевага порівняно з іншими регресійними моделями? Які типи моделей на основі лонгитюдних даних існують та які методи оцінювання необхідно застосовувати? Яким чином підбирати відповідний тип лонгитюдної моделі? Спробуємо поступово відповісти на всі поставлені запитання. Крім того, проілюструємо та прокоментуємо застосування моделей лонгитюдних даних як в нашій країні, так і за кордоном.

Лонгитюдні дані є спостереженнями за певною ознакою об'єктів однієї сукупності (наприклад: індивідууми, домашні господарства, фірми, регіони, країни тощо) протягом рівновіддалених періодів часу. Фактично лонгитюдні дані є комбінацією варіаційних (крос-секційних) та часових рядів, наприклад середня заробітна плата в регіонах за 1994—2001 рр.; щоквартальні за період 1997—2001 рр. надходження податку з прибутку підприємств в розрізі регіонів України; доходи домогосподарств України протягом певних років та інші. Незважаючи на те що лонгитюдні дані, як правило, фіксуються на мікрорівні, останнім часом все поширенішим стає групування часових рядів однакових показників для певних регіонів, країн або галузей промисловості

та їх одночасний аналіз. Можливість використання лонгітюдних даних дозволяє економістам будувати та оцінювати більш складні та більш реалістичні моделі, ніж просто моделі, які будуються окремо на основі часових або варіаційних рядів.

В загальному вигляді модель лонгітюдних даних може бути представлена так:

$$Y_{it} = \alpha + X'_{it}\beta_{it} + \varepsilon_{it}, \quad (5.1)$$

де Y_{it} — значення досліджуваного показника для i -го об'єкта (індивідуума, регіону, фірми тощо) в t -й період часу; $I = 1, 2, 3, \dots, N$; $t = 1, 2, 3, \dots, T$; $X'_{it} = \{X_{1it}, X_{2it}, \dots, X_{kit}\}$ — вектор порядку $(k \times 1)$ пояснюючих змінних (факторів); ε_{it} — збурення для i -го об'єкта (індивідуума, регіону, фірми тощо) в t -й період часу; $I = 1, 2, 3, \dots, N$; $t = 1, 2, 3, \dots, T$; α — скаляр; β_{it} — параметри моделі, що вимірюють часткові ефекти від зміни X_{it} в період t для певного i .

Як можна побачити з (5.1), регресійні моделі лонгітюдних даних відрізняються по зовнішньому вигляду від моделей на основі часових або крос-секційних даних тим, що мають подвійний запис змінних, тобто i -й запис дає інформацію в поперечному розрізі (cross-section dimension — крос-секційний вимір), в той час як t позначає час.

Модель (5.1) є занадто загальною, тому необхідно ввести додаткові обмеження на параметри моделі для її практичного застосування. Стандартним припущенням, дійсним для багатьох емпіричних ситуацій, є припущення постійності параметрів β_{it} для всіх значень t та i . При такому припущенні модель (5.1) набуває вигляду:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it}, \quad (5.2)$$

де Y_{it} — значення залежної змінної для i -го об'єкта у t -й період часу, α та β_j — невідомі параметри моделі, X_{jit} — значення j -го фактора для i -го об'єкта у t -й період часу, ε_{it} — значення випадкової величини; $i = \overline{1, N}$; $t = \overline{1, T}$; $j = \overline{1, k}$.

Модель типу (5.2) можна розглядати як загальну модель лонгітюдних даних (pooled model), яка фактично є регресійною моделлю на множині лонгітюдних даних. Побудова, оцінювання параметрів та дослідження такої моделі нічим не відрізняються від класичних багатофакторних регресійних моделей.

Специфіка моделей лонгітюдних даних надає можливість додатково розбивати збурення на декілька компонентів та відповідно перетворювати модель (5.2) в моделі лонгітюдних даних з одно- та двокомпонентною помилками.

Найбільш поширеними на практиці є моделі з одновимірною складовою помилки (або одновимірними шоками — one-way error component model), тобто моделі, в яких ми можемо виділити один окремих компонент збурення моделі (5.2):

$$\varepsilon_{it} = \mu_i + u_{it}, \quad (5.3)$$

де μ_i позначає неспостережувані специфічні індивідуальні ефекти, а u_{it} — залишкові шуми (remainder disturbances). Наприклад, при дослідженні ринку праці, аналізі впливу різних факторів на заробітну плату голови домогосподарства Y_{it} в моделі (5.2) може відображати дохід, а X_{it} може містити набір змінних, які позначають досвід, освіту, участь у спілках та об'єднаннях, стать, расу тощо. При цьому якщо збурення моделі представлені у вигляді (5.3), то μ_i , які є інваріантними стосовно часу t , можуть пояснювати будь-які індивідуальні специфічні ефекти, не враховані в регресійній моделі. В цьому разі їх можна розглядати як властивості окремого представника сукупності, які не піддаються спостереженню та вимірюванню. Залишкові шуми u_{it} змінюються в залежності від представника сукупності і можуть розглядатись як звичайне збурення в регресійній моделі.

Іншим прикладом необхідності виділення однокомпонентної помилки є приклад виробничої функції. Для її побудови та оцінювання можна використати інформацію про діяльність фірми за певний проміжок часу. При цьому Y_{it} можна розглядати як залежну змінну, що вимірює випуск продукції різних фірм за певні періоди часу (output), а X_{it} — витрати фірм (input). Неспостережувані специфічні ефекти діяльності фірм можна відобразити за допомогою параметрів μ_i , і в цьому випадку вони можуть позначати, наприклад, управлінські навички персоналу різних фірм. Одні з перших моделей з одновимірною складовою помилки (або одновимірними шоками — one-way error component) були застосовані в [21] при дослідженні інвестицій, в [23] та [15] при побудові виробничих функцій, а також в [25] для аналізу попиту на природний газ.

Крім моделей лонгитюдних даних з одновимірною складовою помилки існують моделі з двовимірною складовою помилки (two-way error component). У випадку моделей з двовимірною складовою помилки модель (5.2) можна представити у вигляді:

$$\varepsilon_{it} = \mu_i + \eta_t + u_{it}, \quad (5.4)$$

де μ_i позначає неспостережувані специфічні індивідуальні ефекти; η_t — неспостережувані часові ефекти; а u_{it} — залишкові шуми.

Моделі з двовимірною складовою помилки відрізняються від моделей з одновимірною складовою помилки тим, що дозволяють враховувати додатково специфічні часові ефекти (завдяки параметру η_t), наприклад періоди страйку, ембарго на ввіз певної продукції тощо.

Зауважимо, що ми зосередимо свою увагу лише на лонгітюдних моделях з одновимірною складовою помилки, а моделі з двовимірною складовою помилки розглядати не будемо.

Моделі лонгітюдних даних як з одновимірною, так і з двовимірною складовою помилки, в свою чергу, розподіляються на моделі з фіксованими ефектами та моделі з випадковими ефектами.

Різниця між ними залежить від того, розглядаються відповідно виділені компоненти збурень як фіксовані величини чи як випадкові величини.

Якщо ми розглядаємо μ_i (або μ_i та η_t в моделі з двовимірною складовою помилки) як N невідомих фіксованих параметрів, то модель (5.2) зі збуреннями у вигляді (5.3) належатиме до класу стандартних моделей лонгітюдних даних з фіксованими ефектами (fixed effects models) та може бути записана у вигляді:

$$Y_{it} = \mu_i + X_{it}'\beta + u_{it}, \quad (5.5)$$

де $X_{it}' = \{X_{1it}, X_{2it}, \dots, X_{kit}\}$ — вектор порядку $(k \times 1)$ пояснюючих змінних (факторів) без константи; $\beta = \{\beta_1, \beta_2, \dots, \beta_k\}$ — вектор невідомих параметрів у припущенні, що ефект від зміни X є однаковим для всіх i та всіх періодів часу t ; $i = 1, 2, 3, \dots, N$; $t = 1, 2, 3, \dots, T$. Ще раз підкреслимо, що припущення, за яким середній рівень для i -го об'єкта може відрізнитись від середнього рівня j -го, але є постійним для різних періодів часу, вводиться в модель за допомогою різних значень перетину μ_i , $i = 1, 2, 3, \dots, N$.

Особливість моделі (5.5) полягає в тому, що для різних об'єктів буде оцінено різні значення констант μ_i , які відображають вплив факторів, що є специфічними саме для них. У той же час оцінені параметри β_j будуть однаковими для всіх об'єктів та всіх періодів часу.

Крім того, іншою особливістю моделей лонгітюдних даних з фіксованими ефектами є відсутність спільного перетину α , оскільки при його наявності ми потрапляємо в пастку досконалої мультиколінеарності. Ми розглянемо цю ситуацію детальніше в наступному підрозділі при аналізі моделей лонгітюдних даних з фіксованими ефектами.

В стандартному випадку припускається, що збурення u_{it} моделі (5.5) є незалежними однаково розподіленими випадковими величинами для всіх об'єктів та періодів часу з нульовим математичним сподіванням та постійною дисперсією σ_u^2 .

Якщо припускається, що μ_i можуть тлумачитись як реалізація незалежних від X_{it} випадкових величин з середнім α та дисперсією σ_μ^2 , то модель (5.3) належить до класу стандартних моделей лонгитюдних даних з випадковими ефектами (random effects models). У моделях з випадковими ефектами перетини μ_i тлумачаться як випадкові величини, що мають нульове математичне сподівання, а збурення u_{it} припускаються некорельованими для різних періодів часу. За такими припущеннями модель з випадковими ефектами має вигляд:

$$Y_{it} = \alpha + X_{it}'\beta + \mu_i + u_{it}, \quad (5.6)$$

де α є спільним перетином. Отже, в моделі з випадковими ефектами, на відміну від моделей з фіксованими, можна виділити та окремо оцінити спільний перетин.

Застосування моделей лонгитюдних даних як з фіксованими, так і з випадковими ефектами в емпіричних дослідженнях має значні переваги над іншими підходами. Розглянемо та прокоментуємо їх детальніше.

Насамперед, застосування моделей лонгитюдних даних дозволяє виявити та проаналізувати зміни на індивідуальному рівні, що неможливо ні в рамках окремих моделей часових рядів, ні в рамках моделей варіаційних рядів. Проілюструємо цю перевагу на досить простому прикладі. Припустимо, що середній дохід у домогосподарствах зріс на 2,5 % порівняно з минулим роком. Моделі лонгитюдних даних дозволяють виявити, чим саме викликане таке зростання: наприклад тим, що дохід в кожному домогосподарстві збільшився на 2,5 %, чи тим, що в половині домогосподарств він збільшився на 5 %, а в половині залишився без зміни (можна розглядати й інші комбінації). Іншим прикладом аналізу змін на індивідуальному рівні (індивідуальної динаміки) є приклад аналізу на основі лонгитюдних даних неспостережуваних характеристик. Наприклад, існують два пояснення того факту, чому індивідуум, який мав досвід вирішення певної проблеми в минулому, краще її вирішує в майбутньому. Перше пояснення базується на тому, що в процесі вирішення проблеми індивідуум набуває певного досвіду завдяки покращенню знань, спеціальних навичок тощо, що і допомагає йому в майбутньому. Інше пояснення базується на тому, що індивідууми можуть мати певні неспостережувані характеристики, які саме і збільшують імовірність набуття досвіду у вирішенні проблеми, але самі не змінюються під впливом досвіду. Лонгитюдні дані допомагають виявити причину вирішення проблеми в кожному випадку. Отже, моделі лонгитюдних даних можуть бути використані не тільки для пояснення різної поведінки різних індивідуумів. Вони можуть пояснювати, чому окремі індивідууми по-різному поведуть себе в різні періоди часу.

Другою важливою перевагою застосування лонгітюдних даних є *значне збільшення вибірки моделей*. Ця важлива особливість зменшує загрозу мультиколінеарності та значно збільшує ступені свободи.

Поєднання варіаційного виміру з часовим в лонгітюдних даних приводить також до *підвищення ефективності оцінок* порівняно з оцінюванням на основі окремих моделей часових або варіаційних рядів, причому ця властивість зберігається навіть при однаковому розмірі вибірки. Щоб проілюструвати цю властивість, розглянемо спеціальний випадок моделі з випадковими ефектами (5.6), де тільки включені фіктивні (dummy) часові змінні:

$$Y_{it} = \alpha_t + \mu_i + u_{it}, \quad (5.7)$$

де α_t — невідоме середнє генеральної сукупності (population mean) в період t .

Припустимо, що нас цікавить не окреме значення середнього α в певний період часу, а зміна значення між певними періодами, наприклад $(\alpha_t - \alpha_k)$; $t \neq k$.

Нагадаємо, що дисперсія ефективної оцінки відхилення $(\hat{\alpha}_t - \hat{\alpha}_k)$ для $(\alpha_t - \alpha_k)$; $t \neq k$, дорівнює:

$$\text{Var}(\hat{\alpha}_t - \hat{\alpha}_k) = \text{Var}(\hat{\alpha}_t) + \text{Var}(\hat{\alpha}_k) - 2\text{cov}(\hat{\alpha}_t, \hat{\alpha}_k), \quad (5.8)$$

$$\text{де } \hat{\alpha}_t = \frac{\sum_{i=1}^N Y_{it}}{N} \quad (t = 1, 2, \dots, T).$$

Як правило, якщо використовуються лонгітюдні дані, коваріація між $\hat{\alpha}_t$ та $\hat{\alpha}_k$ є позитивною та в припущеннях моделі (5.6) дорівнює відповідно $\text{cov}(\hat{\alpha}_t, \hat{\alpha}_k) = \sigma_\mu^2/N$.

Але якщо використовуються тільки крос-секційні (варіаційні) дані (наприклад дві незалежні вибірки для різних періодів, що складаються з різних об'єктів), то коваріація між $\hat{\alpha}_t$ та $\hat{\alpha}_k$ дорівнюватиме нулю: $\text{cov}(\hat{\alpha}_t, \hat{\alpha}_k) = 0$.

Таким чином, якщо нас цікавить зміна значень більше, ніж рівні, в такому разі оцінки, отримані на основі лонгітюдних даних, будуть ефективнішими за оцінки, отримані на основі лише крос-секційних даних. Цей факт є зрозумілим і на інтуїтивному рівні. Дійсно, лонгітюдні дані несуть кращу інформацію порівняно з крос-секційними даними завдяки тому, що спостереження організовані за однаковими об'єктами певної сукупності протягом певних інтервалів часу. Крім того, спостереження за однаковими об'єктами на протигагу спостереженням за різними об'єктами можуть привести до зменшення варіації, а відповідно і до зменшення ефективності. Дослідження вчених [5, 10] показали, що якщо екзогенні змінні включені в модель і ми

зацікавлені в оцінці параметрів, що вимірюють ефект від їх зміни, то лонгитюдні дані дають більш ефективні оцінки, ніж крос-секційні дані з однаковою кількістю спостережень.

Іншою значною перевагою використання моделей на основі лонгитюдних даних є *вирішення проблеми пропущених (omitted) змінних* в моделі. При застосуванні лонгитюдних даних ми можемо отримати робасті оцінки параметрів моделі навіть в умовах пропущених змінних. Проблема пропущених змінних є особливо серйозною, якщо виключена з моделі змінна корелює з наявними в моделі факторами. Класичним прикладом такої ситуації є оцінювання виробничої функції. В більшості випадків, особливо при оцінюванні виробничої функції малої фірми, бажано включати як вхідні параметри рівень менеджменту фірми, який в загальному випадку належить до неспостережуваних характеристик. Припустимо, що ми оцінюємо функцію Кобба—Дугласа такого вигляду:

$$y_{it} = \alpha + x'_{it}\beta + \pi_i\beta_{k+1} + \varepsilon_{it},$$

де y_{it} — логарифм випуску; x'_{it} — вектор логарифмів вхідних характеристик розміру $(k \times 1)$ для i -ї фірми в t -й період часу; π_i — рівень менеджменту для фірми i (який вважається незмінним протягом t часу). Неспостережувана змінна π_i , як правило, корелює негативно з вхідними змінними вектора x_{it} , оскільки зазвичай чим кращим є менеджмент, тим ефективніше використовуються фірмами вхідні ресурси. Таким чином, якщо $\beta_{k+1} = 0$, виключення π_i з моделі приводить до зміщених оцінок параметрів моделі при інших змінних. Якщо модель базується на лонгитюдних даних, то проблему зміщеності можна зняти, ввівши специфічний ефект для кожної фірми $\mu_i = \alpha + \pi_i\beta_{k+1}$, який можна розглядати як фіксований параметр. Зауважимо, що без додаткової інформації не можна оцінити параметр β_{k+1} , тому найбільш поширеним обмеженням на параметр є умова постійної віддачі на масштаб (constant returns to scale): $\beta_{k+1} = 1 - (\beta_1 + \beta_2 + \dots + \beta_k)$.

Цей простий приклад ілюструє можливість усунення зміщеності оцінок в результаті пропущених змінних при використанні лонгитюдних даних. Тобто при використанні лонгитюдних даних моделі з неповною специфікацією (пропущеними змінними) мають більш робасті оцінки.

Крім того, застосування лонгитюдних даних вирішує *проблему ендогенних змінних у правій частині за рахунок забезпечення внутрішніх інструментів*. У багатьох випадках на базі лонгитюдних даних можна отримати «внутрішні інструменти» для факторів, які є ендогенними або виміряні з помилкою (measurement error). Такі інструменти в умовах лонгитюдних даних можна отримати за рахунок трансформації наявних даних. При цьому транс-

формовані дані не корелюватимуть з випадковими величинами моделі, але корелюватимуть з факторами, для яких необхідні інструменти, тобто можна обійтись без пошуку зовнішніх інструментів, що часто є великою проблемою. Наприклад, якщо μ_i корелює з X_{it} , то можна показати, що $(X_{it} - \bar{X}_i)$, де \bar{X}_i є середнім значенням по часу для кожного i , не корелює з μ_i та може бути використане як підходящий інструмент для X_{it} . Більш того, при оцінюванні моделей лонгітюдних даних з фіксованим ефектом параметри μ_i елімінуються, що знімає проблему ендогенності при оцінюванні моделей з фіксованим ефектом взагалі. Чому відбувається саме це, розглянемо детальніше в наступному підрозділі.

Ми розглянули основні переваги застосування моделей лонгітюдних даних. Звичайно, їх застосування несе і певні проблеми. До основних недоліків моделей лонгітюдних даних можна віднести той, що оскільки спостереження належать до об'єктів однієї сукупності, то припущення про незалежність різних спостережень, як правило, не справджується, що значно ускладнює аналіз, особливо нелінійних та динамічних моделей. Крім того, відсутність деяких спостережень є типовою ситуацією для лонгітюдних даних. Збір інформації про велику кількість економічних агентів за великий проміжок часу (5—10 років) коштує досить дорого. В результаті часовий вимір найвідоміших опитувань, як правило, не перевищує 3—5 періодів часу.

Звичайно, основні недоліки можна подолати, використовуючи при оцінюванні параметрів лонгітюдних моделей спеціальні методи, які ми розглянемо пізніше.

5.2. ПРИКЛАДИ ЗАСТОСУВАННЯ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ В МОДЕЛЮВАННІ БЮДЖЕТНИХ ПРОЦЕСІВ НА РЕГІОНАЛЬНОМУ РІВНІ

Як зазначалось вище, моделі лонгітюдних даних знаходять широке застосування при аналізі та моделюванні різноманітних процесів, особливо на регіональному рівні. Розглянемо декілька прикладів реального застосування моделей лонгітюдних даних українськими дослідниками.

Модель з фіксованими ефектами при моделюванні надходжень від прибуткового податку з громадян

Моделювання прибуткового податку з громадян на основі лонгітюдних даних детально наведено в роботі [38]. Ми лише відмітимо деякі особливості моделювання цього податку та прокоментуємо саму модель.

В економічно розвинутих країнах прибутковий податок з громадян відіграє істотну роль у формуванні доходів бюджету. В Україні питома вага цього

податку менша, але все ж таки він відіграє значну роль у податковій системі, тому моделювання надходжень цього податку є досить важливим.

До переваг цього податку відносять його соціальну справедливість, оскільки він ставить рівень оподаткування в пряму залежність від обсягу доходів платників. Крім того, за допомогою прогресивних ставок і пільг в оподаткуванні можна зменшувати податковий тягар для малозабезпечених верств населення і посилювати для осіб з високими доходами. А недоліком є значні можливості для платників занижувати оподатковуваний дохід і ухилятися від сплати податку. Особливо це стосується оподаткування доходів від підприємницької діяльності.

Платниками прибуткового податку є громадяни України, іноземні громадяни та особи без громадянства як ті, що мають, так і ті, що не мають постійного місця проживання в Україні.

Об'єктом оподаткування у громадян, які мають постійне місце проживання в Україні, є сукупний оподатковуваний дохід за календарний рік (що складається з місячних сукупних оподатковуваних доходів), одержаний з різних джерел як на території України, так і за її межами. Об'єктом оподаткування у громадян, які не мають постійного місця проживання в Україні, є дохід, одержаний з джерел в Україні. Ставка податку варіюється в залежності від розміру доходу і встановлюється Законом України «Про прибутковий податок з громадян». Але на дохід, отриманий не за місцем основної роботи, встановлюється єдина ставка податку — 20 %.

Отже, прибутковий податок з доходів громадян має доволі складну структуру. Частина надходжень цього податку зараховується до місцевих бюджетів. Законом України «Про бюджетну систему України» не встановлено чіткої частки прибуткового податку з громадян, що зараховується до місцевих бюджетів. Тому ставка встановлювалась законами «Про бюджет України» на відповідний рік. В 1995—1996 рр. до місцевих бюджетів зараховувалось 50 % надходжень цього податку. Протягом цього періоду до місцевих бюджетів Севастополя та АР Крим зараховувалась повна сума надходжень цього податку. Протягом наступних років усі надходження прибуткового податку зараховувались до місцевих бюджетів.

В роботі [38] запропонована модель стягнення прибуткового податку з громадян та його зарахування до місцевих бюджетів.

Для більшості громадян основне місце роботи є єдиним джерелом отримання доходів, тож, знаючи середню зарплату в області та кількість зайнятих, можна легко отримати загальну податкову базу. Звичайно, надходження залежатимуть від розподілу доходів. Наприклад, області з порівняно вищою часткою великих доходів отримуватимуть більші надходження, ніж області з більш рівномірним розподілом. Іншим важливим припущенням є те, що кількість зайнятих щільно пов'язана з чисельністю населення в області, адже

загальна чисельність пов'язана з чисельністю населення працездатного віку, а та, в свою чергу, з чисельністю зайнятих. Зрозуміло, що відмінності у демографічній структурі областей будуть послаблювати цей зв'язок. Щоб зменшити вплив таких відмінностей, достатньо ввести фіксовані перетини, які будуть додатково відображати структуру доходів населення, тобто пропонується використовувати модель лонгітюдних даних з фіксованими ефектами. Таким чином, як наближення до податкової бази можна використовувати добуток чисельності населення та середньої зарплати в області. Цю величину необхідно додатково скорегувати на нормативи відрахувань до Державного бюджету. Отже, в 1995—1996 рр. потрібно ділити податкову базу навпіл, оскільки протягом цього періоду до місцевих бюджетів зараховувалась половина надходжень прибуткового податку з громадян.

Перерахування податку до бюджету здійснюється *одночасно* з виплатою зарплати. Якщо виникає заборгованість із зарплати, то вона виникає і з прибуткового податку у певній пропорції. Підприємства, погашаючи заборгованість із зарплати, будуть погашати й заборгованість з цього податку. Таким чином, динаміка заборгованості із зарплати повинна впливати на динаміку надходжень прибуткового податку.

Крім цього, середня зарплата, як і чисельність зайнятих, має сезонні коливання. Наприклад, наприкінці звітнього року йдуть численні перерахування надмірно стягнутих податків. Очевидно, що в цей період чисті надходження податку будуть меншими. Доцільно включити в регресію індексовані надходження за аналогічний період попереднього року. Це дасть змогу не тільки враховувати сезонні коливання, а й ефективно використовувати обмежену кількість спостережень.

Враховуючи вищенаведений аналіз особливостей формування прибуткового податку з громадян в Україні, в [38] було оцінено таку модель лонгітюдних даних, яка дозволяє прогнозувати та аналізувати надходження цього податку в розрізі регіонів (в дужках наведені значення *t*-статистики):

$$PIT_{ti} = \alpha_i + 0.101AWAGE_{ti} \cdot POP_{ti} \cdot PITSHARE_{ti} + \quad (31.4)$$

$$+ 0.0775PIT_{t-4j} \frac{GDPDEF_i}{GDPDEF_{t-4}} - 21.71D(WAGER_{ti}) \quad (5.95) \quad (1.97)$$

$$R^2 = 0.97; DW = 2.14; F = 3597,$$

де PIT_{ti} — надходження прибуткового податку з громадян у період t в області i (тис. грн); α_i — фіксований коефіцієнт для регіону i ; $AWAGE_{ti}$ — середня заробітна плата у період t в області i (грн); POP_{ti} — населення у період t

в області i (тис. чол.); $PITSHARE_{ti}$ — частка прибуткового податку, що зараховується до місцевих бюджетів у період t в області i ; $GDPDEF_t$ — значення дефлятора ВВП в період t ; $D(WAGEAR_{ti})$ — зміни в обсягах заборгованості із зарплати у період t в області i (млн грн).

Переважає більшість перетинів (отримані значення яких ми не наводимо з міркувань економії місця) від'ємна, що вказує на певні втрати бюджетів від запровадження пільг щодо цього податку. Лише Запорізька, Донецька та Дніпропетровська області мають додатні перетини. Коефіцієнт при базі оподаткування додатний, менший за одиницю та значимий при найвищих рівнях довіри. Він вказує, що чиста ефективна ставка оподаткування доходів громадян становить приблизно 3.36 %¹. Це, звичайно, менше від реальної ефективної ставки (приблизно 11—12 %), бо використовується не чисельність зайнятих, а все населення. Оскільки зміни в заборгованості виміряні у мільйонах гривень, від'ємний значимий коефіцієнт при змінах заборгованості із зарплати вказує на те, що зростання такої заборгованості детермінує зменшення надходжень прибуткового податку на 21 % від зміни обсягу заборгованості.

Розглянемо можливість застосування моделей лонгитюдних даних для моделювання суто місцевого податку, а саме податку з власників транспортних засобів. Ця модель детально наведена в [38].

Модель лонгитюдних даних надходжень податку з власників транспортних засобів

Цей податок є відносно стабільним. Він властивий багатьом країнам та існував ще навіть за часів СРСР. Протягом багатьох десятиріч громадяни — власники транспортних засобів сплачували місцевий збір. З прийняттям у 1981 р. Положення про місцеві податки збір було замінено місцевим податком, і до оподаткування було залучено як міське, так і сільське населення. Починаючи з 1988 р. податок було поширено на транспортні засоби підприємств. Законом України «Про податок з власників транспортних засобів та інших самохідних машин і механізмів» (11.12.1991) встановлюється податок з власників деяких наземних і водних транспортних засобів, самохідних машин і механізмів як джерело фінансування будівництва, реконструкції, ремонту та утримання автомобільних шляхів загального користування та проведення природоохоронних заходів на водоймищах.

Об'єктом оподаткування є переважна більшість транспортних засобів. Платниками цього податку є як юридичні, так і фізичні особи. Ставки по-

¹ База оподаткування розрахована для одного місяця, тоді як надходження — для трьох. Отже, помноживши базу на три, отримаємо квартальну базу оподаткування. Тоді чиста ефективна ставка оподаткування становитиме $0.101/3 = 0.0336 = 3.36\%$.

датку залежать від виду транспортного засобу та об'єму циліндрів двигуна, потужності двигуна, довжини (для водного транспортного засобу). 90 % надходжень цього податку спрямовується на покриття витрат по експлуатації автошляхів та проведення природоохоронних заходів на водоймах. Решта, 10 %, використовується на поліпшення матеріально-технічної бази органів, що здійснюють реєстрацію, перереєстрацію та технічний огляд цих транспортних засобів (переважно ДАІ).

Сплачується податок щорічно, але не пізніше першого півріччя перед реєстрацією, перереєстрацією та технічним оглядом. Нарахування податку здійснюється на основі сумарної потужності двигуна транспортного засобу з урахуванням його різновиду та відповідної ставки, закріпленої в ст. 2 цього закону.

Податок на транспортні засоби є прямим податком та не залежить від результатів господарської діяльності. Характерною особливістю цього податку є значна сезонність. Протягом першого півріччя, і особливо другого кварталу, сплачується більша частина річних надходжень. Очевидно, податкові надходження можна моделювати за допомогою моделей лонгітюдних даних з авторегресійними елементами. Оскільки податок сплачують як фізичні, так і юридичні особи, то надходження будуть залежати від кількості населення. Тут ми припускаємо, що кількість транспортних засобів корелює з кількістю населення. Іншими словами, населення виступає наближенням, або заміником, до невідомої кількості транспортних засобів. Необхідно також враховувати значні зміни в законодавстві щодо ставок цього податку. Сама по собі модель не може передбачати таких екзогенних змін, тому потрібно в регресію ввести димпу-змінну, яка буде дорівнювати одиниці у відповідний період зміни та нулю в інші періоди. Оскільки в даних наявна сильна сезонність, то доцільно використовувати не тільки індексовані надходження за попередній період, а й за аналогічний період протягом минулого року. Регіони, зрозуміло, відрізняються не тільки кількістю, а й якістю транспортних засобів (наприклад переважанням певного типу). Тому необхідно враховувати регіональні відмінності через відповідні коефіцієнти для кожного регіону, тобто використовувати моделі з фіксованим ефектом. Оцінена в [38] модель надходження податку з власників транспортних засобів має вигляд (в дужках наведені значення t -статистики):

$$CT_{it} = \alpha_i + 1.057CT_{t-4,i} + \frac{GDPDEF_t}{GDPDEF_{t-4}} + 13.26POP_{it} -$$

(0.026) (2.44)

$$- 527.8LOG(TREND_t) + 333.3D_CT_i$$

(78.4) (52.31)

$$R^2 = 0.91; DW = 1.91; F = 1031,$$

де CT_{it} — податок з власників транспортних засобів у період t в області i ; α_i — фіксований коефіцієнт для регіону i ; POP_{it} — населення у період t в області i ; $TREND$ — тренд; D_CT_t — фіктивна (dummy) змінна, що відображає зміни у законодавстві.

З коефіцієнтів моделі видно, що надходження цього податку в середньому зростають на 5 % на рік. Впровадження змін у законодавстві призводить до зростання надходжень в середньому на 333 тис. грн на квартал у всіх регіонах. Регіони з більшою кількістю населення мають в середньому більші надходження цього податку. В середньому на 1 тис. населення припадає 13.26 тис. грн податку. Можна й іншим чином включати населення в регресію: використовувати індивідуальні коефіцієнти для кількості населення в кожному регіоні та сталий перетин для всіх регіонів. Таке перетворення дає практично аналогічні результати. З іншого боку, від'ємний знак коефіцієнта при тренді вказує на поступове зменшення надходжень цього податку. Завдяки формі включення тренду це падіння буде все більш малим у часі. Існування самого тренду можна пояснити наданням пільг окремим власникам автотранспортних засобів та поступовим скороченням парку транспортних засобів в Україні.

Модель заборгованості із заробітної плати на регіональному рівні

Розглянемо особливості застосування моделей лонгитюдних даних з фіксованими ефектами для аналізу факторів накопичення заборгованості із заробітної плати у розрізі регіонів в Україні. Детально процес моделювання описаний в роботі [39].

Заборгованість із заробітної плати є явищем специфічним і надзвичайно негативним для економіки України, оскільки загострює економічну кризу, позбавляє працівників стимулів до праці, підвищує соціальну напруженість в суспільстві. Сьогодні це явище практично ліквідовано в Україні, але причини та можливі наслідки цього феномену ще й досі продовжують широко обговорюватись політиками, фахівцями і науковцями.

Загальний обсяг заборгованості із заробітної плати складається з двох частин: заборгованість у бюджетній сфері і заборгованість у небюджетній сфері. Заборгованість у бюджетній сфері включає переважно невиплати заробітної плати у галузі освіти і охорони здоров'я, що фінансується з місцевих бюджетів. Відповідні видатки записані у статті місцевих бюджетів «Фінансування соціально-культурних закладів і заходів». Крім того, значна частина заборгованості із заробітної плати формується на промислових підприємствах і пов'язана зі складним економічним становищем цих підприємств.

Динаміка заборгованості із заробітної плати є різною для різних регіонів України. У зв'язку з цим особливої актуальності набуває моделювання

явища заборгованості на регіональному рівні за допомогою економетричних моделей, які дозволяють дослідити причинно-наслідкові зв'язки накопичення заборгованості із заробітної плати в регіонах України.

Дослідження чинників заборгованості із заробітної плати було здійснено на основі статистичних даних для 26 регіонів України (виключено місто Севастополь). Оцінена модель з фіксованими ефектами мала такий вигляд:

$$\ln \hat{Y}_{it} = \hat{\alpha}_i - 3.751 \ln X_{1it} + 1.423 \ln X_{2it} - 0.997 \ln X_{3it} + 3.665 \ln X_{4it},$$

(-3.616)
(6.092)
(-2.716)
(4.381)

де Y_{it} — обсяги заборгованості на душу населення в i -му регіоні в t -й період часу; X_{1it} — фактична частка фінансування соціально-культурних закладів і заходів у загальних видатках місцевих бюджетів i -го регіону в t -й період часу; X_{2it} — кредиторська заборгованість за виключенням заборгованості із заробітної плати на душу населення в i -му регіоні в t -й період часу; X_{3it} — дебіторська заборгованість на душу населення; X_{4it} — обсяги торгівлі на душу населення i -му регіоні в t -й період часу.

У дужках подано значення t -статистики. Загальна кількість спостережень — 76, коефіцієнт детермінації — 0.756. Оцінені константи для кожної окремої області не аналізувались у ході дослідження і тому не наводяться.

Побудована модель дозволяє оцінити еластичність заборгованості із заробітної плати по частці фінансування соціально-культурних закладів і заходів місцевих бюджетів.

У цілому результати дослідження підтвердили гіпотезу про наявність негативного зв'язку між часткою фінансування соціально-культурних закладів і заходів місцевими бюджетами і заборгованістю із заробітної плати. Більше того, еластичності значно відрізняються для різних груп регіонів, виділених за соціальними ознаками. Наприклад, для кластера регіонів із значним промисловим потенціалом (Дніпропетровська, Донецька, Запорізька, Миколаївська, Одеська, Харківська, Херсонська і Чернівецька області) коефіцієнт еластичності становив -10.37 на рівні значимості 5 %.

Таким чином, загальний висновок дослідження полягає у тому, що заборгованість із заробітної плати на регіональному рівні залежить більшою мірою від ефективності державного управління, яка відображена у структурі видатків місцевих бюджетів, ніж від виробничих характеристик регіонів.

5.3. ТЕОРЕТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ ЗА ДОПОМОГОЮ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ З ФІКСОВАНИМИ ЕФЕКТАМИ

Як ми відмічали вище, існують два базових типи моделей лонгитюдних даних: модель з фіксованими ефектами і модель з випадковими ефектами. Ми розглянемо проблеми побудови та оцінювання цих базових моделей в найбільш спрощеному варіанті, тобто у варіанті статичних моделей з одновимірною складовою помилки. Нагадаємо, що основна різниця між моделями лонгитюдних даних з фіксованими та випадковими ефектами полягає в припущенні щодо перетинів цих моделей. В моделях з випадковими ефектами перетини розглядаються як випадкові величини, а в моделях з фіксованими ефектами — як фіксовані значення, різні для різних об'єктів.

Модель лонгитюдних даних з фіксованими ефектами має такий загальний вигляд:

$$Y_{it} = \mu_i + X'_{it}\beta + u_{it}, u_{it} \sim IID(0, \sigma_u^2), \quad (5.9)$$

де $X'_{it} = \{X_{1it}, X_{2it}, \dots, X_{kit}\}$ — вектор порядку $(k \times 1)$ пояснюючих змінних (факторів) без константи; $\beta = \{\beta_1, \beta_2, \dots, \beta_k\}$ — вектор невідомих параметрів у припущенні, що ефект від зміни X є однаковим для всіх i та всіх періодів часу t ; $i = 1, 2, 3, \dots, N$; $t = 1, 2, 3, \dots, T$. Ще раз підкреслимо, що припущення, за яким середній рівень для i -го об'єкта може відрізнятись від середнього рівня j -го, але є постійним для різних періодів часу, вводиться в модель за допомогою різних значень перетину μ_i ; $i = 1, 2, 3, \dots, N$. Крім того, припускається, що всі X_{it} незалежні від всіх u_{it} , а збурення u_{it} є, в свою чергу, незалежними однаково розподіленими (identical independently distributed — *IID*) випадковими величинами з математичним сподіванням нуль та постійною дисперсією σ_u^2 .

Як легко побачити, (5.9) є фактично лінійною регресійною моделлю, в якій перетини варіюються в залежності від об'єкта i . Тому її можна переписати у вигляді звичайної лінійної регресійної моделі, включивши фіктивну (*dummy*) змінну для кожного об'єкта i :

$$Y_{it} = \sum_{j=1}^N \mu_j d_{ij} + X'_{it}\beta + u_{it}, \quad (5.10)$$

де $d_{ij} = 1$, якщо $i = j$ і нулю у всіх інших випадках. Таким чином, ми маємо додатково N фіктивних (*dummy*) змінних в моделі (5.10).

Модель (5.10) можна записати в більш загальному вигляді. Позначимо через Y_i та X_i вектори розміру $(T \times 1)$ T спостережень для i -ї групи; u_i — вектор випадкових величин розміру $(T \times 1)$; I — одиничний вектор розміру $(T \times 1)$.

Тоді (5.10) ми можемо переписати як

$$Y_{it} = i\mu_i + X_i\beta + u_i, \quad (5.11)$$

тобто

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_N \end{bmatrix} = \begin{bmatrix} i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & i & \cdot & \cdot & \cdot & 0 \\ & & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot & \\ 0 & 0 & \cdot & \cdot & \cdot & i \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \mu_N \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_N \end{bmatrix} \beta + \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_N \end{bmatrix},$$

або у більш загальному виразі:

$$Y = [d_1 \ d_2 \ \cdot \ \cdot \ d_N \ X] \begin{bmatrix} \mu \\ \beta \end{bmatrix} + u, \quad (5.12)$$

де d_i є dummy-змінна, що означає i -ту групу (об'єкт). Позначимо через $D = [d_1 \ d_2 \ \cdot \ \cdot \ d_N]$ матрицю розміром $(NT \times N)$, тоді (5.12) можна переписати у вигляді:

$$Y = D\mu + X\beta + u. \quad (5.13)$$

Модель (5.13) часто називають моделлю фіктивних змінних методу найменших квадратів (least squares dummy variable (LSDV) model), хоча слова «найменших квадратів» стосуються скоріше методу розв'язку, ніж моделі як такої. Зрозуміло, що невідомі параметри такої моделі $\mu_1, \mu_2, \dots, \mu_N$ і β можуть бути оцінені звичайним методом найменших квадратів, який у даному контексті називається методом найменших квадратів з фіктивними змінними (the least squares dummy variable (LSDV) estimator). З представлення моделі з фіксованими ефектами у вигляді (5.13), тобто у вигляді з фіктивними (dummy) змінними, наочно зрозуміло, чому ми не можемо вводити і оцінювати загальний перетин в моделях лонгітюдних даних з фіксованими ефектами. Введення загального перетину у модель передбачає введення одиничного стовпця у матрицю спостережень, який в даному випадку буде лінійною комбінацією всіх фіктивних змінних, що, в свою чергу, приводить до досконалий мультиколінеарності та неможливості оцінювання параметрів моделі.

Звичайно, на практиці незручно оцінювати регресійну модель з великою кількістю додаткових фіктивних змінних, особливо коли N достатньо велике, через те, що ми втрачаємо ступені свободи. На щастя, ми можемо оцінити невідомі параметри моделі (5.10) в дуже простий спосіб. Можна показати, що ті ж самі оцінки параметрів для β можна отримати з регресійної моделі,

де фактори представляються у формі відхилень від середнього окремої групи (об'єкта). Фактично це означає, що ми виключаємо індивідуальні ефекти μ_i з моделі (5.10) за допомогою трансформації даних, а потім оцінюємо їх після оцінки параметрів β .

Розглянемо дану трансформацію детальніше. Зауважимо, що з моделі (5.10) легко отримати модель вигляду:

$$\bar{Y}_i = \mu_i + \bar{X}_i' \beta + \bar{u}_i, \tag{5.14}$$

$$\text{де } \bar{Y}_i = \frac{\sum_{t=1}^T Y_{it}}{T}; \bar{X}_i' = \{\bar{X}_{1i}, \bar{X}_{2i}, \dots, \bar{X}_{ki}\}; \bar{X}_{ki} = \frac{\sum_{t=1}^T X_{kit}}{T}; \bar{u}_i = \frac{\sum_{t=1}^T u_{it}}{T}.$$

Віднявши від обох частин моделі (5.11) відповідні частини моделі (5.14), отримаємо модель зі змінними у відхиленнях від середнього:

$$Y_{it} - \bar{Y}_i = (X_{it} - \bar{X}_i') \beta + (u_{it} - \bar{u}_i), \tag{5.15}$$

Модель (5.16) є трансформованою моделлю, яка не включає індивідуальні ефекти μ_i . Така трансформація називається трансформацією в крос-секційному розрізі (**the within transformation**). Відповідно оцінки параметрів β , отримані з цієї трансформованої моделі методом найменших квадратів, часто називаються оцінками, отриманими методом найменших квадратів для внутрішньогрупової варіації (**the within estimator**), або оцінками фіксованих ефектів, і вони абсолютно ідентичні оцінкам параметрів β , отриманим з моделі (5.10) методом найменших квадратів з фіктивними змінними (LSDV). Оцінки параметрів β , отримані з трансформованої моделі, зазвичай позначаються $\hat{\beta}_{FE}$ та розраховуються за формулою:

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)' \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i). \tag{5.16}$$

За умови, що X_{it} є незалежними від u_{it} , оцінки фіксованих ефектів $\hat{\beta}_{FE}$ є незміщеними оцінками для β . Крім того, якщо збурення u_{it} мають нормальний закон розподілу, то і отримані оцінки $\hat{\beta}_{FE}$ будуть також нормально розподіленими випадковими величинами.

У матричному вигляді оцінку параметрів $\hat{\beta}_{FE}$ можна записати, ввівши матрицю $M_d = I - D(D'D)^{-1}D'$, а саме: матрицю, за допомогою якої можна легко описати трансформацію даних:

$$\hat{\beta}_{FE} = [X'M_d X]^{-1} [X'M_d Y], \quad (5.17)$$

де відповідно $M_d = I - D(D'D)^{-1}D'$. Зазначимо, що за допомогою матриці M_d можна легко описати необхідну нам трансформацію даних, тобто відхилення від середнього. Розпишемо матрицю M_d детальніше. Враховуючи, що стовпці матриці D є ортогональними, матриця M_d має такий специфічний вигляд:

$$M_d = \begin{bmatrix} M^0 & 0 & 0 & \dots & 0 \\ 0 & M^0 & 0 & \dots & 0 \\ & & \vdots & & \\ & & \vdots & & \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & M^0 \end{bmatrix}, \quad (5.18)$$

де кожна матриця на діагоналі є:

$$M^0 = I_T - \frac{1}{T} i i'. \quad (5.19)$$

Матриця M^0 має важливу властивість: результатом добутку будь-якого вектора z_i на матрицю M^0 є вектор відхилень від середнього, тобто $M^0 z_i = z_i - \bar{z}$. Зазначимо, що середнє береться лише по T спостереженнях для групи i . Отже, регресія $M_d Y$ на $M_d X$ еквівалентна регресії (5.15) для трансформованих даних $[Y_{it} - \bar{Y}_i]$ на $[X_{it} - \bar{X}_i]$:

$$Y_{it} - \bar{Y}_i = (X_{it} + \bar{X}_i) \beta + (u_{it} - \bar{u}_i), \quad (5.20)$$

де $\bar{Y}_i = \frac{\sum_{t=1}^T Y_{it}}{T}$; $\bar{X}_i = \{\bar{X}_{1i}, \bar{X}_{2i}, \dots, \bar{X}_{ki}\}$ — вектор середніх X_i по T спостереженнях розміру $(k \times 1)$.

Аби отримані оцінки були консистентними, необхідно, щоб коваріація між трансформованими факторами та випадковими величинами дорівнювала нулю, що аналогічно умові:

$$E\{(X_{it} - \bar{X}_i) u_{it}\} = 0. \quad (5.21)$$

Звичайно, для того, щоб виконувалась умова (5.21), необхідно, щоб ні X_{it} , ні \bar{X}_i не корелювали з u_{it} , а отже, необхідно, щоб

$$E\{X_{it}u_{is}\} = 0 \text{ для всіх } s, t. \tag{5.22}$$

У випадку, коли умова (5.22) виконується, змінні X_{it} називаються строго екзогенними (*strictly exogenous*). Строго екзогенна змінна не може залежати від поточних, майбутніх та минулих значень випадкової величини. У деяких випадках таке припущення може бути досить обмежуючим. Наприклад, воно виключає можливість включення лагових змінних в модель, що є досить обмежуючою умовою для практичних досліджень. Крім того, будь-яка змінна X_{it} , яка залежить від Y_{it} , буде також порушувати цю умову, тому на практиці припущення про строгу екзогенність децю пом'якшується.

На основі знайдених оцінок $\hat{\beta}_{FE}$ для параметрів β та в припущенні, що всі фактори моделі незалежні від випадкових величин, незміщені оцінки для N різних перетинів μ_i моделі (5.10) знаходяться за формулою:

$$\hat{\mu}_i = \bar{Y}_i - \bar{X}_i' \hat{\beta}_{FE}, i = 1, \dots, N. \tag{5.23}$$

Використовуючи раніше введenu матрицю D , формулу (5.23) можна представити в іншому вигляді:

$$D'D\mu + D'X\hat{\beta}_{FE} = D'Y \tag{5.24}$$

або

$$\mu = [D'D]^{-1} D'(Y - \hat{\beta}_{FE}). \tag{5.25}$$

Це означає, що для кожного i μ_i є відхиленням від середніх i -ї групи, тобто (5.25) еквівалентно виразу (5.23): $\hat{\mu}_i = \bar{Y}_i - \bar{X}_i' \hat{\beta}_{FE}$.

Отже, головною особливістю моделей лонгїтюдних даних з фіксованими ефектами є можливість отримання та, відповідно, аналізу різниці «всередині» об'єктів. Формалізованою мовою це означає, що моделі з фіксованими ефектами можуть надати інформацію для пояснення, як Y_{it} відрізняється від \bar{Y}_i , але на їх основі ми не можемо пояснити, чому \bar{Y}_i може відрізнитись від \bar{Y}_s .

При виконанні припущення (5.21) оцінки $\hat{\mu}_i$ фіксованих ефектів μ_i є консистентними, якщо T прямує до нескінченності. Причина неконсистентності μ_i для фіксованих T інтуїтивно зрозуміла, коли T фіксоване, групові середні \bar{Y}_i та \bar{X}_i не сходяться до якоїсь певної величини, якщо кількість груп (об'єктів) зростає.

Важливим є також визначення коваріаційної матриці параметрів для моделі лонгїтюдних даних з фіксованими ефектами.

Коваріаційна матриця для оцінок параметрів $\hat{\beta}_{FE}$ в припущенні, що випадкові величини u_{it} незалежні, однаково розподілені для всіх значень I та

t з постійною дисперсією σ_u^2 , знаходиться за такою формулою:

$$V\{\hat{\beta}_{FE}\} = \sigma_u^2 \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)' \right)^{-1}. \quad (5.26)$$

Оскільки дисперсія трансформованої помилки $(u_{it} - \bar{u}_i)$ дорівнює $[(T-1)/T]\sigma_u^2$, а не просто σ_u^2 , то для отримання консистентної оцінки в припущенні, що N прямує до нескінченності, необхідно внутрішню суму квадратів поділити на $N(T-1)$. Отже,

$$\begin{aligned} \hat{\sigma}_u^2 &= \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \hat{\mu}_i - \bar{X}'_i \hat{\beta}_{FE})^2 = \\ &= \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=1}^T \left([(Y_{it} - \bar{Y}_i) - (X_{it} - \bar{X}_i)]' \hat{\beta}_{FE} \right)^2. \end{aligned} \quad (5.27)$$

В матричному вигляді з урахуванням введених вище позначень відповідну оцінку матриці коваріації для параметрів β можна представити у вигляді:

$$\text{Var}[\hat{\beta}_{FE}] = \hat{\sigma}_u^2 [X' M_d X]^{-1}. \quad (5.28)$$

Зауважимо, що в деяких підручниках замість консистентної оцінки дисперсії випадкової величини використовується незміщена оцінка дисперсії:

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T (Y_{it} - \mu_i - X'_{it} \hat{\beta}_{FE})^2}{NT - N - K}. \quad (5.29)$$

Звичайно, якщо N прямує до нескінченності, то знаменники у виразах (5.27) та (5.29) фактично збігаються, а чисельники є однаковими:

$$\begin{aligned} \hat{u}_{it} &= Y_{it} - \mu_i - X'_{it} \hat{\beta}_{FE} = Y_{it} - (\bar{Y}_i - \bar{X}'_{it} \hat{\beta}_{FE}) - \bar{X}'_{it} \hat{\beta}_{FE} = \\ &= (Y_{it} - \bar{Y}_i) - (X_{it} - \bar{X}_i)' \hat{\beta}_{FE}. \end{aligned} \quad (5.30)$$

5.3.1. F-тест одночасної рівності значень фіксованих ефектів (μ_i)

При використанні моделей лонгітюдних даних з фіксованими ефектами однією з важливих перевірок є перевірка факту, чи статистично значимо відрізняються між собою значення перетинів. При цьому ми можемо засто-

совувати відомий F -тест з нульовою гіпотезою: $H_0: \mu_1 = \mu_2 = \dots = \mu_N = 0$ проти альтернативної, що хоча б один з параметрів не дорівнює нулю. Якщо ми не можемо відкинути нульову гіпотезу, то це означає, що використання моделі лонгитюдних даних з одним перетином (pooled model) має перевагу над використанням моделі лонгитюдних даних з фіксованими ефектами. В даному контексті F -тест є фактично Чоу-тестом (Chow test) порівняння моделі з обмеженнями на параметри (restricted model) та моделі без обмеження на параметри (unrestricted model). Для розрахунку параметрів моделі з обмеженнями застосовують звичайний МНК, а моделі без обмежень — МНК з фіктивними змінними (LSDV). Якщо N є достатньо великим, то можна застосувати крос-секційну трансформацію (within transformation) даних для оцінки параметрів без обмежень.

Для тестування нульової гіпотези розраховується таке значення F -статистики:

$$F(N - 1, NT - N - K) = \frac{(R_f^2 - R_p^2)/(N - 1)}{(1 - R_f^2)/(NT - N - K)}, \quad (5.31)$$

де N — кількість об'єктів (груп); $I = 1, 2, \dots, N$; T — кількість періодів часу; $t = 1, 2, \dots, T$; K — кількість факторів в моделі; R_f^2 — коефіцієнт детермінації моделі з фіксованими ефектами, яка в даному випадку вважається моделлю без обмежень на параметри (unrestricted model); R_p^2 — коефіцієнт детермінації моделі з одним загальним постійним членом (pooled model), яка в даному разі є моделлю з обмеженнями на параметри (restricted model), а саме, що всі вони рівні між собою та дорівнюють певній константі.

Для перевірки нульової гіпотези розрахункове значення F -статистики (5.31) порівнюється з критичним значенням, знайденим за F -таблицями зі ступенями свободи $(N - 1)$ та $(NT - N - K)$ при заданому рівні значимості. Якщо розрахункове значення F -статистики перевищує критичне значення, то нульова гіпотеза відкидається. Це означає, що дійсно краще оцінювати модель лонгитюдних даних з фіксованими ефектами, ніж звичайну регресійну модель зі спільним перетином на множині лонгитюдних даних.

5.4. ТЕОРЕТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ ЗА ДОПОМОГОЮ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ З ВИПАДКОВИМИ ЕФЕКТАМИ

Нагадаємо загальний вигляд моделі лонгитюдних даних:

$$Y_{it} = \alpha + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots + \beta_k X_{kit} + \varepsilon_{it}, \quad (5.32)$$

де Y_{it} — значення залежної змінної для i -го об'єкта у t -й період часу, α та β_j — невідомі параметри моделі, X_{jit} — значення j -го фактора для i -го об'єкта у t -й період часу, ε_{it} — значення випадкової величини; $i = \overline{1, N}$; $t = \overline{1, T}$; $j = \overline{1, k}$.

Як вже зазначалось вище, специфіка моделей лонгітюдних даних надає можливість додатково розбивати збурення на декілька компонентів та відповідно перетворювати модель (5.32) в моделі лонгітюдних даних з одновимірною та двовимірною складовими помилки.

Найбільш поширеними на практиці є моделі з одновимірною складовою помилки (one-way error component model), тобто моделі, в яких ми можемо виділити один окремих компонент збурення моделі (5.32):

$$\varepsilon_{it} = \mu_i + u_{it}, \quad (5.33)$$

де μ_i позначає неспостережувані специфічні індивідуальні ефекти, а u_{it} — залишковий шум.

Якщо припускається, що μ_i можуть тлумачитись як реалізації незалежних від X_{it} випадкових величин з розподілу з середнім α та дисперсією σ_μ^2 , то модель (5.32) належить до класу стандартних моделей лонгітюдних даних з випадковими ефектами (random effects models). У випадку моделей з випадковими ефектами перетини μ_i тлумачаться як випадкові величини, що мають нульове математичне сподівання, а збурення u_{it} припускаються некорельованими для різних періодів часу. За такими припущеннями модель з випадковими ефектами має вигляд:

$$Y_{it} = \alpha + X'_{it}\beta + \mu_i + u_{it}, \quad (5.34)$$

де α є спільним перетином. Отже, в моделі з випадковими ефектами на відміну від моделей з фіксованими ефектами можна виділити та окремо оцінити спільний перетин.

В регресійному аналізі вважається, що вплив неспостережуваних факторів, які в явному вигляді не можуть бути включені до моделі, відображається за допомогою випадкової величини. Тому в моделі з випадковими ефектами припускається, що μ_i є випадковими факторами, незалежно і однаково розподіленими (IID) між індивідуумами (об'єктами). Таким чином, модель з випадковими ефектами (5.34) можна переписати у такому більш повному вигляді:

$$Y_{it} = \alpha + X'_{it}\beta + \mu_i + \varepsilon_{it}, \quad \varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2); \quad \mu_i \sim IID(0, \sigma_\mu^2), \quad (5.35)$$

де $(\mu_i + \varepsilon_{it})$ є випадковою величиною, що складається з двох компонент: індивідуальна специфічна компонента, що не змінюється з часом, та залишкова компонента, яка за припущенням не корельована в часі. Отже, кореляція

випадкових величин в часі є ознакою лише індивідуальних ефектів μ_i . Припускається також, що μ_i та ε_{it} є взаємно незалежними, а також не залежать від зміни X_{js} (для всіх j і s). Це означає, що оцінювання невідомих параметрів α та β моделі (5.34) методом найменших квадратів дає незміщені та консистентні оцінки. Але структура компонент випадкової величини показує, що складна випадкова величина $(\mu_i + \varepsilon_{it})$ являє собою особливу форму автокореляції (якщо тільки $\sigma_\mu^2 \neq 0$). В результаті стандартні помилки, що обчислюються на основі методу найменших квадратів, не є коректними. Більш ефективні оцінки можна отримати методом загальних найменших квадратів (GLS), враховуючи специфічну структуру матриці коваріацій помилок.

Щоб отримати оцінку параметрів моделі (5.35) методом загальних квадратів, спочатку зазначимо, що для кожного індивідуума (об'єкта) i можна згорнути дві випадкові компоненти в одну таким чином:

$$(\mu_i i_T + \varepsilon_i), \tag{5.36}$$

де $i_T = (1, 1, \dots, 1)'$ є вектором розміру $(T \times 1)$, а $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{iT})'$.

Коваріаційна матриця випадкової величини (5.36) дорівнює [7, с. 34]:

$$\text{Var} - \text{Cov}(\mu_i i_T + \varepsilon_i) = \Omega = \sigma_\mu^2 i_T i_T' + \sigma_\varepsilon^2 I_T, \tag{5.37}$$

де I_T є одиничною матрицею порядку $(T \times T)$.

Коваріаційну матрицю (5.37) можна використати для оцінювання невідомих параметрів моделі (5.35) методом узагальнених найменших квадратів (GLS). Для цього проведемо трансформацію даних для кожного i -го об'єкта (індивідуума), помноживши відповідні вектори змінних $Y_i = (Y_{i1}, \dots, Y_{iT})'$ на обернену коваріаційну матрицю Ω^{-1} , яку можна представити у вигляді [14, 31]:

$$\Omega^{-1} = \sigma_\varepsilon^{-2} \left[I_T - \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + T\sigma_\mu^2} i_T i_T' \right], \tag{5.38}$$

або у дещо іншому вигляді:

$$\Omega^{-1} = \sigma_\varepsilon^{-2} \left[\left(I_T - \frac{1}{T} i_T i_T' \right) + \theta \frac{1}{T} i_T i_T' \right], \tag{5.39}$$

де $\theta = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\mu^2}$.

Нагадаємо [10, 14, 31], що множник

$$\left(I_T - \frac{1}{T} i_T i_T' \right)$$

трансформує дані у вигляді відхилень від середнього, тобто помноживши вектор на даний множник, ми отримаємо вектор у вигляді відхилень первісних значень від середнього. Множник

$$\left(\frac{1}{T} i_T i_T' \right)$$

дозволяє отримати індивідуальні середні.

Відповідно оцінку параметрів для β методом узагальнених найменших квадратів (GLS) можна записати як

$$\begin{aligned} \hat{\beta}_{GLS} = & \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)' + \theta T \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})' \right)^{-1} \times \\ & \times \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i) + \theta \sum_{i=1}^N (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y}) \right), \end{aligned} \quad (5.40)$$

$$\text{де } \bar{X} = \frac{\sum_{i=1}^N \sum_{t=1}^T X_{it}}{NT}.$$

Легко побачити, що для $\theta = 0$ вираз (5.40) є фактично оцінкою коефіцієнтів β у випадку моделі з фіксованими ефектами. Оскільки $\theta \rightarrow 0$, коли $T \rightarrow \infty$, то можна зробити висновок, що оцінки коефіцієнтів β є однаковими для моделей з фіксованими та випадковими ефектами при великих T . Якщо $\theta = 1$, оцінка узагальненого методу найменших квадратів (GLS) дорівнює оцінці звичайного методу найменших квадратів (OLS), а варіаційно-коваріаційна матриця Ω є діагональною. Загальну формулу (5.40) для оцінки β методом узагальнених найменших квадратів (GLS) можна записати в такому вигляді (детально див. [14]):

$$\hat{\beta}_{GLS} = D\hat{\beta}_B + (I_k - D)\hat{\beta}_{FE}, \quad (5.41)$$

$$\text{де } \hat{\beta}_B = \left(\sum_{i=1}^N (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})' \right)^{-1} \sum_{i=1}^N (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y}) \quad (5.42)$$

називається оцінкою методу найменших квадратів без крос-секційної (внутрішньогрупової) варіації (between estimator) для β . Тобто вираз (5.42) є оцінкою методу найменших квадратів невідомих параметрів β в моделі середніх значень окремих об'єктів (індивідуумів):

$$\bar{Y}_i = \alpha + \bar{X}_i' \beta + \mu_i + \bar{\varepsilon}_i, \quad i = 1, 2, \dots, N. \quad (5.43)$$

Матриця D є матрицею вагових коефіцієнтів, пропорційною до оберненої коваріаційної матриці параметрів $\hat{\beta}_B$ [14, с. 36]. Таким чином, оцінка параметрів, отримана методом узагальнених найменших квадратів (GLS), є середньозваженою матричною оцінкою двох оцінок: методом найменших квадратів без крос-секційної варіації (between estimator) та методом найменших квадратів для міжгрупової коваріації (within estimator). При цьому вагові коефіцієнти методу узагальнених найменших квадратів у даному випадку залежать від відносних дисперсій оцінок (для більш точних з них ваговий коефіцієнт приймає більше значення).

Як бачимо, оцінки, отримані методом найменших квадратів без крос-секційної (внутрішньогрупової) варіації (between estimator), ігнорують будь-яку інформацію в межах окремих груп (об'єктів).

Таким чином, оцінка параметрів методом узагальнених найменших квадратів (GLS) є оптимальною комбінацією оцінок без внутрішньогрупової варіації та міжгрупової коваріації, причому більш ефективною, ніж кожна з них окремо [10, 14].

Оцінка параметрів методом найменших квадратів (при $\theta = 1$) є також лінійною комбінацією цих двох оцінок, але вона не є ефективною. Метод узагальнених найменших квадратів є переважно ефективнішим за метод звичайних найменших квадратів. Крім того, якщо пояснюючі змінні моделі є незалежними від усіх значень ε_{it} та всіх μ_i , оцінки, отримані узагальненим методом найменших квадратів, є також незміщеними. Вони стають незміщеними для N або T , або для обох, якщо ті прямують до нескінченності, якщо, крім умови $E((X_{it} - \bar{X}_i)\varepsilon_{it}) = 0$, також припускається, що $E(\bar{X}_i \varepsilon_{it}) = 0$ і (що є більш важливим) $E(\bar{X}_i \mu_i) = 0$.

Зверніть увагу, що ці умови також вимагають незміщеності від оцінок, отриманих методом найменших квадратів для міжгрупової коваріації.

Легко обчислити оцінку узагальненим методом найменших квадратів (GLS), якщо зауважити, що вона може бути визначена як МНК-оцінка трансформованої моделі [31]:

$$(Y_{it} - \rho \bar{Y}_i) = \alpha(1 - \rho) + (X_{it} - \rho \bar{X}_i) + u_{it}, \quad (5.44)$$

де $\rho = 1 - \theta^{1/2}$.

Випадкові величини в трансформованій регресійній моделі (5.44) є незалежними, однаково розподіленими величинами для всіх об'єктів (індивідуумів) та всіх періодів часу. Зауважимо, що при $\theta = 0$ ми маємо оцінки, отримані методом найменших квадратів для внутрішньогрупової варіації (the within

estimator ($\rho = 1$)). В загальному випадку фіксоване відношення ρ для індивідуальних середніх віднімається від даних для утворення трансформованої моделі (5.44) ($0 \leq \rho \leq 1$).

Звичайно, компоненти дисперсії σ_μ^2 та σ_ε^2 є невідомими на практиці. В такому разі ми можемо використати здійснений метод узагальнених найменших квадратів (GLS), де невідомі дисперсії обчислюються консистентно на першому кроці. Оцінку для σ_ε^2 легко отримати, використовуючи формулу (5.30). Для проміжної регресії, регресії на основі середніх у крос-секційному розрізі (between regression) помилка дисперсії дорівнює:

$$\sigma_\mu^2 + (1/T)\sigma_\varepsilon^2, \quad (5.45)$$

її можна оцінити консистентно з виразу:

$$\hat{\sigma}_B^2 = \frac{1}{N} \sum_{i=1}^N (\bar{Y}_i - \hat{\alpha}_B - \bar{X}_i' \hat{\beta}_B)^2, \quad (5.46)$$

де $\hat{\sigma}_B$ є оцінкою методу найменших квадратів без крос-секційної варіації (between estimator) для α . Звідси консистентна оцінка для σ_μ^2 має такий вигляд:

$$\hat{\sigma}_\mu^2 = \hat{\sigma}_B^2 - \frac{1}{T} \hat{\sigma}_\varepsilon^2. \quad (5.47)$$

Зазначимо, що оцінку (5.47) можна скорегувати на ступені свободи. В такому випадку від знаменника виразу (5.46) необхідно відняти число $(K + 1)$ — кількість регресорів (regressors) в моделі [14, с. 38], [5, с. 15]. Кінцева оцінка, отримана здійсненим методом узагальнених найменших квадратів (EGLS), вважається оцінкою випадкових ефектів для β (і α) і, як правило, позначається $\hat{\beta}_{RE}$. За умов слабкої регулярності (weak regularity conditions) оцінки випадкових ефектів є асимптотично нормально розподіленими з коваріаційною матрицею, що має вигляд:

$$V(\hat{\beta}_{RE}) = \sigma_\varepsilon^2 \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)' + \theta T \sum_{i=1}^N (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})' \right)^{-1}. \quad (5.48)$$

З виразу (5.48) можна побачити, що оцінка параметрів у разі випадкових ефектів є ефективнішою за оцінку фіксованих ефектів, доки $\theta > 0$. Збільшення ефективності відбувається завдяки міжгруповій варіації (between variation) в даних $(\bar{X}_i - \bar{X})$.

Коваріаційна матриця (5.48) обраховується в сучасних пакетах автоматично при оцінюванні методом найменших квадратів (OLS) трансформованої моделі (5.44).

Таким чином, можна зробити висновок про ефективність застосування різних методів оцінювання вектора невідомих параметрів β в моделях лонгітюдних даних.

При оцінюванні методом найменших квадратів без крос-секційної (внутрішньогрупової) варіації (between estimator) використовується властивість лонгітюдних даних, що базується на наявності різниці між об'єктами (індивідуумами). При цьому модель трансформується і є регресією індивідуальних середніх Y від індивідуальних середніх X (та константи). Оцінка такої трансформованої моделі можлива на основі звичайного МНК. При цьому консистентні оцінки отримуються за умов, що $N \rightarrow \infty$ та $E = (\bar{X}_i \alpha_i) = 0$, $E = (\bar{X}_i \varepsilon_i) = 0$. Фактично це означає, що пояснюючі змінні є строго екзогенними та не корелюють з індивідуальними специфічними ефектами μ_i .

При оцінюванні на основі методу найменших квадратів для міжгрупової коваріації (within estimation) використовується властивість лонгітюдних даних, що базується на наявності різниці всередині об'єктів (індивідуумів), тобто експлуатується ідея використання різниці *внутрішнього виміру* даних (відмінностей між окремими елементами вибірки). При цьому модель, що оцінюється методом звичайних МНК, є регресією відхилень від індивідуальних середніх. Метод дає консистентні оцінки для β при $T \rightarrow \infty$ чи $N \rightarrow \infty$, за умови, що $E\{(X_{it} - \bar{X}_i)\varepsilon_{it}\} = 0$. Це, в свою чергу, вимагає, щоб змінні X були строго екзогенними, але не накладає жодних обмежень на зв'язки між μ_i та X_{it} .

Модель лонгітюдних даних можна також оцінювати звичайним МНК, при цьому використовуються обидва виміри (як внутрішній, так і між елементами вибірки (within and between)). Оцінки первісної моделі, отримані в результаті оцінювання, вже не будуть ефективними. Вони є консистентними за умови, що $T \rightarrow \infty$ чи $N \rightarrow \infty$ та $E\{X_{it}(\varepsilon_{it} + \mu_i)\} = 0$. Це означає, що пояснюючі змінні не повинні корелювати з μ_i , але не накладає умови, щоб вони були строго екзогенними. Достатньо, щоб μ_i та ε_{it} були одночасно некорельованими.

Модель лонгітюдних даних можна оцінювати також спеціальним здійсненим методом узагальнених найменших квадратів (EGLS) для моделей з випадковими ефектами. Даний метод комбінує оцінки МНК для міжгрупової коваріації та оцінки МНК без урахування крос-секційної варіації (within and between), але в ефективніший, ніж в попередньому випадку, спосіб. Оцінки є консистентними при $T \rightarrow \infty$ чи $N \rightarrow \infty$, але лише при поєднанні умов для методів як без урахування крос-секційної варіації (between estimator), так і з урахуванням міжгрупової коваріації (within estimator). При цьому оцінку, отриману здійсненим методом узагальнених найменших квадратів (EGLS), можна визначити як зважену середню between- та within-оцінок, або як МНК-оцінку в регресійній моделі, де змінні трансформовані таким чином [10, 31]:

$$(Y_{it} - \rho \bar{Y}_i) \text{ де } \hat{\rho} \text{ є оцінкою } \rho = 1 - \theta^{1/2} \text{ з } \theta = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\mu^2}.$$

5.5. ПРОБЛЕМИ ВИБОРУ МІЖ МОДЕЛЯМИ З ФІКСОВАНИМИ ТА ВИПАДКОВИМИ ЕФЕКТАМИ

Яка модель є кращою — з фіксованими чи з випадковими ефектами? Це питання найчастіше виникає при відборі специфікації моделі лонгітюдних даних. Дійсно, питання, розглядати індивідуальні ефекти μ_i як фіксовані чи випадкові, є досить важливим. Насамперед тому, що при невірній специфікації ми можемо отримати вражаючі відмінності в оцінках β -параметрів в тому випадку, коли T є невеликим, а N — значним. При невеликій часовій вибірці дуже важливо якнайефективніше використати лонгітюдні дані. Найбільш поширеним є погляд, що не варто дискутувати про «справжню природу» μ_i -ефектів. Підходящою інтерпретацією є те, що підхід з точки зору фіксованих ефектів є умовним, накладеним на величину, для μ_i . Це саме той випадок, коли розподіл Y_{it} вважається умовним відносно μ_i , які можна оцінити. Все це має сенс тільки в тому випадку, якщо всі окремі елементи вибірки є однорідними і не можуть бути інтерпретовані як випадкові реалізації з певної генеральної сукупності. Така інтерпретація, мабуть, є найбільш підходящою, коли різними i -ми об'єктами виступають країни, великі підприємства чи компанії, галузі, а прогноз, який нам необхідно зробити відносно окремих об'єктів, повинен базуватись на всій доступній вибірковій інформації.

На відміну від підходу, що базується на фіксованих ефектах, підхід, що базується на випадкових, не є умовним стосовно окремих μ_i , але інтегрує їх вплив. В цьому випадку ми, як правило, не зацікавлені у конкретній величині μ_i певного об'єкта (індивідуума), наша увага скоріше зосереджена на довільних представниках, які мають певні характеристики. Даний підхід дозволяє нам робити припущення та висновки стосовно характеристик генеральної сукупності. Одним із способів формалізації цього є те що, моделі з випадковими ефектами оцінюють залежну змінну як

$$E\{Y_{it}|X_{it}\} = X'_{it}\beta, \quad (5.49)$$

в той час як оцінка залежної змінної в моделі з фіксованими ефектами дорівнює

$$E\{Y_{it}|X_{it}, \mu_i\} = X'_{it}\beta + \mu_i. \quad (5.50)$$

Зауважимо, що коефіцієнти β у виразах (5.49) та (5.50) є однаковими лише за умови $E(\mu_i|X_{it}) = 0$. Отже, основною причиною того, що на практиці перевага часто віддається моделям з фіксованими ефектами, є певний інтерес до значень μ_i , який має сенс, якщо число об'єктів відносно мале і вони мають специфічну природу. Хоча навіть якщо ми зацікавлені у висновках

щодо великої кількості об'єктів і модель з випадковими ефектами здається більш підходящою, оцінка параметрів на основі фіксованих ефектів все одно може мати певні переваги. Причиною цього є можливість виникнення такої ситуації, коли μ_i та X_{it} є корельованими. В такому разі підхід на основі випадкових ефектів з ігноруванням цієї кореляції призведе до неконсистентних оцінок. Прикладом можливої кореляції між μ_i та X_{it} може слугувати розглянута нами вище модель виробничої функції, де μ_i , які відображали управлінські якості, корелювали з іншими змінними, включеними у виробничу функцію. Крім того, можна навести й інші приклади, які досить поширені на практиці [5, 10, 31, 36, 38, 39].

Проблему кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} можна усунути, використовуючи підхід з фіксованими ефектами, при якому μ_i виключаються з моделі при оцінюванні, що автоматично знімає проблему кореляції, або, іншими словами, знімає проблему залежних регресорів (факторів) у моделі. Отже, при наявності кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} перевагу слід віддати специфікації та наступному оцінюванню моделі з фіксованими, а не випадковими ефектами. Але тут, звичайно, постає інше питання: як визначити, що дана кореляція наявна? Існує декілька формальних тестів, які дозволяють виявити наявність кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} , одним з найуживаніших є тест Хаусмана, який ми розглянемо детальніше.

5.5.1. Тест Хаусмана для перевірки кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} в моделях лонгitudних даних

Основним тестом у моделях лонгitudних даних, що дозволяє виявити наявність кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} , а отже, й виявити, яку модель краще застосовувати — з фіксованими чи з випадковими ефектами, є тест Хаусмана (Hausman test). Нульовою гіпотезою тесту Хаусмана є припущення щодо відсутності кореляції між індивідуальними ефектами і пояснюючими змінними моделі.

При цьому основною ідеєю є порівняння двох типів оцінок невідомих параметрів моделі. Першим типом є консистентні оцінки, що є як в умовах нульової, так і в умовах альтернативної гіпотез. Другий тип — це оцінки, консистентні (як правило, й ефективні) тільки при нульовій гіпотезі. Значима різниця між цими двома оцінками означає, що ми повинні відкинути нульову гіпотезу про відсутність кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} моделі.

Розглянемо ідею тесту Хаусмана детальніше.

Припустимо, що $E\{\varepsilon_{it}X_{is}\} = 0$ для всіх s, t . В такому випадку оцінка параметрів $\hat{\beta}_{FE}$ за моделлю з фіксованими ефектами є консистентною для β , незалежно від того, корелюють μ_i з пояснюючими змінними X_{it} чи ні. В той же час оцінка параметрів у моделях з випадковими ефектами є консистентною та ефективною для β лише за умови, що μ_i та X_{it} не корелюють між собою.

Розглянемо тепер величину $(\hat{\beta}_{FE} - \hat{\beta}_{RE})$, яка є вектором різниці двох оцінок. Щоб оцінити значимість цієї різниці, необхідно знати її коваріаційну матрицю. В загальному випадку це вимагало б від нас оцінювання коваріації між $\hat{\beta}_{FE}$ та $\hat{\beta}_{RE}$, але через те, що остання оцінка є ефективною при нульовій гіпотезі, можна показати, що в такому випадку коваріаційна матриця різниці двох оцінок дорівнює різниці коваріаційних матриць відповідних оцінок [5, 31]:

$$V(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = V(\hat{\beta}_{FE}) - V(\hat{\beta}_{RE}). \quad (5.51)$$

Знаючи коваріаційну матрицю (5.51), статистику для тесту Хаусмана можна розрахувати так:

$$\chi_H^2 = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' [\hat{V}(\hat{\beta}_{FE}) - \hat{V}(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}), \quad (5.52)$$

де \hat{V} є оцінкою справжніх коваріаційних матриць. При нульовій гіпотезі, коли $\text{plim}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = 0$, статистика χ_H^2 має асимптотичний χ^2 -розподіл з K ступенями свободи, де K — число параметрів β , які необхідно оцінити.

Фактично тест Хаусмана перевіряє, чи значимо різняться між собою оцінки параметрів за моделями з випадковими та фіксованими ефектами. Обчислення при цьому не є важким тому, що коваріаційна матриця задовольняє умові (5.51). Однією з найважливіших причин різниці оцінок за двома моделями може бути наявність кореляції між μ_i та X_{it} , хоча й інші причини також можливі, зокрема неправильна специфікація моделей. На практиці проблема може виникнути при розрахунку статистики (5.52), тому що коваріаційна матриця у квадратних дужках може не бути позитивно визначеною в кінцевих вибірках, тож обчислити обернену до неї матрицю неможливо. В таких випадках можливо використовувати тест Хаусмана лише для підмножини параметрів β .

Підсумовуючи досить коротко проблему вибору специфікації моделей лонгітюдних даних, відзначимо, що при наявності кореляції між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} -моделей перевагу необхідно віддати моделі з фіксованими ефектами, а при її відсутності — моделі з випадковими ефектами, що, однак, не виключає й можливості використання моделей з фіксованими ефектами. Існує також просте правило пальця, яке дозволяє

попередньо обрати необхідну специфікацію. Так наприклад, якщо на основі моделей лонгитюдних даних нам необхідно зробити висновки з огляду на характеристики генеральної сукупності, краще використовувати моделі з випадковими ефектами, а якщо потрібно зробити висновки лише з огляду на впливи, що є у вибірці, краще обрати модель з фіксованими ефектами.

Звичайно після того, як потрібна специфікація обрана, оцінено параметри відповідним до обраної специфікації методом, постає ще одне, не менш важливе запитання: якою є якість розрахованої моделі? Як ми знаємо, одним з найбільш поширених критеріїв якості в моделях багатофакторної регресії є коефіцієнт детермінації. Він також використовується і в моделях лонгитюдних даних, але має свої особливості, тому розглянемо проблему оцінки якості моделей лонгитюдних даних детальніше.

5.5.2. Оцінка параметрів лонгитюдних даних методом інструментальних змінних

Як вже відмічалось раніше, вибір між двома різними специфікаціями моделей лонгитюдних даних, а саме з фіксованими чи випадковими ефектами, певною мірою залежить від того, чи існує кореляція між індивідуальними ефектами μ_i і пояснюючими змінними X_{it} . При відсутності кореляції між індивідуальними ефектами і пояснюючими змінними можна застосовувати та відповідно оцінювати моделі з випадковими ефектами, а при її наявності перевагу необхідно віддати моделям з фіксованими ефектами. Але при оцінюванні параметрів моделей з фіксованими ефектами ілюмінується вся інформація, пов'язана зі значеннями, інваріантними щодо часу. Це може бути дуже дорогою ціною, яка сплачується за те, щоб дозволити змінним X_{it} корелювати з індивідуальними ефектами μ_i . Наприклад, ми можемо бути зацікавлені в ефекті, що спричиняють інваріантні щодо часу змінні (такі, наприклад, як стать) на заробітну плату індивідуума.

Взагалі ж немає потреби концентрувати всю увагу тільки на виборі між фіксованими та випадковими ефектами, оскільки існує можливість застосування інших підходів оцінювання моделей лонгитюдних даних у випадку наявності кореляції між індивідуальними ефектами та пояснюючими змінними, а саме застосування інструментальних оцінок, або, іншими словами, методу інструментальних змінних, який можна розглядати як щось середнє між підходами оцінювання у випадку фіксованих та випадкових ефектів.

Щоб це показати, згадаємо спочатку загальний вигляд оцінки параметрів у моделях з фіксованими ефектами, який розпишемо таким чином:

$$\begin{aligned}\hat{\beta}_{FE} &= \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)' \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i) = \\ &= \left(\sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)X_{it}' \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T (X_{it} - \bar{X}_i)Y_{it}.\end{aligned}\quad (5.53)$$

Оцінка фіксованих ефектів вигляду (5.53) легко може бути інтерпретована з позицій оцінок інструментальних змінних для параметрів β в моделі лонгітюдних даних:

$$Y_{it} = \alpha + X'_{it} \mu_i + \varepsilon_{it}, \quad (5.54)$$

де для кожної пояснюючої змінної інструментом є відхилення від специфічного індивідуального середнього, тобто для кожного X_{it} інструментами обрані величини $(X_{it} - \bar{X}_i)$. Зауважте, що за конструкцією $E\{(X_{it} - \bar{X}_i)\mu_{it}\} = 0$ (якщо ми розглядаємо математичні сподівання по i та t) метод інструментальних змінних дає консистентні оцінки за умови, коли $E\{(X_{it} - \bar{X}_i)\varepsilon_{it}\} = 0$, що випливає із строгої екзогенності X_{it} . Точніше, якщо відомо, що конкретна змінна X_{it} не корелює з μ_i , то немає потреби знаходити для неї інструмент. В такому разі сама ця змінна може бути застосована як свій власний інструмент. Діючи таким чином, можна оцінити ефект, спричинений інваріантними стосовно часу змінними.

Для пояснення загального підходу використання методу інструментальних змінних для моделей лонгітюдних даних розглянемо як приклад лінійну модель лише з чотирма групами пояснюючих змінних [13]:

$$Y_{it} = \alpha + X'_{1,it} \beta_1 + X'_{2,it} \beta_2 + Z'_{1i} \gamma_1 + Z'_{2i} \gamma_2 + \mu_i + \varepsilon_{it}, \quad (5.55)$$

де X є змінними, що змінюються з часом, а Z є інваріантними стосовно часу змінними. Припустимо, що змінні з індексом 1 не корелюють ані з μ_i , ані з ε_{it} . Змінні $X_{2,it}$ та Z_{2i} корелюють з μ_i , але не корелюють з жодним ε_{it} . За таких припущень оцінка фіксованих ефектів була б консистентною для β_1 та β_2 , але вона не визначала б коефіцієнти для змінних, які є інваріантними щодо часу. Дж. Хаусман і У. Тейлор [13] запропонували оцінити модель (5.55) за допомогою методу інструментальних змінних, використовуючи як інструменти змінні $X_{1,it}, Z_{1i}, X_{2,it} - \bar{X}_{2,i}, \bar{X}_{1,i}$. В цьому випадку екзогенні змінні служать власними інструментами, інструментами для $X_{2,it}$ є відхилення цих змінних від індивідуальних середніх (так само, як і у підході оцінювання моделей з фіксованими ефектами), а інструментами для Z_{2i} є індивідуальні середні для $X_{1,it}$. Відмітимо, що ідентифікація (детальніше про проблему ідентифікації (ототожнення) див. розділ 4) вимагає, щоб число змінних в $X_{1,it}$

було не меншим за число змінних в Z_{2i} . Описаний вище метод оцінювання (метод Хаусмана—Тейлора) дозволяє оцінити ефект, спричинений інваріантними щодо часу змінними, навіть якщо змінні, що залежать від часу, корелюють з μ_i . Якщо виявиться, що інваріантні змінні також корелюють з μ_i , то для них також необхідно знаходити інструменти. Крім того, важливим моментом є включення в модель достатньої кількості змінних, які змінюються з часом, та не корелюють з μ_i . Звичайно, застосування додаткових інструментів, які не базуються на змінних, що ввійшли до моделі, є прямим ускладненням процесу оцінювання. Але це те, що ми змушені робити, коли жодна з трансформацій не може дати дієвих інструментів.

Значною перевагою підходу Хаусмана—Тейлора є те, що він не потребує використання жодних зовнішніх інструментів. Достатні інструменти можна отримати зі змінних самої моделі. Але незважаючи на цю важливу перевагу, оцінка Хаусмана—Тейлора відіграє на диво незначну роль у сучасних емпіричних дослідженнях.

Хаусман і Тейлор також довели, що застосування запропонованого вище набору інструментів еквівалентне використанню змінних $X_{1,it} - \hat{X}_{1i}, X_{2,it} - \bar{X}_{2i}$ та $X_{1,it}, Z_{1i}$. Це впливає з того факту, що використання різноманітних лінійних комбінацій первинних інструментів не впливає на метод оцінювання. Також було показано, як недіагональна коваріаційна матриця помилок (error term) моделі (4.58) може бути використана з метою поліпшення ефективності оцінки. Такий підхід використовується в методі загальних моментів (GMM).

У [2] була зроблена спроба поліпшити ефективність оцінки інструментальних змінних Хаусмана—Тейлора за допомогою значно більшого набору інструментів. Було також запропоновано використовувати інваріантні стосовно часу інструменти $X_{1,i1} - \bar{X}_{1i}, \dots, X_{1,iT} - \bar{X}_{1i}$, що вимагає, аби $E\{(X_{1,it} - \bar{X}_{1i})\mu_i\} = 0$ для кожного t . Це припущення має сенс тоді, коли кореляція між μ_i та $X_{1,it}$ залежить від інваріантної щодо часу компоненти $X_{1,it}$, такої, що $E\{X_{1,it}\mu_i\}$ для конкретного t не залежить від цього t . В роботі [7] було певним чином підсумовано вищезгадані підходи до оцінювання методом інструментальних змінних та запропоновано використовувати змінні $X_{2,i1} - \bar{X}_{2i}, \dots, X_{2,iT} - \bar{X}_{2i}$ як додаткові інструменти.

Сьогодні методи інструментальних змінних є досить поширеними на практиці, оскільки дають змогу використовувати специфікацію моделей лонгітюдних даних з випадковими ефектами навіть у випадку наявності кореляції між індивідуальними ефектами та пояснюючими змінними.

Таким чином, практично знімається вплив проблем оцінювання при вирішенні питання, якій моделі віддати перевагу — з фіксованими чи з випадковими ефектами. Та чи інша специфікація моделі повинна обиратись, виходячи з економічного змісту проблеми, що аналізується.

Звичайно, після того як оцінено параметри моделей лонгітюдних даних відповідним до обраної специфікації методом, постає ще одне, не менш важ-

ливе запитання: якою є якість розрахованої моделі? Як ми знаємо, одним з найбільш поширених критеріїв якості в моделях багатофакторної регресії є коефіцієнт детермінації. Він також використовується і в моделях лонгітюдних даних, але має свої особливості, тому розглянемо проблему оцінки якості моделей лонгітюдних даних детальніше.

5.6. КРИТЕРІЇ ЯКОСТІ ОЦІНЮВАННЯ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ

Визначення критеріїв якості оцінювання моделей лонгітюдних даних є певною мірою відмінним від класичного багатофакторного аналізу. Одна з причин цього полягає в тому, що в залежності від досліджуваних проблем важливість внутрішньогрупової та міжгрупової варіації даних може бути різною. Іншою причиною є те, що як звичайний, так і скорегований коефіцієнти детермінації (R^2) можуть застосовуватися лише у випадку оцінювання моделей звичайним методом найменших квадратів. Тому при визначенні якості моделей лонгітюдних даних береться до уваги значення коефіцієнта детермінації R^2 , розрахованого як квадрат коефіцієнта кореляції між фактичними та теоретично розрахованими значеннями досліджуваного показника. При цьому значення коефіцієнта детермінації знаходяться в інтервалі $[0, 1]$ незалежно від того, який метод оцінювання було застосовано для отримання розрахованих (теоретичних) значень. Воно також відповідає стандартному визначенню R^2 (в термінах суми квадратів), якщо модель оцінюється за допомогою звичайного методу найменших квадратів (за умови, що перетин включено до моделі).

В такому контексті сумарна варіація в значеннях Y_{it} може бути записана як сума внутрішньогрупової та міжгрупової варіації (within- and between variation):

$$\frac{1}{NT} \sum_{i,t} (Y_{it} - \bar{Y})^2 = \frac{1}{NT} \sum_{i,t} (Y_{it} - \bar{Y}_i)^2 + \frac{1}{N} \sum_i (\bar{Y}_i - \bar{Y})^2, \quad (5.56)$$

де \bar{Y} позначає загальну середню по всій вибірці. Виходячи з виразу (5.56), можна визначити альтернативні версії коефіцієнта детермінації R^2 в моделях лонгітюдних даних в залежності від того, якому типу варіації в даних ми віддаємо перевагу.

Наприклад, якщо модель з фіксованими ефектами була обрана з тим, щоб пояснити внутрішньогрупову варіацію в даних якомога краще, то в такому разі бажано розглядати внутрішньогруповий коефіцієнт детермінації (within — R^2) як один з критеріїв якості моделі:

$$R_{within}^2(\hat{\beta}_{FE}) = corr^2\left(\left(\hat{Y}_{it}^{FE} - \hat{Y}_i^{FE}\right), \left(Y_{it} - \bar{Y}_i\right)\right), \quad (5.57)$$

де $\hat{Y}_{it}^{FE} - \hat{Y}_i^{FE} = (X_{it} - \bar{X}_i)' \hat{\beta}_{FE}$, а $corr^2$ означає квадрат коефіцієнта кореляції.

При використанні проміжних оцінок параметрів моделі методом найменших квадратів без крос-секційної (внутрішньогрупової) варіації (between estimator), який фактично є методом найменших квадратів у трансформованій моделі середніх значень, краще використовувати коефіцієнт детермінації без врахування крос-секційної варіації (between — R^2), який визначається за формулою:

$$R_{between}^2(\hat{\beta}_B) = corr^2(\hat{Y}_i^B, Y_i), \quad (5.58)$$

де $\hat{Y}_i^B = \bar{X}_i' \hat{\beta}_B$.

При оцінці параметрів за допомогою звичайного методу найменших квадратів загальна якість оцінки моделі визначається за допомогою загального коефіцієнта детермінації, який визначається так:

$$R_{overall}^2(\hat{\beta}) = corr^2(\hat{Y}_{it}, Y_{it}), \quad (5.59)$$

де $\hat{Y}_{it} = X_{it}' \hat{\beta}$.

Зауважимо, що в деяких пакетах прикладних програм можливо розраховувати коефіцієнти детермінації і в разі відсутності перетину в моделі, тобто можна визначати внутрішньогруповий, без врахування крос-секційної варіації, та загальний коефіцієнти детермінації в моделях лонгитюдних даних для довільної оцінки $\hat{\beta}$, використовуючи теоретично обчислені значення

$$\hat{Y}_{it} = X_{it}' \hat{\beta}, \hat{Y}_i = \frac{1}{T} \sum_{t=1}^T \hat{Y}_{it}, \hat{Y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{Y}_{it}$$

за моделями з опущеним перетином. При оцінюванні моделей з фіксованим ефектом такий підхід фактично ігнорує варіацію індивідуальних ефектів μ_i .

Отже, підсумовуючи вищесказане, можна зробити висновок, що в термінах коефіцієнта детермінації (а точніше, в термінах квадрата коефіцієнта кореляції) три міри якості можуть бути розраховані для будь-яких розглянутих нами методів оцінювання моделей лонгитюдних даних, а саме: внутрішньогруповий, без врахування крос-секційної варіації, та загальний коефіцієнти детермінації.

Зауважимо, що коефіцієнти детермінації у випадку моделей лонгитюдних даних не можуть слугувати індикатором вибору між двома типами моделей з випадковими або з фіксованими ефектами. Вони дають лише додатну оцінку для вибору альтернативної з множини можливих специфікацій моделі.

5.7. ОСОБЛИВИ ВИПАДКИ В МОДЕЛЯХ ЛОНГІТЮДНИХ ДАНИХ. ТЕСТУВАННЯ НАЯВНОСТІ АВТОКОРЕЛЯЦІЇ ТА ГЕТЕРОСКЕДАСТИЧНОСТІ

Можливість наявності автокореляції в моделях лонгитюдних даних, як правило, є менш поширеною, ніж можливість порушення припущення про постійне значення дисперсії, тобто можливість гетероскедастичності, яка призводить до досить неприємних наслідків. Однак як наявність автокореляції, так і наявність гетероскедастичності необхідно тестувати, щоб відповідно відреагувати на порушення будь-якого з цих припущень, тому розглянемо детально особливості тестування автокореляції та гетероскедастичності в моделях лонгитюдних даних.

Відмітимо, що більшість тестів, які застосовують для виявлення гетероскедастичності та автокореляції в моделях з випадковими ефектами, досить важко розраховувати. Але їх розрахунок не є складним в моделях з фіксованими ефектами, які оцінюються звичайним методом найменших квадратів. На щастя, оскільки оцінка фіксованих ефектів може бути застосована навіть тоді, коли ми припускаємо існування випадкових ефектів (тобто коли індивідуальні ефекти μ_i є однаково розподіленими незалежними випадковими величинами, які, крім того, не залежать від пояснюючих змінних), тести, які зазвичай застосовують для моделей з фіксованими ефектами, також можуть бути застосовані до моделей з випадковими ефектами. Розглянемо деякі з них більш детально.

Як ми знаємо з багатofакторного аналізу, найбільш поширеним тестом перевірки моделі на наявність автокореляції є тест Дарбіна—Уотсона. У випадку моделей з фіксованими ефектами перевірка на автокореляцію також базується на цьому тесті, причому базове рівняння даного тесту має вигляд:

$$\varepsilon_{it} = \rho \varepsilon_{i,t-1} + v_{it}, \quad (5.60)$$

де v_{it} — незалежні однаково розподілені випадкові величини для всіх об'єктів та часових періодів. Як можна побачити з базового рівняння (5.60), тест на автокореляцію в моделях з фіксованими ефектами враховує її в часових періодах, але накладає припущення про те, що кожний об'єкт вибірки має однаковий коефіцієнт автокореляції ρ .

Нульовою гіпотезою тесту Дарбіна—Уотсона є гіпотеза про рівність коефіцієнта автокореляції ρ нулю: $H_0: \rho = 0$ проти односторонньої альтернативної гіпотези $\rho < 0$ чи $\rho > 0$.

Для того, щоб проілюструвати цей тест детальніше, позначимо через $\hat{\varepsilon}_{it}$ помилки (залишки) від регресії зі змінними у відхиленнях від середнього (крос-секційна трансформація) (5.15), або, що є еквівалентним, від регресійної

моделі вигляду (5.10). На основі даних залишків розраховується статистика Дарбіна—Уотсона для моделей лонгитюдних даних, запропонована в [6]:

$$DW_{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T (\hat{\varepsilon}_{it} - \hat{\varepsilon}_{i,t-1})^2}{\sum_{i=1}^N \sum_{t=2}^T \hat{\varepsilon}_{it}^2}. \tag{5.61}$$

Використовуючи таку статистику, автори [6], так само, як і Дарбін та Уотсон, змогли отримати нижні та верхні межі критичних значень даної статистики для моделей лонгитюдних даних, які залежать тільки від N , T і K . На відміну від класичного випадку, область невизначеності тесту Дарбіна—Уотсона для лонгитюдних (панельних) даних є дуже малою, особливо при великій кількості елементів вибірки. У таблиці 5.1 представлено певні нижні та верхні межі критичних значень статистики Дарбіна—Уотсона для моделей лонгитюдних даних при рівні значимості 5%. Дані критичні значення [6] можуть бути застосовані для тестування наявності автокореляції. Наприклад, для моделі з трьома пояснюючими змінними, оціненими по 6 періодах для 100 об'єктів, ми відхиляємо нульову гіпотезу $H_0: \rho = 0$ при рівні значимості 5%, якщо DW_{ρ} є меншим за 1,895 для $N = 100$, проти альтернативної гіпотези, що $\rho < 0$. Для лонгитюдних даних з досить великою кількістю об'єктів у [6] запропоновано просто протестувати, чи розрахована статистика DW_{ρ} є меншою за 2, а вже після цього тестувати на наявність позитивної автокореляції. Оскільки оцінка фіксованих ефектів є консистентною і для моделі з випадковими ефектами, то можна використовувати тест Дарбіна—Уотсона для лонгитюдних даних і в останній моделі.

Таблиця 5.1. 5%-ні верхні та нижні межі для тесту Дарбіна—Уотсона [6]

T		N = 100		N = 500		N = 1000	
		d_L	d_U	d_L	d_U	d_L	d_U
T = 6	K = 3	1.859	1.880	1.939	1.943	1.957	1.959
	K = 9	1.839	1.902	1.935	1.947	1.954	1.961
T = 10	K = 3	1.891	1.904	1.952	1.954	1.967	1.968
	K = 9	1.878	1.916	1.949	1.957	1.965	1.970

Тестування гетероскедастичності в моделях лонгитюдних даних за логікою мало відрізняється від тестування гетероскедастичності в багатofакторних регресійних моделях. Найбільш поширеним тестом є варіант відомого тесту Бройша—Пагана (Breusch—Pagan) на гетероскедастичність. Для тестування гетероскедастичності випадкової величини ε_{it} , як і у випадку тестування автокореляції, використовуються залишки (помилки) $\hat{\varepsilon}_{it}$ від регресії зі змінними у відхиленнях від середнього (крос-секційна трансформація)

(5.15), або, що є еквівалентним, від регресійної моделі вигляду (5.10). На основі даних залишків оцінюється допоміжна регресійна модель залежності $\hat{\varepsilon}_{it}^2$ від константи та факторів X_{it} , які, з погляду дослідників, можуть провокувати гетероскедастичність.

Базове рівняння найпоширенішого варіанту даного тесту має вигляд:

$$V(\varepsilon_{it}) \approx \hat{\varepsilon}_{it}^2 = \alpha_0 + \alpha_1 X_{1it} + \dots + \alpha_L X_{Lit} + v_{it}, \quad (5.62)$$

де v_{it} — незалежні однаково розподілені випадкові величини для всіх об'єктів та часових періодів. Базове рівняння (5.62) можна переписати у такому матричному вигляді [31, с. 325]:

$$V\{\varepsilon_{it}\} = \sigma^2 h(X'_{it}\alpha), \quad (5.63)$$

де h є невідомою неперервною диференційованою функцією, такою, що $h(0) = 1$. Дана умова дозволяє сформулювати нульову гіпотезу тесту Бройша—Пагана у вигляді: $H_0: \alpha = 0$, що рівнозначно умові $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_L = 0$, проти альтернативної, що хоча б одне з цих значень не дорівнює нулю.

Можна показати, що величина $[N(T-1)R^2]$ має χ^2 -розподіл з L ступенями свободи, де T — кількість періодів часу, N — кількість об'єктів, L — кількість факторів, що можуть провокувати гетероскедастичність, R^2 — коефіцієнт детермінації допоміжної регресії.

Якщо розрахована χ^2 -статистика перевищує критичне значення, знайдене за таблицями χ^2 -розподілу при заданому рівні значимості та відповідних ступенях свободи, то нульова гіпотеза відкидається.

Зауважимо, що існують і інші модифікації тесту Бройша—Пагана, які, наприклад, можуть базуватись на залишках з регресії на основі середніх у крос-секційному розрізі (between regression) тощо.

5.8. КОРЕГУВАННЯ СЕРЕДНЬОКВАДРАТИЧНИХ ВІДХИЛЕНЬ ПАРАМЕТРІВ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ ПРИ НАЯВНОСТІ АВТОКОРЕЛЯЦІЇ ТА (АБО) ГЕТЕРОСКЕДАСТИЧНОСТІ

Наявність автокореляції та (або) гетероскедастичності приводить до неможливості застосовувати середньоквадратичні відхилення параметрів моделі для тестування різноманітних гіпотез. Проте завдяки сучасним пакетам прикладних програм корекція середньоквадратичних помилок параметрів моделей лонгitudних даних при наявності автокореляції та гетероскедастичності не є проблемою. Отже, практична сторона корегування автокореляції та ге-

тероскедастичності не викликає жодних труднощів, тому розглянемо проблему корегування з теоретичної точки зору уважніше.

Відмітимо, що моделі як з випадковими, так і з фіксованими ефектами побудовані на припущенні, що наявність μ_i абсорбує всю кореляцію між змінними, що описують неспостережувані (unobservables) властивості в різні проміжки часу. Фактично це означає, що припускається некорельованість випадкових величин ε_{it} і для об'єктів (індивідуумів), і для будь-яких проміжків часу.

За умови, що X_{it} є строго екзогенними, наявність автокореляції в ε_{it} не призводить до неконсистентності оцінок, але робить недійсними їх стандартні помилки та результуючі тести. Більш того, це означає, що оцінки більше не є ефективними. Наприклад, якщо справжня коваріаційна матриця Ω не задовольняє умові [31]

$$\text{Var-Cov}(\mu_i i_T' + \varepsilon_i) = \Omega = \sigma_\mu^2 i_T i_T' + \sigma_\varepsilon^2 I_T, \quad (5.64)$$

то оцінка параметрів моделі з випадковими ефектами більше не відповідає можливій оцінці, отриманій здійсненим методом узагальнених найменших квадратів (EGLS) для β .

Як відомо, наявність гетероскедастичності в ε_{it} або для моделей з випадковими ефектами в μ_i має однакові наслідки [5, 10, 31].

Одним із способів уникнути помилкових висновків у разі наявності автокореляції або гетероскедастичності, не накладаючи при цьому альтернативних припущень на структуру коваріаційної матриці Ω , є використання оцінок β , отриманих звичайним методом найменших квадратів, та наступне корегування стандартних помилок на наявність загальних форм гетероскедастичності та автокореляції.

Щоб проілюструвати це положення, розглянемо загальну модель лонгитюдних даних

$$Y_{it} = X'_{it} \beta + u_{it}, \quad (5.65)$$

причому слід зазначити, що ми не висуваємо жодного припущення щодо розкладання u_{it} на певні складові. Консистентність оцінок β даної моделі, отриманих за допомогою звичайного методу найменших квадратів (МНК):

$$b = \left(\sum_{i=1}^N \sum_{t=1}^T X_{it} X'_{it} \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T X_{it} Y_{it} \quad (5.66)$$

вимагає, щоб

$$E\{X_{it} u_{it}\} = 0. \quad (5.67)$$

Припустивши, що помилки різних елементів вибірки не корелюють між собою ($E\{u_{it}u_{js}\} = 0$ для всіх $i \neq j$), коваріаційна матриця звичайного методу найменших квадратів може бути обчислена за допомогою одного з різновидів оцінки Невея—Веста (Newey—West), яка є практично в усіх сучасних статистичних пакетах:

$$\hat{V}\{b\} = \left(\sum_{i=1}^N \sum_{t=1}^T X_{it} X'_{it} \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \sum_{s=1}^T \hat{u}_{it} \hat{u}_{is} X_{it} X'_{is} \left(\sum_{i=1}^N \sum_{t=1}^T X_{it} X'_{it} \right)^{-1}, \quad (5.68)$$

де \hat{u}_{it} позначають помилки (залишки). Дана оцінка враховує загальні форми гетероскедастичності так само, як і автокореляцію (всередині окремо взятих груп елементів вибірки). Якщо припущення стосовно гетероскедастичності є апріорним, матриця всередині виразу (5.68) може бути замінена на:

$$\sum_{i=1}^N \sum_{t=1}^T \sum_{s=1}^T \left(\frac{1}{N} \sum_{i=1}^N \hat{u}_{it} \hat{u}_{is} \right) X_{it} X'_{is}, \quad (5.69)$$

де $\frac{1}{N} \sum_{i=1}^N \hat{u}_{it} \hat{u}_{is}$ є консистентною оцінкою для $\Omega_{is} = E\{u_{it}u_{is}\}$.

Якщо u_{it} має інваріантну щодо часу компоненту μ_i , яка може корелювати з пояснюючими змінними, оцінка фіксованих ефектів у цьому випадку може бути більш прийнятною, аніж оцінка, отримана за допомогою звичайного методу найменших квадратів, для якої також є можливість застосувати корекцію на гетероскедастичність та автокореляцію [3]. Вираз, отриманий в результаті підстановки, є ідентичним виразу (5.68), але кожна змінна X_{it} замінюється на її внутрішньогрупову трансформацію $X_{it} - \bar{X}_i$, а МНК-помилки замінюються на внутрішньогрупові помилки (залишки) [5, с. 13].

Якщо у дослідника виникне бажання накласти специфічні припущення стосовно форм гетероскедастичності чи автокореляції, то є можливість отримати ефективніші оцінки, аніж оцінки фіксованих ефектів чи ті, які отримані за допомогою звичайного методу найменших квадратів. Це можна зробити, використовуючи структуру коваріаційної матриці помилок чи застосувавши прийнятний узагальнений метод найменших квадратів або підхід з точки зору максимальної правдоподібності. Огляд таких оцінок, які, втім, не є привабливими через велику складність їх розрахунку, зробив Б. Балтажі (B. Baltagi) [5, розділ 5]. Я. Кмента (J. Kmenta) [18] запропонував відносно простий і цілком прийнятний спосіб оцінювання за допомогою узагальненого методу найменших квадратів, котрий враховує автокореляцію першого порядку u_{it} , що поєднується з індивідуальною специфічною гетероскедастичністю, але не бере до уваги інваріантну щодо часу компоненту u_{it} [5].

Отже, при наявності автокореляції та (або) гетероскедастичності в моделях лонгитюдних даних необхідно застосувати корекцію середньоквадратичних відхилень (стандартних помилок) оцінених параметрів, що робить можливим подальше тестування гіпотез та побудову інтервалів довіри.

5.9. ЗАСТОСУВАННЯ МОДЕЛЕЙ ЛОНГІТЮДНИХ ДАНИХ НА ПРАКТИЦІ: МОДЕЛЮВАННЯ ПОДАТКОВИХ НАДХОДЖЕНЬ НА РЕГІОНАЛЬНОМУ РІВНІ. ПРОГНОЗУВАННЯ ВИТРАТ МІСЦЕВИХ ОРГАНІВ ВЛАДИ ДЛЯ УКРАЇНИ

Переваги та недоліки моделей лонгитюдних даних

Використання лонгитюдних даних в емпіричних дослідженнях дає важливий інструментарій для економічного аналізу. В роботі одного з найвідоміших західних фахівців з моделювання на основі лонгитюдних даних Балтажі [5, с. 3—7] дається класифікація переваг та додаткових можливостей використання лонгитюдних даних.

- **Можливість урахування неоднорідності об'єктів.** Фірми і люди є неоднорідними і багатовимірними агентами. Очевидно, що деякі характеристики є неспостережуваними і/або невимірюваними, наприклад якість використуваних ресурсів, якість управління та робочої сили. У такій ситуації виключення важливих, але неспостережуваних факторів призводить до зміщених оцінок параметрів регресії. Проте, якщо ці фактори є незмінними в часі, ми можемо отримати незміщену оцінку, застосовуючи моделі на основі лонгитюдних даних, а відповідно, і спеціальні методи оцінювання невідомих параметрів.

- **Збільшення інформаційної бази.** Лонгитюдні дані забезпечують ширшу варіацію в даних. Ця важлива особливість зменшує загрозу мультиколінеарності. Крім того, використання лонгитюдних даних значно збільшує ступені свободи і підвищує ефективність оцінок параметрів.

- **Динаміка процесів.** Деякі процеси приховані на макрорівні. Наприклад, оборот праці важко виявити на макрорівні, оскільки агреговані змінні можуть бути порівняно стабільними, тоді як на мікрорівні вони набувають значних змін: звільнення можуть бути компенсовані високим наймом, залишаючи рівень безробіття незмінним. Тому варто брати до уваги і мікрорівень, на якому саме економічні агенти приймають рішення, наприклад, залишитись на поточній роботі чи змінити свій статус. Ця крос-варіаційна (тобто варіація у розрізі економічних одиниць: робітників, фірм тощо) особливість інформації доповнюється ще й часовим виміром у випадку лонгитюдних да-

них. Таким чином, ми можемо слідкувати за результатами впливу різноманітних факторів на прийняття рішень окремих економічних агентів протягом певних періодів часу, а відповідно, і простежити економічні процеси.

■ **Можливість побудови складніших моделей.** Оскільки лонгітюдні дані поєднують крос-варіаційний вимір з часовим, ми можемо створити складніші моделі економічних процесів і поведінки людей, ніж при окремому застосуванні варіаційних чи часових рядів.

■ **Зменшення зміщень в агрегованих даних.** Інколи вимірювання даних на мікрорівні є набагато інформативнішими та точнішими, ніж на макрорівні. Наприклад, дуже часто гіпотезу про постійний дохід життєвого циклу (life-cycle/permanent income) економісти перевіряють на макрорівні. Проте на макрорівні практично неможливо розподілити неоднорідні домашні господарства на більш-менш однорідні групи (але, за гіпотезою, поведінка цих груп повинна значно відрізнятись, наприклад, за гіпотезою: молодь та старші люди не заощаджують, тоді як ті, що працюють, заощаджують). Тому перевірка такої гіпотези на мікрорівні з використанням лонгітюдних даних дає набагато точніші результати.

Звичайно, метод використання лонгітюдних даних має свої недоліки. Балтажі розглядає також небезпечні пастки, у які ми можемо потрапити, застосовуючи дану економетричну техніку.

■ **Проблема відбору.** У лонгітюдних даних деякі об'єкти інформації можуть представлятись частіше, ніж інші. Наприклад, лише фірми, що ефективно працюють, як правило, представлені у вибірці протягом тривалого періоду часу, тоді як збиткові фірми часто банкрутують і, таким чином, вибувають із вибірки. Іншим прикладом може служити опитування населення через пошту. Лише той, хто насправді зацікавлений, витратить свій час на написання листа і відправлення його поштою. Тому такий вид опитування часто може бути нерепрезентативним. Такі систематичні зміщення у вибірці призводять, в свою чергу, до зміщених оцінок моделей і зводять результат застосування економетричного моделювання до мінімуму. Формами такого упередженого відбору можуть бути: уникнення відповіді (багаті індивіди уникають відповідей щодо власного доходу), самовідбір (лише певні групи можуть потрапити до вибірки через наявність певної характеристики), відмирання (збиткові фірми банкрутують та вибувають із вибірки).

■ **Помилка виміру.** Переважно помилки на макрорівні знищують одну одну і ми отримуємо більш-менш точну оцінку даних. На жаль, на мікрорівні такі помилки трапляються досить систематично, тому такі дані не можуть бути «достатньо корисними» в прикладних економетричних дослідженнях.

■ **Збір даних.** Ця проблема виникає власне через природу досліджень на мікрорівні, оскільки завжди важко розробити репрезентативну вибірку та неупереджено зібрати інформацію (наприклад, інтерв'юери можуть ви-

кривляти дані, оскільки інтерпретують відомості від об'єктів вибірки через призму власних цінностей).

■ **Короткий період часу.** Збір інформації про велику кількість економічних агентів за великий проміжок часу (5—10 років) коштує досить дорого. В результаті часовий вимір набору лонгитюдних даних, як правило, не перевищує 3—5 спостережень в часовому розрізі.

Отже, використовуючи лонгитюдні дані, ми маємо як певні вигоди, так і проблеми. Проте у більшості випадків вигоди переважають, і цей метод стає досить важливим у прикладних економічних дослідженнях.

Для того, щоб проілюструвати практичне застосування моделей лонгитюдних даних та їх особливості, розглянемо два завдання: 5.1 та 5.2. Перше стосується моделювання податкових надходжень для регіонів України, друге — моделювання витрат місцевих органів влади України на регіональному рівні.

Завдання 5.1.

Необхідно специфікувати та оцінити модель лонгитюдних даних для прогнозу податкових надходжень для регіонів України. Розглянемо прогнозування податкових надходжень на прикладі прибуткового податку з громадян.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. Зауважимо, що в процесі виконання цього завдання ми будемо використовувати пакет прикладних програм E.VIEWS.

Крок 1. Постановка задачі моделювання надходжень від прибуткового податку з громадян на регіональному рівні та специфікація моделі.

Детально опишіть постановку та особливості моделювання надходжень від прибуткового податку з громадян на регіональному рівні. Прокоментуйте перевагу від застосування моделей лонгитюдних даних. Обґрунтуйте, який тип моделі — з фіксованими чи з випадковими ефектами — краще обрати для розв'язання поставленого завдання.

Крок 1.1. Формалізований вигляд моделі та позначення.

Запишіть специфіковану вами модель у формалізованому вигляді, введіть позначення для залежної змінної та екзогенних змінних (факторів).

Крок 2. Оцінювання моделі лонгитюдних даних.

Крок 2.1. Введення даних для розробки моделей лонгитюдних даних в пакеті E.VIEWS.

Введіть лонгитюдні дані для подальшого використання їх для моделювання. Поясніть особливості введення лонгитюдних даних при застосуванні пакета E.VIEWS.

Крок 2.2. Створення спеціального пул-об'єкта (pool) для розрахунків за моделлю лонгітюдних даних.

Прокоментуйте, що являє собою об'єкт *pool* в моделюванні на основі лонгітюдних даних. Як ми його створюємо в пакеті E.VIEWS?

Крок 2.3. Оцінювання невідомих параметрів моделі лонгітюдних даних.

Проведіть оцінювання специфікованої вами моделі надходжень від прибуткового податку з громадян. Детально прокоментуйте отримані результати.

Крок 2.4. Тест Дарбіна—Вю—Хаусмана для вибору специфікації моделі у вигляді моделі з фіксованими чи з випадковими ефектами.

Проведіть тест Дарбіна—Вю—Хаусмана, який дозволяє вирішити, чи є модель з випадковими ефектами більш підходящою в даному випадку, чи можна залишити модель з фіксованими ефектами. Які висновки можна зробити за отриманими вами результатами? Поясніть².

Крок 2.5. Тест на перевірку загального перетину.

Проведіть F-тест на перевірку факту, чи буде модель із загальним перетином (Pooled model) кращою за специфіковану вами модель. Зауважимо, що модель із загальним перетином є фактично багатofакторною регресійною моделлю, побудованою на всій множині лонгітюдних даних. Прокоментуйте отримані вами результати. Які висновки щодо остаточної специфікації моделі лонгітюдних даних ви можете зробити?

Крок 3. Загальні висновки та коментарі.

Зробіть загальні висновки щодо моделювання надходжень від прибуткового податку з громадян. Яку модель краще використовувати — із загальним перетином, з фіксованими ефектами чи з випадковими ефектами — та чому? Зробіть загальні висновки щодо отриманих результатів. Обґрунтуйте їх.

Розв'язок завдання 5.1 та розширений коментар моделювання надходжень від прибуткового податку з громадян на регіональному рівні

Зауважимо, що моделювання надходжень прибуткового податку з громадян на регіональному рівні ми проводимо за допомогою економетричного

² Зауважимо, що тесту Дарбіна—Вю—Хаусмана на зміщеність оцінок в моделях з випадковими ефектами немає в пакеті E.VIEWS, він повинен проводитись або покроково, або в командному режимі, тому ми також наводимо написану нами програму для застосування тесту Дарбіна—Вю—Хаусман в пакеті E.VIEWS.

пакета E.VIEWS. Для побудови моделі використовувались щомісячні дані в розрізі регіонів України, наведені в таблиці 5.1.

Крок 1. Постановка задачі моделювання надходжень від податку з прибутку громадян на регіональному рівні та специфікація моделі.

Огляд економічної літератури показує, що один з найважливіших факторів, що впливає на кількість зібраних податків, — це місцевий дохід, який може бути апроксимовано валовим регіональним випуском, або, точніше, доданою вартістю для регіонів. На жаль, підрахунків регіональної доданої вартості для перших років незалежної України немає. Щоб вирішити цю проблему, пропонуємо приблизно оцінити додану вартість кількістю випущеної промислової продукції в регіонах. Основна проблема, що виникає при цьому, пов'язана з подвійним рахунком в даних. Ми припускаємо, що області є достатньо великими територіальними об'єктами і відмінності у виробничих ланцюгах вирівнюються на обласному рівні. Щоб зменшити можливі невідповідності в апроксимації місцевого доходу, ми також пропонуємо використовувати сільськогосподарський та промисловий випуск, оскільки деякі області отримують відносно більшу частину доходу від аграрної продукції, а не від промислової і навпаки. Наприклад, багато західних областей отримують прибуток від продажу сільськогосподарської продукції, тоді як більшість східних областей заробляють на промисловому випуску. Крім того, ми включимо середню заробітну плату як додатковий фактор в регресію³. Очевидно, всі ці змінні повинні мати додатний градієнт, тобто збільшення цих показників, що апроксимують місцевий дохід, повинні збільшувати місцеві податкові надходження.

Оскільки безробітні не сплачують податку з доходу, можна очікувати, що високий рівень безробіття знизить податкову базу (заробітну плату) в області, а це, відповідно, зменшить доходи від оподаткування. Дані про безробіття — предмет гострої критики, оскільки їх обрахунок на сьогодні ще погано розроблений. Офіційний рівень безробіття обчислюється на основі зареєстрованих безробітних. В Україні допомога з безробіття дуже низька, а бюрократична тяганина надто виснажлива. Тому багато безробітних офіційно не зареєстровані. Тим не менше припускаємо, що зміни офіційного рівня безробіття відповідають змінам фактичного. Іншими словами, кореляція між змінами в офіційному та фактичному безробітті повинна бути більшою, ніж

³ Оскільки ми зважуємо всі змінні по чисельності населення (тобто розглядаємо показники на душу населення), середня заробітна плата є, по суті, апроксимацією фонду заробітної плати в регіоні. Варто також зауважити, що, відповідно до статистики доходів, заробітна плата становить лівову частку в доходах населення.

кореляція між рівнями зареєстрованого і фактичного безробіття. Тому ми очікуємо негативний знак для коефіцієнта цієї змінної.

У розвинутих країнах заборгованість із заробітної плати рідко відхиляється від запланованої у бюджеті. Однак в Україні, як і в багатьох колишніх республіках Радянського Союзу, заборгованість із зарплати стала практично постійним явищем [9]. Зростання заборгованості із заробітної плати зменшує базу оподаткування податку з прибутку громадян, який надходить до місцевих бюджетів. Таким чином, існує зворотний зв'язок між заборгованістю із заробітної плати та доходом місцевих органів влади. Заборгованість є показником запасу, а дохід є показником потоку. Щоб мати змогу порівнювати дані показники, ми будемо використовувати зміну (приріст) *заборгованості*. Оскільки приріст заборгованості ще не використовувався для аналізу доходів від оподаткування в інших країнах, цей термін можна вважати специфічним для України.

Так само ми розглядаємо зміну в запасах заборгованості між підприємствами. Цей тип заборгованості може прямо чи опосередковано впливати на здатність підприємства сплачувати податки, наприклад податок на землю. Прострочена заборгованість привертає особливу увагу, оскільки заборгованість може бути простим «здоровим» торговим кредитом. На жаль, на сьогодні немає статистики простроченої заборгованості між підприємствами, тому ми використовуємо кредиторську заборгованість між підприємствами як заміник для простроченої (*overdue arrears*) заборгованості. Згідно з принципом подвійного запису, існує тісний зв'язок між кредиторською та дебіторською заборгованостями в регіонах. Відмінність в основному стосується процедури обліку, яка вимагає, щоб дебіторська заборгованість записувалась в балансі підприємств як вартість факторів, у той час як кредиторська заборгованість відображається у кінцевих цінах, включаючи податки та інші збори та платежі [16]. Тому не має особливого значення, який фактор ми використовуємо в моделі — дебіторську чи кредиторську заборгованість.

На жаль, ми не можемо чітко передбачити вплив заборгованості між підприємствами на дохід, оскільки існує дві протилежні точки зору. З одного боку, заборгованість послаблює стійкість балансу, оскільки іммобілізує неліквідні активи, причому послаблюється здатність підприємства платити податки, з другого — українська економіка характеризується високим рівнем торгових кредитів і бартером [46, розділ 4]. Заборгованість, як результат такої діяльності, може вказувати не на неліквідність активів підприємств, а на високий рівень економічної активності, в результаті якої можна очікувати зростання податків, зібраних з даного підприємства. Таким чином, чистий ефект досить невизначений, хоча деякі дослідження [35] вказують на позитивну кореляцію між випуском продукції та заборгованістю. Ця проблема теж може вважатися специфічною для України.

Застосування моделей на основі лонгитюдних даних в цьому дослідженні має певні переваги, оскільки може пом'якшувати проблему невключених факторів. Ми можемо отримати консистентні параметри моделей на основі лонгитюдних даних, навіть якщо не маємо їх для неспостережуваних та невимірюваних змінних, таких, наприклад, як якість управління в регіонах та розподіл ресурсів. Припускаючи незмінність таких факторів, як земля, капітал, уподобання та управлінські якості, або будь-яких інших, ми можемо успішно вирішити проблему виключених (omitted) з моделі факторів, використовуючи фіксовані чи випадкові ефекти, тобто специфічні регіональні ефекти в моделях лонгитюдних даних.

Оскільки області варіюють за розмірами, ми використовуємо зважування на населення, щоб уникнути, або принаймні пом'якшити, неоднорідність в даних. Інфляція 1993—2000 рр. була досить значною, тому ми використовуємо показники в реальному вимірі, розраховані на основі дефлятора ВВП. Припускається, що рівень цін ідентичний для всіх областей. Оскільки фактично будь-яка функція може бути лінеаризована на певному проміжку, ми використовуємо лінійну функціональну форму регресій цих рівнянь. Крім того, ми специфікуємо модель на основі лонгитюдних даних з фіксованими ефектами через те, що вибірка вичерпує генеральну сукупність.

Крок 1. Формалізований вигляд моделі та позначення.

Беручи до уваги всі попередні міркування, специфікацію моделі лонгитюдних даних для моделювання податку на прибуток з громадян можна записати в такому формалізованому вигляді:

$$RPCPIT_{it} = \alpha_i + \beta_1 \cdot RPCIO_{it} + \beta_2 \cdot RWAGE_{it} + \beta_3 \cdot D(RPCWA_{it}) + \beta_4 \cdot D(RPCPAY_{it}) + \beta_5 \cdot D(UNEMPL_{it}),$$

де i — індекс областей; t — індекс проміжку часу; $RPCPIT$ — реальні надходження податку на прибуток з громадян на душу населення⁴ (млн грн, Держскарбниця); $RPCIO$ — реальний випуск промислової продукції на душу населення (млн грн, UEPLAC); $RWAGE$ — середня заробітна плата (грн, UEPLAC); $RPCWA$ — реальна заборгованість із зарплати в обрахунку на душу населення (млн грн, кінець періоду, UEPLAC); $RPCPAY$ — реальна кредиторська заборгованість між підприємствами на душу населення (млн грн, кінець періоду, UEPLAC); $UNEMPL$ — офіційний рівень зареєстрованого безробіття (% , UEPLAC); $D(\cdot)$ — оператор різниць.

⁴ Нагадаємо, що ми використовуємо дефлятор ВВП з базою: перший квартал 1997 р. = 1.000.

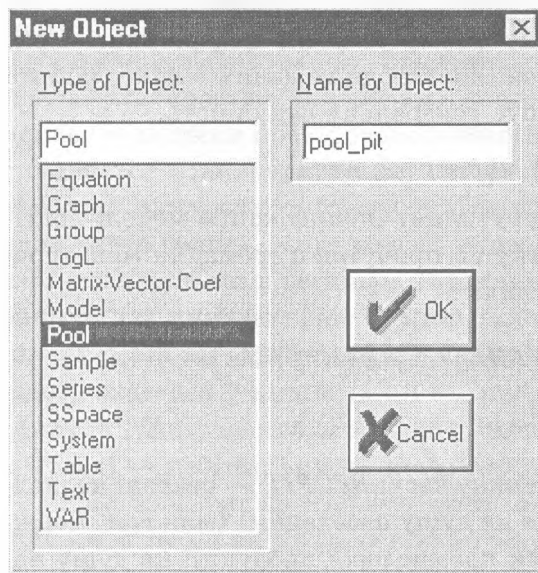
Крок 2. Оцінювання моделі лонгітюдних даних.

Крок 2.1. Введення даних для моделі лонгітюдних даних в пакеті E.VIEWS.

Щоб оцінити специфіковану вище модель для прогнозування податкових надходжень з прибутку громадян, нам необхідно перш за все ввести дані. Зауважимо, що ми використовуємо для розрахунків пакет E.VIEWS, тому прокоментуємо особливості введення даних саме для цього пакета. Лонгітюдні дані краще імпортувати, наприклад з Excel (див. E.VIEWS Help and Manual).

Крок 2.2. Створення спеціального пул-об'єкта (pool) для розрахунків за моделлю лонгітюдних даних.

Для проведення розрахунків за моделями лонгітюдних даних E.VIEWS вимагає створити спеціальний об'єкт: pool. Щоб створити такий об'єкт, виберіть **Objects/New Object**. У вікні, що відкрилось, виберіть **pool** і дайте ім'я цьому об'єкту. Натисніть **OK**.



Тепер потрібно задати ідентифікатори, які є ім'ям крос-секційних варіантів (у нашому випадку — регіонів). Ідентифікатори визначають, як відрізнити ряди для різних об'єктів. Переважно варіаційним рядам (тобто послідовності даних у часі для певного об'єкта в крос-секційному розрізі) присвоюються різні номери. Наприклад, якщо ми маємо справу з країнами, ми можемо позначити Україну цифрою «1», Німеччину цифрою «2», Францію цифрою «3» і т. д. Відповідно матриця часових рядів показників X включає вектор X_1 для України, вектор X_2 для

Німеччини і вектор X_3 для Франції. E.VIEWS також дозволяє використовувати слова (префікси [напр. 1_X, 2_X, 3_X], закінчення [напр. X_Ukr, X_Ger, X_Fra] і навіть слова-назви змінних [напр. real_Ukr_X, real_Ger_X, real_Fra_X]) як ідентифікатори. Взагалі, E.VIEWS забезпечує достатню гнучкість у визначенні ідентифікаторів.

Щоб визначити ідентифікатори, натисніть кнопку **Define** у меню об'єкта **pool**. У вікні, що відкрилось, введіть ідентифікатори. Позначимо в нашому випадку кожен регіон (region) номером від 1 до 25, відповідно список ідентифікаторів матиме такий вигляд:

CROSS SECTION IDENTIFIERS: (ENTER IDENTIFIERS BELOW THIS LINE)

_1
_2
_3
:
_25,

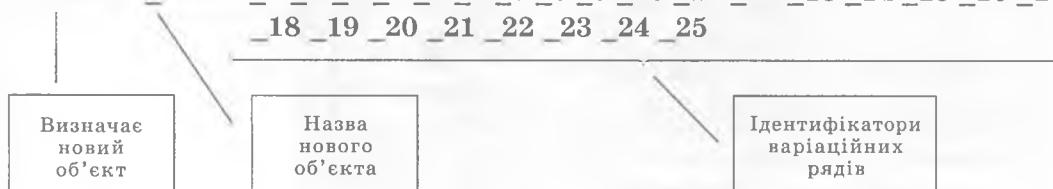
де кожен номер відповідає певній області (зауважимо, що ми виключаємо з розгляду міста Київ та Севастополь). Можна записати ідентифікатори і в рядок:

CROSS SECTION IDENTIFIERS: (ENTER IDENTIFIERS BELOW THIS LINE)

_1 _2 _3 _4 _5 _6 _7 _8 _9 _10 _11 _12 _13 _14 _15 _16 _17 _18 _19 _20 _21
_22 _23 _24 _25

Ми можемо також задати ідентифікатори, використовуючи команди E.VIEWS:

POOL PIT POOL _1 _2 _3 _4 _5 _6 _7 _8 _9 _10 _11 _12 _13 _14 _15 _16 _17
_18 _19 _20 _21 _22 _23 _24 _25



Крок 2.3. Оцінювання невідомих параметрів моделі лонгитюдних даних.

Для оцінювання моделі лонгитюдних даних натисніть Estimate в меню об'єкта pool.

Відкриється вікно, яке необхідно заповнити (див. наступну сторінку).

Зауважте, що ми використовуємо знак «?» для того, щоб визначити для пакета E.VIEWS, що ми використовуємо лонгитюдні дані. Під час розрахунків E.VIEWS замінь «?» на попередньо визначений ідентифікатор крос-секційного ряду. Наприклад, POP? означає, що POP_1 використовуватиметься для першої області (ідентифікатор _1), POP_2 — для другої (ідентифікатор _2) і т. д. Якщо ви хочете використовувати лаг для змінної X, то правильний запис виглядатиме так: X?(-1).

Ми також можемо зазначити, які з регресорів (факторів) матимуть однаковий перетин для всіх крос-секційних об'єктів, а які можуть мати відмінний. У нашому випадку ми накладаємо обмеження, що всі показники в крос-секційному розрізі мають однакові нахили.

Стосовно перетину в моделі ми також можемо вибрати різні опції: опцію (None) — модель без перетину, задати однаковий перетин в моделі в розрізі крос-секційних та часових рядів (Common) або задати різні перетини в розрізі крос-секційних даних, задавши, крім того, додатково, як їх розглядати — як константи (модель з фіксованими ефектами) чи як випадкові величини

Pooled Estimation

Dependent Variable: Sample: Balanced Sample

Regressors and AR() terms:

Common coefficients:

Cross section specific coefficients:

Intercept: None Common Fixed effects Random effects

Weighting: No weighting Cross section weights SUR Iterate to convergence

Random effects uses Variance Components model. Weights, AR() and CS specific coefs not allowed.

Buttons:

Callout 1: Регресори з однаковим нахилом для всіх об'єктів у крос-секційному розрізі

Callout 2: Регресори з відмінним нахилом для всіх об'єктів у крос-секційному розрізі

Callout 3: Специфікує використання фіксованих ефектів (тобто дозволяються відмінні перетини для кожного об'єкта у крос-секційному розрізі)

(модель з випадковими ефектами) (відповідно **Fixed effects** або **Random effects**).

Estimation Options

Iteration Options: Max Iterations: Convergence:

White Heteroskedasticity Consistent Covariance
Not available with SUR or RE

Buttons:

Іноколи нам потрібно збалансувати модель. Модель вважається незбалансованою, якщо деякі крос-секційні ряди не мають значень (спостережень) для певних періодів часу. Коли ми задаємо таку опцію (**Balanced sample**), всі крос-секційні змінні з неповними рядами виключаються з розгляду.

В опціях ми можемо визначити максимальне число ітерацій (**Max Iterations**), точність розрахунків (**Convergence**) та застосовувати корекції у випадку наявності гетероскедастичності: **White Heteroscedasticity Consistent Covariance**⁵.

⁵ Як і у випадку простого МНК, це корегування є бажаним при гетероскедастичності залишків.

Після визначення всіх необхідних опцій нажміть ОК та отримайте результат розрахунків (таблиця 5.1).

Таблиця 5.1. Результати розрахунків за моделлю надходжень податку з прибутку громадян

Dependent Variable: PIT?/POP?/ GDPDEF
 Method: Pooled Least Squares
 Sample(adjusted): 1998:1 2000:2
 Included observations: 10 after adjusting endpoints
 Total panel (balanced) observations 250

Кількість об'єктів у крос-секційному розрізі, помножена на кількість спостережень у крос-секції

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AWAGE?/GDPDEF	0.068148	0.007275	9.367537	0.0000
INDOUT?/POP?/GDPDEF*1000	0.013781	0.002816	4.894078	0.0000
D(WARR?/POP?/GDPDEF)*1000	-0.047309	0.007953	-5.948457	0.0000
D(PAY?/POP?/GDPDEF)*1000	0.000777	0.000471	1.648966	0.1006
D(UNEMPL?)	-0.693005	0.304201	-2.278115	0.0237
Fixed Effects				
_1—C	0.545572			
_2—C	0.440655			
_3—C	0.742452			
_4—C	-1.698630			
_5—C	-1.497778			
_6—C	0.213036			
_7—C	1.651691			
_8—C	-1.177737			
_9—C	-0.809000			
_10—C	-1.097848			
_11—C	0.426006			
_12—C	-1.669807			
_13—C	1.744995			
_14—C	0.772140			
_15—C	5.099823			
_16—C	-0.449834			
_17—C	-1.477738			
_18—C	-0.043418			
_19—C	-0.146904			
_20—C	1.368714			
_21—C	1.361789			
_22—C	0.078448			
_23—C	0.611447			
_24—C	1.046667			
_25—C	0.607629			
R-squared	0.917354	Mean dependent var	11.67577	
Adjusted R-squared	0.906459	S.D. dependent var	3.974097	
S.E. of regression	1.215454	Sum squared resid	325.0122	
F-statistic	610.4866	Durbin-Watson stat	2.289609	
Prob(F-statistic)	0.000000			

Крос-секційні (регіон) специфічні перетини

Відповідно до нашої специфікації:
 $RSPIT = PIT / POP / GDPDEF$

```

RPCIO=INDOUT/POP/GDPDEF
RWAGE=AWAGE/POP
RPCWA=WARR/POP/GDPDEF
RPCPAY=PAY/POP/GDPDEF.

```

Той же результат можна отримати, застосовуючи команди E.VIEWS:

```

PIT_POOL.LS(F) PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/
GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/GDPDEF)*1000
D(UNEMPL?)

```

Крім моделей з фіксованими та випадковими ефектами, E.VIEWS пропонує цікаву комбінацію фіксованих і випадкових ефектів: крос-секційне зважування (у вікні Pool опція **Cross Section Weights**). Ця опція досить корисна, коли наявна гетероскедастичність у крос-секційному розрізі. E.VIEWS застосовує здійснений метод узагальнених найменших квадратів (feasible GLS) для оцінювання моделі. Технічно ми припускаємо, що залишки гетероскедастичні, але не корелюють між собою. Іншими словами, матриця вагів Ω має позитивні значення по діагоналі, а всі інші значення такої матриці дорівнюють нулю, тобто вона має вигляд:

$$\Omega = \begin{bmatrix} \sigma_{11}I & 0 & 0 & \dots & 0 \\ 0 & \sigma_{22}I & 0 & \dots & 0 \\ 0 & 0 & \sigma_{33}I & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \sigma_{nn}I \end{bmatrix}$$

Як і метод оцінювання моделі з випадковими ефектами, даний метод оцінювання також двокроковий. На першому етапі E.VIEWS обраховує помилки (залишки) на основі незміщеного методу найменших квадратів. На другому — використовує оцінені залишки для розрахунку Ω та наступного оцінювання здійсненим методом узагальнених найменших квадратів.

Крок 2.4. Тест Дарбіна—Вю—Хаусмана (Durbin—Wu—Hausman) для моделей на основі лонгітюдних даних.

Оцінена модель була специфікована нами як модель з фіксованими ефектами, невідомі параметри якої розраховувались методом найменших квадратів для внутрішньогрупової варіації, який не бере до уваги кореляцію помилок у крос-секційному розрізі (across cross sections). Ефективнішим методом оцінювання є метод, що використовується для оцінки параметрів моделі з випадковими ефектами. Але ми знаємо з економетричної теорії, що оцінювання в разі випадкових ефектів дає зміщені результати, коли перетини, специфічні для крос-секційних об'єктів (individual cross section specific inter-

sept), корелюють з випадковою величиною. В той же час оцінювання у випадку фіксованих ефектів (метод найменших квадратів для внутрішньогрупової варіації) є невразливим до такої проблеми, оскільки коваріація між константою та випадковою величиною дорівнює нулю. Таким чином, оцінювання параметрів моделі у випадку фіксованих ефектів дає консистентні, але не-ефективні оцінки. Для того, щоб вирішити, яку з моделей краще застосовувати — з фіксованими чи з випадковими ефектами, використовується тест Дарбіна—Вю—Хаусмана, який перевіряє, чи дає оцінювання в разі випадкових ефектів незміщені оцінки, чи ні. Формально тест виконується так: нульова гіпотеза: $H_0: \hat{\beta}^{FE} = \hat{\beta}^{RE}$, де $\hat{\beta}^{FE}$ — параметри моделі, оцінені на основі моделі з фіксованими ефектами, $\hat{\beta}^{RE}$ — параметри моделі, оцінені на основі моделі з випадковими ефектами. Відповідно альтернативна гіпотеза: $H_0: \hat{\beta}^{FE} \neq \hat{\beta}^{RE}$.

Далі можна показати, що статистика $(\hat{\beta}^{FE} - \hat{\beta}^{RE})'(V(\hat{\beta}^{FE}) - V(\hat{\beta}^{RE}))^{-1} \times (\hat{\beta}^{FE} - \hat{\beta}^{RE})$ розподілена як χ^2 зі ступенем свободи, що дорівнює кількості параметрів у векторі β (в нашому випадку 5). $V(\hat{\beta}^{FE})$ — оцінена дисперсійно-коваріаційна матриця для параметрів моделі з фіксованими ефектами, $V(\hat{\beta}^{RE})$ — оцінена дисперсійно-коваріаційна матриця для параметрів моделі з випадковими ефектами. Для прийняття рішень застосовується таке правило:

- якщо розрахована величина $(\hat{\beta}^{FE} - \hat{\beta}^{RE})'(V(\hat{\beta}^{FE}) - V(\hat{\beta}^{RE}))^{-1}(\hat{\beta}^{FE} - \hat{\beta}^{RE})$ більша за критичне значення χ^2 -статистики з відповідним ступенем свободи (в нашому випадку 5) для заданого рівня значимості (наприклад 5%), ми відкидаємо нульову гіпотезу (тобто статистично оцінки $\hat{\beta}^{FE}$, $\hat{\beta}^{RE}$ відмінні);
- якщо розрахована величина $(\hat{\beta}^{FE} - \hat{\beta}^{RE})'(V(\hat{\beta}^{FE}) - V(\hat{\beta}^{RE}))^{-1}(\hat{\beta}^{FE} - \hat{\beta}^{RE})$ менша за критичне значення χ^2 -статистики з відповідним ступенем свободи для заданого рівня значимості, ми не відкидаємо нульову гіпотезу (тобто статистично оцінки $\hat{\beta}^{FE}$, $\hat{\beta}^{RE}$ однакові).

Цей тест можна використовувати для перевірки зміщеності оцінок для значно ширшого класу методів, наприклад в контексті порівняння оцінок МНК та 2МНК (дивись практичне завдання в розділі симультаивних рівнянь).

Хоча даний тест може бути зроблений вручну, ми пропонуємо застосувати програму, що виконує тест Дарбіна—Вю—Хаусмана в командному режимі.

**Тест Дарбіна—Вю—Хаусмана в командах E.VIEWS
(на прикладі вищеоціненої моделі)**

‘визначте ідентифікатори

POOL PIT_POOL _1 _2 _3 _4 _5 _6 _7 _8 _9 _10 _11 _12 _13 _14 _15 _16 _17
_18 _19 _20 _21 _22 _23 _24 _25

‘оцініть модель з фіксованими ефектами

PIT_POOL.LS(F) PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/ GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/GDPDEF)*1000 D(UNEMPL?)

Ця опція вказує на використання фіксованих ефектів

‘обчисліть варіаційно-коваріаційну матрицю для оцінених коефіцієнтів моделі.
MATRIX_COVVAR_FE=PIT_POOL.@COEFCOV

‘обчисліть коефіцієнти моделі.
MATRIX_COEF_FE=PIT_POOL.@COEFS

‘оцініть модель з випадковими ефектами тієї самої специфікації
PIT_POOL.LS(R) PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/ GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/GDPDEF)*1000 D(UNEMPL?)

Ця опція вказує на використання випадкових ефектів

‘обчисліть варіаційно-коваріаційну матрицю для оцінених коефіцієнтів моделі.
MATRIX_COVVAR_RE1=PIT_POOL.@COEFCOV

‘обчисліть коефіцієнти моделі.
MATRIX_COEF_RE1=PIT_POOL.@COEFS

‘знайдіть варіаційно-коваріаційну матрицю для оцінених коефіцієнтів моделі (тільки для нахилів, оскільки в моделях з фіксованими ефектами загальний перетин відсутній)

MATRIX_COVVAR_RE=@SUBEXTRACT(_COVVAR_RE1,2,2)

‘знайдіть матрицю коефіцієнтів для оціненої моделі (тільки для нахилів, оскільки в моделях з фіксованими ефектами загальний перетин відсутній)

MATRIX_COEF_RE=@SUBEXTRACT(_COEF_RE1,2)

‘розрахуйте обернену матрицю для різниці між варіаційно-коваріаційною матрицею моделі з фіксованими ефектами і варіаційно-коваріаційною матрицею моделі з випадковими ефектами

MATRIX_INV_COV=@INVERSE(_COVVAR_FE-_COVVAR_RE)

‘розрахуйте матрицю різниці між варіаційно-коваріаційною матрицею моделі з фіксованими ефектами і варіаційно-коваріаційною матрицею моделі з випадковими ефектами

MATRIX_COEF_DIF=_COEF_FE-_COEF_RE

‘розрахуйте χ^2 -статистику тесту

MATRIX_CHI_STAT=@TRANSPPOSE(_COEF_DIF)*_INV_COV*_COEF_DIF

‘розрахуйте p -value даної статистики

MATRIX_RESULT=@CHISQ(_CHI_STAT(1,1),PIT_POOL.@NCOEF-1)

У нашому випадку p -value дорівнює 0.177, тобто ми не можемо відхилити нульову гіпотезу на 10 %-му рівні значимості. Отже, модель з випадковими ефектами для нашого прикладу дає незміщені оцінки. Таким чином, можемо також застосовувати модель з випадковими ефектами без побоювання отримати незміщені оцінки.

Таблиця 5.2. Результати розрахунків за моделлю на основі лонгитюдних даних з випадковим ефектом

Dependent Variable: PIT?/POP?/GDPDEF
 Method: GLS (Variable Components)
 Sample(adjusted): 1998:1 2000:2
 Included observations: 10
 Total panel (balanced) observations 250

Загальний метод найменших квадратів показує, що ми оцінюємо модель з випадковими ефектами

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.890090	0.706410	1.260020	0.2089
AWAGE?/GDPDEF	0.071497	0.006690	10.68747	0.0000
INDOUT?/POP?/GDPDEF*1000	0.010031	0.001130	8.877898	0.0000
D(WARR?/POP?/GDPDEF)*1000	-0.050676	0.007871	-6.438216	0.0000
D(PAY?/POP?/GDPDEF)*1000	0.000722	0.000479	1.509758	0.1324
D(UNEMPL?)	-0.799641	0.302803	-2.640796	0.0088
Random Effects				
_1—C	-0.283010			
_2—C	-0.101682			
_3—C	-0.076021			
_4—C	-0.221999			
_5—C	-0.330988			
_6—C	-0.409508			
_7—C	0.591406			
_8—C	0.204891			
_9—C	-1.116390			
_10—C	-1.224455			
_11—C	-0.303385			
_12—C	-1.252991			
_13—C	0.996950			
_14—C	0.446009			
_15—C	3.775200			
_16—C	0.171740			
_17—C	-1.566540			
_18—C	-0.069675			
_19—C	-0.862061			
_20—C	1.173899			
_21—C	0.433607			
_22—C	-0.378017			
_23—C	0.157831			
_24—C	0.149506			
_25—C	0.095683			

GLS Transformed Regression

R-squared	0.905002	Mean dependent var	11.67577
Adjusted R-squared	0.903055	S.D. dependent var	3.974097
S.E. of regression	1.237372	Sum squared resid	373.5861
Durbin-Watson stat	2.013970		

Unweighted Statistics including
Random Effects

R-squared	0.915015	Mean dependent var	11.67577
Adjusted R-squared	0.913273	S.D. dependent var	3.974097
S.E. of regression	1.170347	Sum squared resid	334.2100
Durbin-Watson stat	2.251253		

Крок 2.5. Тест на перевірку загального перетину.

Розглянемо інший важливий тест. Зауважимо, що використання крос-секційних ефектів (індивідуальних перетинів для кожного крос-секційного об'єкта) призводить до значного зменшення ступенів свободи, тому вводити їх в моделі лонгitudних даних слід лише в разі явної необхідності. Якщо її немає, то краще використовувати загальний перетин в моделі. Яким чином це вирішити? Для перевірки, чи є дійсно необхідним використання крос-секційних специфічних перетинів (ефектів cross section specific intercept), можна застосувати стандартний F -тест (див. розділ 5.1) або тест Волда (Wald). В обох випадках нульова гіпотеза є такою: немає статистичної різниці між використанням крос-секційно специфічних перетинів⁶ та використанням загального перетину. Розглянемо застосування кожного тесту окремо.

Тест Волда на перевірку загального (common) перетину в моделях на основі лонгitudних даних

Опишемо поетапно алгоритм застосування тесту Волда в E.VIEWS:

- оцініть модель звичайним МНК, обравши опцію «без перетину» (none intercept), та визначте константу C як змінну з крос-секційним специфічним перетином.

Після цього виберіть **View\Wald Coefficient tests...** в меню об'єкта pool. Потім у вікні, що відкривається, задайте обмеження (наприклад, всі коефіцієнти однакові).

Для нашого прикладу отримаємо такі результати:

Wald Test:
Equation: PIT_POOL

Null Hypothesis:

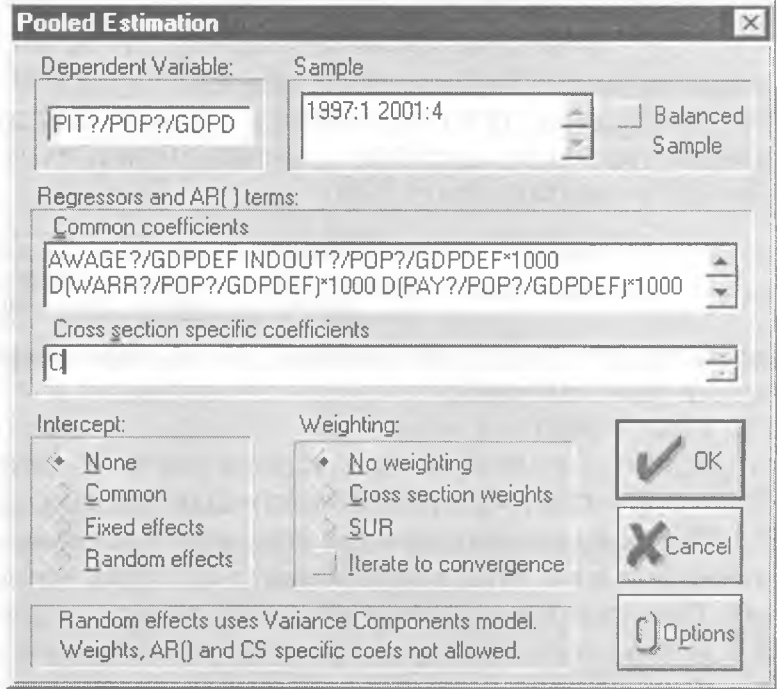
$$C(7)=C(30)$$

$$C(8)=C(30)$$

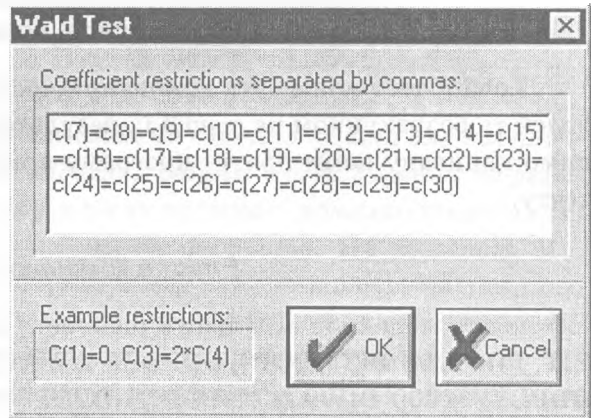
$$C(9)=C(30)$$

⁶ Тобто окремий перетин для кожного об'єкта в крос-секційному розрізі.

- C(10)=C(30)
- C(11)=C(30)
- C(12)=C(30)
- C(13)=C(30)
- C(14)=C(30)
- C(15)=C(30)
- C(16)=C(30)
- C(17)=C(30)
- C(18)=C(30)
- C(19)=C(30)
- C(20)=C(30)
- C(21)=C(30)
- C(22)=C(30)
- C(23)=C(30)
- C(24)=C(30)
- C(25)=C(30)
- C(26)=C(30)
- C(27)=C(30)
- C(28)=C(30)
- C(29)=C(30)



F=statistic 9.1325 Probability 0.00
 Chi-square 210.04 Probability 0.00



Згідно з отриманим значенням *F*-статистики, ми відкидаємо нульову гіпотезу (*p*-значення (*p*-value)⁷ *F*-статистики дорівнює 0.0000) про те, що модель із загальним перетином гірша, ніж модель з фіксованими ефектами, в тому розумінні, що накладені обмеження погіршують точність моделі.

Той самий результат можна отримати, використовуючи команди E.VIEWS.

⁷ Нагадаємо, що *p*-значення показує імовірність нульової гіпотези.

Тест Волда в командах E.VIEWS

‘обчисліть модель з фіксованими ефектами

```
PIT_POOL.LS(N) PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/
GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/
GDPDEF)*1000 D(UNEMPL?) @ C
```

Немає
перетину

Після символу @ задається список змінних із
крос-секційним нахилом для того, щоб обчислити
стандартні помилки крос-секційних ефектів

‘проведіть Wald-тест з обмеженням, що всі крос-секційні специфічні ефекти рівні між собою

```
PIT_POOL.WALD
```

```
C(7)=C(8)=C(9)=C(10)=C(11)=C(12)=C(13)=C(14)=C(15)=C(16)=C(17)=C(18)=C(19)=
C(20)=C(21)=C(22)=C(23)=C(24)=C(25)=C(26)=C(27)=C(28)=C(29)=C(30)
```

Тест Волда пропонується для лонгітюдних даних з малою кількістю об’єктів (варіантів або груп) в крос-секційному розрізі, оскільки він досить простий. Проте у випадку великої кількості варіантів та при використанні моделі з випадковими ефектами цей тест не може бути застосований.

Іншим, більш потужним тестом, особливо при наявності великої кількості варіантів, є *F*-тест, який детально описано в розділі 5.1.

***F*-тест на перевірку загального (common) перетину в моделях на основі лонгітюдних даних**

Хоча *F*-тест може бути виконаний послідовною оцінкою лонгітюдних моделей та розрахунком на основі їх результатів *F*-статистики вручну, ми наводимо написану нами просту програму в командах E.VIEWS для проведення тесту.

***F*-тест в командах E.VIEWS**

‘обчисліть модель із загальним (common) перетином

```
PIT_POOL.LS PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/
GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/
GDPDEF)*1000 D(UNEMPL?)
```

‘отримайте *R*² (коефіцієнт детермінації) для моделі із загальним (common) перетином

```
SCALAR _R2_C=PIT_POOL.@R2
```

‘обчисліть модель з крос-секційно специфічними перетинами (тут ми використовуємо фіксовані ефекти)

```
PIT_POOL.LS(F) PIT?/POP?/GDPDEF AWAGE?/GDPDEF INDOUT?/POP?/
```

GDPDEF*1000 D(WARR?/POP?/GDPDEF)*1000 D(PAY?/POP?/
GDPDEF)*1000 D(UNEMPL?)

‘отримайте R^2 (коефіцієнт детермінації) для моделі з крос-секційними специфічними перетинами

SCALAR_R2_F=PIT_POOL.@R2

‘обчисліть F -статистику, 24 є кількістю ступенів свободи для чисельника (на скільки збільшується кількість параметрів, які необхідно оцінити), команда: pit_pool.@totalobs - pit_pool.@ncoef дає ступені свободи для знаменника (кількість параметрів у розширеній моделі (з крос-секційно специфічними ефектами)

MATRIX_F_STAT=(R2_F-R2_C)/24/((1-R2_F)/(PIT_POOL.@TOTALOBS -
PIT_POOL.@NCOEF))

‘обчисліть p -value F -статистики:

MATRIX_RESULT_F=@FDIST(F_STAT,24,PIT_POOL.@TOTALOBS -
PIT_POOL.@NCOEF)

Результат цих команд: значення F -статистики дорівнює 9.004, а p -value 0.000. Результати досить подібні до тих, що ми отримали у випадку застосування тесту Волда.

Крок 3. Загальні висновки та коментарі.

Зробіть їх самостійно.

Завдання 5.2.

Тепер дослідимо, як лонгитюдні дані можна застосовувати для вирішення проблеми прогнозування витрат місцевих органів влади України на регіональному рівні. Розробимо відповідну модель на основі лонгитюдних даних, щоб вивчити залежність витрат від низки факторів. На прикладі цієї моделі розглянемо проблему ендогенності в моделях на основі лонгитюдних даних та особливості її подолання. Крім того, розглянемо особливості використання інструментальних змінних для моделей лонгитюдних даних.

Розіб'ємо наше завдання на декілька кроків. Детально опишемо та прокоментуємо кожний з них. Зауважимо, що в процесі виконання цього завдання ми використовуємо пакет прикладних програм E.VIEWS.

Проблема ендогенності та застосування двокрокового методу найменших квадратів (2SLS) в моделях лонгитюдних даних

Опишіть, в чому полягає сутність проблеми ендогенності в моделях на основі лонгитюдних даних. Чому наявність ендогенності є проблемою для оцінювання параметрів моделі? В чому полягають переваги застосування двокрокового методу найменших квадратів у випадку наявності ендогенності?

Крок 1. Постановка задачі моделювання витрат місцевих органів влади та специфікація моделі лонгitudних даних.

Детально опишіть постановку та особливості моделювання витрат місцевих органів влади на регіональному рівні. Прокоментуйте перевагу від застосування моделей на основі лонгitudних даних. Обґрунтуйте, який тип моделі — з фіксованими чи з випадковими ефектами — краще обрати для розв'язання поставленого завдання.

Крок 2. Оцінювання моделі лонгitudних даних.

Крок 2.1. Введення даних для розробки моделей лонгitudних даних в пакеті E.VIEWS.

Введіть лонгitudні дані для подальшого використання їх для моделювання. Поясніть особливості введення лонгitudних даних при застосуванні пакета E.VIEWS.

Крок 2.2. Створення спеціального пул-об'єкта (pool) для розрахунків за моделлю лонгitudних даних.

Прокоментуйте, що являє собою об'єкт pool в моделюванні на основі лонгitudних даних. Як ми його створюємо в пакеті E.VIEWS?

Крок 2.3. Оцінювання невідомих параметрів моделі лонгitudних даних.

Проведіть оцінювання специфікованої вами моделі витрат місцевих органів влади. Детально прокоментуйте отримані результати.

Крок 3. Проблема ендогенності дотацій в моделі витрат місцевих органів влади та її тестування.

Детально поясніть, чому в моделях витрат місцевих органів влади дотації, що часто включаються в модель як регресори (пояснюючі змінні), можуть бути ендогенними відносно витрат. Обґрунтуйте ваші пояснення. Як формально можна протестувати дане твердження?

Крок 4. Перевірка ендогенності в моделях лонгitudних даних за допомогою тесту Хаусмана. Поняття інструментальних змінних.

Проведіть тест Хаусмана для перевірки ендогенності дотацій. Які інструментальні змінні ви використовуєте? Поясніть чому, обґрунтуйте їх використання. Які висновки ви можете зробити за результатом тесту Хаусмана. Поясніть їх.

Зауважимо, що тест Хаусмана на перевірку ендогенності не автоматизований в пакеті E.VIEWS, він повинен проводитись або покроково, або в ко-

мандному режимі, тому ми наводимо написану нами програму для застосування тесту Хаусмана в пакеті E.VIEWS.

Крок 5. Тест на перевірку правильності обраних інструментів, або тест переототожнених обмежень.

Для того, щоб бути впевненим, що інструменти для ендогенних змінних (в нашому випадку для дотацій) обрані коректно, можна застосувати спеціальний тест, так званий тест переототожнених обмежень (*overidentifying restrictions test*). Цей тест, нагадаємо, перевіряє, чи існує кореляція між інструментами та залишками. Застосуйте цей тест для перевірки коректності обраних інструментів для дотацій. Прокоментуйте отримані результати. Які остаточні висновки ви можете зробити?

Зауважимо, що тесту на перевірку якості обраних інструментів (*overidentifying restrictions test*) немає в пакеті E.VIEWS, він має проводитись або покроково, або в командному режимі, тому ми наводимо також написану нами програму для застосування тесту переототожнених обмежень в пакеті E.VIEWS.

Крок 6. Загальні висновки та коментарі.

Зробіть загальні висновки щодо усунення проблем ендогенності в моделях на основі лонгитюдних даних, зокрема в моделі витрат місцевих органів влади. Зробіть загальні висновки щодо отриманих результатів. Обґрунтуйте їх.

Розв'язок завдання 5.2 та розширений коментар моделювання витрат місцевих органів влади на регіональному рівні

Моделювання витрат місцевих органів влади України на регіональному рівні ми проводимо за допомогою економетричного пакета E.VIEWS. Для побудови моделі використовувались річні лонгитюдні (*panel*) фіскальні дані Держскарбниці в розрізі областей, а саме в розрізі 25 адміністративно-територіальних одиниць (тобто 24 області та Автономна республіка Крим) за 1995—1999 рр. За аналогічний період та в аналогічному розрізі використовувались також дані для інших економічних показників. Джерела цих даних: Державний комітет статистики, Міністерство економіки та UEPLAC/TACIS. На жаль, ми не можемо перевірити точність та надійність даних, хоча фіскальні дані вимірюються доволі точно, оскільки є документами офіційної звітності перед Міністерством фінансів, Держскарбницею, Контрольно-ревізійним управлінням, Рахунковою палатою Верховної Ради та місцевими радами. Економічні дані, очевидно, менш надійні. (Дані наведені додатку 5.)

Проблема ендогенності та застосування двокрокового методу найменших квадратів в моделях на основі лонгітюдних даних

У багатьох економетричних прикладних дослідженнях ми маємо справу з ендогенністю змінних, що включаються як фактори в моделі. В такому випадку ми вже не можемо застосувати звичайний метод найменших квадратів (МНК) для оцінки невідомих параметрів моделі, бо він дає зміщені та неконсистентні оцінки. Одним з виходів з такої ситуації є застосування двокрокового методу найменших квадратів. Основна ідея цього методу вже розглядалась нами в попередньому розділі у випадку симультивних систем рівнянь. Нагадаємо її. З самої назви очевидно, що цей метод розбивається на два етапи.

На першому етапі ми оцінюємо допоміжне регресійне рівняння залежності ендогенної змінної (що виявилась в правій частині початкового рівняння) від інструментальних змінних та інших екзогенних змінних первісного рівняння. Потім на основі даного економетричного рівняння знаходимо розраховані (fitted) значення ендогенної змінної. Ці розраховані значення не містять в собі ендогенності, яка наявна у вихідних даних для досліджуваної ендогенної змінної. Іншими словами, ми очищаємо досліджувану змінну від ендогенності. На другому етапі ми оцінюємо початкове рівняння, замінивши ендогенний регресор на його розраховані значення, отримані з допоміжного рівняння, оціненого на першому етапі.

Проте може виявитися, що ендогенність/симультивність незначна, і ми можемо застосувати набагато простіші методи, щоб обчислити параметри моделей. Можна застосувати вже згадуваний нами в попередньому розділі тест Хаусмана, щоб перевірити, чи дійсно ендогенність/симультивність є проблемою для оцінки рівняння.

Для того, щоб провести цей тест, нам необхідно оцінити первісне рівняння, замінивши не тільки ендогенну змінну правої частини на саму себе, «очищену від ендогенності» (тобто на розраховані значення цієї змінної (див. Перший крок двокрокового МНК)), а й увівши нову змінну, а саме залишки з останнього оціненого на другому кроці рівняння. Якщо коефіцієнт при змінній залишків статистично значимо відрізняється від нуля, ендогенність є дійсно проблемою в нашому випадку і ми повинні застосовувати двокроковий метод найменших квадратів для оцінки моделі. В протилежному випадку (коефіцієнт статистично незначимо відрізняється від нуля) ендогенність не є значною проблемою, і ми можемо використовувати МНК для оцінки параметрів моделі. Для більш детальної інформації див. [28, с. 303—304].

Продемонструємо особливості застосування методу двокрокових найменших квадратів та тестування можливої наявності ендогенності в регресорах для моделей на основі лонгітюдних даних.

Розглянемо приклад оцінювання функції витрат місцевого уряду для різних регіонів України.

Крок 1. Постановка задачі моделювання витрат місцевих органів влади та специфікація моделі лонгитюдних даних.

Оскільки місцеві громадські товари (і послуги) є нормальними товарами, ми можемо очікувати, що градієнт доходу позитивний. Тим не менше ми повинні пам'ятати про автоматичні (вбудовані) стабілізатори, які обертають витрати в зворотному напрямку від доходу. Місцеві уряди не впливають прямо на макроекономічну стабілізацію, тому можна передбачити додатні знаки у змінних, що є апроксимацією показників доходу на місцевому рівні (замінники доходу).

Очевидно, що одним з факторів витрат місцевого уряду є дотації національного уряду місцевому. Коефіцієнт цієї змінної цікавить нас особливо, оскільки вказує на існування і важливість «ефекту липучки» (the flypaper effect) [8, 33]. Априорі дотації стимулюють витрати.

Вищенаведені міркування дозволяють специфікувати рівняння витрат місцевих урядів у такому загальному вигляді:

$$RPCEXP_{it} = \alpha_i + \beta_1 \cdot RPCIO_{it} + \beta_2 \cdot RPCAGRO_{it} + \beta_3 \cdot RWAGE_{it} + \beta_4 \cdot RPCGRANT_{it} + \beta_5 \cdot D(RPCWARR_{it}) + \beta_6 \cdot D(RPCPAY_{it}), \quad (5.70)$$

де i — індекс області; t — індекс для часового періоду; $RPCEXP$ — реальні витрати місцевого уряду на душу населення; $RPCIO$ — реальний випуск промислової продукції на душу населення; $RPCAGRO$ — реальний випуск сільськогосподарської продукції на душу населення; $RWAGE$ — середня заробітна плата; $RPCGRANT$ — реальні дотації області на душу населення; $RPCWARR$ — реальна заборгованість із заробітної плати на душу населення; $RPCPAY$ — реальні платежі між підприємствами на душу населення.

Крок 2. Оцінювання моделі лонгитюдних даних.

Крок 2.1. Введення даних для розробки моделей лонгитюдних даних в пакеті E.VIEWS (аналогічний кроку 2.1 для завдання 5.1).

Крок 2.2. Створення спеціального пул-об'єкта (pool) для розрахунків за моделлю лонгитюдних даних (аналогічний кроку 2.2 для завдання 5.1).

Крок 2.3. Оцінювання невідомих параметрів моделі лонгитюдних даних.

Процедура детально описана в завданні 5.1 (крок 2.3), тож ми повторювати її не будемо. Наведемо лише результати розрахунків, тобто результати

оцінки параметрів даної моделі (функції витрат місцевих урядів) за допомогою МНК.

Таблиця 5.3. Результати оцінки параметрів моделі витрат місцевих урядів

Dependent Variable: RPCEXP?

Method: Pooled Least Squares

Sample(adjusted): 1995 1999

Included observations: 5 after adjusting endpoints

Total panel (balanced) observations 125

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RPCAGRO?	33.30678	21.31255	1.562778	0.1215
RPCINDOUT?	32.65855	24.81946	1.315844	0.1914
RWAGE?	0.749691	0.353064	2.123385	0.0363
RPCWARRF?	-1376.437	278.9595	-4.934180	0.0000
RPCPAYF?	10.27353	17.04259	0.602815	0.5481
RPCGRANT?	0.927084	0.226467	4.093682	0.0001
Fixed Effects				
1—C	293.2535			
⋮				
25—C	210.5423			
R-squared	0.710502	Mean dependent var	457.5042	
Adjusted R-squared	0.618109	S.D. dependent var	106.1124	
S.E. of regression	65.57463	Sum squared resid	404203.1	
F-statistic	46.14004	Durbin-Watson stat	1.927785	
Prob(F-statistic)	0.000000			

Крок 3. Проблема ендогенності дотацій в моделі витрат місцевих органів влади та її тестування.

Як свідчать певні дослідження, дотації є ендогенними по відношенню до рішень місцевих урядів. Наприклад, в [1] розглядаються фактичні схеми дотацій в розвинутих країнах та в країнах, що розвиваються. На основі даного дослідження можна зробити висновок, що в багатьох випадках обсяги дотацій виявляються ендогенними відносно місцевих урядів. Отже, місцеві уряди повинні бути зацікавлені в тому, щоб отримувати дешеві ресурси від органів державної влади та здатні їх здобувати. Гіпотеза про ендогенність дотацій була перевірена в [17] на лонгітюдних даних муніципальної влади Онтаріо, Канада. Автори застосовували формальні статистичні тести для перевірки екзогенності в традиційних моделях «липучки» (flypaper models). Вони дійшли висновку, що всі тести на ендогенність неспроможні відхилити гіпотезу про те, що дотації визначаються симультативно, тобто є ендогенними. Вони також показали, що у випадку, який розглядався, дотації не стимулюють зростання власного доходу місцевих органів влади, а замінюють його. Іншими словами, «ефекту липучки» не спостерігається. В роботі [19] також

показано, як екзогенні (практично за припущенням) дотації можуть бути ендогенними⁸. Щоб розширити теорію ендогенної природи дотацій, звернемося також до роботи [30]. Її автори вважають, що фактори, які впливають на прийняття рішень місцевих урядів, впливають і на міжурядову допомогу. У їхній праці показано, що дотації місцевим урядам ендогенні до тої міри, до якої вони корисні. Коротко суть даної теорії полягає в тому, що слід остерегатися можливої ендогенності показника дотацій.

Українська система міжурядових фіскальних зв'язків характеризується дотаціями, які даються на покриття дисбалансів між доходами та витратами (gap filling grants). Це наслідок старого принципу складання планів на основі досягнутих результатів (the ratchet effect [29, с. 205—213]). Якщо результати (наприклад, доходи від оподаткування) в області високі, плановані надходження в наступному році підвищуються, а заплановані для наступного фіскального року дотації відповідно зменшуються. І навпаки, якщо надходження в поточному році є низькими, то планові доходи для наступного року будуть знижені, а дотації зростуть. Таким чином, ми спостерігаємо те, що називається «gap-filling» у визначенні грантів. Такий ефект храповика⁹ призводить до невизначеності щодо величини доходів та до зменшення стимулів для розвитку місцевої бази оподаткування. В результаті це заважає розвитку територіальних громад і призводить до неефективного розподілу ресурсів. Деякі економетричні підтвердження цього ефекту можна знайти на прикладі Росії [32]. В результаті ми можемо очікувати, що дотації ендогенні. Для того, щоб перевірити цей факт, застосуємо тест Хаусмана для нашого прикладу.

Крок 4. Перевірка ендогенності в моделях лонгитюдних даних за допомогою тесту Хаусмана.

Поняття інструментальних змінних.

Щоб застосувати тест Хаусмана для перевірки, чи дійсно змінна дотацій є ендогенною, ми повинні спочатку знайти для неї інструменти: змінні, що корелюють з дотаціями, але не корелюють із залишками. Як такі інструменти ми пропонуємо деякі демографічні фактори, причому, щоб врахувати особливості змін демографічних змінних, ми пропонуємо використати показники частки пенсіонерів і дітей у загальному населенні. Більша частка таких категорій населення, як правило, вказує на більші потреби у фінансуванні

⁸ Б. Кнайт (B. Knight) має на увазі гранти Федеральної допомоги для автомагістралей (the Federal Highway Aid Program) [19].

⁹ Хоча ідея ефекту храповика характерна для опису урядових зв'язків у планових економіках, вона легко застосовується і до української міжурядової фіскальної політики.

дитячих садків, лікарень і т. д. Ми очікуємо отримати спадаючу криву середніх витрат і, отже, негативний коефіцієнт для змінної населення. Змінна доходу повинна мати негативний градієнт, тому що гранти забезпечують зменшення регіональної нерівності.

Практично ми застосовуємо ті самі апроксимації доходу, що і в попередньому прикладі моделювання податкових надходжень від податку з прибутку громадян. Таким чином, допоміжна регресія для дотацій має загальний вигляд:

$$\begin{aligned} RPCGRANT_{it} = & \alpha_i + \beta_1 \cdot RPCIO_{it} + \beta_2 \cdot RPCAGRO_{it} + \beta_3 \cdot RWAGE_{it} + \\ & + \beta_4 \cdot PTWR_{it} + \beta_5 \cdot YTWR_{it} + \beta_6 \cdot RPCWARRF_{it} + \beta_7 \cdot RPCWARRF_{it} + \\ & + \beta_8 \cdot POP_{it} + \beta_9 \cdot SMSHARE_{it} + \beta_{10} \cdot RPCSMEXP_{it}, \end{aligned} \quad (5.71)$$

де i — індекс для області; t — індекс для часового періоду; $RPCGRANT$ — реальні дотації області на душу населення; $RPCIO$ — реальний випуск промислової продукції на душу населення; $RPCAGRO$ — реальний випуск сільськогосподарської продукції на душу населення; $RWAGE$ — середня заробітна плата; POP — населення області; $RPCWARRF$ — зміна реальної заборгованості заробітної плати на душу населення; $RPCPAYF$ — зміна в реальній (кредиторській) заборгованості між підприємствами на душу населення; $SMSHARE$ — частка витрат в державному управлінні; $RPCSMEXP$ — частка витрат в державному управлінні на душу населення; $PTWR$ — частка пенсіонерів (вік_{чол.} > 60, вік_{жін.} > 55) в населенні [=PENSION/POP]; $YTWR$ — частка дітей (вік < 16) в населенні [=YOUTH/POP].

Оцінка даної регресії наведена в таблиці 5.4.

Таблиця 5.4. Оцінка регресії від змінних доходу та демографічних змінних

Dependent Variable: RPCGRANT?

Method: Pooled Least Squares

Sample(adjusted): 1995 1999

Included observations: 5 after adjusting endpoints

Total panel (balanced) observations 125

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RPCAGRO?	-48.47853	10.67969	-4.539320	0.0000
RPCIO?	-11.43563	12.86085	-0.889181	0.3763
RWAGE?	-0.047757	0.182873	-0.261148	0.7946
RPCWARRF?	-627.4708	131.2722	-4.779921	0.0000
RPCPAYF?	-0.037360	7.903620	-0.004727	0.9962
POP?	0.284119	0.175223	1.621477	0.1084
SMSHARE_?	-3.175988	5.420442	-0.585928	0.5594

RPCSMEXP?	1.264533	1.796734	0.703795	0.4834
YOUTH?/POP?	-3574.883	1633.386	-2.188633	0.0312
PENSION?/POP?	-3184.487	2211.951	-1.439673	0.1534
Fixed Effects				
1—C	921.6005			
⋮				
25—C	1413.131			
R-squared	0.867982	Mean dependent var	68.64035	
Adjusted R-squared	0.818109	S.D. dependent var	68.51407	
S.E. of regression	29.22037	Sum squared resid	76844.70	
F-statistic	65.74733	Durbin-Watson stat	2.274844	
Prob(F-statistic)	0.000000			

В цілому, як ми й очікували, зростання доходів призводить до зменшення дотацій. На жаль, інші змінні не мають очікуваних знаків. Тепер розрахуємо залишки цієї регресії та включимо їх як змінну і отримаємо таку модель:

$$\begin{aligned}
 RPCEXP_{it} = & \alpha_i + \beta_1 \cdot RPCIO_{it} + \beta_2 \cdot RPCAGRO_{it} + \beta_3 \cdot RWAGE_{it} + \\
 & + \beta_4 \cdot RPCGRANT_{it} + \beta_5 \cdot D(RPCWARR_{it}) + \beta_6 \cdot D(RPCPAY_{it}) + \beta_7 \hat{u}_{it},
 \end{aligned}
 \tag{5.72}$$

де \hat{u}_{it} — оцінені залишки (помилки) з першої регресії.

Нульовою гіпотезою тесту Хаусмана є припущення, що дотації не є ендогенною змінною, тобто $\beta_7 = 0$. Альтернативна гіпотеза: дотації є ендогенними, тому для оцінювання параметрів моделі витрат місцевих урядів слід використати двокроковий метод найменших квадратів (2МНК).

Щоб перевірити нульову гіпотезу, оцінимо (5.72) та отримаємо такі результати:

Таблиця 5.5. Оцінка моделі (5.72)

Dependent Variable: RPCEXP?

Method: Pooled Least Squares

Date: 06/19/01 Time:09:27

Sample(adjusted): 1995 1999

Included observations: 5 after adjusting endpoints

Total panel (balanced) observations 125

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RPCAGRO?	113.7564	46.50244	2.446246	0.0163
RPCIO?	43.64617	25.11055	1.738160	0.0855
RWAGE?	0.938222	0.361318	2.596664	0.0109
RPCWARRF?	-467.5583	543.3985	-0.860434	0.3918
RPCPAYF?	15.23678	16.99163	0.896723	0.3722
RPCGRANT?	2.361487	0.772658	3.056314	0.0029
RESID?	-1.565015	0.807070	-1.939131	0.0555

Fixed Effects			
1—C	110.3431		
⋮			
25—C	-78.68506		
<hr/>			
R-squared	0.721752	Mean dependent var	457.5042
Adjusted R-squared	0.629003	S.D. dependent var	106.1124
S.E. of regression	64.63256	Sum squared resid	388495.2
F-statistic	40.20579	Durbin-Watson stat	1.997732
Prob(F-statistic)	0.000000		

З цієї регресії ми бачимо, що β_7 статистично значимо відрізняється від нуля при 5 %-му рівні значимості. Таким чином, ми відкидаємо нульову гіпотезу на користь альтернативної, тобто дотації є ендогенною змінною. Отже, як показав тест Хаусмана, наявність ендогенності є серйозною проблемою в моделі витрат місцевих урядів, тому необхідно застосувати двокроковий метод найменших квадратів для її оцінювання.

Для застосування двокрокового методу найменших квадратів підставимо в модель витрат уряду замість змінної дотацій (RPCGRANT) її розраховані (fitted) значення з моделі (5.71) (RPCGRANTHAT). Після цього оцінимо параметри моделі методом найменших квадратів (ми використовуємо модель з фіксованими ефектами). Результати оцінювання наведені в таблиці 5.6.

Таблиця 5.6. Оцінка параметрів моделі витрат уряду

Dependent Variable: RPCEXP?

Method: Pooled Least Squares

Date: 06/19/01 Time: 09:36

Sample(adjusted): 1995 1999

Included observations: 5 after adjusting endpoints

Total panel (balanced) observations 125

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RPCAGRO?	113.7564	49.07066	2.318216	0.0226
RPCINDOUT?	43.64617	26.49735	1.647190	0.1029
RWAGE?	0.938222	0.381273	2.460762	0.0157
RPCWARRF?	-467.5583	573.4091	-0.815401	0.4169
RPCPAYF?	15.23678	17.93003	0.849791	0.3976
RPCGRANTHAT?	2.361487	0.815331	2.896355	0.0047
Fixed Effects				
1—C	110.3431			
⋮				
25—C	-78.68506			
<hr/>				
R-squared	0.686838	Mean dependent var	457.5042	
Adjusted R-squared	0.586893	S.D. dependent var	106.1124	
S.E. of regression	68.20206	Sum squared resid	437243.0	
F-statistic	41.23290	Durbin-Watson stat	2.062396	
Prob(F-statistic)	0.000000			

Тест Хаусмана та 2МНК (TSLs) в командах E.VIEWS

Запишемо всі попередні кроки в командах E.VIEWS:

‘створіть об’єкт pool із 25 варіантами

POOL ___POOL2A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
24 25

‘оцініть регресійну модель першого кроку: змінна дотацій є функцією від контрольованих змінних

___POOL2A.LS(F) RPCGRANT? RPCAGRO? RPCIO? RWAGE? RPCWARRF?
RPCPAYF? POP? SMSHARE_? RPCSMEXP? YOUTH?/POP? PENSION?/POP?

‘обчисліть залишки для моделі першого кроку

___POOL2A.MAKERESID RESIDA?

‘створіть об’єкт pool із 25 варіантами, щоб протестувати ендогенність дотацій

POOL ___POO1A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
24 25

‘оцініть модель другого кроку, ввівши в неї залишки з попередньої моделі

‘нульова гіпотеза полягає в тому, що коефіцієнт при залишках дорівнює нулю
___POO1A.LS(F) RPCEXP? RPCAGRO? RPCIO? RWAGE? RPCWARRF?
RPCPAYF? RPCGRANT? RESIDA?

‘обчисліть t -статистику коефіцієнта при залишку

SCALAR ___TSTAT=___POO1A.@TSTAT(7)

‘обчисліть p -значення для розрахованої t -статистики коефіцієнта при залишку

MATRIX ___P_VALUE=@TDIST(___TSTAT,___POO1A.@TOTALOBS-
___POO1A.@NCOEF)

‘обчисліть розраховані (fitted) значення змінної дотацій

___POOL2A.GENR RPCGRANTHAT?=RPCGRANT?-RESIDA?

‘оцініть модель з «ендогенно очищеними» дотаціями

‘створіть об’єкт pool із 25 варіантами

POOL ___POOL4A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
24 25

‘оцініть модель з розрахованими (fitted) значеннями для дотацій («оцінка-ми дотацій»)

___POOL4A.LS(F) RPCEXP? RPCAGRO? RPCIO? RWAGE? RPCWARRF?
RPCPAYF? RPCGRANTHAT?

Крок 5. Тест на перевірку якості обраних інструментів: тест перетотожнених обмежень (overidentifying restrictions test).

У дискусії про ендогенність ми майже не звертали уваги на якість застосованих інструментів. Як ми можемо протестувати, підходять обрані нами

інструменти чи ні? Для перевірки якості інструментів розроблено спеціальний тест: тест переототожнених обмежень (*overidentifying restrictions test*).

Нульовою гіпотезою цього тесту є припущення, що обрані інструменти не корелюють із залишками первісної¹⁰ (тобто кінцевої) регресії. Відповідно, альтернативною гіпотезою є припущення, що вони корелюють. Якщо вони корелюють, то це може бути наслідком або того, що 1) інструменти не підходять самі по собі (тобто є поганими інструментами), або що 2) інструменти повинні застосовуватись у кінцевій регресії, тому що вони пояснюють залежну змінну в кінцевій регресії [31]. На жаль, ми не можемо розділити ці два випадки, що є суттєвим недоліком тесту. В цьому плані цей тест дуже подібний до тесту специфікації, який не має чітко визначеної альтернативної гіпотези. До того ж, ми не можемо визначити, який саме з інструментів не можна застосовувати.

Для того, щоб застосувати цей тест, ми повинні побудувати модель залежності залишків (помилки) кінцевої регресії (зі змінними, очищеними від ендогенності) від визначеного нами набору інструментів, застосованих на першому етапі (для розрахунку «ендогенно очищених» змінних). Зверніть увагу, що в цій допоміжній регресії не повинно бути константи. Можна показати, що величина, яка дорівнює добутку кількості спостережень на коефіцієнт детермінації ($n \times R^2$), є розподіленою асимптотично за χ^2 -розподілом з ступенями свободи ($R - K$), де R — кількість інструментів (*instrument variables*), а K — число змінних, для яких застосовуються інструменти (*instrumented variables*).

Тепер застосуємо тест переототожнених обмежень (*overidentifying restrictions test*) до перевірки якості інструментів, обраних нами для змінної дотацій.

На першому кроці обчислимо залишки регресії з рядами дотацій, очищеними від ендогенності. Після цього побудуємо та оцінимо регресію, в якій ці обчислені залишки є функцією від набору інструментів (частка осіб пенсійного віку *PTWR*, частка дітей *YTWR*, населення *POP*, частка місцевих витрат на державне управління *SMSHARE*, реальні витрати на державне управління на душу населення *RPCSMEXP*).

Результати розрахунків наведені в таблиці 5.7.

Результати оціненої допоміжної регресії показують, що інструменти можна застосовувати, вони не корелюють із залишками: розраховане значення χ^2 -статистики (кількість ступенів свободи = $5 - 1 = 4$) дорівнює 2.84 ($\chi^2 = nR^2 = 0.02 \cdot 125 = 2.84$, $p\text{-value} = 0.6$), тобто ми не можемо відкинути нульову гіпотезу.

¹⁰ Не слід плутати первісну регресію (тобто ту, яка нас, власне, цікавить) з допоміжною, яку ми оцінюємо на першому кроці.

Таблиця 5.7. Оцінка допоміжної регресії

Dependent Variable: RESIDB?

Method: Pooled Least Squares

Sample(adjusted): 1995 1999

Included observations: 5 after adjusting endpoints

Total panel (balanced) observations 125

Variable	Coefficient	Std. Error	t-Statistic	Prob.
POP?	-0.000757	0.005511	-0.137309	0.8910
SMSHARE?	-10.25147	6.206937	-1.651615	0.1012
RPCSMEXP?	2.533864	1.871617	1.353837	0.1783
YOUTH?/POP?	38.26813	136.1329	0.281109	0.7791
PENSION?/POP?	-16.01301	133.6618	-0.119802	0.9048
R-squared	0.022729	Mean dependent var	2.11E-14	
Adjusted R-squared	-0.009847	S.D. dependent var	59.38142	
S.E. of regression	59.67307	Sum squared resid	427305.0	
F-statistic	0.697722	Durbin-Watson stat	2.083004	
Prob(F-statistic)	0.594982			

Ми можемо отримати той самий результат, застосовуючи такі команди **E.VIEWS:**

‘створіть новий pool object

POOL __POO4A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

‘оцініть модель з «ендогенно очищеними» дотаціями

__POO4A.LS(F) RPCEXP? RPCAGRO? RPCIO? RWAGE? RPCWARRF? RPCPAYF? RPCGRANTHAT?

‘розрахуйте залишки

__POOL4.MAKERESID RESIDB?

‘створіть новий pool object

POOL __POOL5A 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

‘оцініть регресію залишків від набору інструментів

__POOL5A.LS(N) RESIDB? POP? SMSHARE? RPCSMEXP? YOUTH?/POP? PENSION?/POP?

‘обчисліть χ^2 -статистику: R^2 помножене на кількість спостережень

SCALAR __CHI_STAT=__POOL5A.@R2*__POOL5A.@TOTALOBS

‘обчисліть *p-value*

SCALAR __P_CHI=@CHISQ(__CHI_STAT,__POOL5A.@NCOEF-1)

Крок 6. Загальні висновки та коментарі.

Зробіть коментарі та загальні висновки щодо усунення проблеми ендогенності в моделях на основі лонгitudних даних самостійно.

Питання для самоперевірки

1. Наведіть приклади застосування моделей лонгitudних даних на практиці. Які переваги мають моделі лонгitudних даних в порівнянні з іншими моделями? Детально прокоментуйте та поясніть модель.

2. Наведіть основні проблеми, які можуть виникнути при застосуванні моделей лонгitudних даних на практиці.

3. Запишіть модель лонгitudних даних в загальному вигляді. Детально прокоментуйте та поясніть її.

4. Поясніть, в чому полягає основна відмінність між моделями лонгitudних даних з фіксованими та випадковими ефектами.

5. Чому моделі з фіксованими та випадковими ефектами ще називають моделями з однокомпонентною помилкою?

6. В чому полягає особливість оцінювання невідомих параметрів моделей лонгitudних даних з фіксованими ефектами?

7. Поясніть детально, чому ми не можемо оцінювати невідомі параметри моделей лонгitudних даних з випадковими ефектами методом найменших квадратів. Який метод нам потрібно використовувати для оцінювання та чому?

8. Поясніть, коли ми повинні застосовувати моделі з фіксованими ефектами, а коли з випадковими. Яка принципова різниця між ними?

9. Наведіть основний формальний критерій перевірки ефективності застосування моделей лонгitudних даних з випадковими ефектами. Прокоментуйте його особливості.

10. В чому полягає суть F -критерію Фішера в моделях лонгitudних даних? Що ми тестуємо за допомогою цього критерію? Відповідь на яке питання ми отримуємо? Які особливості його застосування?

11. За допомогою якого критерію ми можемо протестувати, що модель із загальним перетином є кращою для застосування на практиці в порівнянні з моделями з випадковими або фіксованими ефектами? Детально поясніть, що таке модель із загальним перетином. Запишіть її у загальному вигляді та прокоментуйте її основні особливості.

12. В чому полягає проблема ендогенності в моделях лонгitudних даних?

13. Що таке інструментальні змінні? Детально поясніть особливості їх застосування в моделях лонгitudних даних.

14. Наведіть алгоритм перевірки коректності обраних інструментів у моделях лонгitudних даних. За допомогою якого тесту ми можемо протестувати коректність інструментів?

15. Чи є проблема наявності можливої автокореляції істотною в моделях лонгитюдних даних? Детально прокоментуйте вашу відповідь.

16. Яка з проблем — автокореляції чи гетероскедастичності — є більш поширеною в моделях лонгитюдних даних? Чому?

17. Наведіть основні особливості оцінювання невідомих параметрів у моделях лонгитюдних даних з випадковими ефектами. Яким чином ми знімаємо проблему кореляції між випадковими величинами та неявними факторами?

18. Коротко поясніть, яким чином моделі лонгитюдних даних можуть використовуватись для прогнозу.

19. Наведіть чотири основних, на ваш погляд, критерії прогнозової якості моделей лонгитюдних даних.

Вправи

Вправа 5.1.

1. Припустимо, що ми маємо дані про інвестиційні витрати (Y) та прибутки (X) для трьох фірм протягом 10 періодів, які наведено в таблиці (дані умовні в умовних одиницях).

Таблиця 5.8. Дані про діяльність трьох фірм

t	$I = 1$		$I = 2$		$I = 3$	
	Y	X	Y	X	Y	X
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

Використовуючи інформацію, наведену в таблиці, розрахуйте коефіцієнти регресійної моделі лонгитюдних даних методом найменших квадратів (МНК):

$$Y_{it} = \alpha + \beta X_{it} - \varepsilon_{it}$$

Прокоментуйте отримані результати.

Вправа 5.2.

Використовуючи дані завдання 5.1, оцініть модель лонгітюдних даних з фіксованими ефектами та протестуйте гіпотезу про те, що для всіх трьох фірм константа у регресійній моделі має однакові значення. Обґрунтуйте ваші висновки.

Вправа 5.3.

Використовуючи дані завдання 5.1, оцініть модель лонгітюдних даних з випадковими ефектами, а потім перевірте гіпотезу про можливість використання класичної моделі із загальним перетином. За допомогою якого тесту це можна зробити? Прокоментуйте результати тестування. Обґрунтуйте ваші висновки.

Вправа 5.4.

Використовуючи дані завдання 5.1, перевірте за допомогою тесту Хаусмана, якій із специфікацій моделі, з фіксованими чи з випадковими ефектами, краще віддати перевагу? Прокоментуйте ваші висновки.

Список літератури

1. *Ahmad Enthisham*. Financing Decentralized Expenditures: An international Comparison of Grants.— Cheltenham, UK : Edward Elgar, 1997.
2. *Amemiya T., MaCurdy T. E.* Instrumental-Variable Estimation of an Error Components Model // *Econometrica*, 1986, № 54.— P. 869—881.
3. *Arellano M.* Computing Robust Standard Errors for Within-Groups Estimators // *Oxford Bulletin of Economics and Statistics*, 1987.— № 49, P. 431—434.
4. *Arellano M.* Some Testing for Autocorrelation in Dynamic Random Effects Models // *Review of Economic Studies*, 1990, № 57.— P. 127—134.
5. *Baltagi Badi H.* Econometric Analysis of Panel Data.— New York : John Wiley & Sons, Inc., 1995.— P. 253.
6. *Bhargava A., Franzini L., Narendranathan W.* Serial Correlation and Fixed Effects Model // *Review of Economic Studies*, 1982, № 49.— P. 533—549.
7. *Breusch T. S., Mizon G. E., Schmidt P.* Efficient Estimation Using Panel Data // *Econometrica*, 1989, № 57.— P. 695—700.
8. *Cullis John, Jones Philip.* Public Finance and Public Choice.— 2nd ed.— Oxford : Oxford University Press, 1998.
9. *Earle John S., Klara Z. Sabirianova.* Equilibrium Wage Arrears: A Theoretical and Empirical Analysis of Institutional Lock-In. The Institute for

the Study of Labor (IZA). Discussion Paper, № 196. Available at <ftp://mail-server.iza.org/RePEc/Discussionpaper/dp196.pdf>., 2000.

10. *Greene William*. Econometric analysis.— 4th ed.— Upper Saddle River.— New Jersey : Prentice-Hall, Inc., 2000.

11. *Gujara Damodar N*. Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, Inc., 1995.

12. *Hausman J. A*. Specification Tests in Econometrics // *Econometrica*, 1978, № 43.— P. 1251—1271.

13. *Hausman J. A., Taylor W. E*. Panel Data and Unobservable Individual Effects // *Econometrica*, 1981, № 49.— P. 1377—1398.

14. *Hsiao Cheng*. Analysis of Panel Data. UK.— Cambridge : Cambridge University Press, 1986.

15. *Hoch I*. Estimation of Production Function Parameters Combining Time-Series and Cross-Sectional data // *Econometrica*, 1962, № 30.— P. 34—53.

16. International Monetary Fund. «Ukraine: Recent Economic Developments». IMF Staff Country Paper, № 99/42, 1999.

17. *Islam Muhammed N., Saud A. Choudhury*. The Flypaper Effect and the Revenue Impact of Grants-in-Aid // *Economics Letters*, 1989, Vol. 30 (4).— P. 351—356.

18. *Kmenta J*. Elements of Econometrics.— New York : MacMillan, 1986.

19. *Knight Brian*. The Flypaper Effect Unstuck: Evidence on Endogenous Grants from The Federal Highway Aid Program. Board of Governors of the Federal Reserve System. Washington DC. Available at <http://netec.mcc.ac.uk>., 2000.

20. *Корнаи Янош*. Дефіцит.— М. : Наука, 1991.

21. *Kuh E*. The Validity of Cross-Sectionally Estimated Behavior Equations in Time-Series Applications // *Econometrica*, 1959, № 27.— P. 197—214.

22. *Matyas L., Sevestre P*. The Econometrics of Panel Data: Handbook of Theory and Applications.— Dordrecht : Kluwer Academic Publishers, 1992.

23. *Mundlak Y*. Empirical Production Function Free of Management Bias // *Journal of Farm Economics*, 1969, № 43.— P. 44—56.

24. *Mundlak Y*. On the Pooling of Time-Series and Cross-Section Data // *Econometrica*, 1978, № 46.— P. 44—56.

25. *Nerlove M., Balestra P*. Formulation and Estimation of Econometric Models for Panel Data // *Matyas L., Sevestre P*. The Econometrics of Panel Data: Handbook of Theory and Applications.— Dordrecht : Kluwer Academic Publishers, 1992.— P. 3—18.

26. *Nijman Th. E., Verbeek M*. Estimation of Time Dependent Parameters in Linear Models Using Cross-Section Panels or Both // *Journal of Econometrics*, 1990, № 46.— P. 333—346.

27. *Nijman Th. E., Verbeek M*. Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function // *Journal of Applied Econometrics*, 1992, № 7.— P. 243—257.

28. *Pindyck Robert S., Rubinfeld Daniel L.* Econometric Models and Economic Forecasts.— 3rd ed.— New York : McGraw-Hill, Inc., 1991.
29. *Roland Gerard.* Transition and Economics: Politics, Markets and Firms.— Cambridge, Massachusetts : MIT Press, 2000.
30. *Schneider Mark, Byung Moon Ji.* The Political Economy of Intergovernmental Grant Seeking: Targeting and Suburbs, 1977 and 1982 // American Journal of Political Science, 1990, May, Vol. 34(2).— P. 408—420.
31. *Verbeek Marno.* A Guide to Modern Econometrics.— Chichester : John Wiley and Sons, Ltd., 2000.
32. *Zhuravskaya Ekatherina.* Incentives to Provide Local Public Goods: Fiscal Federalism, Russian Style. EERC Russia Working Papers, 1999. Available at http://www.eerc.ru/publications/workpapers/WP_99-15E.pdf.
33. *Городніченко Ю. О.* Вплив трансферів на податкові зусилля областей // Наукові записки НаУКМА. Спец. випуск. Т. 19, ч. II.— К., 2001.— С. 371—375.
34. *Городніченко Ю. О.* Розщеплення податків як складова підвищення ефективності регіональної політики // Наукові записки. НаУКМА. Серія «Економічні науки», № 19.— К., 2001.— С. 34—40.
35. *Кваджа С., Лук'яненко І. Г., Городніченко Ю. О.* Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К. : Видавничий дім «КМ Academia», 2000.— 204 с.
36. *Лук'яненко І. Г.* Аналіз західного досвіду політики вирівнювання регіонів // Система державних фінансів України: шляхи забезпечення прозорості та оптимальної децентралізації фінансових потоків. Матеріали форуму експертів, Київ, 2 березня 2001 р.: К.І.С.— К., 2001.— С. 64—70.
37. *Лук'яненко І. Г., Городніченко Ю. О.* Моделювання розподілу дотацій між регіонами України // Вісник НБУ, 1999, № 9.— С. 16—21.
38. *Лук'яненко І. Г., Городніченко Ю. О., Краснікова Л. І.* Економетричні підходи до аналізу фінансової програми місцевих органів влади України.— К. : Видавничий дім «КМ Academia», 2000.— 204 с.
39. *Лук'яненко І. Г., Ситник Т. М.* Моделювання заборгованості по заробітній платі на регіональному рівні // Наукові записки НаУКМА. Спец. випуск. Т. 19, ч. II.— К., 1999.— С. 111—114.

ДОДАТКИ

Додаток 1

Вхідні дані для ідентифікації ARIMA-моделей

Рік: місяць	CPI (індекс споживчих цін)	ЕРТ (податок з прибутку підприємств, млрд грн)	РІТ (прибутковий податок з грома- дян, млрд грн)	VAT (податок на додану вартість, млрд грн)
1993:01	135.6773	1.255700	0.363200	1.480900
1993:02	174.7523	1.665300	0.304500	2.074300
1993:03	213.3726	1.325700	0.394000	3.432900
1993:04	263.7285	8.401500	0.531800	4.513100
1993:05	336.5176	3.591900	0.504300	5.291100
1993:06	577.8007	3.134300	0.993100	6.722100
1993:07	795.0538	11.50000	1.125700	8.115600
1993:08	967.5805	5.460900	1.538400	11.72850
1993:09	1744.548	7.645600	1.703500	19.79200
1993:10	2897.694	35.01890	3.530200	28.96980
1993:11	4210.349	16.83300	6.916800	31.58840
1993:12	8033.346	48.90300	10.92870	48.35640
1994:01	9575.748	33.17800	11.44470	40.64070
1994:02	10782.29	46.96390	14.45900	38.95960
1994:03	11396.88	54.97540	16.43530	49.99660
1994:04	12080.70	121.0824	21.19300	72.00770
1994:05	12708.89	84.88670	20.50170	85.72140
1994:06	13204.54	81.37010	23.74070	95.63410
1994:07	13481.83	176.7875	28.74360	106.1917
1994:08	13832.36	127.7269	33.14930	126.9336
1994:09	14842.12	124.2752	36.20120	136.9325
1994:10	18196.44	227.9410	39.43910	149.5776
1994:11	31352.47	132.1347	42.19750	161.5469
1994:12	40256.58	218.2426	52.44520	235.2331
1995:01	48790.97	164.6000	52.30000	257.8000
1995:02	57622.14	305.2000	62.40000	239.5000
1995:03	64191.06	366.5000	69.30000	280.3000
1995:04	67914.14	339.4000	80.50000	289.9000
1995:05	71038.19	340.1000	99.50000	309.8000
1995:06	74448.02	352.5000	128.3000	382.2000
1995:07	78319.32	431.7000	144.5000	368.3000
1995:08	81922.01	464.4000	164.3000	420.2000
1995:09	93554.94	474.4000	185.9000	441.0000
1995:10	102068.4	473.9000	212.1000	493.8000
1995:11	108396.7	485.5000	211.0000	503.4000
1995:12	113382.9	662.4000	185.2000	543.7000
1996:01	124040.9	208.9040	172.8760	431.3320
1996:02	133219.9	467.3030	186.3510	393.4370
1996:03	137216.5	608.0160	184.1270	471.9870
1996:04	140509.7	375.8200	205.9880	508.6920

Закінчення додатку 1

Рік: місяць	CPI (індекс споживчих цін)	ЕРТ (податок з прибутку підприємств, млрд грн)	РІТ (прибутковий податок з грома- дян, млрд грн)	VAT (податок на додану вартість, млрд грн)
1996:05	141493.3	371.4090	200.9370	428.1910
1996:06	141634.8	390.0190	203.5530	413.5210
1996:07	141776.4	490.6560	247.6620	611.2590
1996:08	149857.7	393.6240	231.6950	508.1840
1996:09	152854.9	437.9670	216.8970	424.6480
1996:10	155147.7	425.0690	270.9790	648.0190
1996:11	157009.4	525.1460	214.9510	559.8630
1996:12	158422.5	802.6400	257.1250	847.1050
1997:01	161907.8	147.8000	218.9000	460.5000
1997:02	163850.7	296.8000	220.1000	406.3000
1997:03	164014.6	412.1000	247.2000	631.1000
1997:04	165326.7	382.9000	268.8000	617.7000
1997:05	166649.3	420.7000	243.9000	651.7000
1997:06	166816.0	387.9000	265.8000	659.6000
1997:07	166982.8	444.9000	296.0000	697.2000
1997:08	166982.8	570.1000	319.6000	600.2000
1997:09	168986.6	462.6000	301.6000	693.8000
1997:10	170507.4	698.8000	293.5000	672.1000
1997:11	172042.0	634.9000	253.2000	732.7000
1997:12	174450.6	932.6000	367.1000	1419.400
1998:01	176718.5	469.0000	237.1000	365.4000
1998:02	177071.9	315.9000	240.4000	295.3000
1998:03	177426.0	473.3000	307.9000	546.1000
1998:04	179732.6	427.9000	290.5000	555.9000
1998:05	179732.6	314.8000	249.8000	713.5000
1998:06	179732.6	350.3000	289.7000	596.1000
1998:07	178115.0	442.5000	303.3000	433.5000
1998:08	178471.2	401.3000	273.5000	522.5000
1998:09	185253.1	437.9000	284.7000	574.9000
1998:10	196738.8	617.8000	312.9000	602.9000
1998:11	202641.0	472.8000	312.3000	669.4000
1998:12	209328.1	971.0000	468.5000	1584.600
1999:01	212468.0	427.9000	236.9000	538.6000
1999:02	214593.0	269.9000	254.1000	484.9000
1999:03	216739.0	368.5000	349.1000	612.1000
1999:04	221724.0	581.7000	357.6000	736.3000
1999:05	227045.0	497.4000	313.1000	509.7000
1999:06	227272.0	419.5000	374.2000	652.2000
1999:07	224999.0	579.6000	378.3000	776.8000
1999:08	227249.0	531.8000	386.8000	710.0000
1999:09	230431.0	463.9000	401.6000	829.3000
1999:10	232966.0	601.5000	399.4000	634.6000
1999:11	239722.0	533.3000	403.0000	733.4000
1999:12	249550.0	752.0000	544.8000	1084.100

Додаток 2

Вхідні дані для побудови VAR-моделі споживання та доходу

Рік: квартал	DLNC	DLNY	LNC	LNy
1947:2	0.016226	-0.018592	NA	NA
1947:3	0.001866	0.026948	0.018092	0.008356
1947:4	-0.009583	-0.011123	0.008509	-0.002767
1948:1	0.006645	0.019108	0.015154	0.016341
1948:2	0.011583	0.029364	0.026737	0.045706
1948:3	-0.000591	0.018500	0.026145	0.064206
1948:4	0.011641	0.002013	0.037786	0.066219
1949:1	0.005029	-0.018951	0.042816	0.047269
1949:2	0.003967	0.003021	0.046782	0.050290
1949:3	-0.005881	0.004854	0.040901	0.055144
1949:4	0.009702	0.008870	0.050604	0.064014
1950:1	0.013000	0.070234	0.063603	0.134247
1950:2	0.017782	-0.006103	0.081386	0.128144
1950:3	0.016311	0.015809	0.097697	0.143953
1950:4	-0.008918	0.011892	0.088779	0.155845
1951:1	0.021293	-0.002718	0.110073	0.153127
1951:2	-0.004635	0.015595	0.105437	0.168722
1951:3	0.016917	0.007578	0.122355	0.176300
1951:4	0.008089	0.001800	0.130444	0.178100
1952:1	0.001582	-8.56E-05	0.132026	0.178014
1952:2	0.019166	0.011072	0.151193	0.189087
1952:3	0.015287	0.019793	0.166480	0.208879
1952:4	0.014305	0.018675	0.180785	0.227555
1953:1	0.007989	0.012876	0.188775	0.240431
1953:2	0.008333	0.015569	0.197107	0.255999
1953:3	-0.001598	-0.003492	0.195509	0.252508
1953:4	-0.002760	-0.002866	0.192749	0.249642
1954:1	0.009016	0.002627	0.201765	0.252269
1954:2	0.006772	-0.002787	0.208536	0.249482
1954:3	0.014426	0.012914	0.222963	0.262396
1954:4	0.013829	0.021878	0.236792	0.284274
1955:1	0.010384	0.013387	0.247176	0.297661
1955:2	0.012536	0.022986	0.259712	0.320647
1955:3	0.008498	0.018468	0.268210	0.339115
1955:4	0.019864	0.011957	0.288074	0.351072
1956:1	0.010250	0.008321	0.298324	0.359393
1956:2	0.003695	0.010305	0.302019	0.369699
1956:3	0.005731	0.003529	0.307750	0.373228
1956:4	0.010622	0.014757	0.318372	0.387985
1957:1	0.004430	0.001180	0.322802	0.389164
1957:2	0.005993	0.006842	0.328795	0.396006
1957:3	0.010743	0.005974	0.339538	0.401980
1957:4	0.002513	-0.003566	0.342051	0.398413
1958:1	-0.005551	-0.012653	0.336501	0.385761
1958:2	0.011941	0.006865	0.348442	0.392625
1958:3	0.015253	0.022075	0.363695	0.414701
1958:4	0.008637	0.014961	0.372332	0.429662
1959:1	0.011649	0.006902	0.383981	0.436563
1959:2	0.012507	0.017763	0.396488	0.454327
1959:3	0.009569	-0.002146	0.406057	0.452180

Продовження додатку 2

Рік: квартал	DLNC	DLNY	LNC	LNy
1959:4	0.010052	0.007976	0.416109	0.460157
1960:1	0.005472	0.010537	0.421581	0.470694
1960:2	0.010347	0.003955	0.431928	0.474649
1960:3	-0.003819	0.000764	0.428108	0.475413
1960:4	0.005067	-0.001719	0.433175	0.473694
1961:1	0.008184	0.007744	0.441359	0.481438
1961:2	0.013918	0.015311	0.455277	0.496748
1961:3	0.002648	0.014036	0.457925	0.510785
1961:4	0.015820	0.020120	0.473745	0.530905
1962:1	0.008858	0.008628	0.482602	0.539534
1962:2	0.010011	0.010092	0.492614	0.549625
1962:3	0.008219	0.006593	0.500833	0.556218
1962:4	0.008933	0.007191	0.509766	0.563410
1963:1	0.004601	0.007718	0.514366	0.571128
1963:2	0.007511	0.009264	0.521877	0.580391
1963:3	0.013562	0.010993	0.535439	0.591385
1963:4	0.007114	0.017130	0.542553	0.608515
1964:1	0.017178	0.019248	0.559731	0.627762
1964:2	0.016294	0.025828	0.576025	0.653591
1964:3	0.015804	0.014008	0.591828	0.667598
1964:4	0.008382	0.012260	0.600210	0.679858
1965:1	0.010513	0.009803	0.610723	0.689662
1965:2	0.012739	0.012653	0.623462	0.702315
1965:3	0.013767	0.027158	0.637229	0.729473
1965:4	0.025978	0.020657	0.663207	0.750129
1966:1	0.009341	0.008231	0.672548	0.758360
1966:2	0.010384	0.002681	0.682933	0.761041
1966:3	0.007678	0.013014	0.690610	0.774055
1966:4	0.004043	0.012335	0.694653	0.786390
1967:1	0.010504	0.014438	0.705157	0.800828
1967:2	0.008578	0.007826	0.713735	0.808654
1967:3	0.007134	0.008398	0.720870	0.817051
1967:4	0.006388	0.006488	0.727258	0.823540
1968:1	0.017106	0.017452	0.744364	0.840992
1968:2	0.014669	0.017153	0.759033	0.858145
1968:3	0.014237	0.002600	0.773270	0.860744
1968:4	0.006331	0.006728	0.779600	0.867472
1969:1	0.010051	0.001975	0.789652	0.869447
1969:2	0.007793	0.010752	0.797445	0.880199
1969:3	0.006125	0.019583	0.803569	0.899782
1969:4	0.009941	0.006306	0.813510	0.906087
1970:1	0.010073	0.003921	0.823583	0.910008
1970:2	0.003990	0.016099	0.827573	0.926107
1970:3	0.009312	0.011125	0.836884	0.937232
1970:4	0.007301	-0.003252	0.844185	0.933980
1971:1	0.006001	0.018370	0.850187	0.952350
1971:2	0.006699	0.014075	0.856886	0.966425
1971:3	0.004557	0.003381	0.861443	0.969806
1971:4	0.011842	0.006729	0.873285	0.976535
1972:1	0.011418	0.006224	0.884703	0.982759
1972:2	0.016886	0.011122	0.901589	0.993881
1972:3	0.012960	0.021879	0.914549	1.015760

Продовження додатку 2

Рік: квартал	DLNC	DLNY	LNC	LNy
1972:4	0.017961	0.034599	0.932510	1.050359
1973:1	0.009592	0.013829	0.942102	1.064188
1973:2	0.001563	0.010291	0.943665	1.074480
1973:3	0.006925	0.006340	0.950590	1.080820
1973:4	0.001031	0.016154	0.951621	1.096974
1974:1	-0.004894	-0.016328	0.946727	1.080646
1974:2	0.003145	-0.011105	0.949872	1.069540
1974:3	0.003055	0.003646	0.952927	1.073186
1974:4	-0.002189	-0.009280	0.950738	1.063906
1975:1	0.006746	-0.005418	0.957484	1.058488
1975:2	0.016437	0.046313	0.973922	1.104801
1975:3	0.006708	-0.014100	0.980629	1.090701
1975:4	0.007474	0.010670	0.988104	1.101371
1976:1	0.015950	0.015493	1.004054	1.116865
1976:2	0.009852	0.007242	1.013906	1.124107
1976:3	0.009903	0.007322	1.023809	1.131429
1976:4	0.012596	0.006088	1.036405	1.137517
1977:1	0.007515	0.001410	1.043920	1.138927
1977:2	0.002482	0.011337	1.046401	1.150264
1977:3	0.008738	0.015429	1.055139	1.165693
1977:4	0.013843	0.016608	1.068982	1.182301
1978:1	0.010137	0.010144	1.079119	1.192445
1978:2	0.012361	0.014032	1.091480	1.206477
1978:3	0.007976	0.007110	1.099456	1.213587
1978:4	0.007491	0.012299	1.106947	1.225886
1979:1	0.008507	0.009000	1.115454	1.234886
1979:2	0.001798	-0.002563	1.117252	1.232323
1979:3	0.005480	0.004852	1.122731	1.237175
1979:4	0.008626	0.004977	1.131357	1.242152
1980:1	0.000709	0.005540	1.132065	1.247692
1980:2	-0.009401	-0.019166	1.122664	1.228526
1980:3	0.005246	0.009184	1.127910	1.237710
1980:4	0.009164	0.018643	1.137074	1.256353
1981:1	8.38E-05	0.004099	1.137158	1.260452
1981:2	0.005868	-0.005265	1.143026	1.255187
1981:3	0.000559	0.019697	1.143585	1.274884
1981:4	0.001269	-0.002577	1.144854	1.272306
1982:1	0.002618	-0.006674	1.147471	1.265632
1982:2	0.002620	0.012193	1.150091	1.277825
1982:3	0.005702	-0.002398	1.155793	1.275427
1982:4	0.012004	0.003937	1.167797	1.279363
1983:1	0.008192	0.003013	1.175989	1.282377
1983:2	0.012730	0.009436	1.188719	1.291813
1983:3	0.011359	0.014793	1.200078	1.306605
1983:4	0.010774	0.019996	1.210853	1.326601
1984:1	0.006520	0.023325	1.217373	1.349927
1984:2	0.011324	0.018271	1.228697	1.368198
1984:3	0.007343	0.015971	1.236040	1.384169
1984:4	0.009373	0.003225	1.245413	1.387394
1985:1	0.011454	0.001047	1.256867	1.388441
1985:2	0.007902	0.018314	1.264769	1.406755
1985:3	0.010294	-0.004522	1.275063	1.402233

Закінчення додатку 2

Рік: квартал	DLNC	DLNY	LNC	LNy
1985:4	0.009417	0.009522	1.284480	1.411755
1986:1	0.007293	0.013968	1.291772	1.425724
1986:2	0.006812	0.008766	1.298585	1.434490
1986:3	0.005470	0.004937	1.304055	1.439427
1986:4	0.009378	-0.001967	1.313432	1.437460
1987:1	0.010625	0.009507	1.324057	1.446967
1987:2	0.008563	-0.009921	1.332620	1.437046
1987:3	0.005443	0.014365	1.338063	1.451412
1987:4	0.005710	0.015177	1.343773	1.466589
1988:1	0.012651	0.015927	1.356424	1.482516
1988:2	0.007694	0.002089	1.364118	1.484605
1988:3	0.011016	0.007232	1.375134	1.491836
1988:4	0.008012	0.005647	1.383146	1.497484
1989:1	0.002410	0.010950	1.385556	1.508434
1989:2	0.001532	-0.004243	1.387088	1.504191
1989:3	0.006997	0.004401	1.394085	1.508593
1989:4	0.007837	0.004134	1.401922	1.512727
1990:1	0.003898	0.012056	1.405820	1.524783
1990:2	0.006281	0.004090	1.412101	1.528873
1990:3	0.005199	0.001175	1.417300	1.530048
1990:4	-0.004859	-0.006848	1.412440	1.523200
1991:1	-0.003798	-0.003576	1.408643	1.519624
1991:2	0.006480	0.008916	1.415122	1.528540
1991:3	0.000702	0.000355	1.415824	1.528895
1991:4	-0.001242	0.002570	1.414582	1.531465
1992:1	0.013589	0.012883	1.428171	1.544348
1992:2	0.004745	0.007464	1.432916	1.551812
1992:3	0.005236	0.000303	1.438152	1.552116
1992:4	0.011361	0.019042	1.449513	1.571158
1993:1	0.001441	-0.017202	1.450954	1.553956
1993:2	0.005495	0.017500	1.456450	1.571456
1993:3	0.008683	0.002760	1.465132	1.574216
1993:4	0.005057	0.011718	1.470189	1.585934
1994:1	0.008503	-0.018803	1.478692	1.567131
1994:2	0.007098	0.016370	1.485791	1.583501
1994:3	0.005113	0.007179	1.490904	1.590679
1994:4	0.005402	0.011796	1.496306	1.602476
1995:1	0.005643	0.006348	1.501949	1.608824
1995:2	0.007555	-0.001435	1.509505	1.607389
1995:3	0.004561	0.006765	1.514066	1.614155
1995:4	0.004875	0.009529	1.518941	1.623683
1996:1	0.008301	0.007217	1.527242	1.630901
1996:2	0.008914	0.005275	1.536157	1.636175
1996:3	0.005878	0.010782	1.542035	1.646957
1996:4	0.005745	0.003188	1.547780	1.650145
1997:1	0.007917	0.008180	1.555697	1.658326
1997:2	0.005065	0.007127	1.560762	1.665453
1997:3	0.011852	0.005962	1.572614	1.671415
1997:4	0.006754	0.007169	1.579368	1.678584
1998:1	0.011692	0.009750	1.591060	1.688334
1998:2	0.013069	0.006485	1.604129	1.694819
1998:3	0.010601	0.007973	1.614730	1.702793

Продовження додатку 3

Рік: місяць	AWAGE (середня ЗП), грн	CPI (ІСЦ)	ЕРТ (податок з прибутку підприємств), млрд грн	GDP (ВВП), млрд грн	NTI (неподат- ковуваний мінімум), грн	РАУ (кред. забор- гованість між підприєм- ствами), млн грн	РІТ (податок з прибутку громадян), млрд грн	ВАТ (ПДВ), млрд грн	WARR (заборо- ваність із зарплати), млн грн
1995:06	79.03000	74448.02	352.5000	4258.290	1.400000	20477.00	128.3000	382.2000	N.A
1995:07	83.93000	78319.32	431.7000	4461.140	1.400000	19427.00	144.5000	368.3000	N.A
1995:08	90.03000	81922.01	464.4000	5026.770	1.400000	20361.00	164.3000	420.2000	N.A
1995:09	97.83000	93554.94	474.4000	5124.940	1.400000	26943.00	185.9000	441.0000	N.A
1995:10	105.23000	102068.4	478.9000	5476.210	1.700000	25911.00	212.1000	493.8000	N.A
1995:11	111.19000	108396.7	485.5000	5587.220	1.700000	28282.00	211.0000	503.4000	N.A
1995:12	134.3800	113382.9	662.4000	5577.360	1.700000	30543.00	185.2000	543.7000	575.0000
1998:01	114.5200	124040.9	208.9040	5896.240	1.700000	33801.00	172.8760	431.3320	852.0000
1998:02	120.9300	133219.9	467.3030	6168.290	1.700000	37233.00	186.3510	393.4370	947.0000
1998:03	129.3300	137216.5	608.0160	6266.610	1.700000	45152.00	184.1270	471.9870	1300.0000
1998:04	129.5900	140509.7	375.8200	6646.400	1.700000	45884.00	205.9880	508.6920	1771.0000
1998:05	132.1000	141493.3	371.4090	6414.470	1.700000	45935.00	200.9370	428.1910	2197.0000
1998:06	137.7400	141634.8	390.0190	5903.560	1.700000	55533.00	203.5530	413.5210	2616.0000
1998:07	143.7300	141776.4	490.6560	6934.990	1.700000	53838.00	247.6620	611.2590	2778.0000
1998:08	143.0500	149857.7	393.6240	6745.410	1.700000	55747.00	231.6950	508.1840	3000.0000
1998:09	146.0900	152854.9	437.9670	6326.270	1.700000	60289.00	216.8970	424.6480	3078.0000
1998:10	148.2100	15147.7	425.0690	6998.240	1.700000	59226.00	270.9790	648.0190	3251.0000
1998:11	144.7600	157009.4	525.1460	6502.690	1.700000	60574.00	214.9510	559.8680	3559.0000
1998:12	163.6900	158422.5	802.6400	7296.840	1.700000	73168.00	257.1250	847.1050	3739.0000
1997:01	139.5000	161907.8	147.8000	5919.990	1.700000	72156.00	218.9000	460.5000	4151.0000
1997:02	138.9000	163850.7	296.8000	5959.850	1.700000	72784.00	220.1000	406.3000	4288.0000
1997:03	147.7300	164014.6	412.1000	6947.750	1.700000	84968.00	247.2000	631.1000	4371.0000
1997:04	146.7500	165326.7	382.9000	7164.710	1.700000	86805.00	268.8000	617.7000	4428.0000
1997:05	153.3900	166649.3	420.7000	7154.290	1.700000	86134.00	243.9000	651.7000	4593.0000
1997:06	158.3300	166816.0	387.9000	7046.670	1.700000	95149.00	265.8000	659.6000	4665.0000
1997:07	165.0600	166982.8	444.9000	7358.920	1.700000	92814.00	296.0000	697.2000	4852.0000
1997:08	159.9200	166982.8	570.1000	7271.780	1.700000	93271.00	319.6000	600.2000	4735.0000
1997:09	163.3600	168986.6	462.6000	8377.290	1.700000	98428.00	301.6000	693.8000	4723.0000
1997:10	162.9200	170507.4	698.8000	7970.580	1.700000	99138.00	293.5000	672.1000	4771.0000
1997:11	160.2400	172042.0	634.9000	7485.660	1.700000	99847.00	253.2000	732.7000	4931.0000
1997:12	178.3200	174450.6	932.6000	8294.500	1.700000	102507.0	367.1000	1419.400	4908.0000

Закінчення додатку 3

Рік: місяць	AWAGE (середня ЗП), грн	СРІ (ІСЦ)	БРТ (податок з прибутку підприємств), млрд грн	GDP (ВВП), млрд грн	NTI (неоподат- ковуваний мінімум), грн	PAY (кред. забор- гованість між підпри- ємствами), млн грн	PIT (податок з прибутку громадян), млрд грн	VAT (ПДВ), млрд грн	WARR (заборго- ваність із зарплати), млн грн
1998:01	152.1700	176718.5	469.0000	6195.730	1.700000	93979.00	237.1000	365.4000	5156.000
1998:02	153.7300	177071.9	315.9000	6398.020	1.700000	97826.00	240.4000	295.3000	5330.000
1998:03	165.8200	177426.0	473.3000	8080.080	1.700000	103873.0	307.9000	546.1000	5337.000
1998:04	161.5800	179732.6	427.9000	7172.880	1.700000	104902.0	290.5000	555.9000	5469.000
1998:05	163.7300	179732.6	314.8000	6955.090	1.700000	108729.0	249.8000	713.5000	5732.000
1998:06	172.9800	179732.6	350.3000	7527.140	1.700000	114015.0	289.7000	596.1000	5785.000
1998:07	173.3300	178115.0	442.5000	7497.610	1.700000	113886.0	303.3000	433.5000	5997.000
1998:08	166.4300	178471.2	401.3000	7726.620	1.700000	114247.0	273.5000	522.5000	6201.000
1998:09	169.9800	185253.1	437.9000	8348.900	1.700000	121125.0	284.7000	574.9000	6423.000
1998:10	171.0100	196738.8	617.8000	9036.400	1.700000	133221.0	312.9000	602.9000	6542.000
1998:11	170.5400	202641.0	472.8000	8512.250	1.700000	134311.0	312.3000	669.4000	6567.000
1998:12	188.3000	209328.1	971.0000	11078.44	1.700000	137614.0	468.5000	1584.600	6518.000
1999:01	162.3400	212468.0	427.9000	7500.000	1.700000	138988.0	236.9000	538.6000	6635.000
1999:02	167.5600	214593.0	269.9000	7887.000	1.700000	147575.0	254.1000	484.9000	6755.000
1999:03	184.1000	216739.0	368.5000	9097.000	1.700000	163394.0	349.1000	612.1000	6814.000
1999:04	182.0600	221724.0	581.7000	8866.000	1.700000	168446.0	357.6000	736.3000	6798.000
1999:05	186.2800	227045.0	497.4000	8474.000	1.700000	172668.0	313.1000	509.7000	6888.000
1999:06	198.5900	227272.0	419.5000	9595.000	1.700000	182614.0	374.2000	652.2000	6886.000
1999:07	200.0900	224999.0	579.6000	10278.00	1.700000	183363.0	378.3000	776.8000	6973.000
1999:08	197.8000	227249.0	531.8000	10509.00	1.700000	188003.0	386.8000	710.0000	6990.000
1999:09	203.8900	230431.0	463.9000	11387.00	1.700000	192078.0	401.6000	829.3000	6830.000
1999:10	204.9500	232966.0	601.5000	11495.00	1.700000	193200.0	399.4000	634.6000	6645.000
1999:11	209.1600	239722.0	533.3000	11187.00	1.700000	195567.0	403.0000	733.4000	6602.000
1999:12	235.3100	249550.0	752.0000	12740.00	1.700000	196364.0	544.8000	1084.100	6462.000

Вхідні дані для спрощеної макромоделі України

Рік: квартал	CPI	M2	mb	NX	r	rCONS	rDI	rG	RGDP	rINV	rM2	rNBU	rTAX
1994-1	7,268	0,681			30,100						9,869		
1994-2	8,693	1,082			24,500						12,446		
1994-3	9,646	1,863			13,500						19,314		
1994-4	20,548	3,216			15,200						15,651		
1995-1	39,036	3,935	1,910	-0,808	17,300	13,540	20,030	12,633	30,397	5,031	10,081	20,000	10,886
1995-2	49,107	5,309	2,679	-0,601	10,200	14,006	21,450	14,306	34,173	6,461	10,811	9,200	13,360
1995-3	58,290	6,387	3,165	-0,516	6,500	16,773	26,149	15,486	39,316	7,574	10,957	5,600	13,826
1995-4	74,553	6,846	3,538	-0,671	8,400	16,182	24,232	15,846	37,171	5,813	9,183	8,100	13,586
1996-1	90,805	7,086	3,769	-1,491	9,000	12,915	17,762	11,278	27,357	4,655	7,804	8,600	9,595
1996-2	97,615	7,522	4,074	0,018	7,000	13,277	18,774	9,634	27,339	4,409	7,706	5,400	8,565
1996-3	102,399	7,592	4,302	-0,147	5,400	14,255	20,947	10,496	30,675	6,070	7,414	3,300	9,728
1996-4	108,440	9,024	4,882	-1,025	5,300	14,306	20,814	13,671	33,115	6,164	8,322	3,300	12,301
1997-1	112,886	9,514	5,199	-1,390	5,000	12,432	16,094	9,003	23,855	3,810	8,428	2,700	7,761
1997-2	114,918	10,971	6,122	-0,915	4,300	13,481	17,865	11,625	28,118	3,927	9,547	2,000	10,253
1997-3	115,837	12,222	6,877	-0,021	3,600	14,330	23,050	14,795	34,983	5,880	10,551	1,400	11,983
1997-4	119,081	12,447	7,058	-0,530	3,400	13,627	21,188	14,319	33,822	6,406	10,453	2,100	12,634
1998-1	122,296	12,836	7,096	-1,258	3,900	12,573	16,509	9,906	24,918	3,696	10,496	3,300	8,409
1998-2	124,130	13,256	7,259	-0,211	4,000	14,414	18,092	10,273	27,335	2,859	10,679	3,800	9,243
1998-3	124,751	14,142	7,534	-0,597	4,800	15,525	20,323	9,707	29,862	5,228	11,336	6,700	9,539
1998-4	140,270	15,432	8,604	-0,480	5,400	12,707	18,574	11,364	30,019	6,428	11,002	6,600	11,445
1999-1	148,274	15,631	8,562	-0,186	5,050	12,768	16,187	7,725	23,736	3,429	10,542	5,000	7,548
1999-2	155,688	18,258	10,334	1,774	4,680	13,188	18,563	9,333	27,526	3,232	11,727	4,175	8,963
1999-3	157,514	20,019	11,182	-0,078	4,200	15,299	20,833	10,336	30,829	5,272	12,709	3,750	9,996

Додаток 5

Змінна YOUTH (тис. чол.)

Змінна PENSION (тис. чол.)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	438.9	444.3	450.1	455.8	454.4	496.2	481.9	463.2	445.0	424.2
Вінницька	498.4	495.8	494.2	494.9	491.9	394.9	390.1	384.2	378.0	369.3
Волинська	236.9	237.1	236.6	235.3	233.1	270.3	266.1	261.4	256.9	251.6
Дніпропетровська	845.9	851.0	858.8	867.2	868.2	827.1	803.8	776.8	751.0	723.0
Донецька	1,227.9	1,230.8	1,234.6	1,236.5	1,228.9	070.1	031.4	986.8	946.4	903.7
Житомирська	367.9	366.7	366.4	366.8	364.2	331.6	325.4	319.3	313.1	305.2
Закарпатська	118.2	220.7	222.7	223.8	223.8	340.4	335.2	329.2	323.4	316.1
Запорізька	448.4	451.4	456.1	461.3	462.0	446.5	433.8	419.7	406.2	389.8
Івано-Франківська	301.9	305.4	307.2	307.0	306.1	361.2	357.4	351.6	345.3	337.1
Київська	437.6	437.9	439.9	442.4	441.5	416.9	407.2	396.7	386.3	373.6
Кіровоградська	297.6	295.9	295.0	295.3	293.7	264.2	259.5	253.3	246.7	239.0
Луганська	653.1	656.9	658.0	657.5	652.5	596.1	574.5	549.4	525.9	500.8
Львівська	573.6	578.0	580.9	581.1	579.7	640.6	630.7	618.1	605.5	591.0
Миколаївська	274.2	275.2	277.1	279.5	279.4	311.4	304.3	296.2	287.9	277.0
Одеська	532.8	533.5	539.3	543.0	543.9	567.5	555.2	541.1	526.1	509.3
Полтавська	446.3	447.5	447.0	448.1	446.8	357.1	350.0	341.6	333.3	323.1
Рівненська	239.7	241.3	242.1	242.1	240.8	308.7	305.4	301.1	297.1	291.6
Сумська	362.3	360.9	360.6	359.0	354.8	289.1	282.0	273.7	265.6	256.1
Тернопільська	285.0	286.5	286.6	285.2	282.9	272.6	269.2	265.2	260.9	255.3
Харківська	709.3	707.7	709.8	710.7	706.8	623.5	604.1	582.7	561.8	539.6
Херсонська	248.6	250.6	253.3	256.8	257.3	299.8	293.1	284.9	277.1	267.8
Хмельницька	380.6	379.2	377.6	376.7	373.9	331.3	326.2	320.3	314.1	306.3
Черкаська	387.3	386.6	386.7	387.8	386.0	318.0	312.3	305.4	299.0	290.2
Чернівецька	201.2	201.8	203.1	203.5	203.3	227.3	224.3	220.7	217.2	212.8
Чернігівська	399.5	395.8	393.5	390.4	385.0	267.2	259.8	252.6	245.3	236.9
Середня	440.5	441.5	443.1	444.3	442.4	425.2	415.3	403.8	392.6	379.6
Середньоквадратичне відхилення	225.2	225.8	227.0	227.9	226.9	196.2	188.6	179.9	171.9	163.8
Коефіцієнт варіації	0.511	0.511	0.512	0.513	0.513	0.461	0.454	0.445	0.438	0.431

Продовження додатку 5

Зміна RWAGE (грн)

Зміна RPCWARR (тис. грн на душу населення)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	51.3	133.8	153.3	154.8	165.8	0.0057	0.0700	0.0807	0.1234	0.0858
Вінницька	47.4	119.7	130.6	126.6	127.4	0.0191	0.0689	0.0839	0.1104	0.1205
Волинська	43.7	107.4	122.1	114.1	115.3	0.0170	0.0812	0.1021	0.1278	0.1057
Дніпропетровська	65.4	172.5	194.2	198.6	204.0	0.0169	0.0965	0.1343	0.1631	0.1548
Донецька	69.0	171.6	190.9	202.0	217.1	0.0168	0.0994	0.1444	0.1831	0.1974
Житомирська	49.0	121.9	130.8	128.5	130.5	0.0110	0.0722	0.0875	0.1204	0.1127
Закарпатська	38.5	97.6	110.6	116.5	127.6	0.0027	0.0271	0.0277	0.0425	0.0379
Запорізька	61.7	164.6	178.8	194.4	210.0	0.0128	0.0818	0.0981	0.1188	0.1186
Івано-Франківська	45.7	118.4	130.1	128.4	139.0	0.0062	0.0611	0.0782	0.0948	0.0852
Київська	60.8	158.0	176.7	173.1	184.6	0.0096	0.1154	0.1617	0.2152	0.1217
Кіровоградська	47.8	124.5	133.6	129.7	134.2	0.0150	0.0797	0.1184	0.1620	0.1620
Луганська	58.5	143.2	161.7	167.7	181.8	0.0233	0.0978	0.1576	0.2204	0.2167
Львівська	46.2	118.6	135.7	141.0	149.0	0.0013	0.0604	0.0770	0.1060	0.0810
Миколаївська	54.0	136.5	150.4	154.7	167.0	0.0108	0.0718	0.0915	0.1430	0.1567
Одеська	52.0	141.8	159.7	156.9	173.6	0.0002	0.0534	0.0574	0.0777	0.0824
Полтавська	54.5	147.2	161.2	160.5	169.2	0.0095	0.0623	0.0884	0.1373	0.1413
Рівненська	50.1	127.0	138.1	129.1	133.6	0.0134	0.0824	0.1016	0.1167	0.1070
Сумська	51.5	132.1	143.5	140.6	146.8	0.0210	0.0820	0.1127	0.1477	0.1534
Тернопільська	43.3	105.9	117.9	115.1	111.0	0.0134	0.0802	0.0987	0.1268	0.1403
Харківська	53.0	136.2	161.2	168.5	179.2	0.0147	0.0853	0.0871	0.1214	0.1458
Херсонська	45.7	118.5	134.7	134.6	140.5	0.0007	0.0657	0.0937	0.1333	0.1404
Хмельницька	47.5	117.5	128.0	125.3	122.7	0.0132	0.0692	0.0963	0.1312	0.1340
Черкаська	50.6	129.7	139.4	139.7	146.1	0.0090	0.0640	0.0854	0.1311	0.1442
Чернівецька	49.8	113.3	124.4	118.1	120.5	0.0064	0.0568	0.0744	0.0947	0.0783
Чернігівська	44.8	131.6	142.0	132.7	139.1	0.0103	0.0648	0.0929	0.1342	0.1208
Середня	51.3	131.6	146.0	146.1	153.4	0.011	0.074	0.097	0.131	0.126
Середньоквадратичне відхилення	7.1	19.3	21.9	25.6	29.6	0.006	0.017	0.029	0.038	0.038
Коефіцієнт варіації	0.138	0.147	0.150	0.176	0.193	0.546	0.236	0.296	0.286	0.305

Продовження додатку 5

Зміна РРСРАУ (тис. грн на душу населення) Зміна РРСГРАНТ (грн на душу населення)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	0.200	0.515	0.594	1.275	1.518	0.00	0.00	15.53	47.13	35.22
Вінницька	0.231	0.346	0.552	1.440	1.588	0.00	29.60	95.73	92.80	77.06
Волинська	0.127	0.299	0.460	1.020	1.491	38.58	45.02	119.94	122.70	136.93
Дніпропетровська	0.584	2.238	3.060	4.623	6.952	0.00	0.00	0.00	0.32	5.47
Донецька	0.788	1.361	2.036	4.275	5.925	0.00	0.00	0.00	0.27	7.87
Житомирська	0.184	0.322	0.408	0.985	1.118	62.37	48.92	158.37	161.26	149.87
Закарпатська	0.050	0.241	0.313	0.557	0.700	38.30	49.99	127.84	115.47	144.92
Запорізька	0.474	1.141	1.799	3.305	5.108	0.00	0.00	0.00	0.56	11.36
Івано-Франківська	0.487	0.732	0.907	1.255	1.822	3.30	0.00	58.71	59.14	66.21
Київська	0.281	0.859	1.056	2.173	2.805	4.70	0.00	43.85	33.64	84.37
Кіровоградська	0.166	0.439	0.668	1.445	2.591	20.47	48.22	100.05	102.85	125.94
Луганська	0.563	1.062	1.366	3.208	4.210	0.00	0.00	24.48	35.44	21.76
Львівська	0.521	1.122	1.438	2.178	2.728	0.00	0.00	28.88	23.20	27.65
Миколаївська	0.416	0.781	1.071	2.050	2.802	7.81	6.19	54.61	42.88	16.31
Одеська	0.195	0.624	0.476	1.093	1.782	0.00	0.00	0.00	9.15	13.11
Полтавська	0.375	0.839	1.560	2.352	2.404	0.00	0.00	0.00	0.18	0.37
Рівненська	0.267	0.587	0.884	1.561	2.306	11.45	16.95	68.49	89.17	48.00
Сумська	0.246	0.551	0.839	1.809	3.080	0.00	0.00	20.24	37.09	15.52
Тернопільська	0.113	0.280	0.416	1.119	1.451	9.96	38.93	112.54	110.36	123.67
Харківська	0.540	1.245	1.748	3.117	2.890	0.00	0.00	0.00	0.11	4.19
Херсонська	0.212	0.514	0.730	1.632	2.137	15.21	50.54	101.53	92.97	90.55
Хмельницька	0.210	0.488	0.740	1.425	1.662	12.09	29.27	92.54	99.98	75.26
Черкаська	0.261	0.551	0.882	1.452	1.637	19.00	0.00	60.39	96.44	55.38
Чернівецька	0.132	0.289	0.423	0.760	0.888	22.95	34.18	94.96	83.81	29.14
Чернігівська	0.203	0.427	0.468	1.059	1.280	0.00	0.00	61.49	60.72	32.65
Середня	0.313	0.714	0.996	1.887	2.515	10.648	15.912	57.606	60.706	55.952
Середньоквадратичне відхилення	0.181	0.445	0.638	1.042	1.520	15.518	20.209	46.959	45.685	47.788
Коефіцієнт варіації	0.578	0.623	0.641	0.552	0.604	1.457	1.270	0.815	0.753	0.854

Продовження додатку 5

Змінна РРСЕХР (грн на душу населення)

Змінна РРСРЕУ (грн на душу населення)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	235.61	277.73	331.99	310.17	502.54	218.94	265.39	185.98	194.51	292.46
Вінницька	134.76	208.52	298.35	296.28	409.08	120.63	166.46	131.64	127.28	149.29
Волинська	157.22	189.06	348.67	334.61	600.17	102.70	124.64	108.66	117.75	137.30
Дніпропетровська	221.16	236.69	507.69	349.64	494.26	188.32	217.15	329.98	298.24	261.16
Донецька	189.65	218.09	356.05	372.77	450.18	173.76	214.13	285.95	301.04	252.61
Житомирська	205.90	253.41	422.66	441.72	539.78	130.01	177.56	128.23	135.72	170.39
Закарпатська	134.69	186.74	353.94	333.27	590.56	83.37	125.55	88.77	117.36	174.14
Запорізька	188.48	210.32	327.95	388.38	551.76	174.32	207.37	278.61	323.36	296.34
Івано-Франківська	181.18	214.86	311.65	358.10	627.73	165.95	207.05	163.25	186.01	250.44
Київська	186.23	237.68	329.55	392.48	1004.37	157.57	221.04	170.46	297.20	425.68
Кіровоградська	179.61	243.54	317.08	323.38	583.25	123.71	183.21	132.75	126.61	163.25
Луганська	149.13	216.33	280.62	363.94	520.54	134.08	194.99	202.65	192.62	242.02
Львівська	212.49	198.67	262.31	284.33	600.62	212.84	179.68	186.48	210.04	268.61
Миколаївська	180.25	222.55	381.84	429.13	426.72	156.90	212.09	227.00	282.55	255.75
Одеська	149.50	199.29	291.64	293.30	373.05	147.14	196.34	238.66	248.32	278.14
Полтавська	178.15	234.62	445.73	568.86	548.14	163.91	231.28	342.46	470.07	329.40
Рівненська	145.70	178.80	311.03	379.65	654.99	128.84	159.08	153.98	168.58	335.65
Сумська	168.57	231.73	325.32	321.18	367.06	162.96	213.36	219.74	216.39	237.17
Тернопільська	135.24	185.14	324.59	310.94	531.57	107.45	133.05	107.22	99.16	123.94
Харківська	219.44	272.48	471.36	509.50	494.76	211.78	263.31	326.03	372.86	253.42
Херсонська	146.49	244.07	323.56	316.03	446.24	112.58	166.63	148.74	145.07	158.34
Хмельницька	156.04	230.91	309.93	336.53	466.31	130.41	183.99	143.70	157.87	189.84
Черкаська	111.57	224.97	414.54	435.34	567.27	177.75	217.13	216.22	205.50	278.34
Чернівецька	137.69	187.22	352.78	340.12	354.38	103.56	134.23	121.73	137.98	152.37
Чернігівська	164.00	222.33	346.11	301.62	520.31	156.86	210.70	182.16	159.33	218.67
Середня	174.750	221.030	349.877	363.651	529.025	149.855	192.216	192.841	211.656	235.789
Середньоквадратичне відхилення	29.761	25.778	58.900	67.802	126.030	35.648	37.393	71.798	90.250	71.172
Коефіцієнт варіації	0.170	0.117	0.168	0.186	0.238	0.238	0.195	0.372	0.426	0.302

Продовження додатку 5

Змінна RPSAGRO (тис. грн на душу населення) Змінна RPSIO (тис. грн на душу населення)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	0.5306	0.4717	0.4668	0.4286	0.3939	0.5176	0.5994	0.5524	0.5894	0.7869
Вінницька	1.0586	1.0256	0.9976	0.8953	0.7864	0.6283	0.8003	0.8090	0.8127	1.0319
Волинська	0.8335	0.8598	0.7390	0.7319	0.7390	0.3223	0.3795	0.3994	0.4097	0.5333
Дніпропетровська	0.4702	0.4018	0.4169	0.3748	0.3389	2.1205	2.7483	2.9123	3.3582	4.4731
Донецька	0.3012	0.2497	0.2609	0.2153	0.2312	2.2851	2.5877	2.6297	2.9552	3.9943
Житомирська	0.8768	0.8358	0.7479	0.6892	0.6621	0.5289	0.6525	0.6178	0.6204	0.7747
Закарпатська	0.4731	0.4760	0.4649	0.4228	0.4447	0.2051	0.2411	0.2379	0.2282	0.3329
Запорізька	0.5654	0.4319	0.5001	0.4388	0.4593	2.2388	2.8188	2.9541	3.2779	4.3674
Івано-Франківська	0.5673	0.5629	0.5463	0.4463	0.4669	1.0940	0.9981	0.9661	1.0014	1.1293
Київська	1.0782	0.9999	0.8908	0.8661	0.7603	0.9589	1.3157	1.0918	1.3260	1.6159
Кіровоградська	0.9608	0.7219	0.8609	0.8090	0.7016	0.4703	0.5299	0.4967	0.5126	0.6986
Луганська	0.3307	0.2860	0.3029	0.2630	0.2748	1.7276	1.8424	1.6421	1.7205	2.4732
Львівська	0.4698	0.4702	0.4490	0.4241	0.4352	0.6666	0.7526	0.7516	0.7678	1.0342
Миколаївська	0.6953	0.5488	0.6467	0.5541	0.5438	1.1393	1.1894	1.2468	1.2815	1.5271
Одеська	0.4938	0.4431	0.5957	0.4829	0.4954	0.4989	0.6444	0.5411	0.6970	0.8519
Полтавська	0.9863	0.8687	0.8514	0.7475	0.6509	1.3748	1.6815	1.7934	2.4417	3.0461
Рівненська	0.7788	0.7840	0.6832	0.6389	0.6406	0.7333	1.1272	1.1987	1.0894	1.3884
Сумська	0.8541	0.8265	0.7403	0.7060	0.5976	0.9348	1.2343	1.3567	1.3748	1.8890
Тернопільська	0.9258	0.9065	0.8095	0.6938	0.6783	0.3720	0.4810	0.4728	0.3952	0.4745
Харківська	0.5382	0.4733	0.5098	0.4309	0.4087	0.8465	1.1095	1.3227	1.5716	2.1393
Херсонська	0.7303	0.6613	0.8138	0.7644	0.7088	0.5229	0.4696	0.4337	0.4568	0.5188
Хмельницька	0.9709	0.9174	0.8688	0.7542	0.7397	0.7139	0.8602	0.8983	0.8930	1.0960
Черкаська	0.9742	0.9164	0.8952	0.8446	0.7381	0.7166	0.8198	0.8501	0.9375	1.2400
Чернівецька	0.6572	0.5801	0.6062	0.5157	0.5231	0.3537	0.3984	0.4133	0.4074	0.4528
Чернігівська	1.1004	1.0584	0.9118	0.8863	0.7965	0.6310	0.8163	0.8209	0.8365	1.0979
Середня	0.729	0.671	0.663	0.601	0.569	0.902	1.084	1.096	1.199	1.559
Середньоквадратичне відхилення	0.239	0.237	0.201	0.197	0.163	0.585	0.715	0.752	0.885	1.193
Коефіцієнт варіації	0.328	0.353	0.303	0.327	0.287	0.648	0.660	0.686	0.738	0.765

Продовження додатку 5

Зміна РОР (тис. чол.)

Зміна WARR (млн грн)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	2,192.0	2,177.2	2,150.6	2,129.3	2,106.3	12.5	152.3	173.5	262.8	180.8
Вінницька	1,877.0	1,864.2	1,850.4	1,835.3	1,819.7	35.9	128.5	155.2	202.7	219.3
Волинська	1,075.0	1,072.4	1,069.0	1,065.1	1,061.1	18.3	87.1	109.2	136.2	112.2
Дніпропетровська	3,877.0	3,840.6	3,799.8	3,764.0	3,733.6	65.6	370.8	510.4	614.0	578.1
Донецька	5,246.0	5,177.9	5,104.8	5,043.8	4,987.3	88.4	514.8	737.0	923.5	984.7
Житомирська	1,485.0	1,472.8	1,460.1	1,449.3	1,437.7	16.3	106.3	127.7	174.5	162.1
Закарпатська	1,281.0	1,281.4	1,281.9	1,281.5	1,280.7	3.4	34.7	35.5	54.5	48.6
Запорізька	2,087.0	2,070.0	2,051.4	2,034.7	2,016.0	26.7	169.4	201.3	241.7	239.0
Івано-Франківська	1,456.0	1,456.8	1,455.3	1,453.3	1,450.3	9.1	89.0	113.8	137.8	123.6
Київська	1,906.0	1,890.2	1,874.8	1,858.4	1,844.0	18.2	218.2	303.2	399.9	224.4
Кіровоградська	1,224.0	1,213.5	1,199.9	1,186.5	1,172.5	18.4	96.7	142.0	192.2	190.0
Луганська	2,821.0	2,782.8	2,737.3	2,700.7	2,668.1	65.7	272.1	431.4	595.3	578.2
Львівська	2,750.0	2,741.2	2,730.3	2,719.3	2,708.3	3.6	165.5	210.3	288.4	219.5
Миколаївська	1,349.0	1,341.0	1,329.8	1,320.2	1,307.6	14.6	96.3	121.7	188.8	204.9
Одеська	2,588.0	2,568.1	2,548.4	2,529.4	2,510.2	0.5	137.1	146.3	196.6	206.8
Полтавська	1,748.0	1,734.8	1,719.6	1,704.0	1,688.8	16.6	108.0	152.0	234.0	238.7
Рівненська	1,189.0	1,188.8	1,187.1	1,186.8	1,184.9	16.0	98.0	120.6	138.5	126.8
Сумська	1,405.0	1,392.7	1,379.1	1,364.6	1,349.3	29.5	114.2	155.4	201.6	207.0
Тернопільська	1,172.0	1,170.5	1,167.4	1,163.5	1,159.0	15.8	93.9	115.3	147.6	162.6
Харківська	3,103.0	3,068.1	3,034.8	3,004.1	2,977.6	45.6	261.6	264.3	364.7	434.1
Херсонська	1,272.0	1,262.7	1,252.1	1,243.8	1,234.1	0.9	83.0	117.3	165.8	173.3
Хмельницька	1,511.0	1,503.2	1,492.8	1,480.1	1,468.4	20.0	104.0	143.7	194.1	196.7
Черкаська	1,513.0	1,500.4	1,486.9	1,474.5	1,459.5	13.7	96.1	127.0	193.3	210.5
Чернівецька	948.0	946.4	943.6	941.3	938.2	6.1	53.8	70.2	89.1	73.5
Чернігівська	1,364.0	1,346.4	1,330.3	1,315.4	1,299.8	14.1	87.3	123.6	176.5	157.0
Середня	1,937.6	1,922.6	1,905.5	1,890.0	1,874.5	23.0	149.5	196.3	260.6	250.1
Середньоквадратичне відхилення	981.2	967.4	952.7	940.5	930.0	21.4	104.9	151.2	188.9	196.2
Коефіцієнт варіації	0.506	0.503	0.500	0.498	0.496	0.931	0.702	0.770	0.725	0.734

Продовження додатку 5

Зміна РАУ (млн грн)

Зміна GRANT (тис. грн)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	438.0	1,121.1	1,276.7	2,715.0	3,198.4	0.0	0.0	33,393.0	100,361.7	74,186.9
Вінницька	434.0	644.1	1,020.6	2,642.0	2,890.1	0.0	55,176.0	177,148.0	170,321.6	140,226.6
Волинська	136.0	320.6	491.4	1,086.0	1,582.0	41,478.0	48,282.0	128,211.0	130,689.9	145,301.0
Дніпропетровська	2,263.0	8,593.5	11,625.5	17,400.0	25,955.7	0.0	0.0	0.0	1,189.0	20,407.5
Донецька	4,134.0	7,049.3	10,395.6	21,562.0	29,551.3	0.0	0.0	0.0	1,376.0	39,230.3
Житомирська	273.0	474.5	595.8	1,427.0	1,607.3	92,622.0	72,053.0	231,234.0	233,707.1	215,472.6
Закарпатська	64.0	309.1	400.8	714.0	896.0	49,067.6	64,053.0	163,873.0	147,971.9	185,594.1
Запорізька	990.0	2,361.5	3,690.1	6,724.0	10,297.7	0.0	0.0	0.0	1,132.4	22,896.9
Івано-Франківська	709.0	1,066.3	1,320.1	1,824.0	2,642.7	4,805.0	0.0	85,445.0	85,952.4	96,018.9
Київська	536.0	1,624.3	1,980.6	4,038.0	5,171.5	8,959.0	0.0	82,201.0	62,512.3	155,583.5
Кіровоградська	203.0	532.3	801.3	1,714.0	3,037.7	25,057.0	58,517.0	120,046.0	122,028.9	147,670.0
Луганська	1,588.0	2,955.1	3,739.7	8,664.0	11,232.5	0.0	0.0	66,999.0	95,716.3	58,065.6
Львівська	1,434.0	3,076.3	3,925.6	5,923.0	7,389.2	0.0	0.0	78,859.0	63,094.3	74,889.1
Миколаївська	561.0	1,047.1	1,424.4	2,707.0	3,663.4	10,535.0	8,300.0	72,623.0	56,606.7	21,323.2
Одеська	505.0	1,602.3	1,213.3	2,764.0	4,472.7	0.0	0.0	0.0	23,156.5	32,901.8
Полтавська	656.0	1,455.9	2,683.1	4,008.0	4,060.1	0.0	0.0	0.0	313.0	622.7
Рівненська	318.0	697.9	1,049.8	1,853.0	2,732.9	13,611.0	20,148.0	81,304.0	105,829.8	56,880.9
Сумська	346.0	767.5	1,156.7	2,468.0	4,156.0	0.0	0.0	27,916.0	50,611.5	20,938.6
Тернопільська	133.0	328.1	485.2	1,302.0	1,681.4	11,671.0	45,567.0	131,376.0	128,403.2	143,336.9
Харківська	1,676.0	3,820.0	5,304.5	9,363.0	8,606.5	0.0	0.0	0.0	326.0	12,479.8
Херсонська	270.0	649.5	914.2	2,030.0	2,631.2	19,345.0	63,813.0	127,122.0	115,633.9	111,749.0
Хмельницька	317.0	734.0	1,104.0	2,109.0	2,441.2	18,274.0	44,004.0	138,138.0	147,975.7	110,517.5
Черкаська	395.0	826.3	1,311.6	2,141.0	2,389.1	28,751.0	0.0	89,795.0	142,202.1	80,827.3
Чернівецька	125.0	273.7	398.7	715.0	833.4	21,754.8	32,345.0	89,601.0	78,888.7	27,341.7
Чернігівська	277.0	575.5	622.3	1,393.0	1,664.3	0.0	0.0	81,798.0	79,872.0	42,441.1
Середня	751.2	1,716.2	2,357.3	4,371.4	5,791.6	13,837.2	20,490.3	80,283.3	85,834.9	81,476.1
Середньоквадратичне відхилення	881.8	2,034.1	2,848.8	5,025.8	7,032.2	21,063.6	26,232.6	61,724.3	59,967.1	59,188.7
Коефіцієнт варіації	1.174	1.185	1.209	1.150	1.214	1.522	1.280	0.769	0.699	0.726

Продовження додатку 5

Зміна ЕХР (тис. грн)

Зміна OWNREV (тис. грн)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	516,460.8	604,680.0	713,971.0	660,444.5	1,058,494.0	479,926.1	577,799.0	399,975.0	414,170.4	615,998.9
Вінницька	252,998.8	388,714.0	552,058.0	543,762.3	744,409.9	226,430.0	310,322.0	243,589.0	233,592.0	271,661.2
Волинська	169,012.3	202,746.0	372,724.0	356,388.4	636,837.2	110,400.7	133,660.0	116,158.0	125,418.1	145,686.3
Дніпропетровська	857,428.6	909,035.0	1,929,103.0	1,316,053.6	1,845,385.5	730,125.6	833,976.0	1,253,850.0	1,122,557.3	975,057.2
Донецька	994,918.4	1,129,231.0	1,817,541.0	1,880,199.5	2,245,188.5	911,526.8	1,108,725.0	1,459,721.0	1,518,409.6	1,259,846.1
Житомирська	305,768.7	373,228.0	617,122.0	640,185.6	776,036.7	193,088.8	261,508.0	187,234.0	196,692.3	244,968.5
Закарпатська	172,532.8	239,287.0	453,722.0	427,080.2	756,328.3	106,794.3	160,876.0	113,792.0	150,403.1	223,022.3
Запорізька	393,366.8	435,370.0	672,760.0	790,242.9	1,112,351.3	363,803.8	429,248.0	571,588.0	657,943.5	597,421.1
Івано-Франківська	263,795.7	313,002.0	453,543.0	520,424.4	910,390.0	241,617.2	301,636.0	237,573.0	270,324.9	363,212.5
Київська	354,953.3	449,258.0	617,847.0	729,382.9	1,852,049.3	300,330.7	417,810.0	319,571.0	552,317.2	784,949.5
Кіровоградська	219,843.0	295,536.0	380,461.0	383,693.8	683,861.0	151,424.5	222,322.0	159,292.0	150,217.9	191,411.6
Луганська	420,683.2	602,009.0	768,150.0	982,904.1	1,388,858.6	378,248.8	542,626.0	554,722.0	520,215.4	645,722.7
Львівська	584,352.0	544,598.0	716,181.0	773,182.7	1,626,652.3	585,300.3	492,531.0	509,138.0	571,168.4	727,482.2
Миколаївська	243,163.4	298,444.0	507,769.0	566,540.7	557,975.4	211,656.0	284,413.0	301,859.0	373,018.0	334,423.5
Одеська	386,903.0	511,799.0	743,219.0	741,869.1	936,425.4	380,804.9	504,223.0	608,192.0	628,104.3	698,183.7
Полтавська	311,403.0	407,016.0	766,477.0	969,332.3	925,692.9	286,521.9	401,229.0	588,899.0	800,995.3	556,293.7
Рівненська	173,242.7	212,554.0	369,227.0	450,570.6	776,099.9	153,192.6	189,117.0	182,787.0	200,065.9	397,715.9
Сумська	236,837.0	322,728.0	448,651.0	438,282.9	495,270.9	228,987.1	297,151.0	303,038.0	295,284.8	320,018.1
Тернопільська	158,497.0	216,707.0	378,925.0	361,774.1	616,086.9	125,929.0	155,735.0	125,167.0	115,371.7	143,647.7
Харківська	680,924.0	835,986.0	1,430,475.0	1,530,577.1	1,473,192.8	657,147.6	807,849.0	989,421.0	1,120,094.4	754,582.8
Херсонська	186,341.0	308,186.0	405,130.0	393,081.0	550,710.4	143,207.0	210,399.0	186,237.0	180,434.4	195,413.3
Хмельницька	235,775.0	347,108.0	462,659.0	498,091.5	684,734.3	197,055.3	276,570.0	214,510.0	233,670.3	278,765.7
Черкаська	320,110.0	337,538.0	616,381.0	641,911.8	827,932.6	268,942.4	325,780.0	321,503.0	303,016.9	406,240.5
Чернівецька	130,528.0	177,188.0	332,879.0	320,154.9	332,478.4	98,177.9	127,040.0	114,867.0	129,878.0	142,950.2
Чернігівська	223,697.0	299,345.0	460,431.0	396,757.5	676,300.9	213,963.3	283,685.0	242,321.0	209,580.8	284,229.2
Середня	351,739.0	430,451.5	679,496.2	692,515.5	979,589.7	309,732.1	386,249.2	412,198.2	442,917.8	462,356.2
Середньоквадрантальне відхилення	216,536.7	229,923.3	415,499.3	380,601.8	478,237.4	207,788.5	236,150.9	346,251.3	353,052.7	283,350.2
Коефіцієнт варіації	0.616	0.534	0.611	0.550	0.488	0.671	0.611	0.840	0.808	0.613

Продовження додатку 5

Зміна ARGO (млн грн)

Зміна INDOUT (млн грн)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	1,163.0	1,027.0	1,004.0	912.7	829.6	1,134.5	1,305.0	1,188.0	1,255.0	1,657.4
Вінницька	1,987.0	1,912.0	1,846.0	1,643.1	1,431.1	1,179.3	1,492.0	1,497.0	1,491.6	1,877.8
Волинська	896.0	922.0	790.0	779.5	784.2	346.5	407.0	427.0	436.4	565.9
Дніпропетровська	1,823.0	1,543.0	1,584.0	1,410.7	1,265.4	8,221.3	10,555.0	11,066.0	12,640.4	16,700.8
Донецька	1,580.0	1,293.0	1,332.0	1,085.8	1,153.1	11,725.5	13,399.0	13,424.0	14,905.2	19,920.9
Житомирська	1,302.0	1,231.0	1,092.0	998.8	951.9	782.4	961.0	902.0	899.1	1,113.8
Закарпатська	606.0	610.0	596.0	541.9	569.5	262.7	309.0	305.0	292.5	426.4
Запорізька	1,180.0	894.0	1,026.0	892.9	925.9	4,672.3	5,835.0	6,060.0	6,669.5	8,804.6
Івано-Франківська	826.0	820.0	795.0	648.7	677.2	1,592.8	1,454.0	1,406.0	1,455.4	1,637.8
Київська	2,055.0	1,890.0	1,670.0	1,609.6	1,402.0	1,827.6	2,487.0	2,047.0	2,464.3	2,979.8
Кіровоградська	1,176.0	876.0	1,033.0	959.9	822.6	575.7	643.0	596.0	608.2	819.1
Луганська	933.0	796.0	829.0	710.4	733.1	4,873.5	5,127.0	4,495.0	4,646.6	6,598.7
Львівська	1,292.0	1,289.0	1,226.0	1,153.3	1,178.7	1,833.2	2,063.0	2,052.0	2,088.0	2,801.0
Миколаївська	938.0	736.0	860.0	731.6	711.1	1,536.9	1,595.0	1,658.0	1,691.8	1,996.9
Одеська	1,278.0	1,138.0	1,518.0	1,221.5	1,243.5	1,291.0	1,655.0	1,379.0	1,762.9	2,138.4
Полтавська	1,724.0	1,507.0	1,464.0	1,273.7	1,099.2	2,403.2	2,917.0	3,084.0	4,160.6	5,144.3
Рівненська	926.0	932.0	811.0	758.2	759.0	871.9	1,340.0	1,423.0	1,292.9	1,645.1
Сумська	1,200.0	1,151.0	1,021.0	963.4	806.4	1,313.4	1,719.0	1,871.0	1,876.1	2,548.8
Тернопільська	1,085.0	1,061.0	945.0	807.2	786.2	435.9	563.0	552.0	459.8	550.0
Харківська	1,670.0	1,452.0	1,547.0	1,294.5	1,216.8	2,626.8	3,404.0	4,014.0	4,721.2	6,370.0
Херсонська	929.0	835.0	1,019.0	950.8	874.7	665.1	593.0	543.0	568.2	640.3
Хмельницька	1,467.0	1,379.0	1,297.0	1,116.3	1,086.2	1,078.7	1,293.0	1,341.0	1,321.8	1,609.4
Черкаська	1,474.0	1,375.0	1,331.0	1,245.4	1,077.3	1,084.2	1,230.0	1,264.0	1,382.3	1,809.8
Чернівецька	623.0	549.0	572.0	485.5	490.8	385.3	377.0	390.0	383.5	424.8
Чернігівська	1,501.0	1,425.0	1,213.0	1,165.9	1,035.3	860.7	1,099.0	1,092.0	1,100.3	1,427.0
Середня	1,265.4	1,145.7	1,136.8	1,014.4	956.4	2,141.2	2,552.9	2,563.0	2,822.9	3,688.4
Середньонадратичне відхилення	357.6	354.8	332.0	297.6	247.5	2,623.7	3,109.7	3,174.7	3,593.9	4,810.1
Коефіцієнт варіації	0.306	0.310	0.292	0.293	0.259	1.225	1.218	1.239	1.273	1.304

Продовження додатку 5

Змінна SMEXP (тис. грн)

Змінна RPSEXP (грн)

Область	1995	1996	1997	1998	1999	1995	1996	1997	1998	1999
АР Крим	13,107.2	29,106.0	21,716.0	25,991.4	29,850.0	5.98	13.37	10.10	12.21	14.17
Вінницька	10,317.3	18,095.0	11,645.0	11,590.8	16,171.1	5.50	9.71	6.29	6.32	8.89
Волинська	6,517.5	9,650.0	7,126.0	6,385.2	10,022.3	6.06	9.00	6.67	5.99	9.45
Дніпропетровська	17,822.6	28,128.0	24,421.0	25,849.9	35,978.5	4.60	7.32	6.43	6.87	9.64
Донецька	25,878.7	35,027.0	31,026.0	33,954.6	41,577.8	4.93	6.76	6.08	6.73	8.34
Житомирська	10,757.2	9,798.0	10,686.0	9,849.3	15,611.0	7.24	6.65	7.32	6.80	10.86
Закарпатська	6,267.4	11,646.0	9,203.0	9,561.5	14,860.5	4.89	9.09	7.18	7.46	11.60
Запорізька	11,191.2	27,305.0	15,400.0	16,960.5	19,229.8	5.36	13.19	7.51	8.34	9.54
Івано-Франківська	8,437.9	10,255.0	11,106.0	12,529.9	16,364.8	5.80	7.04	7.63	8.62	11.28
Київська	13,428.5	8,351.0	18,362.0	20,601.5	29,865.5	7.05	4.42	9.79	11.09	16.20
Кіровоградська	15,015.5	89,875.0	18,721.0	9,333.5	14,230.0	6.91	9.10	8.43	7.87	12.14
Луганська	13,554.1	18,968.0	13,463.0	16,247.0	26,938.3	5.32	32.30	6.84	7.23	10.10
Львівська	8,144.2	15,408.0	11,939.0	10,328.5	11,334.2	4.93	6.92	4.93	5.97	9.13
Миколаївська	13,073.0	30,472.0	17,838.0	21,011.6	26,607.5	6.04	11.49	8.98	7.82	8.67
Одеська	12,237.7	16,799.0	13,779.0	17,555.3	17,571.7	5.05	11.87	7.00	8.31	10.60
Полтавська	6,930.9	5,004.0	8,572.0	8,525.4	11,746.5	7.00	9.68	8.01	10.30	10.40
Рівненська	8,179.3	8,777.0	11,363.0	11,567.3	14,010.8	5.83	4.21	7.22	7.18	9.91
Сумська	6,962.2	5,186.0	8,101.0	7,439.3	10,144.1	5.82	6.30	8.24	8.48	10.38
Тернопільська	18,559.1	46,055.0	21,074.0	23,366.1	23,370.2	5.94	4.43	6.94	6.39	8.75
Харківська	7,482.2	10,766.0	9,538.0	7,691.8	9,201.6	5.98	15.01	6.94	7.78	7.85
Херсонська	10,585.6	21,671.0	8,844.0	8,986.7	12,413.3	5.88	8.53	7.62	6.18	7.46
Хмельницька	9,739.3	12,653.0	11,004.0	10,718.4	15,774.7	7.01	14.42	5.92	6.07	8.45
Черкаська	4,559.9	8,975.0	6,974.0	7,537.2	8,449.6	6.44	8.43	7.40	7.27	10.81
Чернівецька	11,161.6	13,700.0	12,628.0	10,947.6	13,132.1	4.81	9.48	7.39	8.01	9.01
Чернігівська	11,134.7	20,108.5	13,786.0	14,561.9	18,767.0	8.18	10.18	9.49	8.32	10.10
Середня						5.942	9.956	7.454	7.744	10.149
Середньоквадратичне відхилення	4,578.7	17,436.0	5,857.1	7,087.2	8,635.7	0.881	5.415	1.196	1.535	1.890
Коефіцієнт варіації	0.411	0.367	0.425	0.487	0.460	0.148	0.544	0.160	0.198	0.186

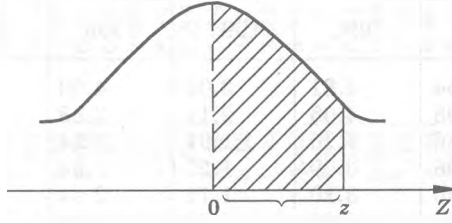
Закінчення додатку 5

Змінна SMSHARE (%)

Область	1995	1996	1997	1998	1999
АР Крим	2.54	4.81	3.04	4.00	5.82
Вінницька	4.08	4.66	2.11	2.83	4.46
Волинська	3.86	4.76	1.91	2.24	3.90
Дніпропетровська	2.08	3.09	1.27	2.53	3.25
Донецька	2.60	3.10	1.71	2.44	2.77
Житомирська	3.52	2.63	1.73	2.18	3.62
Закарпатська	3.63	4.87	2.03	2.19	5.06
Запорізька	2.84	6.27	2.29	2.31	2.93
Івано-Франківська	3.20	3.28	2.45	2.70	4.61
Київська	3.78	1.86	2.97	2.23	4.99
Кіровоградська	3.85	3.74	2.66	2.91	5.35
Луганська	3.57	14.93	2.44	2.83	4.68
Львівська	2.32	3.48	1.88	2.11	3.95
Миколаївська	3.35	5.16	2.35	2.60	2.74
Одеська	3.38	5.95	2.40	2.84	4.21
Полтавська	3.93	4.13	1.80	2.96	2.21
Рівненська	4.00	2.35	2.32	1.87	3.97
Сумська	3.45	2.72	2.53	3.41	4.05
Тернопільська	4.39	2.39	2.14	2.58	4.00
Харківська	2.73	5.51	1.47	2.39	2.11
Херсонська	4.02	3.49	2.35	2.58	3.14
Хмельницька	4.49	6.24	1.91	2.28	3.25
Черкаська	3.04	3.75	1.79	2.22	3.97
Чернівецька	3.49	5.07	2.10	3.99	4.07
Чернігівська	4.99	4.58	2.74	3.43	4.64
Середня	3.49	4.51	2.18	2.67	3.91
Середньоквадратичне відхилення	0.69	2.45	0.43	0.54	0.93
Коефіцієнт варіації	0.198	0.544	0.198	0.203	0.238

СТАТИСТИЧНІ ТАБЛИЦІ

Графік і таблиця нормального закону розподілу



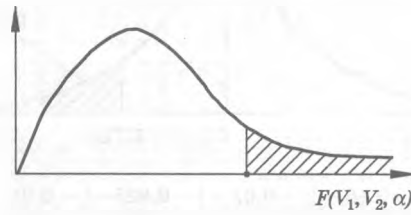
Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

За таблицею знаходяться ймовірності значень стандартної нормально розподіленої випадкової величини в межах від нуля до z . Функція щільності нормального закону розподілу симетрична відносно осі $z = 0$, тому вся площа під кривою зліва від нуля також дорівнює 0.5. Це дозволяє шляхом досить простих перетворень отримувати ймовірності, не передбачені в наведеній таблиці.

$$\text{Приклад. } P(z \leq 1.84) = 0.5 + P(0 \leq z \leq 1.84) = 0.5 + 0.4671 = 0.9671$$

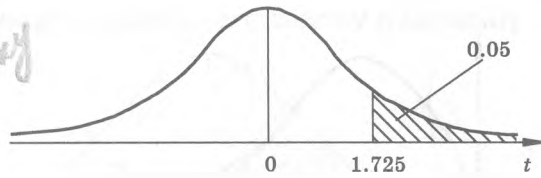
$$P(z > 1.64) = 0.5 - P(0 \leq z \leq 1.64) = 0.5 - 0.45 = 0.05$$

Графік і таблиця F-розподілу Фішера



v_2/v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.8	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.99	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.64	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.28	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.392	1.32	1.22	1.00

Примітка. За таблицею знаходяться критичні значення F-розподілу Фішера; v_1 — ступені свободи для чисельника; v_2 — ступені свободи для знаменника; α — рівень значимості.

Графік і таблиця t -розподілу Ст'юдента

df/Pr	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.861	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.227	3.143	3.707	3.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.434	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.081	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.167	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Приклад. Для $df = 20$ (df — ступені свободи)

$Pr(t > 2,089) = 0,025,$

$Pr(t > 1,725) = 0,05,$

$Pr(|t| > 1,725) = 0,10.$

DW-статистика.

Критичні точки d_L та d_U при рівні значимості $\alpha = 0.05$

n	p = 1		p = 2		p = 3		p = 4		p = 5		p = 6		p = 7		p = 8		p = 9		p = 10		
	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
7	0.700	1.356	0.467	1.896	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
8	0.763	1.332	0.559	1.777	0.368	2.287	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	---	---	---	---	---	---	---	---	---	---	---	---	---
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	---	---	---	---	---	---	---	---	---	---	---
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005	---	---	---	---	---	---	---	---	---
12	0.971	1.311	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149	---	---	---	---	---	---	---
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266	---	---	---	---	---
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.290	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	---	---	---
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.226	0.447	2.472	0.343	2.727	0.251	3.079	0.175	3.216	0.111	3.438	---
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	1.624	0.304	2.860	0.222	3.090	0.155	3.304	---
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	1.537	0.356	2.677	0.272	2.975	0.198	3.184	---
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	1.407	0.407	2.667	0.321	2.873	0.244	3.073	---
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	1.396	0.456	2.589	0.369	2.783	0.290	2.974	---
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	1.389	0.502	2.521	0.416	2.704	0.336	2.885	---
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	1.290	0.547	2.460	0.461	2.633	0.380	2.806	---
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	1.246	0.588	2.407	0.504	2.571	0.424	2.734	---
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	1.208	0.632	2.360	0.545	2.514	0.465	2.670	---
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	1.174	0.666	2.318	0.584	2.464	0.506	2.613	---
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	1.144	0.702	2.280	0.621	2.419	0.544	2.560	---
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	1.117	0.735	2.246	0.657	2.379	0.581	2.513	---
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	1.093	0.767	2.216	0.691	2.342	0.616	2.470	---
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	1.071	0.798	2.188	0.723	2.309	0.650	2.431	---
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.875	1.944	0.900	1.052	0.826	2.164	0.753	2.278	0.682	2.386	---
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.921	0.926	1.034	0.854	2.141	0.782	2.251	0.712	2.363	---
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.823	1.020	1.920	0.950	1.018	0.879	2.120	0.810	2.226	0.741	2.333	---
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	1.004	0.904	2.102	0.836	2.203	0.769	2.306	---
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	0.991	0.927	2.085	0.861	2.181	0.795	2.281	---
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	0.979	0.950	2.069	0.885	2.162	0.821	2.257	---
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	0.967	0.971	2.054	0.908	2.144	0.845	2.236	---
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	0.957	0.991	2.041	0.930	2.127	0.868	2.216	---
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	0.948	1.011	2.029	0.951	2.112	0.891	2.198	---
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	0.939	1.029	2.017	0.970	2.098	0.912	2.180	---
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	0.932	1.047	2.007	0.990	2.085	0.932	2.164	---
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	0.924	1.064	1.997	1.008	2.072	0.952	2.149	---
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	0.895	1.139	1.958	1.039	2.022	1.038	2.088	---
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	0.875	1.201	1.930	1.056	1.986	1.110	2.044	---
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	0.861	1.253	1.909	1.099	1.912	1.170	2.010	---
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	0.850	1.285	1.894	1.126	1.939	1.222	1.984	---
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	0.843	1.336	1.882	1.301	1.923	1.266	1.964	---
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	0.837	1.369	1.873	1.337	1.910	1.305	1.948	---
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.760	1.458	1.801	1.428	0.834	1.399	1.867	1.369	1.901	1.339	1.935	---
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	0.831	1.425	1.861	1.397	1.893	1.369	1.925	---
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.518	1.801	1.474	0.829	1.448	1.857	1.422	1.886	1.396	1.916	---
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	0.827	1.469	1.854	1.445	1.881	1.420	1.909	---
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	0.827	1.489	1.852	1.465	1.877	1.442	1.903	---
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.569	1.780	1.550	1.803	1.528	0.826	1.506	1.852	1.484	1.874	1.462	1.898	---
150	1.720	1.746	1.706	1.789	1.693	1.774	1.679	1.788	1.665	1.802	1.617	1.817	1.637	0.832	1.622	1.847	1.608	1.862	1.554	1.877	---
200	1.758	1.778	1.748	1.799	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	0.841	1.686	1.852	1.675	1.863	1.664	1.874	---

n	$p = 11$		$p = 12$		$p = 13$		$p = 14$		$p = 15$		$p = 16$		$p = 17$		$p = 18$		$p = 19$		$p = 20$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
16	0.098	3.503	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
17	0.188	3.378	0.087	3.557	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
18	0.177	3.265	0.123	3.441	0.078	3.603	---	---	---	---	---	---	---	---	---	---	---	---	---	---
19	0.220	3.159	0.160	3.335	0.111	3.496	0.070	3.642	---	---	---	---	---	---	---	---	---	---	---	---
20	0.263	3.063	0.200	3.234	0.145	3.395	0.100	3.542	0.063	3.676	---	---	---	---	---	---	---	---	---	---
21	0.307	2.976	0.240	3.141	0.182	3.300	0.132	3.448	0.091	3.583	---	---	---	---	---	---	---	---	---	---
22	0.349	2.897	0.281	3.057	0.220	3.211	0.166	3.358	0.120	3.495	0.058	3.619	0.052	3.731	---	---	---	---	---	---
23	0.391	2.826	0.322	2.979	0.259	3.128	0.202	3.272	0.153	3.409	0.110	3.535	0.076	3.650	0.048	3.753	---	---	---	---
24	0.431	2.761	0.362	2.908	0.297	3.053	0.239	3.193	0.186	3.327	0.141	3.454	0.101	3.672	0.070	3.678	0.044	3.773	---	---
25	0.470	2.702	0.400	2.844	0.335	2.983	0.275	3.119	0.221	3.251	0.172	3.376	0.130	3.494	0.094	3.604	0.065	3.702	0.041	3.790
26	0.508	2.649	0.438	2.784	0.373	2.919	0.312	3.051	0.256	3.179	0.205	3.303	0.160	3.420	0.120	3.531	0.087	3.582	0.060	3.724
27	0.544	2.600	0.475	2.730	0.409	2.859	0.348	2.987	0.291	3.112	0.238	3.233	0.191	3.349	0.149	3.460	0.112	3.563	0.081	3.658
28	0.578	2.555	0.510	2.680	0.445	2.805	0.383	2.928	0.325	3.050	0.271	3.168	0.222	3.283	0.178	3.392	0.138	3.495	0.104	3.592
29	0.612	2.515	0.544	2.634	0.479	2.755	0.418	2.874	0.359	2.992	0.305	3.107	0.254	3.219	0.208	3.327	0.166	3.431	0.129	3.528
30	0.643	2.477	0.577	2.592	0.512	2.708	0.451	2.823	0.392	2.937	0.337	3.050	0.286	3.160	0.238	3.266	0.195	3.368	0.156	3.465
31	0.674	2.443	0.608	2.553	0.545	2.655	0.484	2.776	0.425	2.887	0.370	2.996	0.317	3.103	0.269	3.224	0.224	3.309	0.183	3.406
32	0.703	2.411	0.638	2.517	0.576	2.625	0.515	2.733	0.457	2.840	0.401	2.946	0.349	3.050	0.299	3.153	0.253	3.252	0.211	3.348
33	0.731	2.382	0.668	2.484	0.606	2.598	0.546	2.692	0.488	2.796	0.432	2.899	0.379	3.000	0.329	3.100	0.283	3.198	0.239	3.293
34	0.758	2.355	0.695	2.454	0.634	2.554	0.575	2.654	0.518	2.754	0.462	2.854	0.409	2.954	0.359	3.051	0.312	3.147	0.267	3.240
35	0.783	2.330	0.722	2.425	0.662	2.521	0.604	2.619	0.547	2.716	0.492	2.813	0.439	2.910	0.388	3.005	0.340	3.099	0.295	3.190
36	0.808	2.306	0.748	2.398	0.689	2.492	0.631	2.585	0.575	2.680	0.520	2.774	0.467	2.868	0.417	2.961	0.369	3.053	0.323	3.142
37	0.831	2.285	0.772	2.374	0.714	2.464	0.657	2.555	0.602	2.646	0.548	2.738	0.495	2.829	0.445	2.920	0.397	3.009	0.351	3.097
38	0.854	2.265	0.796	2.351	0.739	2.438	0.683	2.526	0.628	2.614	0.575	2.703	0.522	2.792	0.472	2.880	0.424	2.968	0.378	3.054
39	0.875	2.246	0.819	2.329	0.763	2.413	0.707	2.499	0.653	2.585	0.600	2.671	0.549	2.757	0.499	2.863	0.451	2.929	0.404	3.013
40	0.896	2.228	0.840	2.309	0.785	2.391	0.731	2.473	0.678	2.557	0.626	2.641	0.575	2.724	0.525	2.808	0.477	2.892	0.430	2.974
45	0.988	2.156	0.988	2.295	0.887	2.296	0.838	2.367	0.788	2.439	0.740	2.512	0.692	2.586	0.644	2.659	0.598	2.733	0.553	2.807
50	1.064	2.103	1.019	2.163	0.973	2.225	0.927	2.287	0.882	2.350	0.836	2.414	0.792	2.479	0.747	2.544	0.703	2.610	0.660	2.675
55	1.129	2.062	1.087	2.116	1.045	2.170	1.003	2.225	0.961	2.281	0.919	2.338	0.877	2.396	0.836	2.454	0.795	2.512	0.754	2.571
60	1.184	2.031	1.145	2.079	1.106	2.127	1.068	2.177	1.029	2.227	0.990	2.278	0.951	2.330	0.913	2.382	0.874	2.434	0.836	2.487
65	1.231	2.006	1.195	2.049	1.160	2.093	1.124	2.138	1.088	2.183	1.052	2.229	1.016	2.276	0.980	2.323	0.944	2.371	0.908	2.419
70	1.272	1.986	1.239	2.029	1.206	2.066	1.172	2.106	1.139	2.148	1.105	2.189	1.072	2.232	1.038	2.275	1.005	2.318	0.971	2.362
75	1.308	1.970	1.277	2.006	1.247	2.043	1.215	2.080	1.184	2.118	1.153	2.156	1.121	2.195	1.098	2.235	1.058	2.275	1.027	2.315
80	1.340	1.957	1.311	1.991	1.283	2.024	1.253	2.059	1.224	2.093	1.195	2.129	1.165	2.165	1.136	2.201	1.106	2.238	1.076	2.275
85	1.369	1.946	1.342	1.977	1.315	2.009	1.287	2.040	1.260	2.073	1.232	2.105	1.205	2.139	1.177	2.172	1.149	2.206	1.121	2.241
90	1.395	1.937	1.369	1.966	1.344	1.995	1.318	2.025	1.292	2.055	1.266	2.085	1.240	2.116	1.213	2.148	1.187	2.179	1.160	2.211
95	1.418	1.929	1.394	1.956	1.370	1.984	1.345	2.012	1.321	2.040	1.296	2.068	1.271	2.100	1.247	2.122	1.216	2.156	1.197	2.186
100	1.439	1.923	1.416	1.948	1.393	1.974	1.371	2.000	1.347	2.026	1.324	2.053	1.301	2.080	1.277	2.108	1.253	2.135	1.229	2.164
150	1.579	1.892	1.564	1.908	1.550	1.924	1.535	1.940	1.519	1.956	1.504	1.972	1.489	1.989	1.474	2.006	1.458	2.023	1.443	2.040
200	1.654	1.885	1.643	1.896	1.632	1.908	1.621	1.919	1.610	1.931	1.599	1.943	1.588	1.955	1.576	1.967	1.565	1.979	1.554	1.991

Примітка. n — кількість спостережень; p — кількість факторів у моделі.

DW-статистика.

Критичні точки d_L та d_U при рівні значимості $\alpha = 0.01$

n	p = 1		p = 2		p = 3		p = 4		p = 5		p = 6		p = 7		p = 8		p = 9		p = 10		
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	
6	0.390	1.142	0.294	1.676	2.102																
7	0.435	1.086	0.345	1.489	0.229	1.875															
8	0.497	1.003	0.408	1.389	0.279	1.875	0.183	2.433													
9	0.554	0.998	0.466	1.333	0.340	1.733	0.230	2.193	0.150	2.690											
10	0.604	1.001	0.519	1.297	0.396	1.640	0.286	2.030	0.193	2.453	0.124	2.892									
11	0.653	1.010	0.569	1.274	0.449	1.575	0.339	1.913	0.244	2.280	0.164	2.665	0.105	3.053							
12	0.697	1.023	0.616	1.261	0.499	1.526	0.391	1.826	0.294	2.150	0.211	2.490	0.140	2.838	0.090	3.182					
13	0.738	1.038	0.660	1.254	0.547	1.490	0.441	1.757	0.343	2.049	0.257	2.354	0.188	2.667	0.122	2.981	0.078				
14	0.776	1.054	0.700	1.252	0.591	1.464	0.488	1.704	0.391	1.967	0.308	2.244	0.236	2.530	0.161	2.817	0.107				
15	0.811	1.070	0.737	1.252	0.633	1.446	0.532	1.663	0.437	1.900	0.349	2.153	0.269	2.416	0.200	2.681	0.142	2.944			
16	0.844	1.086	0.773	1.252	0.672	1.432	0.574	1.630	0.480	1.847	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127		
17	0.874	1.102	0.805	1.255	0.708	1.422	0.613	1.604	0.522	1.803	0.435	2.015	0.355	2.238	0.282	2.467	0.216	2.697	0.160		
18	0.902	1.118	0.835	1.265	0.742	1.415	0.650	1.584	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.381	0.255	2.597	0.196		
19	0.928	1.132	0.863	1.271	0.773	1.411	0.685	1.567	0.598	1.737	0.515	1.918	0.436	2.110	0.362	2.308	0.294	2.510	0.232		
20	0.952	1.147	0.890	1.277	0.803	1.408	0.718	1.554	0.633	1.712	0.552	1.881	0.474	2.059	0.400	2.244	0.331	2.434	0.268		
21	0.975	1.161	0.914	1.284	0.831	1.407	0.748	1.543	0.667	1.691	0.587	1.849	0.510	2.015	0.437	2.188	0.368	2.367	0.304		
22	0.997	1.174	0.938	1.291	0.856	1.407	0.777	1.534	0.698	1.673	0.620	1.821	0.545	1.977	0.473	2.140	0.404	2.308	0.340		
23	1.018	1.189	0.960	1.298	0.882	1.407	0.805	1.528	0.728	1.658	0.652	1.797	0.578	1.944	0.507	2.097	0.439	2.255	0.375		
24	1.037	1.199	0.981	1.305	0.906	1.409	0.831	1.523	0.756	1.645	0.682	1.776	0.610	1.915	0.540	2.059	0.473	2.209	0.409		
25	1.055	1.211	0.981	1.312	0.928	1.411	0.855	1.518	0.783	1.635	0.711	1.759	0.640	1.889	0.572	2.026	0.505	2.168	0.441		
26	1.072	1.222	1.001	1.319	0.949	1.413	0.878	1.515	0.808	1.626	0.738	1.743	0.669	1.867	0.602	1.997	0.536	2.131	0.473		
27	1.089	1.233	1.019	1.319	0.968	1.415	0.900	1.513	0.835	1.618	0.768	1.729	0.696	1.847	0.630	1.970	0.566	2.098	0.504		
28	1.104	1.244	1.034	1.322	0.989	1.418	0.921	1.512	0.865	1.611	0.788	1.718	0.723	1.830	0.658	1.947	0.595	2.068	0.533		
29	1.119	1.254	1.054	1.332	1.006	1.421	0.941	1.511	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562		
30	1.133	1.263	1.070	1.339	1.006	1.425	0.960	1.510	0.897	1.601	0.834	1.698	0.772	1.800	0.710	1.906	0.649	2.017	0.589		
31	1.147	1.273	1.085	1.345	1.023	1.425	0.960	1.510	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615		
32	1.160	1.282	1.100	1.352	1.040	1.428	0.979	1.510	0.917	1.597	0.856	1.690	0.794	1.788	0.734	1.889	0.674	1.995	0.615		
33	1.172	1.291	1.114	1.358	1.055	1.432	0.996	1.510	0.936	1.594	0.876	1.683	0.816	1.776	0.757	1.874	0.698	1.975	0.641		
34	1.184	1.299	1.128	1.364	1.070	1.435	1.012	1.511	0.954	1.591	0.896	1.677	0.837	1.766	0.779	1.860	0.722	1.957	0.665		
35	1.195	1.307	1.140	1.370	1.085	1.439	1.028	1.512	0.971	1.589	0.914	1.671	0.857	1.757	0.800	1.847	0.744	1.940	0.689		
36	1.206	1.315	1.153	1.376	1.098	1.442	1.043	1.513	0.988	1.588	0.932	1.666	0.879	1.742	0.821	1.836	0.766	1.925	0.711		
37	1.217	1.323	1.165	1.382	1.112	1.446	1.058	1.514	1.004	1.586	0.950	1.662	0.895	1.742	0.841	1.825	0.787	1.911	0.733		
38	1.227	1.330	1.176	1.388	1.124	1.449	1.072	1.515	1.019	1.585	0.966	1.658	0.913	1.735	0.860	1.816	0.807	1.899	0.754		
39	1.237	1.337	1.187	1.393	1.137	1.453	1.085	1.516	1.034	1.584	0.982	1.655	0.930	1.729	0.878	1.807	0.826	1.887	0.774		
40	1.246	1.344	1.198	1.398	1.148	1.457	1.098	1.518	1.048	1.584	0.997	1.652	0.946	1.724	0.895	1.799	0.844	1.876	0.749		
45	1.288	1.376	1.245	1.423	1.201	1.474	1.156	1.528	1.111	1.584	1.065	1.643	1.011	1.704	0.974	1.768	0.927	1.834	0.881		
50	1.324	1.403	1.285	1.446	1.254	1.491	1.205	1.538	1.164	1.587	1.123	1.639	1.081	1.692	1.039	1.748	0.997	1.805	0.955		
55	1.356	1.427	1.320	1.466	1.284	1.506	1.247	1.548	1.209	1.592	1.172	1.638	1.134	1.685	1.095	1.748	1.057	1.785	1.018		
60	1.383	1.449	1.350	1.484	1.317	1.520	1.283	1.568	1.249	1.598	1.214	1.639	1.179	1.682	1.144	1.726	1.108	1.771	1.072		
65	1.407	1.468	1.377	1.500	1.346	1.534	1.315	1.568	1.283	1.604	1.251	1.642	1.218	1.680	1.186	1.720	1.153	1.761	1.120		
70	1.429	1.485	1.400	1.515	1.372	1.546	1.343	1.578	1.313	1.611	1.283	1.645	1.253	1.680	1.223	1.716	1.192	1.754	1.162		
75	1.448	1.501	1.422	1.529	1.395	1.557	1.368	1.587	1.340	1.624	1.313	1.649	1.284	1.682	1.256	1.714	1.227	1.748	1.199		
80	1.466	1.515	1.441	1.541	1.416	1.563	1.390	1.595	1.364	1.624	1.338	1.663	1.313	1.683	1.285	1.714	1.259	1.745	1.232		
85	1.482	1.528	1.458	1.553	1.435	1.578	1.411	1.603	1.386	1.630	1.362	1.687	1.337	1.685	1.312	1.714	1.287	1.743	1.262		
90	1.496	1.540	1.474	1.563	1.452	1.587	1.429	1.611	1.406	1.636	1.383	1.661	1.361	1.687	1.336	1.714	1.312	1.741	1.288		
95	1.510	1.552	1.489	1.573	1.468	1.596	1.446	1.618	1.425	1.642	1.403	1.666	1.381	1.690	1.358	1.715	1.336	1.741	1.313		
100	1.522	1.562	1.503	1.583	1.482	1.604	1.462	1.625	1.441	1.647	1.421	1.670	1.400	1.693	1.378	1.717	1.357	1.741	1.335		
150	1.611	1.637	1.598	1.651	1.584	1.665	1.571	1.679	1.557	1.693	1.543	1.708	1.530	1.722	1.515	1.737	1.501	1.752	1.486		
200	1.664	1.684	1.653	1.693	1.643	1.704	1.633	1.715	1.623	1.725	1.613	1.735	1.603	1.746	1.592	1.757	1.582	1.768	1.571		

n	p = 11		p = 12		p = 13		p = 14		p = 15		p = 16		p = 17		p = 18		p = 19		p = 20	
	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U	d _L	d _U
16	0.060	3.446																		
17	0.094	3.286	0.053	3.506																
18	0.113	3.146	0.075	3.358	0.047	3.357														
19	0.145	3.023	0.102	3.227	0.067	3.420	0.043	3.601												
20	0.178	2.914	0.131	3.109	0.092	3.297	0.061	3.474	0.038	3.639										
21	0.212	2.817	0.162	3.004	0.119	3.185	0.084	3.355	0.055	3.521	3.671									
22	0.246	2.729	0.194	2.909	0.148	3.084	0.109	3.252	0.077	3.412	0.035	0.035								
23	0.281	2.651	0.227	2.824	0.178	2.991	0.136	3.165	0.100	3.311	0.070	0.050	0.032	3.700						
24	0.315	2.580	0.260	2.744	0.209	2.906	0.165	3.065	0.125	3.218	0.092	0.065	0.046	3.597	0.029					
25	0.348	2.517	0.292	2.674	0.240	2.829	0.194	2.982	0.152	3.131	0.116	0.085	0.039	3.501	0.068	0.027	3.747			
26	0.381	2.460	0.324	2.610	0.272	2.758	0.224	2.906	0.180	3.050	0.141	0.107	0.055	3.410	0.080	0.048	3.629			
27	0.413	2.409	0.356	2.552	0.303	2.694	0.253	2.836	0.208	2.976	0.167	0.131	0.079	3.325	0.079	0.055	3.572			
28	0.444	2.363	0.387	2.499	0.333	2.635	0.283	2.772	0.237	2.907	0.194	0.156	0.093	3.245	0.100	0.073	3.490			
29	0.474	2.321	0.417	2.451	0.363	2.582	0.313	2.713	0.266	2.843	0.222	0.182	0.114	3.169	0.122	0.093	3.412			
30	0.503	2.283	0.447	2.407	0.393	2.533	0.342	2.659	0.294	2.875	0.249	0.208	0.137	3.098	0.146	0.114	3.338			
31	0.531	2.248	0.475	2.367	0.422	2.487	0.371	2.609	0.322	2.800	0.277	0.234	0.160	3.032	0.171	0.137	3.267			
32	0.558	2.216	0.503	2.330	0.450	2.446	0.399	2.563	0.350	2.730	0.304	0.261	0.184	2.970	0.196	0.160	3.201			
33	0.586	2.187	0.530	2.296	0.477	2.408	0.426	2.520	0.377	2.663	0.331	0.287	0.209	2.912	0.221	0.184	3.137			
34	0.610	2.160	0.556	2.266	0.503	2.373	0.452	2.481	0.404	2.590	0.357	0.313	0.233	2.858	0.246	0.209	3.078			
35	0.634	2.136	0.581	2.237	0.529	2.340	0.478	2.444	0.430	2.550	0.383	0.339	0.257	2.808	0.272	0.233	3.022			
36	0.658	2.113	0.605	2.210	0.554	2.310	0.504	2.410	0.455	2.512	0.409	0.364	0.282	2.761	0.297	0.257	2.969			
37	0.680	2.092	0.628	2.186	0.578	2.282	0.528	2.379	0.480	2.477	0.434	0.389	0.306	2.717	0.322	0.282	2.919			
38	0.702	2.073	0.651	2.164	0.601	2.256	0.552	2.350	0.504	2.445	0.458	0.414	0.330	2.675	0.347	0.306	2.872			
39	0.723	2.055	0.673	2.143	0.623	2.232	0.575	2.323	0.528	2.414	0.482	0.438	0.354	2.637	0.371	0.330	2.828			
40	0.744	2.039	0.694	2.123	0.645	2.210	0.597	2.297	0.551	2.386	0.505	0.461	0.377	2.600	0.395	0.354	2.787			
45	0.835	1.972	0.790	2.044	0.744	2.218	0.700	2.193	0.655	2.269	0.612	0.570	0.468	2.424	0.418	0.377	2.748			
50	0.913	1.925	0.871	1.987	0.829	2.051	0.787	2.116	0.746	2.182	0.705	0.665	0.586	2.318	0.625	0.586	2.652			
55	0.979	1.891	0.940	1.945	0.902	2.002	0.863	2.059	0.825	2.117	0.786	0.748	0.674	2.237	0.711	0.674	2.559			
60	1.037	1.865	1.001	1.914	0.965	1.964	0.929	2.015	0.893	2.067	0.857	0.822	0.751	2.173	0.786	0.751	2.483			
65	1.087	1.845	1.053	1.889	1.020	1.934	0.986	1.980	0.953	2.027	0.919	0.886	0.819	2.122	0.852	0.819	2.421			
70	1.131	1.831	1.099	1.870	1.068	1.911	1.037	1.953	1.005	1.995	0.974	0.938	0.880	2.082	0.911	0.880	2.372			
75	1.170	1.819	1.141	1.856	1.111	1.893	1.082	1.931	1.052	1.970	1.023	0.969	0.934	2.049	0.964	0.934	2.331			
80	1.205	1.810	1.177	1.844	1.150	1.878	1.122	1.918	1.094	1.949	1.066	1.009	0.983	2.022	1.011	0.983	2.297			
85	1.236	1.803	1.210	1.834	1.184	1.866	1.158	1.898	1.132	1.931	1.106	1.066	1.080	1.999	1.053	1.053	2.268			
90	1.264	1.798	1.240	1.827	1.215	1.856	1.191	1.886	1.166	1.917	1.141	1.094	1.116	1.973	1.091	1.091	2.242			
95	1.290	1.793	1.267	1.821	1.244	1.848	1.221	1.876	1.197	1.905	1.174	1.141	1.150	1.958	1.126	1.126	2.217			
100	1.314	1.790	1.292	1.816	1.270	1.841	1.248	1.868	1.225	1.895	1.203	1.181	1.181	1.949	1.158	1.158	2.197			
150	1.473	1.783	1.458	1.799	1.444	1.814	1.429	1.830	1.414	1.847	1.407	1.385	1.385	1.880	1.370	1.370	1.857			
200	1.581	1.791	1.550	1.801	1.539	1.813	1.528	1.824	1.518	1.836	1.507	1.495	1.495	1.860	1.484	1.484	1.842			

Примітка. n — кількість спостережень; p — кількість факторів у моделі.

ГЛОСАРІЙ

Autocorrelation coefficients	— коефіцієнти кореляції
Autocorrelation function	— автокореляційна функція
Autocovariance coefficients	— коефіцієнти коваріації
Autoregressive (AR) model	— авторегресійна модель
Autoregressive integrated moving average (ARIMA) model	— інтегрована модель авторегресії та ковзного середнього
Between estimator	— метод найменших квадратів без крос-секційної (внутрішньогрупової) варіації
Between regression	— регресія на основі середніх у крос-секційному розрізі
Between variation	— міжгрупова варіація (варіація в часовому розрізі)
Characteristic equation	— характеристичне рівняння
Characteristics roots	— характеристичні корені
Choleski decomposition	— декомпозиція Чолеського
Cointegrating relationship	— коінтеграційний зв'язок
Cointegrating vector	— коінтеграційний вектор
Cointegration error	— відхилення від рівноважного зв'язку (рівноваги)
Cointegration	— коінтеграція
Correlogram	— корелограма
Cross section specific intercept	— крос-секційно специфічний перетин; перетин, специфічний для об'єкта в крос-секційному розрізі
Detrended	— очищений від тренду
Disturbances	— випадкові величини, збурення
Drift	— зміщення (зсув)
Dummy variables	— фіктивні змінні
Eigenvalues	— характеристичні значення, власні значення
Endogeneity purged estimate	— оцінка коефіцієнта змінної, очищеної від ендогенності
Equilibrium relationship	— рівноважна взаємозалежність, рівноважний зв'язок

Error variance	— дисперсія помилки
Estimator	— метод оцінювання, функція вибірко- вих значень
Feasible GLS (EGLS)	— здійснений метод узагальнених най- менших квадратів
Fixed effects model	— модель з фіксованими ефектами
Flypaper effect	— ефект «липучки»
Forecast	— прогноз
Gap filling grants	— гранти, що визначаються як різниця між власними доходами та запланова- ними (нормативними) видатками
Goodness of fit	— якість (точність) відтворення моделлю вхідних (фактичних) даних
Granger causality	— причинність по Гренджеру
Identical independently distributed	— незалежні випадкові величини з одна- ковим розподілом
Identifications	— ідентифікація
Identified equation	— ототожене (або ідентифіковане) рівняння
Impact multipliers	— миттєві мультиплікатори
Impulse response function	— функція імпульсного відгуку
Innovation accounting	— декомпозиція інновацій (шоків), аналіз впливу шоків
Innovations	— шоки
Instantaneous (impact) effect	— миттєвий вплив
Invertibility	— оберненість
Irregular components	— випадкові компоненти
Lag operator	— лаговий оператор
Long-run multipliers	— довгострокові мультиплікатори
Maximum likelihood estimator	— метод максимальної правдоподібності
Mean	— середнє
Measurement error	— помилка вимірювання
Moving average (MA) model	— модель ковзного середнього
Non-nested models	— невикладені моделі, моделі без угрупу- вання
Omitted variables	— невиключені (пропущені) змінні (регре- сори, що не включені в модель)
One-way error components model	— модель з одновимірною складовою помилки (або одновимірними шоками)
Overidentified equation	— переототожене (переідентифіковане) рівняння

Overidentifying restrictions test	— тест переототожнених обмежень (моментів)
Partial autocorrelation function	— часткова автокореляційна функція
Polynomial in the lag operator	— поліном лагового оператора
Population mean	— середня генеральної сукупності
Random effects model	— модель з випадковими ефектами
Random walk	— нестационарний випадковий процес, мартінгал
Rank condition	— умова рангу
Ratchet effect	— ефект храповика
Reduced form	— приведена форма
Regressors	— регресори, фактори моделі, пояснювальні змінні
Regressand	— регресанд, залежна змінна
Remainder disturbances	— залишковий шум (або «помехи»)
Responsiveness	— відгук, чутливість
Robustness	— робастність, стабільність, нечутливість до невеликих змін (як правило, в специфікації моделі)
Speed of adjustment	— швидкість пристосування
S-step ahead forecast	— прогноз на s періодів наперед
Standard deviation	— середньоквадратичне відхилення
Stationarity	— стаціонарність
Strictly exogenous	— строго екзогенний
Super-consistent	— суперконсистентний
Seemingly unrelated regressions (SUR)	— напівзв'язана регресія
The least squares dummy variable (LSDV) estimator	— метод найменших квадратів з фіктивними змінними
Three-stage Least squares (3SLS)	— трикроковий метод найменших квадратів (ЗМНК)
Time series modeling	— моделювання на основі часових рядів
Trace	— слід матриці
Two-stage Least squares (2SLS)	— двокроковий метод найменших квадратів (2МНК)
Two-way error components mode	— модель з двовимірною складовою помилки (або двовимірними шоками)
Unconditional mean	— безумовне середнє
Underidentified (or unidentified) equation	— недоототожене (недоідентифіковане) рівняння
Unit root	— одиничний корінь

Univariate time series analysis	— аналіз часових рядів однієї змінної
Unobservables (variables)	— неспостережувані (змінні), неспостережні
Unrestricted model	— модель, в якій не накладено обмеження на значення параметрів моделі
Variance	— дисперсія
Variance decomposition	— декомпозиція дисперсій
Vector moving average (VMA) models	— вектор моделей ковзної середньої
Within estimator	— метод найменших квадратів для внут- рішньогрупової варіації
Within transformation	— трансформація в крос-секційному розрізі
Within variation	— внутрішньогрупова варіація
Yule-Walker equations	— рівняння Юла—Уокера
Individual dynamics	— індивідуальна динаміка
Order condition	— умова порядку
Panel data model	— модель на основі лонгitudних (панельних) даних
Policy maker	— особа, що приймає рішення
Policy variables	— керовані змінні
Predetermined (variables)	— предетерміновані, попередньо визна- чені (змінні)
Constant returns to scale	— постійна віддача від масштабів (вироб- ництва)
Cross-sectional data	— дані з крос-секційним та часовим вимірами

ПРЕДМЕТНИЙ ПОКАЖЧИК

- Автокореляційна функція (Autocorrelation function) 22
- Авторегресійна модель (autoregressive (AR) model) 11, 20
- Аналіз впливу шоків (Innovation accounting) 97
- Аналіз часових рядів однієї змінної (Univariate time series analysis) 11, 76
- Безумовне середнє (Unconditional mean) 21
- Вектор-авторегресійні моделі (VAR-models) 76
- Вектор моделей ковзної середньої (Vector moving average (VMA) models) 91
- Випадковий процес (Random work) 41, 120
- Випадкові величини (збурення) (Disturbances) 11, 76
- Відгук (чутливість) (Responsiveness) 122
- Відхилення від рівноважного зв'язку (рівноваги) (Cointegration error) 118
- Внутрішньогрупова варіація (Within variation) 255, 262
- Гранти, що визначаються як різниця між власними доходами та запланованими (нормативними) видатками (Gap filling grants) 293
- Двокроковий метод найменших квадратів (Two-stage Least squares, TSLS) 175, 197, 290
- Декомпозиція дисперсій (Variance decomposition) 94, 107, 157
- Декомпозиція Чолеського (Choleski decomposition) 81
- Динамічне прогнозування (Dynamic Forecasting) 68, 207
- Дисперсія (Variance) 21
- Довгострокові мультиплікатори (Long-run multipliers) 94
- Ефект «липучки» (Flypaper effect) 291, 293
- Ефект храповика (The ratchet effect) 293

- Залишковий шум («помехи») (Remainder disturbances) 231
- Здійснений метод узагальнених найменших квадратів (Feasible GLS (EGLS)) 254, 280
- Зміщення (зсув) (Drift) 41
- Ідентифікація (Identifications) 40, 68
- Імпульсна функція відгуків (Impulse response function) 90, 106, 155
- Інструментальні змінні (Instrument variables) 293, 298
- Інтегрована модель авторегресії та ковзного середнього (autoregressive integrated moving average (ARIMA) model) 11, 38, 40
- Інформаційний критерій Акайка (Akaike Information Criterion) 45, 84
- Інформаційний критерій Шварца (Schwarz Information Criterion) 45
- Керовані змінні (Policy variables) 213
- Коефіцієнти автоковаріації (Autocovariance coefficients) 24, 37, 39
- Коефіцієнти автокореляції (Autocorrelation coefficients) 24, 39
- Кointegraційний вектор (Cointegrating vector) 118, 123
- Кointegraційний зв'язок (Cointegrating relationship) 116
- Кointegraція (Cointegration) 117
- Корелограма (Correlogram) 25
- Крос-секційні (варіаційні) дані (Cross-sectional data) 229
- Крос-секційно специфічний перетин; перетин, специфічний для об'єкта в крос-секційному розрізі (Cross-section specific intercept) 232, 279, 284
- Лаговий оператор (Lag operator) 15
- Лонгітюдні (панельні) дані (Panel data) 229
- Метод інструментальних змінних (Instrument variables estimator) 197, 259, 290, 293
- Метод максимальної правдоподібності (Maximum likelihood estimation) 135
- Метод найменших квадратів без крос-секційної (внутрішньогрупової) варіації (Between estimator) 252
- Метод найменших квадратів для внутрішньогрупової варіації (The within estimator) 245, 253
- Метод найменших квадратів з фіктивними змінними

- (The least squares dummy variable (LSDV) estimator) 244
- Метод непрямих найменших квадратів (Indirect Least squares) 175
- Метод оцінювання, функція вибірових значень (Estimator) 245, 252
- Миттєвий вплив (Instantaneous impact) 93
- Миттєві мультиплікатори (Impact multipliers) 93
- Міжгрупова варіація (варіація в часовому розрізі) (Between variation) 254, 262
- Модель, в якій не накладено обмеження на значення параметрів моделі (Unrestricted model) 135
- Модель з випадковими ефектами (Random effects model) 233, 250
- Модель з двовимірною складовою помилки (або двовимірними шоками) (Two-way error components model) 231
- Модель з одновимірною складовою помилки (або одновимірними шоками) (One-way error components model) 231
- Модель з фіксованими ефектами (Fixed effects model) 232, 243
- Модель ковзного середнього (Moving average (MA) model) 13, 36
- Модель лонгітюдних даних з одним перетином (Nooled model) 230, 249
- Модель на основі лонгітюдних (панельних) даних (Panel data model) 230
- Моделювання на основі часових рядів (Time series modeling) 10, 56
- Невключені (пропущені) змінні (регресори, що не включені в модель) (Omitted variables) 235, 275
- Недоототожене (недоідентифіковане) рівняння (Underidentified (or unidentified) equation) 174
- Незалежні ідентично розподілені випадкові змінні (незалежні випадкові величини з однаковим розподілом) (Identical independently distributed) 232, 243
- Неспостережувані (Unobservables) 231, 267
- Оберненість (Invertibility) 38
- Одиничний корінь (Unit root) 32, 41, 44
- Особа, що приймає рішення (Policy maker) 213
- Ототожене (ідентифіковане) рівняння (Identified equation) 173
- Оцінка коефіцієнта змінної, очищеної від ендогенності («Endogeneity purged» estimate) 177, 290

Очищені від тренду (Detrended) 132

Переототожене (переідентифіковане) рівняння
(Overidentified equation) 173

Поліном лагового оператора (Polynomial in the lag operator) 16, 26

Помилка вимірювання (Measurement error) 235

Порядок інтеграції (Order of integration) 17

Постійна віддача від масштабів (виробництва)
(Constant returns to scale) 235

Предетерміновані, попередньо визначені (Predetermined) 169

Приведена форма (Reduced form) 14, 79, 170

Причинність по Гренджеру (Granger causality) 129

Прогноз на s періодів наперед (S-step ahead forecast) 51

Прогноз (Forecast) 48, 68, 87, 157, 205

Регресія на основі середніх в крос-секційному розрізі
(Between regression) 252

Регресори, фактори моделі, пояснюючі змінні (Regressors) 168

Рівноважний зв'язок (Equilibrium relationship) 117

Рівняння Юла—Уокера (Yule—Walker equations) 29

Робастність, стабільність (Robustness) 235

Середнє (Mean) 21

Середньоквадратичне відхилення (Standard deviation) 90

Середня генеральної сукупності (Population mean) 234

Слід матриці (Trace) 137

Стабільність (стаціонарність) VAR-моделі 85

Статичне прогнозування (Static Forecasting) 68, 207

Стаціонарність (Stationarity) 17

Строго екзогенні (Strictly exogenous) 247

Суперконсистентні (Super-consistent) 127

Тест Дікі—Фуллера (Dickey—Fuller test) 18, 58

Тест переототожнених обмежень (моментів)
(Overidentifying restrictions test) 289, 297

Трансформація в крос-секційному розрізі
(The within transformation) 245

Трикроковий метод найменших квадратів
(Three-stage Least squares) 171, 197

Умова порядку (Order condition) 173

Умова рангу (Rank condition) 173

Фіктивні змінні (Dummy variables) 234, 240, 243

Характеристичне рівняння (Characteristic equation) 31, 86

Характеристичні (власні) значення (Eigenvalues) 85

Характеристичні корені (Characteristics roots) 31, 85, 134

Часткова автокореляційна функція (Partial Autocorrelation function) 33

Швидкість пристосування (Speed of adjustment) 122, 157

Шоки (Innovations) 78, 83, 156

Якість (точність) відтворення моделлю вхідних (фактичних) даних
(Goodness of fit) 56, 159, 208

p -значення (p-value) 285

SUR-метод 126

ЗАГАЛЬНИЙ СПИСОК ЛІТЕРАТУРИ

Abraham B., Ledolter J. Statistical Methods for Forecasting.— New York : Wiley, 1989.

Ahmad Enthisham. Financing Decentralized Expenditures: An international Comparison of Grants.— Cheltenham, UK : Edward Elgar, 1997.

A Guide to FRB/US // A Macroeconomic Model of the United States — Federal Reserve Board.— Washington, D. C., 1996.— 45 p.

Amemiya T., MaCurdy T. E. Instrumental-Variable Estimation of an Error Components Model // *Econometrica*, 1986, № 54.— P. 869—881.

Anderson T. W. The Statistical Analysis of Time Series.— New York : Wiley, 1983.

Anderson T. W., Vahid F. Testing Multiple Equation System for Common Non-linear Components // *Journal of Econometrics*, 1998, № 84.— P. 1—36.

Arellano M. Computing Robust Standard Errors for Within-Groups Estimators // *Oxford Bulletin of Economics and Statistics*, 1987.— № 49, P. 431—434.

Arellano M. Some Testing for Autocorrelation in Dynamic Random Effects Models // *Review of Economic Studies*, 1990, № 57.— P. 127—134.

Box G. E. P., Jenkins G. M. Time Series Analysis: Forecasting and Control / Revised Edition.— San Francisco : Holden Day, 1976.

Baillie R. T. Long Memory Processes and Fractional Integration in Econometrics // *Journal of Econometrics*, 1996, № 73.— P. 5—60.

Balke N. S., Fomby T. Threshold Cointegration // *International Economic Review*, 1997, № 38.— P. 627—646.

Baltagi Badi H. Econometric Analysis of Panel Data.— New York : John Wiley & Sons, Inc., 1995.— P. 253.

Banerjee A., Dolado J., Galbraith J. W., Hendry D. F. Cointegration, Error Correction, and the Econometric Analysis of Nonstationary Data.— Oxford : Oxford University Press, 1993.

Begg D. Monetary Policy During Transition: Progress and Pitfalls in Central and Eastern Europe // *Oxford Review of Economic Policy*, 1990, № 6, Vol. 13 (2).— P. 33—46.

Beguín J.-M., Gourieroux C., Montfort A. Identification of a Mixed Autoregressive-Moving Average Process: The Corner Method. Time Series.— Amsterdam : North Holland, 1980.— P. 423—436.

Bhargava A., Franzini L., Narendranathan W. Serial Correlation and Fixed Effects Model // *Review of Economic Studies*, 1982, № 49.— P. 533—549.

Blanchard Oliver J. What is Left of the Multiplier Accelerator // *The American Economic Review*, 1981, May, Vol. 71, Issue 2.— P. 150—154.

Boswijk H. P. Asymptotic Theory for Integrated Processes.— Oxford : Oxford University Press, 1998.

Breusch T. S., Mizon G. E., Schmidt P. Efficient Estimation Using Panel Data // *Econometrica*, 1989, № 57.— P. 695—700.

Campbell John Y., Perron Pierre. Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Working Papers Series, 1991 (Available at www.nber.org).

Charemza W. W., Deadman D. F. New Direction in Econometric Practice.— Brookfield : Edward Eglar, 1992.

Clements M. P., Hendry D. F. Forecasting in Cointegrated Systems // Journal of Forecasting, 1995, № 10.— P. 127—146.

Cullis John, Jones Philip. Public Finance and Public Choice.— 2nd ed.— Oxford : Oxford University Press, 1998.

Davidson J. E. H. Problems with the Estimation of Moving Average Processes // Journal of Econometrics, 1981, № 16.— P. 295—310.

Dickey D. A., Bell W. R., Miller R. B. Unit Roots in Time Series Models: Tests and Implications // American Statistician, 1986, № 40.— P. 12—26.

Dickey D. A., Fuller W. A. Distribution of the Estimators for the Autoregressive Time Series with a Unit Root // Journal of the American Statistical Association, 1979, № 74.— P. 427—431.

Dickey D. A., Fuller W. A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root // Econometrica, 1981, № 49.— P. 1057—1072.

Dornbush R., Fisher S., Startz R. Macroeconomics.— 7th ed.— Boston : Irwin McGraw-Hill, 1998.— 550 p.

Earle John S., Sabirianova Klara Z. Equilibrium Wage Arrears: A Theoretical and Empirical Analysis of Institutional Lock-In. The Institute for the Study of Labor (IZA). Discussion Paper. № 196. Available at <ftp://mailserver.iza.org/RePEc/Discussionpaper/dp196.pdf>, 2000.

Enders W. Applied Econometric Time Series.— New York : Wiley, 1995.

Engle R. F., Granger C. W. Cointegration and Error Correction: Representation. Estimation and Testing // Econometrica, 1987, № 55.— P. 251—276.

Engle R. F., Yoo B. S. Forecasting and Testing in Cointegrated System // Journal of Econometrics, 1987, № 35.— P. 143—159.

Frances P. H. Time Series Models for Business and Economic Forecasting.— Cambridge : Cambridge University Press, 1999.— 280 p.

Frances P. H., Hylleberg S., Lee H. S. Spurious Deterministic Seasonality // Economics Letters, 1995, № 48.— P. 249—256.

Frumkin N. Guide to Economic Indicators.— 2nd ed.— M. E. Sharpe, 1994.— 430 p.

Fuller W. A. Introduction to Statistical Time Series.— New York : Wiley, 1976.

Geiger Linwood T. Macroeconomic Analysis and Transitional Economy.— Pennsylvania : St. Davids, 1992.

Gonzalo J., Granger C. W. Estimation of Common Long — Memory Components in Cointegrated System // Journal of Business and Economic Statistics, 1995, № 13.— P. 27—36.

Granger C. W. Some Recent Developments in a Concept of Causality // Journal of Econometrics, 1988, № 39.— P. 199—211.

Granger C. W., Andersen A. P. Introduction to Bilinear Time Series Models.— Goettingen : Vandenhoeck and Ruprecht, 1978.

Granger C. W. J., Newbold P. Spurious Regressions in Econometrics // Journal of Econometrics, 1974, № 2.— P. 111—120.

Granger C. W. J., Terasvirta T. Modelling Nonlinear Economic Relationships.— Oxford : Oxford University Press, 1993.

Greene William H. Econometric Analysis.— 4th ed.— New Jersey : Prentice Hall, 2000.— 1004 p.

Gronicki Mirek, Piktko Katarzyna. Interest Rates and Real Economy in Ukraine // Ukrainian Economic Trends, 1998, № 1.— P. 4—5

Gronicki Miroslaw, Piktko Katarzyna. Macroeconomic Model for Ukraine // CASE Working Paper, 1999.— 56 p.

Gros D., Steinherr A. Winds of Change: Economic Reforms in Central and Eastern Europe.— London : Longman, 1995.

Gujarati Damodar N. Basic Econometrics.— 3rd ed.— New York : McGraw-Hill, 1995.— 838 p.

Hall Robert E. Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence // The Journal of Political Economy, 1978, Vol. 86(6).— P. 971—987.

Hall Robert E., Mishkin Frederic S. The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households // Econometrica, 1982, Vol. 50(2).— P. 461—482.

Hamilton James D. Time Series Analysis.— Princeton, New Jersey : Princeton University Press, 1994.

Hannan E. J. Multiple Time Series.— New York : Wiley, 1970.

Hannan E. J. The Estimation of the Order of an ARMA Process // Annals of Statistics, 1980, № 8.— P. 1071—1081.

Hansen L. P., Sargent T. J. Seasonality and Approximation Errors in Rational Expectations Models // Journal of Econometrics, 1993, № 55.— P. 21—56.

Harvey A. C. Time Series Models.— Oxford : Philip Allan, 1981.

Hatanaka M. Time-Series-Based Econometrics. Unit Roots and Cointegration.— Oxford : Oxford University Press, 1996.

Hausman J. A. Specification Tests in Econometrics // Econometrica, 1978, № 43.— P. 1251—1271.

Hausman J. A., Taylor W. E. Panel Data and Unobservable Individual Effects // Econometrica, 1981, № 49.— P. 1377—1398.

Hendry D. F. Dynamic Econometrics.— Oxford : Oxford University Press, 1995.

Hoch I. Estimation of Production Function Parameters Combining Time-Series and Cross-Sectional data // Econometrica, 1962, № 30.— P. 34—53.

Hsiao Cheng. Analysis of Panel Data. UK.— Cambridge : Cambridge University Press, 1986.

Hylleberg S. Modelling Seasonality.— Oxford : Oxford University Press, 1992.

Hylleberg S., Engle R. F., Granger C. W. J., Yoo B. S. Seasonal Integration and Cointegration // Journal of Econometrics, 1990, № 44.— P. 215—238.

International Monetary Fund. "Ukraine: Recent Economic Developments". IMF Staff Country Paper, № 99/42, 1999.

Islam Muhammed N., Saud A. Choudhury. The Flypaper Effect and the Revenue Impact of Grants-in-Aid // Economics Letters, 1989, Vol. 30 (4).— P. 351—356.

Johansen S. Cointegration in Partial Systems and the Efficiency of Single — Equation Analysis // Journal of Econometrics, 1992, № 52.— P. 389—402.

Johansen S. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.— Oxford : Oxford University Press, 1995.

Johansen S. Statistical Analysis of Cointegration Vector // Journal of Economics Dynamics and Control, 1988, № 12.— P. 231—254.

Johansen S., Juselius K. Maximum Likelihood Estimation and Inference on Cointegration with Application to the Demand for Money // Oxford Bulletin of Economics and Statistics, 1990, № 52.— P. 169—210.

Johnston J., Dinardo J. Econometric Methods.— New York : McGraw-Hill, 1997.

Kaufmann D., Kaliberda A. Integrating the Unofficial Economy into Dynamics of Post-Socialist Economies: a Framework of Analysis and Evidence // Economic Transition in Russia and the New States of Eurasia, 1996, August.

Kennedy P. A Guide to Econometrics.— Cambridge, Massachusetts : The MIT Press, 1998.— 468 p.

Klein L. (editor) Comparative Performance of U. S. Econometric Models.— Oxford : Oxford University Press, 1991.

Kmenta J. Elements of Econometrics.— New York : MacMillan, 1986.

Knight Brian. The Flypaper Effect Unstuck: Evidence on Endogenous Grants from The Federal Highway Aid Program. Board of Governors of the Federal Reserve System. Washington DC. Available at <http://netec.mcc.ac.uk>, 2000.

Krugman Paul R. Obstfeld Maurice International Economics: Theory and policy.— 3rd ed.— Harper Collins College Publishers, 1994.

Kuh E. The Validity of Cross-Sectionally Estimated Behavior Equations in Time-Series Applications // Econometrica, 1959, № 27.— P. 197—214.

Lee T.-H., Tse Y. Cointegration Tests with Conditional Heteroskedasticity // Journal of Econometrics, 1996, № 73.— P. 401—410.

Lin J.-L., Tsay R. S. Cointegration Constraints and Forecasting: An Empirical Examination // Journal of Applied Econometrics, 1996, № 11.— P. 519—538.

Lucas R. E. Econometric Policy Evaluation: A Critique // *K. Brunner, A. H. Meltzer.* The Philips Curve and Labor Markets. Carnegie-Rochester Conference Series on Public Policy, 1976, № 1. — P. 19—46.

Lukepohl H. Introduction to Multiple Time Series Analysis.— Berlin : Springer-Verlag, 1991.

Lukepohl H., Reimers H.-E. Impulse Response Analysis of Cointegrated Systems // Journal of Economic Dynamics and Control, 1992, № 16.— P. 53—78.

Macroeconometric Modeling and Its Use for the Conduct of Policy in Ukraine // TACIS, Ukrainian Economic Trends, 1998, № 9.— P. 82—86.

Makridakis S., Wheelwright S. C., McGee V. E. Forecasting: Method and Application.— 2nd ed.— New York : Wiley, 1983.

Mankiw N. Gregory. Macroeconomics.— 2nd ed.— New York : Worth Publishers, Inc., 1992.— 514 p.

Mankiw Gregory, Campbell John. Consumption, Income, and Interest rates. Reinterpreting the Time Series Evidence // NBER Working Papers, 1989, № 2924.

Matyas L., Sevestre P. The Econometrics of Panel Data: Handbook of Theory and Applications.— Dordrecht : Kluwer Academic Publishers, 1992.

McPherson Malcolm F., Rakovski Tzvetana S. A Small Econometric Model of the Zambian Economy // HIID Development Discussion Papers, 1999, January, № 672.— 30 p.

Meade N., Islam T. Prediction Intervals for Growth Curve Forecasts // Journal of Forecasting, 1995, № 14.— P. 413—430.

Mills Terence C. Time Series Techniques for Economists.— Cambridge : Cambridge University Press, 1998.

Mitchell W. F. Testing for Unit Roots and Persistence in OECD Unemployment Rates // Applied Econometrics, 1993, № 25.— P. 1489—1501.

Mundlak Y. Empirical Production Function Free of Management Bias // Journal of Farm Economics, 1969, № 43.— P. 44—56.

Mundlak Y. On the Pooling of Time-Series and Cross-Section Data // Econometrica, 1978, № 46.— P. 44—56.

Nerlove M., Balestra P. Formulation and Estimation of Econometric Models for Panel Data // *Matyas L., Sevestre P.* The Econometrics of Panel Data: Handbook of Theory and Applications.— Dordrecht : Kluwer Academic Publishers, 1992.— P. 3—18.

Nijman Th. E., Verbeek M. Estimation of Time Dependent Parameters in Linear Models Using Cross-Section Panels or Both // Journal of Econometrics, 1990, № 46.— P. 333—346.

Nijman Th. E., Verbeek M. Nonresponse in Panel Data: The Impact on Estimates of a Life Cycle Consumption Function // Journal of Applied Econometrics, 1992, №, 7.— P. 243—257.

Ooms M. Empirical Vector Autoregressive Modeling.— Berlin : Springer-Verlag, 1994.

Osborn D. R., Smith J. P. The performance of Periodic Autoregressive Models in Forecasting Seasonal UK Consumption // Journal of Business and Economic Statistics, 1989, № 7.— P. 117—127.

Pagan A. R. The Econometrics of Financial Markets // Journal of Empirical Finance, 1996, № 3.— P. 15—102.

Patterson Kerry. An Introduction to Applied Econometrics: A Time Series Approach.— London : St. Martin Press, 2000.

Paulsen J. Order Determination of Multivariate Autoregressive Time Series with Unit roots // Journal of Time Series Analysis, 1984, № 5.— P. 115—127.

Pindyck R. S., Rubinfeld D. L. Econometric Models and Economic Forecasts.— 3rd ed.— New York : McGraw-Hill, 1991.

Poirier D. J. Intermediate Statistics and Econometrics: A Comparative Approach.— Cambridge, Massachusetts : MIT Press, 1995.

Poskitt D. S., Tremayne A. R. Diagnostic Tests for Multiple Time Series Models // Annals of Statistics, 1982, № 10.— P. 114—120.

Gaunes A., Thakur S. Macroeconomic Accounting and Analysis in Transition Economies.— Washington, D.C. : International Monetary Fund, 1997.— 183 p.

Ray C. Fair Testing Macroeconometric Models.— Harvard University Press, 1994.— 421 p.

Roland Gerard. Transition and Economics: Politics, Markets and Firms.— Cambridge, Massachusetts : MIT Press, 2000.

Ross S. A., Westerfield R. W., Jaffe J. Corporate Finance.— 4th ed.— Chicago : Irwin, 1996.— 899 p.

Ruud P. A. An Introduction to Classical Econometric Theory.— Oxford : Oxford University Press, 2000.

Sachs Jeffrey D., Larrain Felipe B. Macroeconomics in the global economy.— New Jersey : Prentice-Hall, 1993.— 3778 p.

Schneider Mark, Byung Moon Ji. The Political Economy of Intergovernmental Grant Seeking: Targeting and Suburbs, 1977 and 1982 // American Journal of Political Science, 1990, May, Vol. 34(2).— P. 408—420.

Steinherr A. Banking Reforms in Eastern European Countries // Oxford Review of Economic Policy, Vol. 13, № 2.— P. 106—125.

Stock J., Watson M. W. Testing for Common Trend // Journal of the American Statistical Association, 1989, № 83.— P. 1097—1107.

TACIS. Ukrainian Economic Trends. Monthly Update, 1999, № 5.

Taylor L. Income Distribution, Inflation and Growth: Lectures on Structuralist Macroeconomic Theory.— Cambridge, Massachusetts : MIT, 1991.— 543 p.

Taylor S. J. Modelling Financial Time Series.— Chichester : Wiley, 1986.

Tiao G. C., Box G. E. P. Modelling Multiple Time Series with Applications // Journal of the American Statistical Association, 1981, № 76.— P. 802—816.

Tong H. Non-linear Time Series. A Dynamical System Approach.— Oxford : Oxford University Press, 1990.

Ukraine: Recent Economic Developments. IMF Staff Country Paper, 1999, № 99/42.

Varian Hal R. Intermediate Microeconomics: A Modern Approach.— 5th ed.— New York : W. W. Norton & Company, 1999.— 711 p.

Veenstra A. W., Frances P. H. A Cointegration Approach to Forecasting Freight Rates in the Dry Bulk Shipping Sector // Transportation Research A, 1997, № 31.— P. 447—458.

Verbeek Marno. A Guide to Modern Econometrics.— Chichester : John Wiley and Sons, Ltd., 2000.

Zhuravskaya Ekatherina. Incentives to Provide Local Public Goods: Fiscal Federalism, Russian Style. EERC Russia Working Papers, 1999. Available at http://www.eerc.ru/publications/workpapers/WP_99-15E.pdf.

Аллен Р. Математическая экономия / Пер. с англ; Под ред. А. Л. Вайнштейна.— М. : Изд-во иностранной литературы, 1963.— 599 с.

Беседін В. Ф., Богдан І. В., Варениченко Л. Л. та ін. Основні прогностичні показники економічного і соціального розвитку України на 2000 рік // За ред. В. Ф. Беседіна.— К. : Інтелект, 1999.— 159 с.

Бураковський І. Теорія міжнародної торгівлі.— К. : Основи, 1996.— 241 с.

Бураковський В. В. та ін. Податки. Навчальний посібник.— Дніпропетровськ : Пороги, 1998.— 646 с.

Бюджетний і податковий огляд. Третій квартал 1998 р.: Аналітичні статті, коментарі, таблиці: Група фіскального аналізу при Комітеті Верховної Ради України з питань бюджету.

Бюджетний і податковий огляд. Четвертий квартал 1998 р.: Аналітичні статті, коментарі, таблиці: Група фіскального аналізу при Комітеті Верховної Ради України з питань бюджету.

Гесць В., Скрипниченко М., Соколик М., Шумська С. Секторальні макромоделі прогнозування економіки України // Економіст, 1998, № 5.— С. 58—67.

Гончар В. Моделювання інфляції в Україні // Вісник НБУ, 1998, № 2.— С. 52—56.

Городніченко Ю. О. Вплив трансферів на податкові зусилля областей // Наукові записки НаУКМА. Спец. випуск. Т. 19, ч. II.— К., 2001.— С. 371—375.

Городніченко Ю. О. Розщеплення податків як складова підвищення ефективності регіональної політики // Наукові записки. НаУКМА. Серія «Економічні науки», № 19.— К., 2001.— С. 34—40.

Гроніцікі М., Ширмер Я., П'єнтка К. Альтернативні сценарії розвитку економіки України у 1998—1999 роках // Економіст, 1998, № 5.— С. 72—76.

Декрет Кабінету Міністрів України «Про прибутковий податок з громадян» // Відомості Верховної Ради, 1993, № 10.

Джонстон Дж. Эконометрические методы / Пер. с англ. и предисловие Рывкина А. А.— М. : Статистика, 1980.— 444 с.

Слейко В. Основи економетрії.— Львів : Марка, 1995.

Закон України «Про бюджетну систему України» // ВВР, 1991, № 1.

Закон України «Про Державний бюджет України на 2000 рік» // ВВР, 2000, № 8.

Закон України «Про оподаткування прибутку підприємств» // ВВР, 1995, № 4.

Закон України «Про податок на додану вартість» від 3 квітня 1997 р. // ВВР, 1997, № 21.

Льчук С. Погані борги в Україні створює держава // Перспективні дослідження. I квартал 1999 р., № 1.— С. 17—33.

Кваджа С., Лук'яненко І. Г., Городніченко Ю. О. Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К. : Видавничий дім «KM Academia», 2000.— 204 с.

Ковальчук Т., Коваль Н. Основні чинники інфляції в Україні // Економіка України, 1996, № 3, 4.— С. 31—40, 22—32.

Корнаи Я. Дефицит.— М. : Наука, 1992.— 608 с.

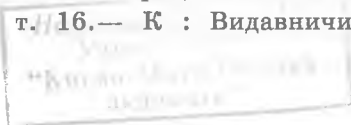
Корольов О. А. Економетрія. Навчальний посібник: КНТУ.— К., 2000.— 660 с.

Крючкова І. В. Прогноз основних макроекономічних показників України до 2000 року: Інститут економічного прогнозування НАН України.— К., 1999.— 66 с.

Крючкова И. П. Инфляция и процентные ставки // Деньги и кредит, 1995, № 3.— С. 20—25

Лук'яненко І. Г. Аналіз західного досвіду політики вирівнювання регіонів // Система державних фінансів України: шляхи забезпечення прозорості та оптимальної децентралізації фінансових потоків. Матеріали форуму експертів, Київ, 2 березня 2001 р.: К.І.С.— К., 2001.— С. 64—70.

Лук'яненко І. Г. Попит на гроші. Ілюстрація класичного підходу // Наукові записки НаУКМА. Економіка, т. 16.— К. : Видавничий дім «Педагогіка», 2000.— С. 38—42.



16387b

Лук'яненко І. Г., Городніченко Ю. О. Економетричне моделювання процесів у грошово-кредитній сфері // Наукові записки НаУКМА. Економіка, т. 15.— К.: Видавничий дім «Педагогіка», 1999.

Лук'яненко І. Г., Городніченко Ю. О. Моделювання розподілу дотацій між регіонами України // Вісник НБУ, 1999, № 9.— С. 16—21.

Лук'яненко І. Г., Городніченко Ю. О. Прогнозування податкових надходжень за допомогою моделей корегування помилки // Фінанси України, 2001, № 7.— С. 89—99.

Лук'яненко І. Г., Городніченко Ю. О., Краснікова Л. І. Економетричні підходи до аналізу фінансової програми місцевих органів влади України.— К.: Видавн. дім «KM Academia», 2000.— 204 с.

Лук'яненко І. Г., Краснікова Л. І. Економетрика. Теорія та практика.— К.: Знання, 1998.— 493 с.

Лук'яненко І. Г., Краснікова Л. І. Економетрика. Практикум з використанням комп'ютера.— К.: Знання, 1998.— 217 с.

Лук'яненко І. Г., Ситник Т. М. Моделювання заборгованості по заробітній платі на регіональному рівні // Наукові записки НаУКМА. Спец. випуск. Т. 19, ч. II.— К., 1999.— С. 111—114.

Наконечний С. І., Терещенко Т. О., Романюк Т. П. Економетрія. Підручник. КНЕУ.— К., 2000.— 296 с.

Райская Н. Временные лаги в динамике инфляции // Вопросы экономики, 1996, № 8.— С. 34—41

Рывок в рыночную экономику.— К.: Феникс, 1997.— 304 с.

Султан К., Лук'яненко І. Г., Городніченко Ю. О. Методологічні аспекти розробки та практичного застосування макроеконометричної моделі України.— К.: Видавничий дім «KM Academia», 2000.— 204 с.

Указ Президента України № 857/98 «Про деякі зміни в оподаткуванні» від 7 серпня 1998 р.

Холден К., Піл Д. А., Томпсон Дж. Л. Економічне прогнозування. Вступ.— К.: Інформтехніка-ЕМЦ, 1996.— 215 с.

Черваньов Д. М., Комашко О. В. Економетрика. Курс лекцій.— РВЦ КІЕМБСС, 1998.— 61 с.

Черняк О. І., Ставицький А. В. Динамічна економетрика.— К.: КВЦ, 2000.— 120 с.

Экономическая статистика / Под ред. Иванова Ю. Н.— М.: Инфра-М, 1998.— 480 с.

Навчальне видання

Лук'яненко Ірина Григорівна
Городніченко Юрій Олександрович

Сучасні економетричні методи у фінансах

Навчальний посібник

Рекомендовано Міністерством освіти і науки України

Відповідальна за випуск *С. В. Хрустальова*
Коректор *О. Л. Рудь*
Відповідальна за верстку *Л. Ф. Усаненко*

Підписано до друку 16.09.2002. Формат 70 × 100/16.
Гарнітура Schoolbook. Папір офсетний. Друк офсетний.
Умов. друк. арк. 28,4. Обл. вид. арк. 28,5.
Зам. 2-650.

Видавництво «Літера ЛТД».
03057, Київ, вул. Нестерова, 3, к. 408.
Телефон для довідок 456-40-21.
Свідоцтво про реєстрацію № 923 від 22.05.2002 р.

Віддруковано на ВАТ «Поліграфкнига».
03057, Київ, вул. Довженка, 3.