

МОДЕЛЬ СИСТЕМИ ОДНОЧАСНИХ РІВНЯНЬ ІЗ ЛАГОВИМ ЕФЕКТОМ ДЛЯ ОЦІНКИ ЯКОСТІ РЕКЛАМНОЇ КАМПАНІЇ

У цій статті описано більш узагальнену систему одночасних рівнянь для прогнозування рівня продажів залежно від рекламних кампаній за різними каналами та інших факторів. У середовищах RStudio та Google Colab описано створення моделі на основі реальних даних деякого товару. Застосовано специфікаційний тест Хаусмана для визначення методу оцінки моделі. Оскільки показники рекламних кампаній виявились ендогенними змінними, зроблено висновок про важливість використання двокрокового методу найменших квадратів (2МНК). Виявлено, що обсяг реклами є причиною для обсягу продажів за Гренджером, що не можна сказати про зворотне припущення, а саме, причиновість обсягу реклами від продажів за Гренджером. Також визначено «глибину» лагів. Після проведеного дослідження виявили вагомні лаги, тобто по одному лагу для обох каналів реклами. Було оцінено залежності обсягів продажів від різних факторів, у тому числі дистрибуції товару, цінового індексу, впливу реклами та її лагів, вплив рекламної діяльності конкурентів. Коефіцієнти утвореної більш узагальненої системи одночасних рівнянь були оцінені за допомогою двокрокового методу найменших квадратів. Усі статистичні показники свідчать про адекватність моделі. Показники ефективності (ROI – return on investment) рекламних кампаній показали, що реклама і на телебаченні, і в інтернеті є прибутковою для розглянутого товару фірми. Актуальність цієї статті полягає у створенні більш узагальненої системи одночасних рівнянь із включенням моделі прогнозу продажів товару з урахуванням впливу реклами.

Ключові слова: модель системи одочасних рівнянь, лаговий ефект, тест Хаусмана, тест Гренжера, двокроковий метод найменших квадратів.

Вступ

Медіа, як один із важливих рушіїв та посилювачів попиту – сфера, яка потребує якісних досліджень із застосуванням математичних методів. Із появою реклами на ринку сформувався запит на збір даних та оцінку рекламних кампаній. Проведення аналітики дає змогу обирати рекламні канали, на яких та чи інша фірма могла би краще залучати свою цільову аудиторію, виявляти оптимальні дні та час для розміщення рекламних роликів, рівень та архітектуру рекламного тиску, довжину ролику, характер передавання інформації і т.д.

У статті застосовано дані продажів та рекламних кампаній фірми для одного з її товарів, побудовано систему одночасних рівнянь та застосовано до них специфікаційний тест Хаусмана та тест причинності за Гренджером у RStudio та Google Colab. Отримавши висновки з тестів, обсяги продажів було оцінено за допомогою 2МНК у RStudio, оцінку проаналізовано, якість рекламних кампаній перевірено за допомогою ROI. Створено програмний продукт,

який призначений для експериментального дослідження впливу реклами на продаж продукту.

Предметом дослідження у статті є моделювання і оцінка якості впливу рекламних кампаній на продажі товарів.

Метою цієї статті є створення узагальненої системи одночасних рівнянь і застосування до неї тесту Хаусмана задля пошуку залежностей та оцінка моделі залежності обсягів продажів від різних факторів, у тому числі реклами та її лагів, кількість яких визначена тестом Гренжера. Основними завданнями статті є:

- застосувати специфікаційний тест Хаусмана для визначення методу оцінки моделі продажів;
- застосувати метод причиновості за Гренджером, щоб виявити причинові зв'язки та «глибину» лагів реклами;
- побудувати і оцінити модель, виходячи з розробленого загального рівняння, попередньо розділивши дані на навчальну та тестову вибірки;
- порахувати ROI для рекламних кампаній та проаналізувати їх.

Система одночасних рівнянь обсягів продажів та реклами

Якісний аналіз у сфері реклами неможливий без математичного, а у напшому випадку економічного, підходу. Для оцінки впливу реклами пропонуємо систему одночасних рівнянь, що визначатимуть модель.

По-перше, система складається з рівняння, яке визначатиме долю продажів (s_t) серед усіх продажів товарів ринку групою конкуруючих фірм. Доля на ринку – це відсоткове чи абсолютне значення частки продажів конкретної фірми від продажів ринку в цілому:

$$s_t = \frac{S_t}{S},$$

де s_t – доля продажів товару t на ринку,

S_t – сума продажів товару t за даний період,

S – сума продажів усіх товарів категорії (ринку) за даний період.

Використання долі ринку на відміну від використання обсягів продажів зумовлена бажанням зменшити вплив сезонних чинників. Рекламні кампанії часто проводять тоді, коли є більший попит на певну категорію товару. Доля ринку показує, наскільки продажі фірми змінюються щодо інших фірм, що значно згладжує вплив сезонності.

У моделі залежності долі ринку фірми використані такі незалежні змінні:

- **дистрибуція** – відсоток магазинів, торгових точок, у яких можна придбати товар певної фірми;

- **ціновий індекс** – відношення ціни одиниці товару фірми до середньої ціни на товари конкурентів на ринку.

До них долучаються дані про:

- **рекламні кампанії** як фактор з відкладеним ефектом (лагами);

- **інші фактори** – активність конкурентів, святкові дні, температура повітря тощо.

Рівняння долі ринку з урахуванням усіх факторів набуває вигляду:

$$s_t = \alpha_0 + \sum_{i=1}^I \alpha_i d_{i,t} + \alpha_j P_{s,t} + \sum_{k=1}^K \alpha_k M_{k,t} + \sum_{l=1}^L \alpha_l o_{l,t} + e_t,$$

де $d_i, i \in [1, \dots, I]$ – дистрибуції усіх одиниць випуску товару (SKU),

P_s – ціновий індекс товару,

$M_k, k \in [1, \dots, K]$ – рекламні кампанії та їх лаги,

де,

$o_l, l \in [1, \dots, L]$ – інші фактори,

e – похибка.

По-друге, система містить інформацію про проведені рекламні кампанії. Залежна змінна – це обсяги рекламної кампанії, а фактори – доля ринку та ціна на цю рекламну кампанію:

$$M_t = \beta_0 + \beta_1 s_t + \beta_2 P_{M,t} + e_t,$$

де M – вартість рекламна кампанія,

P_M – вартість одиниці рекламної кампанії,

e – похибка.

Отже, отримуємо таку систему одночасних рівнянь обсягів продажів та реклами:

$$\begin{cases} s_t = \alpha_0 + \sum_{i=1}^I \alpha_i d_{i,t} + \alpha_j P_{s,t} + \sum_{k=1}^K \alpha_k M_{k,t} + \sum_{l=1}^L \alpha_l o_{l,t} + e_t; \\ M_{1,t} = \beta_{10} + \beta_{11} s_t + \beta_{12} P_{M,t} + e_{1,t}; \\ \dots \\ M_{k,t} = \beta_{k0} + \beta_{k1} s_t + \beta_{k2} P_{M,t} + e_{k,t}. \end{cases}$$

Варто зазначити, що рівень лагів визначаємо шляхом використання тесту Гренжера, а метод для оцінювання моделі та визначення доречності використання системи рівнянь – за тестом Хаусмана.

Результати застосування моделі для оцінки ефективності реклами

Даними є: обсяг продажів, ціна товару, дистрибуція, ціна на рекламні кампанії товару фірми та конкурентів. Розглянуті дані представлені потижнево у часових рамках із першого тижня 2019 року до вісімнадцятого 2021 року.

Упродовж 2019–2021 років частка ринку товару змінюється від 4% до 7,5% (див. рис. 1). Можемо припустити, відповідно до візуальної оцінки, що наявна залежність між «піками» росту частки ринку та рекламними розміщеннями на ТБ та в інтернеті.

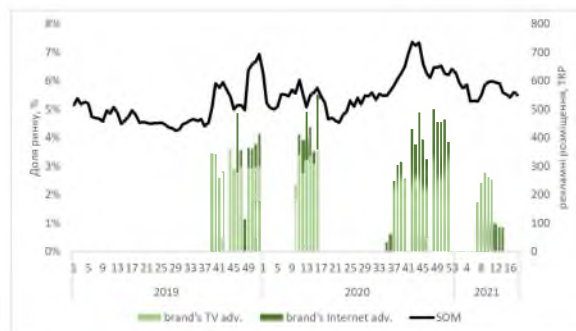


Рис. 1. Частка ринку товару та його рекламні розміщення

0,30

Система попередньо мала такий вигляд:

$$\begin{cases} s_t = & \alpha_0 + \alpha_1 d_{1,t} + \alpha_2 d_{2,t} + \alpha_3 P_{s,t} + \\ & + \alpha_4 M_{TV,t} + \alpha_5 M_{int,t} + \alpha_6 o_{TV,t} + \\ & + \alpha_7 o_{int,t} + e_{1,t}; \\ M_{TV,t} = & \beta_{10} + \beta_{11} s_t + \beta_{12} P_{TV,t} + e_{2,t}; \\ M_{int,t} = & \beta_{20} + \beta_{21} s_t + \beta_{22} P_{int,t} + e_{3,t}; \end{cases}$$

де d_1, d_2 – дистрибуції усіх одиниць випуску товару (SKU),

P_s – ціновий індекс товару,

M_{TV} – обсяги реклами у рекламній кампанії фірми на телебаченні,

M_{int} – обсяги реклами у рекламній кампанії фірми в інтернеті,

o_{TV} – обсяги реклами у рекламній кампанії конкурентів на телебаченні,

o_{int} – обсяги реклами у рекламній кампанії конкурентів в інтернеті,

e_1, e_2, e_3 – похибки.

Ендогенними змінними системи вважаємо частку ринку, рекламну кампанію на телебаченні та в інтернеті.

Застосування тесту Хаусмана

За допомогою **тесту Хаусмана** перевіримо, чи M_t є ендогенною чи екзогенною змінною. Від цього залежить, який метод для оцінки краще використати – метод двокрокового МНК (надалі – 2МНК) чи звичайний МНК.

Виразимо потенційну ендогенну змінну $M_{TV,t}$ та $M_{int,t}$ через друге та третє рівняння у першому рівнянні системи. Після перетворення s_t набуває вигляду:

$$s_t = \alpha'_0 + \alpha'_1 d_{1,t} + \alpha'_2 d_{2,t} + \alpha'_3 P_t + \alpha'_4 P_{TV,t} + \alpha'_5 P_{int,t} + \alpha'_6 o_{TV} + \alpha'_7 o_{int,t} + e'_{1,t}.$$

Отримане рівняння підставили у друге і третє рівняння системи та отримаємо:

$$M_{TV,t} = \beta'_{10} + \beta'_{11} s_t + \beta'_{12} d_{1,t} + \beta'_{13} d_{2,t} + \beta'_{14} P_t + \beta'_{15} P_{TV,t} + \beta'_{16} P_{int,t} + \beta'_{17} o_{TV} + \beta'_{18} o_{int,t} + e'_{2,t};$$

$$M_{int,t} = \beta'_{20} + \beta'_{21} s_t + \beta'_{22} d_{1,t} + \beta'_{23} d_{2,t} + \beta'_{24} P_t + \beta'_{25} P_{TV,t} + \beta'_{26} P_{int,t} + \beta'_{27} o_{TV} + \beta'_{28} o_{int,t} + e'_{3,t}.$$

Оцінимо $M_{TV,t}$ та $M_{int,t}$ за допомогою МНК. Оцінені \widehat{M}_{TV} та \widehat{M}_{int} підставимо у першу модель та застосуємо МНК.

$$s_t = \alpha''_0 + \alpha''_1 d_{1,t} + \alpha''_2 d_{2,t} + \alpha''_3 P_{s,t} + \alpha''_4 M_{TV,t} + \alpha''_5 M_{int,t} + \alpha''_6 s_{TV,t} + \alpha''_7 s_{int,t} + \alpha''_8 \widehat{M}_{TV,t} + \alpha''_9 \widehat{M}_{int} + e_{1,t}.$$

Обчислені **t**-статистики є значущі для \widehat{M}_{TV} та \widehat{M}_{int} за рівня значущості 5%, що підтверджує, що їх не можна відкинути. Як висновок, краще застосувати 2МНК.

Застосування тесту причиновості за Гренжером

Визначимо, чи реклама є причиною за Гренжером для продажів та навпаки. У разі наявності причинових зв'язків визначимо «глибину» лагів для моделі. Для створення моделей із даних обрали: M_{TV}, M_{int}, s_t .

Короткі (S – short) та довгі (L – long) моделі будуть схожого вигляду:

$$L : M_t = \alpha + \sum_{i=1}^I \beta_i M_{t-i} + \sum_{j=1}^J \gamma_j s_{t-j} + \varepsilon_t;$$

$$S : M_t = \alpha + \sum_{i=1}^I \beta_i M_{t-i} + \varepsilon_t.$$

Модель набуває вигляду:

$$L : M_{1,t} = \alpha_1 + \sum_{i=1}^1 \beta_{1,i} M_{t-i} + \sum_{j=1}^1 \gamma_{1,j} s_{t-j} + \varepsilon_{1,t};$$

...

$$M_{k,t} = \alpha_k + \sum_{i=1}^K \beta_{k,i} M_{t-i} + \sum_{j=1}^K \gamma_{k,j} s_{t-j} + \varepsilon_{k,t};$$

$$S : M_{1,t} = \alpha_1 + \sum_{i=1}^1 \beta_{1,i} M_{t-i} + \varepsilon_{1,t};$$

...

$$M_{k,t} = \alpha_k + \sum_{i=1}^K \beta_{k,i} M_{t-i} + \varepsilon_{k,t}.$$

Розглянули моделі з лагом від 1 до 5 для змінної s_t . F-статистика для моделі, починаючи з першої лагової змінної, підтвердила, що обсяг продажів не є причиною для реклами за Гренжером.

Перевірка зворотного твердження (чи є обсяг реклами причиною за Гренжером для обсягу продажів) виявила, що модель із лагом реклами на телебаченні, за рівня значущості 5%, задовольняють перші 7 лагів, за рівня 1% – 3.

F-статистики для моделі з лагом реклами в інтернеті свідчать, що за рівня значущості 5% впливовими є перші 14 лагів, за рівня 1% – 8. Це підтверджує, що обсяг реклами в інтернеті є причиною за Гренжером для обсягів продажів.

На графіках (див. рис. 2) видно, що після першого лагу відбувається зменшення F-статистики більше ніж удвічі для обох каналів

реклами, а далі відбуваються набагато менші зміни. Для моделі взято тільки перші лаги для обох змінних.

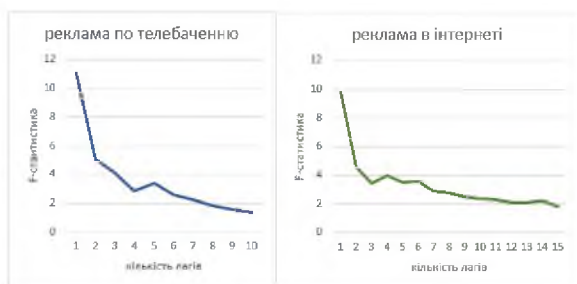


Рис. 2. Графіки значень F-статистик причинової за Гренджером залежно від кількості лагів у моделі

Отже, обсяг реклами є причиною для обсягу продажів за Гренджером, але зворотна причиновість не підтверджується. До розгляду в модель візьмемо залежності обсягів продажів від обсягів реклами та по одному лагу для обох каналів реклами.

Формування моделі та її оцінка за допомогою 2МНК

Узагальнена система рівнянь має такий вигляд:

$$\begin{cases} s_t = \alpha_0 + \alpha_1 d_{1,t} + \alpha_2 d_{2,t} + \alpha_3 P_{s,t} + \\ \quad + \alpha_4 M_{TV,t} + \alpha_5 M_{int,t} + \alpha_6 M_{TV,t-1} + \\ \quad + \alpha_7 M_{int,t-1} + \alpha_8 s_{TV,t} + \\ \quad + \alpha_9 s_{int,t} + e_{1,t}; \\ M_{TV,t} = \beta_{10} + \beta_{11} s_t + \beta_{12} P_{TV,t} + e_{11,t}; \\ M_{int,t} = \beta_{20} + \beta_{21} s_t + \beta_{22} P_{int,t} + e_{21,t}. \end{cases}$$

Застосуємо метод 2МНК до системи одночасних рівнянь та оцінимо коефіцієнти ($\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8, \alpha_9$).

На графіку (див. рис. 3) представлені змодельовані частки ринку товару для тренувального набору даних 2019–2020 років.

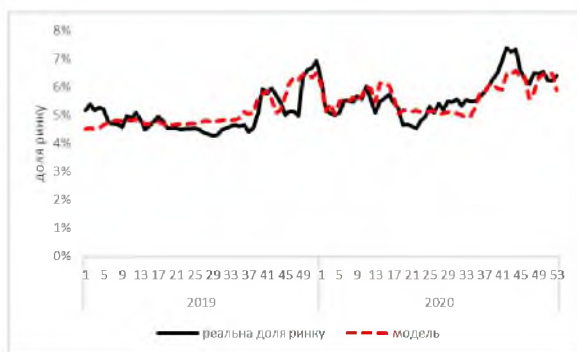


Рис. 3. Реальна та оцінена частки ринку товару

На графіку (див. рис. 4) прогнозування тестових даних представлено модель, що показує хороші результати і достатньо точно описує поведінку реальної частки ринку товару. Середня квадратична похибка становить 0,000002.

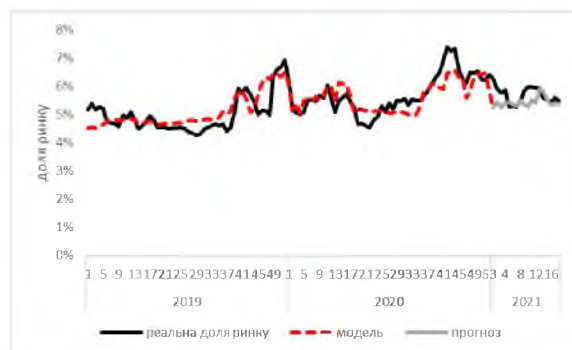


Рис. 4. Прогноз на наступний період

Для рекламних кампаній на телебаченні та в інтернеті відповідні ROI наведено в табл. 1.

Таблиця 1

ROI рекламних кампаній товару на телебаченні та в інтернеті

| Період | ROI ТБ | ROI Інтернет |
|---------|--------|--------------|
| 2019 | 2,04 | 1,34 |
| 2020 | 1,64 | 1,33 |
| 2021 | 1,58 | 0,42 |
| загалом | 1,76 | 1,14 |

Отже,

- майже всі рекламні канали є прибутковими й генерують більший дохід, ніж закладали у їхній бюджет;
- реклама в інтернеті працює гірше, ніж реклама на телебаченні;
- найкращі рекламні кампанії були в 2019 році.

Висновки

Розроблено загальну модель системи одночасних рівнянь – залежності продажів від різних факторів, а саме, від реклами, обсягу продажів та ціни на рекламу.

Застосовано специфікаційний тест Хаусмана для визначення методу оцінки моделі. Як висновок, показники рекламних кампаній виявились ендогенними змінними, що вказує на важливість використання методу 2МНК.

Виявлено, що обсяг реклами є причиною для обсягу продажів за Гренджером, що не можна сказати про зворотне припущення – причиновість обсягу реклами від продажів за Гренджером. Також визначено «глибину» лагів, а саме, по одному лагу для обох каналів реклами.

Оцінено коефіцієнти моделі за допомогою двокрокового методу найменших квадратів. Усі статистичні показники свідчать про адекватність моделі. ROI рекламних кампаній показали, що реклама є прибутковою для розглянутого типу товару фірми.

тність моделі. ROI рекламних кампаній показали, що реклама є прибутковою для розглянутого типу товару фірми.

Список літератури

1. Лук'яненко І., Краснікова Л. Економетрика : підручник. Київ : Знання, 1998. 493 с.
2. Берндт Эрнст Роберт. Практика эконометрики: классика и современность : учебник / пер. с англ. под ред. проф. С.А. Айвазяна. Москва : ЮНИТИ-ДАНА, 2005. 863 с. (Серия «Зарубежный учебник»).
3. Назайкин А. Н. Медиапланирование на 100 процентов. Москва, 2007.
4. Broadbent S. One Way TV Advertisements Work. *Journal of the Market Research Society*. 1979. Vol. 23, No. 3. Pp. 139–166.
5. Leone R. P. Generalizing what is known about temporal aggregation and advert-over. *Marketing Science*. 1995. Vol. 14, issue 3. G141–G150.

References

1. I. Luk'ianenko and L. Krasnikova, *Ekonometryka: pidruchnyk* (Kyiv: Znannia, 1998).
2. Jernst Robert Berndt, *Praktika jekonometriki: klassika i sovremennost': Uchebnik* (Moskva: JuNITI-DANA, 2005).
3. A. N. Nazajkin, *Mediaplanirovanie na 100 procentov* (Moskva, 2007).
4. S. Broadbent, "One Way TV Advertisements Work", *Journal of the Market Research Society*. **23** (3) (1979).
5. R. P. Leone, "Generalizing what is known about temporal aggregation and advert-over", *Marketing Science*. **14**. G141–G150 (1995).

S. Drin, Ye. Reznichenko

A MODEL OF A SYSTEM OF SIMULTANEOUS EQUATIONS WITH A LAG EFFECT FOR ESTIMATING THE QUALITY OF AN ADVERTISING CAMPAIGN

This article describes the creation of a more generalized system of simultaneous equations for forecasting the level of sales depending on advertising campaigns on different channels and other factors. RStudio (R programming language) and Google Colab (Python programming language) environments describe the creation of a model based on real data of a product. The Hausman specification test was applied to determine the model estimation method. As a conclusion, the indicators of advertising campaigns turned out to be endogenous variables, which indicates the importance of using the 2MNK method. It was found that the volume of advertising is the cause of the volume of sales according to Granger, which cannot be said about the reverse assumption - the causality of the volume of advertising from sales according to Granger. The "depth" of lags is also determined, namely, one lag for both advertising channels. The dependence of sales volumes on various factors, including product distribution, the price index, the influence of advertising and its lags, and the influence of competitors' advertising activities, was evaluated. The coefficients of the resulting more generalized system of simultaneous equations were estimated using the two-step least squares method. All statistical indicators testify to the adequacy of the model. Performance indicators (ROI - return on investment) of advertising campaigns showed that advertising both on television and on the Internet is profitable for the company's product in question. The relevance of this article lies in the creation of a more general system of simultaneous equations with the inclusion of a product sales forecast model taking into account the influence of advertising.

Keywords: model of the system of current equations, lag effect, Hausman test, Granger test, two-stage least squares method.

Матеріал надійшов 25.09.2022



Creative Commons Attribution 4.0 International License (CC BY 4.0)