

*Stanisława Bartosiewicz\**

## **PROCEDURA „STOPNIOWEGO” KONSTRUOWANIA LINIOWYCH MODELI EKONOMETRYCZNYCH O WIELU ZMIENNYCH OBJAŚNIAJĄCYCH**

Motto:

*Budowa modelu jest jednocześnie sztuką i nauką  
Do stosowania dobrego modelu niezbędne są inspiracja,  
poszukiwania, częściowe potwierdzenia oraz wiele  
źródeł informacji*

Władysław Welfe, Lawrence R. Klein, 1982, s. 16.

### **1. PRZEDMOWA – WSPOMNIENIA**

Było to dawno, dawno temu. Być może odpowiednią datą jest rok 1950, kiedy poznałam profesora Władysława Welfe na kursie ekonomistów organizowanym przez Ministerstwo (chyba) Szkolnictwa Wyższego początkowo w Cieplicach na Dolnym Śląsku, a potem w Przesiece, dokąd zesłano sekcję statystyków, zresztą ku ich ogólnemu zadowoleniu. Zadowolenie to wynikało z powodu braku jakiegokolwiek kontroli ze strony urzędniczej ministerstwa, a także ze względu na przeżycie prawdziwego „komunizmu” w sensie zaopatrzenia w żywność naszego zespołu. Po pierwszym posiłku, kiedy to wszystko będące na stole znikło, po następnych półmiski schodziły ze stołu z dużymi resztkami. Przeżywaliśmy też tam ucztę duchową (co ma niewątpliwie znacznie większe znaczenie), albowiem szkolili nas wspaniali nauczyciele, tacy jak np. profesorowie Stefan Szulc, Edward Rosset i inni. Profesor Władysław Welfe (wtedy piękny i młody), jako niezwykle zdolny i pracowity szybko kroczył po drodze kariery naukowej, podczas gdy ja dreptałam ku wiedzy wolno, również z powodu wrodzonego lenistwa. Zdarzyło się więc nie raz tak, że młodszy od mnie o siedem długich lat dzisiejszy osiemdziesięciolatek stawał się moim mistrzem recenzując prace zbiorowe, w których występowałam jako współautor i zawsze dając mi więcej niż pozytywne recenzje. Stąd wnoszę, że Profesor Władysław Welfe

---

\* Prof. dr hab., Akademia Ekonomiczna im. Oskara Langego we Wrocławiu.

to mój przyjaciel, którego podziwiam przede wszystkim za stworzenie prawdziwej Szkoły Naukowej Makroekonomii. Potem spotykaliśmy się sporadycznie przy okazji łódzkich konferencji związanych z tematami finansowanymi przez poprzedników KBN, na spotkaniach z okazji okrągłych rocznic dzisiejszego Jubilata, a także na konferencjach naukowych organizowanych co roku w Zakopanem przez nieodżałowanej pamięci profesora Aleksandra Zeliasia, gdzie słuchaliśmy wzajemnie swoich wystąpień i mieliśmy okazję do przyjacielskich rozmów głównie przy stole podczas uroczystych kolacji. Myślę przeto, że znamy się dobrze, że Profesor czuje moją sympatię, szacunek, a nawet uwielbienie, które to wyznaczenie upoważnia mnie do zakończenia tych trochę przydługich wspomnień (jak na naukowy a może tylko „naukowy” artykuł).

## 2. PROCEDURA „STOPNIOWEGO” KONSTRUOWANIA MODELI EKONOMETRYCZNYCH

### 2.1. Wprowadzenie

Metodę zasygnalizowaną tytułem rozdziału 2 opisałam w książeczce pt. *Ekonometria z przymrużeniem oka* (2005), którą spełniłam zobowiązanie złożone trzydzieści, a może nawet czterdzieści lat temu świeżo upieczonym magistrom, uczestnikom prowadzonego przeze mnie seminarium. Obietnica ta brzmiała mniej więcej tak: jak pójdę na emeryturę, to napiszę książkę pt. *Ekonometria na wesoło*. Pijąc z nimi tradycyjną wtedy szklankę wody rysowałam na serwetkach papierowych rysunki, z których jeden tylko zapamiętałam i wykorzystałam w zaprezentowanej wyżej książeczce (to ten, na którym ekonometryk usiłuje dopchać do przedstawionego na korelacyjnym rysunku zbioru danych jedną nietypową informację, która odbiega od pozostałych znacznie psując ich „elipsoidalny” rozrzut).

### 2.2. Opis procedury

Oto kolejne etapy procedury „stopniowego” konstruowania modelu:

- określić zmienną objaśnianą oraz korzystając z wiedzy ekonomicznej i intuicji naukowej wybrać zespół zmiennych objaśniających;
- z natury rzeczy (wobec wszechzwiązku zjawisk - jednej z zasad materializmu dialektycznego) znajdują się w tym zespole zmienne ze sobą mocno, średnio lub słabo skorelowane, użycie przeto ich wszystkich może sprawić, że model nie będzie sensowny, tzn. niektóre z ocen parametrów mogą wskazać, że np. zmiana ich dotycząca wpływa na zmienną objaśnianą ujemnie, chociaż zgodnie z naszą wiedzą powinna działać dodatnio. Następnym krokiem w procedurze jest

więc wybór zmiennych objaśniających takich, które są względem siebie *quasi*-ortogonalne, tzn. słabo skorelowane. W tym celu należy opracować macierz współczynników korelacji i zastosować jedną ze znanych metod usuwania zmiennych zbyt mocno z innymi skorelowanych. Można też postąpić inaczej, a mianowicie oszacować model ze wszystkimi zmiennymi objaśniającymi pochodzącymi z pierwotnego zespołu i jako ostateczny zespół potraktować te, dla których oceny parametrów są istotne na obranym poziomie istotności;

- wybrane zmienne objaśniające uporządkować według współczynników korelacji wiążących zmienne objaśniające ze zmienną objaśnianą tak, że pierwsze miejsce w porządku otrzymuje zmienna o najwyższym współczynniku korelacji, a ostatnie ta o najniższym;

- oszacować model uzależniający zmienną objaśnianą z pierwszą zmienną objaśniającą i policzyć reszty ( $e_1$ );

- oszacować model uzależniający reszty ( $e_1$ ) z następną (występującą w opisanym wyżej porządku) zmienną objaśniającą i znowu policzyć reszty ( $e_2$ ); i tak dalej do ostatniej zmiennej;

- dodać do siebie oszacowane cząstkowe modele, uzyskując w ten sposób ostateczną konstrukcję liniowego modelu ekonometrycznego.

Uwaga: po każdym kroku szacującym cząstkowe modele sprawdzić należy, czy ocena parametru kierunkowego jest istotna na obranym poziomie istotności. Jeżeli wystąpi nieistotność, należy opuścić tę zmienną objaśniającą i przejść do następnej z ich ustalonego porządku.

### 3. PRZYKŁAD

A) Zmienna objaśniana to produkcja dodana budownictwa w województwach Polski w roku 2004 (symbol zmiennej  $Y$ ). Jako wstępny zbiór zmiennych objaśniających przyjęto następujące:

- $X_{t-1}$  wartość inwestycji w budownictwie w roku 2003 (zmienna opóźniona o rok, bo inwestycje działać powinny z pewnym opóźnieniem) w cenach bieżących w milionach zł;
- $X_2$  wartość środków trwałych w budownictwie w cenach bieżących w milionach zł;
- $X_3$  liczba pracujących w budownictwie na koniec 2004 roku w tysiącach osób;
- $X_4$  średnia płaca w budownictwie w roku 2004 w zł;
- $X_5$  wskaźnik urbanizacji województwa (% ludności miejskiej).

Ma więc to być funkcja produkcji dodanej objaśniania przez wartość inwestycji, wartość środków trwałych, wielkość zatrudnienia w budownictwie, rozmiar średniej płacy. Zmienna określająca stopień urbanizacji została dodana,

moja intuicja bowiem podpowiedziała mi, iż ten stopień może mieć niejaki wpływ na wielkość produkcji budownictwa (w mieście buduje się więcej).

Dane pochodzą z roczników statystycznych GUS'u z roku 2004 i dla  $X_{t-1}$  z roku 2003 z działu „Ważniejsze dane o województwach”. Zwalnia to mnie od ich przytaczania, jakkolwiek niektóre z nich powstały przez dokonanie na nich prostych operacji arytmetycznych (trzeba było bowiem zastąpić procentowe udziały wielkościami wyrażonymi w złotych).

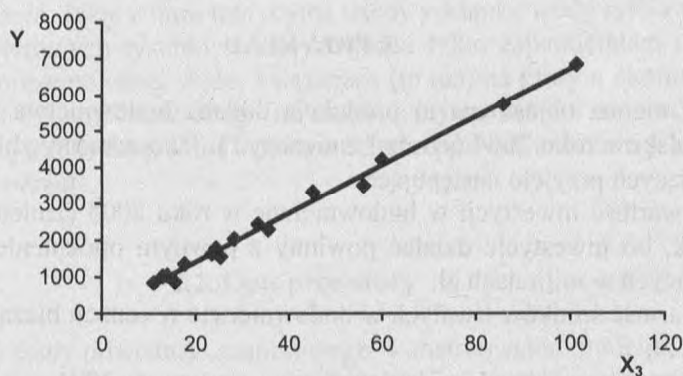
B) Ostateczny zespół zmiennych objaśniających wyznaczyłam metodą badania istotności ocen parametrów, szacując model dla wszystkich pierwotnie obranych zmiennych. W modelu tym istotne oceny parametrów wystąpiły jedynie przy zmiennych  $X_3$ ,  $X_4$  i  $X_5$ .

C) Współczynniki korelacji ze zmienną objaśnianą  $Y$  wskazały, że kolejność zmiennych objaśniających dla „stopniowego” konstruowania modelu jest następująca:  $X_3$ ,  $X_4$ ,  $X_5$ .

D) Zgodnie z podaną procedurą powstaje pierwszy model:

$$Y = 76,635 + 66,567 X_3 + e_1 \quad (1)$$

Ocena parametru kierunkowego przy  $X_3$  jest istotna na poziomie istotności 0,10, współczynnik determinacji  $R^2 = 0,992$ . Poniższy rysunek przedstawia tę zależność.

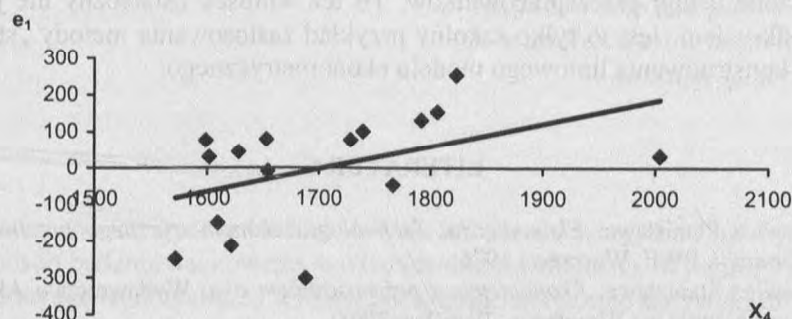


Rys. 1.

E) Model uzależniający reszty  $e_1$  od  $X_4$  jest następujący:

$$e_1 = -1004,2 + 0,674 X_4 + e_2 \quad (2)$$

Ocena parametru kierunkowego przy zmiennej  $X_4$  istotna na poziomie 0,10.  $R^2 = 0,209$ , dopasowanie modelu jest więc słabe, ale to nic dziwnego i złego, bo przecież w resztach mieści się także składnik losowy. Rys. 2 przedstawia związek opisany równaniem (2).



Rys. 2.

F) Uzależnienie reszt  $e_2$  od zmiennej  $X_5$  nie dało dobrych rezultatów bowiem ocena parametru kierunkowego przy tej zmiennej nie była istotna. W związku z tym procedura „stopniowego” konstruowania modelu liniowego zakończona została na drugim kroku i w rezultacie po dodaniu obu otrzymanych modeli (wzory 1 i 2) otrzymujemy taką postać funkcji produkcji dodanej w budownictwie:

$$Y = -927,565 + 66,567 X_3 + 0,674 X_4 + u \quad (3)$$

Model jest dobry, bo  $R^2 = 0,994$  (dopasowanie modelu do danych poprawiło się zatem nieco w porównaniu z modelem (1), a istotność ocen parametrów została w kolejnych krokach procedury sprawdzona). Ciąg reszt ( $u$ ) spełnia warunki symetrii, albowiem jest dziewięć reszt dodatnich a siedem ujemnych (Sprawdzenia hipotezy o symetrii dokonano opracowanym przeze mnie testem dwumianowym dla małej próby opublikowanym w książce mego autorstwa pt. *Ekonometria. Technologia ekonometrycznego przetwarzania informacji*, 1976, s. 145). Reszty ( $u$ ) uporządkowane według wiodącej zmiennej  $X_3$  nie wykazują też długich serii o jednakowym znaku (najdłuższa seria jest trzy-elementowa), a więc są losowe. Sprawdziłam też skonstruowaną prze mego mistrza profesora

Jana Falewicza miarę wyrazistości modelu  $W = \text{odchylenie standardowe resztu} / \text{podzielone przez średnią arytmetyczną zmienną } Y$ ). Piszę o tym mierniku w cytowanej parę wierszy wyżej książce na s. 141–142. Tu podam jedynie informację, że im ten miernik ma mniejszą wartość, tym model jest wyrazistszy. W omawianym przykładzie  $W = 0,053$  jest więc zadowalający.

Ostateczny merytoryczny wniosek to: **produkcja dodana budownictwa w województwach zależy od liczby zatrudnionych i ich średniej płacy**. Ten wniosek trzeba przyjąć z przymrużeniem oka jako dość oczywisty, bo przecież produkcja dodana jest tworzona przez pracę żywą, a wydajność tej ostatniej jest psychologicznie uzasadniona wysokością wynagrodzenia uzyskiwanego za świadczone usługi przez pracowników. To też wniosek ostateczny nie jest jakimś odkryciem. Jest to tylko szkolny przykład zastosowania metody „stopniowego” konstruowania liniowego modelu ekonometrycznego.

#### LITERATURA

- Bartosiewicz Stanisława: *Ekonometria. Technologia ekonometrycznego przetwarzania informacji*, PWE, Warszawa 1976;  
Bartosiewicz Stanisława: *Ekonometria z przymrużeniem oka*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2005;  
Klein Lawrence R.: *Wykłady z ekonometrii*, PWE 1982, cytat pochodzi z punktu 1.2 pt. *Model makroekonomiczny – nowe podejście*, autorstwa Władysława Welfe.

Stanisława Bartosiewicz

#### “STEP BY STEP” PROCEDURE OF CONSTRUCTING LINEAR ECONOMETRIC MODELS WITH MANY ENDOGENOUS VARIABLES

The “step by step” procedure of constructing linear econometric models with many endogenous variables presented in the article consists in arranging them according to their importance measured with the strength of the relationship with the exogenous variable and the follow-on connection 1. of the first in the established succession endogenous variable with the exogenous variable and calculating the model residuals, so that 2. they would be model-connected with the next variable and once again calculate the residuals of the model received. One deals with them as with the residuals in point 2 and one goes on with the procedure until the set of endogenous variables is exhausted. The final model is received by adding all partial models. This procedure is explained in terms of an empirical example.