

Матвійчук А.В.,

кандидат економічних наук,
доцент кафедри економіко-математичних методів
ДВНЗ “Київський національний економічний університет
імені Вадима Гетьмана”

ПРОГНОЗУВАННЯ НАДХОДЖЕНЬ ПОДАТКУ НА ДОДАНУ ВАРТІСТЬ

Побудовано економіко-математичні моделі прогнозування надходжень податку на додану вартість, що ґрунтуються на економетричному інструментарії, нейронних мережах та нечіткій логіці. Проведено порівняльний аналіз результатів експериментів з прогнозування ПДВ на базі цих моделей та надано відповідні рекомендації з їх використання при плануванні дохідної частини бюджету.

There are constructed the economical and mathematical models of value added tax incoming forecasting, which are based on econometric tools, neural networks and fuzzy logic. There are carried out the comparative experiments for VAT forecasting on the basis of these models and given corresponding recommendations for using these models when planning the income part of budget.

Для стабільного соціально-економічного розвитку країни важливим є збільшення надходжень коштів до бюджетів усіх рівнів. У зв'язку з цим особливої актуальності набуває питання підвищення ефективності формування дохідної частини бюджету, яке безпосередньо пов'язано з прогнозуванням бюджетних надходжень та розподілом бюджетних призначень за різними типами податків у розрізі регіонів із урахуванням галузевих особливостей.

Найпростішим підходом до оцінювання податкового потенціалу регіону та відповідного прогнозування бюджетних надходжень є коригування попередніх зборів податків на інфляцію та розвиток регіону за цей період. Проте такий підхід дещо некоректний, оскільки він не враховує великої кількості важливих факторів та базується здебільшого на адміністративних можливостях виконання попередніх бюджетних призначень, які, вірогідно, були встановлені певною мірою суб'єктивно.

Щоб пересвідчитись у недостатній коректності такого підходу, автором було проведено порівняння розрахунків надходжень ПДВ, скоригованих з урахуванням темпів зростання надходжень у попередньому періоді, із реальними обсягами надходжень ПДВ у 2005 та 2006 роках. У результаті було виявлено суттєву залежність змін обсягів надходжень ПДВ не від показників розвитку економіки, а передусім від політичної ситуації (дуже відчутного суб'єктивного впливу людського фактора), що взагалі значно ускладнює можливість моделювання економічних систем із застосуванням економетричних підходів.

В Україні вже робилися спроби прогнозувати податкові надходження, зокрема, із застосуванням авторегресійної моделі ARIMA¹. Як відмічається окремими

¹ Скрипник А.В., Терещенко Л.А. Динаміка податкових надходжень та їх прогноз за допомогою моделі ARIMA // Моделювання та інформаційні системи в економіці. — К.: КНЕУ, 2001. — Вип. 66. — С. 7—12.

науковцями, для використання цієї моделі необхідними є щомісячні спостереження за податком, що прогнозується, протягом принаймні 5–6 років. Проте зауважимо, що з огляду на відсутність статистичної однорідності відповідних часових рядів в умовах української економіки перехідного періоду (тим більше протягом такого значного відрізка часу) можливість здійснення адекватного прогнозування на основі екстраполяційної регресійної моделі виключається. Зазначимо, що навіть у наукових працях, де використовуються моделі ARIMA², йдеться про відсутність нормального розподілу серед статистичних даних. Відповідно, застосування подібних підходів є недостатньо обґрунтованим. Тому ARIMA або інші однофакторні моделі прогнозування у дослідженні не розглядаються. Автором зроблено вибір на користь багатфакторного прогнозування. А з метою збільшення статистики запропоновано аналізувати одночасно всі часові ряди надходжень податку за різними регіонами України. Відповідно, у такому разі здійснити прогнозування часових рядів шляхом їх екстраполяції вже не вдасться.

Що стосується багатфакторного прогнозування на основі економетричних моделей, то для їх використання є обов'язковим дотримання гіпотези щодо нормального або логнормального розподілу випадкових величин показників, які є в моделі. Однак аналіз на відповідність нормальному закону розподілу практично для будь-якого показника в умовах української економіки перехідного періоду дасть негативний результат щодо підтвердження даної гіпотези. Проведений нами аналіз наявної статистики показав, що рівні доходів різних суб'єктів господарювання різняться на порядки, причому відповідна доходам кількість підприємств змінюється згідно з розподілом, близьким до розподілу Парето, який властивий початковим стадіям становлення ринкових відносин в економіці. Це підтверджують і дослідження інших науковців у цій галузі³. Таким чином, для розв'язання завдання прогнозування бюджетних надходжень доцільно використовувати математичний апарат, який не вимагатиме обов'язкового дотримання гіпотези щодо нормального розподілу або статистичної однорідності відповідних випадкових процесів.

Отже, автором розроблено методологічний підхід до розв'язання завдання прогнозування бюджетних надходжень, який відповідає вказаним вище вимогам. У рамках цього підходу було запропоновано перелік найбільш вагомих факторів впливу, що є визначальними для бюджетних надходжень за обраними типами податків. На другому етапі побудови моделей прогнозування — сформовано статистику за обраними показниками.

На третьому етапі побудови економіко-математичних моделей у рамках відповідного методологічного підходу було проведено специфікацію моделей згідно

² Скрипник А.В. Моделювання податкової політики у трансформаційній економіці: дис. ... доктора. екон. наук: 08.03.02. — К., 2004. — 421 с.; Ситник В.Ф. Інформаційні системи в економіці та бізнесі: сучасний стан і перспективи // Моделювання та інформаційні системи в економіці. — К.: КНЕУ, 2000. — Вип. 64. — С. 5–19.

³ Скрипник А.В. Зазнач. праця.

з відібраними факторами. Спочатку побудовано кількісну економетричну модель прогнозування податкових надходжень регресійного типу, що базується на статистичних характеристиках випадкових величин пояснюючих змінних. Використання такої моделі передбачає отримання можливості порівняння якості прогнозування з більш складними економіко-математичними моделями (хоча вище і обґрунтовувалася проблематичність застосування моделей цього типу). Відповідно, у рамках методологічного підходу також було побудовано нелінійні моделі прогнозування надходжень за різними типами податків, що не вимагають дотримання випадковими величинами пояснюючих змінних гіпотези щодо нормального розподілу та ґрунтуються на методах теорій нейронних мереж та нечіткої логіки. Тобто паралельно було побудовано три типи моделей — регресійну, нейромережеву та нечітку.

Четвертий етап — навчання моделей на відібраних статистичних даних. Всі вказані моделі характеризуються здатністю до оптимізації власних параметрів згідно з наявною статистикою. На п'ятому етапі було проведено експерименти з визначення адекватності побудованих моделей, точності прогнозування бюджетних надходжень, а також здійснено необхідне корегування моделей у відповідності з отриманими результатами. Після тестування розроблених моделей, вибору адекватної моделі та проведення додаткових узгоджень її було взято за базову для прогнозування податкових надходжень.

Викладемо методологічний підхід та результати порівняльного аналізу ефективності побудованих моделей на прикладі статистичних даних щодо надходжень ПДВ у розрізі регіонів.

Проведений аналіз динаміки надходжень ПДВ, а також теоретичне узагальнення власного досвіду та наукових праць дослідників у цій галузі дали змогу дійти висновку щодо важливості врахування при прогнозуванні надходжень податку на додану вартість таких показників: надходження ПДВ у попередньому періоді; втрати від надання податкових пільг; сума податкових зобов'язань з ПДВ; сума податкового кредиту з ПДВ; сума ПДВ, що підлягає відшкодуванню; відшкодування ПДВ; валова додана вартість; обсяг експорту; обсяг імпорту; ВВП; переплата ПДВ. Проте це є первинним переліком факторів впливу, який у процесі побудови економіко-математичних моделей може зазнати певних змін. Крім того, значення самих показників проходитимуть первинну обробку перед їх поданням на входи моделей прогнозування. Зокрема, важливим є перехід від абсолютних значень цих показників до відносних, що буде далі обґрунтовано.

Варто зазначити, що суттєва проблема прогнозування пов'язана з тим, що на момент прогнозу значення пояснюючих змінних ще невідомі (прогноз здійснюється наприкінці поточного року на рік вперед). Отже, для прогнозування, наприклад, надходжень на III квартал, невідомі ані показники виконання плану за II квартал року прогнозування, ані поточні значення якихось екзогенних факторів. Можна лише застосовувати дані за попередні роки. Причому наявна статистика є нерепрезентативною (це при тому, що дані офіційно було підготовлено та

передано для розрахунків ДПА України та Держаналітінформом). Статистика представлена для різних факторів за різні роки (часто не більше двох років спостережень). Для одних показників є дані помісячні, других — поквартальні, третіх — щорічні. Оскільки статистика подана нерівномірно, то побудова моделей потребує початкової обробки інформації щодо вхідних факторів.

Для побудови економіко-математичних моделей прогнозування надходжень ПДВ необхідно здійснити перехід від абсолютних значень відібраних показників до їх відносних змін. Важливість цього продемонструємо на прикладі взаємозв'язку між щорічними показниками надходжень ПДВ та обсягом імпорту для країни в цілому. Якщо подивитися на відносні зміни цих показників, то кореляція між ними є від'ємною ($-0,66$) — тобто відповідно до цього значення можна зробити висновок, що темпи зростання надходжень ПДВ зменшуються через рік після того, як темпи імпорту збільшуються. Проте, якщо розглядати кореляцію між цими показниками в абсолютних величинах, то вона є додатною із дуже високим показником взаємозв'язку ($+0,98$). Аналогічна залежність спостерігається між усіма показниками в абсолютних величинах.

Пояснюється це тим, що всі показники з часом зростають, хоч і різними темпами, проте кореляція виявляє саме цю загальну тенденцію. При побудові економетричних моделей зазначеного виявляється достатньо, щоб відібрати для аналізу або незначимі показники (наприклад, коефіцієнт кореляції між надходженнями ПДВ та обсягом експорту в абсолютних значеннях складає $0,85$, а у відносних — лише $0,08$), або показники, які мають мультиколінеарний зв'язок між собою.

Аналіз залежності надходжень ПДВ від значень низки показників у попередньому періоді, для яких є поквартальна статистика, показує, що кореляція між відносними значеннями всіх показників та надходженнями ПДВ наступного періоду є на дуже низькому рівні (максимальне значення кореляції у межах $0,37$), що викликає сумніви відносно можливості побудови адекватної економетричної моделі на основі цих даних. Для таких показників, як сума податкових зобов'язань з ПДВ, сума податкового кредиту з ПДВ, сума ПДВ, що підлягає відшкодуванню, взагалі не можна розрахувати коефіцієнт кореляції з величинами податкових надходжень ПДВ наступного періоду через відсутність часової відповідності між вхідними та вихідною змінною у відносних величинах. Цікаво, що коефіцієнти автокореляції із лагом у рік між річними значеннями надходжень ПДВ виявилися значно вищими, ніж між їх кварталними значеннями, що вказує на меншу залежність річних показників від впливу суб'єктивних чинників, ніж кварталних надходжень.

Таким чином, проведений аналіз статистичних даних та пошук прихованих закономірностей у відповідних часових рядах виявив низьку залежність вихідної змінної “Надходження ПДВ” від змін інформативних факторів у попередніх періодах. Високі коефіцієнти кореляції між вхідними та вихідною змінними з'являються лише в абсолютних величинах, що пояснюється загальною тенденцією зростання усіх фінансових та економічних показників, пов'язаною з розвитком економіки країни та супутніми інфляційними процесами.

Відповідно, щоб не допустити залучення до економіко-математичних моделей незначимих пояснюючих змінних із високим необґрунтованим рівнем шільності зв'язку як із вихідною, так і з рештою вхідних змінних, доречно здійснювати прогнозування, базуючись на відносних змінах показників, хоча між ними і не виявлено значимих функціональних залежностей.

Також необхідно враховувати абсолютно недостатній обсяг статистичних даних для побудови та налаштування економіко-математичних моделей. Так, якщо ми прагнемо побудувати регресійну модель прогнозування надходжень, яка буде складатися з 11 факторів (що попередньо було відібрано до аналізу), то для цього необхідно володіти статистикою за всіма цими факторами мінімум з 12 спостережень, представлених в одні й ті самі моменти часу. А краще, щоб вибірка складалася хоча б із кількох десятків спостережень. Окрім того, що дані представлені для різних часових проміжків, їх абсолютно недостатньо, аби коректно побудувати навіть трьохфакторну модель для кожної області окремо, що дало б змогу однозначно врахувати галузеві та інші економічні особливості регіону.

Про нейромережеву модель тут взагалі не йдеться, якщо за мету поставити побудову окремої моделі для кожного регіону із відповідним налаштуванням її параметрів. Адже, зважаючи на значну кількість параметрів нейронної мережі (вага міжнейронних зв'язків, параметри функцій активацій нейронів), навіть за умови її досить простої конфігурації, необхідно мати приблизно втричі більший обсяг статистичної вибірки за кількість вхідних змінних.

Розв'язанням проблеми браку статистики є об'єднання даних щодо надходжень ПДВ з усіх областей, звісно, з переведенням їх у відносну форму. При такому підході ми втрачаємо чітку прив'язку до особливостей функціонування кожного окремого регіону, але можемо виявити загальні тенденції впливу пояснюючих факторів на зміни податкових надходжень. На основі цієї узагальненої моделі, в якій врахована одночасно специфіка кожної територіальної одиниці країни, можна здійснювати прогнозування надходжень ПДВ як для України в цілому, так і за регіонами.

Як зазначалося вище, складність побудови економіко-математичних моделей прогнозування надходжень ПДВ полягає в тому, що відповідну статистику за різними показниками подано нерівномірно. Так, наприклад, відносні зміни показників “Сума податкових зобов'язань з ПДВ”, “Сума податкового кредиту з ПДВ”, “Сума ПДВ, що підлягає відшкодуванню” за всіма областями країни є лише для 2006 року та перших двох кварталів 2007 року, відповідне значення вихідної змінної — відносної зміни надходжень ПДВ цього ж кварталу наступного року на час здійснення аналізу наявне лише для двох періодів — I та II кварталів 2007 року. Для III та IV кварталів 2007 року необхідної статистики надходжень ПДВ немає. Зауважимо, що значення ВДВ та ВВП для 2006 року за регіонами у статистиці відсутні.

Також слід вказати на недостатність статистики за показниками “Відшкодування ПДВ” та “ВВП” порівняно з іншими пояснюючими змінними. Зважаючи,

що показники ВВП за регіонами були представлені лише для 2005 року, то окремо будувалися моделі за усіма показниками (крім вилучених вище) з метою прогнозування надходжень ПДВ у чотирьох кварталах 2006 року. У такому випадку статистика за всі інші роки при побудові цієї моделі вже не бралася для розрахунків.

Відповідно, у проведеному дослідженні було побудовано значну кількість економетричних моделей прогнозування надходжень ПДВ, які ґрунтувалися на різних множинах пояснюючих змінних, та, зважаючи на брак статистичної інформації, налаштовувалися згідно з даними щодо цих змінних за різні роки. Тобто для деяких моделей використовувалися статистичні дані лише за 2006 рік (для тих пояснюючих змінних, для яких була наявна статистика за 2006 рік) з відтворенням при цьому надходжень ПДВ у відповідних кварталах 2007 року. Окремі моделі будувалися на статистиці за 2004—2005 роки, прогножуючи надходження ПДВ у 2005—2006 роках, відповідно. Також було побудовано економіко-математичні моделі й на інших часових інтервалах.

Оцінку адекватності моделей прогнозування можна здійснювати кількома способами. Одним з напоширеніших та достовірних підходів є перевірка точності прогнозування із застосуванням різних критеріїв ефективності (наприклад, дисперсії відхилень прогнозу від реальних значень часових рядів, середньоквадратичної помилки прогнозу або інших показників). Середньоквадратична помилка прогнозу дає змогу визначити ступінь розкиду прогнозованих оцінок від реальних значень фінансового показника:

$$\sigma_{err} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n [\hat{y}(i) - y(i)]^2}, \quad (1)$$

де n — кількість елементів часового ряду, на які робиться прогноз;

$\hat{y}(i)$, $y(i)$ — прогнозоване та реальне значення i -го елементу часового ряду.

Проте один лише показник середньоквадратичної помилки (1) не є достатньо інформативним щодо якості прогнозу, оскільки для стійких часових рядів здійснювати прогноз набагато легше, ніж для показників, курс яких є більш волатильним. Тому при оцінюванні якості прогнозу важливо враховувати інтенсивність коливання аналізованого показника на визначеному відрізьку часу.

Для того, щоб можна було порівнювати якість прогнозу за різноманітними методиками для різних часових рядів на різних ділянках часу, доцільно використовувати показник нормалізованої середньоквадратичної помилки (normalized root-mean-square error), що являє собою відношення середньоквадратичної помилки (1) до середньоквадратичного відхилення значень випадкової величини прогнозованого показника відносно свого математичного сподівання на прогнозованому інтервалі часу:

$$\sigma_{norm} = \sqrt{\frac{\sum_{i=1}^n [\hat{y}(i) - y(i)]^2}{\sum_{i=1}^n [E(y) - y(i)]^2}}, \quad (2)$$

де $E(y)$ — математичне сподівання.

Подібним чином можна здійснювати порівняльний аналіз точності прогнозування на основі різних підходів. Перевірку адекватності лінійних моделей можна також проводити шляхом оцінки можливості відтворення вихідної змінної на основі відповідних значень вхідних змінних із застосуванням F -критерія Фішера чи коефіцієнта детермінації R^2 .

Усі побудовані вище економетричні моделі показали коефіцієнт детермінації на рівні до 0,15 та відповідні низькі значення F -критерія (які значно нижчі за табличні значення), що вказує на їх недостатню придатність для прогнозування надходжень ПДВ. Нормалізована середньоквадратична помилка прогнозу для вказаних моделей відмічалася у межах від 0,9 до 1,0. Такі значення помилки прогнозу (2) показують, що середньоквадратична помилка прогнозу (1) близька до середньоквадратичного відхилення часового ряду, а прогноз, відповідно, є ненабагато кращим за звичайне середнє арифметичне.

У результаті проведеного порівняльного аналізу точності відтворення надходжень ПДВ на основі низки пояснюючих змінних попередніх періодів було відібрано економіко-математичні моделі, що будувалися на основі статистики за роками, з 2002 по 2005, для показників: “Надходження ПДВ”, “Втрати від надання податкових пільг”, “Валова додана вартість”, “Обсяг експорту”, “Обсяг імпорту” та “Переплата”. На рис. 1 показано відтворення відносних змін надходжень ПДВ за регіонами для I кварталу із застосуванням відповідної економетричної моделі на основі вказаних показників.

Отже, прогноз надходжень ПДВ, отриманий на базі регресійних моделей, суттєво розходиться з реальними даними. Такі низькі характеристики відтворення вихідного показника на основі пояснюючих змінних обумовлені низькою залежністю між ними. Так, коефіцієнт автокореляції між значеннями надходжень ПДВ із лагом в один рік для I кварталу становить 0,064; для другого — 0,102; для третього — 0,035; для четвертого — 0,005. Коефіцієнти кореляції для чотирьох кварталів між “Надходженнями ПДВ” та показником “Втрати від надання податкових пільг” — 0,118, -0,02, -0,063 та -0,031; “Валова додана вартість” — 0,034, 0,03, 0,055 та 0,092; “Обсяг експорту” — 0,059, 0,22, -0,293 та -0,06; “Обсяг імпорту” — 0,113, 0,095, -0,005 та -0,08; “Переплата” — -0,036, 0,12, -0,075 та 0,039.

Подібна ситуація з прогнозуванням (див. рис. 1) спостерігається й для інших кварталів. Практично всі побудовані економетричні моделі виявилися нездатними відтворювати надходження ПДВ на базі значень низки різних показників попереднього періоду. Пояснити це можна лінійною природою моделей, а також наявністю та суттєвою важливістю деяких неврахованих суб’єктивних чинників, таких, як, скажімо, політична складова, спрямованість керівництва податкової служби на активізацію певного виду діяльності в кожному окремий період часу тощо. Ці чинники вносять настільки вагому частку невизначеності до відповідних випадкових процесів, що їх моделювання із застосуванням класичного економетричного інструментарію викликає значні сумніви.

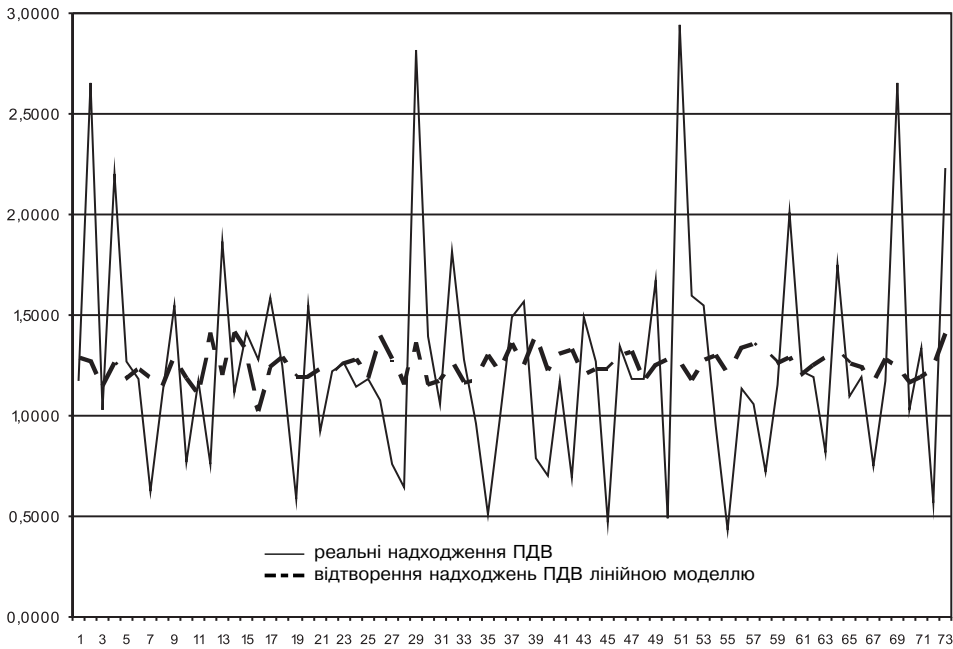


Рис. 1. Відтворення відносних змін надходжень ПДВ для I кварталу із застосуванням лінійної моделі

Продемонструємо можливість відтворення надходжень ПДВ із застосуванням економіко-математичних моделей, побудованих на основі методів нейронних мереж на цій же статистиці та з відібраними до регресійних моделей пояснюючими змінними. Формування цих моделей зводиться до вибору оптимальної конфігурації нейронних мереж, визначення кількості внутрішніх шарів мереж та нейронів у цих шарах.

Налаштування параметрів економіко-математичних моделей на нейронних мережах здійснюватимемо на базі тих пояснюючих змінних, за якими будувались останні економетричні моделі, оскільки для них існує найбільш повна статистика за всіма регіонами. Обсяг навчальної вибірки особливо важливий для коректного налаштування нейронної мережі, що дає змогу уникнути ефекту перенавчання (сутність чого розкрито, зокрема, у авторській роботі⁴).

Перша модель, побудована на основі інструментарію нейронних мереж, мала таку конфігурацію: повнозв'язна нейронна мережа типу “багатошаровий персептрон з одним внутрішнім шаром”, що складався з трьох нейронів; перший шар складається з шести нейронів (за кількістю вхідних змінних моделі) та один нейрон вихідного шару (на якому отримується прогнозоване значення майбутніх надходжень ПДВ). Перша ж побудована нелінійна модель на базі інструментарію

⁴ *Матвійчук А.В.* Виявлення і запобігання ефекту перенавчання нейронної мережі // Збірник наукових праць Черкаського державного технологічного університету. Серія: Економічні науки. — 2002. — Випуск 3. — С. 124–131.

нейронних мереж виявила високу точність відтворення вихідної змінної на основі множини вхідних показників. Це підтверджується і значенням нормалізованої середньоквадратичної помилки прогнозу $\sigma_{norm} = 0,043$ ($\sigma_{err} = 0,023$), яке вказує на підвищення його точності у $(1-0,043)/(1-0,983) = 56,3$ раза порівняно з лінійною економетричною моделлю, побудованою на цьому ж статистичному матеріалі. Результат відтворення показника “Надходження ПДВ” для I кварталу із застосуванням цієї моделі наведено на рис. 2.

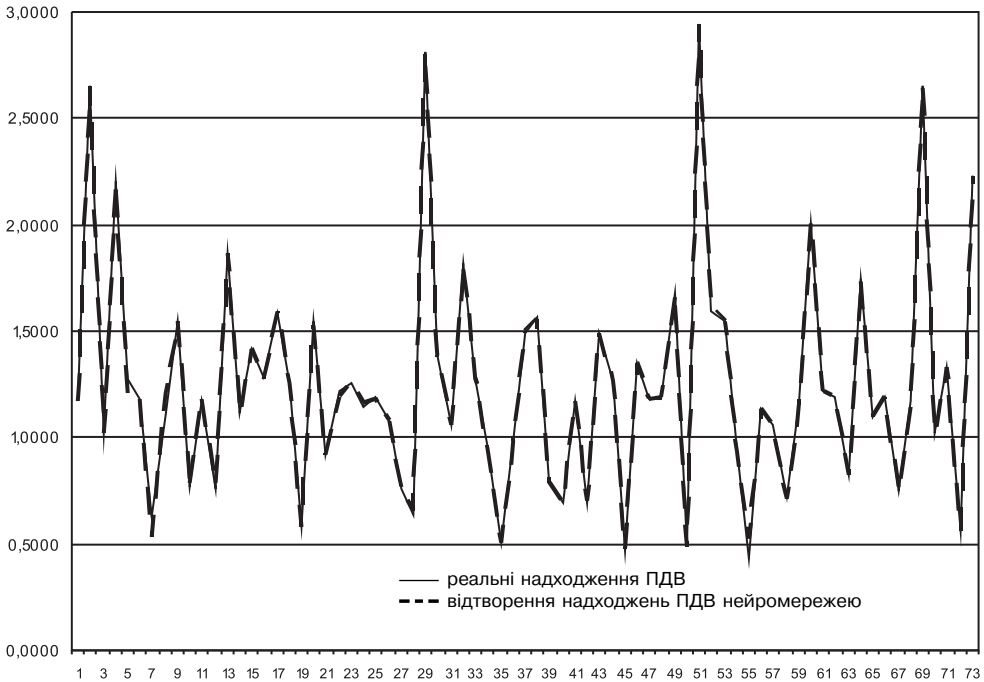


Рис. 2. Відтворення відносних змін надходжень ПДВ для I кварталу із застосуванням нейромережевої моделі

Тепер відтворимо вихідний показник “Надходження ПДВ” із застосуванням нейронних мереж іншої конфігурації. Вже зі структурою мережі у три шари, де на внутрішньому шарі було п’ять нейронів, модель виявила ефект перенавчання — кількість її внутрішніх параметрів перебільшувала обсяг статистичної вибірки. У результаті налаштування моделі помилка прогнозу прямує до нуля, відповідно, модель точно відтворює дані з навчальної вибірки. Але для того, щоб уникнути екстремальної поведінки моделей, доцільно, аби обсяг вибірки перевищував кількість її параметрів, модель не точно апроксимувала вихідний часовий ряд, натомість мала місто незначна похибка навчання. Відповідно, при побудові нейромережевих моделей прогнозування надходжень ПДВ за кожним кварталом відбиралася така нейронна мережа, яка найточніше відтворювала наявну статистику й водночас мала простішу структуру та уникала ефекту перенавчання.

При побудові моделі для II кварталу було використано описану вище структуру нейронної мережі типу “персептрон”: шість нейронів вхідного шару, три нейрони середнього шару та один нейрон на виході моделі. Нормалізована середньоквадратична помилка відтворення показника “Надходження ПДВ” для II кварталу із застосуванням зазначеної моделі склала $\sigma_{norm} = 0,030$ ($\sigma_{err} = 0,013$). Для III та IV кварталу можна було застосувати мережу, на середньому шарі якої чотири нейрони, оскільки статистичних даних тут виявилось більше. У цьому випадку для III кварталу $\sigma_{norm} = 0,096$ ($\sigma_{err} = 0,063$), а для IV — $\sigma_{norm} = 0,124$ ($\sigma_{err} = 0,094$).

Отримані результати відтворення надходжень ПДВ із застосуванням методів нейронних мереж підтверджують високу ефективність цього інструментарію та доцільність його використання для розв’язання завдання прогнозування бюджетних надходжень. Нарешті, після побудови лінійних економетричних та нелінійних нейромережевих моделей покажемо можливість відтворення відносних змін надходжень ПДВ із застосуванням економіко-математичних моделей, побудованих на нечіткій логіці.

Для побудови економіко-математичних моделей прогнозування надходжень ПДВ було застосовано метод Сугено побудови нечітких моделей, який полягає в автоматичній екстракції правил прийняття рішень з наявної статистики. Зауважимо, що для моделей на нечіткій логіці цього типу також є важливим, як і для нейронних мереж, обсяг навчальної вибірки через необхідність виявлення прихованих (латентних) закономірностей у досліджуваних часових рядах, що реалізується в процесі налаштування їх параметрів на наявній статистиці.

Так, навіть за умови побудови моделі типу Сугено з найпростішою структурою (коли кожна змінна описується лише двома лінгвістичними термами, функції належності яких мають по два параметри налаштування) у випадку шести входів та одного виходу кількість параметрів моделі сягає вже 472. Зазначимо, що обсяг статистичних даних по кожному кварталу не перевищував і 100 значень, що унеможлиблює застосування методу Сугено для моделювання надходжень ПДВ на наявній статистичній базі. Відповідно, вже після десяти епох навчання моделі похибка прогнозування наближається до нуля (надзвичайно висока точність відтворення вихідного часового ряду), виявляючи при цьому ефект перенавчання.

Висновки

У статті розкрито сутність наукового дослідження розробки методології та побудови системи відповідних економіко-математичних моделей прогнозування надходжень податку на додану вартість. У результаті проведення модельних експериментів було встановлено, що практично всі побудовані економетричні моделі виявилися нездатними відтворювати надходження ПДВ на основі значень низки різних показників попередніх періодів. Проте економіко-математичні моделі прогнозування надходжень ПДВ, побудовані на базі інструментарію нейронних мереж, продемонстрували високу ефективність та обумовили доцільність їх застосування для розв’язання завдання розподілу бюджетних призначень у територіальному розрізі.

Для побудови економіко-математичних моделей на нечіткій логіці за методом Сугено наявної статистики виявилось недостатньо — сконструйована в подібний спосіб модель із найпростішою конфігурацією одразу виявила ефект перенавчання. Отже, з урахуванням результатів проведеного аналізу модельних експериментів можна зробити висновок, що у випадку, коли обсяг навчальної вибірки є незначним, доречно або скористатися нейронною мережею із максимально спрощеною структурою, або застосувати економіко-математичні моделі на нечіткій логіці типу Мамдані. У разі, якщо серед пояснюючих змінних є якісні показники, або за необхідності врахування експертно встановлених правил прийняття рішень, підхід Мамдані до побудови моделей аналізу та прогнозування розвитку економічних систем є найбільш прийнятним⁵.

Нагадаємо, що нечіткі моделі типу Мамдані здатні працювати навіть без налаштування відповідно до реальних даних — лише базуючись на закладених у них наборах логічних правил та експертно встановлених параметрах функцій належності. Ці моделі мають усі переваги, властиві нейронним мережам, та на відміну від інших методів здатні поєднувати можливість налаштування своїх параметрів на реальних даних із урахуванням при проведенні аналізу як кількісних, так і якісних факторів впливу на обсяги надходжень ПДВ.

⁵ *Матвійчук А.В.* Прогнозування розвитку фінансових показників із використанням апарату нечіткої логіки // *Фінанси України*. — 2006. — № 1. — С. 107—115.