

Modele panelowe w analizach sektorowych na przykładzie działań przetwórstwa przemysłowego

Prof. dr hab. Wacława Starzyńska
Uniwersytet Marii Curie-Skłodowskiej
w Lublinie

Dr hab. Maria Magdalena Grzelak
Uniwersytet Łódzki

We współczesnych badaniach społecznych i ekonomicznych zauważalny jest wzrost zapotrzebowania na nowoczesne metody analityczne. Należą do nich metody analiz danych czasowo-przekrojowych oraz odpowiednie modele ekonometryczne.

Szczególną klasę danych łączących wymiar czasowy i przekrojowy stanowią dane panelowe.

- Celem referatu jest próba wykorzystania modeli panelowych do analizy konkurencyjności działów przetwórstwa przemysłowego w Polsce w latach 1999 - 2011.

Dane panelowe w modelowaniu ekonometrycznym

Dane czasowo-przekrojowe powstają z połączenia w jeden zbiór szeregów czasowych obserwacji, pochodzących z różnych obiektów.

Obiektami, czyli jednostkami przekroju, mogą być:

- ❑ państwa będące członkami tej samej organizacji gospodarczej,
- ❑ regiony administracyjno-gospodarcze w obrębie jednego lub kilku państw,
- ❑ powiązane ze sobą makroregiony gospodarcze,
- ❑ ale także przedsiębiorstwa lub działy przemysłu.

Obiekty takie funkcjonują w ścisłym powiązaniu ze sobą, z drugiej jednak strony pozostają one odrębnymi jednostkami. Na zachowanie każdego obiektu wpływają zarówno czynniki indywidualne, wynikające wyłącznie z jego specyfiki, działające na funkcjonowanie tylko tego jednego obiektu, jak i inne czynniki, determinujące działalność wszystkich obiektów w jednakowy sposób i w jednakowym stopniu. Obecność tych drugich, „wspólnych” czynników, uzasadnia łączenie obserwacji pochodzących z różnych obiektów w jeden szereg i prowadzenie analiz ekonometrycznych na podstawie połączonych danych.

Dane panelowe stanowią szczególną podgrupę danych, łączących wymiar czasowy i przekrojowy.

Od typowych danych czasowo-przekrojowych różnią się tym, że liczba obserwowanych obiektów N jest duża w stosunku do liczby punktów w czasie T .

Modele ekonometryczne estymowane w oparciu o dane panelowe, nazywane modelami panelowymi, są zwykle zorientowane na analizę przekrojową, a zadaniem ich jest wyodrębnienie różnic pomiędzy obiektami, które są nierozdzielnie związane z czynnikami specyficznymi dla poszczególnych obiektów. Heterogeniczność obiektów jest więc integralną częścią lub nawet centralnym zagadnieniem analizy.

- W modelach ekonometrycznych, szacowanych na podstawie danych panelowych zakłada się z reguły, że na kształtowanie się zmiennej objaśnianej wpływają, oprócz zmiennych objaśniających, niemierzalne, stałe w czasie i specyficzne dla danego obiektu czynniki, zwane efektami grupowymi i/lub stałe względem obiektów specyficzne dla danego okresu czynniki, zwane efektami czasowymi. Uwzględnienie w modelach panelowych efektów grupowych oraz czasowych powoduje konieczność zastosowania specyficznych metod estymacji.

Zalety wykorzystania danych panelowych w praktyce

- wzrost liczby stopni swobody i redukcja problemu współliniowości danych;
- ułatwienie identyfikowalności modeli ekonomicznych i możliwość dokonania wyboru między konkurencyjnymi hipotezami ekonomicznymi;
- umożliwienie eliminacji lub redukcji obciążenia estymatorów.

Charakterystyka badania

Głównym celem badania empirycznego jest ilościowa ocena wpływu czynników produkcji na konkurencyjność działów przetwórstwa przemysłowego w Polsce w latach 1999-2011.

Miernikami konkurencyjności, w zależności od modelu, są:

- wartość dodana brutto,
- produkcja sprzedana,
- wydajność pracy.

Badanie opiera się na analizie funkcji produkcji Cobb-Douglasa poszerzonej o zmienne opisujące innowacje (wewnętrzne nakłady na działalność badawczo-rozwojową lub alternatywnie nakłady na działalność innowacyjną).

Źródło danych

W badaniu wykorzystano publikowane dane statystyczne Głównego Urzędu Statystycznego na poziomie działów przetwórstwa przemysłowego w Polsce w latach 1999-2011.

Jednak, ze względu na zmiany w układzie Polskiej Klasyfikacji Działalności (PKD), badany okres podzielono na dwa podokresy 1999-2007 (PKD 2004) i 2008-2011 (PKD 2007).

Zgodnie z PKD 2004 badanie w latach 1999-2007 obejmuje następujące działy przetwórstwa przemysłowego (sekcja D):

PKD 2004	Nazwa działu
15	Produkcja artykułów spożywczych i napojów
16	Produkcja wyrobów tytoniowych
17	Włókiennictwo
18	Produkcja odzieży i wyrobów futrzarskich
19	Produkcja skór wyprawionych i wyrobów ze skór wyprawionych
20	Produkcja drewna i wyrobów z drewna oraz z korka (z wyłączeniem mebli), wyrobów ze słomy i materiałów używanych do wypalania
21	Produkcja masy włóknistej, papieru oraz wyrobów z papieru

22	Działalność wydawnicza, poligrafia i reprodukcja zapisanych nośników informacji
23	Wytwarzanie koksu, produktów rafinacji ropy naftowej i paliw jądrowych
24	Produkcja wyrobów chemicznych
25	Produkcja wyrobów gumowych i z tworzyw sztucznych
26	Produkcja wyrobów z pozostałych surowców niemetalicznych
27	Produkcja metali

28	Produkcja metalowych wyrobów gotowych, z wyłączeniem maszyn i urządzeń
29	Produkcja maszyn i urządzeń, gdzie indziej niesklasyfikowana
30	Produkcja maszyn biurowych i komputerów
31	Produkcja maszyn i aparatury elektrycznej, gdzie indziej niesklasyfikowana
32	Produkcja sprzętu i urządzeń radiowych, telewizyjnych i telekomunikacyjnych

33	Produkcja instrumentów medycznych, precyzyjnych i optycznych, zegarów i zegarków
34	Produkcja pojazdów samochodowych, przyczep i naczep
35	Produkcja pozostałego sprzętu transportowego
36	Produkcja mebli; działalność produkcyjna, gdzie indziej niesklasyfikowana
37	Przetwarzanie odpadów
	<i>Źródło: opracowanie własne na podstawie Rocznika statystycznego przemysłu 2008, GUS.</i>

W okresie 2008-2011, zgodnie z PKD 2007,
badaniem objęto następujące działy
przetwórstwa przemysłowego (sekcja C):

PKD 2007	Nazwa działu
10	Produkcja artykułów spożywczych
11	Produkcja napojów
12	Produkcja wyrobów tytoniowych
13	Produkcja wyrobów tekstylnych
14	Produkcja odzieży
15	Produkcja skór i wyrobów ze skór wyprawionych

16	Produkcja wyrobów z drewna oraz korka, z wyłączeniem mebli; produkcja wyrobów ze słomy i materiałów używanych do wyplatania
17	Produkcja papieru i wyrobów z papieru
18	Poligrafia i reprodukcja zapisanych nośników informacji
19	Wytwarzanie i przetwarzanie koksu i produktów rafinacji ropy naftowej
20	Produkcja chemikaliów i wyrobów chemicznych
21	Produkcja podstawowych substancji farmaceutycznych oraz leków i pozostałych wyrobów farmaceutycznych
22	Produkcja wyrobów z gumy i tworzyw sztucznych

23	Produkcja wyrobów z pozostałych mineralnych surowców niemetalicznych
24	Produkcja metali
25	Produkcja metalowych wyrobów gotowych, z wyłączeniem maszyn i urządzeń
26	Produkcja komputerów, wyrobów elektronicznych i optycznych
27	Produkcja urządzeń elektrycznych
28	Produkcja maszyn i urządzeń, gdzie indziej niesklasyfikowana
29	Produkcja pojazdów samochodowych, przyczep i naczep, z wyłączeniem motocykli

30	Produkcja pozostałego sprzętu transportowego
31	Produkcja mebli
32	Pozostała produkcja wyrobów
33	Naprawa, konserwacja i instalowanie maszyn i urządzeń
	<i>Źródło: opracowanie własne na podstawie Rocznika statystycznego przemysłu 2012, GUS.</i>

Specyfikacja modeli

Modele panelowe mogą mieć postać: modeli z dekompozycją wyrazu wolnego (FEM - *Fixed Effects Model*) lub modeli z dekompozycją składnika losowego (REM - *Random Effects Model*), przy czym dekompozycja może uwzględniać tylko jeden czynnik (modele jednoczynnikowe) lub dwa czynniki równocześnie (modele dwuczynnikowe).

Modele FEM i REM można ogólnie zapisać następująco:

$$y_{it} = m_i + bx_{it} + e_{it}$$

gdzie: m_i – ogólny wyraz wolny,
 b – parametr strukturalny wyrażający wpływ zmiennej objaśniającej X ,
 x_{it} – realizacja zmiennej objaśniającej dla i -tego obiektu w t -tym okresie,
 e_{it} – reszty, spełniające klasyczne założenia: $E(e_{it}) = 0$ i $\text{Var}(e_{it}) = S_e^2$

Dla uproszczenia wykorzystano modele z jedną zmienną objaśniającą, niemniej jednak modele te mogą mieć oczywiście postać wielozmiennowych.

W modelu FEM m_i jest dekomponowany w wyrazy wolne (stałe) dla poszczególnych grup oddzielnie. Model ma zatem postać [Sucheckki 2000]:

$$y_{it} = a_1 d_{1it} + a_2 d_{2it} + \dots + a_k d_{kit} + b x_{it} + e_{it} = a_i + b x_{it} + e_{it}$$

gdzie: a_i – specyficzne wyrazy wolne, zaś d_j zmienne zero-jedynkowe, przyjmujące wartość 1, gdy $j = i$.

W modelu REM m_i wyraża specyficzne składniki losowe. Model ten można zapisać następująco [Green 2008]:

$$y_{it} = a + bx_{it} + e_{it} + u_i$$

gdzie: $E(u_i) = 0$, $\text{Var}(u_i) = \sigma_e^2$ $\text{Cov}(e_{it}, u_i) = 0$.

Wybór między modelem FEM i REM dokonywany jest z wykorzystaniem testu Hausmana (przy $p < 0,05$ model FEM uznawany jest za bardziej wiarygodny niż REM) [Hausman 1978; Hausman, Taylor 1981].

Wyniki empiryczne modelu wartości dodanej brutto działań przetwórstwa przemysłowego

W badaniu uwzględniono następujące funkcje wartości dodanej:

model 1

$$\ln Wdodbr_{it} = a_i + b_0 + b_1 \ln Zatr_{it} + b_2 \ln Ninw_{it} + b_3 \ln B + R_{it} + b_4 t$$

model 2

$$\ln Wdodbr_{it} = a_i + b_0 + b_1 \ln Zatr_{it} + b_2 \ln Ninw_{it} + b_3 \ln Innow_{it} + b_4 t$$

gdzie:

$\text{LnWdodbr}(it)$ - logarytm naturalny wartości dodanej brutto w cenach stałych w mln zł z 2005 roku (do urealnienia danych użyto wskaźnika cen PKB GUS) dla i -tego działu przemysłu w roku t ;

$\text{LnZatr}(it)$ - logarytm naturalny przeciętnego zatrudnienia w tysiącach osób dla i -tego działu przemysłu w roku t ;

$\text{LnNinw}(it)$ -logarytm naturalny nakładów inwestycyjnych w cenach stałych w mln zł z 2005 roku (do urealnienia danych użyto wskaźnika cen inwestycyjnych GUS);

$\text{LnB+R}(it)$ - logarytm naturalny wartości nakładów na B+R w cenach stałych w mln zł z 2005 roku (do urealnienia użyto wskaźnika cen PKB GUS);

$\text{LnInnow}(it)$ - logarytm naturalny wartości nakładów na innowacje w cenach stałych w mln zł z 2005 roku (do urealnienia danych użyto wskaźnika cen inwestycyjnych GUS);

t –zmienna czasowa, przyjmująca wartość 1 w roku 1999, rosnąca o 1 rocznie;
 α specyficzny wyraz wolny, stały w czasie dla danego działu, zmienny między działami (tzw. efekt indywidualny).

Wyniki estymacji parametrów modeli jednoczynnikowych opisujących kształtowanie się wartości dodanej brutto w przetwórstwie przemysłowym w latach 1999-2007 (LnWdodbr) - model 1

Zmienna	współczynnik	<i>t</i>	<i>p</i>	współczynnik	<i>t</i>	<i>p</i>
	model (1.1.a)			model (1.1.b)		
<i>LnZatr</i>	0,392	2,480	0,0142	0,427	2,842	0,0050
<i>lnNinw</i>	0,132	2,138	0,0340	0,131	2,123	0,0353
<i>LnB+R</i>	0,242	0,741	0,4600	-	-	-
<i>t</i>	0,352	5,163	0,0000	0,353	5,188	0,0000
<i>Stała</i>	-	-	-	-	-	-
R^2	0,9063			0,9060		
Ocena istotności efektów grupowych	LRT= 142,983, p=0,0000; F= 12,282, p=0,0000			LRT=144,265, p=0,0000; F=12,529, p=0,0000		

t - wartość statystyki t-Studenta, R^2 – współczynnik determinacji, *LRT* – statystyka testu LRT (*Likelihood Ratio Test*), *F* – statystyka testu Fishera-Snedecora, *p* – prawdopodobieństwo testowe ($p \in [0,1]$).

Źródło: własne obliczenia na podstawie danych GUS (*Roczniki Statystyczne Przemysłu dla lat 2000-2008*)

Wartości specyficznych wyrazów wolnych oraz odpowiadające im statystyki t

Dział przetwórstwa przemysłowego	Model 1b	
	Współczynnik	t
Produkcja artykułów spożywczych i napojów	6,202	7,085
Wytwarzanie koksu, produktów rafinacji ropy naftowej i paliw	6,118	12,398
Produkcja wyrobów chemicznych	6,097	8,861
Działalność wydawnicza, poligrafia i reprodukcja zapisanych nośników informacji	6,075	9,462
Produkcja metalowych wyrobów gotowych, z wyłączeniem maszyn i urządzeń	6,022	7,859
Produkcja maszyn i urządzeń, gdzie indziej niesklasyfikowana	5,945	7,841
Produkcja wyrobów z pozostałych surowców niemetalicznych	5,891	8,154
Produkcja wyrobów gumowych i z tworzyw sztucznych	5,799	8,167
Produkcja pojazdów samochodowych, przyczep i naczep	5,714	8,260
Produkcja maszyn i aparatury elektrycznej, gdzie indziej niesklasyfikowana	5,682	8,626
Produkcja mebli; działalność produkcyjna, gdzie indziej niesklasyfikowana	5,631	7,659
Produkcja instrumentów medycznych, precyzyjnych i optycznych, zegarów i zegarków	5,601	10,455
Produkcja drewna i wyrobów z drewna oraz z korka (z wyłączeniem mebli), wyrobów ze słomy i materiałów używanych do wypalania	5,592	8,096
Produkcja metali	5,575	8,686
Produkcja pozostałego sprzętu transportowego	5,420	8,777
Włókiennictwo	5,279	8,276
Produkcja sprzętu i urządzeń radiowych, telewizyjnych i telekomunikacyjnych	5,205	10,001

Wyniki estymacji parametrów modeli jednoczynnikowych opisujących kształtowanie się wartości dodanej brutto w przetwórstwie przemysłowym w latach 1999-2007 (LnWdodbr) - model 2

Zmienna	współczynnik	<i>t</i>	<i>p</i>	współczynnik	<i>t</i>	<i>p</i>
	model (2.2.a)			Model (2.2.b)		
<i>LnZatr</i>	0,607	4,532	0,0000	0,607	4,540	0,0000
<i>lnNinw</i>	0,166	2,646	0,0088	0,156	2,719	0,0071
<i>LnInnow</i>	-0,168	-0,385	0,704	-	-	-
<i>t</i>	0,311	5,118	0,0000	0,312	5,139	0,0000
<i>Stała</i>						
R^2	0,9374			0,9374		
Ocena istotności efektów grupowych	LRT=156,793, $p=0,0000$; F=10,266, $p=0,0000$			LRT=168,222, $p=0,0000$; F= 11,419, $p=0,0000$		

t - wartość statystyki t-Studenta, R^2 – współczynnik determinacji, *LRT* – statystyka testu LRT (*Likelihood Ratio Test*), *F* – statystyka testu Fishera-Snedecora, *p* – prawdopodobieństwo testowe ($p \in [0,1]$).

Źródło: własne obliczenia na podstawie danych GUS (*Roczniki Statystyczne Przemysłu dla lat 2000-2008*)

Wyniki empiryczne modelu produkcji sprzedanej

Funkcja wartości produkcji sprzedanej działów przetwórstwa przemysłowego przyjmuje postać:

model 1

$$\ln Ps_{it} = a_i + b_0 + b_1 \ln Zatr_{it} + b_2 \ln Ninw_{it} + b_3 \ln B + R_{it} + b_4 t$$

model 2

$$\ln Ps_{it} = a_i + b_0 + b_1 \ln Zatr_{it} + b_2 \ln Ninw_{it} + b_3 \ln Innow_{it} + b_4 t$$

gdzie: oznacza logarytm naturalny wartości produkcji sprzedanej w mln zł w cenach stałych z 2005 roku (do urealnienia danych użyto wskaźnika cen produkcji sprzedanej przetwórstwa przemysłowego) dla i-tego działu przemysłu w okresie t. Pozostałe oznaczenia są takie same jak w modelu wartości dodanej.

Wyniki estymacji parametrów modeli jednoczynnikowych opisujących kształtowanie się wartości produkcji sprzedanej w przetwórstwie przemysłowym w latach 1999-2007 (LnPs) - model 3

Zmienna	współczynnik	<i>t</i>	<i>p</i>	współczynnik	<i>t</i>	<i>p</i>
	model (3.1.a)			Model (3.1.b)		
<i>LnZatr</i>	0,490	6,077	0,0000	0,471	6,132	0,0000
<i>lnNinw</i>	0,117	3,701	0,0003	0,117	3,728	0,0003
<i>LnB+R</i>	-0,132	-0,792	0,4293	-	-	-
<i>t</i>	0,639	18356	0,0000	0,638	18366	0,0000
<i>Stała</i>						
R^2	0,9797			0,9796		
Ocena istotności efektów grupowych	LRT=327,090, p=0,0000; F=54,466, p=0,0000			LRT=327,606, p=0,0000; F= 55,027, p=0,0000		

t - wartość statystyki t-Studenta, R^2 – współczynnik determinacji, *LRT* – statystyka testu LRT (*Likelihood Ratio Test*), *F* – statystyka testu Fishera-Snedecora, *p* – prawdopodobieństwo testowe ($p \in [0,1]$).

Źródło: własne obliczenia na podstawie danych GUS (*Roczniki Statystyczne Przemysłu dla lat 2000-2008*)

Wyniki estymacji parametrów modeli jednoczynnikowych opisujących kształtowanie się wartości produkcji sprzedanej w przetwórstwie przemysłowym w latach 1999-2007 (LnPs) - model 4

Zmienna	współczynnik	<i>t</i>	<i>p</i>	współczynnik	<i>t</i>	<i>p</i>
	model (4.2.a)			model (4.2.b)		
<i>LnZatr</i>	0,652	8,599	0,0000	0,653	8,621	0,0000
<i>lnNinw</i>	0,149	4,221	0,0000	0,156	4,794	0,0000
<i>LnInnow</i>	0,113	0,459	0,6468	-	-	-
<i>t</i>	0,537	15,591	0,0000	0,537	15,618	0,0000
<i>Stała</i>						
R^2	0,9837					
Ocena istotności efektów grupowych	LRT=338,052, p=0,0000; F=37,016, p=0,0000			LRT=378,904, p=0,0000; F=47,204, p=0,0000		

t - wartość statystyki t-Studenta, R^2 – współczynnik determinacji, *LRT* – statystyka testu LRT (*Likelihood Ratio Test*), *F* – statystyka testu Fishera-Snedecora, *p* – prawdopodobieństwo testowe ($p \in [0,1]$).

Źródło: własne obliczenia na podstawie danych GUS (*Roczniki Statystyczne Przemysłu dla lat 2000-2008*)

Wnioski

- Wykorzystanie danych czasowo-przekrojowych w modelowaniu ekonometrycznym zyskuje obecnie coraz większą popularność.
- Dla zapewnienia dobrych własności estymatorów konieczne jest stosowanie specyficznych metod estymacji, niejednokrotnie bardziej skomplikowanych niż metody używane do estymacji modeli na podstawie danych „jednowymiarowych”. Mimo wynikających stąd problemów natury metodologicznej i technicznej (konieczność wykorzystania specjalistycznego oprogramowania) liczba analiz prowadzonych na podstawie danych czasowo-przekrojowych rośnie.

Fakt ten wydaje się świadczyć o tym, że wnioski uzyskiwane na podstawie danych, łączących wymiar czasowy i przekrojowy, są unikalne i niemożliwe do uzyskania na podstawie danych „jednowymiarowych”.

- W przeprowadzonych badaniach udało się zbudować modele wartości dodanej brutto i wartości produkcji sprzedanej działów przetwórstwa przemysłowego dla okresu 1999-2007.
- Nie powiodły się próby zbudowania analogicznych modeli dla lat 2008-2011. Jedną z przyczyn tych niepowodzeń jest zapewne światowy kryzys gospodarczy, którego efektem są znaczące spadki nakładów inwestycyjnych, w tym nakładów na działalność badawczo-rozwojową i innowacyjną.

- Zastosowanie funkcji produkcji Cobba-Douglasa umożliwia interpretację ocen parametrów w kategoriach elastyczności wskaźników produkcyjnych względem nakładów poszczególnych czynników wytwórczych.
- Nie stwierdzono istotnego statystycznie wpływu aktywności innowacyjnej (nakładów na B+R /nakładów na innowacje) na wartość dodaną oraz wartość produkcji sprzedanej.
- Nie udało się zbudować poprawnego w sensie statystycznym i ekonomicznym modelu wydajności pracy. Trudno jednak uwierzyć w rzeczywisty brak wpływu innowacyjności na wydajność pracy w przedsiębiorstwach przemysłu przetwórczego.

Przyczyną niepowodzeń związanych z modelowaniem wydajności pracy nie jest brak zależności między nakładami a wydajnością pracy tylko prawdopodobnie niski poziom tych nakładów, próba obejmująca zbyt krótki przedział czasowy oraz brak właściwej metody badawczej.

Bibliografia

- Dańska-Borsiak B., (2011), *Dynamiczne modele panelowe w badaniach ekonomicznych*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Hausman J. A., (1978), *Specification Tests in Econometrics*, *Econometrica*, vol. 46
- Hausman J. A., Taylor W. E., [1981], *Panel Data and Unobservable Individual Effects*, *Econometrica*, vol. 49
- Nauka i Technika (2009), GUS, Warszawa

- Porter M., (1990), *The Competitive Advantage of Nations*, New York, The Free Press
- Roczniki Statystyczne Przemysłu (2000- 2012), GUS, Warszawa
- Suchecki B., [2000], *Panel Data and Multivariate Models In the Economic Researches*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź
- Welfe W., (2000), *Empiryczne modele wzrostu gospodarczego*, Ekonomista nr 4, Warszawa, s. 483-497.
- Witkowski B., Weresa M. A., (2006), *Wpływ innowacji na konkurencyjność branż polskiego przemysłu*, [w:] Polska. Raport o konkurencyjności 2006. Rola innowacji w kształtowaniu przewag konkurencyjnych, (red.) Weresa M. A., Instytut Gospodarki Światowej SGH, Warszawa