

УДК 330.3:519.86

## ІДЕНТИФІКАЦІЯ МОДЕЛЕЙ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПРОЦЕСІВ ЗАСОБАМИ АФІНІТИВНОГО АНАЛІЗУ

О. Белз

*Львівський національний університет імені Івана Франка,  
Проспект Свободи 18, 79008 Львів, Україна*

*Досліджено вплив способів відбору економічних показників, які входять до економетричної моделі як пояснювальні змінні, на прогностні характеристики таких моделей. Запропоновано для вибору суттєвих факторів моделі застосовувати методи інтелектуального аналізу даних.*

*Ключові слова: ідентифікація моделей, моделювання соціально-економічних процесів, прогнозування значень економічних показників, відбір суттєвих факторів.*

**Постановка проблеми.** Побудова моделей соціально-економічних процесів є необхідною умовою ефективного управління економікою країни. Існує ряд моделей та методів дослідження економічної динаміки, а саме: ряди динаміки, моделі автоматного типу, мережі Петрі тощо. Окрему групу складають методи моделювання багатомірних процесів із застосуванням апарату регресійного аналізу.

**Аналіз останніх досліджень і публікацій та невирішені частини проблеми.** На якість побудованих статистичних моделей вагомий вплив має вибір набору пояснювальних (екзогенних) змінних, які описують процес функціонування досліджуваних об'єктів. Тому процес ідентифікації моделі системи (побудови моделі системи за даними її функціонування) є надзвичайно складним. Дж. Бокс та Г. Дженкінс зазначають, що “питання про те, які типи моделей зустрічаються на практиці і у яких обставинах – це риса поведінки фізичного світу, і вона не може бути розв’язана чисто математичними міркуваннями” [1, с. 193]. С. Наконечний, Т. Терещенко та Т. Романюк виокремлюють такі етапи вибору пояснювальних змінних: формування початкової гіпотези про набір незалежних (пояснювальних) змінних; експертна оцінка цього набору; аналіз взаємозв’язків; добір і звуження кола істотних для моделювання змінних [6, с. 16].

Ю. Толбатов зазначає, що суттєвість впливу факторів на показник можна визначити з використанням  $F$ -статистики. Підбір факторів, які необхідно включити в модель, можна здійснювати поступово, розглядаючи один або декілька факторів, і на кожному етапі перевіряти суттєвість впливу на показник включених у розгляд факторів [9].

В. Т. Доля вказує, що “рішення щодо включення незалежних змінних (факторів) у рівняння регресії приймається за наслідками тестування на значущість (невипадковість) їхнього впливу на залежну змінну за  $t$ -статистикою Стьюдента,

автономність впливу за  $\gamma$ -тестом і достатність вкладу у множинний коефіцієнт детермінації за  $\rho$ -тестом” [2, с. 57].

**Мета дослідження.** Метою цієї роботи є дослідження впливу способів відбору економічних показників, які входять до економетричної моделі як пояснювальні змінні, на прогнозні характеристики таких моделей.

**Вхідними даними** для дослідження є помісячні значення чотирнадцяти макроекономічних показників з січня 2006 р. по серпень 2013 р. (на підставі даних Держкомстату України [8]), а саме таких показників:

1. Обсяг реалізованої промислової продукції (товарів, послуг), млн грн.
2. Обсяг продукції сільського господарства, млн грн.
3. Обсяг роздрібного товарообороту, млн грн.
4. Грошова маса МЗ (залишки коштів на кінець періоду), млн грн.
5. Доходи зведеного бюджету, млн грн.
6. Видатки зведеного бюджету, млн грн.
7. Індекс споживчих цін (ІСЦ) (до попереднього місяця), %.
8. Індекс цін виробників промислової продукції (до попереднього місяця), %.
9. Чисельність громадян, зареєстрованих як безробітні, тис. осіб.
10. Рівень безробіття (на кінець періоду), %.
11. Середньомісячна номінальна заробітна плата працівників, грн.
12. Середньозважена ставка рефінансування за всіма інструментами.
13. Внутрішній кредит (залишки коштів на кінець періоду), млн грн.
14. Депозити резидентів, залучені депозитними корпораціями окрім Національного банку України (залишки коштів на кінець періоду), млн грн.

**Інструментарієм дослідження** є програмні засоби аналітичної платформи Deductor Academic та MS Excel.

**Основні результати дослідження.** У загальному випадку підхід щодо визначення суттєвості впливу на результатний показник окремих факторів з використанням  $F$ -статистики, описаний Ю. Толбатовим у [9], вимагає побудови ряду економетричних моделей на підставі повного перебору всеможливих кандидатів у екзогенні змінні. Так у випадку трьох незалежних змінних необхідно апроксимувати 15 економетричних залежностей, а у випадку десяти незалежних змінних – 1023 економетричних залежностей. Так як у процесі побудови статистичних моделей дослідники оперують зі значно більшими масивами незалежних змінних, підхід, описаний Ю. Толбатовим, у практичних задачах застосувати неможливо. Прийняття рішень щодо включення незалежних змінних у рівняння регресії на підставі  $\rho$ -тестом також вимагає побудови ряду економетричних моделей на підставі повного перебору всеможливих кандидатів у екзогенні змінні [2]. Тому такий підхід є неприйнятним для застосування, коли потрібно оперувати із великими масивами незалежних змінних.

В основі відбору суттєвих факторів за  $t$ -статистикою Стьюдента і  $\gamma$ -тестом лежить розрахунок парних коефіцієнтів кореляції [2]. Такий підхід не вимагає значних обчислювальних ресурсів і може бути застосований у практичних задачах.

Для проведення дослідження у роботі запропоновано порівняти ефективність відбору екзогенних змінних на підставі методів статистичного та інтелектуального аналізу даних.

З метою дослідження прогнозних характеристик побудованих моделей із статистичної вибірки виокремлено дві частини (навчальну та тестову). На підставі

даних навчальної вибірки розраховують коефіцієнти моделей, а на підставі даних тестової вибірки оцінюють якість прогнозування. Навчальна вибірка містить дані з січня 2006 р. по грудень 2011 р., а тестова вибірка – з січня 2012 р. по грудень 2012 р.

Для ідентифікації моделей соціально-економічних процесів методами статистичного аналізу даних використано підхід Сімса, запропонований у 1980 р., який має назву VAR або векторна авторегресія [4]. Сімс запропонував подавати кожен із компонент багатовимірною випадковою процесу як лінійну комбінацію від попередніх значень усіх змінних. Цей підхід передбачає побудову системи рівнянь такого виду:

$$\begin{cases} \tilde{X}_t^1 = \alpha_1 + \beta_{11} \cdot X_{t-1}^1 + \beta_{12} \cdot X_{t-1}^2 + \dots + \beta_{1m} \cdot X_{t-1}^m \\ \tilde{X}_t^2 = \alpha_2 + \beta_{21} \cdot X_{t-1}^1 + \beta_{22} \cdot X_{t-1}^2 + \dots + \beta_{2m} \cdot X_{t-1}^m \\ \dots \\ \tilde{X}_t^m = \alpha_m + \beta_{m1} \cdot X_{t-1}^1 + \beta_{m2} \cdot X_{t-1}^2 + \dots + \beta_{mm} \cdot X_{t-1}^m \end{cases}, \quad (1)$$

де  $\tilde{X}_t^1, \tilde{X}_t^2, \dots, \tilde{X}_t^m$  – теоретичні значення макроекономічних показників у часі  $t$ ;  $X_{t-1}^1, X_{t-1}^2, \dots, X_{t-1}^m$  – фактичні значення макроекономічних показників у часі  $(t-1)$ ;  $\alpha_1, \alpha_2, \dots, \alpha_m$  та  $\beta_{11}, \beta_{12}, \dots, \beta_{mm}$  – коефіцієнти побудованих моделей, розраховані за допомогою методу найменших квадратів;  $m$  – кількість макроекономічних показників.

Оскільки вхідними даними є чотирнадцять макроекономічних показників, то система (1) містить чотирнадцять рівнянь. Для подальшого дослідження запропоновано введення пояснювальних змінних у кожне рівняння системи (1) здійснювати поступово по мірі спадання абсолютних значень парних коефіцієнтів кореляції між фактичним значенням кожного з макроекономічних показників у часі  $(t-1)$  та фактичним значенням обраного макроекономічного показника у часі  $t$ . Як результат для кожного з чотирнадцяти макроекономічних показників побудовано по одній парній лінійній регресійній моделі та по тринадцять множинних лінійних регресійних моделей (усього 196 економетричних моделей).

Для ідентифікації моделей соціально-економічних процесів методами інтелектуального аналізу даних обрано алгоритми афінитивного аналізу – комбінація алгоритму Кохонена та алгоритму G-means, що реалізована в аналітичній платформі Deductor Academic [7]. Як і у випадку статистичного аналізу даних за основу обрано економетричні моделі вигляду (1). Введення пояснювальних змінних у кожне рівняння системи (1) будемо здійснювати поступово по мірі спадання рівня значимості показників. Для визначення рівня значимості показників виконано кластеризацію станів макроекономічного середовища. Стан макроекономічного середовища описують такі атрибути:

- період – інформаційні дані, тобто ідентифікатор стану макроекономічного середовища;
- фактичні значення кожного з макроекономічних показників у часі – вхідні дані;
- фактичні значення обраного макроекономічного показника у часі – результатні дані.



			вибірки)				вибірки)	
1	2	3	4	5	6	7	8	9
<b>Модель залежності обсягу реалізованої промислової продукції від макроекономічних показників</b> (максимальне $R^2 = 0,6074$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,5952; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,3422)								
Статистичний	7	0,4723	-21,1307	-70,3619	2	0,3498	-22,9324	-29,7897
Інтелектуальний	6	0,3094	-8,2347	-10,8477	7	0,3104	-9,7314	-9,2773
<b>Модель залежності обсягу продукції сільського господарства від макроекономічних показників</b> (максимальне $R^2 = 0,5476$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,5438; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,3461)								
Статистичний	6	0,3430	0,3507	0,2849	5	0,3424	0,3418	0,2885
Інтелектуальний	11	0,4603	0,4308	-0,0824	6	0,3322	0,3248	0,2732
<b>Модель залежності обсягу роздрібного товарообороту від макроекономічних показників</b> (максимальне $R^2 = 0,9429$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9335; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,6487)								
Статистичний	9	0,9184	0,4306	-5,6066	2	0,8262	0,4129	-0,5110
Інтелектуальний	11	0,9329	0,4770	-7,8919	8	0,8160	-0,0070	-0,3895
<b>Модель залежності грошової маси М3 від макроекономічних показників</b> (максимальне $R^2 = 0,9976$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9979; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,6825)								
Статистичний	8	0,9974	0,8042	0,9174	1	0,9949	0,7460	0,9372
Інтелектуальний	12	0,9976	0,7707	0,8554	4	0,9958	0,7215	0,9102
<b>Модель залежності доходів зведеного бюджету від макроекономічних показників</b> (максимальне $R^2 = 0,7477$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,8576; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,5712)								
Статистичний	1	0,6665	0,1694	-3,562	1	0,6665	0,1694	-3,562
Інтелектуальний	10	0,7314	0,1766	-8,9621	4	0,6317	-0,2014	0,4463
<b>Модель залежності видатків зведеного бюджету від макроекономічних показників</b> (максимальне $R^2 = 0,8385$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,8159; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,5533)								
Статистичний	14	0,8385	0,3991	-7,7385	1	0,6089	0,2009	-2,1283
Інтелектуальний	11	0,8143	0,5463	-8,140	2	0,5538	-0,1461	-0,0038
<b>Модель залежності індексу споживчих цін від макроекономічних показників</b> (максимальне $R^2 = 0,5068$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,6834; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,3023)								
Статистичний	2	0,4202	-0,1086	-0,2452	3	0,4517	-3,8807	-0,1680
Інтелектуальний	7	0,1456	-1,9742	0,5230	7	0,1456	-1,9742	0,5230
<b>Модель залежності індексу цін виробників промислової продукції від макроекономічних показників</b> (максимальне $R^2 = 0,4632$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,5158; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,2445)								
Статистичний	4	0,3354	-0,0990	-0,2339	2	0,3226	-0,1428	-0,1584
Інтелектуальний	5	0,1295	0,2144	-0,3232	1	0,015	-0,4050	-0,1015
<b>Модель залежності чисельності громадян, зареєстрованих як безробітні, від макроекономічних показників</b> (максимальне $R^2 = 0,9275$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,6399; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,3295)								
Статистичний	2	0,7748	0,1697	-13,7568	2	0,7748	0,1697	-13,7568
Інтелектуальний	10	0,9217	0,3277	-61,3475	6	0,8547	0,2556	-19,7948
<b>Модель залежності рівня безробіття від макроекономічних показників</b> (максимальне $R^2 = 0,9467$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9502; усереднене абсолютних значень								

парних коефіцієнтів кореляції становить 0,5323)								
Статистичний	14	0,9467	0,6389	0,4349	3	0,9027	0,5487	0,7128
Інтелектуальний	8	0,9244	0,7029	0,4640	2	0,8973	0,5598	0,6350
<b>Модель залежності середньомісячної номінальної заробітної плати працівників від макроекономічних показників</b> (максимальне $R^2 = 0,9971$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9989; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,6876)								
Статистичний	4	0,9966	0,9461	0,9684	9	0,9967	0,9246	0,9723
Інтелектуальний	7	0,9968	0,9581	0,9074	11	0,9969	0,9464	0,9324
<b>Модель залежності середньозваженої ставки рефінансування за всіма інструментами від макроекономічних показників</b> (максимальне $R^2 = 0,8607$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9163; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,2478)								
Статистичний	2	0,8182	-1,925	-0,9470	2	0,8182	-1,925	-0,9470
Інтелектуальний	3	0,8086	-2,9202	-1,8323	3	0,8086	-2,9202	-1,8323
<b>Модель залежності внутрішнього кредиту від макроекономічних показників</b> (максимальне $R^2 = 0,9985$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9986; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,6523)								
Статистичний	1	0,9966	0,8505	0,9072	1	0,9966	0,8505	0,9072
Інтелектуальний	9	0,9973	0,8479	0,7561	13	0,9980	0,8241	0,9518
<b>Модель залежності депозитів резидентів, залучених депозитними корпораціями окрім Національного банку України, від макроекономічних показників</b> (максимальне $R^2 = 0,9969$ ; максимальне абсолютне значення парного коефіцієнта кореляції становить 0,9980; усереднене абсолютних значень парних коефіцієнтів кореляції становить 0,6820)								
Статистичний	9	0,9967	0,8806	0,9440	7	0,9961	0,8117	0,9878
Інтелектуальний	12	0,9968	0,8697	0,8712	12	0,9968	0,8697	0,8712

Як видно з табл. 1, висока якість моделі на підставі даних навчальної вибірки не гарантує хороших прогностичних характеристик такої моделі. В економічних дослідженнях зв'язок між пояснювальними (екзогенними) змінними та ендогенною змінною вважається тісним, якщо коефіцієнт множинної детермінації більший за величину 0,7 [3]. Проведене дослідження підтвердило це твердження. Для моделей макроекономічних показників, у яких значення коефіцієнта множинної детермінації  $R^2$  (розраховане на підставі даних навчальної вибірки) було меншим за величину 0,7, не вдалося досягнути прийнятних прогностичних характеристик як на підставі даних тестової, так і на підставі даних контрольної вибірок. Для моделей макроекономічних показників, у яких значення коефіцієнта множинної детермінації модифіковане  $R^{2m}$  (розраховане на підставі даних тестової вибірки) було меншим за величину 0,7, не вдалося досягнути прийнятних прогностичних характеристик на підставі даних контрольної вибірки. Якщо значення коефіцієнта множинної детермінації модифіковане  $R^{2m}$  (розраховане на підставі даних тестової вибірки) було більшим за величину 0,7, то тісний зв'язок між пояснювальними (екзогенними) змінними та ендогенною змінною відслідковувався і на підставі даних контрольної вибірки. Серед моделей чотирнадцяти макроекономічних показників хороші прогностичні характеристики мають моделі, що відповідають найкращим значенням  $R^{2m}$  на даних тестової вибірки як за методами статистичного аналізу даних, так і за методами інтелектуального аналізу даних, а саме: моделі залежності грошової маси М3 від макроекономічних показників; моделі залежності рівня безробіття від макроекономічних показників; моделі залежності середньомісячної номінальної заробітної плати працівників від макроекономічних показників; моделі залежності внутрішнього

кредиту від макроекономічних показників; моделі залежності депозитів резидентів, залучені депозитними корпораціями окрім Національного банку України, від макроекономічних показників. Тому перелічені моделі є коректно ідентифікованими.

Проведемо порівняння ефективності застосування методів статистичного та інтелектуального аналізу даних (див. табл. 2).

Таблиця 2

**Порівняння ефективності застосування методів статистичного та інтелектуального аналізу даних**

Моделі макроекономічних показників	Значення $R^{2m}$ , що відповідає максимальному значенню $R^2$ (на підставі даних навчальної вибірки)	Максимальне значення $R^{2m}$ (на підставі даних тестової вибірки) для моделі, побудованої методами статистичного аналізу даних	Максимальне значення $R^{2m}$ (на підставі даних тестової вибірки) для моделі, побудованої методами інтелектуального аналізу даних	Краща модель досягається на підставі методів	
				статистичного аналізу даних	інтелектуального аналізу даних
Обсягу реалізованої промислової продукції	-103,154	-21,1307	-8,2347		+
Обсягу продукції сільського господарства	0,2127	0,3507	0,4308		+
Обсягу роздрібного товарообороту	-0,0015	0,4306	0,4770		+
Грошової маса М3	0,7526	0,8042	0,7707	+	
Доходів зведеного бюджету	-0,0956	0,1694	0,1766		+
Видатків зведеного бюджету	0,3991	0,3991	0,5463		+
Індексу споживчих цін	-7,4785	-0,1086	-0,1680	+	
Індексу цін виробників промислової продукції	-0,6971	-0,0990	0,2144		+
Чисельності громадян, зареєстрованих як безробітні	0,1140	0,1697	0,3277		+
Рівня безробіття	0,6389	0,6389	0,7029		+
Середньомісячної номінальної заробітної плати працівників	0,8480	0,9461	0,9581		+
Середньозваженої	-6,6682	-1,925	-2,9202	+	

ставки рефінансування за всіма інструментами					
Внутрішнього кредиту	0,7783	0,8505	0,8479	+	
Депозитів резидентів, залучених депозитними корпораціями окрім Національного банку України	0,8603	0,8806	0,8697	+	

Як видно з табл. 2, застосування методів статистичного та інтелектуального аналізу даних у всіх випадках покращує прогностні характеристики моделей. У більшості випадків краща модель досягається на підставі методів інтелектуального аналізу даних.

Для більш повного аналізу було проведено перевірку на наявність мультиколінеарності між пояснювальними змінними за допомогою алгоритму Фаррара-Глобера. Для усунення мультиколінеарності виключено з розгляду вісім макроекономічних показників і залишено такі: доходи зведеного бюджету; видатки зведеного бюджету; індекс споживчих цін; індекс цін виробників промислової продукції; рівень безробіття; внутрішній кредит.

Після проведення необхідних розрахунків отримано такі прогностні характеристики побудованих економетричних моделей на підставі пояснювальних змінних, що не є мультиколінеарними (див. табл. 3).

Таблиця 3

**Прогностні характеристики моделей макроекономічних показників, побудованих на підставі пояснювальних змінних, що не є мультиколінеарними**

Моделі макроекономічних показників	Значення $R^2$ (на підставі даних навчальної вибірки)	Значення $R^{2m}$ (на підставі даних тестової вибірки)	Значення $R^{2m}$ (на підставі даних контрольної вибірки)
Обсягу реалізованої промислової продукції	0,4918	-21,6231	-23,8584
Обсягу продукції сільського господарства	0,3463	0,2098	-0,2184
Обсягу роздрібного товарообороту	0,8509	-0,5692	-1,8647
Грошової маса МЗ	0,9507	-4,2074	-24,3777
Доходів зведеного бюджету	0,6002	-0,7022	-1,3145
Видатків зведеного бюджету	0,7563	0,0338	-1,6373
Індексу споживчих цін	0,4356	-2,2971	-2,319
Індексу цін виробників промислової продукції	0,3408	-0,0484	-0,2741
Чисельності громадян, зареєстрованих як безробітні	0,3957	-0,6357	-82,1071
Рівня безробіття	0,9395	0,3267	0,2890



Середньомісячної номінальної заробітної плати працівників	0,9472	-18,2782	-42,2771
Середньозваженої ставки рефінансування за всіма інструментами	0,2358	-78,9184	-307,269
Внутрішнього кредиту	0,9979	0,8486	0,7913
Депозитів резидентів, залучених депозитними корпораціями окрім Національного банку України	0,9303	-6,8231	-56,8511

Як видно з табл. 3, моделі макроекономічних показників, побудовані на підставі немультіколінеарних пояснювальних змінних, не є коректно ідентифікованими окрім моделі залежності внутрішнього кредиту від макроекономічних показників.

**Висновки.** Підсумовуючи зазначимо, що коректна ідентифікація моделей соціально-економічних процесів є можливою, якщо коефіцієнт множинної детермінації  $R^2$ , розрахований на підставі даних навчальної вибірки моделі, у яку введено всі можливі претенденти на екзогенні змінні, є більший за величину  $0,7$ . У випадку виконання такої умови є доцільним проводити подальший пошук коректно ідентифікованих моделей. Для подальшого вибору суттєвих факторів моделі можна застосовувати як методи статистичного аналізу даних, так і методи інтелектуального аналізу даних.

1. Бокс Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс ; пер. с англ. А. Л. Левшина ; под ред. В. Ф. Писаренко. – [вып. 1]. – М. : Мир, 1974. – 408 с.
2. В. Т. Доля. Економетрія: навч. посібник / В. Т. Доля; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2010. – 171 с.
3. Слейко В. Основи економетрії: У 2 ч. / В. Слейко. – [Частина 1]. – Львів: ТЗОВ "Марка Лтд", 1995. – 192 с.
4. Канторович Г. Г. Анализ временных рядов / Г. Г. Канторович // Экономический журнал ВШЭ. – 2002. – № 4. – С. 498–523.
5. Канторович Г. Г. Анализ временных рядов / Г. Г. Канторович // Экономический журнал ВШЭ. – 2002. – № 2. – С. 251–273.
6. Наконечний С. І. Економетрія: підручник / С. І. Наконечний, Т. О. Терещенко, Т. П. Романюк. – [вид. 2-ге, допов. та перероб.]. – К. : КНУ, 2000. – 296 с.
7. Паклин Н. Б. Бизнес-аналитика: от данных к знаниям (+CD): Учеб. пособие / Н. Б. Паклин, В. И. Орешков. – [2-е изд., перераб. и доп.]. – СПб.: Питер, 2010. – 704 с.
8. Статистичний бюлетень (електронне видання). – Режим доступу: [http://www.bank.gov.ua/control/uk/publish/category?cat\\_id=57897](http://www.bank.gov.ua/control/uk/publish/category?cat_id=57897)
9. Толбатов Ю. А. Економетрика: Підручник для студентів екон. спеціальн. вищ. навч. закл / Ю. А. Толбатов. – К.: Четверта хвиля, 1997. – 320 с.

**IDENTIFYING PATTERNS OF SOCIO-ECONOMIC PROCESSES BY THE  
RESOURCE AFFINITYVNOHO ANALYSIS**

**O. Belz**

*Ivan Franko National University of Lviv,  
Prospekt Svobody 18, UA – 79008, Lviv, Ukraine*

The author discovers problem of influence of methods of selection of economic indicators that are included in the econometric model as explanatory variables on forecast of characteristic of such models. The author suggests to use data mining techniques for choice significant factors of model.

Key words: identification of models, modeling social and economic processes, forecast values of economic indicators, selection of significant factors.

**ИДЕНТИФИКАЦИЯ МОДЕЛЕЙ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ  
ПРОЦЕССОВ СРЕДСТВАМИ АФФИНТИВНОГО АНАЛИЗА**

**А. Белз**

*Львовский национальный университет имени Ивана Франка,  
Проспект Свободы 18, 79008 Львов, Украина*

Исследовано влияние способов отбора экономических показателей, входящих в эконометрической модели как объяснительные переменные, на прогнозные характеристики таких моделей. Предложено для выбора существенных факторов модели применять методы интеллектуального анализа данных.

Ключевые слова: идентификация моделей, моделирование социально-экономических процессов, прогнозирование значений экономических показателей, отбор существенных факторов.