

Рядно О.А.,

д.т.н.,

професор кафедри прикладної математики та інформатики,  
Університет митної справи та фінансів

Беркут О.В.,

старший викладач кафедри прикладної математики та інформатики,  
Університет митної справи та фінансів

## ЕКОНОМЕТРИЧНЕ МОДЕЛЮВАННЯ ФОРМУВАННЯ ВАЛОВОГО РЕГІОНАЛЬНОГО ПРОДУКТУ НА ОСНОВІ ПАНЕЛЬНИХ ДАНИХ

**Анотація.** У статті наведено результати дослідження впливу основних соціально-економічних показників на величину валового регіонального продукту України. Доведено доцільність використання методів обробки панельних даних. Моделювання проведено за допомогою економетричних методів з урахуванням панельної структури даних. У результаті побудови та аналізу моделей різних видів за допомогою статистичних тестів здійснено вибір найбільш адекватної моделі.

**Ключові слова:** регіон, диференціація соціально-економічного розвитку, економіко-математичне моделювання, панельні дані, аналіз панельних даних.

**Постановка проблеми.** Для ефективного прийняття рішень на регіональному та державному рівнях необхідно визначити загальні фактори, що впливають на рівень соціально-економічного розвитку, за якими спостерігається найбільша диференціація.

Для вирішення цього завдання можна використовувати методи обробки панельних даних. Панельні дані складаються з спостережень одних і тих самих об'єктів, які здійснюються в послідовні періоди часу. Таким чином, панельні дані поєднують у собі дані просторового типу та дані типу часових рядів, дані цього типу дають змогу враховувати індивідуальні відмінності між економічними одиницями. Ці методи дають можливість урахувати ефекти, які неможливо простежити, залишаючись у рамках звичайних регресійних моделей. Показник валового регіонального продукту (ВРП) виберемо як узагальнюючу характеристику економічного розвитку регіону. Під час моделювання обсягу валового регіонального продукту на душу населення є можливість урахувати, крім спостережуваних макроекономічних показників, фактори, які або не спостерігаються, або не можуть бути представлені в чисельній формі, але які можуть чинити істотний вплив на досліджуваний показник: географічне положення, природні ресурси тощо.

**Аналіз останніх досліджень і публікацій.** Питанню моделювання залежності валового регіонального (або внутрішнього) продукту від соціально-економічних показників присвячено багато наукових праць як вітчизняних, так і закордонних економістів. Автори розглядають вплив роздрібною торгівлі на розвиток економіки, а саме її внесок у валовий внутрішній продукт, залежність валового регіонального продукту від розвитку підприємництва в регіоні, показників якості життя населення, факторів трудового потенціалу. Всі ці роботи поєднують використання кореляційно-регресійного аналізу просторових даних для побудови моделі [1–4].

**Виділення невирішених раніше частин загальної проблеми.** Такі моделі дають можливість визначити вплив окремих соціально-економічних показників на обсяг ВРП, але не дають змоги виявити індивідуальні відмінності між регіонами. Цього недоліку позбавлені моделі, які враховують панельну структуру даних.

**Мета статті** полягає у дослідженні впливу соціально-економічних факторів на валовий регіональний продукт з урахуванням панельної структури даних.

**Виклад основного матеріалу дослідження.** Для моделювання величини ВРП на душу населення розглядалися щорічні статистичні дані по 25 регіонах України в період з 2005 по 2015 р. Джерелом статистичних даних є статистичні збірники «Регіони України» за 2009–2015 рр., офіційні сайти Державної служби статистики України та головних управлінь статистики областей України.

Дані сформовано у вигляді панелі, яка складається із спостережень за соціально-економічними показниками в означений період. Панель даних є незбалансованою у зв'язку з відсутністю даних за 2014 та 2015 рр. по АР Крим; окрім того, дані по Донецькій та Луганській областях є неповними та надаються без урахування частини зони проведення антитерористичної операції. Всі розрахунки, які будуть наведені далі у цій роботі, проведено за допомогою системи статистичної обробки даних та програмування R.

Уведемо умовні позначення, які будуть використовуватися далі. Нехай  $Y_{it}$  – залежна змінна економічного об'єкта (значення валового регіонального продукту по 25 регіонах України) в момент часу  $t$ ,  $X_{it}$  – набір незалежних змінних (вектор розмірності  $k$ , що містить значення соціально-економічних показників за період спостереження),  $\beta$  – оцінки параметрів,  $\alpha$  – вільний член,  $\varepsilon_{it}$  – відповідна помилка,  $i = 1, \dots, N = 1, \dots, 25$  – номер регіону,  $t = 1, \dots, T = 2005, \dots, 2015$  – момент часу.

Базовим варіантом розглядається така загальна модель:

$$Y_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it}, \quad (1)$$

Величина  $\alpha_i$  виражає індивідуальний ефект об'єкта  $i$ , який не залежить від часу, при цьому  $x_{it}$  не містять константу.

Відносно індивідуальних ефектів економічних одиниць існують такі основні підходи, які відрізняються один від одного припущенням щодо статусу змінних  $\alpha_i$  в моделі.

Об'єднана модель панельних даних або наскрізна регресія (pooled model):

$$Y = \alpha + X\beta + \varepsilon \quad (2)$$

Ця модель, по суті, не враховує панельну структуру даних. При цьому передбачається, що всі помилки  $\varepsilon_{it}$  некорельовані

між собою, та некорельовані з усіма пояснюючими змінними  $x_{it}$ . Окрім того, модель передбачає відсутність істотних відмінностей між об'єктами вибірки. Неоднорідність об'єктів моделі була доведена на попередньому етапі дослідження [5], тому обмеження, які накладаються на дані в разі використання моделі цього типу, не виконуються. Враховуючі наведене вище, модель цього типу буде використовуватися для визначення оптимального переліку змінних тільки як допоміжна на початковому етапі дослідження. Для оцінки векторів коефіцієнтів  $\beta$  використовується звичайний метод найменших квадратів.

**Модель із фіксованими ефектами (fixed effects model)** передбачає, що ефекти  $\alpha_i$  – це фіксовані невідомі параметри моделі.

Під час оцінки параметрів економетричної моделі панельних даних використовують так звані within-оцінки (або внутрішньогрупові оцінки), які можна отримати, побудувавши модель за відхиленнями від групових середніх та between-оцінки (або міжгрупові оцінки), побудувавши модель за груповими середніми.

Модель панельних даних, що побудована на середніх за часом значеннях змінних для кожного  $i$ -го об'єкта (between regression):

$$\frac{1}{T} \sum_{t=1}^T Y_{it} = \alpha + \frac{1}{T} \sum_{t=1}^T X_{it} \beta + \varepsilon_i. \quad (3)$$

Модель панельних даних, що побудована на відхиленнях від групових середніх, або модель із фіксованими ефектами (within regression):

$$Y_{it} - Y_i = (X'_{it} - X'_i) \beta + \varepsilon_{it} - \varepsilon_i. \quad (4)$$

Цей вид оцінки називають оцінкою з фіксованим ефектом (fixed effects model).

Якщо необхідно врахувати в моделі неспостережувані фактори, які мають різні значення для кожного моменту часу, але однакові для всіх об'єктів сукупності, то фіктивні змінні, що визначають фіксовані ефекти, вводяться для моменту часу (period fixed effects).

**Модель панельних даних із випадковими ефектами (random effects model)** передбачає, що ефекти  $\alpha_i$  – це випадкові величини. При цьому передбачається, що всі помилки  $\varepsilon_{it}$  не корельовані з  $\alpha_i$ .

$$Y_{it} = X_{it} \beta + const + \alpha_i + \varepsilon_{it}, \quad (5)$$

де const – це деякий параметр, який є спільним для всіх об'єктів у всі моменти часу.

Для оцінки векторів коефіцієнтів  $\beta$  використовується оцінка з випадковим ефектом, яка отримується за допомогою узагальненого методу найменших квадратів [6, с. 357–374].

Спочатку було проведено оцінювання об'єднаної моделі панельних даних, до якої увійшли показники, що входять до перших чотирьох компонент, тобто узагальнених факторів, за якими спостерігається найбільша регіональна диференціація без урахування панельної структури даних [7]. Показники, що увійшли до початкового набору, наведено в табл. 1.

У цілому, згідно з F-критерієм Фішера, отримана модель є статистично значимою з коефіцієнтом детермінації 61,27%. Однак у силу мультиколінеарності багато факторів й цій моделі виявилися незначущими, тому, враховуючі мультиколінеарність, значимість і силу впливу ознак-факторів на результуючий фактор, методом покрокового виключення отримана модель, до якої увійшли показники: KAPINV, PROM та KPEI як показники, що характеризують розвиток промисловості в регіоні; ROZTORG як показник, який непрямо чинить вплив на рівнем економічного розвитку регіону та якістю життя населення; CHNPM та MPN як показники, що відображають якість життя населення регіону.

Моделювання з новим набором показників показало не гірші результати статистичної значимості отриманої моделі. При цьому введення у модель фіктивних змінних *Region i* для кожного регіону підвищило значення F-критерію на 10%, що підтверджує доцільність проведення дослідження в напрямі врахування панельної структури даних.

Далі побудовано чотири типи моделей, що враховують панельну структуру даних [8; 10–15]. Слід відзначити, що коефіцієнти детермінації в моделях із панельними даними недоцільно використовувати для того, щоб визначити, яка з моделей краще. Так, якщо одну й ту саму модель, наприклад, оцінити за допомогою звичайного методу найменших квадратів та за допомогою випадкового ефекту, то об'єднаний коефіцієнт детермінації в першому випадку завжди буде меншим, ніж відповідний об'єднаний коефіцієнт детермінації для випадкового ефекту, навіть якщо використання випадкового ефекту є більш доцільним. Коефіцієнти детермінації можна використовувати для порівняння моделей, що оцінюються одним методом.

Проаналізуємо отримані результати моделювання та проведемо попарне порівняння моделей за основними характеристиками та спеціальними статистичними тестами. У табл. 2 наведено характеристики моделей, що були побудовані.

Таблиця 1

Показники соціально-економічного розвитку регіонів України

Назва показника	Умовне позначення	Одиниця виміру
Капітальні інвестиції на одну особу	KAPINV	грн.
Обсяг реалізованої промислової продукції (товарів, послуг) на одну особу населення	PROM	грн.
Фінансові результати підприємств до оподаткування	FR	млн. грн.
Оборот роздрібною торгівлі	ROZTORG	млн. грн.
Прямі іноземні інвестиції (акціонерний капітал) на одну особу населення	ININVEST	дол. США
Розподіл загального обсягу витрат за напрямками інноваційної діяльності	INOV	тис. грн.
Коефіцієнт покриття експортом імпорту	KPEI	
Міграційний приріст (скорочення) населення	MPN	тис. осіб
Середньомісячна номінальна заробітна плата працівників	SZP	грн.
Частка населення із середньодушовими еквівалентними загальними доходами на місяць нижче прожиткового мінімуму	CHNPM	відсотків
Коефіцієнт зайнятості	KZAN	

Таблиця 2  
Результати моделювання з урахуванням панельної структури даних

Модель	Pooling	Between	Within	Random
(Intercept)	5037,32 * (2178,39)	-113,94 (7164,37)		5318,98 * (2226,52)
KAPINV	0,81 * (0,35)	1,06 (0,61)	0,28 (0,47)	0,80 * (0,37)
PROMPR	0,41 *** (0,05)	0,43 *** (0,08)	0,02 (0,08)	0,38 *** (0,05)
ROZTORG	0,25 * (0,11)	0,76 (0,52)	0,84 *** (0,15)	0,27 * (0,12)
CHNPM	60,71 (42,89)	-22,57 (173,90)	91,50 * (42,39)	60,75 (42,86)
KPEI	1105,28 ‘ (649,30)	742,48 (1092,01)	-1869,31 ‘ (1009,64)	1035,61 (699,52)
MPN	850,55 *** (205,68)	215,56 (515,13)	1252,72 *** (272,81)	878,38 *** (217,44)
R <sup>2</sup>	0,61	0,92	0,45	0,56

Коефіцієнти за змінних, що відмічено \*\*\*, \*\*, \* є значимими на 1-, 5- та 10% рівнях відповідно, ‘ – значимість коефіцієнтів наближується до 10% рівня.

Джерело: розраховано авторами за допомогою мови статистичної обробки даних R

Вибір найбільш адекватної моделі здійснювався шляхом попарного порівняння моделей із використанням відповідних тестів [6, с. 375–379; 8, с. 20–22].

За результатами моделювання найвищі значення коефіцієнта детермінації  $R^2 = 0,92$  отримано для моделі з міжгруповими оцінками (between regression). Таке значення коефіцієнта детермінації відображає високу якість підгонки моделі та свідчить про те, що вплив на обсяг ВРП середніх за часом показників для кожного регіону є більш суттєвим, аніж коливання цих показників навколо середніх протягом часу. Але, незважаючи на високу якість підгонки моделі, значимим виявився коефіцієнт  $\beta$  тільки для одного фактора, а саме для PROMPR. Усі інші фактори в межах цієї моделі є незначимими; крім того, спостерігаються більш високі сумарні значення стандартних помилок коефіцієнтів для моделі з міжгруповими оцінками порівняно з іншими моделями, тобто отримані оцінки  $\beta$  є менш ефективними, тому далі ця модель буде використовуватися тільки як допоміжна.

Для моделі з фіксованими ефектами (within regression) значущими виявилися коефіцієнти за факторів ROZTORG, CHNPM, KPEI та MPN. При цьому показники ROZTORG, CHNPM та MPN характеризують рівень економічного розвитку регіону та якість життя населення, а показник KPEI є відображенням торгового балансу регіону. Всі показники, крім KPEI, мають позитивні коефіцієнти та рівень значення помилки розрахунку, що є співставним з іншими моделями (крім моделі з міжгруповими оцінками). Тобто зростання цих показників позитивно впливає на зростання обсягу ВРП. Коефіцієнт для показника KPEI є від’ємним, причому така поведінка коефіцієнту властива тільки для моделі з фіксованими ефектами, що більш реально відображає дійсність, тому що протягом останніх років Україна мала пасивний торговий баланс.

Розглянемо отримані результати для моделі з випадковими ефектами. Порівняно з моделлю з фіксованими ефектами змінилися кількість та склад значущих коефіцієнтів  $\beta$  для факторів, що розглядаються. Значущими виявилися такі фактори:

KAPINV та PROMPR як показники, що характеризують розвиток промисловості в регіоні; ROZTORG та MPN як відображення якості життя населення. Всі показники мають позитивні коефіцієнти та рівень значення помилки розрахунку дещо менший ніж для моделі з фіксованими ефектами.

Проведемо попарне порівняння моделей за допомогою F-критерію Фішера, Хаусмана та Бройша-Погана.

Для вибору між моделлю наскрізної регресії та моделлю з фіксованими ефектами проведемо F-тест, що перевіряє гіпотезу  $H_0: \alpha_1 = \alpha_2 = \alpha_1 \dots = \alpha_n$  для моделі з фіксованими ефектами, тобто проведемо перевірку на відсутність фіксованих ефектів. У нашому випадку  $F_{\text{табл}} = 1,57$ , що менше ніж  $F_{\text{розр}} = 3,0771$  для моделі з фіксованими ефектами, тому гіпотеза про відсутність фіксованих ефектів відкидається.

Розглянемо результати тесту Хаусмана для вибору між моделями з фіксованими та випадковими ефектами. У моделі з випадковими ефектами передбачається, що індивідуальні ефекти не корелюють з іншими пояснюючими факторами. Таким чином необхідно перевірити гіпотезу  $H_0: \text{Cov}(\alpha_i, x_{it}) = 0$ . Альтернативна гіпотеза, що  $\text{Cov}(\alpha_i, x_{it})$  відмінно від 0, тобто нульовою гіпотезою є те, що кращою є модель із випадковими ефектами. За тестом Хаусмана значення  $p_{\text{value}} = 1,549e-07 < 0,05$ , тому нульова гіпотеза відкидається, тобто модель із фіксованими ефектами є кращою, ніж модель із випадковими ефектами.

Для оцінки необхідності врахування часових ефектів проведено тест Бройша-Погана. Розраховане значення  $p_{\text{value}} < 2,2e-16$ , що значно менше, ніж 0,05, тому в нашому випадку використання моделей із фіксованими часовими ефектами є недоцільним.

Таким чином, рівняння, що відображає вплив соціально-економічних показників на обсяг валового регіонального продукту з урахуванням індивідуальних фіксованих ефектів регіону, має вигляд:

$$VRP = ER + 0,84*ROZTORG + 91,5*CHNPM - 1869,31*KPEI + 1252,72*MPN, \quad (6)$$

де ER – індивідуальний фіксований ефект регіону, коефіцієнти за змінних є значимими на 10%-му рівні.

Розраховані коефіцієнти еластичності для показників моделі наведено в табл. 3.

Таблиця 3  
Значення показників еластичності моделі

Показник	Коефіцієнт еластичності
ROZTORG	0,471
CHNPM	0,084
KPEI	0,131
MPN	-0,009

Джерело: розраховано авторами за допомогою мови статистичної обробки даних R

Найбільш істотний вплив на зростання ВРП має показник обороту роздрібною торгівлі (ROZTORG), а саме збільшення цього показника на 1% приведе до збільшення ВРП на 0,471%, тоді як зміни показника міграційного приросту (скорочення) населення (MPN) на 1% приведуть до змін ВРП тільки на 0,009%, причому напрям цих змін буде протилежним.

Фіксовані ефекти відображають вплив на ВРП факторів, індивідуальних для кожного регіону. Це можуть бути фактори, пов’язані з природно-кліматичними умовами, наявністю або відсутністю на території регіону корисних копалин або інших природних ресурсів, особливим економічним становищем регі-

## Фіксовані ефекти для регіонів України

№ за/п	Область	Ефект	№ за/п	Область	Ефект
1	Донецька	34407,7359	14	Житомирська	7595,8959
2	Дніпропетровська	29056,8144	15	Київська	7474,6568
3	Полтавська	22612,1719	16	Херсонська	7184,3642
4	Луганська	19526,678	17	Харківська	6969,1566
5	Запорізька	15396,4484	18	Івано-Франківська	6721,3765
6	Кіровоградська	12827,9952	19	Хмельницька	6669,4525
7	Миколаївська	11922,9092	20	Закарпатська	5102,683
8	Сумська	11137,9993	21	Тернопільська	4320,8824
9	Черкаська	10190,8524	22	Волинська	2682,6499
10	Львівська	9644,5036	23	Одеська	2382,2548
11	Вінницька	9608,2653	24	АР Крим	2340,6711
12	Чернігівська	8485,6245	25	Чернівецька	921,5629
13	Рівненська	7796,4542			

Джерело: розраховано авторами за допомогою мови статистичної обробки даних R

ону, якістю управління в регіоні та ін. Чим більше індивідуальний фіксований ефект, тим більше можливостей у регіону збільшити ВРП за рахунок факторів, що не включені в модель.

Регіони було проранжовано за величиною фіксованого ефекту (табл. 4). До п'ятірки регіонів із максимальними фіксованими ефектами належать Донецька, Дніпропетровська, Полтавська, Луганська, Запорізька області. На попередньому етапі дослідження було проведено кластерний аналіз диференціації соціально-економічного розвитку регіонів України [7]. За його результатами було виділено п'ять кластерів, причому вище наведені області увійшли до складу п'ятого та третього кластерів. Регіони третього кластера (Дніпропетровська, Полтавська та Запорізька області) займають лідируючі позиції за економічним розвитком та якістю життя населення. Области п'ятого кластера (Донецька та Луганська області) характеризуються поєднанням високого економічного розвитку, зменшення коефіцієнту покриття експорту імпортом та від'ємним міграційним приростом.

**Висновки.** Результати проведеного дослідження підтверджують наявність індивідуальних особливостей регіонів України та доцільність використання під час аналізу моделей, що враховують панельну структуру даних. За результатами F-критерію Фішера, Хаусмана та Бройша-Погана для опису залежності валового регіонального продукту та соціально-економічних показників для об'єктів сукупності (регіонів) кращою є модель із фіксованими ефектами.

До п'ятірки регіонів із максимальними фіксованими ефектами належать Донецька, Дніпропетровська, Полтавська, Луганська, Запорізька області. Ці регіони займають лідируючі позиції за економічним розвитком (насамперед розвитком промисловості) та якістю життя населення. Крім того, Донецька та Луганська області характеризуються поєднанням високого економічного розвитку, зменшенням коефіцієнту покриття експорту імпортом та від'ємним міграційним приростом.

Загальними факторами, що впливають на величину валового регіонального продукту, є: оборот роздрібною торгівлі, частка населення із середньодушовими еквівалентними загаль-

ними доходами на місяць нижче прожиткового мінімуму, коефіцієнт покриття експортом імпорту та коефіцієнт міграційного приросту (скорочення) населення. Найбільш істотний вплив на зростання валового регіонального продукту має показник обороту роздрібною торгівлі, найменший – показник міграційного приросту (скорочення) населення.

Результати кількісних змін показників соціально-економічного розвитку регіонів України, що отримані за допомогою моделей панельних даних, можуть бути основою для подальшої експертної та аналітичної оцінки якісних змін у регіональному розвитку.

**Література:**

1. Емцева Е.Д. Моделирование взаимосвязи валового регионального продукта и показателей качества жизни / Е.Д. Емцева, А.Л. Мазелис // Вектор науки ТГУ. Экономика и управление. – 2016. – № 3(26). – С. 24–28.
2. Кармелюк Г.І. Економетричні моделі залежності валового внутрішнього продукту від роздрібною товарообороту / Г.І. Кармелюк // Економічний аналіз. – 2010. – Вип. 5. – С. 107–110.
3. Новіков Д.А. Економіко-математичне моделювання впливу показників-факторів трудового потенціалу на валовий регіональний продукт промислових регіонів України / Д.А. Новіков // Вісник економіки транспорту і промисловості. – 2014. – № 47. – С. 19–23.
4. Турський І.В. Моделювання залежності валового регіонального продукту від розвитку підприємництва в регіонах України / І.В. Турський // Науковий вісник Херсонського державного університету. – 2017. – Вип. 25. – Ч. 2. – С. 77–81.
5. Беркут О.В. Структура та динаміка диференціації соціально-економічного розвитку регіонів України / О.В. Беркут // Інфраструктура ринку. – 2017. – № 6. – С. 182–188.
6. Економетрика : [учебник] / Я.Р. Магнус, П.К. Катъшев, А.А. Пересецкий ; 6-е изд., перераб. и доп. – М. : Дело, 2004. – 576 с.
7. Рядно О.А. Дослідження структури та динаміки диференціації соціально-економічного розвитку регіонів України на основі кластерного аналізу / О.А. Рядно, О.В. Беркут // Економічний вісник Донбасу. – 2016. – № 1(43). – С. 60–67.
8. Croissant Y. Panel Data Econometrics in R: The plm Package / Y. Croissant, G. Millo // Journal of Statistical Software published by the American Statistical Association. – 2008. – Volume 27. – Issue 2. – P. 43.

**Рядно А.А., Беркут Е.В. Эконометрическое моделирование формирования валового регионального продукта на основе панельных данных**

**Аннотация.** В статье приведены результаты исследования влияния основных социально-экономических показателей на величину валового регионального продукта Украины. Доказана возможность использования методов обработки панельных данных. Моделирование проведено с помощью эконометрических методов с учетом панельной структуры данных. В результате построения и анализа моделей различных видов с помощью статистических тестов осуществлен выбор наиболее адекватной модели.

**Ключевые слова:** регион, дифференциация социально-экономического развития, экономико-математическое моделирование, панельные данные, анализ панельных данных.

**Ryadno O.A., Berkut O.V. Econometric modeling formation of gross regional product based on panel data**

**Summary.** The article presents the results of the study of the influence of the main socio-economic indicators on the gross regional product of Ukraine. The possibility of using panel data processing methods is proved. Simulation is performed using econometric methods taking into account the panel structure of the data. As a result of the construction and analysis of models of different species, the most adequate model was selected using statistical tests.

**Keywords:** region, differentiation of socio-economic development, economic-mathematical modeling, panel data, analysis of panel data.