DYNAMICZNE MODELE EKONOMETRYCZNE
X Ogólnopolskie Seminarium Naukowe, 4–6 września 2007 w Toruniu
Katedra Ekonometrii i Statystyki, Uniwersytet Mikołaja Kopernika w Toruniu

Marek Szajt
Politechnika Częstochowska

Techniczne aspekty dekompozycji wyrazu wolnego w modelach przestrzenno-czasowych

1. Wprowadzenie

Celem artykułu jest porównanie wyników otrzymywanych w drodze estymacji modeli przestrzenno-czasowych, w których wymagana jest eliminacja nieistotnych statystycznie części dekomponowanego wyrazu wolnego zwanych dalej „wyrazami wolnymi". Ze względu na trudności w pozyskiwaniu danych – zwłaszcza makroekonomicznych – korzystanie z modeli przestrzenno czasowych zdobywa sobie coraz większą popularność. Ich niewątpliwą zaletą jest możliwość wskazania podobieństw i różnic pomiędzy wykorzystywanymi w próbie łączonej obiektami. Tak więc oceny parametrów stojących przy zmiennych możemy traktować jako wspólne, natomiast dzięki dekompozycji wyrazu wolnego lub składnika losowego (porównaj: Dańska, 2000) jesteśmy w stanie wskazać na indywidualne cechy badanych obiektów – państw, regionów, przedsiębiorstw.

2. Problemy związane z dekompozycja wyrazu wolnego

O ile w przypadku dekompozycji wyrazu wolnego w etapie wstępnym oczekujemy pierwotnych wyników otrzymanych „mechanicznie", o tyle w kolejnych etapach pojawia się wiele problemów natury decyzyjnej. Zwykle w procesie estymacji okazuje się, iż nie wszystkie elementy dekomponowanego wyrazu wolnego są istotne statystycznie. W związku z tym musimy podjąć odpowiednie kroki, by uzyskać satysfakcjonujące nas wyniki. Najprostszym sposobem jest pozostawienie jednego wspólnego wyrazu wolnego oraz wprowadzenie odpowiedniej liczby zmiennych zero-jedynkowych. Zwykle działania te mają charakter sekwencyjny i odbywają się w ten sposób, iż każdemu (poza jed-

nym) obiektowi w etapie pierwszym odpowiada zmienna zero-jedynkowa. Po wstępnej weryfikacji eliminujemy zmienne zero-jedynkowe, które okazały się nieistotne. Otrzymujemy więc wspólny wyraz wolny – będący na swój sposób reprezentantem grupy obiektów nieposiadających odpowiadających im „własnych" zmiennych zero-jedynkowych (typowych), oraz informację na temat charakterystycznych właściwości poszczególnych – specyficznych – obiektów.

Innym sposobem, jest estymacja bez wspólnego wyrazu wolnego. W pierwszym etapie otrzymujemy tyle dekomponowanych części wyrazu wolnego (specyficznych „wyrazów wolnych"), ile obiektów wziętych zostało do budowy modelu. Następnie po wstępnej estymacji dane odpowiadające poszczególnym obiektom – w przypadku nieistotności danego wyrazu wolnego – są łączone. Pojawia się tu problem sposobu łączenia danych, możemy:

– łączyć w grupy (traktować jako jeden obiekt) dane dla kilku nieistotnych statystycznie obiektów,
– łączyć w grupy (traktować jako jeden obiekt) dane dla nieistotnych statystycznie obiektów o tych samych znakach przy wstępnych wartościach „wyrazów wolnych"
– łączyć w grupy (traktować jako jeden obiekt) dane dla nieistotnych statystycznie obiektów z innymi im najbliższymi co do wstępnej wartości otrzymanych „wyrazów wolnych"

Przedstawione sposoby dotyczą aspektów technicznych, należy wskazać również na względy merytoryczne. Tu pojawia się problem. Jeżeli łączymy dane dla nieistotnych statystycznie dekomponowanych wyrazów wolnych, otrzymujemy w efekcie jeden „wyraz wolny" charakterystyczny dla grupy „słabo wyindywidualizowanych obiektów", który często również nie jest wysoce istotny.

Łączenie wyrazów wolnych „słabych", ale o tych samych znakach, jest zbliżone w efekcie do swojego rodzaju poprawy segregacji wstępnej. Przy czym w efekcie w porównaniu z poprzednim podejściem, otrzymujemy dwie „słabe grupy" zwykle charakteryzujące się również słabo istotnymi „wyrazami wolnymi".

Łączenie nieistotnych z najbliższymi im dotyczy analizy nie wartości statystyk t – jak miało to miejsce w poprzednich ujęciach. W tym ujęciu po wstępnej estymacji z pełną grupą zmiennych, na podstawie statystyki t wskazujemy obiekty, które powinny zostać dołączone, po czym obserwując wartość parametru, wskazujemy z którym wydają się być najbliższe. Tu również poprawiamy segregację wstępną. Odgórny podział ze względu na elementy szeregu odpowiadające poszczególnym obiektom jest sztuczny. W niektórych przypadkach dane są tak podobne i zarazem o tak zbliżonym przebiegu i wzajemnych zależnościach, iż ich podział mija się z celem. W związku z tym w procesie grupowania niekiedy połączeń dokonuje się wcześniej – często bezpodstawnie. Optymalnym wydaje się być przeprowadzenie estymacji i „odrzucenie" (połączenie) tych obiektów, które nie wnoszą żadnych specyficznych informacji. Pojawia się jeszcze jeden problem. Czy wzajemne sąsiedztwo rozpatrywać z

uwzględnieniem znaków czy nie. Znaki znajdujące się bowiem przy ocenach parametrów wskazują nie na siłę reakcji, ale przede wszystkim na jej kierunek lub odległość od pewnego punktu odniesienia.

3. Sposoby weryfikacji właściwych procedur eliminacji nieistotnych efektów przestrzennych

Warto zwrócić jednak uwagę na kilka ważnych aspektów związanych z wykorzystaniem współczynnika determinacji R^2 w badaniu podobieństwa wyników. Do głównych jego wad zaliczyć należy:

- wyznaczanie R^2 przy danych o dużej zmienności lub - jak w naszym przypadku - wyraźnej segmentacji wynikającej z konstrukcji próby, jest nie w pełni prawidłowe ze względu na właściwości wykorzystywanej średniej arytmetycznej,
- wykorzystanie R^2 w przypadku zależności o charakterze innym niż prostoliniowym nie zawsze jest prawidłowe (por. Deadman i Charemza, 1997),
- charakteryzuje się niewielką wrażliwością na różny kształt przebiegu badanego procesu w przypadku odpowiednio wysokiej zmienności kilku badanych obiektów.

Biorąc pod uwagę powyższe można – zdaniem autora – uzupełnić wnioskowanie na podstawie współczynnika determinacji o dodatkowe wnioskowanie w oparciu o współczynnik rozbieżności Theila. Można go wyznaczyć korzystając nie z wartości empirycznych i prognoz, lecz jako rozbieżność (podobieństwo) dwóch ciągów danych odpowiadających analogicznym kategoriom wartości teoretycznych y_0 i prognozowanych y_i:

$$I^2 = \frac{\sum_{t=1}^{n}(y_{0t} - y_{it})^2}{\sum_{t=1}^{n} y_{0t}^2}. \tag{1}$$

Niewątpliwą zaletą tego współczynnika jest możliwość rozłożenia go na trzy składniki:

$$I^2 = I_1^2 + I_2^2 + I_3^2. \tag{2}$$

Składniki te pozwalają na bliższe określenie błędów aproksymacji – czego nie można było powiedzieć wspomnianym wcześniej współczynniku korelacji.

$$I_1^2 = \frac{(\bar{y}_0 - \bar{y}_i)^2}{\frac{1}{n}\sum_{t=1}^{n} y_{0t}^2}. \tag{3}$$

Miernik pierwszy wskazuje na błąd wynikający z obciążenia metody predykcji. We wskazanym zastosowaniu można co najwyżej mówić o dopasowaniu, a nie o predykcji. Kolejny składnik wyznaczamy według wzoru:

$$I_2^2 = \frac{(S_{y_0} - S_{y_i})^2}{\frac{1}{n}\sum_{t=1}^{n} y_{0t}^2}, \tag{4}$$

gdzie:

S_{y_0} – to odchylenie standardowe dla wartości prognozowanych zmiennej y,

S_{y_i} – odchylenie standardowe dla wartości rzeczywistych zmiennej y.

Informuje on o ewentualnym niedostosowaniu elastyczności prognoz do stanu faktycznego. Miara wyznaczona w ten sposób odgrywa we wskazanym zastosowaniu bardzo ważną rolę. Podobną rolę odgrywa również ostatni z mierników:

$$I_3^2 = \frac{2S_{y_0}S_{y_i}(1 - r_{y_0y_i})}{\frac{1}{n}\sum_{t=1}^{n} y_{0t}^2}. \tag{5}$$

Miernik ten wskazuje na rozbieżności pomiędzy kierunkiem zmian wartości teoretycznych i empirycznych (por. Kukuła, 2000).

4. Wyniki badań empirycznych

W badaniach empirycznych wykorzystano przestrzenno–czasową próbę składającą się z obserwacji rocznych dotyczących kształtowania się PKB w 28 państwach europejskich w zależności od czynników innowacyjnych takich jak wydatki brutto na działalność B+R, liczba badaczy, liczba patentów zgłoszonych przez rezydentów, liczba patentów zgłoszonych przez nierezydentów oraz liczba mieszkańców. Próba miała charakter niezbilansowany, a długości szeregów czasowych dla poszczególnych państw sięgały od 5 do 23 okresów. Istotą badań nie było jednak rzeczywiste modelowanie PKB, a zbadanie kilku możliwości dekompozycji wyrazu wolnego, jakie autor arbitralnie przyjął za możliwe.

Estymacji poddano cztery warianty modelu:

A – dekompozycja z eliminacją nieistotnych „wyrazów wolnych" poprzez ich łączenie – otrzymano w efekcie 23 statystycznie istotne oceny „wyrazów wolnych".

B – dekompozycja z eliminacją nieistotnych „wyrazów wolnych" poprzez ich łączenie z innymi nieistotnymi o tych samych znakach - otrzymano w efekcie 20 statystycznie istotnych ocen „wyrazów wolnych".

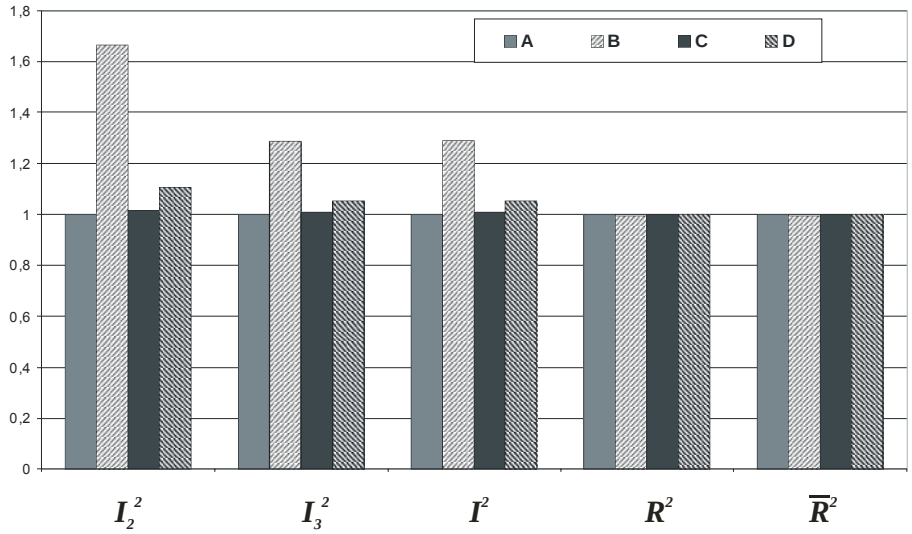
C – dekompozycja z eliminacją nieistotnych „wyrazów wolnych" poprzez ich zamianę na zmienne zero-jedynkowe i wprowadzenie wspólnego wyrazu wolnego - otrzymano w efekcie 21 statystycznie istotnych ocen dla zmiennych zero–jedynkowych + statystycznie istotny wyraz wolny.

D – dekompozycja z eliminacją nieistotnych „wyrazów wolnych" poprzez ich łączenie z innymi najbliższymi im co do oceny (obiektami) - otrzymano w efekcie 22 statystycznie istotne oceny „wyrazów wolnych".

Każdy z otrzymanych szeregów wartości teoretycznych został porównany za pomocą wskazanych wcześniej miar. Ocena pierwszego wskaźnika (I_1^2)dla procedury A okazała się nawet kilkadziesiąt krotnie większa od pozostałych, jednakże jej wartość absolutna ($5{,}516 * 10^{-20}$) podpowiada, iż nie jest to ocena dyskwalifikująca. Co więcej ten wskaźnik informuje o obciążeniu wynikającym z zastosowanej metody predykcji, która jest w zasadzie, co do podstawowych założeń, jednakowa we wszystkich przypadkach. W pozostałych przypadkach najlepsze oceny współczynników I_1^2, I_2^2, I_3^2 uzyskała procedura A, po niej C, D i B.

We wszystkich przypadkach grupa zmiennych objaśniających pozostała w tym samym zestawieniu, jednakże najsilniejsze determinanty, były jednocześnie najbardziej wrażliwe na ewentualne zmiany. W zasadzie wskazano na bardzo wysoką zależność pomiędzy zmiennością parametrów a ich wielkością absolutną. Współczynnik korelacji policzony dla odpowiednich odchyleń standardowych w parze ze średnimi ocenami parametrów wyniósł 0,993, trudno więc pominąć tę kwestię w analizie wyników. Warto zatem pamiętać, iż - jak to za chwilę zostanie wskazane – błędny sposób eliminacji nieistotnych efektów przestrzennych może powodować przeszacowanie ocen głównych determinant badanego zjawiska.

W celu dalszej weryfikacji zestawiono w sposób porównawczy najlepsze oceny współczynników I_1^2, I_2^2, I_3^2 - przypisując im jeden - i pozostałe oceny - dzieląc ich wartość przez ocenę najlepszą (Rys. 1).



Rys. 1. Porównanie wartości poszczególnych wskaźników weryfikujących modele
Źródło: opracowanie własne.

Powyższy rysunek wyraźnie wskazuje na najlepsze wyniki procedury A i praktycznie takie same procedury C. Pamiętać należy, iż połączenie nieistotnych „wyrazów wolnych" i wprowadzenie wspólnego wyrazu wolnego jest działaniem o podobnej idei. Stąd otrzymane wyniki charakteryzują się bardzo wysokim stopniem podobieństwa. Co się tyczy ocen parametrów, zastosowanie procedury A wskazuje poprzez oceny parametrów –„wyrazów wolnych" – na ich specyficzny wpływ na zmienną endogeniczną z punktem odniesienia na poziomie 0. W procedurze C mamy natomiast możliwość wskazania tego wpływu również w odniesieniu do ogólnego poziomu (punktu odniesienia) równego ogólnemu wyrazowi wolnemu.

Otrzymane wartości dekomponowanych wyrazów wolnych przedstawiono w kolejnej tabeli. Dla wartości połączonych wskazano tę samą wartość oceny (dla identyfikacji pogrubiono czcionkę). W przypadku procedury B i D istniały dwie wartości powtarzalne (drugą dodatkowo pochylono). Co również warte podkreślenia, większość ocen indywidualnych – niepowtarzalnych – dla procedury A i C osiągnęła podobny lub jednakowy poziom. Analogiczne oceny dla pozostałych procedur są mimo wszystko wyraźnie inne, zwłaszcza w przypadku procedury B. Średnie różnice tych ocen dla procedur A i C wynosiło ok. 12%, podczas gdy dla procedur A i B – 43%, A i D – 34%.

Tabela 1. Wartości ocen „wyrazów wolnych" w poszczególnych procedurach

Nr	A	B	C	D
1	0.0133	0.0120	0.0134	0.0126
2	0.0111	0.0097	0.0111	0.0103
3	-0.0031	-0.0024	-0.0032	-0.0025
4	0.0140	0.0169	0.0138	0.0152
5	**-0.0007**	**-0.0006**	**-0.0010**	**-0.0007**
6	0.0160	0.0145	0.0160	0.0155
7	**-0.0007**	***0.0035**	**-0.0010**	***0.0027**
8	0.0064	***0.0035***	0.0064	0.0063
9	**-0.0007**	**-0.0006**	**-0.0010**	-0.0019
10	-0.0053	-0.0051	-0.0057	-0.0064
11	0.0047	0.0050	0.0046	0.0048
12	-0.0014	-0.0011	**-0.0010**	**-0.0007**
13	0.0128	0.0123	0.0127	0.0130
14	0.0100	0.0093	0.0100	0.0099
15	-0.0023	**-0.0006**	-0.0026	-0.0028
16	**-0.0007**	**-0.0006**	**-0.0010**	**-0.0007**
17	-0.0021	**-0.0006**	-0.0023	**-0.0007**
18	0.0086	0.0070	0.0086	0.0075
19	0.0148	0.0135	0.0148	0.0145
20	-0.0079	-0.0063	-0.0082	-0.0075
21	0.0041	0.0044	0.0040	0.0042
22	-0.0053	**-0.0006**	-0.0055	-0.0049
23	**-0.0007**	**-0.0006**	**-0.0010**	**-0.0007**
24	0.0032	0.0030	0.0032	***0.0027***
25	**-0.0007**	**-0.0006**	**-0.0010**	**-0.0007**
26	0.0065	0.0035	0.0067	0.0052
27	-0.0125	-0.0100	-0.0129	-0.0125
28	-0.0074	-0.0071	-0.0077	-0.0081

Źródło: obliczenia własne.

Jak łatwo zauważyć, wspólne „wyrazy wolne" dotyczą bez względu na procedurę zwykle tych samych obiektów. Co ważne, w przypadku obiektu 7 (wartości zaznaczone gwiazdką) obserwujemy niezgodność poziomu w postaci odmiennych znaków. O ile w badanym przypadku domniemujemy zależność o charakterze liniowym (poziom może być zaniżony ze względu na zawyżone wartości ocen parametrów przy zmiennych), o tyle w przypadku chociażby modeli potęgowych, sytuacja ta może powodować błędne wnioskowanie.

5. Podsumowanie

Prezentowany powyżej tekst wskazuje kilka możliwości postępowania w przypadku pojawienia się nieistotnych statystycznie ocen zdekomponowanego wyrazu wolnego w procesie estymacji modeli przestrzenno–czasowych. Powyższe wyniki empiryczne wskazują stosowanie zmiennych zero-jedynkowych w połączeniu ze wspólnym wyrazem wolnym jako najlepsze – obok łączenia nieistotnych statystycznie „wyrazów wolnych" jako najlepszą metodę. Próby łączenia „wyrazów wolnych" o podobnych ocenach lub znakach wydają się dawać gorsze rezultaty. Omówione wyniki są jedynie wskazówką, która w przypadku potwierdzenia większą liczbą badań mogłaby pomóc w badaniach wykorzystujących próby mieszane.

Literatura

Charemza, W., Deadman, D. (1997), *Nowa ekonometria*, PWE, Warszawa.

Dańska, B. (2000), *Przestrzenno-czasowe modelowanie zmian w działalności produkcyjnej w Polsce*, *Zastosowanie modeli panelowych*, Absolwent, Łódź, 25–27.

Kukuła, K. [red.] (2000), *Wprowadzenie do ekonometrii w przykładach i zadaniach*, PWN, Warszawa, 129–130.