

МАТЕМАТИЧНІ МЕТОДИ, МОДЕЛІ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ЕКОНОМІЦІ

УДК 330.43

Моделювання цін на квартири в множинній регресії

Ковпак Е.О.

кандидат економічних наук,
доцент кафедри економічної кібернетики та прикладної економіки
Харківського національного університету імені В.Н. Каразіна

Малець В.В.

студент факультету економічної інформатики і менеджменту
Національного технічного університету
«Харківський політехнічний інститут»

У статті проаналізовано ціноутворюючі фактори на ринку нерухомості. Показано результати кластеризації історичних районів м. Харкова за рівнем цін на житлову нерухомість. Запропоновано вид множинної регресії із використанням фіктивних змінних для опису характеристик поточного стану та місцезнаходження об'єкту нерухомості. Наведено результати оцінки параметрів моделі множинної регресії для цін пропозиції на однокімнатні квартири у м. Харкові.

Ключові слова: прогнозування, множинна регресія, фіктивні змінні, ціна квартири, первинний ринок житла, гетероскедастичність.

Ковпак Э.А., Малец В.В. МОДЕЛИРОВАНИЕ ЦЕН НА КВАРТИРЫ В МНОЖЕСТВЕННОЙ РЕГРЕССИИ

В статье проанализированы ценообразующие факторы на рынке недвижимости. Показаны результаты кластеризации исторических районов г. Харькова по уровню цен на жилую недвижимость. Предложен вид множественной регрессии с использованием фиктивных переменных для описания характеристик текущего состояния и местонахождения объекта недвижимости. Приведены результаты оценки параметров модели множественной регрессии для цен предложения на однокомнатные квартиры в г. Харькове.

Ключевые слова: прогнозирование, множественная регрессия, фиктивные переменные, цена квартиры, первичный рынок жилья, гетероскедастичность.

Kovpak E.O., Malets V.V. MODELLING RESIDENTIAL REAL ESTATE PRICES IN MULTIPLE REGRESSION

The pricing factors in the real estate market have been analyzed in the article. The results of clustering the historical districts of Kharkov by level of residential real estate prices have been shown. The specification of the multiple regression including dummy variables to describe the characteristics of the current status and location of the property has been proposed. The results of the estimation for multiple regression for the bid prices for one-bedroom apartments in Kharkov have been shown.

Keywords: forecasting, multiple regression, dummy variables, price of the apartment, primary housing market, heteroscedasticity.

Постановка проблеми в загальному вигляді. Завдання моделювання та прогнозу цін на нерухомість постає в щоденній практичній діяльності професійних учасників ринку нерухомості: інвестиційних консультантів та аналітиків, оцінювачів, ріелторів тощо. В умовах високого ступеню невизначеності широко затребувані експертні методи прогнозування на ринку нерухомості. Серед інших методів, що можуть використовуватись для цілей прогнозування цін на об'єкти нерухомості, можна виділити методи економіко-математичного моделювання, апарат трендових моделей

росту, різноманітних моделей часових рядів, нейро-нечіткі методи, комбіновані методи. Адекватні регресійні моделі цін на ринку нерухомості при їх відносній легкості побудови та оцінки окрім здатності до прогнозування виступають ще й зручним інструментарієм аналізу ступеню впливу різних ціноутворюючих факторів на об'єкти нерухомості.

Аналіз останніх досліджень і публікацій. Питанням аналізу динаміки та ціноутворення на ринку нерухомості присвячено роботи Г.М. Стерніка, А. Асаула, І. Балабанової, Я. Маркуса, Ю. Манцевича, І. Геллера,

Н. Ордуея, Дж. Фрідмена, Г. Харрісона та ін. Питання застосування нейронних мереж для прогнозування цін на нерухомість показано в роботах В. Григорківа, О. Ярошенко, Н. Філіпчука, В. Тазетдінова, В. Вороніна. Застосування методів нечіткого моделювання для вибору інструментарію прогнозування на ринку нерухомості розглянуто в роботі Н. Максишко [1]. Однозначної переваги одних методів моделювання цін на нерухомість над іншими немає, вибір інструментарію моделювання цілком залежить від мети дослідження та структури даних.

Формулювання цілей статті (постановка завдання). Мета статті – проаналізувати застосовність моделей множинної регресії із включенням фіктивних змінних для прогнозування цін на квартири в багатоквартирних будинках, що відносяться до новобудов.

Викладення основного матеріалу дослідження. Всі об'єкти нерухомості мають одночасно такі характеристики, що відрізняють їх від інших товарів [2]:

- корисність,
- фіксоване місцерозташування,
- фундаментальність,
- унікальність,
- двокомпонентність (об'єкт нерухомості складається із землі та будівель),
- довговічність.

Всі перелічені характеристики об'єктів нерухомості можна розглядати як ціноутворюючі фактори. Так, корисність об'єкта нерухомості визначається такими його характеристиками, як розмір приміщення, планування, поточний стан, благоустрій навколишньої території, матеріал основних конструкцій, – від всіх перелічених характеристик залежить ціна нерухомості. Місцезнаходження в основному обумовлює унікальність кожного об'єкта нерухомості, багато в чому визначає його позиціонування на ринку. Навіть дві однакові квартири в одному і тому ж багатоквартирному будинку мають різне розташування в термінах поверху, орієнтації вікон, розташування в будинку і так далі. Властивість двокомпонентності нерухомості суттєво впливає на ціну об'єкта нерухомості у випадку одноквартирних житлових домів. Властивість унікальності нерухомості як товару проявляється в тому, що кожен об'єкт нерухомості має певні, властиві лише йому ознаки, що відрізняють його від інших об'єктів нерухомості. Ознаками об'єкта нерухомості, що роблять його унікальним можуть бути будь-які обставини або характеристики: наявність в зоні пішохід-

ної доступності від будинку парку або упорядкованої набережної річки збільшать ціни на квартири, а, навпаки, близькість промислового об'єкта з великою кількістю забруднюючих атмосферу викидів або шум та вібрація у сусідніх з квартирою приміщеннях знизять ціну на неї.

Саме властивість унікальності об'єктів нерухомості ускладнює як процедуру оцінки житла, так і розробку моделей прогнозування цін на житло. Для різних об'єктів нерухомості властивість унікальності забезпечується дією різноманітних характеристик та обставин. Але брати до уваги при побудові моделей прогнозування велику кількість факторів недоцільно, бо, у випадку, моделей регресії це може знизити якість прогнозування та розширити довірчі інтервали для прогнозів, що зробить їх неінформативними для прийняття рішень.

В даному дослідженні моделі прогнозування цін на квартири розроблялися лише для новобудов. Якщо звернутись до нормативних актів, то в Порядку державної реєстрації речових прав на нерухоме майно та їх обтяжень, що відповідає Закону України [3], дається визначення лише для новозбудованого об'єкта – «завершений будівництвом об'єкт нерухомого майна, державна реєстрація права власності на який проводиться вперше». Тобто слово «новобудова» не є нормативно-правовим чи законодавчо визначеним терміном, і повинен застосовуватися по відношенню до об'єктів нерухомості на первинному ринку житла. Первинний ринок житла представляє собою сегмент ринку нерухомості, на якому обертаються тільки створені об'єкти житлової нерухомості, або такі, що вперше пропонуються до продажу (наприклад, у результаті приватизації). Вторинний ринок нерухомості – це ринок об'єктів, права на які вже були зареєстровані у єдиному державному реєстрі прав власності. Тобто квартири в новобудовах реалізуються забудовником багатоквартирного будинку (юридичною особою), або фізичними особами, що продають квартири за угодами пайової участі, тобто в порядку переуступки, до того моменту, як будинок побудований повністю [4]. Але часто учасники ринку нерухомості відносять до квартир у новобудовах і ті, права на які вже зареєстровані в єдиному державному реєстрі прав власності (тобто формально ці квартири вже належать до вторинного ринку житла), але багатоквартирні будинки, в яких знаходяться ці квартири, ще або не здані в експлуатацію, або не заселені повністю. В різних джерелах містяться реко-

мендації щодо віднесення багатоквартирних будинків до новобудов, якщо після здачі будинку в експлуатацію не пройшло 2 роки, або, наприклад, 10 років [5].

Статистичною вибіркою для побудови моделі регресії слугували дані про ціни пропозиції з сайту [6] в м. Харкові на квартири, первинного або вторинного ринку нерухомості в об'єктах, після здачі яких в експлуатацію не пройшло більше 4 років.

Серед характеристик корисності квартир в моделі регресії будемо враховувати дві властивості:

- загальна площа квартири,
- поточний стан.

Стосовно характеристики загальної площі квартири – вона кількісна і проблем обліку її в моделі регресії нема.

Стосовно характеристики поточного стану квартири – вона якісна і для обліку її в моделі регресії повинна бути переведена в кількісний вид. Використання якісних (індикаторних) змінних в моделях регресії можливо завдяки інструментарію фіктивних змінних. Фіктивні змінні (англ. «dummy» variables) найчастіше вводяться до моделі регресії як булевська змінна, що приймає значення 0 «неправда» або 1 «істина».

Якщо квартира у новобудові продається фізичною особою, її корисні властивості як товару можуть змінитися порівняно із такою ж квартирою від забудовника (перший влас-

ник може здійснити в квартирі певні оздоблювальні роботи, змінити планування, конструкції та ін.). Самі забудовники теж пропонують квартири в новобудовах у різному стані: без оздоблюваних робіт (квартири, в яких не здійснено штукатурних робіт, не проведено електричних кабелів та ін.), з реалізованим циклом штукатурних робіт або вже в житловому стані.

При моделюванні цін на квартири можна обрати будь-яке число значень якісної змінної «поточний стан квартири», число цих значень залежить від необхідного ступеню деталізації опису. Наприклад, можна враховувати в моделі два види стану квартири в новобудові – без оздоблювальних робіт та з виконаними оздоблювальними роботами, а можна класифікувати види виконаних оздоблювальних робіт, для чого вводяться більш ніж два стани якісної змінної «поточний стан квартири». В даному дослідженні було обрано три значення змінної «поточний стан квартири»: «без оздоблювальних робіт», «розпочаті оздоблювальні роботи», «завершено оздоблювальні роботи». Для обліку такої змінної в моделі треба ввести 2 фіктивні змінні:

$$D_1 = \begin{cases} 1, & \text{якщо квартира з завершеними оздоблювальними роботами} \\ 0, & \text{інакше} \end{cases}, \quad (1)$$

$$D_2 = \begin{cases} 1, & \text{якщо квартира з розпочатими оздоблювальними роботами} \\ 0, & \text{інакше} \end{cases}. \quad (2)$$

Якщо поточний стан квартири з вибірки було ідентифіковано як «без оздоблювальних робіт», то фіктивним змінним D_1 та D_2 присвоюються значення 0.

Таблиця 1

Результати кластеризації історичних районів та місцевостей м. Харкова за значенням цін на квартири

Номер кластеру	Історичні райони	Середня ціна на квартири, дол. США		
		Однокімнатні	Двокімнатні	Трикімнатні
1	Нагорний, Сокольники, Держпром, Соснова гірка, Шатіловка	85600	131580	191000
2	Південний вокзал, Центральний ринок, Кінний ринок, Левада, станція метро Спортивна, Цирк	28130	39379	59920
	Олексіївка	29140	45300	59980
	Лісопарк, Жуковського, П'ятихатки, Шишківка	32875	45625	59750
3	Салтівка	23480	29320	36400
	Холодна Гора, Баварія, Залютіно, Лиса гора, Філіповка	25667	35433.3	42000
4	Нові дома	19071	24730	33200
	Новожанівка, Новоселівка, Диканьовка, Москальовка	18500	23750	29500
	Роганьський напрям	17943	23643	30785
	Аеропорт, Одеська, Жихарь	14400	22075	29825

Джерело: складено авторами на основі [6]

Ще одну якісну змінну – «місцезнаходження», можна теж врахувати в моделі ціни на квартири завдяки інструментарію фіктивних змінних. Для визначення числа значень, що приймає фактор «місцезнаходження» необхідний попередній аналіз подібностей та відмінностей цін на квартири у різних районах міста. Зазвичай виділення однорідних груп об'єктів здійснюється в задачах кластеризації. Інструментарій кластерного аналізу підтримує комплексний підхід до сегментації ринку нерухомості.

Кластерний аналіз (англ. Data clustering) – задача розбиття заданої вибірки об'єктів (ситуацій) на підмножини, які називаються кластерами так, щоб кожен кластер складався із схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися.

У результаті вирішення задачі кластеризації історичних районів м. Харкова було виділено 4 групи районів (див. табл. 1).

Для обліку змінної «Місцезнаходження» в модель введено 3 фіктивні змінні:

$$R_1 = \begin{cases} 1, \text{ якщо квартира з 1 кластеру} \\ 0, \text{ інакше} \end{cases}, \quad (3)$$

$$R_2 = \begin{cases} 1, \text{ якщо квартира з 2 кластеру} \\ 0, \text{ інакше} \end{cases}, \quad (4)$$

$$R_3 = \begin{cases} 1, \text{ якщо квартира з 3 кластеру} \\ 0, \text{ інакше} \end{cases}. \quad (5)$$

Таким чином, модельне значення ціни \hat{Y}_i на i -ту однокімнатну квартиру в м. Харкові може оцінюватись у вигляді такої множинної регресії:

$$\hat{Y}_i = \alpha_0 + \alpha_1 X_{1i} + \alpha_2 D_{1i} + \alpha_3 D_{2i} + \alpha_4 R_{1i} + \alpha_5 R_{2i} + \alpha_6 R_{3i}, \quad (6)$$

де α_i – коефіцієнти множинної регресії,

X_i – площа i -ї квартири;

D_{1i} , D_{2i} – фіктивні змінні, що описують стан i -ї квартири згідно формул (1)-(2);

R_{1i} , R_{2i} , R_{3i} – фіктивні змінні, що є індикаторами місцезнаходження i -ї квартири згідно формул (3)-(5).

Параметри моделі (6) було оцінено методом найменших квадратів на виборці зі 114 однокімнатних квартир у новобудовах м. Харкова, та отримано такий вид моделі:

$$\hat{Y}_i = -8679 + 730_1 X_{1i} + 11130 D_{1i} + 6548 D_{2i} + 7786 R_{1i} + 4826 R_{2i} + 1676 R_{3i}. \quad (7)$$

Коефіцієнт детермінації для моделі (7) склав 0,78, емпіричне значення статистики Фішера 61,8. Перевірка якості моделі (7) виявила незначущість за критерієм Стюдента оцінки коефіцієнта регресії біля R_{3i} , загалом адекватність моделі за критерієм Фішера (з рівнем довіри не менш 95%). Але проведений для моделі (7) тест Уайта виявив похибку гетероскедастичності. Помилка гетероскедастичності є частою для моделей регресії, що будуються для перехресних даних, як і в даному випадку з квартирами.

Для усунення помилки гетероскедастичності для отримання ефективних оцінок параметрів регресії (6) можна застосувати метод зважених найменших квадратів. Модельне значення ціни \tilde{Y}_i на i -ту однокімнатну квартиру в м. Харкові за методом зважених найменших квадратів для оцінки параметрів моделі виду (6) дав такі результати:

$$\tilde{Y}_i = -12672 + 788_1 X_{1i} + 13293 D_{1i} + 8901 D_{2i} + 7025 R_{1i} + 5767 R_{2i} + 3178 R_{3i}. \quad (8)$$

Коефіцієнт детермінації за моделлю (8) складає 0,76. За критерієм Стюдента значущими є всі оцінки параметрів з рівнем довіри не менш 95%. Однією з переваг отриманої моделі є можливість змістовного аналізу значень отриманих коефіцієнтів регресії (всі оцінки, окрім вільного члену,



Рис. 1. Графічне зіставлення модельного та фактичного рядів даних

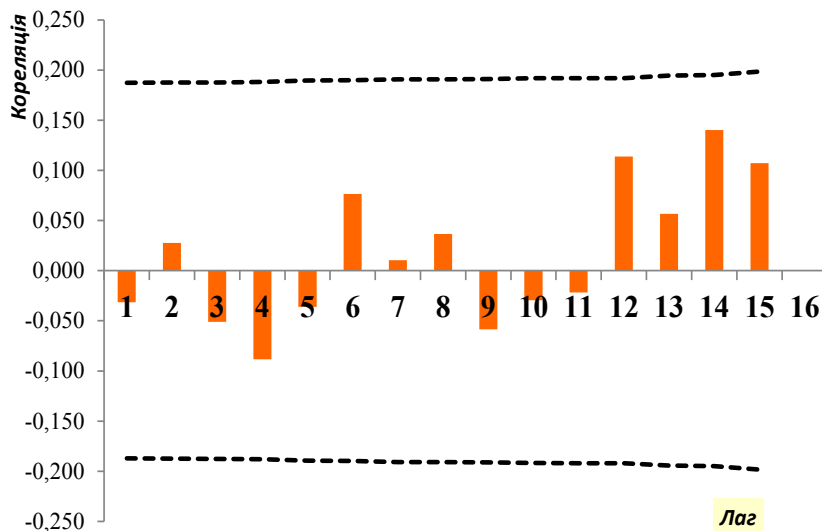


Рис. 2. Автокореляційна функція для залишків моделі регресії

позитивні, що не суперечить економічному змісту впливу ціноутворюючих факторів на ціну квартири та відповідає введеним фіктивним змінним).

Помилка гетероскедастичності в моделі (8) вже відсутня (на підставі результатів тесту Уайта). Середня відносна похибка апроксимації за моделлю становить 11,72%, візуально порівняння модельних та фактичних цін на квартири зображено на рис. 1.

Автокореляція залишків в моделі (8) відсутня за тестом Дарбіна-Уотсона (емпіричне

значення статистики Дарбіна-Уотсона 2,02), а також судячи з автокореляційної функції (див. рис. 2).

Результати оцінки якості побудованої моделі свідчать про можливість її використання для цілей прогнозування цін на однокімнатні квартири у новобудовах м. Харкова.

Висновки з цього дослідження. Для моделювання цін на ринку нерухомості більшість ціноутворюючих факторів має якісний (індикативний) характер. Індикаторні змінні в моделях регресії можна застосувати завдяки інструментарію фіктивних змінних.

У дослідженні на прикладі вибірки однокімнатних квартир в м. Харкові показано можливість застосування моделі множинної регресії з використанням фіктивних змінних для моделювання цін на квартири. Для розглянутої моделі проведено аналіз значущості параметрів та моделі в цілому, усунуто наслідки гетероскедастичності збурень.

Запропонована специфікація моделі може бути змінена завдяки введенню додаткових регресорів, що можуть деталізувати корисні характеристики об'єкту нерухомості.

ЛІТЕРАТУРА:

1. Максишко Н. Нечітка модель ідентифікації фаз на ринку нерухомості / Н. Максишко, В. Шаповалова // Нейронечіткі технології моделювання в економіці: науково-аналітичний журнал – К.: Київський національний університет ім. В. Гетьмана. – 2014. – № 3. – С. 94–119.
2. Паттури Я. Економіка нерухомості: [учеб. пособие] / Я. Паттури; НовГУ ім. Ярослава Мудрого. – Великий Новгород, 2002. – 243 с.
3. Про державну реєстрацію речових прав на нерухоме майно та їх обтяжень: Закон України від 10.11.2015 р., зі змінами та доповненнями [Електронний ресурс]. – Режим доступу: <http://zakon3.rada.gov.ua/laws/show/1952-15>.
4. Какой дом можно считать новостройкой? [Електронний ресурс]. – Режим доступу: http://ack1.ru/Library/kakoj_dom_mozhno_schitat_novostrojkoj.html.
5. Новостройка или вторичное жилье [Електронний ресурс]. – Режим доступу: <http://stroypryam.ru/home/njedvizhivosti/1292-novostroyka-ili-vtorichnoe-zhile.html>.
6. Новостройки Украины [Електронний ресурс]. – Режим доступу: <http://novostroyki.lun.ua>.