

Міністерство освіти і науки України
Національний університет «Києво-Могилянська академія»
Факультет економічних наук
Кафедра фінансів

Кваліфікаційна робота
освітній ступінь – бакалавр
на тему: « **ФУНДАМЕНТАЛЬНІ ФАКТОРИ ДОХІДНОСТІ АКЦІЙ
ПУБЛІЧНИХ КОМПАНІЙ: ПРИКЛАД США** »

Спеціальності:

072 Фінанси, банківська справа та страхування

Масловської Єлизавети Миколаївни

Керівник: Якуненко К. В.

Кандидат економічних наук, старший викладач

Рецензент: Харламова Г.О.

Кваліфікаційна робота захищена

з оцінкою «_____»

Секретар ЕК _____ Антонова Є.І.

«_____» _____ 2026 р.

ЗМІСТ

ВСТУП	3
РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ ФАКТОРНОГО АНАЛІЗУ ДОХІДНОСТІ АКЦІЙ ПУБЛІЧНИХ КОМПАНІЙ.....	7
1.1. Теоретичні підходи до трактування дохідності акцій	7
1.2. Еволюція підходів до пояснення дохідності акцій	9
1.3. Емпіричні дослідження зв'язку фінансових характеристик компаній із майбутньою дохідністю акцій	12
Висновки до розділу 1	16
РОЗДІЛ 2. ФОРМУВАННЯ ДАНИХ ТА ОЦІНЮВАННЯ МОДЕЛІ ДОХІДНОСТІ АКЦІЙ	18
2.1. Формування інформаційної бази та вибірки дослідження.....	18
2.2. Побудова змінних моделі та перевірка емпіричної вибірки	20
2.3. Економетрична специфікація та оцінювання моделі	26
Висновки до розділу 2	31
РОЗДІЛ 3. АНАЛІЗ СТІЙКОСТІ ТА ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ ЕМПІРИЧНОЇ МОДЕЛІ	34
3.1. Аналіз причин отриманих результатів	34
3.2. Уточнення причин слабких і нестійких результатів на основі літературних підходів	41
3.3. Сегментована переоцінка центральної моделі	45
Висновки до розділу 3	48
ВИСНОВКИ	50
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	54
Додаток А.....	56
Додаток Б.....	57

ВСТУП

Актуальність теми. Фондовий ринок відображає очікування інвесторів щодо майбутніх результатів компаній, а дохідність акцій є одним із головних показників такої оцінки. Від того, які фундаментальні фінансові показники пов'язані з майбутньою дохідністю, залежить оцінка інвестиційної привабливості компаній, формування портфелів і розуміння ринкової реакції на фінансові показники емітентів. У фінансовій літературі це питання розглядається через різні підходи: від класичних моделей ризику до багатофакторних моделей і характеристик окремих компаній. Значний внесок у цю проблематику зробили Sanjoy Basu (1977), Eugene F. Fama та Kenneth R. French (1992, 1993, 2015), Robert Novy-Marx (2013), Sheridan Titman, K. C. John Wei та Feixue Xie (2004), Narasimhan Jegadeesh і Sheridan Titman (1993), Jonathan Lewellen (2015), Jeremiah Green, John R. M. Hand та X. Frank Zhang (2017).

У вітчизняній літературі близькі до теми аспекти розглядаються переважно через український фондовий ринок. *Д. Митько (2025)* досліджує чинники довгострокової дохідності акцій на прикладі українського фондового ринку, *Є. Сова (2020)* аналізує зв'язок між монетарними індикаторами та показниками фондового ринку, а *І. Волошина (2015)* розглядає поєднання фундаментального й технічного аналізу для прогнозування цін українських акцій.

Результати попередніх досліджень відрізняються залежно від періоду, вибірки, способу побудови змінних і специфікації моделі. Для США це питання має окреме значення, оскільки він є одним із найбільш досліджених і водночас мінливих фондових ринків. Актуальність роботи полягає у перевірці зв'язку між фундаментальними факторами публічних компаній США та дохідністю їхніх акцій у наступному періоді, а також в оцінюванні стійкості цього зв'язку за різних специфікацій моделі й сегментації вибірки.

Дослідницьке питання роботи полягає в тому, які фундаментальні фінансові показники публічних компаній США пов'язані з дохідністю їхніх акцій у

наступному кварталі і наскільки стійким є цей зв'язок у різних специфікаціях моделі та сегментах вибірки.

Гіпотеза дослідження полягає в тому, що фундаментальні фінансові характеристики публічних компаній США можуть мати різний за силою та стійкістю зв'язок із дохідністю їхніх акцій у наступному кварталі. Передбачається, що частина показників зберігатиме статистичну значущість у різних специфікаціях моделі, тоді як інші можуть бути чутливими до часових ефектів, складу предикторів або структури вибірки.

Метою роботи є емпірично оцінити зв'язок між фундаментальними фінансовими показниками публічних компаній США та дохідністю їхніх акцій у наступному кварталі, а також пояснити стійкість або нестійкість отриманих результатів у різних специфікаціях моделі та сегментах вибірки.

Для досягнення поставленої мети визначено такі **задачі**:

1. Сформувати теоретичну рамку дослідження дохідності акцій, уточнивши її зміст, складові та зв'язок із характеристиками компаній.
2. Систематизувати підходи, у яких фундаментальні характеристики компаній розглядаються як можливі предиктори майбутньої дохідності акцій.
3. Обґрунтувати вибір фундаментальних показників для емпіричної перевірки з урахуванням попередніх досліджень і змісту кожної змінної.
4. Сформувати панельну вибірку публічних компаній США за 2015Q1–2025Q4 та побудувати змінні моделі відповідно до логіки зіставлення характеристик компанії в періоді t з дохідністю акцій у періоді $t+1$.
5. Оцінити зв'язок між фундаментальними показниками компаній і майбутньою дохідністю акцій у базовій специфікації та специфікації з часовими фіксованими ефектами.
6. Перевірити стійкість отриманих результатів через додаткові аналітичні процедури, зокрема індивідуальні моделі, квартильне групування, вкладені специфікації та сегментацію вибірки.
7. Інтерпретувати стійкі й нестійкі результати з урахуванням структури вибірки та підходів попередніх емпіричних досліджень.

Об'єктом дослідження виступає процес формування дохідності акцій публічних компаній на фондовому ринку США.

Предметом дослідження є зв'язок між фундаментальними фінансовими показниками публічних компаній США та дохідністю їхніх акцій у наступному кварталі.

Методи дослідження. У роботі використано аналіз наукової літератури, систематизацію теоретичних і емпіричних підходів та порівняльний аналіз попередніх досліджень. Ці методи застосовано для формування теоретичної рамки дослідження, визначення змісту дохідності акцій і обґрунтування набору фундаментальних показників моделі.

Емпірична частина ґрунтується на формуванні панельного датасету, point-in-time підході до узгодження фінансової звітності з датами спостереження, розрахунку фінансових коефіцієнтів і ринкових змінних. Для попередньої перевірки даних використано описову статистику, кореляційний аналіз, вінсоризацію екстремальних значень на рівнях 1% і 99%, а також аналіз мультиколінеарності через кореляційну матрицю та VIF.

Економетричне оцінювання проведено за допомогою Pooled OLS та моделі з часовими фіксованими ефектами. Для стандартних похибок використано кластеризацію за компаніями. Додатково застосовано діагностичні тести специфікації, зокрема Pesaran CD, Breusch–Pagan LM, Ramsey RESET, тест Бройша–Годфрі/Вулдріджа на автокореляцію в панельних даних і тести на надлишкові часові фіксовані ефекти. Стійкість і причини отриманих результатів перевірено через одновимірні регресійні моделі, кuartильне та терцильне групування спостережень, вкладені специфікації з послідовним вилученням предикторів, сегментацію вибірки за GPA і LMC та порівняльну інтерпретацію результатів у різних специфікаціях і сегментах вибірки.

Інформаційну базу дослідження становлять статті у провідних фінансових журналах, зокрема The Journal of Finance, Journal of Financial Economics, The Review of Financial Studies, Journal of Financial and Quantitative Analysis та Critical Finance Review; робочі матеріали й препринти, розміщені на SSRN та NBER; матеріали

CFA Institute; вітчизняні академічні джерела з репозитарію eKMAIR, Наукових записок НаУКМА та Ukrainian Journal Емпіричну базу роботи сформовано на основі квартальної фінансової звітності та ринкових даних публічних компаній США за 2015Q1–2025Q4, отриманих із Financial Modeling Prep. Фінальний панельний датасет охоплює 417 компаній і 17 234 квартальні спостереження.

Практичне значення одержаних результатів полягає в обґрунтуванні підходу до аналізу неоднозначних і нестійких результатів багатофакторного моделювання дохідності акцій. У роботі показано, як за допомогою додаткових перевірок оцінювати робастність коефіцієнтів, виявляти залежність показників від специфікації моделі та розкривати неоднорідність структури вибірки. Запропонована послідовність дій може використовуватися студентами, молодими дослідниками й аналітиками-початківцями під час роботи з панельними фінансовими даними та побудови моделей майбутньої дохідності акцій. Одержані результати можуть бути орієнтиром для подальшого уточнення моделі, зокрема через перегляд набору фундаментальних характеристик, використання скоригованих цін, зміну часових горизонтів і перевірку нелінійних зв'язків.

Наукова новизна одержаних результатів полягає в уточненні підходу до емпіричної перевірки зв'язку між фундаментальними характеристиками компаній і майбутньою дохідністю акцій на прикладі публічних компаній США за 2015–2025 рр. У роботі конкретизовано, що статистична значущість окремих фундаментальних показників залежить від специфікації моделі та структури вибірки. Набуло подальшого розвитку використання сегментованої перевірки результатів, що дозволило виявити неоднорідність зв'язку показників балансової вартості до ринкової вартості та боргу до активів із майбутньою дохідністю, а також відсутність стійкого результату для темпу зростання активів у межах сформованої вибірки.

Ключові слова: дохідність акцій, фундаментальні фінансові показники, публічні компанії, фондовий ринок США, майбутня дохідність, панельні дані, Pooled OLS, часові фіксовані ефекти, робастність результатів, фінансові показники, ринкова капіталізація, моментум, боргове навантаження.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ЗАСАДИ ФАКТОРНОГО АНАЛІЗУ ДОХІДНОСТІ АКЦІЙ ПУБЛІЧНИХ КОМПАНІЙ

1.1 Теоретичні підходи до трактування дохідності акцій

Дохідність акцій є одним із ключових показників, через який оцінюється результат інвестування на фондовому ринку. У фінансовій теорії вона відображає відносну зміну вартості інвестиції за певний період і виступає базовою змінною, відносно якої аналізують вплив ризику, ринкових умов і характеристик окремих компаній. У найпростішому вигляді дохідність акції формується за рахунок двох компонентів: поточного доходу, який інвестор отримує у формі дивідендів, і зміни ринкової ціни самої акції. У такому розумінні дохідність відображає сукупний результат володіння активом [4, с. 14]. Тому в аналізі акцій важливо враховувати зміну ринкової ціни і виплати, які супроводжують інвестицію протягом періоду.

Дохідність акцій не має фіксованого характеру. Інвестор не може наперед визначити, якою буде ринкова ціна акції в наступному періоді та як ринок відреагує на зміни у фінансових показниках компанії. Майбутній результат залежить від прибутковості бізнесу, темпів його розвитку, структури фінансування, змін у конкурентному середовищі й загального стану фондового ринку. Аналіз дохідності акцій є складнішим порівняно зі звітними показниками компанії, адже вона відображає внутрішній стан емітента та реакцію ринку на нього.

Цей показник одночасно відображає вже реалізовані ринкові зміни та очікування щодо майбутніх перспектив компанії. З одного боку, дохідність акцій відображає результат уже реалізованих процесів, що знайшли відображення в ринковій ціні. З іншого боку, це пов'язано з оцінкою майбутніх перспектив, оскільки інвестори купують акції, орієнтуючись на очікуваний фінансовий результат у наступних періодах. У дослідженнях *stock returns* дохідність використовується як залежна

змінна, за допомогою якої перевіряється, чи мають характеристики компанії систематичний зв'язок із подальшою ринковою динамікою її акцій.

У фінансових дослідженнях дохідність акцій розглядають у двох взаємопов'язаних вимірах: як реалізовану та як очікувану. Реалізована дохідність відображає фактичний результат інвестування після завершення певного періоду. Очікувана дохідність характеризує прогнозований рівень майбутнього результату, який інвестори пов'язують із ризиком, доступною інформацією та характеристиками конкретної компанії. Теоретичні моделі ціноутворення активів зосереджуються на *expected return*, оскільки саме вона дає змогу пояснити, чому різні активи забезпечують різний інвестиційний результат. Емпірична перевірка таких взаємозв'язків будується інакше: вона спирається на спостережувану дохідність, яка вже реалізувалася, і через неї тестує, чи були певні характеристики компанії пов'язані з її подальшою ринковою поведінкою [6, с. 427–428].

Дохідність акцій є складним об'єктом емпіричного аналізу також через те, що її короткострокова динаміка значною мірою визначається новинами, змінами ринкових очікувань і загальними ринковими шоками. Унаслідок цього зв'язок між характеристиками компанії та подальшою дохідністю не є механічним і не завжди демонструє стійкість у різних вибірках і специфікаціях. *Shihao Gu, Bryan Kelly and Dacheng Xiu (2020)* зазначають, що варіація дохідностей значною мірою зумовлена непередбачуваними новинами, а оцінювання *expected returns* є складним емпіричним завданням [13, с. 2224]. *Amit Goyal, Ivo Welch and Aleksander Zafirov (2024)* також показують, що багато змінних, запропонованих у літературі як предиктори фондової дохідності, втрачають стійкість результатів після жорсткіших перевірок і розширення вибірки [11, с. 3490–3491].

Для емпіричного дослідження важливе значення має і часовий горизонт аналізу. На коротких інтервалах сильніше проявляються ситуативні ринкові реакції, тоді як на довших горизонтах зростає ймовірність того, що в динаміці дохідності чіткіше відобразяться більш стійкі характеристики самої компанії. Це робить вибір горизонту спостереження важливою частиною дослідницької постановки, оскільки саме від нього залежить, які взаємозв'язки можуть бути виявлені в подальшому аналізі.

Не менш важливим є правильне часове впорядкування змінних. Якщо характеристика компанії використовується для пояснення дохідності, вона має бути зафіксована до моменту реалізації цієї дохідності. Інакше в аналіз потрапляє інформація, яка на момент прогнозування ще не була доступною, що призводить до спотворення результатів. Залежною змінною виступає дохідність акції в наступному періоді. У сформованому датасеті вона має технічну назву “*RET_LEAD*”. Емпірична постановка зіставляє характеристики компанії, зафіксовані в момент t , з дохідністю акції в періоді $t+1$. Такий підхід дає змогу перейти від загального трактування дохідності до питання про те, які саме підходи використовуються у фінансовій літературі для пояснення відмінностей у майбутній дохідності акцій.

Дохідність акцій у цьому дослідженні розглядається як результативна змінна, що відображає ринкову оцінку компанії та реакцію ринку на інформацію, доступну інвесторам. Її емпіричний аналіз потребує чіткого розмежування між очікуваною та реалізованою дохідністю, правильного вибору часового горизонту й коректного впорядкування змінних у часі. Це створює підґрунтя для розгляду теоретичних підходів, у межах яких фінансова література пояснює відмінності в дохідності акцій.

1.2 Еволюція підходів до пояснення дохідності акцій

Питання про те, що визначає дохідність акцій, посідає важливе місце у фінансовій літературі. Перші узагальнені підходи пов’язували відмінності в дохідності активів із загальним ринковим ризиком. У моделі CAPM *William F. Sharpe (1964)* очікувана надлишкова дохідність активу визначається через його чутливість до систематичного ринкового ризику, тобто через зв’язок між дохідністю активу та дохідністю ринкового портфеля [20, с. 439–442]. Відповідно до цієї логіки вищу очікувану дохідність мають активи, які сильніше реагують на коливання ринку.

У такій постановці інвестор отримує винагороду за прийняття ринкового ризику, який не може бути усунутий диверсифікацією. CAPM сформувала базову рамку для подальших досліджень дохідності акцій, оскільки запропонувала просте пояснення зв'язку між ризиком та очікуваним інвестиційним результатом. Її значення полягало в утвердженні ідеї, що дохідність акцій має пояснюватися систематичними чинниками ризику [20, с. 440–442].

Подальші емпіричні дослідження показали, що пояснення дохідності акцій лише через ринковий фактор не охоплює всіх відмінностей у середніх дохідностях. В працях *Eugene Fama and Kenneth French (1992)* продемонстровано, що характеристики, пов'язані з розміром компанії та співвідношенням book-to-market, мають вищу пояснювальну силу, ніж ринкова beta у її класичному трактуванні [6, с. 427–429]. Це змінило напрям аналізу: увага змістилася від одного ринкового параметра до відмінностей між самими компаніями.

Багатофакторний підхід закріпився у фінансовій літературі після появи трифакторної моделі *Eugene Fama and Kenneth French (1993)*. Ця модель поряд із ринковою премією враховує розмір компанії та співвідношення балансової і ринкової вартості [7, с. 5–7]. У такій постановці дохідність акцій пов'язується з рухом ринку та відмінностями між емітентами.

Розвиток багатофакторного підходу змінив спосіб постановки питання про дохідність акцій. Якщо в однофакторній моделі головну роль відігравав зв'язок активу з ринком, то в багатофакторній логіці важливими стали стійкі характеристики компаній. Це створило основу для емпіричних робіт, у яких дохідність акцій аналізується через поєднання ринкових факторів і характеристик емітентів.

Подальші дослідження розширили цю рамку. *Eugene F. Fama and Kenneth R. French (2015)* включили до п'ятифакторної моделі прибутковість та інвестиції [8, р. 2]. *Kewei Hou, Chen Xue and Lu Zhang (2015)* запропонували q-факторну модель, у якій дохідність пояснюється через ринковий фактор, розмір, інвестиції та прибутковість [14, с. 650]. Такі підходи показали, що пояснення stock returns дедалі більше спирається на характеристики компаній, які можна спостерігати в даних і перевіряти емпірично.

Поступовий перехід від ринкового фактора до характеристик компаній привів до ширших емпіричних робіт, у яких дослідники аналізують не одну змінну, а набори характеристик. Ці праці важливі для цієї роботи, оскільки вони показують, як у літературі формується зв'язок між *company characteristics* і майбутньою дохідністю акцій. Узагальнення трьох ключових досліджень подано в Таблиці 1.2.

Таблиця 1.2 Емпіричні дослідження характеристик компаній у поясненні майбутньої дохідності акцій

Дослідження	Дослідницький фокус	Методологія	Вибірка	Основні результати
Eugene F. Fama та Kenneth R. French (1992)	Аналіз факторів, що пояснюють перехресні відмінності в очікуванні дохідності акцій	Крос-секційні регресії Fama–MacBeth із включенням показників beta, size, book-to-market, leverage та earnings-price ratio	Акції NYSE, AMEX і NASDAQ за 1963 - 1990 рр.	Показники size та <i>book-to-market equity</i> спільно пояснюють значну частину variation у середній дохідності акцій. Водночас зв'язок між market beta та average return виявився слабким у традиційній CAPM-постановці.
Jonathan Lewellen (2014)	Оцінювання очікуваної дохідності акцій через поєднання численних характеристик компаній	Регресії Fama–MacBeth і побудова composite expected-return forecasts на основі firm characteristics	Акції США за 1964 - 2013 рр.	Характеристики компаній, зокрема <i>size</i> , <i>book-to-market</i> , <i>past returns</i> та <i>investment</i> , статистично пов'язані з майбутньою дохідністю акцій. Композитні прогнози демонструють суттєву predictive power щодо фактичних returns.
Jeremiah Green, John R. M. Handra X. та Frank Zhang (2017)	Визначення характеристик компаній, що містять незалежну інформацію про середню місячну дохідність акцій США	Одночасне включення 94 firm characteristics у регресії Fama–MacBeth із корекцією впливу microcaps і data-snooping bias	Акції США за 1980 - 2014 рр.	Лише обмежена кількість характеристик зберігає незалежну explanatory power після одночасної перевірки. Для non-microcap акцій predictive ability характеристик суттєво слабшає після 2003 року.

Джерело: складено автором на основі джерел [6;16;12]

Наведені в таблиці 1.2 праці показують, що емпірична література перейшла від перевірки окремих факторів до аналізу ширших наборів характеристик компаній. У роботі *Jonathan Lewellen (2014)* характеристики компаній використовуються разом для побудови прогнозів очікуваної дохідності, а не як окремі ізольовані показники [16, с. 2–3]. *Jeremiah Green, John R. M. Hand та X. Frank Zhang (2017)*

показують, що після одночасної перевірки великої кількості змінних незалежну пояснювальну силу зберігає обмежена частина характеристик [12, с. 1–2].

У вітчизняній літературі близькі питання розглядаються переважно через призму українського фондового ринку, макроекономічних чинників і прогнозування цін акцій. *Д. Митько (2025)* досліджує чинники довгострокової дохідності акцій на прикладі українського фондового ринку, *Є. Сова (2020)* аналізує взаємозв'язок між монетарними індикаторами та показниками фондового ринку, а *І. Волошина (2015)* розглядає можливості поєднання фундаментального й технічного аналізу для прогнозування цін українських акцій [17;21;23]. Ці праці близькі до теми цієї роботи через увагу до факторів ринкової динаміки акцій, однак відрізняються за ринком, рівнем аналізу та набором змінних.

Дані результати важливі для постановки подальшого дослідження. Вони показують, що характеристики компаній можуть містити інформацію про майбутню дохідність акцій. Сила такого зв'язку залежить від конкретної змінної, вибірки та специфікації моделі. Загальний висновок літератури не зводиться до твердження, що всі характеристики однаково важливі. Скоріше йдеться про потребу перевірити, які з них мають статистичний зв'язок із майбутньою дохідністю в конкретній емпіричній постановці.

Дослідницьке питання роботи стосується того, які характеристики компанії пов'язані з дохідністю її акцій у наступному періоді. Наступний підпункт присвячений працям, на основі яких формується набір характеристик для подальшої емпіричної перевірки.

1.3. Емпіричні дослідження зв'язку фінансових характеристик компаній із майбутньою дохідністю акцій

Після формулювання дослідницького питання необхідно визначити, які саме характеристики компанії доцільно включити до емпіричної перевірки. У фінансовій літературі такі характеристики охоплюють кілька груп показників:

ринкову оцінку компанії, прибутковість, інвестиційну активність, розмір, попередню ринкову динаміку та фінансову структуру. У межах цієї роботи ці групи представлені показниками прибутку до ціни (EP), балансової вартості до ринкової вартості (BM), валової прибутковості до активів (GPA), зростання активів (AG), логарифма ринкової капіталізації (LMC), шестимісячного моментуму ($MOM6$) та боргу до активів (DTA).

Одну з ранніх груп таких показників становлять характеристики ринкової оцінки. *Sanjoy Basu (1977)* досліджував зв'язок між коефіцієнтом ціни до прибутку (P/E) та інвестиційною результативністю звичайних акцій. У його дослідженні портфелі з нижчим значенням P/E демонстрували вищу абсолютну та скориговану на ризик дохідність порівняно з портфелями з вищим P/E [2, с. 667–668, 681].

У подальшій моделі показник прибутку до ціни (EP) використовується як характеристика співвідношення між прибутком компанії та її ринковою оцінкою. *Eugene Fama and Kenneth French (1992)* показали, що балансова вартість до ринкової вартості разом із розміром компанії пояснює відмінності в середніх дохідностях акцій, тоді як зв'язок ринкової β із середньою дохідністю є слабким у класичній постановці [6, с. 427–429]. В емпіричній частині роботи показник балансової вартості до ринкової вартості (BM) використовуватиметься як характеристика ринкової оцінки компанії відносно її бухгалтерської вартості.

Інший напрям пов'язаний із прибутковістю та інвестиційною активністю компаній. *Robert Novy-Marx (2013)* показав, що валова прибутковість, виміряна як відношення валового прибутку до активів, має приблизно таку саму силу в поясненні перехресного розрізу середніх дохідностей, як і співвідношення балансової та ринкової вартості, а також містить додаткову інформацію порівняно з показниками ринкової оцінки [18, с. 1–2, 7]. Цьому напрямку відповідає показник прибутковості валового прибутку до активів (GPA). *Sheridan Titman, K. C. John Wei and Feixue Xie (2004)* встановили, що компанії, які збільшують капітальні інвестиції, у наступні періоди демонструють від'ємну дохідність, скориговану відносно бенчмарку [22, с. 2, 13, 33]. Зростання активів (AG) використовується як характеристика інвестиційної активності компанії.

Окремий блок становлять характеристики, пов'язані з ринковим розміром, попередньою динамікою дохідності та фінансовою структурою. *Rolf Banz (1981)* показав, що менші компанії мали вищу скориговану на ризик дохідність, ніж більші компанії, а цей зв'язок був найсильнішим серед дуже малих фірм [1, с. 3–4, 12, 16]. Логарифм ринкової капіталізації (*LMC*) відображає цю групу результатів як показник розміру компанії. *Narasimhan Jegadeesh and Sheridan Titman (1993)* задокументували, що стратегії купівлі акцій минулих переможців і продажу акцій минулих аутсайдерів генерують позитивну дохідність на горизонтах 3–12 місяців [15, с. 65–66]. Шестимісячний моментум (*MOM6*) репрезентує цей напрям як показник попередньої ринкової динаміки. *Hitesh Doshi, Kris Jacobs, Praveen Kumar and Ramon Rabinovitch (2018)* аналізують роль фінансового левериджу у перехресному розрізі дохідностей акцій і показують, що леверидж змінює інтерпретацію закономірностей дохідності, зокрема закономірностей, пов'язаних із розміром компанії та ринковою оцінкою [5, с. 1–4]. Боргове навантаження представлено показником боргу до активів (*DTA*).

Узагальнення ключових емпіричних праць, які використовуються для обґрунтування вибору характеристик у подальшому аналізі, наведено в Таблиці 1.3.

Таблиця 1.3 Емпіричні праці, що обґрунтовують вибір характеристик компаній для подальшого аналізу

Дослідження	Група характеристик	Характеристика у фокусі	Методологія та вибірка	Використання в емпіричній моделі
Basu (1977)	Ринкова оцінка компанії	Коефіцієнт ціни до прибутку (P/E)	Формування портфелів за P/E; акції NYSE, квітень 1957 - березень 1971	Показує зв'язок між P/E та подальшою інвестиційною результативністю акцій; використовується для характеристики прибутку до ціни (EP)
Fama & French (1992)	Ринкова оцінка та розмір компанії	Балансова вартість до ринкової вартості; розмір компанії	Fama-MacBeth regressions; акції NYSE, AMEX і NASDAQ, 1963 - 1990	Показує роль балансової вартості до ринкової вартості та розміру компанії у поясненні відмінностей у середніх дохідностях; використовується для характеристики BM

Novy-Marx (2013)	Прибутковість компанії	Валова прибутковість до активів (GPA)	Fama-MacBeth regressions і портфельні сортування; акції США, липень 1963 - грудень 2010	Показує, що валова прибутковість до активів має пояснювальну силу щодо відмінностей у середніх дохідностях; використовується для характеристики GPA
Titman, Wei & Xie (2004)	Інвестиційна активність	Капітальні інвестиції; зростання активів як близький інвестиційний показник	Портфелі за abnormal capital investment; NYSE, AMEX і NASDAQ, липень 1973 - червень 1996	Показує зв'язок між зростанням капітальних інвестицій і нижчою подальшою дохідністю; використовується для характеристики зростання активів (AG)
Banz (1981)	Розмір компанії	Ринкова вартість компанії	Крос-секційний аналіз returns і market value; акції NYSE, 1936 - 1975	Показує зв'язок між розміром компанії та скоригованою на ризик дохідністю; використовується для характеристики логарифма ринкової капіталізації (LMC)
Jegadeesh & Titman (1993)	Попередня ринкова динаміка	Минулі дохідності акцій; momentum	Портфельні стратегії на основі минулих дохідностей; акції США, 1965-1989	Показує, що попередня динаміка дохідності може бути пов'язана з наступною дохідністю; використовується для характеристики шестимісячного моментуму (MOM6)
Doshi et al. (2018)	Фінансова структура	Фінансовий леверидж; боргове навантаження	Порівняння levered і unlevered equity returns; акції США, CRSP/Compustat 1963 - 2016	Показує, що боргове навантаження пов'язане з неоднорідною інтерпретацією закономірностей дохідності; використовується для характеристики боргу до активів (DTA)

Джерело: складено автором на основі джерел [2;6;18;22;1;15;5]

Таблиця 1.3 показує, що обрані характеристики охоплюють різні сторони діяльності компанії: її ринкову оцінку, прибутковість, інвестиційну активність, розмір, попередню ринкову динаміку та фінансову структуру. Цей набір змінних створює можливість переходу від загального дослідницького питання до емпіричної специфікації моделі, у межах якої перевіряється зв'язок між характеристиками компанії та дохідністю її акцій у наступному періоді.

Результати попередніх досліджень не означають, що всі зазначені характеристики мають однакову пояснювальну силу. Частина показників може

бути стійко пов'язана з майбутньою дохідністю, тоді як інші можуть втрачати значущість після включення додаткових змінних або зміни специфікації моделі. Подальша емпірична перевірка має оцінити наявність зв'язку між окремими характеристиками та дохідністю акцій і визначити, які з них зберігають значущість у межах спільної панельної моделі.

Отже, у подальшому аналізі до моделі включаються показники прибутку до ціни (*EP*), балансової вартості до ринкової вартості (*BM*), прибутковості валового прибутку до активів (*GPA*), зростання активів (*AG*), логарифма ринкової капіталізації (*LMC*), шестимісячного моментуму (*MOM6*) та боргу до активів (*DTA*). Наступний розділ присвячений опису даних, побудові цих змінних і вибору економетричної специфікації для перевірки їхнього зв'язку з дохідністю акцій у наступному періоді.

Висновки до Розділу 1

У першому розділі було визначено теоретичну основу дослідження дохідності акцій. Дохідність акцій розглянуто як результативну змінну, що відображає зміну вартості інвестиції, ринкову оцінку компанії та реакцію інвесторів на нову інформацію. Окрему увагу приділено розмежуванню очікуваної та реалізованої дохідності, оскільки емпірична частина роботи спирається на фактично спостережувану дохідність акцій у наступному періоді.

Завданням розділу було показати, як у фінансовій літературі сформувався перехід від пояснення дохідності через ринковий ризик до аналізу характеристик компаній. Для цього розглянуто розвиток підходів від CAPM до багатофакторних моделей Fama and French, а також до сучасних емпіричних досліджень, у яких характеристики компаній перевіряються як потенційні предиктори майбутньої дохідності акцій.

Проведений огляд літератури показав, що характеристики компаній можуть містити інформацію про майбутню дохідність, однак їхня пояснювальна сила не є однаковою. Частина показників зберігає статистичний зв'язок із дохідністю після

ширших перевірок, тоді як інші можуть бути чутливими до вибірки, періоду дослідження або специфікації моделі. Це формує дослідницьке питання роботи: які характеристики компаній мають значущий зв'язок із дохідністю акцій у наступному періоді.

Результатом розділу стало обґрунтування набору змінних для подальшої емпіричної перевірки. До моделі включено показники прибутку до ціни (*EP*), балансової вартості до ринкової вартості (*BM*), прибутковості валового прибутку до активів (*GPA*), зростання активів (*AG*), логарифма ринкової капіталізації (*LMC*), шестимісячного моментуму (*MOM6*) та боргу до активів (*DTA*). Така постановка відповідає логіці дослідження, у межах якої характеристики компанії в момент часу t зіставляються з дохідністю її акцій у наступному періоді $t+1$.

РОЗДІЛ 2

ФОРМУВАННЯ ДАНИХ ТА ОЦІНЮВАННЯ МОДЕЛІ ДОХІДНОСТІ АКЦІЙ

2.1. Формування інформаційної бази та вибірки дослідження

Інформаційну базу дослідження становлять панельні дані публічних компаній США за період 2015Q1–2025Q4. Вибір американського фондового ринку зумовлений кількома причинами. По-перше, це один із найбільш ліквідних і інформаційно насичених ринків капіталу, для якого доступний широкий набір фінансових і ринкових показників. По-друге, ринок США найчастіше використовується в емпіричних дослідженнях дохідності акцій, тому отримані результати можна співвідносити з попередніми зарубіжними працями. По-третє, велика кількість публічних емітентів і тривалий часовий горизонт спостережень дають змогу сформувати вибірку, придатну для панельного економетричного аналізу.

Джерелом даних виступають квартальна фінансова звітність компаній та ринкові показники, отримані з бази *Financial Modeling Prep* [9]. Для побудови датасету використовувалися дані про ринкову капіталізацію, ринкову ціну акцій, показники прибутковості, балансову вартість, активи, сукупний борг, власний капітал і динаміку котирувань. На основі цих даних розраховано залежну змінну та пояснювальні фінансові характеристики компаній. Кожне спостереження у вибірці відповідає парі «компанія–квартал», що формує панельну структуру даних.

Базова логіка формування спостережень будувалася за принципом *point-in-time*. Для кожної дати сортування використовувалася остання доступна квартальна звітність, для якої дата прийняття (*acceptedDate*) не перевищувала відповідну дату спостереження. Такий підхід важливий для уникнення *look-ahead bias*, тобто ситуації, коли в модель потрапляє інформація, яка на момент прийняття інвестиційного рішення ще не була доступною ринку. Інакше оцінки зв'язку між

фінансовими характеристиками та майбутньою дохідністю були б зміщеними в бік завищення пояснювальної сили моделі.

Часова структура вибірки побудована на квартальній частоті. Для кожної компанії аналізується набір фінансових характеристик на момент завершення кварталу, а далі оцінюється дохідність акцій у наступному кварталі. Така схема дозволяє поєднати два типи інформації: бухгалтерсько-фінансову, яка оновлюється зі звітністю, і ринкову, яка відображає реакцію цін акцій на інформацію про компанію та ринкове середовище. Квартальна частота є зручною для дослідження, оскільки вона не є надто грубою, як річні дані, але й не настільки шумною, як денні чи тижневі спостереження.

При формуванні залежної змінної використовувалася логіка “поточні характеристики компанії → дохідність у наступному періоді”, обґрунтована в першому розділі. Майбутня дохідність акцій визначалася як дохідність між датою сортування та завершенням наступного кварталу. У фінальному датасеті ця залежна змінна має технічну назву RET_LEAD. Через обмеження доступного набору даних цінова база будувалася на основі звичайної ціни закриття (*close*), а не скоригованої ціни (*adjClose*). Це було враховано під час підготовки вибірки та зафіксовано як особливість інформаційної бази дослідження.

Пояснювальні змінні формувалися з квартальних фінансових і ринкових показників компаній. До фінального набору увійшли EP, BM, GPA, AG, LMC, MOM6 і DTA. Їхній економічний зміст і спосіб розрахунку розкриваються в наступному підпункті.

На етапі очищення вибірки було вилучено спостереження, які не дозволяли коректно розрахувати ключові показники моделі, а також компанії з надто короткою історією квартальних даних. Фінальний варіант вибірки включає ті компанії, для яких доступно не менше 8 кварталів спостережень. Такий поріг було введено для того, щоб уникнути включення фірм із надто коротким часовим рядом, де оцінювання фінансових характеристик і їхнього зв'язку з майбутньою дохідністю є малозмістовним. Після застосування цього обмеження вибірка стала

більш однорідною з точки зору панельної структури й краще придатною для економетричного аналізу.

Фінальна вибірка містить 417 компаній і 17 234 квартальні спостереження за період 2015Q1–2025Q4. Панель є незбалансованою, оскільки тривалість спостереження за окремими компаніями відрізняється. Основні характеристики сформованої вибірки наведено в Таблиці 2.1.

Таблиця 2.1 Характеристика вибірки

Показник	Значення
Період спостереження	2015Q1–2025Q4
Тип панелі	незбалансована
Кількість компаній	417
Кількість спостережень	17 234
Мінімальна кількість кварталів на компанію	8
Кількість часових періодів	44

Джерело: складено автором на основі сформованої вибірки компаній США за даними *Financial Modeling Prep* [9]

Сформована база даних потребує подальшої перевірки перед економетричним оцінюванням моделі. Наступний підпункт присвячений формуванню змінних моделі та перевірці їхніх статистичних властивостей.

2.2. Побудова змінних моделі та перевірка емпіричної вибірки

Після формування інформаційної бази потрібно перевірити, чи придатний датасет для економетричного оцінювання, і уточнити спосіб побудови змінних моделі. У даному підпункті визначено залежну та пояснювальні змінні, подано порядок їх розрахунку, проаналізовано описову статистику та кореляції між предикторами. Окремо розглянуто екстремальні значення, які можуть впливати на результати оцінювання.

Залежна змінна моделі – RET_LEAD, що відображає дохідність акції у наступному кварталі відносно поточної дати спостереження. Її розрахунок здійснюється як відношення ціни закриття на кінець наступного кварталу до ціни закриття на дату сортування мінус одиниця:

$$\text{RET LEAD}_{i,t+1} = \frac{\text{Close}_{i,t+1}}{\text{Close}_{i,t}} - 1 \quad (2.1)$$

де $\text{Close}_{i,t}$ – ціна закриття акції компанії (i) у момент формування спостереження, $\text{Close}_{i,t+1}$ – ціна закриття цієї самої акції наприкінці наступного кварталу.

RET_LEAD використовується як залежна змінна моделі та фіксує майбутній результат з яким зіставляються характеристики компанії в поточному періоді.

До набору пояснювальних змінних включено сім показників: EP, BM, GPA, AG, LMC, MOM6 і DTA. Їхню міжперіодну структуру подано в Таблиці 2.2.1.

Таблиця 2.2.1. Характеристика змінних і міжперіодна структура моделі

Показник	Характеристика	Час у моделі	Що відображає
RET_LEAD	Дохідність акцій у наступному кварталі	t+1	Залежну змінну моделі
EP	Співвідношення прибутку до ціни акції	t	Ринкову оцінку компанії
BM	Співвідношення балансової та ринкової вартості	t	Вартісну характеристику компанії
GPA	Валова прибутковість активів	t	Операційну ефективність
AG	Темп зростання активів	t	Інвестиційну активність
LMC	Логарифм ринкової капіталізації	t	Розмір компанії
MOM6	Шестимісячний моментум	t	Попередню ринкову динаміку
DTA	Співвідношення сукупного боргу до сукупних активів	t	Частку активів, профінансованих боргом

Джерело: складено автором на основі емпіричної бази дослідження

Таблиця 2.2.1 подає загальну структуру змінних моделі та їхнє часове розміщення. Коректність подальшого оцінювання залежить від того, як побудовано кожен пояснювальну змінну, тому далі розглянуто порядок їх розрахунку та економічне трактування.

Показник *EP* у датасеті задається через *earnings yield*, він доступний у джерелі даних. Він відображає співвідношення прибутку на акцію до ринкової ціни акції. За економічним змістом *EP* дає змогу оцінити, наскільки високо або низько ринок оцінює компанію відносно її прибуткового потенціалу. У моделі *EP* використовується як один із показників ринкової оцінки компанії.

Змінна BM розраховується як відношення балансової вартості власного капіталу до ринкової капіталізації компанії на дату сортування. Ця змінна відображає співвідношення між бухгалтерською оцінкою компанії та її ринковою вартістю. Високі значення BM характерні для компаній, ринкова оцінка яких є нижчою порівняно з балансовою вартістю, тоді як низькі значення характерні для компаній із вищою ринковою премією.

Показник GPA визначається як відношення валового прибутку до сукупних активів. У дослідженні цей показник використовується як індикатор прибутковості активів компанії. Його зміст полягає в оцінці того, наскільки ефективно компанія генерує валовий прибуток на одиницю залучених ресурсів. На відміну від чистого прибутку, валовий прибуток меншою мірою залежить від окремих облікових рішень, пов'язаних із витратами нижчих рівнів, що робить GPA зручним показником для порівняння компаній у межах великої панельної вибірки.

Показник AG відображає темп зростання активів компанії й розраховується як відношення поточного обсягу активів до їхнього значення чотири квартали тому мінус одиниця:

$$AG_{i,t} = \frac{Total\ Assets_{i,t}}{Total\ Assets_{i,t-4}} - 1 \quad (2.2)$$

де $Total\ Assets_{i,t}$ – сукупні активи компанії (i) у поточному кварталі (t),

$Total\ Assets_{i,t-4}$ – сукупні активи тієї самої компанії чотири квартали тому.

Показник відображає річний темп зростання активів, розрахований на кварталних даних.

AG використовується як проксі інвестиційної активності компанії. Високі значення показника відповідають періодам інтенсивного розширення, а від'ємні значення свідчать про скорочення активів.

Показник LMC задається як натуральний логарифм ринкової капіталізації компанії на дату сортування. Використання логарифмічної форми дає можливість зменшити масштабну неоднорідність між великими та малими компаніями й

зробити розподіл змінної більш зручним для регресійного оцінювання. У моделі LMC виступає індикатором розміру компанії.

Показник *MOM6* відображає шестимісячний моментум, що обчислюється як дохідність між кінцем сьомого та кінцем першого місяця перед датою сортування:

$$MOM6_{i,t} = \frac{Price_{i,t-1}}{Price_{i,t-7}} - 1 \quad (2.3)$$

де $Price_{i,t-1}$ – ціна акції компанії (*i*) на кінець місяця, що передує даті сортування, а $Price_{i,t-7}$ – ціна цієї самої акції на кінець сьомого місяця перед датою сортування.

Під час побудови *MOM6* найближчий місяць перед датою сортування виключається, щоб зменшити вплив короткострокового ефекту reversal. Цей показник використовується для врахування попередньої ринкової динаміки акцій компанії перед оцінюванням її майбутньої квартальної дохідності.

Показник *DTA* визначається як відношення сукупного боргу до сукупних активів компанії:

$$DTA_{i,t} = \frac{Total\ Debt_{i,t}}{Total\ Assets_{i,t}} \quad (2.4)$$

Ця змінна характеризує частку активів компанії, профінансованих за рахунок боргових зобов'язань. Її економічний зміст пов'язаний з оцінкою ролі позикового фінансування у структурі ресурсів компанії та загального рівня фінансового левереджу. *DTA* використовується як базовий індикатор боргового навантаження, придатний для порівняння компаній у межах великої панельної вибірки.

Після формування всіх змінних датасет було перевірено перед економетричним оцінюванням. Першим кроком став аналіз описової статистики змінних моделі, що дає змогу оцінити центральні тенденції, розсіювання та межі варіації змінних. Основні результати наведено в Таблиці 2.2.2.

Таблиця 2.2.2 Описова статистика змінних моделі

Показник	ret_lead	ep	bm	gra	ag	lmc	mom6	dta
Mean	0,042438	0,010154	0,342822	0,063171	0,115492	24,229984	0,07804	0,314932
Median	0,033722	0,01048	0,250723	0,051049	0,056705	24,200155	0,05622	0,30354
Max	3,779176	0,062746	1,564208	0,217259	1,428265	27,520312	0,885395	0,988343
Min	-0,830762	-0,078409	-0,119962	0	-0,229292	21,035692	-0,373749	0
Std. Dev.	0,180133	0,017198	0,324908	0,047496	0,245383	1,200004	0,216162	0,20003

Джерело: створено автором за допомогою Eviews на основі самостійно сформованого масиву даних

Дані Таблиці 2.2.2 показують, що залежна змінна RET_LEAD має середнє значення 0,042438 тобто в середньому квартальна дохідність акцій у вибірці є додатною. Значна варіація дохідності узгоджується з природою фондового ринку. Серед пояснювальних змінних найбільший розмах значень зберігається для AG, що вказує на високу чутливість цього показника до змін активів компанії. Для LMC розподіл є більш стабільним, що пояснюється використанням логарифмічної трансформації. Показник GRA має нижню межу 0, а боргова змінна DTA характеризується додатним і економічно зрозумілим діапазоном варіації: від 0 до 0,988343. Такий діапазон підтверджує коректність DTA як показника левереджу для подальшого економетричного оцінювання.

Наступним етапом перевірки стала побудова кореляційної матриці пояснювальних змінних, яка дозволяє попередньо оцінити характер взаємозв'язку між ними та виявити потенційні ознаки мультиколінеарності. Результати наведено в Таблиці 2.2.3.

Таблиця 2.2.3 Кореляційна матриця пояснювальних змінних

Показник	ep	bm	gra	ag	lmc	mom6	dta
ep	1	0,140875	0,04657	-0,092901	0,13612	0,023785	-0,089312
bm	0,140875	1	-0,532161	-0,065014	-0,166561	-0,207511	-0,291821
gra	0,04657	-0,532161	1	0,023895	0,073081	0,088633	0,044289
ag	-0,092901	-0,065014	0,023895	1	-0,083415	0,084964	-0,066883
lmc	0,13612	-0,166561	0,073081	-0,083415	1	0,03881	-0,004165
mom6	0,023785	-0,207511	0,088633	0,084964	0,03881	1	-0,013974
dta	-0,089312	-0,291821	0,044289	-0,066883	-0,004165	-0,013974	1

Джерело: створено автором за допомогою Eviews на основі самостійно сформованого масиву даних

Кореляційна матриця свідчить про відсутність критично високих парних кореляцій між пояснювальними змінними. Найбільший за модулем зв'язок спостерігається між BM та GPA (-0,532161), що вказує на помірний взаємозв'язок між цими показниками, але не свідчить про критичну мультиколінеарність. Показник DTA має найпомітніший зв'язок із BM (-0,291821), однак і це значення не є достатнім, щоб вважати мультиколінеарність визначальною проблемою моделі. Для більшості інших пар значення кореляції залишаються низькими або помірними. Отже, сильного лінійного дублювання інформації між предикторами не виявлено.

Окрему увагу на етапі перевірки датасету було приділено екстремальним значенням змінних. Первинний аналіз показав, що частина предикторів, зокрема EP, BM, GPA, AG, MOM6 і DTA, має довгі хвости або поодинокі великі спостереження. Такі значення можуть впливати на оцінки параметрів і робити результати моделі надто чутливими до невеликої кількості спостережень. Для зменшення цього впливу до пояснювальних змінних EP, BM, GPA, AG, LMC, MOM6 і DTA застосовано вінсоризацію на рівнях 1% і 99%. Залежна змінна RET_LEAD не вінсоризувалася, оскільки її варіація є безпосереднім об'єктом дослідження.

Після обробки екстремальних значень датасет зберіг достатню кількість спостережень і залишився придатним для панельного оцінювання. Для AG розподіл залишився складнішим порівняно з іншими предикторами, що пов'язано з чутливістю цього показника до різких змін активів компанії.

Проведена перевірка показала, що змінні моделі сформовано коректно, критичної мультиколінеарності між предикторами не виявлено, а вплив екстремальних значень зменшено через вінсоризацію. Це дозволяє перейти до специфікації економетричної моделі та оцінювання зв'язку між характеристиками компаній і дохідністю їхніх акцій у наступному періоді.

2.3. Економетрична специфікація та оцінювання моделі

Після формування змінних і перевірки датасету дослідження переходить до специфікації економетричної моделі. У цьому підпункті оцінюється зв'язок між фінансовими характеристиками компаній і дохідністю їхніх акцій у наступному періоді. Залежною змінною виступає RET_LEAD , яка відображає дохідність акції у кварталі, наступному за датою спостереження. Пояснювальними змінними є EP , BM , GPA , AG , LMC , $MOM6$ і DTA .

Базова економетрична специфікація моделі має такий вигляд:

$$RET_LEAD_{i,t+1} = \beta_0 + \beta_1 EP_{i,t} + \beta_2 BM_{i,t} + \beta_3 GPA_{i,t} + \beta_4 AG_{i,t} + \beta_5 LMC_{i,t} + \beta_6 MOM6_{i,t} + \beta_7 DTA_{i,t} + \varepsilon_{i,t} \quad (2.5)$$

де (i) позначає компанію, t – квартал спостереження, β_0 – константу моделі, β_1 , β_2 , β_3 , β_4 , β_5 , β_6 , β_7 – коефіцієнти при відповідних пояснювальних змінних, що відображають силу та напрям зв'язку між фінансовими характеристиками компанії та дохідністю її акцій у наступному періоді, $\varepsilon_{i,t}$ – випадкову похибку.

Для оцінювання цієї специфікації в роботі використовуються дві моделі. Перша – *Pooled OLS*, яка розглядає всі спостереження панелі як єдину сукупність і оцінює середній лінійний зв'язок між характеристиками компаній та майбутньою дохідністю акцій. Друга – модель з часовими фіксованими ефектами (*Period Fixed Effects*), яка використовується як перевірка робастності та враховує вплив факторів, спільних для всіх компаній у певному кварталі. До таких факторів можуть належати зміни ринкової кон'юнктури, процентних ставок, макроекономічного фону, інфляційних очікувань або загального інвестиційного настрою.

Оцінювання проводилося на сформованій незбалансованій панелі. Для обох моделей застосовано робастні стандартні похибки з кластеризацією за компаніями. Це враховує можливу неоднорідність дисперсії залишків і залежність похибок усередині окремих компаній у часі, що є важливим для коректного оцінювання панельних даних. Використання кластеризованих стандартних похибок

узгоджується з *Petersen (2009)*, який показує, що у фінансових панельних даних звичайні стандартні похибки OLS можуть бути зміщеними за наявності залежності залишків у межах компаній, тоді як кластеризація стандартних похибок є коректнішим способом оцінювання статистичної значущості. Результати оцінювання базової моделі Pooled OLS наведено в Таблиці 2.3.1.

Таблиця 2.3.1 Результати оцінювання моделі Pooled OLS

Змінна	Коефіцієнт	Std. Error	t-Statistic	Prob.
C	0,40752	0,043912	9,280414	0,0000
EP	-0,677038	0,106257	-6,371731	0,0000
BM	0,017626	0,006283	2,805344	0,0050
GPA	0,127998	0,038156	3,354593	0,0008
AG	-0,000519	0,007661	-0,067692	0,9460
LMC	-0,01559	0,001756	-8,878527	0,0000
MOM6	0,028810	0,008406	3,427414	0,0006
DTA	0,010266	0,008878	1,156311	0,2476

Джерело: створено автором за допомогою Eviews

За результатами Pooled OLS статистично значущий зв'язок із майбутньою дохідністю акцій мають EP, BM, GPA, LMC і MOM6. Коефіцієнт при EP є від'ємним і статистично значущим (-0,677038; p-value = 0,0000), що вказує на від'ємний зв'язок між *earnings yield* та дохідністю акцій у наступному кварталі. Такий знак відрізняється від класичної логіки вартісної премії, однак не є унікальним для емпіричних досліджень останнього десятиліття. Blitz (2020) показав, що у 2010–2019 рр. earnings-to-price та інші value-показники мали від'ємну премію [3, с. 2–4]. Тому від'ємний знак EP у цій роботі не трактується як технічна помилка, а розглядається як результат, можливий для сформованої вибірки та періоду дослідження. Коефіцієнт при LMC також є від'ємним і статистично значущим (-0,015590; p-value = 0,0000), що вказує на зв'язок між розміром компанії та її майбутньою дохідністю. Змінні BM, GPA і MOM6 мають додатні коефіцієнти та залишаються статистично значущими. Для AG та DTA статистично значущого зв'язку в моделі Pooled OLS не виявлено.

Для перевірки того, наскільки ці результати зберігаються після врахування спільних часових шоків, модель було додатково оцінено у специфікації з часовими фіксованими ефектами. Результати такого оцінювання наведено в Таблиці 2.3.2.

Таблиця 2.3.2 Результати оцінювання моделі Period Fixed Effects

Змінна	Коефіцієнт	Std. Error	t-Statistic	Prob.
EP	-0,530894	0,09552	-5,557926	0,0000
BM	-0,000283	0,005527	-0,051266	0,9591
GPA	0,098652	0,034139	2,889741	0,0039
AG	0,001381	0,006807	0,202920	0,8392
LMC	-0,016669	0,001858	-8,970507	0,0000
MOM6	0,031432	0,00888	3,539542	0,0004
DTA	-0,004216	0,008243	-0,511466	0,6090

Джерело: створено автором за допомогою Eviews

Модель Period Fixed Effects використовується як перевірка робастності базової специфікації. Після врахування часових ефектів статистично значущими залишаються EP, GPA, LMC і MOM6 (p -value = 0,0000, 0,0039, 0,0000 і 0,0004 відповідно). Для BM, AG і DTA стійкого статистично значущого ефекту не виявлено (p -value = 0,9591, 0,8392 і 0,6090). Порівняно з Pooled OLS, головна зміна стосується BM, її зв'язок із майбутньою дохідністю зникає після контролю за спільними квартальними шоками. BM виглядає чутливою до зміни специфікації, тоді як EP, GPA, LMC і MOM6 зберігають значущість в обох моделях.

Після оцінювання двох моделей наведено їхні основні показники якості. Вищий R^2 у специфікації з часовими фіксованими ефектами потребує обережного трактування, оскільки частина пояснювальної сили пов'язана з включенням часових ефектів. У цій роботі Period Fixed Effects використовується як перевірка робастності, тому таблиця не слугує автоматичним критерієм вибору моделі з вищим R^2 .

Таблиця 2.3.3 Показники якості базової моделі та специфікації з часовими фіксованими ефектами

Показник якості моделі	Pooled OLS	Period Fixed Effects
Коефіцієнт детермінації (R^2)	0,018229	0,230696
Скоригований коефіцієнт детермінації (Adjusted R^2)	0,01783	0,228457
F-статистика	45,6910 (0,0000)	103,0553 (0,0000)
Статистика Дарбіна–Вотсона	2,073013	2,003818

Джерело: створено автором за допомогою Eviews

Таблиця 2.3.3 фіксує різницю в основних показниках якості між базовою моделлю та специфікацією з часовими фіксованими ефектами. Модель Period FE має вищий коефіцієнт детермінації, ніж *Pooled OLS* (0,230696 проти 0,018229). Така різниця очікувана, оскільки часові ефекти поглинають спільні квартальні коливання дохідності. Через це вищий R^2 не трактується як самостійне покращення моделі. У межах однієї базової формули модель оцінюється у двох специфікаціях: Pooled OLS та Period Fixed Effects. Друга специфікація враховує спільні квартальні шоки й уможливорює порівняння результатів після додавання часових ефектів. Обидві моделі загалом є статистично значущими, що підтверджується значеннями F-статистики (45,6910 і 103,0553) та відповідними p-value (0,0000 в обох випадках). Статистика Дарбіна–Вотсона в обох моделях є близькою до 2 (2,073013 для Pooled OLS і 2,003818 для Period FE). Водночас у панельних фінансових даних цей показник не є основним критерієм коректності стандартних похибок, тому в роботі використано кластеризацію стандартних похибок за компаніями.

Для оцінки коректності специфікацій в роботі проведено додаткові діагностичні тести. Їх результати наведено в Таблиці 2.3.4.

Таблиця 2.3.4 Результати діагностичних тестів для моделей Pooled OLS та Period Fixed Effects

Показник перевірки специфікації	Pooled OLS	Period Fixed Effects
Тест Песарана на крос-секційну залежність (Pesaran CD)	499,174 (0,0000)	13,052 (0,0000)
Тест Бройша–Пагана на крос-секційну залежність (Breusch–Pagan LM)	404 911,639 (0,0000)	186 212,511 (0,0000)

Тест Рамсея на правильність функціональної форми (Ramsey RESET)	41,046 (0,0000)	181,507 (0,0000)
Максимальне значення VIF	1,748	1,748
Тест Бройша–Годфрі/Вулдріджа на автокореляцію в панельних даних	174,401 (0,0000)	66,547 (0,0000)
Кількість компаній-кластерів	417	417
Тест на надлишкові часові фіксовані ефекти (Period F)	—	110,363 (0,0000)
Тест на надлишкові часові фіксовані ефекти (Period Chi-square)	—	153 456,443 (0,0000)

Джерело: створено автором за допомогою Eviews

Діагностичні тести фіксують наявність крос-секційної залежності у вибірці. Для *Pooled OLS* значення *Pesaran CD* становить 499,174, а *Breusch–Pagan LM* – 404 911,639; для *Period Fixed Effects* відповідні значення дорівнюють 13,052 і 186 212,511. В усіх випадках $p\text{-value} = 0,0000$. Для вибірки публічних компаній така залежність є очікуваною, оскільки компанії можуть одночасно реагувати на спільні ринкові та макроекономічні події. Тому в роботі додатково використано модель із часовими фіксованими ефектами, яка відокремлює спільні квартальні умови. Кластеризація стандартних похибок за компаніями застосована для врахування можливої залежності похибок у межах однієї компанії в часі.

Тест Ramsey RESET для обох моделей є статистично значущим, 41,046 для *Pooled OLS* і 181,507 для *Period FE*, в обох випадках $p\text{-value} = 0,0000$. Лінійна форма специфікації моделі не повністю відображає структуру зв'язків у даних. Можливими причинами є нелінійності, пропущені змінні або складніша форма залежності між фінансовими характеристиками компаній і майбутньою дохідністю їхніх акцій.

Значення VIF залишаються низькими в обох специфікаціях, причому максимальне значення становить 1,748. Значить, критичної мультиколінеарності між пояснювальними змінними у вибірці не спостерігається. Найбільш помітний зв'язок між змінними *BM* і *GPA* недостатній для того, щоб вважати мультиколінеарність визначальною проблемою моделі.

Тест Бройша–Годфрі-Вулдріджа на автокореляцію в панельних даних також є статистично значущим для обох моделей: 174,401 для *Pooled OLS* і 66,547 для *Period FE*, в обох випадках $p\text{-value} = 0,0000$. Це вказує на наявність серійної

залежності залишків у межах панельної структури даних. Такий результат узгоджується з аргументацією *Petersen (2009)*, який зазначає, що у фінансових панельних даних залишки можуть бути пов'язаними в межах однієї компанії в часі, а звичайні стандартні похибки OLS у такому випадку можуть бути зміщеними [19, с. 435–437, 440, 475]. Тому використання робастних стандартних похибок із кластеризацією за компаніями є обґрунтованим.

Щодо тестів на надлишкові часові фіксовані ефекти, є статистично значущими: *Period F* = 110,363, *p-value* = 0,0000; *Period Chi-square* = 153 456,443, *p-value* = 0,0000. Це може вказувати на наявність спільних квартальних ефектів, які впливають на дохідність компаній у вибірці. Модель *Period Fixed Effects* використовується як перевірка робастності базових результатів *Pooled OLS* і не розглядається як заміна базової специфікації.

За результатами порівняння двох моделей слід підсумувати, що найбільш стійкими пояснювальними змінними в обох специфікаціях виявилися EP, GPA, LMC і MOM6. Змінна BM зберігає значущість в *Pooled OLS*, але втрачає її після врахування часових фіксованих ефектів, що вказує на нестійкий характер цього зв'язку. Для AG та DTA підтвердження стійкого статистично значущого зв'язку не отримано. Це означає, що за результатами порівняння двох специфікацій найпереконливіший зв'язок із майбутньою дохідністю акцій демонструють показники ринкової оцінки, прибутковості, розміру компанії та попередньої ринкової динаміки. Отримані результати у наступному розділі розглядатимуться детальніше через додаткові аналітичні перевірки: одновимірні моделі, квантильне групування, вкладені специфікації та сегментацію вибірки за окремими характеристиками компаній. Такі перевірки потрібні для пояснення стійкості або нестійкості окремих коефіцієнтів.

Висновки до Розділу 2

У другому розділі сформовано емпіричну базу дослідження та підготовлено датасет для оцінювання зв'язку між фінансовими характеристиками компаній і дохідністю їхніх акцій у наступному періоді. Фінальна вибірка охоплює 417

публічних компаній США, 17 234 квартальні спостереження та 44 часові періоди за 2015Q1–2025Q4. Вибірка є незбалансованою, а до її фінального складу включено лише компанії, для яких доступно не менше 8 кварталів спостережень.

У межах підготовки датасету визначено залежну змінну `RET_LEAD`, сформовано набір пояснювальних змінних `EP`, `BM`, `GPA`, `AG`, `LMC`, `MOM6` і `DTA`, а також уточнено порядок їх розрахунку. Описова статистика показала, що найбільш складний розподіл серед предикторів має змінна `AG`, тоді як `DTA` характеризується додатним і стабільнішим діапазоном значень. Кореляційна матриця засвідчила відсутність критично високих парних кореляцій між пояснювальними змінними, тому мультиколінеарності не має.

Для зменшення впливу екстремальних значень на результати оцінювання до незалежних змінних `EP`, `BM`, `GPA`, `AG`, `LMC`, `MOM6` і `DTA` застосовано *winsorization* на рівнях 1% і 99%, залежна змінна `RET_LEAD` не вінсоризувалася, бо її варіація є об'єктом дослідження. Після цього оцінено дві специфікації моделі – *Pooled OLS* та *Period Fixed Effects*, для обох моделей використано робастні стандартні похибки з кластеризацією за компаніями.

Результати показали, обох моделях найбільш стійкий статистично значущий зв'язок з майбутньою дохідністю акцій демонструють `EP`, `GPA`, `LMC` і `MOM6`. Причому від'ємний коефіцієнт при `EP` суперечить класичній логіці вартісної премії, але для періоду 2015–2025 рр. на ринку США це може бути пов'язано зі слабшою динамікою value-сегмента та сильнішими позиціями growth-компаній. `BM` є значущою в *Pooled OLS*, але втрачає значущість після врахування часових фіксованих ефектів. Для `AG` та `DTA` стійкого статистично значущого зв'язку не отримано.

Діагностичні тести підтвердили наявність крос-секційної залежності, автокореляції в панельних даних і складнішої структури зв'язків у даних, ніж передбачає проста лінійна форма. Значення *VIF* залишаються низькими, тому критичної мультиколінеарності між предикторами не виявлено. Це обґрунтовує використання часових фіксованих ефектів і робастних стандартних похибок із кластеризацією за компаніями. Дві розглянуті специфікації, виконують різні

функції: *Pooled OLS* використовується як базова модель, *Period Fixed Effects* – для перевірки стійкості результатів після врахування спільних часових шоків. Вищий R^2 у *Period Fixed Effects* трактується обережно, оскільки частина пояснювальної сили пов'язана з включенням часових ефектів. Отримані результати створюють основу для подальшого аналізу стійких і нестійких коефіцієнтів у третьому розділі.

РОЗДІЛ 3

АНАЛІЗ СТІЙКОСТІ ТА ІНТЕРПРЕТАЦІЯ РЕЗУЛЬТАТІВ ЕМПІРИЧНОЇ МОДЕЛІ

3.1 Аналіз причин отриманих результатів

Результати другого розділу встановили, які фінансові характеристики компаній виявилися статистично значущими в базовій моделі, а які не продемонстрували стійкого зв'язку. Проте самі по собі оцінки коефіцієнтів і *p-value* ще не пояснюють, чому одні змінні зберігають статистичну силу в різних специфікаціях, тоді як інші втрачають її після врахування часових ефектів або зміни складу предикторів. З цієї причини в даному підпункті увагу зосереджено на змістовному поясненні поведінки окремих змінних вже в уточненій емпіричній постановці.

За результатами другого розділу найбільш стійкий зв'язок із майбутньою дохідністю акцій продемонстрували EP, GPA, LMC і MOM6. Показник BM виявився чутливим до врахування часових ефектів, тоді як AG і DTA не дали переконливого статистичного сигналу. Для пояснення такої відмінності між змінними проведено додатковий аналіз на основі фінального датасету. Перед виконанням додаткових перевірок центральну модель було відтворено в R та зіставлено з результатами, отриманими в EViews. Основні висновки щодо знаків, статистичної значущості та складу стійких змінних збіглися, що дало підстави використовувати R для подальшого аналізу. Такий перехід пов'язаний із технічними обмеженнями студентської версії EViews щодо автоматизованого групування спостережень, повторного оцінювання допоміжних специфікацій і зведення результатів. У межах додаткового аналізу досліджено три аспекти: індивідуальний сигнал змінних, кватильну структуру майбутньої дохідності та чутливість коефіцієнтів до спрощення специфікації моделі. Технічну реалізацію розрахунків наведено в Додатку Б.

Спочатку перевіряється, чи має кожна пояснювальна змінна власний окремий зв'язок із RET_LEAD без урахування інших фінансових контролів. Для цього, для

кожної змінної було обчислено парну кореляцію з майбутньою дохідністю, після чого оцінено дві одновимірні моделі – Pooled OLS та модель з часовими фіксованими ефектами, для того щоб побачити, чи є в змінної самостійний інформаційний сигнал і чи не зникає він після врахування спільних квартальних шоків. Результати наведено в Таблиці 3.1.1.

Таблиця 3.1.1 Індивідуальний сигнал пояснювальних змінних щодо RET_LEAD

Змінна	Парна кореляція з RET_LEAD	Одновимірна pooled-модель $RET_LEAD_{i,t+1} = \beta_0 + \beta_1 X_{i,t} + \varepsilon_{i,t}$			Одновимірна модель з часовими ефектами $RET_LEAD_{i,t+1} = \beta_0 + \beta_1 X_{i,t} + \lambda_t + \varepsilon_{i,t}$		
		Коефіцієнт	p-value	R ²	Коефіцієнт	p-value	R ²
EP	-0,072858	-0,763104	0,0000	0,005308	-0,665549	0,0000	0,217851
BM	0,01157	0,006414	0,1288	0,000134	-0,006761	0,0738	0,214002
GPA	0,009783	0,037102	0,1991	0,000096	0,069434	0,0069	0,214189
AG	0,014877	0,010921	0,0508	0,000221	0,014272	0,0043	0,214229
LMC	-0,114135	-0,017133	0,0000	0,013027	-0,017361	0,0000	0,226229
MOM6	0,025179	0,020982	0,0009	0,000634	0,032116	0,0000	0,215152
DTA	0,009387	0,008453	0,2179	0,000088	0,002307	0,7060	0,213862

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

Таблиця 3.1.1 показує, що найвиразніший окремий сигнал мають EP і LMC. Їхній зв'язок із RET_LEAD простежується в парній кореляції та в обох одновимірних моделях. Означає, що їхній зв'язок не виникає штучно всередині повної регресії, а спостерігається вже на базовому рівні.

GPA і MOM6 дають слабший, але все ж помітний сигнал. Для GPA проста кореляція майже нічого не показує, проте після врахування часових ефектів коефіцієнт стає значущим. Це вказує на те, що роль прибутковості краще проявляється після контролю спільних квартальних коливань. MOM6 поводить інакше, проста кореляція невелика, але в обох одновимірних моделях змінна лишається значущою. Тож, попередня ринкова динаміка теж має самостійний зміст.

BM вже на даному етапі виглядає нестійкою. Виразного окремого зв'язку з RET_LEAD вона не показує, тому її роль у повній моделі, ймовірно, залежить від сусідніх предикторів. AG дає лише слабкий сигнал. DTA виглядає ще слабше,

зв'язок із майбутньою дохідністю майже не читається ні в кореляції, ні в простих регресіях.

Окремого пояснення потребує показник EP. Його від'ємний знак зберігається і в одновимірних моделях, тому такий результат не можна пояснити лише впливом повної специфікації моделі. Він відрізняється від класичної логіки вартісної премії, однак не є унікальним для емпіричних досліджень останнього десятиліття. Blitz (2020) показав, що у 2010–2019 рр. премія за earnings-to-price була від'ємною (-0,72), як і премії за інші value-показники [3, с. 2–4]. Тому від'ємний зв'язок між EP і майбутньою дохідністю у цій роботі не трактується як технічна помилка, а розглядається як результат, можливий для сформованої вибірки та періоду дослідження.

Цього ще недостатньо, щоб зрозуміти, як змінна поводить себе в самих даних. Тому далі слід перевірити, чи змінюється майбутня дохідність акцій при переході від нижчих значень показника до вищих. Для цього всі спостереження типу «компанія–квартал» було поділено на квартилі за значенням кожної змінної, після чого для кожної групи обчислено середнє значення RET_LEAD. Таблиця 3.1.2 допомагає простежити, чи результат проявляється тільки як окремий регресійний ефект.

Таблиця 3.1.2 Середня майбутня дохідність акцій за квартилями пояснювальних змінних

Змінна	Q1	Q2	Q3	Q4	Q4 – Q1
EP	0,060892	0,037808	0,035214	0,035835	-0,025058
BM	0,049742	0,044246	0,033224	0,042538	-0,007204
GPA	0,0395	0,039753	0,044296	0,046203	0,006703
AG	0,043239	0,035026	0,036825	0,054658	0,011418
LMC	0,078312	0,03816	0,027822	0,025453	-0,052859
MOM6	0,050576	0,035168	0,032911	0,051093	0,000518
DTA	0,046414	0,039665	0,035958	0,047713	0,001299

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

У Таблиці 3.1.2 Q1 відповідає спостереженням із найнижчими значеннями змінної, а Q4 - із найвищими. Різниця Q4 - Q1 фіксує розрив між крайніми групами. Послідовна зміна середньої дохідності від квартиля до квартиля свідчить про більш

стійкий зв'язок, тоді як нерівномірний профіль вказує на слабший або нестабільний ефект.

Тож, для EP спостерігається помітний спад середньої майбутньої дохідності: від 0,060892 у першому квартилі до 0,035835 у четвертому. Різниця між крайніми групами становить -0,025058. Отже, від'ємний зв'язок, зафіксований у регресійних моделях, видно і в самому розподілі даних.

Ще чіткіше, ця закономірність проявляється для LMC. Середня майбутня дохідність знижується від 0,078312 у першому квартилі до 0,025453 у четвертому, а розрив $Q4 - Q1$ дорівнює -0,052859. Це найбільше абсолютне відхилення серед усіх змінних у таблиці. Зв'язок розміру компанії з майбутньою дохідністю в цій вибірці виглядає найбільш послідовним. Менші компанії в середньому мали вищу майбутню дохідність, ніж більші.

Для GPA картина слабша, але напрям не суперечить результатам регресій. Середня дохідність зростає від 0,039500 у першому квартилі до 0,046203 у четвертому, а різниця між крайніми групами становить 0,006703. Це не дуже великий розрив, проте він узгоджується з додатним знаком коефіцієнта в моделі.

Для BM послідовного градієнта немає. Різниця між крайніми квартилями невелика (-0,007204), а проміжні значення розташовані нерівномірно. Це добре узгоджується з попереднім висновком про те, що BM не має сильного самостійного ефекту й поводить себе нестійко.

МOM6 теж не дає чіткого квартильного профілю. Попри значущість у регресіях, розрив між $Q4$ і $Q1$ становить лише 0,000518. Отже, її сигнал слабо читається в простому групуванні спостережень і краще проявляється всередині регресійної моделі. Водночас додатна статистично значуща залежність для MOM6 у регресійних моделях узгоджується з результатами Blitz (2020), який показав, що фактор momentum залишався одним із небагатьох факторів із додатною премією протягом 2010-х років [3, с. 4].

Для AG і DTA квартильна картина теж не виглядає переконливо. У AG різниця між крайніми квартилями становить 0,011418, але внутрішня структура лишається

нерівною. Для DTA розрив майже відсутній (0,001299). Це ще раз показує, що борговий показник і темп зростання активів не формують простого й послідовного зв'язку з майбутньою дохідністю.

Після цього варто перевірити, наскільки результати окремих змінних залежать від складу самої моделі. З цією метою було побудовано кілька спрощених специфікацій із послідовним вилученням BM, GPA, а також AG і DTA. Таблиця 3.1.3 показує, які коефіцієнти зберігають стійкість після таких змін, а які помітно залежать від набору предикторів.

Таблиця 3.1.3 Чутливість коефіцієнтів у вкладених специфікаціях моделі

Специфікація моделі	Змінна	Коефіцієнт	p-value
Pooled OLS (базова)	EP	-0,677038	0,0000
Pooled OLS (базова)	BM	0,017626	0,0014
Pooled OLS (базова)	GPA	0,127998	0,0002
Pooled OLS (базова)	AG	-0,000519	0,9267
Pooled OLS (базова)	LMC	-0,015590	0,0000
Pooled OLS (базова)	MOM6	0,028810	0,0000
Pooled OLS (базова)	DTA	0,010266	0,1566
Pooled OLS (без BM)	EP	-0,623713	0,0000
Pooled OLS (без BM)	GPA	0,067532	0,0194
Pooled OLS (без BM)	AG	-0,001795	0,7496
Pooled OLS (без BM)	LMC	-0,016312	0,0000
Pooled OLS (без BM)	MOM6	0,024571	0,0001
Pooled OLS (без BM)	DTA	0,002770	0,6862
Pooled OLS (без GPA)	EP	-0,632402	0,0000
Pooled OLS (без GPA)	BM	0,006371	0,1684
Pooled OLS (без GPA)	AG	-0,000817	0,8847
Pooled OLS (без GPA)	LMC	-0,015815	0,0000
Pooled OLS (без GPA)	MOM6	0,027738	0,0000
Pooled OLS (без GPA)	DTA	0,006574	0,3600
Pooled OLS (без AG і DTA)	EP	-0,678485	0,0000
Pooled OLS (без AG і DTA)	BM	0,015132	0,0038
Pooled OLS (без AG і DTA)	GPA	0,121274	0,0004
Pooled OLS (без AG і DTA)	LMC	-0,015673	0,0000
Pooled OLS (без AG і DTA)	MOM6	0,028001	0,0000
Period FE (базова)	EP	-0,530894	0,0000
Period FE (базова)	BM	-0,000283	0,9543
Period FE (базова)	GPA	0,098652	0,0014
Period FE (базова)	AG	0,001381	0,7836

Специфікація моделі	Змінна	Коефіцієнт	p-value
Period FE (базова)	LMC	-0,016669	0,0000
Period FE (базова)	MOM6	0,031432	0,0000
Period FE (базова)	DTA	-0,004216	0,5141
Period FE (без BM)	EP	-0,531763	0,0000
Period FE (без BM)	GPA	0,099623	0,0001
Period FE (без BM)	AG	0,001403	0,7797
Period FE (без BM)	LMC	-0,016659	0,0000
Period FE (без BM)	MOM6	0,031501	0,0000
Period FE (без BM)	DTA	-0,004095	0,5025
Period FE (без GPA)	EP	-0,497119	0,0000
Period FE (без GPA)	BM	-0,008990	0,0295
Period FE (без GPA)	AG	0,001156	0,8183
Period FE (без GPA)	LMC	-0,016760	0,0000
Period FE (без GPA)	MOM6	0,030908	0,0000
Period FE (без GPA)	DTA	-0,006942	0,2786
Period FE (без AG і DTA)	EP	-0,531770	0,0000
Period FE (без AG і DTA)	BM	0,000698	0,8812
Period FE (без AG і DTA)	GPA	0,101296	0,0009
Period FE (без AG і DTA)	LMC	-0,016646	0,0000
Period FE (без AG і DTA)	MOM6	0,031923	0,0000

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

Таблиця 3.1.3 узагальнює результати перевірки того, наскільки оцінки окремих змінних залежать від складу моделі. Якщо коефіцієнт зберігає знак і значущість після видалення частини предикторів, такий зв'язок виглядає стійкішим. Якщо результат різко змінюється, це вказує на залежність від специфікації.

Із таблиці видно, що найстійкіше поведуться EP, LMC і MOM6. Їхні коефіцієнти лишаються значущими в усіх розглянутих варіантах як у Pooled OLS, так і в Period FE. Їхній зв'язок із майбутньою доходністю не зникає після спрощення моделі і не тримається на одному конкретному наборі контролів.

GPA теж виглядає достатньо стійкою. У Pooled OLS після видалення BM її коефіцієнт зменшується до 0,067532, але лишається значущим (p-value = 0,0194). У моделі Period FE GPA зберігає значущість в усіх розглянутих специфікаціях (p-value = 0,0014, 0,0001 і 0,0009). Її ефект не виглядає випадковим, хоча частина сигналу справді перетинається з BM.

ВМ виявляється найменш стійкою серед змінних, які в повній моделі ще давали певний результат. У Pooled OLS вона є значущою (0,017626, p -value = 0,0014). Після вилучення GPA ця значущість зникає (0,006371, p -value = 0,1684). Для Period FE результат відрізняється. У повній специфікації ВМ не є значущою (p -value = 0,9543). Після вилучення GPA вона стає значущою з від'ємним знаком (-0,008990, p -value = 0,0295). Її поведінка помітно залежить від присутності GPA в моделі та врахування часових ефектів.

Для AG та DTA картина лишається слабкою в усіх варіантах. Їхні коефіцієнти не набувають стійкої значущості ні в Pooled OLS, ні в Period FE. Проблема не зводиться до конкретного поєднання предикторів у базовій моделі. Для DTA цей висновок особливо важливий, бо навіть після переходу до коректнішого показника левереджу на основі активів переконливого зв'язку з майбутньою дохідністю виявити не вдалося. Для AG така поведінка разом із нерівною кuartильною структурою радше вказує на слабкий або складніший за формою ефект, ніж на стійкий лінійний зв'язок.

Додатковий розбір показав, що стійкість EP і LMC пояснюється трьома речами: вони мають власний окремий сигнал, дають чіткий градієнт у самих даних і майже не втрачають сили після зміни специфікації. GPA теж не є випадковою, але її ефект слабший і краще проявляється після врахування часових умов. MOM6 зберігає значущість у регресіях, проте в кuartильному розподілі виглядає менш переконливо, тому її зв'язок радше вловлюється всередині моделі, ніж у простому групуванні даних. ВМ виявилася нестійкою через залежність від складу моделі, передусім від присутності GPA. Для AG і DTA причина слабких результатів інша: у них немає виразного окремого сигналу, послідовної структури в даних і стійкості після спрощення моделі. Надалі перевірятиметься, чи мають слабкі та нестійкі результати ВМ, AG і DTA пояснення в межах підходів, запропонованих попередніми емпіричними дослідженнями.

3.2 Уточнення причин слабких і нестійких результатів на основі літературних підходів

Підпункт 3.1 показав, які змінні базової моделі дають стійкий результат, а які лишаються слабкими або нестійкими. Проте цього недостатньо, щоб пояснити причини такої поведінки. Далі перевіряється, чи узгоджуються результати центральної моделі з підходами, що використовуються в зарубіжній літературі. Увагу зосереджено на трьох змінних: BM, AG і DTA. BM втратила стійкість, AG не дала переконливого лінійного зв'язку, а DTA лишилася слабкою як показник боргового навантаження.

Для BM використано підхід *Robert Novy-Marx (2013)*, де value-характеристики розглядаються разом із profitability. Для AG взято логіку робіт про asset growth anomaly, де аналіз зміщується від середнього коефіцієнта до крайніх груп спостережень. Для DTA використано підхід із літератури про leverage, де слабкий середній результат може приховувати різний зв'язок боргового навантаження з доходністю в окремих типах компаній.

і. Нестійкість BM і її зв'язок із profitability

Нестійкість BM перевіряється через логіку *Robert Novy-Marx (2013)*. У цій роботі value-характеристики розглядаються разом із profitability. Автор показує, що gross profitability має порівнянну силу з book-to-market у поясненні очікуваних доходностей. Подвійні сортування за *profitability* і *book-to-market* використовуються для перевірки value-сигналу в групах із різним рівнем прибутковості [18, с. 2–3, 25–27]. Для цієї роботи така логіка важлива через нестійкість BM і більш стабільну поведінку GPA. Тому далі BM перевіряється окремо в групах компаній із нижчим і вищим рівнем GPA.

Для такої перевірки вибірку поділено на дві підгрупи за медіанним значенням GPA. У кожній підгрупі повторно оцінено центральну модель, але в Таблиці 3.2.1 наведено тільки коефіцієнт при BM. Даний формат фокусує перевірку на зміні value-сигналу залежно від рівня прибутковості компанії.

Таблиця 3.2.1 Поведінка BM у групах за GPA

Група компаній	Кількість спостережень	Pooled OLS		Period FE	
		Коефіцієнт BM	p-value	Коефіцієнт BM	p-value
Компанії з нижчою GPA	8617	0,02824	0,0016	0,011767	0,1268
Компанії з вищою GPA	8617	-0,002932	0,8474	-0,022349	0,1343

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних.

Примітка: вибірку поділено за медіаною GPA (0,0510). У кожній підвибірці окремо центральну модель $RET_LEAD \sim EP + BM + GPA + AG + LMC + MOM6 + DTA$; для специфікації Period FE додано часові ефекти. Стандартні похибки кластеризовано за компаніями.

Таблиця 3.2.1 добре передає вибірковий характер BM. У групі компаній із нижчим рівнем GPA коефіцієнт при BM у моделі Pooled OLS є додатним і статистично значущим (0,028240; p-value = 0,0016). У групі з вищим GPA цей зв'язок майже зникає (-0,002932; p-value = 0,8474). У моделі Period FE статистична значущість не зберігається в жодній із підгруп, хоча знак і масштаб коефіцієнта також відрізняються. Такий результат узгоджується з логікою *Robert Novu-Marx (2013)*, за якою value-сигнал доцільно розглядати разом із profitability [18, с. 25–27]. У цій вибірці BM проявляється переважно серед фірм із нижчою profitability, а серед більш прибуткових компаній майже втрачає силу.

ii. Слабкість AG: крайні групи та структура зв'язку

Змінна AG у центральній моделі не дала переконливого лінійного ефекту. Тому потрібно перевірити, чи не приховується її можливий вплив у крайніх частинах розподілу, а не в середньому коефіцієнті повної регресії.

Така логіка відповідає літературі про asset growth anomaly. У ній слабкий середній коефіцієнт не розглядається як достатній аргумент для висновку про відсутність ефекту. *Fangjian Fu (2014)* показує, що поведінку AG доцільно аналізувати через крайні групи спостережень, де найнижчий і найвищий децилі asset growth суттєво відрізняються за майбутньою дохідністю [10, с. 2]. Автор також пов'язує низькі дохідності компаній із високим asset growth зі значними обсягами зовнішнього фінансування та показує, що після контролю за delisting bias і external financing незалежний ефект asset growth зникає [10, с. 1, 3, 15–20].

Для такої перевірки всі спостереження було ранжовано за значенням AG і поділено на три рівні за чисельністю підгрупи відповідно до терцилів розподілу. До групи з низьким AG увійшли спостереження зі значеннями $AG \leq 0,023108$, до середньої групи – спостереження в інтервалі $0,023108 < AG \leq 0,102554$, а до групи з високим AG – спостереження зі значеннями $AG > 0,102554$. Значення 0,023108 і 0,102554 є межами терцилів, розрахованими на основі фактичних значень AG у вибірці. Після поділу для кожної групи обчислено середню майбутню дохідність, середній рівень DTA, GPA та LMC. Це дає можливість зіставити групи за дохідністю і фінансовим профілем компаній із різним темпом зростання активів.

Таблиця 3.2.2 Середня майбутня дохідність у групах за AG

Група AG	Середня дохідність	Середній AG	Середній DTA	Середня GPA	Середній LMC	Кількість спостережень
Низький AG	0,041255	-0,040209	0,334536	0,057817	24,259452	5745
Середній AG	0,034174	0,058705	0,318759	0,061244	24,292046	5744
Високий AG	0,051883	0,327969	0,291501	0,070452	24,138464	5745

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

За даними Таблиці 3.2.2 класична конфігурація *asset growth anomaly* у цій вибірці не спостерігається. Найнижча середня майбутня дохідність зафіксована в середній групі (0,034174), тоді як група з високим AG має найвище середнє значення дохідності (0,051883). У вибірці не формується спадний профіль, за якого швидше зростання активів відповідало б нижчій майбутній дохідності.

Цей результат відрізняється від класичної логіки *asset growth anomaly*, на яку вказує *Fangjian Fu (2014)* [10, с. 15–20]. У його роботі проблема високого *asset growth* пов'язується з крайніми децилями розподілу, делістингами та зовнішнім фінансуванням. У цій вибірці вже на рівні групового розподілу не виникає базової конфігурації, з якої починається класична *asset growth anomaly*. Слабкість AG у центральній моделі, ймовірно, пов'язана не тільки з лінійною формою специфікації. У досліджуваному періоді AG не працює як універсальний окремий фактор, а його зміст залежить від інших характеристик компаній.

Це видно і з профілю груп. Компанії з високим AG мають вищу середню GPA (0,070452) та нижчий середній DTA (0,291501), ніж компанії з низьким AG. Тому AG складно трактувати просто як сигнал надмірного розширення активів із

подальшим погіршенням дохідності. Слабкий результат цієї змінної виглядає закономірним, оскільки в межах досліджуваної вибірки стійка *asset growth anomaly* не відтворюється.

iii. Слабкість DTA і неоднорідність зв'язку боргового навантаження з майбутньою дохідністю

Для DTA важливо перевірити, чи слабкий середній результат не пов'язаний із різним характером зв'язку боргового навантаження з дохідністю в окремих групах компаній. Обраний хід перевірки узгоджується з *Doshi et al. (2018)*, де leverage ускладнює пряме трактування середнього зв'язку в загальній регресії через неоднорідність і нелінійність залежностей [5, с. 1–4]. Автори також показують, що після врахування leverage value-ефект суттєво слабшає, тоді як size-ефект зберігається, хоч і у слабшій формі [5, с. 4, 14–17].

Для перевірки цього зв'язку вибірку поділено на дві підгрупи за медіанним значенням LMC. До групи малих компаній увійшли спостереження зі значеннями $LMC \leq 24,200155$, до групи великих компаній – спостереження зі значеннями $LMC > 24,200155$. У кожній підвибірці повторно оцінено центральну модель, але в Таблиці 3.2.3 наведено тільки коефіцієнт при DTA. Обраний формат таблиці дозволяє сфокусувати перевірку на тому, чи змінюється зв'язок DTA з майбутньою дохідністю залежно від розміру компанії.

Таблиця 3.2.3 Поведінка DTA у групах за розміром компаній

Група компаній	Кількість спостережень	Pooled OLS		Period FE	
		Коефіцієнт DTA	p-value	Коефіцієнт DTA	p-value
Малі компанії	8617	0,031573	0,0213	0,010389	0,391
Великі компанії	8617	-0,007716	0,3293	-0,015408	0,0358

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних.

Примітка: вибірку поділено за медіаною LMC (24,200155). У кожній підвибірці оцінено центральну модель $RET_LEAD \sim EP + BM + GPA + AG + LMC + MOM6 + DTA$; для специфікації Period FE додано часові ефекти. Стандартні похибки кластеризовано за компаніями.

Таблиця 3.2.3 передає неоднорідність DTA. Для малих компаній у моделі Pooled OLS коефіцієнт при DTA є додатним і статистично значущим (0,031573; p-value =

0,0213). Для великих компаній у Pooled OLS такий зв'язок не підтверджується, а в моделі Period FE коефіцієнт стає від'ємним і статистично значущим (-0,015408; p-value = 0,0358). Слабкий середній результат DTA в центральній моделі не означає повної відсутності зв'язку між борговим навантаженням і майбутньою дохідністю. У цій вибірці він пов'язаний з тим, що для малих і великих компаній зв'язок має різний напрям. Отримані результати узгоджуються із загальною логікою *Doshi et al. (2018)*, оскільки боргове навантаження не проявляється як однорідний середній ефект для всієї вибірки, а змінюється залежно від розміру компанії.

Отже, проведені перевірки уточнюють причини слабких і нестійких результатів базової моделі. Для ВМ важливим виявився зв'язок із *profitability*, оскільки value-сигнал зберігається переважно серед компаній із нижчим рівнем GPA. Для АГ у межах досліджуваної вибірки не відтворюється класична *asset growth anomaly*. Для DTA слабкий середній результат приховує неоднорідний зв'язок боргового навантаження з майбутньою дохідністю, який відрізняється між малими та великими компаніями. У наступному підпункті ці висновки використовуються для перевірки всієї центральної моделі в сегментах вибірки.

3.3. Сегментована переоцінка центральної моделі

Після результатів, отриманих у попередньому підпункті щодо причин слабкості ВМ, АГ і DTA, доцільно перевірити, як ці змінні поведуться вже не окремо, а в межах усієї центральної моделі після поділу вибірки на сегменти. Для сегментації використано GPA і LMC, оскільки в попередньому підпункті ці характеристики допомогли пояснити неоднорідність ВМ і DTA. Основні результати наведено в Таблицях 3.3.1 і 3.3.2. Додаткові характеристики сформованих підвибірок і показники якості моделей подано в Таблицях А.1 і А.2 Додатка А.

Таблиця 3.3.1 Сегментована оцінка центральної моделі в групах за GPA

Змінна	Компанії з нижчою GPA				Компанії з вищою GPA			
	Pooled OLS		Period FE		Pooled OLS		Period FE	
	Коефіцієнт	p-value	Коефіцієнт	p-value	Коефіцієнт	p-value	Коефіцієнт	p-value
EP	-0,693819	0,0000	-0,496975	0,0001	-0,493467	0,0064	-0,397105	0,0253
BM	0,028240	0,0016	0,011767	0,1268	-0,002932	0,8474	-0,022349	0,1343

GPA	0,220403	0,2530	0,112710	0,5164	-0,035373	0,4546	-0,030140	0,4965
AG	-0,012085	0,2606	-0,005033	0,5993	0,014109	0,2095	0,009852	0,3283
LMC	-0,020057	0,0000	-0,022620	0,0000	-0,013409	0,0000	-0,013169	0,0000
MOM6	0,026793	0,0292	0,038073	0,0053	0,029630	0,0133	0,028906	0,0238
DTA	0,020864	0,1294	0,010154	0,4196	-0,001279	0,9048	-0,017760	0,0833

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

Примітка: вибірку поділено за медіаною GPA (0,0510). У кожній підвибірці окремо оцінено центральну модель.

Таблиця 3.3.1 показує, як змінюється центральна модель після поділу вибірки за GPA. У двох підвибірках найстійкішими залишаються EP, LMC і MOM6. Вони зберігають статистичну значущість як у групі компаній із нижчою GPA, так і в групі з вищою GPA. Це означає, що поділ за прибутковістю не змінює основне ядро моделі.

Найпомітніша відмінність стосується BM. У групі компаній із нижчою GPA ця змінна є значущою в Pooled OLS (0,028240; p-value = 0,0016), тоді як у групі з вищою GPA її зв'язок із майбутньою дохідністю не підтверджується (-0,002932; p-value = 0,8474). AG не стає значущою в жодній із підвбірок. DTA також не формує стійкого результату, хоча в Period FE для компаній із вищою GPA її p-value наближається до 10% рівня значущості (-0,017760; p-value = 0,0833).

У порівнянні із загальною моделлю з підпункту 2.3 поділ за GPA не змінює роль EP, LMC і MOM6. Ці змінні залишаються основою моделі в обох сегментах. Головна зміна стосується BM. У загальній моделі її значущість була нестійкою, а після поділу вона локалізується в групі компаній із нижчою GPA.

Сильнішої окремої конфігурації після поділу за GPA не виникає. Обидві підвбірки зберігають близьку структуру значущих коефіцієнтів, а AG і DTA не переходять до стійко значущих змінних. Сегментація за GPA корисна передусім для уточнення ролі BM, оскільки ця змінна проявляється переважно серед компаній із нижчою GPA.

Після поділу за GPA аналіз переходить до сегментації за LMC. Розмір компанії вже був одним із найстійкіших показників у центральній моделі, тому важливо подивитися, як інші коефіцієнти поведуться в групах нижчої та вищої капіталізації.

Таблиця 3.3.2 Сегментована оцінка центральної моделі за рівнем ринкової капіталізації

Змінна	Компанії з нижчим рівнем ринкової капіталізації				Компанії з вищим рівнем ринкової капіталізації			
	Pooled OLS		Period FE		Pooled OLS		Period FE	
	Коефіцієнт	p-value	Коефіцієнт	p-value	Коефіцієнт	p-value	Коефіцієнт	p-value
EP	-0,669246	0,0000	-0,515912	0,0000	-0,256820	0,0616	-0,107677	0,4050
BM	0,015149	0,1297	-0,017329	0,0297	0,009291	0,1712	0,000085	0,9887
GPA	0,092166	0,0887	0,041944	0,3991	0,093577	0,0306	0,076292	0,0640
AG	-0,005625	0,6439	0,000122	0,9906	-0,003281	0,6992	-0,004161	0,5773
LMC	-0,033183	0,0000	-0,038751	0,0000	-0,000698	0,7511	-0,000536	0,7951
MOM6	0,012374	0,2691	0,016877	0,1595	0,048177	0,0000	0,045790	0,0000
DTA	0,031573	0,0213	0,010389	0,3910	-0,007716	0,3293	-0,015408	0,0358

Джерело: розраховано автором у R на основі самостійно сформованого масиву даних

Примітка: вибірку поділено за медіаною LMC (24,200155). У кожній підвибірці оцінено центральну модель.

Таблиця 3.3.2 показує, що після поділу за LMC центральна модель працює по-різному в групах нижчої та вищої капіталізації. У групі компаній із нижчим рівнем ринкової капіталізації в Pooled OLS значущими є EP, LMC і DTA, а в Period FE – EP, BM і LMC. У групі компаній із вищим рівнем ринкової капіталізації у Pooled OLS значущими є GPA і MOM6, тоді як у Period FE значущими залишаються MOM6 і DTA. Це вказує на зміну складу значущих коефіцієнтів після сегментації за розміром компанії.

Порівняно із загальною моделлю з підпункту 2.3, після поділу за LMC стійкість EP і LMC зберігається насамперед у групі компаній із нижчою капіталізацією. У групі компаній із вищою капіталізацією головним стійким показником стає MOM6. GPA має значущість тільки в Pooled OLS, тому її результат у цьому сегменті не можна вважати стійким для обох специфікацій.

Окремо важливо врахувати поведінку самого LMC. У групі нижчої капіталізації він залишається значущим в обох специфікаціях, тоді як у групі вищої капіталізації втрачає значущість. Це не суперечить базовій моделі, оскільки після поділу за LMC частина варіації розміру вже винесена в саму структуру підвбірок.

Роль DTA після поділу за LMC стає помітнішою. У групі нижчої капіталізації коефіцієнт при DTA є додатним і значущим у Pooled OLS (0,031573; p-value = 0,0213). У групі вищої капіталізації він стає від'ємним і значущим у Period FE (-0,015408; p-value = 0,0358). Отже, борговий показник не має однакового напрямку зв'язку для компаній різного розміру.

Поділ за LMC сильніше змінює конфігурацію центральної моделі, ніж поділ за GPA. Він не створює універсально кращої специфікації, але чіткіше розділяє показники, які проявляються в групах нижчої та вищої капіталізації. AG не набуває значущості в жодному сегменті, тому навіть після сегментації не формує переконливого результату.

Сегментована переоцінка центральної моделі дала нерівномірний результат. Поділ за GPA не змінив загальної структури моделі, але допоміг локалізувати роль ВМ. Ця змінна зберігає значущість лише в групі компаній із нижчою GPA. Поділ за LMC виявився змістовнішим для всієї специфікації, оскільки після нього змінюється склад значущих коефіцієнтів у двох сегментах. DTA також набуває різного напрямку зв'язку залежно від рівня капіталізації. EP, LMC і MOM6 підтвердили відносну стійкість, тоді як AG не дала переконливих результатів навіть після сегментації.

Отримані результати показують, що сегментація вибірки корисна для уточнення ролі ВМ і DTA, але не усуває всі нестійкі результати моделі. Частина показників залишається нестійкою навіть у більш однорідних групах компаній.

Висновки до Розділу 3

Третій розділ був спрямований на пояснення результатів, отриманих у другому розділі. Центральна модель показала стійкий зв'язок із майбутньою дохідністю для EP, GPA, LMC і MOM6, але ВМ виявилася нестійкою, а AG і DTA не дали переконливого самостійного результату. Тому основним завданням розділу було з'ясувати причини такої різниці та перевірити, як змінюється робота моделі після додаткових перевірок і сегментації вибірки.

Для цього спочатку було перевірено індивідуальний сигнал кожної змінної, квартильну структуру майбутньої дохідності та чутливість коефіцієнтів до зміни складу моделі. Ці перевірки підтвердили, що EP, LMC, GPA і MOM6 мають найстійкіший зв'язок із майбутньою дохідністю. BM залежить від складу моделі, передусім від присутності GPA. AG і DTA не формують переконливого самостійного ефекту в загальній постановці.

Подальший аналіз уточнив причини слабких і нестійких результатів. Для BM визначальним виявився зв'язок із profitability, оскільки її значущість зберігається переважно в групі компаній із нижчою GPA. Для AG не підтверджено класичну логіку asset growth anomaly: компанії з високим AG у сформованій вибірці не мали нижчої майбутньої дохідності. Для DTA слабкий середній результат пов'язаний із неоднорідністю зв'язку між борговим навантаженням і майбутньою дохідністю в групах компаній із різним рівнем капіталізації.

Сегментована переоцінка центральної моделі показала, що поділ вибірки дає різний результат залежно від обраної характеристики. Поділ за GPA допоміг уточнити роль BM, оскільки після нього її значущість збереглася лише в групі компаній із нижчою GPA. Поділ за LMC виявився змістовнішим для всієї специфікації, бо змінив склад значущих коефіцієнтів у групах нижчої та вищої капіталізації й чіткіше виявив неоднорідність DTA. EP, LMC і MOM6 підтвердили відносну стійкість у більшості перевірок, тоді як AG не дала переконливого результату навіть після додаткового аналізу та сегментації. Отже, третій розділ показав, що найбільш інформативними виявилися ті перевірки, які дозволили пов'язати поведінку окремих коефіцієнтів зі структурою вибірки без зміни центральної специфікації моделі.

ВИСНОВКИ

У кваліфікаційній роботі було розглянуто зв'язок між фундаментальними фінансовими показниками публічних компаній США та дохідністю їхніх акцій у наступному кварталі. Основне дослідницьке питання полягало в тому, які з цих показників мають статистично значущий і стійкий зв'язок із майбутньою дохідністю, а які дають нестійкий або слабкий результат. Для цього в роботі поєднано теоретичний огляд, формування власного панельного датасету, економетричне оцінювання та додаткові перевірки стійкості результатів.

У першому розділі було сформовано теоретичну рамку дослідження дохідності акцій. Дохідність розглянуто як результат зміни ринкової вартості акції та як показник, що відображає реакцію ринку на фінансовий стан, очікування та ризики компанії. Аналіз попередніх досліджень показав, що фундаментальні характеристики компаній часто використовуються для пояснення відмінностей у майбутній дохідності акцій. До таких характеристик належать показники ринкової оцінки, прибутковості, інвестиційної активності, розміру компанії, попередньої ринкової динаміки та боргового навантаження. Водночас результати попередніх робіт не є однаковими для всіх періодів і вибірок, тому ці зв'язки потребують емпіричної перевірки в конкретному наборі даних.

На основі теоретичного огляду було обґрунтовано вибір змінних для емпіричної моделі. До моделі включено EP, BM, GPA, AG, LMC, MOM6 і DTA, а залежною змінною обрано RET_LEAD. Така побудова відповідає логіці зіставлення фінансових характеристик компанії в періоді t з дохідністю її акцій у періоді $t+1$.

У другому розділі сформовано емпіричну базу дослідження. Фінальна вибірка охоплює 417 публічних компаній США, 17 234 квартальні спостереження та 44 часові періоди за 2015Q1–2025Q4. Вибірка є незбалансованою, оскільки не всі компанії мають однакову кількість квартальних спостережень. До фінального датасету включено лише компанії, для яких доступно не менше 8 кварталів даних. Незалежні змінні було вінсоризовано на рівнях 1% і 99%, щоб зменшити вплив

екстремальних значень. Залежна змінна RET_LEAD не вінсоризувалася, оскільки вона є результатом, який потрібно пояснити.

Економетричне оцінювання проведено у двох специфікаціях: Pooled OLS та моделі з часовими фіксованими ефектами. Для обох моделей використано робастні стандартні похибки з кластеризацією за компаніями. Це було потрібно через панельну структуру даних і можливу залежність похибок у межах однієї компанії в часі.

Результати Pooled OLS показали, що статистично значущий зв'язок із майбутньою дохідністю мають EP, BM, GPA, LMC і MOM6. Змінні AG та DTA не дали статистично значущого результату. У моделі з часовими фіксованими ефектами значущими залишилися EP, GPA, LMC і MOM6, тоді як BM втратила значущість. Це стало першим сигналом, що BM є чутливою до специфікації моделі, а EP, GPA, LMC і MOM6 мають стійкіший зв'язок із майбутньою дохідністю.

Окремої уваги потребує знак коефіцієнта при EP. У двох основних специфікаціях він є від'ємним і статистично значущим. Такий результат відрізняється від класичної логіки *value premium*, однак не є унікальним для емпіричних досліджень останнього десятиліття. Зокрема, у 2010–2019 рр. премія за *earnings-to-price* була від'ємною (-0,72). Тому в межах цієї роботи від'ємний знак EP не трактується як технічна помилка, а розглядається як результат, можливий для сформованої вибірки та періоду дослідження.

Діагностичні тести підтвердили, що структура даних є складнішою за просту лінійну модель. Тести Pesaran CD і Breusch–Pagan LM вказали на крос-секційну залежність у вибірці, а тест Бройша–Годфрі/Вулдріджа – на автокореляцію в панельних даних. Ramsey RESET був статистично значущим для обох моделей, що свідчить про можливі нелінійності або пропущені зв'язки. Водночас максимальне значення VIF становило 1,748, тому критичної мультиколінеарності між пояснювальними змінними не виявлено. Тести на часові фіксовані ефекти були статистично значущими, тому врахування спільних квартальних шоків є обґрунтованим. При цьому Period Fixed Effects у роботі не розглядалася як

автоматично краща модель лише через вищий R^2 , а використовувалася для перевірки стійкості результатів.

У третьому розділі було проведено додаткові аналітичні перевірки. Спочатку оцінено індивідуальний сигнал кожної змінної через парні кореляції, одновимірні моделі, кuartильне групування та вкладені специфікації. Ці перевірки підтвердили, що EP і LMC мають найчіткіший власний сигнал. GPA і MOM6 також зберігають змістовний зв'язок із майбутньою дохідністю, хоча їхній результат менш однорідний у додаткових перевірках. BM поводить себе нестійко і залежить від складу моделі. AG і DTA не формують переконливого самостійного результату в загальній постановці.

Для BM було встановлено, що її нестійкість пов'язана з profitability. Після поділу вибірки за GPA коефіцієнт при BM залишився додатним і статистично значущим лише в групі компаній із нижчою GPA в моделі Pooled OLS. У групі компаній із вищою GPA цей зв'язок зник. Це означає, що BM не працює рівномірно в усій вибірці, а її value-сигнал проявляється переважно серед менш прибуткових компаній. Така нестійкість BM також не суперечить результатам Blitz (2020), який показав слабку ефективність value-показників упродовж 2010-х років.

Для AG не підтверджено класичну asset growth anomaly. У груповому розподілі найвища середня майбутня дохідність була зафіксована не в групі з низьким AG, а в групі з високим AG. Крім того, компанії з високим AG мали вищу середню GPA та нижчий середній DTA. Через це AG у сформованій вибірці не можна трактувати як простий сигнал надмірного зростання активів із подальшим зниженням дохідності. Його слабкий результат виглядає закономірним для цієї вибірки.

Для DTA було виявлено неоднорідність зв'язку між борговим навантаженням і майбутньою дохідністю. У загальній моделі DTA не мав стійкої значущості, але після поділу компаній за рівнем ринкової капіталізації результат змінився. Для малих компаній коефіцієнт при DTA у Pooled OLS був додатним і статистично значущим. Для великих компаній у Period FE він став від'ємним і статистично значущим. Отже, слабкий середній результат DTA не заперечує ролі боргового

навантаження. У цій вибірці він пов'язаний з тим, що для малих і великих компаній зв'язок має різний напрям.

Сегментована переоцінка центральної моделі показала, що поділ вибірки не створює нової універсально кращої специфікації. Поділ за GPA допоміг уточнити роль ВМ, але не змінив основне ядро моделі. EP, GPA, LMC і MOM6 загалом зберегли статус найстійкіших змінних, хоча після сегментації їхня значущість проявлялася нерівномірно. Поділ за LMC дав сильніший аналітичний результат, оскільки після нього змінився набір значущих факторів у групах нижчої та вищої капіталізації. У нижчій капіталізації стійкіше проявлялися EP і LMC, а у вищій капіталізації сильніше проявлявся MOM6. DTA після такого поділу перестав виглядати повністю нейтральним показником.

Отже, відповідь на дослідницьке питання є такою: найстійкіший зв'язок із дохідністю акцій у наступному кварталі в межах сформованої вибірки продемонстрували EP, GPA, LMC і MOM6. Показник ВМ має нестійкий характер і залежить від специфікації моделі та рівня прибутковості компаній. AG не отримав переконливого підтвердження ні в загальній моделі, ні після додаткових перевірок. DTA не є стійким фактором у середньому для всієї вибірки, але його роль змінюється залежно від розміру компанії.

Це підтверджує гіпотезу про те, що фундаментальні фінансові характеристики публічних компаній США мають різну силу та стійкість зв'язку з майбутньою дохідністю акцій.

Практичний висновок роботи полягає в тому, що фундаментальні показники не варто автоматично включати до багатфакторної моделі лише через їхнє використання в попередніх дослідженнях. Кожну змінну потрібно перевіряти в конкретній вибірці, у різних специфікаціях і за різних способів групування даних. Для подальших досліджень доцільно розширити набір фундаментальних характеристик, перевірити інші горизонти дохідності, використати скориговані ціни та протестувати нелінійні зв'язки між фінансовими показниками компаній і майбутньою дохідністю акцій.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3–18. URL: https://search.dailystocks.com/Banz_sizeeffect_1980.pdf (81)90018-0
2. Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance*, 32(3), 663–682. URL: <https://endexiresearch.wordpress.com/wp-content/uploads/2020/07/investment-performance-of-common-stock-s.-basu.pdf>
3. Blitz, D. (2020). Factor performance 2010–2019: A lost decade? SSRN. URL: <https://ssrn.com/abstract=3562242>
4. CFA Institute. (2024). CFA Program Curriculum 2024 Level I (Vol. 1). CFA Institute.
5. Doshi, H., Jacobs, K., Kumar, P., & Rabinovitch, R. (2018). Leverage and the cross-section of equity returns. *The Journal of Finance*, 74(3), 1431–1471. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2493903
6. Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427–465. URL: <https://traders.studentorg.berkeley.edu/papers/Fama%20and%20French.pdf>
7. Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. URL: https://www.bauer.uh.edu/rsusmel/phd/Fama-French_JFE93.pdf
8. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. URL: <https://tevgeniou.github.io/EquityRiskFactors/bibliography/FiveFactor.pdf>
9. Financial Modeling Prep. (n.d.). Financial Modeling Prep API documentation. URL: <https://site.financialmodelingprep.com/developer/docs#undefined>
10. Fu, F. (2014). Dissecting the asset growth anomaly. SSRN. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2533316
11. Goyal, A., Welch, I., & Zafirov, A. (2024). A comprehensive 2022 look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 37(11), 3490–3536. URL: <https://doi.org/10.1093/rfs/hhae041>

- 12.Green, J., Hand, J. R. M., & Zhang, X. F. (2017). The characteristics that provide independent information about average U.S. monthly stock. *The Review of Financial Studies*, 30(12), 4389–4436. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2262374
- 13.Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273. URL: <https://doi.org/10.1093/rfs/hhaa009>
- 14.Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3), 650–705. URL: <https://theinvestmentcapm.com/uploads/1/2/2/6/122679606/houxuezhang2015rfs.pdf>
- 15.Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1), 65–91. URL: https://moneytothemasses.com/wp-content/uploads/2014/08/Jegadeesh_Titman_1993.pdf
- 16.Lewellen, J. (2014). The cross-section of expected stock returns. *Critical Finance Review*, 4(1), 1–44. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2511246
- 17.Митько, Д. (2025). Чинники довгострокової дохідності акцій як класу активів [Магістерська робота, Національний університет «Києво-Могилянська академія»]. eKMAIR. <https://ekmair.ukma.edu.ua/items/bb411b16-0727-4a8d-810a-c133d88d10f4>
- 18.Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), 1–28. URL: https://oldschoolvalue-files.s3.amazonaws.com/pdf/Novy-Marx_Gross-Profitability-Anomaly_JFE_2013.pdf
- 19.Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data. Sets: Comparing Approaches. *The Review of Financial Studies*, 22(1), 435–480. URL: <https://bpb-us-w2.wpmucdn.com/u.osu.edu/dist/7/36891/files/2020/03/PetersenRFS2009clusteringfixedeffectspaneldata.pdf>
- 20.Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425–442. URL: <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- 21.Сова, Є. С. (2020). Теоретичні аспекти взаємозв'язку між монетарними індикаторами і показниками фондового ринку. *Наукові записки НаУКМА. Економічні науки*, 5(1), 118–123. URL: <http://spne.ukma.edu.ua/article/view/213569>

22. Titman, S., Wei, K. C. J., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39(4), 677–700. URL: https://www.nber.org/system/files/working_papers/w9951/w9951.pdf
23. Voloshyna, I. (2015). Opportunities for combining fundamental and technical analysis in forecasting Ukrainian stock prices. *Ukrainian Journal Ekonomist*, 1, 24–26. URL: https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=http://jn.as.nbu.gov.ua/j-pdf/econ_2015_1_8.pdf&ved=2ahUKEwiQ1JGwidyUAxUxQ_EDHaEQPGgQFnoECBoQAQ&usg=AOvVaw2nvOKguAxld87HFmtv62iz

ДОДАТОК А

Додаткові характеристики сегментованих підвбірок і моделей

Таблиця А.1 Додаткові характеристики моделей і підвбірок, сформованих за GPA

Показник	Компанії з нижчою GPA	Компанії з вищою GPA
R ² Pooled OLS	0,025077	0,014157
R ² Period FE	0,237767	0,252327
N	8617	8617
Кількість компаній	302	298
Середня дохідність акцій у наступному кварталі	0,039626	0,04525

Джерело: розраховано автором у R на основі самотійно сформованого масиву даних

Таблиця А.2 Додаткові характеристики моделей і підвбірок, сформованих за рівнем ринкової капіталізації

Показник	Компанії з нижчим рівнем ринкової капіталізації	Компанії з вищим рівнем ринкової капіталізації
R ² Pooled OLS	0,023377	0,004842
R ² Period FE	0,250805	0,237013
N	8617	8617
Кількість компаній	323	335
Середня дохідність акцій у наступному кварталі	0,058239	0,026637

Джерело: розраховано автором у R на основі самотійно сформованого масиву даних

ДОДАТОК Б

Технічна реалізація розрахунків розділу 3 у середовищі R

Лістинг Б.1 Код для розрахунків підпункту 3.1

```

library(plm)

path <-
"/Users/mac/PycharmProjects/PythonProject/Diplom_Y/company_quarter_eviews_ready_u
nbalanced_winsorized_combined436_final_us_clean_v4_dta.csv"
out_dir <- "/Users/mac/PycharmProjects/PythonProject/за 3 дні/section3_deep_dive"
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)

df <- read.csv(path, stringsAsFactors = FALSE)
df$quarter_t <- factor(df$quarter_t)
pdf <- pdata.frame(df, index = c("symbol", "quarter_t"))
vars <- c("ep", "bm", "gpa", "ag", "lmc", "mom6", "dta")

# Table 3.1: pairwise correlation + univariate pooled + univariate time FE
rows <- list()
for (v in vars) {
  corr <- cor(df$ret_lead, df[[v]], use = "complete.obs")
  fit_p <- lm(as.formula(paste("ret_lead ~", v)), data = df)
  cp <- summary(fit_p)$coefficients
  fit_fe <- lm(as.formula(paste("ret_lead ~", v, "+ quarter_t")), data = df)
  cf <- summary(fit_fe)$coefficients
  rows[[v]] <- data.frame(
    variable = toupper(v),
    corr_ret = corr,
    pooled_coef = cp[2, 1],
    pooled_p = cp[2, 4],
    pooled_r2 = summary(fit_p)$r.squared,
    fe_coef = cf[2, 1],
    fe_p = cf[2, 4],
    fe_r2 = summary(fit_fe)$r.squared
  )
}
tab1 <- do.call(rbind, rows)
write.csv(tab1, file.path(out_dir, "table_3_1_variable_signal_v4_dta.csv"),
row.names = FALSE)

# Table 3.2: quartile mean returns and spreads
qrows <- list()
for (v in vars) {
  qs <- quantile(df[[v]], probs = c(0.25, 0.5, 0.75), na.rm = TRUE)
  grp <- cut(
    df[[v]],
    breaks = c(-Inf, qs[1], qs[2], qs[3], Inf),
    labels = c("Q1", "Q2", "Q3", "Q4"),
    include.lowest = TRUE
  )
  means <- tapply(df$ret_lead, grp, mean, na.rm = TRUE)
  qrows[[v]] <- data.frame(
    variable = toupper(v),
    Q1 = means["Q1"],
    Q2 = means["Q2"],
    Q3 = means["Q3"],
    Q4 = means["Q4"],
    spread_Q4_Q1 = means["Q4"] - means["Q1"],
    abs_spread = abs(means["Q4"] - means["Q1"])
  )
}
tab2 <- do.call(rbind, qrows)
write.csv(tab2, file.path(out_dir, "table_3_2_quartile_returns_v4_dta.csv"),
row.names = FALSE)

```

```

# Table 3.3: sensitivity of coefficients in nested models
model_specs <- list(
  pooled_base = ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta,
  pooled_no_bm = ret_lead ~ ep + gpa + ag + lmc + mom6 + dta,
  pooled_no_gpa = ret_lead ~ ep + bm + ag + lmc + mom6 + dta,
  pooled_no_ag_dta = ret_lead ~ ep + bm + gpa + lmc + mom6,
  fe_base = ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta + quarter_t,
  fe_no_bm = ret_lead ~ ep + gpa + ag + lmc + mom6 + dta + quarter_t,
  fe_no_gpa = ret_lead ~ ep + bm + ag + lmc + mom6 + dta + quarter_t,
  fe_no_ag_dta = ret_lead ~ ep + bm + gpa + lmc + mom6 + quarter_t
)
coi <- c("ep", "bm", "gpa", "ag", "lmc", "mom6", "dta")
res <- list()
for (nm in names(model_specs)) {
  fit <- lm(model_specs[[nm]], data = df)
  cf <- summary(fit)$coefficients
  for (v in coi) {
    if (v %in% rownames(cf)) {
      res[[length(res) + 1]] <- data.frame(
        model = nm,
        variable = toupper(v),
        coef = cf[v, 1],
        p = cf[v, 4]
      )
    }
  }
}
tab3 <- do.call(rbind, res)
write.csv(tab3, file.path(out_dir,
"table_3_3_nested_model_sensitivity_v4_dta.csv"), row.names = FALSE)

cat("saved", out_dir, "\n")
print(round(tab1, 6))
cat("\n")
print(round(tab2, 6))

```

Лістинг Б.2 Код для розрахунків підпункту 3.2

```

library(lmtest)
library(sandwich)

path <-
"/Users/mac/PycharmProjects/PythonProject/Diplom_Y/company_quarter_eviews_ready_u
nbalanced_winsorized_combined436_final_us_clean_v4_dta.csv"
out_dir <- "/Users/mac/PycharmProjects/PythonProject/за 3 дні/section3_deep_dive"
dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)

df <- read.csv(path, stringsAsFactors = FALSE)
df$quarter_t <- factor(df$quarter_t)

cluster_coef <- function(data, formula, term) {
  fit <- lm(formula, data = data)
  vc <- vcovCL(fit, cluster = data$symbol, type = "HC1")
  ct <- coeftest(fit, vcov. = vc)
  c(
    coef = if (term %in% rownames(ct)) unname(ct[term, 1]) else NA_real_,
    p = if (term %in% rownames(ct)) unname(ct[term, 4]) else NA_real_,
    r2 = summary(fit)$r.squared,
    n = nrow(data)
  )
}

quartile_spread <- function(data, var) {
  qs <- quantile(data[[var]], probs = c(0.25, 0.5, 0.75), na.rm = TRUE)
  grp <- cut(
    data[[var]],
    breaks = c(-Inf, qs[1], qs[2], qs[3], Inf),
    labels = c("Q1", "Q2", "Q3", "Q4"),
    include.lowest = TRUE
  )
  means <- tapply(data$ret_lead, grp, mean, na.rm = TRUE)
  c(
    Q1 = unname(means["Q1"]),
    Q4 = unname(means["Q4"]),
    spread = unname(means["Q4"] - means["Q1"])
  )
}

# 3.2.1 BM inside low/high GPA groups
gpa_med <- median(df$gpa, na.rm = TRUE)
df$gpa_group <- ifelse(df$gpa <= gpa_med, "Низька GPA", "Висока GPA")

bm_rows <- list()
for (grp in c("Низька GPA", "Висока GPA")) {
  sub <- df[df$gpa_group == grp, ]
  pooled <- cluster_coef(sub, ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta,
    "bm")
  fe <- cluster_coef(sub, ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta +
    quarter_t, "bm")
  spread <- quartile_spread(sub, "bm")
  bm_rows[[length(bm_rows) + 1]] <- data.frame(
    subgroup = grp,
    observations = nrow(sub),
    median_gpa_cutoff = gpa_med,
    pooled_bm_coef = pooled["coef"],
    pooled_bm_p = pooled["p"],
    pooled_r2 = pooled["r2"],
    fe_bm_coef = fe["coef"],
    fe_bm_p = fe["p"],
    fe_r2 = fe["r2"],
    bm_q1_ret = spread["Q1"],
    bm_q4_ret = spread["Q4"],
    bm_q4_minus_q1 = spread["spread"]
  )
}
tab_bm <- do.call(rbind, bm_rows)
write.csv(tab_bm, file.path(out_dir, "table_3_2_1_bm_by_gpa_v4_dta.csv"),
  row.names = FALSE)

```

```

# 3.2.2 AG extremes: low / medium / high
ag_q <- quantile(df$ag, probs = c(1/3, 2/3), na.rm = TRUE)
df$ag_group <- cut(
  df$ag,
  breaks = c(-Inf, ag_q[1], ag_q[2], Inf),
  labels = c("Низький AG", "Середній AG", "Високий AG"),
  include.lowest = TRUE
)

ag_summary <- aggregate(
  cbind(ret_lead, ag, dta, gpa, lmc) ~ ag_group,
  data = df,
  FUN = mean
)

ag_counts <- as.data.frame(table(df$ag_group), stringsAsFactors = FALSE)
names(ag_counts) <- c("ag_group", "observations")
ag_table <- merge(ag_summary, ag_counts, by = "ag_group", sort = FALSE)
names(ag_table) <- c("Група AG", "Середня дохідність", "Середній AG", "Середній
DTA", "Середня GPA", "Середній LMC", "Кількість спостережень")
write.csv(ag_table, file.path(out_dir, "table_3_2_2_ag_groups_v4_dta.csv"),
  row.names = FALSE)

# 3.2.3 DTA inside size groups
lmc_med <- median(df$lmc, na.rm = TRUE)
df$size_group <- ifelse(df$lmc <= lmc_med, "Малі компанії", "Великі компанії")

dta_rows <- list()
for (grp in c("Малі компанії", "Великі компанії")) {
  sub <- df[df$size_group == grp, ]
  pooled <- cluster_coef(sub, ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta,
    "dta")
  fe <- cluster_coef(sub, ret_lead ~ ep + bm + gpa + ag + lmc + mom6 + dta +
    quarter_t, "dta")
  spread <- quartile_spread(sub, "dta")
  dta_rows[[length(dta_rows) + 1]] <- data.frame(
    subgroup = grp,
    observations = nrow(sub),
    median_lmc_cutoff = lmc_med,
    pooled_dta_coef = pooled["coef"],
    pooled_dta_p = pooled["p"],
    pooled_r2 = pooled["r2"],
    fe_dta_coef = fe["coef"],
    fe_dta_p = fe["p"],
    fe_r2 = fe["r2"],
    dta_q1_ret = spread["Q1"],
    dta_q4_ret = spread["Q4"],
    dta_q4_minus_q1 = spread["spread"]
  )
}
tab_dta <- do.call(rbind, dta_rows)
write.csv(tab_dta, file.path(out_dir, "table_3_2_3_dta_by_size_v4_dta.csv"),
  row.names = FALSE)

cat("saved literature-style 3.2 tables to", out_dir, "\n")

```

Лістинг Б.2 Код для розрахунків підпункту 3.3

```

library(lmtest)
library(sandwich)

path <- "/Users/mac/PycharmProjects/PythonProject/за 3
дні/company_quarter_eviews_ready_unbalanced_winsorized_combined436_final_us_clean
_v4_dta.csv"
out_dir <- "/Users/mac/PycharmProjects/PythonProject/за 3 дні/section3_deep_dive"

dir.create(out_dir, showWarnings = FALSE, recursive = TRUE)

df <- read.csv(path, stringsAsFactors = FALSE)
df$quarter_t <- factor(df$quarter_t)

segment_summary <- function(data, label) {
  vars <- c("ep", "bm", "gpa", "ag", "lmc", "mom6", "dta")
  var_labels <- c("EP", "BM", "GPA", "AG", "LMC", "MOM6", "DTA")

  pooled_fit <- lm(
    as.formula(paste("ret_lead ~", paste(vars, collapse = " + "))),
    data = data
  )

  fe_fit <- lm(
    as.formula(paste("ret_lead ~", paste(c(vars, "quarter_t"), collapse = " +
  "))),
    data = data
  )

  pooled_ct <- coeftest(
    pooled_fit,
    vcov. = vcovCL(pooled_fit, cluster = data$symbol, type = "HC1")
  )

  fe_ct <- coeftest(
    fe_fit,
    vcov. = vcovCL(fe_fit, cluster = data$symbol, type = "HC1")
  )

  coef_rows <- data.frame(
    segment = label,
    row_type = "coef",
    variable = var_labels,
    pooled_coef = pooled_ct[vars, 1],
    pooled_p = pooled_ct[vars, 4],
    fe_coef = fe_ct[vars, 1],
    fe_p = fe_ct[vars, 4],
    stringsAsFactors = FALSE
  )

  stats_rows <- data.frame(
    segment = label,
    row_type = c("r2", "n"),
    variable = c("R²", "N"),
    pooled_coef = c(summary(pooled_fit)$r.squared, nrow(data)),
    pooled_p = c(NA_real_, NA_real_),
    fe_coef = c(summary(fe_fit)$r.squared, nrow(data)),
    fe_p = c(NA_real_, NA_real_),
    stringsAsFactors = FALSE
  )

  rbind(coef_rows, stats_rows)
}

```

```
# 3.3.1 Поділ вибірки за GPA

gpa_med <- median(df$gpa, na.rm = TRUE)
df_gpa_low <- df[df$gpa <= gpa_med, ]
df_gpa_high <- df[df$gpa > gpa_med, ]

tab_gpa <- rbind(
  segment_summary(df_gpa_low, "Компанії з нижчою GPA"),
  segment_summary(df_gpa_high, "Компанії з вищою GPA")
)

write.csv(
  tab_gpa,
  file.path(out_dir, "table_3_3_1_segmented_model_gpa_v4_dta.csv"),
  row.names = FALSE
)

# 3.3.3 Поділ вибірки за рівнем ринкової капіталізації

lmc_med <- median(df$lmc, na.rm = TRUE)
df_small <- df[df$lmc <= lmc_med, ]
df_large <- df[df$lmc > lmc_med, ]

tab_lmc <- rbind(
  segment_summary(df_small, "Малі компанії"),
  segment_summary(df_large, "Великі компанії")
)

write.csv(
  tab_lmc,
  file.path(out_dir, "table_3_3_2_segmented_model_lmc_v4_dta.csv"),
  row.names = FALSE
)

cat("Saved:\n")
cat(file.path(out_dir, "table_3_3_1_segmented_model_gpa_v4_dta.csv"), "\n")
cat(file.path(out_dir, "table_3_3_2_segmented_model_lmc_v4_dta.csv"), "\n")
```