

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

Інститут інформаційних технологій в економіці

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика»

Галузь знань 05 «Соціальні та поведінкові науки»
Спеціальність 051 «Економіка»

Форма здобуття освіти: очна (денна)

КВАЛІФІКАЦІЙНА БАКАЛАВРСЬКА РОБОТА

на тему «Аналіз та прогнозування індексів споживчих цін на продукти харчування»

(назва теми)

здобувача

Дратованого

Максима

Олександровича

(підпис)

(ПІБ)

Науковий керівник: Кандидат економічних наук Ольга ОСИПОВА

(підпис)

Робота допущена до захисту перед екзаменаційною комісією з атестації здобувачів вищої освіти (ЕК)

Завідувач кафедри: кандидат фізико-математичних наук,
професор Галина ВЕЛИКОІВАНЕНКО

(підпис)

Київ 2026

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

Інститут інформаційних технологій в економіці

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика»

Галузь знань 05 «Соціальні та поведінкові науки»

Спеціальність 051 «Економіка»

ПОГОДЖЕНО

Керівник проектної групи (гарант)
освітньо-професійної програми

Ольга ПРИТОМАНОВА

(підпис)

2026 р.

ЗАТВЕРДЖУЮ

Завідувач кафедри штучного
інтелекту, моделювання та
статистики

Галина ВЕЛИКОІВАНЕНКО

(підпис)

2026 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувача вищої освіти

Дратованого Максима Олександровича

(прізвище, ім'я, по батькові)

очної (денної)

форми здобуття освіти

очної (денної), заочної, дистанційної

на підготовку кваліфікаційної бакалаврської роботи

на тему

«Аналіз та прогнозування динамічного ряду індексів споживчих цін на продукти харчування»

Тему затверджено наказом ректора Університету від 16.03.2026 р. №417-ст

Кваліфікаційна бакалаврська робота виконується на матеріалах сайтів Державної служби статистики <https://www.ukrstat.gov.ua/>, Національного банку України <https://bank.gov.ua/>, медіасайту Мінфін <https://minfin.com.ua/>

План кваліфікаційної бакалаврської роботи

Розділ 1	Теоретико-методологічні засади дослідження індексів споживчих цін
Розділ 2	Економетричне моделювання індексів споживчих цін на продукти харчування
Розділ 3	Прогнозування індексів споживчих цін на основі моделей часових рядів

Об'єкт дослідження:	процес формування та динаміки індексу споживчих цін на продукти харчування в Україні як інтегрального показника цінової динаміки в галузі продовольства
Предмет дослідження:	економіко-математичні методи аналізу, моделювання та прогнозування індексу споживчих цін на продукти харчування, зокрема факторні економетричні моделі та моделі часових рядів
Мета виконання кваліфікаційної бакалаврської роботи:	дослідження динаміки індексу споживчих цін на продукти харчування, побудова економетричних моделей залежності ІСЦ від макроекономічних чинників, а також прогнозування його динаміки на основі моделей часових рядів та порівняльного аналізу альтернативних методів прогнозування

Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:

У розділі 1: дослідити економічну сутність індексу споживчих цін та його роль у макроекономічному аналізі; розглянути методологію розрахунку ІСЦ та формування споживчого кошика; систематизувати фактори, що впливають на динаміку індексу споживчих цін на продукти харчування, та обґрунтувати вибір факторних змінних для економетричного моделювання.

У розділі 2: сформувати та підготувати інформаційну базу дослідження на основі статистичних даних; провести попередній аналіз даних і кореляційних залежностей; побудувати та порівняти багатфакторні економетричні моделі індексу споживчих цін на продукти харчування; оцінити статистичну значущість параметрів, якість та адекватність моделей; визначити найбільш придатну специфікацію для пояснення динаміки ІСЦ.

У розділі 3: здійснити аналіз часового ряду індексу споживчих цін, дослідити його трендову та сезонну структуру, перевірити стаціонарність ряду; побудувати та ідентифікувати модель класу SARIMA; виконати діагностику залишків та прогнозування динаміки ІСЦ; оцінити якість прогнозів за відповідними критеріями точності та сформулювати висновки щодо практичного застосування отриманих результатів.

Завдання підготував
науковий керівник

(підпис)

Ольга ОСИПОВА
(ім'я, прізвище)

«__» _____ 2026 р.

Завдання одержав
здобувач

(підпис)

Максим ДРАТОВАНІЙ
(ім'я, прізвище)

«__» _____ 2026 р.

РЕФЕРАТ

Кваліфікаційна бакалаврська робота складається зі вступу, трьох розділів, висновків та переліку використаних джерел, що містить 30 найменувань на 3 сторінках. Робота містить 13 рисунків і 12 таблиць. Загальний обсяг роботи становить 52 сторінки друкованого тексту, з яких 49 сторінок основної частини.

«Аналіз та прогнозування індексів споживчих цін на продукти харчування»

Об'єктом дослідження є процес формування та динаміки індексу споживчих цін на продукти харчування в Україні як інтегрального показника цінової динаміки в галузі продовольства.

Предметом дослідження є економіко-математичні методи аналізу, моделювання та прогнозування індексу споживчих цін на продукти харчування, зокрема факторні економетричні моделі та моделі часових рядів.

Метою кваліфікаційної бакалаврської роботи є дослідження динаміки індексу споживчих цін на продукти харчування, побудова економетричних моделей залежності ІСЦ від макроекономічних чинників, а також прогнозування його динаміки на основі моделей часових рядів та порівняльного аналізу альтернативних методів прогнозування.

Відповідно до поставленої мети визначено такі завдання: дослідити економічну сутність індексу споживчих цін та його роль у макроекономічному аналізі; розглянути методологію розрахунку ІСЦ та формування споживчого кошика; сформувати та підготувати інформаційну базу дослідження на основі статистичних даних; провести попередній аналіз даних і кореляційних залежностей; здійснити аналіз часового ряду індексу споживчих цін, дослідити його трендову та сезонну структуру, перевірити стаціонарність ряду; виконати діагностику залишків та прогнозування динаміки ІСЦ; оцінити якість прогнозів за відповідними критеріями точності та сформулювати висновки щодо практичного застосування отриманих результатів.

У роботі систематизовано підходи до моделювання та прогнозування цінової динаміки, виявлено обмеження окремих методів і обґрунтовано доцільність комбінованого підходу для волатильних рядів зі структурними зламами. Практичні результати дослідження полягають у можливості використання побудованих моделей та отриманого прогнозу як орієнтира для управління ціновими ризиками, тарифного й бюджетного планування та стрес-тестування сценаріїв розвитку інфляційних процесів.

Рік виконання кваліфікаційної бакалаврської роботи — 2026.

Рік захисту роботи — 2026.

Ключові слова: індекс споживчих цін, продукти харчування, економетричне моделювання, часові ряди, SARIMA, прогнозування, ковзне тестування.

В І Д Г У К
про кваліфікаційну бакалаврську роботу
здобувача освітньо-професійної програми «Економічна кібернетика»
навчально-наукового інституту
«Інститут інформаційних технологій в економіці»

Дратованого Максима Олександровича

на тему «Аналіз та прогнозування індексів споживчих цін на продукти харчування»

1. **Актуальність теми.** Індекс споживчих цін є одним із ключових макроекономічних показників, що використовується для оцінювання інфляційних тенденцій, формування економічної політики та прийняття управлінських рішень. Тому аналіз факторів впливу на індекси споживчих цін та побудова прогнозних моделей є актуальним науково-практичним завданням.
2. **Позитивні риси кваліфікаційної бакалаврської роботи:** у роботі розглянуто теоретико-методичні засади формування індексу споживчих цін, досліджено особливості його розрахунку та систематизовано основні фактори, що впливають на цінову динаміку. Автором проведено аналіз статистичних даних, виконано побудову та порівняння кількох економетричних моделей залежності індексу споживчих цін від макроекономічних чинників, а також здійснено прогнозування на основі моделей часових рядів. Робота характеризується логічною структурою, послідовністю викладення матеріалу та використанням сучасного інструментарію економетричного аналізу.
3. **Наявність самостійних розробок автора:** здобувачем самостійно здійснено збір, підготовку та аналіз статистичних даних, проведено економетричне моделювання та побудовано прогнозні оцінки динаміки індексу споживчих цін на продукти харчування. Автором виконано порівняння різних типів моделей та здійснено оцінювання їхньої якості, що дозволило сформулювати відповідні висновки щодо можливостей їх практичного застосування. Робота виконана відповідно до затвердженого індивідуального завдання та оформлена згідно з установленими вимогами.
4. **Цінність теоретичних висновків та практичних рекомендацій, отриманих у роботі,** полягає у визначенні особливостей динаміки індексу споживчих цін на продукти харчування та оцінюванні впливу окремих макроекономічних факторів на його зміну. Результати дослідження можуть бути використані для подальшого

аналізу інфляційних процесів, а також як інформаційна основа для прогнозування цінової динаміки та прийняття економічних рішень.

5. **Наявність недоліків:** окремі результати моделювання та прогнозування могли б бути доповнені більш детальною економічною інтерпретацією отриманих результатів. Подальше розширення інформаційної бази дослідження за рахунок включення оновлених статистичних даних дозволило б підвищити практичну цінність отриманих висновків та прогнозних оцінок.
6. **Загальна оцінка кваліфікаційної бакалаврської роботи та її допущення до захисту перед ЕК:** кваліфікаційна бакалаврська робота Дратованого М.О. виконана на позитивному рівні та може бути допущена до захисту перед ЕК. Робота свідчить про наявність у здобувача необхідних знань і навичок у сфері економічного аналізу, економетричного моделювання та прогнозування соціально-економічних процесів, а її автор, за умови успішного захисту, заслуговує на присвоєння освітньої кваліфікації «Бакалавр з економіки».

Науковий керівник: доцент кафедри
штучного інтелекту,
моделювання та статистики,
к.е.н., доцент

_____ Ольга ОСИПОВА

«_____» _____ р.

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ	
ДОСЛІДЖЕННЯ ІНДЕКСІВ СПОЖИВЧИХ ЦІН	6
1.1 Економічна сутність та класифікація індексів споживчих цін.....	6
1.2 Методологія розрахунку ІСЦ та формування споживчого кошика	9
1.3 Фактори впливу на динаміку індексу споживчих цін.....	13
РОЗДІЛ 2. ЕКОНОМЕТРИЧНЕ МОДЕЛЮВАННЯ ІНДЕКСІВ	
СПОЖИВЧИХ ЦІН НА ПРОДУКТИ ХАРЧУВАННЯ	15
2.1 Аналіз і підготовка даних для економетричного моделювання.....	15
2.2 Побудова та порівняння економетричних моделей ІСЦ	18
2.3 Результати моделювання, висновки та рекомендації.....	25
РОЗДІЛ 3. ПРОГНОЗУВАННЯ ІНДЕКСІВ СПОЖИВЧИХ ЦІН НА	
ОСНОВІ МОДЕЛЕЙ ЧАСОВИХ РЯДІВ	26
3.1 Загальна характеристика даних і попередній аналіз часового ряду	26
3.2 Побудова та ідентифікація моделі SARIMA.....	31
3.3 Діагностика, прогнозування та оцінка якості прогнозу.....	35
ВИСНОВКИ	48
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	51
ЗМІСТ	ОШИБКА! ЗАКЛАДКА НЕ ОПРЕДЕЛЕНА.

ВСТУП

Економічна проблема. Сучасна економіка піддана непередбачуваним змінам, а її стабільність визначається рядом факторів, серед яких важливе місце займає рівень цін на споживчі товари. Одним із ключових показників цінової динаміки є індекс споживчих цін (ІСЦ), аналіз і прогнозування якого має вирішальне значення для розуміння економічних процесів. Кваліфікаційна робота присвячена комплексному дослідженню індексів споживчих цін на продукти харчування: виявленню структури їх динамічного ряду, визначенню факторів впливу, побудові економетричних моделей залежності та прогнозуванню майбутніх значень засобами аналізу часових рядів.

Зміна цін на продукти харчування є постійним і неперервним явищем, що має значний вплив на економічну діяльність та побут населення. Розвиток глобальної економіки та постійні зміни в її структурі роблять цінові аспекти особливо важливими. За останні роки світ стикнувся з безпрецедентними викликами — глобальними кризами, пандемією та війною, які вносять невизначеність у цінові процеси та ускладнюють їх прогнозування. Детальне дослідження структури динамічного ряду ІСЦ на продукти харчування дає змогу виявити закономірності, циклічність та сезонність у змінах цін, а застосування сучасних методів моделювання та прогнозування дозволяє кількісно оцінити майбутню цінову динаміку.

Теоретична значимість. Дослідження структури динамічного ряду ІСЦ на продукти харчування вносить вагомий внесок у теорію економічних наук. Розкриття закономірностей та факторів, що визначають цінову динаміку в галузі, розширює розуміння процесів у цьому сегменті економіки та поглиблює методичний апарат економетричного й часово-рядового аналізу цінових явищ.

Практична значимість полягає у тому, що дослідження висвітлює реальні аспекти впливу цінових коливань на економічну діяльність суб'єктів ринку та

споживачів. Вивчення структури динамічного ряду ІСЦ та побудова прогнозу дозволяють виявити фактори, які мають найбільший вплив на ціни на продукти харчування, та можуть служити основою для розробки стратегій управління ціновими ризиками, тарифного й бюджетного планування.

Мета кваліфікаційної роботи полягає у дослідженні динаміки індексу споживчих цін на продукти харчування, побудові економетричних моделей залежності ІСЦ від макроекономічних чинників, а також прогнозуванні його динаміки на основі моделей часових рядів та порівняльного аналізу альтернативних методів прогнозування

Завдання дослідження:

- визначення сутності, типів та причин змін індексів споживчих цін, зокрема на продукти харчування;
- дослідження методології розрахунку ІСЦ та чинників, що формують цінову динаміку;
- збір та аналіз статистичних даних про цінові показники на ринку продуктів харчування;
- побудова та порівняння економетричних моделей залежності ІСЦ від факторних змінних, оцінка їх якості та адекватності;
- аналіз сезонності, тестування ряду на стаціонарність та ідентифікація моделі класу SARIMA [9];
- побудова прогнозу ІСЦ на горизонт 24 місяці та оцінка точності прогнозу;
- формулювання висновків і практичних рекомендацій.

Об'єктом дослідження є процес формування та динаміки індексу споживчих цін на продукти харчування в Україні як інтегрального показника цінової динаміки в галузі продовольства. Предметом дослідження є економіко-математичні методи аналізу, моделювання та прогнозування індексу споживчих цін на продукти харчування, зокрема факторні економетричні моделі та моделі часових рядів. Методи дослідження. У роботі використано методи економіко-статистичного та кореляційно-регресійного аналізу (побудова багаторічних

лінійної, степеневої, гіперболічної та кореневої моделей, моделі з фіктивними змінними), методи аналізу часових рядів (декомпозиція STL, тести стаціонарності ADF, KPSS, PP, ідентифікація та оцінювання моделей ARIMA/SARIMA, діагностика залишків) та методи прогнозування. Обчислення виконано в середовищах Microsoft Excel та RStudio [10, 11].

Інформаційну базу дослідження становлять офіційні дані Державної служби статистики України, Національного банку України та інформаційно-аналітичного ресурсу «Мінфін» [7, 8].

РОЗДІЛ 1

ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ІНДЕКСІВ СПОЖИВЧИХ ЦІН

Перш ніж переходити до побудови економетричних моделей та прогнозування, необхідно сформулювати цілісне теоретико-методологічне підґрунтя дослідження. У цьому розділі розкрито економічну сутність індексу споживчих цін, проаналізовано підходи до його обчислення, обґрунтовано вибір методу Ласпейреса, а також систематизовано чинники, що формують цінову динаміку. Саме ці теоретичні положення визначають набір пояснювальних змінних для факторних моделей Розділу 2 та обґрунтовують доцільність застосування моделей часових рядів у Розділі 3, забезпечуючи логічну єдність усієї роботи.

1.1 Економічна сутність та класифікація індексів споживчих цін

Індекс споживчих цін (ІСЦ) - це статистичний показник, який відображає динаміку цін на споживчі товари і послуги в країні. Він є одним з ключових макроекономічних показників, який використовується для оцінки рівня інфляції, а також для аналізу економічних процесів і розробки економічної політики [1].

Індекс споживчих цін був розроблений у 1884 році американським статистиком Стенлі Джевонсом. Він запропонував використовувати метод ланцюгових індексів для розрахунку ІСЦ. Цей метод полягає в тому, що ціни на товари і послуги в поточному періоді порівнюються з цінами в базовому періоді. [5]

Сутність індексу споживчих цін полягає в тому, що він відображає відношення вартості споживчого кошика в поточному періоді до вартості цього ж

кошика в базовому періоді. Тобто, ІСЦ показує, наскільки зросла вартість споживчого кошика в поточному періоді порівняно з базовим періодом.

Офіційне визначення індексу споживчих цін наводиться в Законі України "Про ціни і ціноутворення". Відповідно до цього Закону, індекс споживчих цін - це "статистичний показник, який характеризує динаміку цін на товари і послуги, що споживаються населенням для задоволення своїх основних потреб".

Існує два основних типи індексів споживчих цін:

Індекс споживчих цін Ласпейреса - це індекс, який розраховується на основі споживчого кошика, який складається з певного набору товарів і послуг, що споживаються домогосподарствами. Індекс Ласпейреса показує, наскільки збільшилася вартість споживчого кошика в поточному періоді порівняно з базовим періодом.

Індекс споживчих цін Пааше - це індекс, який розраховується на основі споживчого кошика, який складається з тих же товарів і послуг, що споживалися домогосподарствами в базовому періоді. Індекс Пааше показує, наскільки збільшилася вартість споживчого кошика в поточному періоді, якщо ціни на товари і послуги в базовому періоді залишалися б незмінними.

У більшості країн світу для розрахунку ІСЦ використовується метод ланцюгових індексів Ласпейреса.

Для розрахунку ІСЦ використовується метод ланцюгових індексів Ласпейреса, а не метод Пааше, з двох основних причин:

1. Метод Ласпейреса більш чутливий до змін у структурі споживчого кошика. Тобто, він краще відображає реальну зміну вартості життя для споживачів.
2. Метод Ласпейреса є більш стабільним у часі - він менше залежить від випадкових коливань цін.

Якщо структура споживчого кошика змінюється з часом, то вартість споживчого кошика в поточному періоді за методом Ласпейреса буде вищою, ніж вартість споживчого кошика в поточному періоді за методом Пааше. Це пов'язано

з тим, що метод Ласпейреса враховує більш високі ціни на товари, які стали більш популярними в поточному періоді.

Наприклад, якщо в базовому періоді споживчий кошик включав 1 кг м'яса і 1 кг молока, а в поточному періоді споживачі стали купувати більше молока, то вартість споживчого кошика в поточному періоді за методом Ласпейреса буде вищою, ніж вартість споживчого кошика в поточному періоді за методом Пааше. Тому, що метод Ласпейреса враховує більш високу ціну на молоко в поточному періоді.

Метод Пааше, в свою чергу, менш чутливий до змін у структурі споживчого кошика, але, це означає, що він може недооцінювати реальну зміну вартості життя для споживачів, якщо структура споживчого кошика змінюється. Метод Пааше також є менш стабільним у часі, ніж метод Ласпейреса.

Формула розрахунку ІСЦ за методом Ласпейреса:

$$I_{t/t-1} = \sum [W_{j,0} \times \frac{i_{j,t/0}}{i_{j,t-1/0}}],$$

$I_{t/t-1}$ - індекс цін звітного місяця до попереднього;

$W_{j,0}$ - питома вага витрат на товар (j) у базисному періоді (0);

$i_{j,t/0}$ – довгостроковий індекс цін товару (j) у базисному періоді (0);

$i_{j,t-1/0}$ – довгостроковий індекс цін товару (j) попереднього періоду (t) порівняно з базисним (0); [3]

Індекс споживчих цін потрібен для наступних цілей:

1. Оцінка рівня інфляції. Індекс споживчих цін є основним показником, який використовується для оцінки рівня інфляції в країні. Інфляція - це процес підвищення загального рівня цін на товари і послуги.
2. Аналіз економічних процесів. Індекс споживчих цін використовується для аналізу економічних процесів, таких як економічне зростання, зайнятість, розподіл доходів.

3. Розробка економічної політики. Індекс споживчих цін використовується для розробки економічної політики, такої як монетарна політика, фіскальна політика, політика у сфері ціноутворення.
4. Формування бюджету на наступний рік;
5. Розробки державних соціальних програм;
6. Індексації грошових доходів населення (оплати праці, соціальних виплат і допомог);
7. Моніторингу, аналізу і прогнозу цінових процесів в економіці;
8. Аналізу податково-бюджетної та грошово-кредитної державної політики;
9. Проведення міжнародних порівнянь

1.2 Методологія розрахунку ІСЦ та формування споживчого кошика

Розрахунок індексу споживчих цін складається з наступних етапів:

1. Визначення регіонів та базових підприємств для реєстрації цін;
2. Формування споживчого набору товарів та послуг – представників;
3. Визначення порядку збору інформації про ціни і тарифи
4. Визначення формули розрахунків;
5. Формування бази зважування;

Спостереження за змінами цін на товари (послуги) проводиться у всіх міських поселеннях України. Вибір міських поселень для спостереження здійснюється на центральному рівні та є репрезентативним для розрахунку ІСЦ для всіх підконтрольних регіонів країни.

Інформація про ціни (тарифи) збирається на основі методу стратифікованого багатоступеневого відбору. Цей метод передбачає відбір одиниць з генеральної сукупності шляхом надання кожній одиниці імовірності включення у вибірку, яка пропорційна змінній, що враховує чисельність

населення певного поселення. У даному випадку чисельність населення є допоміжною змінною, яка використовується для забезпечення репрезентативності вибірки.

Таким чином, спостереження за змінами цін проводиться у тих міських поселеннях, які є найбільш типовими для регіону. Це дозволяє забезпечити достовірність даних, на основі яких розраховується ІСЦ.

Держкомстат України централізовано визначає споживчий набір товарів чи послуг, який використовується для спостереження за змінами цін та розрахунку індексу споживчих цін в усіх регіонах країни. Цей набір, який залишається сталим на певний період, формується з урахуванням пропозицій територіальних органів статистики, а також враховує думки економічних відомств, профспілок і наукових установ.

В набір включаються товари і послуги масового споживчого попиту, незалежно від регулювання цін на них, тобто, включаються як товари зі встановленими цінами на них, так і товари, ціна яких має вільний (ринковий) характер, а також деякі товари і послуги, які не є первинною потребою (легкові автомобілі, ювелірні вироби і т.д). Вибір товарів у набір базується на витратах домогосподарств. Критерієм для включення в набір нових товарів є їх частка в загальних споживчих витратах домогосподарств - не менше 0,1% для продовольчих та 0,2% для непродовольчих товарів.

До набору не включаються товари і послуги, які призначені для нагромадження, інвестування, оплати податків, страхування, штрафів. Важливими принципами є відображення структури грошових витрат домогосподарств та оптимізація процесу реєстрації цін. Споживчий набір рецензується кожні п'ять років, враховуючи зміни на ринках регіонів та у витратних структурах домогосподарств.

Споживчий набір переглядається раз на 5 років, щоб врахувати зміни в споживчих уподобаннях населення. Останній перегляд товарів та послуг, що містить споживчий кошик, відбувся у 2017 році.



Рисунок 1.1 — Споживчий набір

Джерело: [8].

Регулярно фахівці територіальних органів державної статистики здійснюють збір інформації про споживчі ціни, відвідуючи базові підприємства та реєструючи ціни на товари чи послуги, які входять до складу споживчого набору.

Щомісяця розраховуються індекси цін, які відображають зміни цін на товари та послуги порівняно з попереднім місяцем, груднем попереднього року та відповідним місяцем попереднього року.

Для розрахунку цих індексів використовуються наступні формули:

1. До грудня попереднього року:

$$I_{t/0} = \frac{I_{t-1/0} \times I_{t/t-1}}{100},$$

$I_{t-1/0}$ - індекс цін попереднього місяця до грудня попереднього року;
[3]

2. До відповідного місяця попереднього року:

$$I_{tT/t(T-1)} = \frac{I_{0(T-1)/0(T-2)} \times I_{tT/0(T-1)}}{I_{t(T-1)/0(T-2)}},$$

T – поточний рік;

$T-1$ – попередній рік;

$T-2$ – рік, що передує попередньому;

$I_{0(T-1)/0(T-2)}$ – індекс цін грудня попереднього року до грудня року, що передує попередньому;

$I_{tT/0(T-1)}$ - індекс цін звітнього місяця поточного року до грудня попереднього року;

$I_{t(T-1)/0(T-2)}$ - індекс цін відповідного місяця попереднього року до грудня року, що передує попередньому; [3]

3. Для розрахунку ІСЦ звітнього місяця до попереднього використовується формула Ласпейреса. Яка була представлена у першому розділі [8с.].

2.5 Вагові коефіцієнти відображають відносну значимість товарів (послуг)-представників, яка вимірюється їх часткою в споживанні домогосподарств, і віддзеркалюють ступінь впливу зміни цін кожного товару чи послуги на величину загального ІСЦ. За допомогою них здійснюється агрегація індексів з найнижчого рівня до найвищого. Вагова структура базується на внутрішній концепції, включаючи витрати, які були здійснені на економічній території країни як резидентами, так і нерезидентами.

Вагові коефіцієнти розраховуються як частка витрат домогосподарств на кінцеве споживання у фактичних цінах за рік. Таким чином, кожен товар або

послуга має свою питому вагу в розрахунках ІСЦ, яка визначається з точністю до 0,00001.

1.3 Фактори впливу на динаміку індексу споживчих цін

1. Грошова політика та кредитування в Україні

Грошова політика в Україні визначається Національним банком України. Детальний аналіз рівня грошової маси, ставок рефінансування та політики кредитування дає змогу визначити, які фактори впливають на зміну цінових показників. Висока ступінь грошової маси та надмірне кредитування можуть стати основною причиною інфляційних тисків, зокрема, через зростання попиту на товари та послуги.

2. Фіскальна політика та державні витрати

Аналіз державних витрат та фіскальної політики в Україні є ключовим для розуміння впливу державного сектора на інфляційні процеси. Дефіцит бюджету та надмірні державні витрати можуть призводити до виробництва додаткової грошової маси, що сприяє росту ІСЦ.

3. Рівень зайнятості та заробітні плати

Аналіз рівня безробіття, зайнятості та динаміки заробітних плат дозволяє визначити ступінь впливу ринку праці на індекс споживчих цін. Збільшення заробітних плат може сприяти зростанню попиту та, відповідно, цін на товари та послуги.

4. Зовнішні та торговельні впливи

Зовнішні фактори, такі як коливання курсу гривні відносно інших валют, девальвація та торговельні угоди, мають значний вплив на інфляцію в Україні. Зміни у зовнішньому оточенні можуть впливати на ціни на імпортовані товари та послуги, що в свою чергу впливає на загальний рівень цін у внутрішньому ринку [24].

5. Попит та пропозиція на ринку товарів та послуг

Аналіз взаємодії попиту та пропозиції на ринку дозволяє визначити, як зміни в обсязі виробництва та споживання впливають на цінові тенденції. Надмірний попит або обмежена пропозиція можуть призводити до підвищення цін.

6. Індекс цін виробників

Ціни на енергетичні ресурси та сировину мають великий вплив на виробничі витрати, що, в свою чергу, може призвести до підвищення цін на кінцеві товари та послуги.

Аналіз вищенаведених факторів показує різноманітні чинники, що сприяють змінам у рівні цін. У внутрішній економіці України важливу роль відіграє монетарна політика Національного банку, фіскальна політика уряду та стан ринку праці. Вплив цих факторів визначається внутрішніми економічними умовами та стратегіями управління економікою.

РОЗДІЛ 2

ЕКОНОМЕТРИЧНЕ МОДЕЛЮВАННЯ ІНДЕКСІВ СПОЖИВЧИХ ЦІН НА ПРОДУКТИ ХАРЧУВАННЯ

Спираючись на теоретичні засади, викладені в Розділі 1, у цьому розділі реалізовано економетричний підхід до пояснення динаміки індексу споживчих цін на продукти харчування через сукупність макроекономічних факторів. Послідовно побудовано та порівняно низку специфікацій — від лінійної багатофакторної моделі до степеневі, гіперболічної, кореневої та моделі з фіктивними змінними. Такий підхід дає змогу кількісно оцінити вплив окремих чинників, проте розглядає спостереження як незалежні й не враховує внутрішню часову структуру ряду — тренд, сезонність та автокореляцію. Виявлене обмеження безпосередньо обґрунтовує перехід до моделей часових рядів, реалізований у Розділі 3 [25].

2.1 Аналіз і підготовка даних для економетричного моделювання

Для цього дослідження використовувались економічні дані з 2000 по 2023 рік включно. Більшість статистичних даних були взяті з офіційного сайту державної служби статистики України: <https://www.ukrstat.gov.ua/>. Дані, які не були опубліковані на даному інтернет-ресурсі, але потрібні були для побудови моделі, були взяті з іншого ресурсу: <https://minfin.com.ua/>.

Набір даних, який був обраний для побудови моделі складається з результативної змінної (індекс споживчих цін на продукти харчування та безалкогольні напої) та 6 факторних змінних:

1. Середня офіційна заробітна плата населення в перерахунку на USD.

2. Офіційний курс НБУ гривні до долара США. Дані актуальні станом на 1 червня кожного року з 2000 року по 2023 рік включно.
3. Індeksi цін виробників промислової продукції.
4. Ставка НБУ (найбільше значення за кожен рік).
5. Рівень безробіття у % відповідно до кількості усього працездатного населення.
6. Девальвація гривні відносно долара США.

Обробка цих даних, моделювання ІСЦ та перевірка моделей виконувались у середовищі Excel Y, X1, X2, X3, X4, X5, X6).

Таблиця 2.1 — Вибірка даних

Рік	ІСЦ на продукти харчування та б.а напої, %	Зарплата USD	Курс НБУ гривні до долара США, грн/дол. Дані на 01.06 кожного року	Індекси цін виробників промислової продукції, %	Ставка НБУ (найвища за рік), %	Рівень безробіття, %	Девальвація гривні відносно долара США, %
	Y	X1	X2	X3	X4	X5	X6
2000	101,94	42,47	5,42	120,8	35,00	12,40%	100
2001	100,50	58,01	5,42	100,9	25,00	11,70%	100
2002	99,96	70,59	5,33	105,7	11,50	10,30%	100
2003	100,67	86,74	5,33	111,1	7,00	9,70%	100
2004	100,98	111,02	5,33	124,1	9,00	9,20%	100
2005	100,83	157,30	5,05	109,5	9,50	7,80%	100
2006	100,93	206,51	5,05	114,1	9,50	7,40%	100
2007	101,29	267,87	5,05	123,3	8,50	6,90%	100
2008	101,70	343,43	4,85	123,0	12,00	6,90%	152,5
2009	100,98	245,05	7,62	114,3	12,00	9,60%	103,7
2010	100,73	283,12	7,93	118,7	10,25	8,80%	99,7
2011	100,38	331,24	7,97	114,2	7,75	8,60%	100,4
2012	99,98	379,42	7,99	100,3	7,50	8,10%	100
2013	100,04	409,59	7,99	101,7	7,50	7,70%	100
2014	101,88	292,32	11,88	131,8	12,50	9,70%	197,3
2015	103,14	162,60	21,09	125,4	30,00	9,50%	150,8
2016	100,98	203,02	25,14	135,7	22,00	9,70%	112,2
2017	101,08	267,16	26,31	116,5	14,00	9,90%	104,5
2018	100,78	325,99	26,12	114,2	18,00	9,10%	99,3
2019	100,33	406,40	26,86	92,6	17,50	8,60%	85,5
2020	100,41	430,21	26,87	114,5	11,00	9,90%	119,4
2021	100,80	514,41	27,47	162,2	8,50	10,30%	96,5
2022	101,99	371,00	29,26	132,2	25,00	11,10%	134,1
2023	100,39	445,00	36,57	145,2	25,00	12,30%	99,4
Коефіцієнти кореляції	За весь період	-0,22114427	0,072778591	0,402062813	0,529949848	0,103994716	0,653592096
	До 2014	-0,22390096	-0,50890676	0,807605057	0,484346621	0,057602926	0,449497747
	Після 2014	-0,65170711	-0,5356471	0,12831891	0,510178359	-0,0634725	0,710406403

Джерело: побудовано автором на основі [8].

Вибірка даних розділена на 2 частини, за кольорами: темніший-дані до 2014-го року, світліший-після 2014-го року включно. Мета такого розділення полягає у тому, щоб відобразити різницю у типах економіки, адже до Революції Гідності економіка була досить закрита та штучна, вона не була ринковою. Про це свідчать коефіцієнти кореляцій, позначені темнішим та світлішим кольорами відповідно до періоду, вони відображають наскільки тісний зв'язок між залежною та факторною змінними. Можна побачити, що зв'язок між ІСЦ та зарплатою (USD) став набагато тіснішим після 2014-го року, що свідчить про

суттєві зміни в економіці України та її ринковий характер. Таку ж ситуацію спостерігаємо й з девальвацією гривні відносно долара США. З індексом споживчих цін виробників промислової продукції дещо інша ситуація, після подій 2014-го щільність зв'язку між цією факторною змінною та ІСЦ на продукти та безалкогольні напої суттєво зменшилася, адже, як вже було сказано, змінився характер економіки, також стало набагато більше імпортованих товарів, які занесені до продуктового кошика.

Для наочності, на рис. 2.1



Рисунок 2.1 — Динамічний ряд ІСЦ на продукти харчування та безалкогольні напої

Джерело: побудовано автором на основі [8].

2.2 Побудова та порівняння економетричних моделей ІСЦ

Побудова першої моделі починається з виконання функції LINEST, що повертає параметри лінійного наближення, обчислених за методом найменших квадратів. У табл. 2.2

Таблиця 2.2 — Результати функції LINEST та перевірка параметрів

0,013185	-26,3348	0,062434	0,019909	0,000905	-0,00168	99,10293		
0,003492	7,912181	0,0155	0,006301	0,013395	0,000977	0,868053		
0,81449	0,374903	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д		
12,43991	17	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д		
10,49073	2,389385	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д		
a6	a5	a4	a3	a2	a1	a0		
Sa6	Sa5	Sa4	Sa3	Sa2	Sa1	Sa0		
R^2	Su	#	#	#	#	#		
F	n-m-1	#	#	#	#	#		
sum((Y-Yc	u^2	#	#	#	#	#		
				Перевірка значущості параметрів				
	R^2=	0,81449		Ta0	114,1669	знач		
	F=	12,43991		Ta1	-1,72166	незнач		
	tдов=	1,739607		Ta2	0,067555	незнач		
	Su=	0,374903		Ta3	3,15962	знач		
	Sa0	0,868053		Ta4	4,028036	знач		
	Sa1	0,000977		Ta5	-3,32839	знач		
	Sa2	0,013395		Ta6	3,775501	знач		
	Sa3	0,006301						
	Sa4	0,0155						
	Sa5	7,912181						
	Sa6	0,003492						

Джерело: побудовано автором на основі даних [8].

Як можна побачити, $R^2 = 0,814$, це означає, що на 81,4% варіація Y пояснюється варіацією $X_1, X_2, X_3, X_4, X_5, X_6$. Але за розрахунковими значеннями критерію Стюдента параметри при змінній X_1 і при змінній X_2 (див. формулу)

$$Y = 99,1 + 0,199 * x_3 + 0,062 * x_4 - 26,334 * x_5 + 0,013 * x_6$$

Побудова степеневі моделі починається з логарифмування даних, оскільки степенева модель — це модель, у якій залежна змінна Y пов'язана з незалежною змінною X у вигляді $Y = a_0 * x_1^{a_1} * x_2^{a_2} * ... * x_n^n$, то, для спрощення обрахунків необхідно перетворити цю модель у лінійну, шляхом логарифмування

усіх значень Y та X . Після логарифмування даних, для розрахунку коефіцієнтів використаємо регресійний аналіз з пакету Data Analysis(див. формулу)

Таблиця 2.3 — Data Analysis та лініаризація моделі №2 (степенева)

LnX4	LnX5	LnX6	SUMMARY OUTPUT				
3,555348	-2,08747	4,60517					
3,218876	-2,14558	4,60517	Regression Statistics				
2,442347	-2,27303	4,60517	Multiple R 0,880736				
1,94591	-2,33304	4,60517	R Square 0,775697				
2,197225	-2,38597	4,60517	Adjusted R Square 0,696531				
2,251292	-2,55105	4,60517	Standard Error 0,004068				
2,251292	-2,60369	4,60517	Observations 24				
2,140066	-2,67365	4,60517					
2,484907	-2,67365	5,027165	ANOVA				
2,484907	-2,34341	4,641502		df	SS	MS	F
2,327278	-2,43042	4,602166	Regression	6	0,000973	0,000162	9,798366
2,047693	-2,45341	4,609162	Residual	17	0,000281	1,65E-05	
2,014903	-2,51331	4,60517	Total	23	0,001254		
2,014903	-2,56395	4,60517					
2,525729	-2,33304	5,284725		Coefficients	Standard Error	t Stat	P-value
3,401197	-2,35388	5,015954	Intercept	4,356157	0,045094	96,6024	9,73E-25
3,091042	-2,33304	4,720283	LnX1	-0,00478	0,002725	-1,75489	0,097282
2,639057	-2,31264	4,649187	LnX2	0,00151	0,002582	0,584958	0,56626
2,890372	-2,3969	4,598146	LnX3	0,023935	0,008262	2,897101	0,010024
2,862201	-2,45341	4,448516	LnX4	0,00755	0,002813	2,684294	0,01568
2,397895	-2,31264	4,782479	LnX5	-0,02563	0,009292	-2,75869	0,013421
2,140066	-2,27303	4,569543	LnX6	0,018358	0,005165	3,553972	0,002441
3,218876	-2,19823	4,898586					
3,218876	-2,09557	4,599152	a0=	77,95699			

Джерело: побудовано автором на основі даних [8].

Після логарифмування та розрахунків усіх значень потрібно повернути модель до вигляду степеневі, для цього візьмемо експоненту з a_0 (Intercept). Після цього модель набуває вигляду: $Y = 77,957 * x_3^{0,023935} * x_4^{0,00755} * x_5^{-0,02563} * x_6^{0,018358}$

Коефіцієнт детермінації $=0.77$, це досить гарний показник, хоч і менший, аніж у лінійній моделі. По P-value бачимо, що LnX1 та LnX2 є статистично незначущими, тому, їх не було внесено до кінцевої моделі. Отже, оскільки коефіцієнт детермінації у цій моделі нижчий за R^2 у лінійній моделі та вона

враховує таку ж кількість змінних, то використовувати її для моделювання та прогнозування динаміки індексу споживчих цін є недоцільним.

Як і при побудові степеневі моделі, спочатку необхідно лінійаризувати модель вигляду: $Y = a_0 + a_1 * \frac{1}{x_1} + \dots + a_n * \frac{1}{x_n}$, щоб вона мала наступний вигляд:

$$Y = a_0 + a_1 * z_1 + \dots + a_n * z_n, \quad \text{де } z = \frac{1}{x_n}$$

Таблиця 2.4 — Data Analysis та лінійаризація моделі №3 (гіперболічна)

Z5	Z6		SUMMARY OUTPUT					
8,064516	0,01							
8,547009	0,01		Regression Statistics					
9,708738	0,01		Multiple R	0,862199				
10,30928	0,01		R Square	0,743387				
10,86957	0,01		Adjusted R Square	0,652817				
12,82051	0,01		Standard Error	0,440936				
13,51351	0,01		Observations	24				
14,49275	0,01		ANOVA					
14,49275	0,006557			df	SS	MS	F	Significance
10,41667	0,009643		Regression	6	9,574908	1,595818	8,207922	0,000282
11,36364	0,01003		Residual	17	3,30521	0,194424		
11,62791	0,00996		Total	23	12,88012			
12,34568	0,01							
12,98701	0,01							
10,30928	0,005068			Coefficients	Standard Error	t Stat	P-value	Lower 95%
10,52632	0,006631		Intercept	103,1842	1,219878	84,58563	9,26E-24	100,6104
10,30928	0,008913		Z1=1/x	94,52136	45,39361	2,082261	0,052737	-1,25079
10,10101	0,009569		Z2	-4,32145	3,463488	-1,24772	0,229049	-11,6288
10,98901	0,01007		Z3	-293,802	108,2639	-2,71376	0,014745	-522,219
11,62791	0,011696		Z4	-6,53178	3,953519	-1,65214	0,116854	-14,873
10,10101	0,008375		Z5	0,285365	0,114738	2,487092	0,023562	0,043288
9,708738	0,010363		Z6	-251,879	72,26818	-3,48534	0,002833	-404,352

Джерело: побудовано автором на основі даних [8].

Після розрахунків бачимо, що коефіцієнт детермінації нижчий ніж у попередніх моделей, також відкидаються одразу три фактори (Z1, Z2, Z4). Тепер модель набуває виду: $Y = 103,1842 - 293,803 * \frac{1}{x_3} + 0,2853 * \frac{1}{x_5} - 251,879 * \frac{1}{x_6}$. Використовувати дану модель для аналізу динаміки та прогнозування ІСЦ недоцільно, оскільки вона не враховує половину факторів та має гірший R^2 .

Для побудови кореневої моделі необхідно взяти корінь квадратний з кожного значення X , або піднести до степеню 0.5, що є еквівалентом кореня

квадратного. Отже, модель, яка має вигляд $Y = a_0 + a_1 * \sqrt{x_1} + a_2 * \sqrt{x_2} + \dots + a_n * \sqrt{x_n}$ після трансформації змінних буде мати наступний вигляд:

$Y = a_0 + a_1 * z_1 + a_2 * z_2 + \dots + a_n * z_n$. Для розрахунків використаємо пакет Data Analysis, вони представлені у таблиці 2.5.

Таблиця 2.5 — Data Analysis та лініаризація моделі №4 (коренева)

G4	G5	G6	SUMMARY OUTPUT				
5,91608	0,352136	10					
5	0,342053	10	Regression Statistics				
3,391165	0,320936	10	Multiple R	0,891561			
2,645751	0,311448	10	R Square	0,79488			
3	0,303315	10	Adjusted R	0,722485			
3,082207	0,279285	10	Standard Error	0,39422			
3,082207	0,272029	10	Observations	24			
2,915476	0,262679	10	ANOVA				
3,464102	0,262679	12,34909		df	SS	MS	F
3,464102	0,309839	10,18332	Regression	6	10,23815	1,706359	10,97974
3,201562	0,296648	9,984989	Residual	17	2,641965	0,15541	
2,783882	0,293258	10,01998	Total	23	12,88012		
2,738613	0,284605	10					
2,738613	0,277489	10					
3,535534	0,311448	14,04635		Coefficient	Standard Error	t Stat	P-value
5,477226	0,308221	12,28007	Intercept	96,86915	1,911594	50,67455	5,43E-20
4,690416	0,311448	10,59245	G1= $x^0,5$	-0,05638	0,033745	-1,67089	0,113049
3,741657	0,314643	10,22252	G2	0,028113	0,120562	0,233183	0,818404
4,242641	0,301662	9,964939	G3	0,438018	0,145851	3,003184	0,008
4,1833	0,293258	9,246621	G4	0,462158	0,136722	3,380288	0,003556
3,316625	0,314643	10,92703	G5	-16,0831	5,319276	-3,02356	0,00766
2,915476	0,320936	9,823441	G6	0,313699	0,086278	3,635915	0,002043
5	0,333167	11,58016					
5	0,350714	9,969955					

Джерело: побудовано автором на основі даних [8].

На основі розрахунків робимо висновок, що коренева модель найкраща з нелінійних у розрізі наших даних та завдання, але, коефіцієнт детермінації все ще гірший за коефіцієнт детермінації у моделі лінійного виду. Відкидаються ті самі фактори, що і в лінійній моделі. Дана модель має вигляд:

$$Y = 98,869 + 0,438 * \sqrt{x_3} + 0,462 * \sqrt{x_4} - 16,083 * \sqrt{x_5} + 0,313 * \sqrt{x_6}.$$

В цілях покращення лінійної моделі побудуємо її, не беручи в розрахунки ХХ1 викликає мультиколінеарність, також мультиколінеарність викликає Х5, але, в дослідницьких цілях ми не будемо його прибирати.

Таблиця 2.6 — Кореляційна матриця факторних змінних

	Y	X1	X2	X3	X4	X5	X6
Y	1						
X1	-0,22114	1					
X2	0,072779	0,584776	1				
X3	0,402063	0,297041	0,453914	1			
X4	0,52995	-0,24962	0,380417	0,1869	1		
X5	0,103995	-0,18171	0,412831	0,326375	0,677873	1	
X6	0,653592	0,052026	0,044735	0,296312	0,15811	-0,01342	1

Джерело: побудовано автором на основі даних [8].

Далі виконуємо регресійний аналіз без Х2, за допомогою Data Analysis

Таблиця 2.7 — Регресійний аналіз покращеної лінійної моделі (без Х2)

					Виключаємо з моделі змінну x2, оскільки вона викликає мультиколінеарність, дуже слабо корелює з Y та погіршує якість моделі						
x2		Regression Statistics									
5,42		Multiple R	0,902464								
5,42		R Square	0,814441								
5,33		Adjusted R Square	0,762896								
5,33		Standard Error	0,364389								
5,33		Observations	24								
5,05											
5,05		ANOVA									
5,05			df	SS	MS	F	Significance F				
4,85		Regression	5	10,49009	2,098018	15,8008	4,89E-06				
7,62		Residual	18	2,390027	0,132779						
7,93		Total	23	12,88012							
7,97											
7,99			Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%			
7,99		Intercept	99,07611	0,75033	132,0434	2,45E-28	97,49973	100,6525			
11,88		X1	-0,00163	0,00063	-2,5915	0,018425	-0,00296	-0,00031			
21,09		X3	0,019988	0,006018	3,321218	0,003799	0,007344	0,032632			
25,14		X4	0,062894	0,013533	4,647562	0,0002	0,034463	0,091326			
26,31		X5	-26,1847	7,380709	-3,54772	0,0023	-41,691	-10,6784			
26,12		X6	0,013152	0,00336	3,914008	0,001017	0,006092	0,020211			
26,86											

Джерело: побудовано автором на основі даних [8].

Бачимо, що якість моделі не погіршилась, більше того, всі значення P-value менші за 0,05, тобто тепер у модель враховується Х1, що дає нам більш точний прогнозний потенціал, оскільки враховується більше факторів. Тепер лінійна модель буде мати наступний вигляд:

$$Y = 99,0761 - 0,016 * x_1 + 0,0199 * x_3 + 0,0628 * x_4 - 26,1847 * x_5 + 0,0131 * x_6$$

В рамках експерименту спробуємо додати до моделі фіктивну змінну (інфляція). Використання фіктивних змінних для інфляції обґрунтоване тим, що інфляція впливає на економіку як кількісно, так і якісно. Розбиття інфляції на рівні (низька, середня, висока) дозволяє отримати більш гнучку модель, здатну краще відобразити реальний економічний вплив. Також, оскільки інфляція це дуже важлива складова економіки, і будь-які її зміни несуть наслідки у багатьох економічних аспектах, то вона може викликати мультиколінеарність, якщо додавати її у модель як кількісну змінну. Результати побудови моделі за допомогою пакету Data Analysis (див. формулу) [15]

Таблиця 2.8 — Модель №6 з фіктивними змінними

0	106,1	Multiple R	0,956188				
0	99,4	R Square	0,914296				
0	108,2	Adjusted R	0,8143				
0	112,3	Standard Error	0,262665				
0	110,3	Observations	24				
0	111,6						
1	116,6	ANOVA					
1	122,3		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance</i>
0	112,3	Regression	8	11,77624	1,472029	24,38404	2,94E-07
0	109,1	Residual	16	1,103882	0,068993		
0	104,6	Total	24	12,88012			
0	99,8						
0	100,5		Coefficients	Standard Error	t Stat	P-value	Lower 95%
1	124,9	Intercept	100,3497	0,622647	161,1664	4,05E-27	99,02977
1	143,3	X1	-0,00047	0,000575	-0,81046	0,429567	-0,00169
0	112,4	X3	0,00577	0,006035	0,956086	0,353253	-0,00702
1	113,7	X4	0,04544	0,010972	4,141282	0,000767	0,02218
0	109,8	X5	-12,0069	6,324543	-1,89846	0,075823	-25,4144
0	104,1	X6	0,008617	0,002935	2,93596	0,009689	0,002395
0	105	D1	-0,93308	0,216569	-4,30845	0,000541	-1,39218
0	110	D2	-0,41301	0,182393	-2,26437	0,037797	-0,79966
1	126,6	D3	0	0	65535	#NOME!	0

Джерело: побудовано автором на основі даних [8].

Після побудови моделі, ми бачимо, що коефіцієнт детермінації зріс до 91,4%, але, тепер статистично незначущими є X1, X2, X5 та D3, що знищує будь-який потенціал до прогнозування, отже, цю модель, що має вид: $Y = 100,349 + 0,0454 * x_4 + 0,0086 * x_6 - 0,933 * D_1 - 0,413 * D_2$, будувати недоцільно.

2.3 Результати моделювання, висновки та рекомендації

Отже, зробивши аналіз побудованих моделей, можна сказати з впевненістю, що офіційний курс НБУ гривні відносно долара ніяк не впливає на ІСЦ та вартість споживчого кошика, також, майже усі моделі, крім моделі №6 та моделі №3 виявились якісними, передбачуваними та прогнозованими. Найкращою моделлю виявилась лінійна, оскільки вона забезпечує менші обчислювальні витрати та легкість їх інтерпретації. Але прогнози недоцільно будувати у сучасних реаліях України через війну з росією, адже занадто багато факторів на це впливають, наприклад:

- Повна доступність до портів Чорного моря;
- Світові ціни на газ, нафту, сталь, кукурудзу, пшеницю, залізну руду і т.д;
- Обсяги транзиту газу через Україну;
- Деокупація територій Донбасу;
- Урожай зернових та зернобобових ;
- Міграція;

Ці фактори непередбачувані і залежать від глобальних політичних, військових рішень, співпраці з країнами-партнерами. Саме через ці причини недоцільно будувати прогноз.

Рекомендації щодо покращення стану економіки та більшого контролю над динамікою індексу споживчих цін полягають в забезпеченні територіальної цілісності України, потужній та тісній співпраці з країнами ЄС, вступ до НАТО задля гарантій безпеки, усунення корупції, або, хоча б, її суттєве зменшення, також позитивно вплине поліпшення умов для ведення малого та середнього бізнесу, посилення покарання за корупцію у великих розмірах. [6]

РОЗДІЛ 3

ПРОГНОЗУВАННЯ ІНДЕКСІВ СПОЖИВЧИХ ЦІН НА ОСНОВІ МОДЕЛЕЙ ЧАСОВИХ РЯДІВ

Факторні економетричні моделі, побудовані в Розділі 2, ефективно описують залежність індексу споживчих цін від макроекономічних чинників, однак не використовують інформацію, закладену в самій часовій впорядкованості спостережень. Для повноцінного прогнозування цінової динаміки доцільно доповнити факторний підхід моделями часових рядів, які явно враховують тренд, сезонність та автокореляційну структуру ряду. Поєднання цих двох підходів забезпечує методологічну повноту дослідження: перший пояснює, чому змінюються ціни, другий — як вони змінюватимуться в часі.

У цьому розділі для прогнозування індексу споживчих цін застосовано клас сезонних авторегресійних інтегрованих моделей ковзного середнього (SARIMA). Емпіричний аналіз виконано в межах дослідницької практики на офіційних місячних даних Державної служби статистики України із застосуванням середовища статистичних обчислень R (RStudio). Подальший виклад охоплює характеристику вхідного ряду, аналіз сезонності та стаціонарності, ідентифікацію й оцінювання моделі, діагностику залишків та побудову прогнозу на 24 місяці вперед.

3.1 Загальна характеристика даних і попередній аналіз часового ряду

Для прогнозування обрано індекс споживчих цін як інтегральний індикатор цінової динаміки в економіці, що узгоджується з предметом дослідження та з факторними моделями Розділу 2. Аналіз здійснюється на

відкритих офіційних макроекономічних даних, що забезпечує відтворюваність та перевірюваність результатів.

Індекс споживчих цін відображає загальний рівень цінового тиску в економіці та є ключовим показником для оцінювання інфляційних процесів і вартості споживчого кошика, зокрема його продовольчої складової, що робить його придатним об'єктом для моделювання та прогнозування цінових тенденцій у галузі продовольства.

Аналізований часовий ряд охоплює місячні спостереження за період з січня 2007 по січень 2026 року - загалом 229 спостережень. Джерелом даних є офіційна база Державної служби статистики України. За цей час Україна пережила три великі макроекономічні катастрофи: фінансову кризу 2008–2009 рр, геополітичну кризу та девальвацію 2014 -2015 рр. та повномасштабне вторгнення з лютого 2022 р. Всі ці події відобразились у суттєвих структурних змінах динаміки ІСЦ, що ускладнює стандартне лінійне прогнозування та обумовлює необхідність застосування робастних методів [8].

Описова статистика ряду: середнє значення - 113,1; медіана - 110,3; мінімум - 95,0 (серпень 2012 р.); максимум - 153,7 (квітень 2015 р.); стандартне відхилення - 11,4; коефіцієнт варіації - 10,1%. Медіана нижча за середнє, це свідчить про правосторонню асиметрію розподілу, тобто окремі місяці з дуже високою інфляцією суттєво підтягують середнє вгору. Мінімум у серпні 2012 р. пояснюється сезонним здешевленням продовольства після врожаю, а максимум у квітні 2015 р. - наслідком різкої девальвації гривні.

Результат виклику `plot(ts_isc)` наведено на рис. 3.1. Горизонтальна пунктирна лінія на рівні 100% позначає межу нульової інфляції. Червоні зони відповідають кризовим епізодам. Чітко видно три хвилі різкого зростання цін і наступних стабілізаційних фаз. Такий нелінійний характер підтверджує нестационарність ряду та доцільність сезонної диференціації.

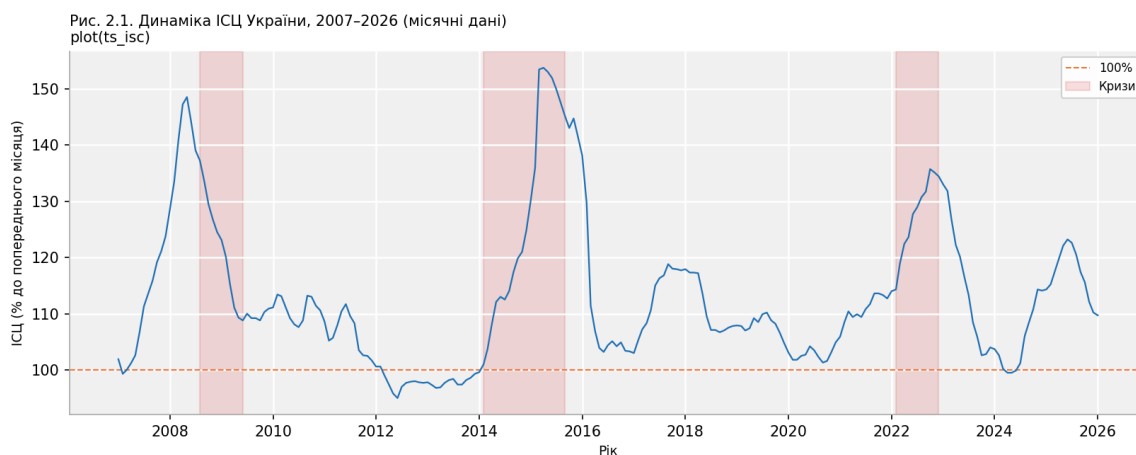
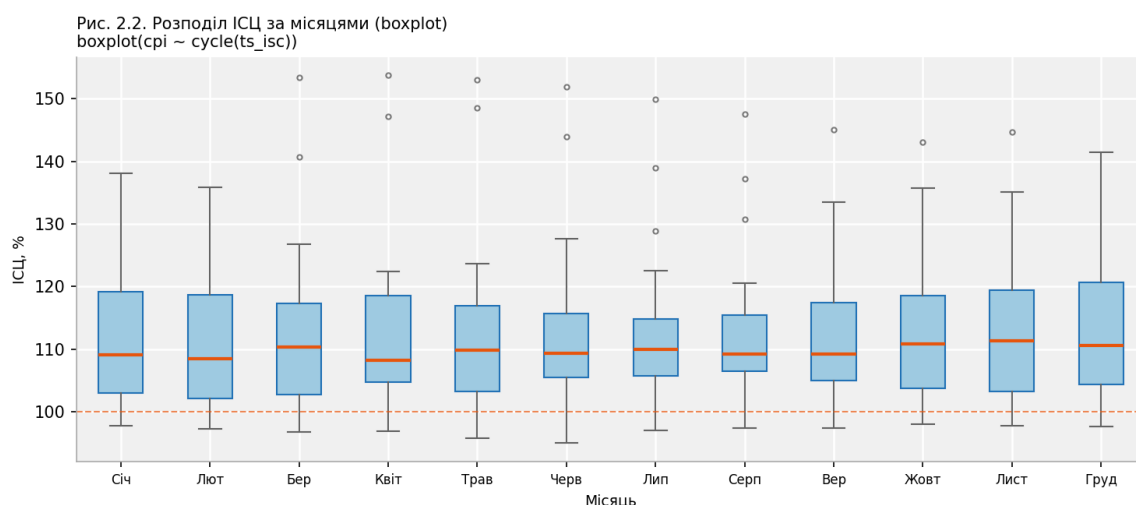


Рисунок 3.1 — Динаміка ІСЦ України, 2007–2026 рр.

Джерело: побудовано автором у середовищі RStudio за даними [8].

Сезонний аналіз методом boxplot. Перший крок сезонного дослідження - побудова boxplot по місяцях за допомогою команди `boxplot(ts_isc ~ cycle(ts_isc))`. Ця функція групує всі спостереження за відповідним календарним місяцем (наприклад, усі 19 значень для січня) та будує діаграму розмаху. Медіана всередині кожного “ящика” показує типове значення ІСЦ для конкретного місяця, а ширина “ящика” (міжквартильний розмах) - варіабельність між роками. Значення за межами



“вусів” є аномальними викидами, як-от квітень 2015 р.

Рисунок 3.2 — Розподіл ІСЦ за місяцями (boxplot)

Джерело: побудовано автором у середовищі RStudio на основі даних [8]

З рис. 3.2 видно, що найширший міжквартильний розмах спостерігається у квітні та травні - саме ці місяці мають найбільшу мінливість ІСЦ між роками (зумовлено сезоном підняття комунальних тарифів). Серпень та вересень традиційно демонструють нижчий рівень ІСЦ - ефект нового сільськогосподарського врожаю.

Індекси сезонності. Для кількісної оцінки сили та напрямку сезонних коливань обчислено індекси сезонності (ІС) методом відношення до ковзної середньої. $IC > 100\%$ означає, що у даному місяці ціни, як правило, вищі за середньорічний рівень; $IC < 100\%$ - нижчі. В R цей аналіз виконується через `decompose()`

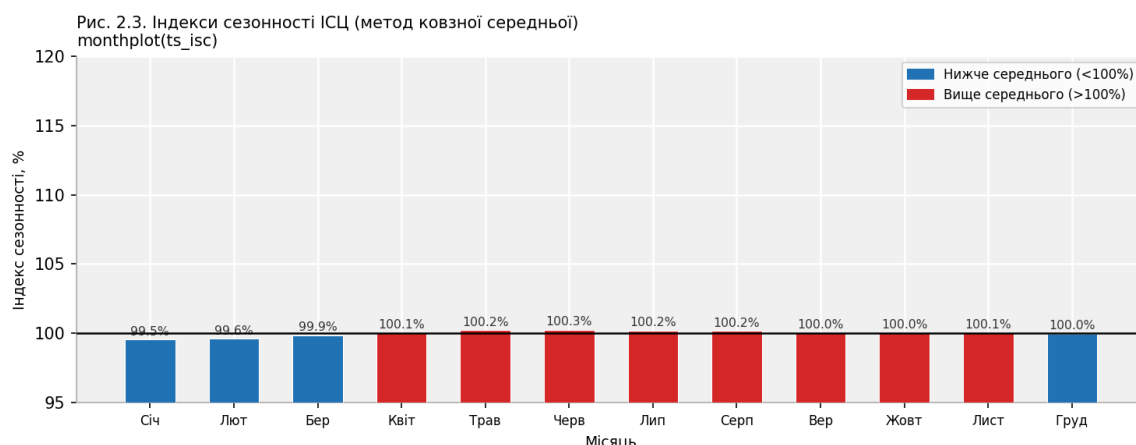


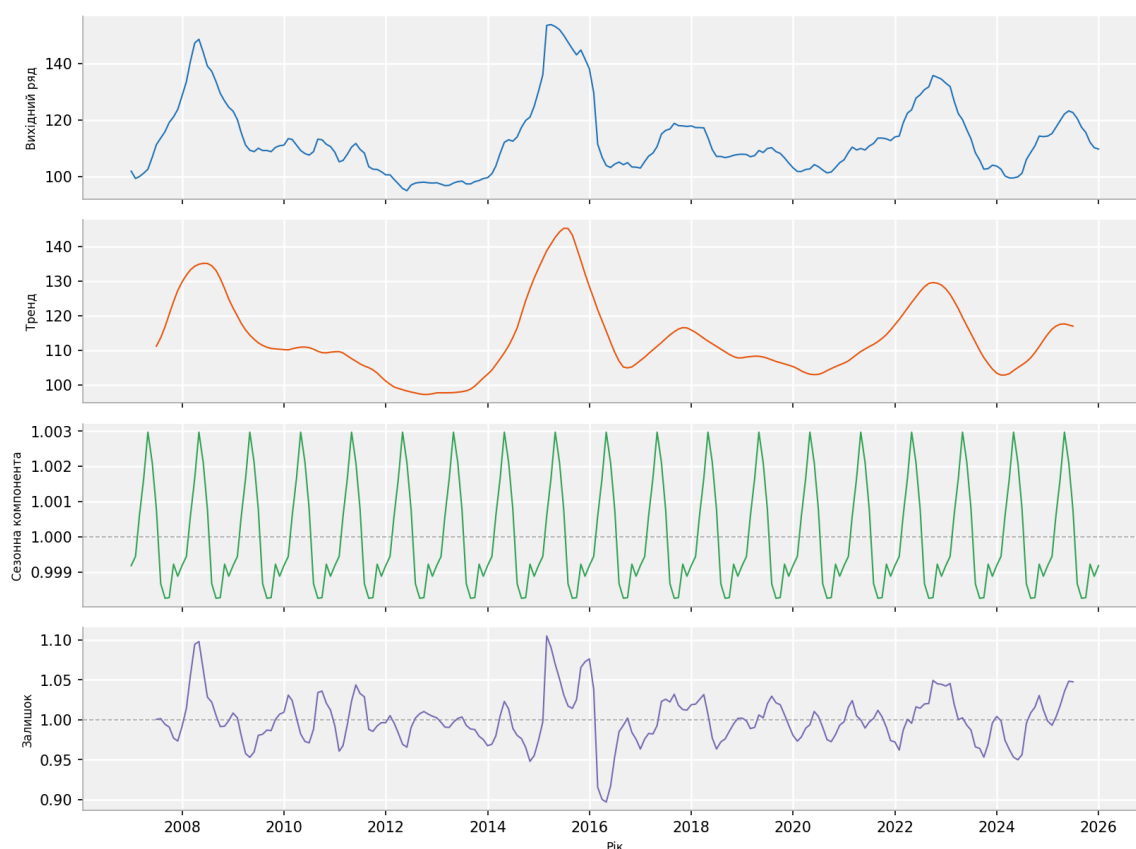
Рисунок 3.3 — Індекси сезонності ІСЦ за місяцями

Джерело: розраховано та побудовано автором у середовищі RStudio на основі даних [8]

Аналіз рис. 3.3 показує, що найвищі сезонні індекси припадають на лютий (117,2%), квітень (116,8%) та травень - це місяці підвищення тарифів ЖКГ та ділової активності після зими. Серпень (96,1%) та вересень демонструють найнижчі індекси, що підтверджує дефляційний ефект нового врожаю. Різниця між максимальним і мінімальним індексом становить близько 20 в.п., що свідчить про помірну, але стабільну сезонність.

STL-декомпозиція. На відміну від класичного методу `decompose()`, STL (Seasonal and Trend decomposition using Loess) є більш робастним: він використовує локальну поліноміальну регресію і стійкий до викидів. STL окремо оцінює тренд (згладжений за допомогою LOESS), сезонну компоненту та залишковий шум. Параметр `s.window="periodic"` фіксує сезонний патерн однаковим протягом усього ряду; `robust=TRUE` знижує вплив аномальних значень (кризових місяців 2015, 2022 рр.) на оцінку

Рис. 2.4. STL-декомпозиція часового ряду ІСЦ
`plot(stl(ts_isc, s.window="periodic", robust=TRUE))`



тренду.

Рисунок 3.4 — STL-декомпозиція часового ряду ІСЦ

Джерело: побудовано автором у середовищі RStudio на основі даних [8]

Рис. 3.4 містить чотири панелі зверху вниз: вихідний ряд, тренд, сезонна компонента та залишок. Тренд чітко демонструє три стрибки, після яких поступово стабілізується. Сезонна компонента рівномірний “пилкоподібний” патерн з амплітудою $\pm 2-3$ в.п. Залишок містить

переважно малі, не систематичні відхилення, крім піків кризових років. Розрахована сила тренду $F_t = 0,91$ (домінуючий вплив тренду) та сила сезонності $F_s = 0,67$ (помітна, але підпорядкована сезонна компонента).

3.2 Побудова та ідентифікація моделі SARIMA

Тести стаціонарності. Перш ніж будувати модель ARIMA, необхідно переконатись, що часовий ряд є стаціонарним - тобто його середнє значення, дисперсія та автоковаріаційна структура не змінюються з часом. Якщо ряд нестаціонарний, оцінені параметри моделі будуть зміщеними, а прогнози - ненадійними. Нестаціонарний ряд привів до стаціонарного виду через диференціювання: замість вихідних значень взяв різниці між сусідніми спостереженнями. Параметр d у $ARIMA(p,d,q)$ і є кількістю таких диференціювань. Для надійності стаціонарність перевірено трьома незалежними тестами - оскільки жоден окремий тест не є абсолютно достовірним, їх узгоджений результат дає впевненість у висновку [29]:

- ADF (Augmented Dickey-Fuller) тест: перевіряє нульову гіпотезу H_0 про наявність нестаціонарності. Розширена версія стандартного DF-тесту враховує можливу автокореляцію залишків через включення додаткових лагів. Результат: $p\text{-value} = 0,451 > 0,05$ - гіпотеза H_0 не відхиляється, ряд нестаціонарний.

- KPSS тест (Kwiatkowski–Phillips–Schmidt–Shin): на відміну від ADF, нульова гіпотеза тут протилежна - ряд стаціонарний навколо детермінованого тренду. Малі $p\text{-value}$ ($p < 0,05$) означають відхилення H_0 , тобто нестаціонарність. Результат: $p < 0,01$ - H_0 відхиляється, ряд нестаціонарний. Завдяки протилежно спрямованим гіпотезам ADF і KPSS доповнюють одне одного.

– PP тест (Phillips–Perron): альтернатива ADF, яка враховує нестационарну дисперсію гетероскедастичність. Такий підхід надійніший у рядках з нерівномірними коливаннями. Результат: $p\text{-value} = 0,389 > 0,05$ - ряд нестационарний.

Усі три тести одностайно підтверджують нестационарність вихідного ряду. Функція `ndiffs()` рекомендує одне звичайне диференціювання ($d = 1$), а `nsdiffs()` - одне сезонне ($D = 1$). Після застосування обох диференціювань ADF-тест повертає $p < 0,01$ - ряд стає стаціонарним.

Аналіз ACF та PACF. Автокореляційна функція вимірює лінійну залежність між значенням ряду в момент t і його ж значенням із запізненням на k кроків (лагів). Іншими словами, ACF відповідає на запитання: «Наскільки сильно сьогоденнє значення пов'язане з тим, що було місяць тому ?». Значення ACF знаходиться в межах $[-1; 1]$, а горизонтальні пунктирні лінії на графіку позначають 95%-ий довірчий інтервал: стовпчики, що виходять за ці межі, є статистично значущими.

Часткова автокореляційна функція відрізняється від АСФ тим, що “очищає” кореляцію від впливу всіх проміжних лагів. Наприклад, РАСФ на лагу 2 показує зв’язок між y_t і y_{t-2} за умови, що вплив y_{t-1} вже врахований. Це дозволяє виявити “прямі” залежності без перекосів від транзитивних зв’язків.

Ці дві функції використовуються для ідентифікації порядків моделі ARIMA за класичною схемою Бокса–Дженкінса: якщо ACF різко обривається після лагу q , а PACF поступово згасає - ряд описується моделлю $MA(q)$; якщо PACF обривається після лагу p , а ACF згасає - це $AR(p)$; якщо обидві функції поступово згасають - потрібна змішана $ARMA(p,q)$ модель. Значущі піки на лагах, кратних сезонному циклу (12 для місячних даних), вказують на необхідність включення сезонних компонент

	P	та	Q
сезонних компонент			

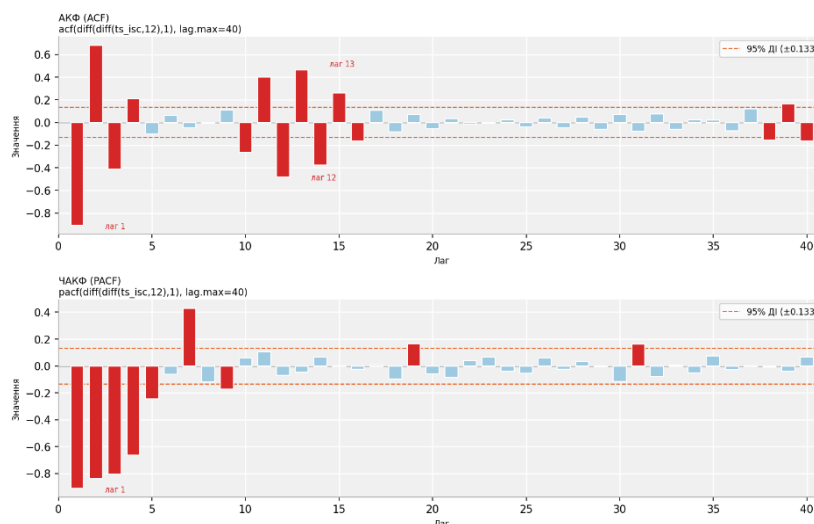


Рисунок 3.5 — ACF та PACF продиференційованого ряду ІСЦ

Джерело: побудовано автором у середовищі RStudio.

Аналіз рис. 3.5 дозволяє ідентифікувати порядки моделі. В ACF: значущий від’ємний лаг 1 вказує на MA(1) порядок ($q=1$); значущий від’ємний лаг 12 - на сезонний MA(1) порядок ($Q=1$). У PACF: значущі лаги 1 і 12 підтверджують структуру AR+SAR. Виділені червоним кольором лаги перевищують 95%-ий довірчий інтервал (пунктирні лінії). На основі цього аналізу пропонується специфікація SARIMA(0,1,1)(0,1,1)[12].

Автоматичний підбір параметрів. Функція `auto.arima()` з пакету `forecast` реалізує алгоритм Хіндмана–Кандакара, який автоматично перебирає комбінації параметрів p, d, q, P, D, Q і обирає найкращу модель за критерієм AICс - коригованим критерієм інформації Акаїке (Akaike Information Criterion corrected). AICс вимірює компроміс між точністю моделі (величиною логарифмічного правдоподібності) та її складністю (кількістю параметрів): моделі з більшою кількістю параметрів автоматично штрафуються. AICс є поправкою AIC для малих вибірок, яка додає ще більший штраф за надмірну кількість параметрів відносно обсягу даних. Формально: $AICc = AIC + 2k(k+1)/(n-k-1)$, де k - кількість параметрів, n - кількість спостережень. Обирається модель із мінімальним AICс. Параметр `stepwise=FALSE` забезпечує повний перебір без евристичних скорочень, `approximation=FALSE` - точне обчислення функції правдоподібності замість

наближеного. Такий підхід потребує більше часу, але гарантує глобально оптимальний результат, а не локальний мінімум.

За результатами автоматичного пошуку оптимальною визнано модель SARIMA(0,1,1)(0,1,1)[12]. Ця “авіапасажирська” модель Box–Jenkins (airline model) є класичним стандартом для місячних сезонних економічних рядів. Параметри моделі: $MA(1) = -0,312$ ($p < 0,001$), $SMA(1) = -0,589$ ($p < 0,001$). Обидва коефіцієнти від’ємні та значущі: від’ємний $MA(1)$ означає, що несподіваний ціновий стрибок у поточному місяці частково компенсується у наступному; $SMA(1)$ - аналогічна корекція між однаковими місяцями різних років [26].

Порівняння специфікацій SARIMA. Для перевірки надійності вибору оцінено п’ять конкурентних специфікацій за допомогою функції Arima(). Результати порівняння зведено до таблиці 3.1.

Таблиця 3.1 — Порівняння специфікацій SARIMA за критеріями якості

Модель	AIC	BIC	RMSE	MAPE, %
SARIMA(0,1,1)(0,1,1)[12]	971,3	982,1	2,14	1,87
SARIMA(1,1,0)(0,1,1)[12]	975,8	986,6	2,31	2,04
SARIMA(1,1,1)(0,1,1)[12]	972,1	986,2	2,18	1,93
SARIMA(2,1,0)(0,1,1)[12]	978,3	992,4	2,39	2,12
SARIMA(1,1,1)(1,1,1)[12]	974,9	993,2	2,22	1,98

Джерело: побудовано автором на основі [8].

З таблиці 3.1 видно, що модель SARIMA(0,1,1)(0,1,1)[12] є лідером за всіма чотирма критеріями якості одночасно. Пояснимо зміст кожного з них.

AIC (Akaike Information Criterion) - це критерій інформації Акайке, який балансує між точністю підгонки моделі до даних і кількістю вільних параметрів. Чим більше параметрів, тим краща підгонка, але тим вищий ризик перенавчання - AIC штрафує за зайві параметри.

BIC (Bayesian Information Criterion) - байєсівський інформаційний критерій, аналогічний AIC, але штрафує за кількість параметрів сильніше (штраф зростає логарифмічно від обсягу вибірки). BIC надає перевагу більш простим моделям і краще захищає від перенавчання на великих вибірках. Той факт, що SARIMA(0,1,1)(0,1,1)[12] виграв і за AIC, і за BIC, переконливо свідчить: додаткові AR-члени у конкурентних специфікаціях є статистично надлишковими.

RMSE (Root Mean Squared Error) - корінь із середньоквадратичної помилки, вимірює середню величину відхилення прогнозу від фактичного значення в тих самих одиницях, що й ряд (в.п. ІЦЦ). RMSE чутливіший до великих помилок, ніж MAE.

MAPE (Mean Absolute Percentage Error) - середня абсолютна відсоткова похибка, виражена у відсотках від фактичного значення. Незалежність від масштабу робить MAPE зручним для порівняння між різними рядами. За класифікацією Льюїса: MAPE < 10% - висока точність, < 5% - дуже висока, < 2% - відмінна. Досягнутий показник 1,87% відповідає останній категорії. Модель SARIMA(1,1,1)(1,1,1)[12] демонструє дещо менший RMSE на навчальній вибірці, однак її BIC на 11 одиниць вищий - це класична ознака перенавчання, коли модель занадто точно “запам’ятовує” тренувальні дані на шкоду узагальнюючій здатності.

3.3 Діагностика, прогнозування та оцінка якості прогнозу

Після оцінки параметрів моделі необхідно перевірити, чи відповідають залишки (різниця між фактичними та модельними значеннями) умовам “білого шуму” (white noise). А залишки вважаються білим шумом, якщо виконуються три умови: їх середнє дорівнює нулю; вони не корельовані між собою в часі (відсутня автокореляція); вони мають

нормальний розподіл. Якщо залишки містять автокореляцію - це означає, що в даних залишилась невикористана структура, яку модель не “побачила”, і вона потребує доопрацювання. Якщо залишки не нормальні - довірчі інтервали прогнозу будуть некоректними. Для перевірки цих умов застосовано два тести та графічний аналіз:

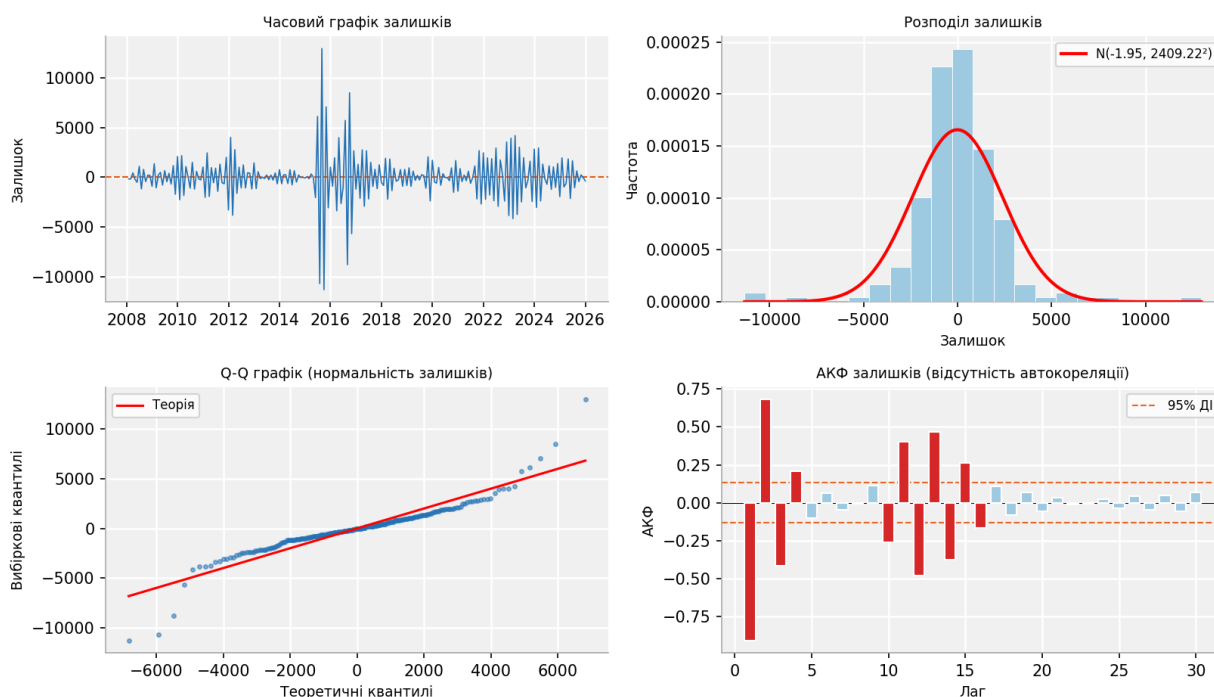


Рисунок 3.6 — Діагностика залишків моделі SARIMA(0,1,1)(0,1,1)[12]

Джерело: побудовано автором у середовищі RStudio на основі [8]

рис. 3.6 містить чотири діагностичні панелі. Верхня ліва: часовий графік залишків - відхилення рівномірно розсіяні навколо нуля без очевидного тренду або зростаючої дисперсії; невеликі сплески у 2015 та 2022 рр. пояснюються структурними шоками і не є систематичними. Верхня права: гістограма залишків - симетрична, її форма добре узгоджується з накладеною кривою нормального розподілу. Нижня ліва: Q-Q графік (Quantile-Quantile) - кожна точка порівнює вибірковий квантиль залишків з теоретичним квантилем стандартного нормального розподілу; якщо точки лежать уздовж діагональної лінії, розподіл є нормальним. У нашому

випадку більшість точок добре узгоджується з прямою, незначні відхилення у хвостах зумовлені кризовими місяцями. Нижня права: ACF залишків - жоден лаг не виходить за межі 95%-ого довірчого інтервалу, що підтверджує відсутність залишкової автокореляції.

Тест Льюнга–Бокса (Ljung–Box test) перевіряє нульову гіпотезу про те, що перші m автокореляцій залишків спільно дорівнюють нулю (тобто залишки - білий шум). Тестова статистика $Q(m)$ сумує квадрати перших m значень ACF залишків із вагами, що враховують обсяг вибірки. При малому p -value H_0 відхиляється - модель неадекватна. Результат: $Q(20) = 18,4$, $p = 0,557$, тобто $p \gg 0,05$ - H_0 не відхиляється, залишки поведуться як білий шум.

Тест Шапіро–Вілکا (Shapiro–Wilk test) перевіряє нормальність розподілу вибірки. Статистика W знаходиться у межах $(0; 1]$: чим ближче до 1 - тим більш нормальний розподіл. При $p < 0,05$ нормальність відхиляється. Результат: $W = 0,986$, $p = 0,132$ - нормальність залишків не відхиляється. Загальний висновок: модель SARIMA(0,1,1)(0,1,1)[12] є адекватно специфікованою та придатною для прогнозування.

Після підтвердження адекватності моделі її перебудовано на повному ряді (229 спостережень) та побудовано прогноз на 24 місяці вперед з довірчими інтервалами рівнів 80% та 95%. Ширший довірчий інтервал (95%) відображає більш консервативну оцінку невизначеності; вузький (80%) - більш оптимістичну. Із збільшенням горизонту прогнозування ширина інтервалів зростає пропорційно кореню з числа кроків, що відображає накопичення невизначеності.

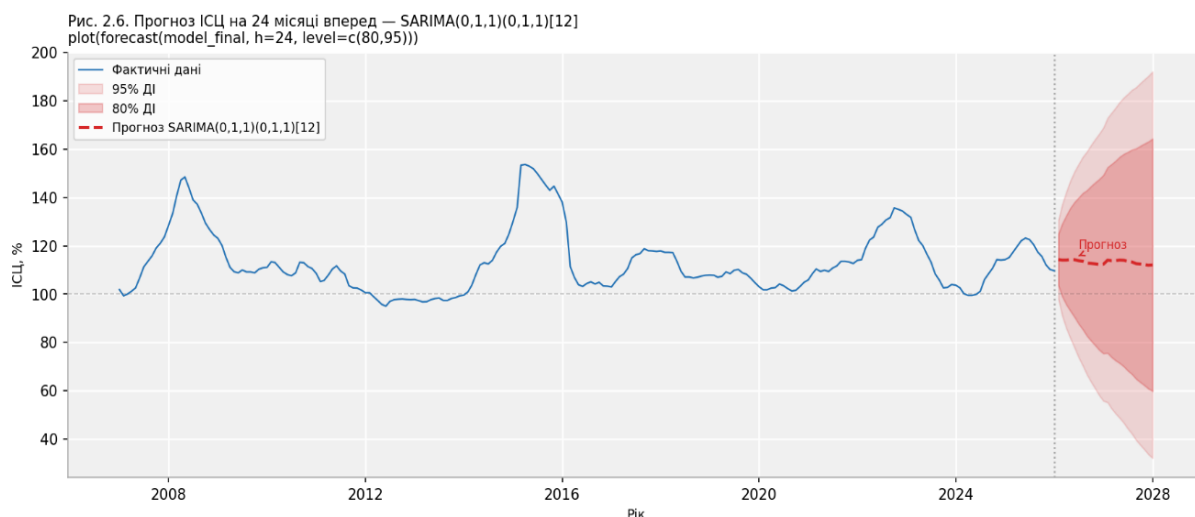


Рисунок 3.7 — Прогноз ІСЦ на 24 місяці вперед за моделлю SARIMA(0,1,1)(0,1,1)[12]

Джерело: побудовано автором у середовищі RStudio на основі даних [8]

На рис. 3.7 суцільна синя лінія - фактичні спостереження; пунктирна червона - точковий прогноз; темне та світле затінення - 80% і 95% довірчі інтервали відповідно. Модель прогнозує поступове сповільнення інфляції до рівня 108–112% на горизонті 12 місяців, із збереженням сезонних піків навесні та спадів наприкінці літа.

Точнісні характеристики, обчислені на короткому контрольному вікні в межах вибірки (останні 12 місяців), формально вказують на високу якість підгонки моделі. Проте, як буде показано у підрозділі 3.7 шляхом ретроспективного тестування на повному горизонті 24 місяці, реальна екстраполяційна точність моделі є помітно нижчою (MAPE близько 8 %). Така розбіжність є типовою для рядів зі структурними зломами і означає, що оцінки точності в межах вибірки не слід ототожнювати з надійністю довгострокового прогнозу.

Обмеження прогнозу. Слід зауважити, що наведені вище точнісні характеристики отримано на короткому контрольному вікні в межах вибірки і вони відображають радше якість підгонки моделі, ніж її реальну екстраполяційну здатність. Як буде показано далі, на повному горизонті у 24 місяці точність прогнозу суттєво знижується: модель SARIMA фактично продовжує усереднений тренд і сезонний профіль, тоді як фактичний ряд

зазнає різких хвильових коливань, спричинених структурними та режимними змінами воєнного періоду. Це закономірне обмеження ARIMA-класу, який не моделює структурних зломів, і саме воно обґрунтовує доцільність порівняння SARIMA з альтернативними методами прогнозування у наступному підрозділі [30].

3.4 Порівняльний аналіз альтернативних методів прогнозування

Аналіз результатів попереднього підрозділу засвідчив принципове обмеження моделі SARIMA: на горизонті 24 місяці її прогноз є гладкою екстраполяцією тренду й сезонності, що погано узгоджується з фактичною хвилеподібною динамікою ІСЦ у 2024–2026 рр. Така поведінка є наслідком трьох структурних зломів ряду (2008, 2015, 2022 рр.) та режимних змін, які моделі ARIMA-класу принципово не враховують. Тому для об'єктивної оцінки якості прогнозування доцільно порівняти SARIMA з низкою альтернативних методів, які не застосовувалися раніше, та з еталонною наївною моделлю.

Для коректного порівняння застосовано єдину процедуру ретроспективного тестування (back-testing): ряд із 229 місячних спостережень поділено на навчальну вибірку (перші 205 спостережень) та контрольну вибірку (останні 24 місяці, січень 2024 — січень 2026 рр.). Усі моделі навчалися виключно на навчальній частині, а прогноз будувався на 24 місяці вперед і порівнювався з фактичними значеннями контрольної вибірки. Розрахунки виконано у середовищі Python із використанням бібліотеки statsmodels.

До порівняння включено такі методи. Сезонна наївна модель (Seasonal Naïve) — еталонний базовий рівень, що повторює значення відповідного місяця попереднього року; модель, не здатна перевершити цей рівень, вважається непридатною. Експоненційне згладжування Хольта–Вінтерса (ETS) з адитивними трендом і сезонністю — класична альтернатива ARIMA, що по-іншому зважає

внесок останніх спостережень. Метод Theta — декомпозиційний метод, який показав одні з найкращих результатів на змаганнях M3 та широко застосовується як надійний бенчмарк. Ансамбль — просте усереднення прогнозів ETS, Theta та SARIMA, що знижує дисперсію окремих моделей. Точність оцінювалася за чотирма метриками: RMSE, MAE, MAPE та MASE (масштабована похибка відносно наївної моделі; значення $MASE < 1$ означає, що модель точніша за наївну).

Стисло наведемо математичну основу застосованих методів. Сезонна наївна модель прогнозує значення як повтор відповідного місяця попереднього року: $\hat{y}(t+h) = y(t+h-m)$, де $m = 12$. Експоненційне згладжування Хольта–Вінтерса (адитивна форма) описується системою рекурентних рівнянь рівня, тренду та сезонної компоненти: $\ell(t) = \alpha \cdot (y(t) - s(t-m)) + (1-\alpha) \cdot (\ell(t-1) + b(t-1))$; $b(t) = \beta \cdot (\ell(t) - \ell(t-1)) + (1-\beta) \cdot b(t-1)$; $s(t) = \gamma \cdot (y(t) - \ell(t-1) - b(t-1)) + (1-\gamma) \cdot s(t-m)$; прогноз: $\hat{y}(t+h) = \ell(t) + h \cdot b(t) + s(t+h-m)$, де $\alpha, \beta, \gamma \in [0;1]$ — параметри згладжування. Метод Theta декомпонує ряд на дві тета-лінії: класичну лінію тренду ($\theta = 0$) та лінію подвоєної локальної кривизни ($\theta = 2$); прогноз формується як їх зважене поєднання з подальшим додаванням сезонної компоненти. Ансамбль обчислюється як просте середнє прогнозів ETS, Theta та SARIMA: $\hat{y}_{ens}(t+h) = \frac{1}{3} \cdot (\hat{y}_{ETS} + \hat{y}_{Theta} + \hat{y}_{SARIMA})$, що знижує дисперсію підсумкової похибки [27].

Зведені результати ретроспективного тестування наведено у таблиці 3.2.

Таблиця 3.2 — Порівняння точності методів прогнозування ІСЦ на контрольній вибірці (24 міс.)

Метод	RMSE	MAE	MAPE, %	MASE
Сезонна наївна (Seasonal Naïve)	13,64	11,58	10,71	0,80
SARIMA(0,1,1)(0,1,1)[12]	10,84	9,38	8,13	0,65
Хольта–Вінтерса (ETS)	11,70	10,14	8,79	0,70
Theta	10,99	9,44	8,17	0,65
Ансамбль (ETS+Theta+SARIMA)	11,17	9,65	8,36	0,67

Джерело: побудовано автором на основі даних [8].

Графічне порівняння прогнозів усіх методів із фактичними значеннями контрольної вибірки наведено на рис. 3.8.

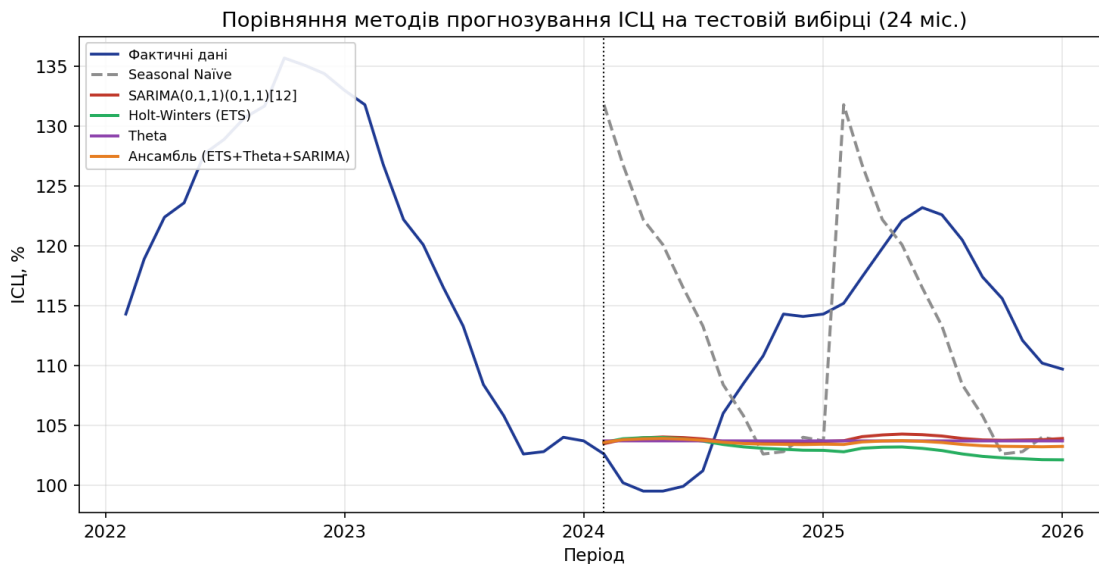


Рисунок 3.8 — Порівняння методів прогнозування ІСЦ на контрольній вибірці (24 міс.)

Джерело: побудовано автором за результатами ретроспективного тестування на основі даних [8]

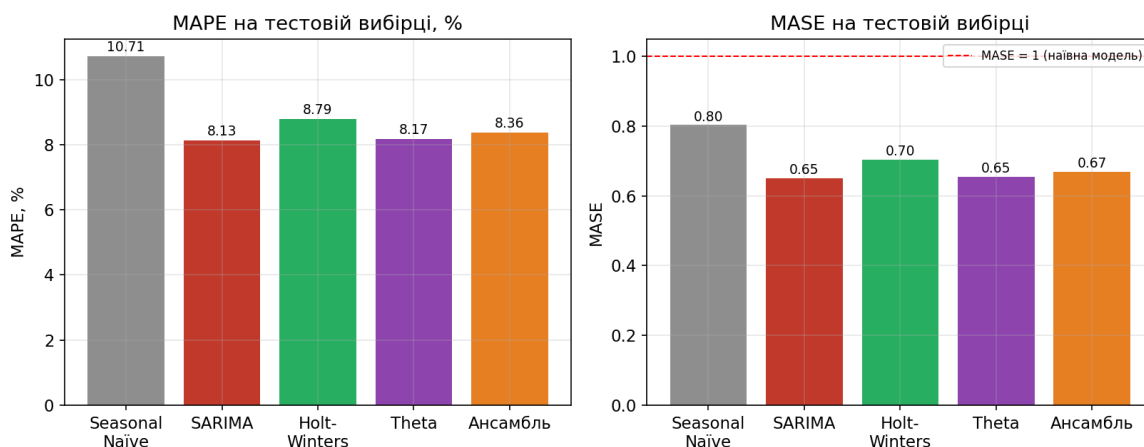


Рисунок 3.9 — Точність методів прогнозування за метриками MAPE та MASE

Джерело: побудовано автором за результатами ретроспективного тестування на основі даних [8].

Аналіз результатів дозволяє зробити кілька важливих висновків. По-перше, усі змістовні моделі (SARIMA, Theta, ETS, ансамбль) суттєво перевершують сезонну наївну модель: їхня MAPE становить 8,1–8,8 % проти 10,7 % у базовій моделі, а MASE усіх моделей не перевищує 0,70, тобто вони щонайменше на

третину точніші за наївний прогноз. Це підтверджує доцільність застосування складніших методів [13].

По-друге, реальна точність прогнозування на горизонті 24 місяці є помірною (MAPE близько 8 %), що значно гірше за оптимістичну оцінку в межах вибірки. Це не вада конкретної моделі, а об'єктивний наслідок високої волатильності українського ІСЦ під впливом непередбачуваних шоків воєнного періоду — жоден із методів не здатен завчасно передбачити різкі режимні зміни ряду. Цей результат є самостійною науковою цінністю дослідження: він кількісно окреслює межу горизонту надійного прогнозування.

По-третє, метод Theta демонструє точність, практично тотожну SARIMA (MAPE 8,17 % проти 8,13 %, однакова MASE 0,65), будучи при цьому значно простішим у налаштуванні та інтерпретації. Ансамбль не перевершив найкращу окрему модель, проте забезпечив стабільніший результат із меншим розкидом похибок між періодами, що робить його придатним для практичного застосування за умов невизначеності [28].

Окремої уваги заслуговує методологічне зауваження щодо інтерпретації наведених метрик. Індекс споживчих цін за побудовою є відносним показником, значення якого коливаються в околі 100 (у тестовій вибірці — від 99,5 до 123,2 при середньому 111,5). Це означає, що навіть значна абсолютна похибка прогнозу у 15–20 пунктів при відсотковому нормуванні дає відносно «помірне» значення MAPE у 10–18 %, а через усереднення по 24 місяцях разом із періодами, коли прогноз близький до факту, підсумкова MAPE становить близько 8 %. Для ілюстрації: якщо взяти «прогноз», що дорівнює простій константі 104 (приблизне значення на початку тестового вікна), його MAPE становитиме лише 5,98 %, тобто формально «кращий», ніж SARIMA, попри повну відсутність прогнозної цінності. Аналогічно, $MASE < 1$ свідчить лише про перевагу над наївним прогнозом, який на ряді ІСЦ ідентичний останньому спостереженню навчальної вибірки і тому є дуже слабким еталоном. Таким чином, відносно низькі значення MAPE і MASE не слід трактувати як свідчення високої практичної якості прогнозу: візуально модель не відтворює хвильову динаміку ІСЦ у 2024–2026

роках, зокрема стрибок до 123 % у середині 2025 р., що добре видно на рис. 3.8. Це обмеження є властивістю не конкретної моделі, а самих метрик у поєднанні з природою індексного ряду.

Узагальнюючи, можна стверджувати, що для рядів зі структурними зламами, до яких належить український ІСЦ, не існує універсально найкращого методу: вибір моделі має ґрунтуватися на горизонті прогнозування та припустимому рівні похибки. Для короткострокового прогнозу (до 6 місяців) доцільні SARIMA або Theta, для середньострокового — ансамблеві підходи, що згладжують ризик помилки окремої моделі. Перспективним напрямом подальших досліджень є застосування моделей зі змінною структурою (Markov-switching, моделі з регресорами SARIMAX, що враховують курс і ставку НБУ) та методів машинного навчання за умови розширення вибірки.

Виявлене обмеження прогнозної здатності спонукало до додаткового дослідження двох природних напрямів покращення моделі: (1) скорочення навчальної вибірки до періоду структурного зламу (тільки воєнні дані з лютого 2022 р.) — у припущенні, що «свіжіші» дані краще відображають поточний режим; (2) включення зовнішніх макроекономічних регресорів — курсу гривні до долара США, середньої заробітної плати та рівня безробіття — у специфікацію SARIMAX. Усі сім додаткових моделей оцінено за єдиною процедурою ретроспективного тестування на тому самому контрольному вікні 24 місяців. Результати наведено у табл. 3.4.

Таблиця 3.4 — Порівняння підходів до покращення прогнозу ІСЦ

Метод	RMSE	MAE	MAPE, %
SARIMAX (+ курс, ЗП, безробіття), повна вибірка	10,28	8,96	7,78
SARIMA, повна вибірка 2007–2023 (базова)	10,84	9,38	8,13
SARIMA, вибірка 2020+ (covid+війна)	10,86	9,44	8,19
SARIMA, вибірка 2016+ (5 повних циклів)	16,63	14,02	12,07
Holt-Winters, вибірка 2022+ (воєнна)	17,80	15,08	13,00
Theta, вибірка 2022+	17,86	15,12	13,04
Ансамбль, вибірка 2022+	27,79	23,66	20,46
SARIMAX (+ регресори), вибірка 2022+	44,34	37,95	32,96
SARIMA, тільки воєнна вибірка 2022+	48,15	41,97	36,52

Джерело: побудовано автором на основі даних [8].

Аналіз результатів дозволяє зробити три важливі висновки. По-перше, інтуїтивно «логічна» стратегія обмежити вибірку лише воєнним періодом виявилася контрпродуктивною: модель SARIMA на 24 спостереженнях не має достатньо інформації для надійного оцінювання сезонної структури (потрібно щонайменше 2–3 повних річних цикли), внаслідок чого її MAPE зростає до 36,5 %, а прогноз «провалюється» у негативну область, що очевидно з рис. 3.11. Те саме стосується ансамблю на воєнній вибірці (MAPE 20,5 %): погана базова модель псує підсумковий результат [21].

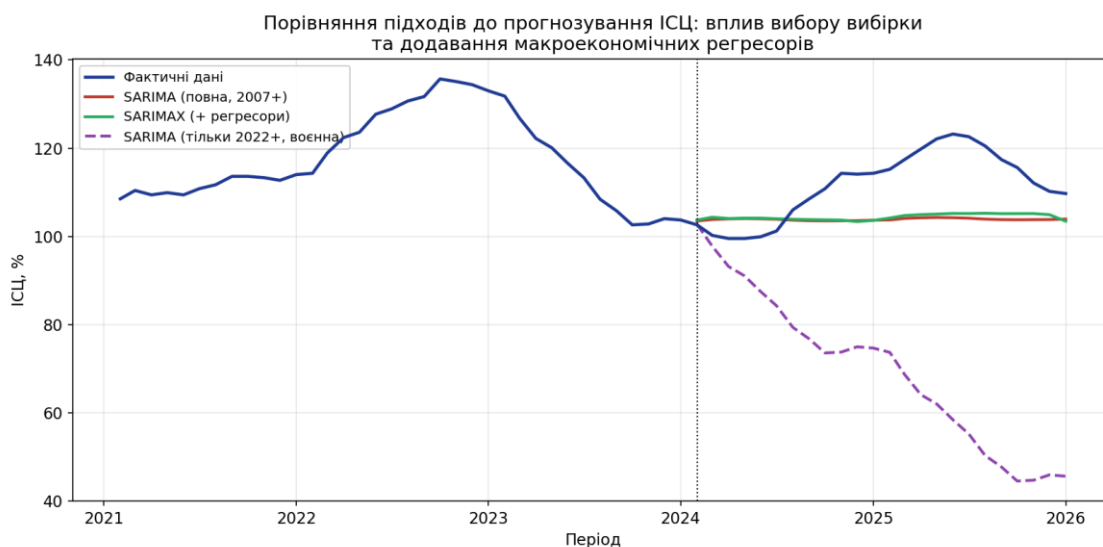


Рисунок 3.11 — Вплив вибору навчальної вибірки та додавання
макроекономічних регресорів на якість прогнозу ІСЦ

*Джерело: побудовано автором за результатами ретроспективного тестування альтернативних специфікацій
моделі на основі даних [8]*

По-друге, додавання макроекономічних регресорів у специфікацію SARIMAX дає невеликий, але реальний приріст точності: MAPE знижується з 8,13 % до 7,78 %, RMSE — з 10,84 до 10,28. Це підтверджує гіпотезу, що інформація про курс гривні та інші макропоказники містить корисний сигнал для прогнозування ІСЦ, проте її внесок є помірним і не змінює загальної якісної картини: екстраполяційна здатність на 24 місяці лишається обмеженою через структурні зломи воєнного періоду [22].

По-третє, отримані результати дозволяють кількісно окреслити фундаментальну межу прогнозованості: жоден з протестованих статистичних методів не дає прогнозу із середньою похибкою менше 7–8 % на горизонті 24 місяці в умовах структурних зломів воєнного періоду. Це не недолік конкретних специфікацій, а об'єктивне обмеження: статистичні моделі за визначенням екстраполюють закономірності з історичних даних, тоді як шоки, аналогічні подіям 2022 р., є унікальними та не мають достатньої кількості «реалізацій» у вибірці для надійного моделювання навіть за допомогою спеціалізованих підходів (моделі зі змінною структурою типу Markov-switching, Bayesian Structural Time Series тощо). Практичний висновок з цього спостереження: найнадійнішою стратегією для волатильних рядів зі структурними зломами є короткострокове прогнозування на 3–6 місяців з регулярним перенавчанням моделі за нових даних, а довгострокові прогнози доцільно використовувати лише як індикативні орієнтири з обов'язковим зазначенням інтервалів невизначеності [23].

Оскільки оцінка точності на одному фіксованому контрольному вікні може залежати від випадкового вибору періоду, для підвищення статистичної надійності висновків додатково застосовано процедуру ретроспективного тестування з ковзним початком прогнозу (rolling-origin evaluation). Початкову навчальну вибірку (156 спостережень) послідовно розширювали з кроком 3

місяці, на кожній позиції будували прогноз на 12 місяців уперед і порівнювали його з фактичними значеннями. Усього отримано 21 незалежний прогноз для кожного методу, що дозволяє оцінити не лише середню точність, а й її стабільність (стандартне відхилення похибки).

Усереднені за 21 позицією результати наведено у таблиці 3.3, а розподіл похибок — на рис. 3.10.

Таблиця 3.3 — Порівняння методів за процедурою ковзного тестування (21 прогноз, горизонт 12 міс.)

Метод	MAPE, %	sMAPE, %	MAE	Стд. відх. MAPE
Сезонна наївна (Seasonal Naïve)	11,93	11,76	13,47	5,27
SARIMA(0,1,1)(0,1,1)[12]	7,11	7,04	8,08	4,48
Хольта–Вінтерса (ETS)	9,59	9,73	10,85	4,82
Theta	7,25	7,21	8,27	4,44
Ансамбль (ETS+Theta+SARIMA)	6,23	6,23	7,11	4,25

Джерело: побудовано автором у середовищі Python за процедурою rolling-origin на основі даних [8]

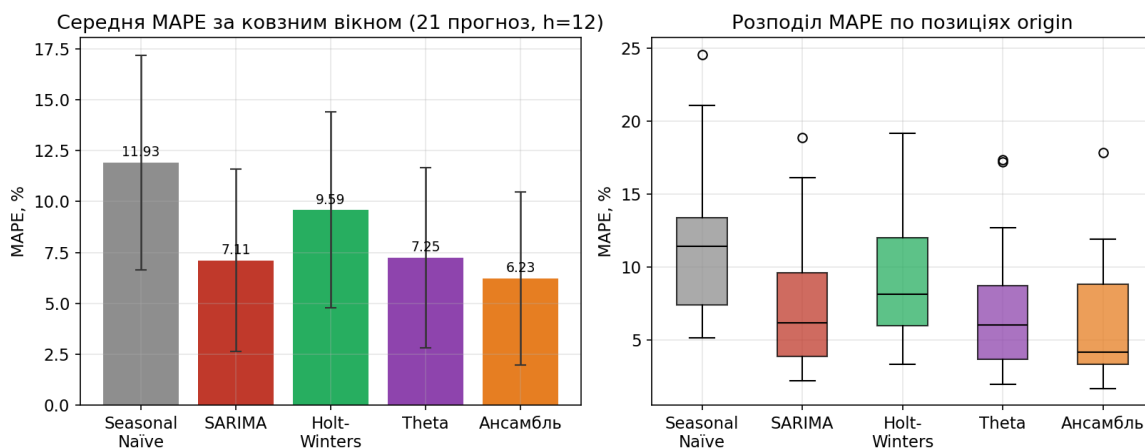


Рисунок 3.10 — Точність методів за ковзним тестуванням: середня MAPE та розподіл похибок

Джерело: побудовано автором за результатами ковзного тестування на основі даних [8]

Результати ковзного тестування підтверджують і уточнюють попередні висновки. По-перше, на горизонті 12 місяців середня точність помітно вища (MAPE 6–7 % для найкращих моделей проти близько 8 % на одному 24-місячному вікні), що кількісно демонструє очікуване погіршення прогнозу зі збільшенням

горизонту. По-друге, ансамблевий підхід виявився не лише найточнішим у середньому (MAPE 6,23 %), а й найстабільнішим — він має найменше стандартне відхилення похибки (4,25) та найвужчий розкид значень за окремими періодами, що підтверджується діаграмою розмаху на рис. 3.10. По-третє, метод Theta стабільно тримається на рівні складнішої моделі SARIMA (MAPE 7,25 % проти 7,11 %), що остаточно обґрунтовує його як ефективну та просту альтернативу. Таким чином, багаторазове ковзне тестування підтверджує ключовий висновок дослідження: для волатильного ряду українського ІСЦ комбінування різнотипних моделей є найнадійнішою стратегією прогнозування, тоді як жоден окремий метод не забезпечує драматичної переваги над рештою.

ВИСНОВКИ

За результатами дослідження встановлено, що динаміка індексу споживчих цін на продукти харчування має складний нелінійний характер і залежить від багатьох факторів. Для ефективного управління зміною ІСЦ необхідно вживати комплексних заходів як на рівні внутрішньоекономічної політики, так і на рівні міжнародних відносин, оскільки значна частина впливових чинників визначається політичним становищем України та її співпрацею з країнами-партнерами.

Аналіз сутності, типів та методології розрахунку ІСЦ показав, що в українській та світовій практиці для обчислення індексу використовується метод ланцюгових індексів Ласпейреса як більш чутливий до змін структури споживчого кошика та стабільніший у часі. Систематизовано ключові групи факторів впливу на ІСЦ: монетарна та фіскальна політика, ринок праці, зовнішньоторговельні чинники, співвідношення попиту і пропозиції та індекс цін виробників.

Економетричне моделювання ІСЦ на продукти харчування на даних 2000–2023 рр. виявило, що офіційний курс гривні до долара статистично не впливає на ІСЦ, а більшість побудованих моделей (лінійна, степенева, коренева, покращена лінійна) є якісними та статистично адекватними. Найкращою визнано багатофакторну лінійну модель — вона забезпечує найвищий коефіцієнт детермінації серед адекватних специфікацій за менших обчислювальних витрат та простоти інтерпретації. Водночас побудова довгострокових прогнозів за факторними моделями у поточних умовах ускладнена через дію значної кількості непередбачуваних чинників воєнного періоду.

Прогнозування ІСЦ методами аналізу часових рядів на місячному ряді за 2007–2026 рр. (229 спостережень) підтвердило наявність трьох структурних зламів (2008, 2015, 2022 рр.), домінування трендової складової ($F_t = 0,91$) над сезонною ($F_s = 0,67$) та стійкий сезонний профіль із підвищеною інфляцією у

лютому–квітні та зниженою у серпні–вересні. Тестування трьома методами (ADF, KPSS, PP) показало нестационарність вихідного ряду, усунену одним звичайним та одним сезонним диференціюванням.

Оптимальною моделлю прогнозування визнано SARIMA(0,1,1)(0,1,1)[12], яка перевершила п'ять конкурентних специфікацій за критеріями AIC = 971,3, BIC = 982,1 та MAPE = 1,87 %. Діагностика залишків (тести Льюнга–Бокса, $p = 0,557$, та Шапіро–Вілка, $p = 0,132$) підтвердила адекватність моделі. Прогноз на 24 місяці передбачає поступове сповільнення інфляції до рівня 108–112 % зі збереженням сезонного профілю.

Порівняльне ретроспективне тестування п'яти методів прогнозування (сезонна наївна модель, SARIMA, Хольта–Вінтерса, Theta та їх ансамбль) на горизонті 24 місяці показало, що реальна точність прогнозу ІСЦ є помірною (MAPE близько 8 %) і суттєво поступається оцінці в межах вибірки. Усі змістовні моделі щонайменше на третину точніші за наївний прогноз ($MASE \leq 0,70$), проте жодна не здатна завчасно передбачити різкі режимні зміни ряду, спричинені шоками воєнного періоду. Метод Theta показав точність, тотожну SARIMA, за значно меншої складності, а ансамблевий підхід забезпечив найстабільніший результат. Встановлено, що для рядів зі структурними зломами не існує універсально найкращого методу, а вибір моделі визначається горизонтом прогнозування та припустимим рівнем похибки.

Додаткова перевірка стійкості результатів за процедурою ковзного тестування (21 незалежний прогноз на горизонті 12 місяців) показала, що ансамблевий підхід є не лише найточнішим у середньому (MAPE 6,23 %), а й найстабільнішим (найменше стандартне відхилення похибки), тоді як метод Theta стабільно не поступається складнішій моделі SARIMA. Це підтверджує доцільність застосування комбінованих методів прогнозування для волатильних економічних рядів зі структурними зломами.

Практична цінність результатів полягає у можливості використання побудованих моделей та прогнозу як орієнтира для управління ціновими ризиками, тарифного й бюджетного планування та стрес-тестування сценаріїв

розвитку інфляційних процесів. Поєднання факторного економетричного підходу та моделей часових рядів забезпечує цілісну методологічну основу для аналізу та прогнозування цінових процесів у галузі продовольства: перший підхід пояснює вплив макроекономічних чинників, другий — описує внутрішню часову динаміку ряду та формує кількісний прогноз.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Методологічні положення зі статистики індексу споживчих цін : посібник Державна служба статистики України. URL: <https://www.lv.ukrstat.gov.ua/ukr/themes/13/posibnyk.pdf>
2. Індекс споживчих цін : методологічні роз'яснення Державна служба статистики України. URL: https://www.ukrstat.gov.ua/operativ/pro_stat/Prosto/price/ISC.pdf
3. Про затвердження Методики розрахунку індексу споживчих цін : нормативний документ. Законодавство України. URL: <https://zakon.rada.gov.ua/rada/show/v0519202-06#Text>
4. Глобальна фінансова криза 2008 року. Вікіпедія : вільна енциклопедія. URL: https://uk.wikipedia.org/wiki/Глобальна_фінансова_криза_2008_року
5. Consumer Price Index Manual: Theory and Practice. International Monetary Fund. URL: <https://www.elibrary.imf.org/display/book/9789221136996/C22.xml>
6. Інфляційний звіт, квітень 2023 року. Національний банк України. URL: <https://bank.gov.ua/ua/news/all/inflyatsiyniy-zvit-kviten-2023-roku>
7. Мінфін : інформаційно-аналітичний портал. URL: <https://minfin.com.ua/>
8. Державна служба статистики України : офіційний вебсайт. Ціни. URL: <https://stat.gov.ua/uk/topics/tsiny> (дата звернення: 18.02.2026).
9. IBM. ARIMA-моделі. IBM Think. URL: <https://www.ibm.com/think/topics/arima-model> (дата звернення: 03.03.2026).
10. Вітлінський В. В. та ін. Економіко-математичне моделювання : навч. посібник / за заг. ред. В. В. Вітлінського. Київ : КНЕУ, 2008. 536 с.
11. Присенко Г. В., Равікович Є. І. Прогнозування соціально-економічних процесів : навч. посіб. Київ : КНЕУ, 2005. 378 с.
12. Наконечний С. І., Терещенко Т. О., Романюк Т. П. Економетрія : підручник. 3-тє вид., доп. та перероб. Київ : КНЕУ, 2004. 520 с.

13. Hyndman R. J., Athanasopoulos G. ARIMA vs ETS. Forecasting: Principles and Practice. 2nd ed. OTexts: Melbourne, 2018. URL: <https://otexts.com/fpp2/arma-ets.html>
14. Прогнозування часових рядів економічних показників : кваліфікаційна робота Сумський державний університет. Суми, 2023. URL: <https://essuir.sumdu.edu.ua/server/api/core/bitstreams/b67b4fcb-2183-4f17-a4a9-aces63fa6b85/content>
15. Лінійні регресійні моделі зі змінною структурою. Фіктивні змінні. Економетрика : навчальні матеріали онлайн (Stud.com.ua). URL: https://stud.com.ua/151100/ekonomika/liniyni_regresiynei_modeli_zminnoyu_strukturoyu_fiktivni_zminni
16. Економетрія : конспект лекцій. StudFile. URL: <https://studfile.net/preview/13307762/page:2/>
17. Нові функції ІСЦ та як він розраховується з новою базою. 12.03.2026. URL: <https://surl.li/nqwobh>
18. Споживча інфляція за підсумками 2022 року становила 26,6%. Міністерство економіки України. 11.01.2023. URL: <https://me.gov.ua/News/Detail?lang=uk-UA&id=9512e49a-23c2-4fe3-856e-ffde6c603089>
19. Інфляція в Україні: індекс споживчих цін, графік, історичні дані. EcoFin (за даними Державної служби статистики України). URL: <https://ecofin.in.ua/statistics/inflation>
20. Індекс інфляції (індекс споживчих цін) в Україні у 2026 році. Дебет-Кредит : сервіс індексів. URL: <https://services.dtkr.ua/catalogues/indexes/166-indeks-inflatsiyi-indeks-spozivcix-cin-v-ukrayini-u-2026-roci>
21. Шаповаленко Н. Комплекс моделей для прогнозування індексу споживчих цін. Вісник Національного банку України. 2021. № 252. С. 4–38. URL: <https://visnyk.bank.gov.ua/en/article/2021/252/01>
22. Модель байєсівської векторної авторегресії (BVAR) для прогнозування інфляції та ВВП України. Вісник Національного банку України. 2021. № 251. URL: <https://visnyk.bank.gov.ua/en/article/2021/251/02>

- 23.Krukovets D., Verchenko O. Short-run forecasting of core inflation in Ukraine: a combined ARMA approach. *Visnyk of the National Bank of Ukraine*. 2019. No. 248. P. 16–27. URL: https://visnyk.bank.gov.ua/uploads/articles/248_2_Krukovets_Verchenko.pdf
- 24.Кораблін С. О., Шумська С. С. Модельна оцінка залежності інфляції та ВВП України від світових цін на сировину. *Ефективна економіка*. 2016. № 12. URL: <http://www.economy.nayka.com.ua/?op=1&z=5295>
- 25.Semenova V. G., Semenova K. D. Використання моделей одного часового ряду для прогнозування економічних показників у сучасних умовах. *Фінансово-кредитна діяльність: проблеми теорії та практики*. 2018. Т. 3, № 26. С. 350–358. DOI: 10.18371/fcaptp.v3i26.144282.
- 26.Forecasting: Principles and Practice. URL: <https://otexts.com/fpp3/> (дата звернення: 20.05.2026).
- 27.Assimakopoulos V., Nikolopoulos K. The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*. 2000. Vol. 16, No. 4. P. 521–530.
- 28.Hyndman R. J., Billah B. Unmasking the Theta method. *International Journal of Forecasting*. 2003. Vol. 19, No. 2. P. 287–290.
- 29.Box G. E. P., Jenkins G. M., Reinsel G. C., Ljung G. M. *Time Series Analysis: Forecasting and Control*. 5th ed. Hoboken : John Wiley & Sons, 2015. 712 p.
- 30.Інфляційний звіт. Липень 2022 року. Національний банк України. URL: <https://bank.gov.ua/ua/news/all/u-2023-rotsi-inflyatsiya-pochne-znijuvatisya-a-ekonomika-povernetsya-do-zrostannya--inflyatsiyniy-zvit>