

Józef Hozer*

Uniwersytet Szczeciński

ZMIENNE LOSOWE CZY NIELOSOWE W EKONOMETRII**STRESZCZENIE**

W literaturze ekonometryczno-statystycznej większość teorii budowy modeli ekonometrycznych oparta jest na założeniu nielosowości zmiennych objaśniających. W artykule wykazano, że realniejsze jest założenie, że zmienne objaśniające w modelu ekonometrycznym mogą być zarówno losowe, jak i nielosowe. Przedstawiono reperkusje wynikające z takiego ujęcia problemu.

Słowa kluczowe: zmienne losowe i nielosowe, ekonometria.

W procesie gospodarowania powstaje masa faktów, które mierzymy, gromadzimy dane, a potem „mielimy” je w komputerach dla celów poznawczych. Staramy się, aby to, co się dzieje, analizować, diagnozować i prognozować, używając przy tym metod statystycznych. Kiedy do analiz stosujemy metody ekonometrii, to czynimy założenia, które pozwalają wyciągać stosowne wnioski, na przykład co do własności używanych estymatorów. Jednym z takich założeń jest nielosowość zmiennych objaśniających w analizie wielowymiarowej za pomocą modelu ekonometrycznego. Znaczna część ekonometrii poświęcona jest sytuacji, w której dla zapisu:

$$Y_t = f(X_{1t}, X_{2t}, \dots, X_{kt}, U_t) \quad (1)$$

* Adres e-mail: hozer@wneiz.pl.

obowiązują założenia, że zmienne objaśniające X_{it} są nielosowe (a zmienne Y_t i U_t są losowe) [Pawłowski, 1976, s. 91; Pawłowski, 1980, s. 265; Glodberger, 1972, s. 209]. Wówczas taki model określany jest jako model regresji opisowej. Jeżeli ta regresja ma charakter liniowy, to mówimy o modelu opisowej regresji liniowej. Zbieramy dane statystyczne i szacujemy parametry modelu. Estymatory takiego modelu mają wymagane własności statystyczne i model daje nadzieję na adekwatne wnioskowanie statystyczne, a to oznacza trafne analizy, diagnozy i prognozy. Najczęściej jednak szacujemy parametry modelu bez weryfikacji założenia o nielosowości zmiennych objaśniających. Powstaje pytanie: jakiego charakteru są zjawiska ekonomiczne – losowego czy nielosowego?

W ekonometrii rozważa się przypadek losowych zmiennych objaśniających [Goldberger, 1972, s. 341; Pawłowski, 1976, s. 276; Pawłowski, 1980, s. 109]. Rozważania nad zagadnieniem poprzedzimy określeniem zmiennej losowej zapożyczonym z podręcznika Z. Hellwiga: „Zmienną losową nazywa się taką wielkość, która w wyniku doświadczenia przyjmuje określoną wartość, znaną po zrealizowaniu doświadczenia, a nie dającą się przewidzieć przed realizacją doświadczenia” [Hellwig, 1977, s. 65]. Innymi słowy, zmienna losowa to taka zmienna, której wartości nie możemy poznać przed jej zrealizowaniem. Jednak w ekonomii dane statystyczne nie pochodzą z eksperymentu [Pawłowski, 1980, s. 106]. Dane dotyczą zrealizowanych procesów gospodarczych. Ekonometryk nie zna zatem wartości zmiennych przed realizacją procesu. Nie oznacza to jednak, że wszystkie zmienne są losowe. Jak wykażemy poniżej, zmienna traktowana przez ekonometryka jako losowa dla menadżera zarządzającego i podejmującego decyzję może być nielosowa. Przy kwalifikacji zmiennych w ekonomii przychodzi nam w sukurs rozróżnienie rodzajów związków w ekonomii. Na ogół ekonometrycy są przekonani, że budują modele dla związków przyczynowych. Tymczasem w rzeczywistości oprócz związków przyczynowych realizują się powiązania celowe i współlistnienia. Świadomość tych różnic jest bardzo istotna z punktu widzenia opisu prawidłowości statystycznych związków w ekonomii. Te trzy rodzaje związków są opisane między innymi w pracy [Hozer, Doszyń, 2004]. W pierwszym przypadku realizuje się proces, w którym pytamy, dlaczego zrealizował się fakt Y_t ? W drugim pytamy, po co kreujemy X_t ? Odpowiedzi są następujące. W pierwszym przypadku odpowiedź jest następująca: ponieważ zrealizował się fakt X_{t-1} . W drugim przypadku odpowiedź brzmi: po to,

aby powstał fakt Y_{t+1} . Są to dwie, diametralnie odmienne sytuacje. W pierwszym przypadku taka relacja przyczynowa:

$$Y_t = f(X_{t-1}, \dots) \quad (2)$$

może być identyfikowana i wykorzystywana do prognozowania na jeden okres.

Czy do przewidywania na wiele okresów w przód możemy wykorzystać relację (2)? Owszem, ale wówczas będziemy budować prognozę zjawiska Y , na podstawie prognozy zjawiska X , a to może dawać niezbyt dobre rezultaty (istotne błędy prognoz). W aplikacjach ekonometrycznych przy dłuższych horyzontach prognozy zwykle bywa, że prognozy te obciążone są istotnymi błędami.

W drugim przypadku relacja celowa:

$$Y_{t+1} = f(X_t, \dots) \quad (3)$$

może z założenia być wykorzystywana do podejmowania decyzji w procesie gospodarowania (zarządzania). Oznacza to, że identyfikujemy prawidłowość statystyczną, którą wykorzystujemy w podejmowaniu decyzji. Czy możemy wykorzystać tę relację do przewidywania? Wydaje się, że tak. Zarządzający decyduje o wartości X_t , a realizacja Y_t jest nadzieją matematyczną warunkowego rozkładu Y_t . Okazuje się więc, że relacja celowości jest bliższa schematowi ekonometrycznego modelu opisowego niż relacja przyczynowa. W drugim przypadku X_t jest nielosowe, a w pierwszym X_{t-1} z definicji jest losowe w myśl pojęcia losowości zmiennej.

Kiedy ta sama zmienna może być traktowana jako losowa i nielosowa jednocześnie? Weźmy jako przykład badanie popytu na jakiś towar, który zależy od jego ceny. Dla obserwatora zewnętrznego i popyt, i cena są zmiennymi losowymi. Przed zrealizowaniem obserwator (ekonometryk) nie zna ich wartości, ale sprzedawca (właściciel) ustala cenę i zna tę wartość przed realizacją. Dla obserwatora obie zmienne są losowe. Dla właściciela cena jest zmienną nielosową. Jak widzimy, ta sama zmienna może być jednocześnie losową i nielosową. Jest to poważne utrudnienie epistemologiczne w badaniach ekonometrycznych. Problem losowości ekonometrycy rozwiązują w ten sposób, że rozpatrują tak zwany uogólniony model opisowy [Goldberger, 1972, s. 341; Pawłowski, 1976, s. 276]. Rozpatrywanie uogólnionego modelu opisowego w ekonometrii wydaje się bardziej uzasadnione ze względu na częstość występowania losowych zmiennych w ekonomii. Z tego powodu ekonometrycy często nie przyjmują rygorystycznych założeń co do bada-

nych procesów, zakładając tylko, że badane relacje mają charakter stochastyczny, a w zbiorze zmiennych objaśniających mogą występować zarówno zmienne nielosowe jak i losowe.

Przykłady

Z. Pawłowski przytacza przykład empirycznego modelu kosztów [Pawłowski, 1980, s. 337]:

$$K_C = 0,2120Q + 45,96X_1 + 15,96 X_2 - 74,09 X_3 - 204,56 + U \quad (4)$$

Zmienna K_C (koszty całkowite) jest niewątpliwie zmienną losową. Pozostałe zmienne oznaczające:

Q – produkcja piwa,

X_1 – udział piwa pełnego w całej produkcji browaru,

X_2 – frakcja produkcji słodu zużywanego do produkcji piwa w danym browarze,

X_3 – udział zużycia surowców zastępczych,

są zmiennymi losowymi i nielosowymi: Q jest zmienną losową, a X_1 , X_2 , X_3 – to zmienne, które są bardziej losowe, niż nielosowe.

Jeżeli wartość produkcji nie jest dokładnie znana przed końcem procesu produkcyjnego (a tak zwykle bywa), to udział piwa pełnego w produkcji całego browaru również jest zmienną losową. Tak samo można interpretować charakter zmiennych X_2 i X_3 .

W pracy [Hozer, 1993, s. 29], przytoczono tak zwany model Wolda:

$$P_t = f_1[(D_{t-1} - S_t), P_{t-1}, U_{1t}] \quad (5)$$

$$D_t = f_2(P_t, U_{2t}) \quad (6)$$

$$S_t = f_3(P_{t-1}, U_{3t}) \quad (7)$$

gdzie:

P_t – cena,

D_t – popyt,

S_t – podaż.

Wszystkie zmienne, zarówno objaśniane jak i objaśniające, wydają się mieć charakter losowy.

W tej samej pracy [Hozer, 1993, s. 59], przytoczono hipotezę modelową do badania utargu w jednostce handlowej:

$$Y_t = f(X_{1t}, X_{2t}, X_{3t}, X_{4t}, Q_{kt}, U_t) \quad (8)$$

gdzie:

Y_t – utarg w poszczególnych dekadach,

X_{1t} – czas pracy punktu w godzinach,

X_{2t} – wielkość dostaw,

X_{3t} – liczba personelu,

X_{4t} – powierzchnia handlowa punktu,

Q_{kt} – zmienna zero-jedynkowa, przyjmująca wartość 1 w k -tej dekadzie, a wartość 0 w pozostałych.

Kwalifikując zmienne występujące w tym modelu, możemy uznać, że zmienne $Y_t, X_{1t}, X_{2t}, X_{3t}$ to zmienne losowe, a zmienne X_{4t}, Q_{kt} to zmienne nielosowe. Okazuje się, że nawet taka zmienna, jak liczba personelu obsługi w sklepie, jest trudna do przewidzenia z powodu urlopów (w tym urlopów losowych), zwolnień itp.

W pracy [Hozer, Zawadzki, 1990, s. 121] przytoczono następujący model produkcji globalnej przemysłu włókienniczego w Polsce (Y_t) na tle zatrudnienia robotników (X_{1t}) i zużycia energii (X_{2t}), czasu (T):

$$Y_t = 0,14155X_{1t} + 0,04065 X_{2t} - 0,37382 T - 39,809 \quad (9)$$

Zmienna Y_t jest zmienną losową, natomiast T, X_{1t}, X_{2t} są zmiennymi nielosowymi. Widzimy, że zmienne objaśniające występujące w modelach przyczynowych w literaturze ekonometrycznej raz mają charakter losowy, a innym razem charakter nielosowy. Z tego wynika ważny wniosek, że częściej mamy do czynienia z uogólnionym modelem opisowym niż z klasycznym modelem regresji.

Są to problemy podstaw epistemologii ekonometrycznej, lecz warto się nimi zajmować, aby poprawić jakość zastosowań ekonometrycznych w opisywaniu i prognozowaniu zjawisk ekonomicznych. W artykule rozważamy dwa rodzaje związków, pomijając trzeci, czyli związki współlistnienia, od których zaczęła się teoria ekonometrii. Trzy związki opisane są przykładowo w pracy [Hozer, 1993].

Jakie mogą być reperkusje losowości i nielosowości zmiennych objaśniających?

1. Ta sama zmienna może być losowa i nielosowa jednocześnie, a zależy to od punktu widzenia.
2. Losowe zmienne objaśniające utrudniają wykorzystanie oszacowanych relacji do prognozowania. W takiej sytuacji prognozy opierają się na innych prognozach. Zmienna nielosowa często występuje w związkach celowych, a więc ich wartości są wynikiem podejmowanych decyzji. Realizacja zmiennej objaśnianej, produkcja, popyt są wynikiem kalkulacji decydena.
3. Przyjęcie założenia o losowości zmiennych objaśniających w modelowaniu ekonometrycznym (oprócz nielosowości innych zmiennych objaśniających) powoduje, że procesy szacowania parametrów nie są obciążone tak rygorystycznymi założeniami. Po prostu przyjęte założenia są bardziej realne, mimo że musimy się zadowolić gorszymi własnościami parametrów.
4. Wykorzystanie modeli związków do prognozowania natrafia na bariery. Jedną z nich jest prognozowanie zmiennej objaśnianej na podstawie prognozowania zmiennych objaśniających, co może dawać niezbyt zadowalające efekty, gdy prognozujemy za pomocą modeli związków przyczynowych. Druga bariera polega na wykorzystaniu do prognozowania modeli związków celowych z nielosowymi zmiennymi i wartości które określa decydena.

Być może w praktyce trudności te często powodują duże błędy prognoz, które zniechęcają do stosowania modeli ekonometrycznych do prognozowania.

Literatura

- Goldberger A. (1972), *Teoria ekonometrii*, PWE, Warszawa.
- Hellwig Z. (1977), *Elementy rachunku prawdopodobieństwa i statystyki matematycznej*, PWN, Warszawa.
- Hozer J. (1993), *Mikroekonometria. Analizy, diagnozy, prognozy*, PWE, Warszawa.
- Hozer J., Doszyń M. (2004), *Ekonometria skłonności*, PWE, Warszawa.
- Hozer J., Zawadzki J. (1990), *Zmienna czasowa i jej rola w badaniach ekonometrycznych*, PWN, Warszawa.
- Pawłowski Z. (1976), *Statystyka matematyczna*, PWN, Warszawa.
- Pawłowski Z. (1980), *Ekonometria*, PWN, Warszawa.

RANDOM AND NON-RANDOM VARIABLES IN ECONOMETRICS**Abstract**

In the econometric and statistical literature, most theories regarding constructing econometric models are based on the assumption that the explanatory variables are non-random. The paper shows that it is more realistic to assume that the explanatory variables in the econometric model can be both random and non-random. It also presents repercussions resulting from such an approach to the problem.

Translated by Elwira Zaorska

Keywords: random and non-random variables, econometrics.

KOD JEL: C1.

