

DYNAMICZNE MODELE EKONOMETRYCZNE

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

Marcin Błazejowski
Wyższa Szkoła Bankowa w Toruniu

Analiza porównawcza wielorównaniowych modeli mikrorynku produktów konsumpcyjnych dla danych tygodniowych i dziennych*

1. Wprowadzenie

Niniejszy artykuł jest kontynuacją i rozwinięciem artykułu z 2005 roku pt. „Model opisu „bąbla promocyjnego”: identyfikacja, analiza oraz wykorzystanie”, opublikowanego w *Dynamiczne Modele Ekonometryczne*, dotyczącego modelowania reakcji konsumenta na impuls wywołany cenową promocją sprzedaży. W 2005 roku prezentowany był model VAR(3), oszacowany na podstawie próby $n=200$ dziennych obserwacji dotyczących transakcji sprzedaży proszków do prania (ilość oraz cena za kilogram). Profil czasowy funkcji odpowiedzi rynku opisywany był przebiegiem funkcji odpowiedzi impulsowej. W podsumowaniu artykułu z 2005 roku stwierdzono, że naturalnym krokiem w dalszych badaniach będzie oszacowanie modelu o równaniach współzależnych i porównanie uzyskanych wyników symulacji dynamicznych z profilami funkcji odpowiedzi impulsowych. Budowa modelu o równaniach współzależnych powinna bowiem pozwolić w lepszym stopniu, niż to ma miejsce w przypadku klasycznego modelu VAR, poznać charakterystykę reakcji konsumentów na zmiany rynkowe.

W niniejszym artykule zaprezentowane zostaną wyniki uzyskane dla 2 modeli o równaniach współzależnych, oszacowanych na podstawie szeregów obserwacji dziennych oraz tygodniowych i dotyczących tego samego rynku.

* Praca naukowa finansowana ze środków na naukę w latach 2006-2008, projekt badawczy nr 1-H02B-007-30.

2. Specyfikacja modeli

Przyjęty schemat specyfikacji równań modelu współzależnego musiał być na tyle uniwersalny, aby mógł spełnić następujące kryteria:

1. identyfikowalność równań,
2. posiadanie dodatniej liczby stopni swobody poszczególnych równań, co wiąże się z uwzględnieniem niezbyt długich rzędów opóźnień,
3. pożądane własności procesów resztowych, w tym przede wszystkim brak autokorelacji reszt,
4. odpowiednie własności dynamiczne modelu tak, aby nadawał się on do symulacji.

W przypadku analizy procesów sprzedażowych jedynymi zmiennymi egzogenicznymi będą zmienne cenowe. Dodatkowo należy włączyć dane opisujące strukturę wewnętrzną tych procesów, w tym przede wszystkim cykliczność sprzedaży. Bardzo ważne, z punktu widzenia dopasowania równań do danych empirycznych, będzie posiadanie wiedzy o czynnikach dodatkowych, takich jak stany magazynowe (np. czy zerowa sprzedaż w danym okresie nie wynikała z braku zapasów w magazynie), kanałach i środkach promocji i reklamy danego produktu oraz danych kalendarzowych, takich jak uwzględnienie dni wolnych od pracy (świąt, długich weekendów, itd...). Niestety, w wykorzystanym materiale statystycznym nie było powyższych informacji.

Specyfikacja obu modeli o równaniach współzależnych, na podstawie których przeprowadzono symulacje, oparta była na powyższym schemacie. W badaniu uwzględniono 11 marek proszków o największym udziale w rynku i sprzedaż każdej z tych marek analizowano w układzie małych i średnich opakowań, co dało liczbę 22 zmiennych endogenicznych. Ponieważ nie można a priori wykluczyć żadnych powiązań jednoczesnych między procesami sprzedaży poszczególnych marek, każdy z modeli posiadał 22 zmienne łącznie współzależne, będące logarytmami naturalnymi ilości sprzedanych opakowań danej marki. Schemat specyfikacji i -tego równania w modelu był następujący:

$$A(L)ilos_{i,w,t} = C_t + \sum_{j=1}^k \sum_{w=1}^3 \sum_{s=0}^{q_{jw}} \beta_{j,w,s} ilosc_{j,w,t-s} + \sum_{j=1}^k \sum_{w=1}^3 \sum_{s=0}^{p_{jw}} \alpha_{j,w,s} ilosc_{j,w,t-s} + \zeta_{it}$$

dla $i = 1, 2, \dots, k$; $w = 1, 2, 3$; $t = 1, 2, \dots, n$, gdzie: k – liczba marek (w badaniu wybrano $k=11$), w – indeks typu wielkości opakowania (małe, średnie oraz duże), $ilos_{i,w,t}$ – logarytm naturalny liczby sprzedanych jednostek małych, średnich oraz dużych opakowań i -tej marki w okresie t , $cena_{j,w,t}$ – logarytm naturalny ceny za kilogram w małych, średnich oraz dużych opakowaniach j -tej marki w okresie t , $A(L)$ – operator autoregresji dla i -tego procesu sprzedażowego, p_{jw} , q_{jw} – rząd opóźnień dla cen oraz ilości j -tej marki w opakowaniu typu w , C_t – część równania opisująca cykliczność, która w modelu dla danych dziennych

była następująca: $\sum_{r=1}^2 \gamma_{i,D,r} D_{r,t} + \sum_{m=1}^5 \gamma_{i,d,m} d_{m,t}$ gdzie: $D_{r,t}$ – zmienna zero-jedynkowa reprezentująca dekadę r w cyklu miesięcznym ($r=1,2$), $d_{m,t}$ – zmienna zero-jedynkowa reprezentująca dzień m w cyklu tygodniowym (w układzie tygodnia sześciodniowego, $m=1,2,\dots,5$), natomiast w modelu dla danych tygodniowych część D miała postać: $\sum_{r=1}^3 \gamma_{i,D,r} D_{r,t}$ gdzie: $D_{r,t}$ – zmienna zero-jedynkowa reprezentująca tydzień r w cyklu miesięcznym ($r=1,2,3$).

Komentarza wymaga tutaj wybór rzędów opóźnień dla zmiennych cenowych oraz ilości sprzedanych opakowań marek konkurencyjnych (p_{jw} , q_{jw}). Ze względu na konieczność zapewnienia dodatniej liczby stopni swobody oraz identyfikowalności równań, w modelu dla danych dziennych przyjęto następujące rzędy opóźnień:

1. dla i -tego procesu sprzedaży (zmiennej endogenicznej równania): 4,
2. dla pozostałych 21 procesów sprzedaży: 3,
3. dla zmiennych cenowych: 6 (co wynika z układu tygodnia sześciodniowego).

W modelu dla danych tygodniowych przyjęto następujące rzędy opóźnień:

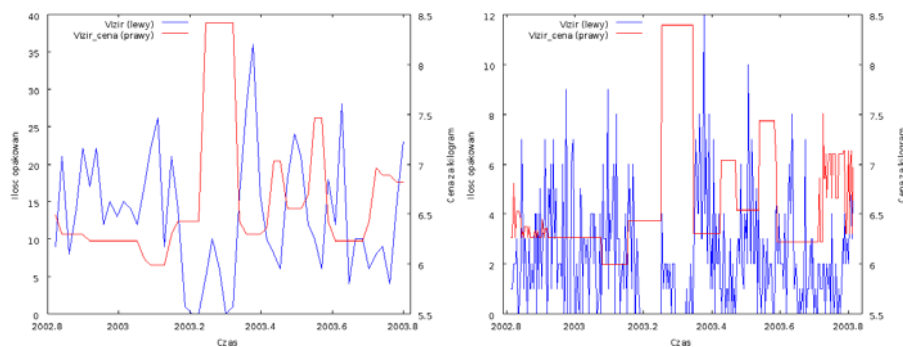
1. dla i -tego procesu sprzedaży (zmiennej endogenicznej równania): 3,
2. dla pozostałych 21 procesów sprzedaży: 2,
3. dla zmiennych cenowych: 2.

Taki wybór rzędów opóźnień pozwolił na zapewnienie identyfikowalności równań z jednej strony oraz własności dynamicznych z drugiej strony. Estymacja parametrów równań obu modeli została wykonana w programie GRETL z wykorzystaniem podwójnej metody najmniejszych kwadratów (2MNK).

Ze względu na ograniczenia redakcyjne nie ma możliwości zaprezentowania dokładnej charakterystyki oszacowanych modeli, warto jedynie wspomnieć, że w modelu dla danych dziennych wartości skorygowanych współczynników R^2 poszczególnych równań wahały się od 0,2788 do 0,5268, w 17 przypadkach należało odrzucić hipotezę zerową o normalności rozkładu reszt (na poziomie istotności 0,1) oraz tylko 3 procesy resztowe charakteryzowały się występowaniem autokorelacji reszt niskiego rzędu, jednocześnie dla 9 procesów resztowych wszystkie wartości PACF okazały się statystycznie istotne (na poziomie istotności 0,05). Natomiast w modelu dla danych tygodniowych wartości skorygowanych współczynników determinacji poszczególnych równań wahały się od 0,3304 do 0,9567, w 12 przypadkach należało odrzucić hipotezę zerową o normalności rozkładu reszt (na poziomie istotności 0,1), oraz tylko 4 procesy resztowe charakteryzowały się występowaniem autokorelacji niskiego rzędu oraz tylko dla 3 procesów resztowych wszystkie wartości PACF okazały się statystycznie istotne (na poziomie istotności 0,05). Należy tutaj stwierdzić, że dla danych dziennych oraz tygodniowych, w których występuje duża liczba danych odstających oraz okresów o zerowej sprzedaży, powyższe wyniki można uznać za zadowalające.

3. Charakterystyka materiału empirycznego

Ponieważ dane statystyczne wykorzystane w niniejszym badaniu są tymi samymi, na podstawie których szacowany był model VAR prezentowany w Błażejowski (2006a) oraz (2006b), w tym miejscu warto jedynie wspomnieć, iż są to szeregi czasowe dotyczące zarejestrowanych transakcji sprzedaży proszków do prania za okres od 13.04.1998 roku do 24.11.2003 roku. Dane o pojedynczych transakcjach zostały zagregowane do danych dziennych oraz tygodniowych, przy czym ilości sprzedanych opakowań zostały zsumowane, natomiast ceny zostały wyznaczone jako średnie arytmetyczne w przeliczeniu na jednostkę masy produktu. Wszystkie typy opakowań występujące na rynku zostały pogrupowane w 3 kategorie: małe opakowania (do 0.72 kg włącznie), średnie opakowania (do 3 kg włącznie) oraz duże opakowania (powyżej 3 kg). Ponadto dla okresów, w których występowała zerowa sprzedaż, przyjęto ostatnią wyznaczoną cenę kilograma. Ponieważ jako postać analityczną równań przyjęto funkcję potęgową, należało wyznaczyć logarytmy naturalne cen oraz ilości sprzedanych opakowań każdego dnia. Ze względu jednak na okresy o zerowej sprzedaży i braku możliwości wyznaczenia logarytmu, dane oryginalne zostały przekształcone w ten sposób, że do ilości dziennej sprzedaży została dodana liczba 1. Na wykresie (1) znajdują się przebiegi ilości sprzedanych opakowań oraz ceny za kilogram proszku Vizir sprzedawanego w średnich opakowaniach. Wykres po prawej stronie przedstawia przebiegi danych dziennych, wykres po lewej stronie dotyczy obserwacji tygodniowych.



Wykres 1. Ilość oraz cena za kilogram proszku Vizir (lewy – tygodniowe, prawy – dzienne)

Źródło: opracowanie własne.

4. Założenia scenariuszy do symulacji oraz uzyskane wyniki

Symulacje, jakie przeprowadzono na obu modelach, zakładały 20% obniżki cen trwające przez okres 1 tygodnia. W związku z tym w symulacjach na modelu dla danych dziennych były to impulsy podtrzymane przez 6 kolejnych okresów, podczas gdy w symulacjach na modelu dla danych tygodniowych były to

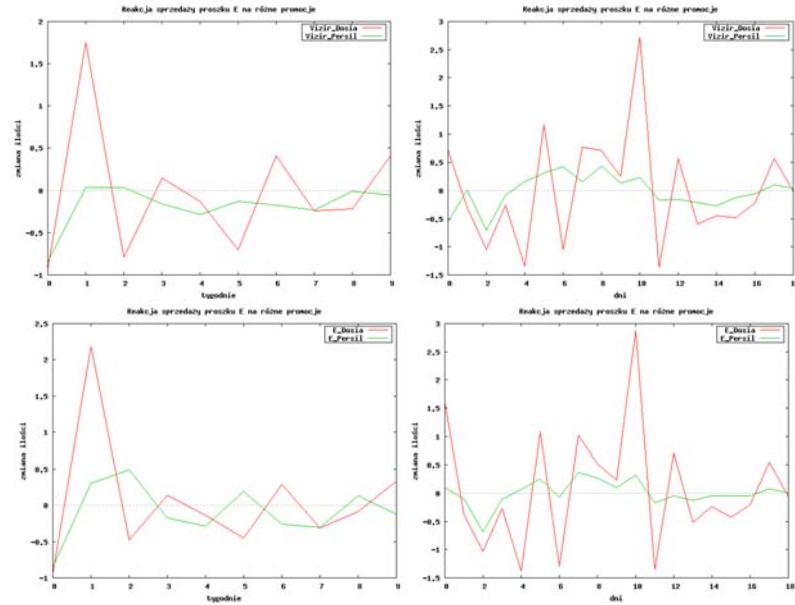
impulsy jedno okresowe. W scenariuszach założono, że takie impulsy, będące efektami cenowych promocji sprzedaży, mogą się pojawić tylko w cenie jednej marki, dwóch marek, trzech marek oraz czterech marek w tym samym okresie¹. Do symulacji wybrano cztery marki o największym udziale w badanym rynku. Ponieważ dysponowano dwoma modelami o równaniach współzależnych, symulacje dotyczyły zmian tych samych cen w identycznych okresach. Dzięki temu możliwe stało się porównanie wyników tych symulacji. Najważniejsze było tutaj uzyskanie, jak się wydaje, podobnych różnic pomiędzy przebiegami rozwiązań zakłóconych i bazowych dla obu modeli oraz podobieństwa przebiegu reakcji na zakłócenie, w tym przede wszystkim okresów, po upływie których efekt zaburzenia zanika.

Przeprowadzono serię symulacji dynamicznych w oparciu o założone scenariusze. Ponieważ liczba uzyskanych ścieżek przebiegów zmiennych endogenicznych jest bardzo duża (6 kombinacji dla symulacji polegających na równoległych promocjach dwóch marek oraz 4 kombinacje polegające na równoległych promocjach trzech marek, co dla analizowanych 4 marek daje 40 przebiegów), zdecydowano się na przedstawienie wyników tylko dla średnich opakowań proszku E. Na wykresie (2) przedstawione są zmiany ilości sprzedanych średnich opakowań proszku E w zależności od tego, jakie 2 marki są w danym momencie promowane, lub inaczej w cenach których marek pojawiło się zaburzenie. Lewe rysunki zawsze dotyczą zmian dziennej sprzedaży, natomiast prawe rysunki zawsze dotyczą zmian tygodniowej sprzedaży.

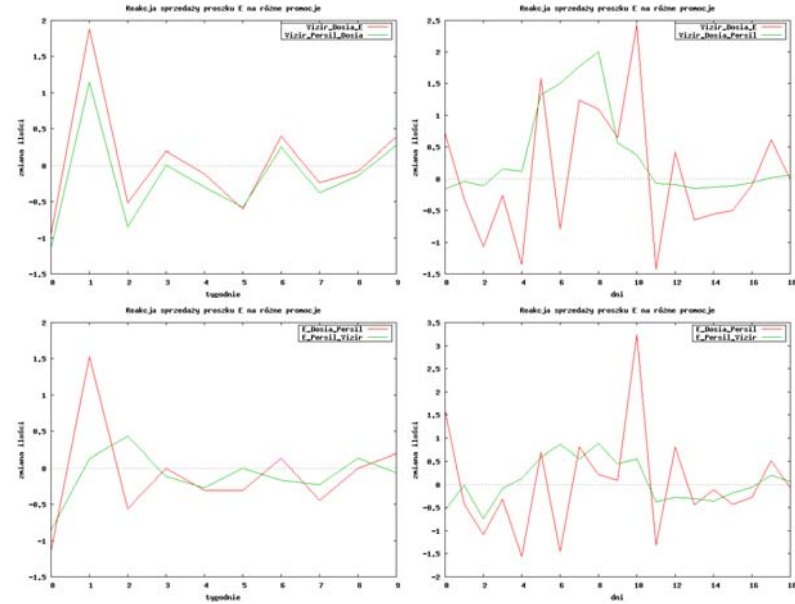
Na wykresie (3) przedstawione są zmiany ilości sprzedanych średnich opakowań proszku E w zależności od tego, jakie 3 marki są w danym momencie promowane, lub inaczej w cenach których marek pojawiło się zaburzenie. Lewe rysunki zawsze dotyczą zmian dziennej sprzedaży, natomiast prawe rysunki zawsze dotyczą zmian tygodniowej sprzedaży. Ponieważ bezpośrednie porównanie wyników symulacji na modelu dla danych dziennych oraz tygodniowych nie jest możliwe, postanowiono porównać zmiany ilości sprzedaży opakowań marek proszków do prania promowanych w różnych scenariuszach po upływie podobnych okresów. Pod uwagę wzięto dwa przypadki:

1. zmiany ilości po upływie 1 tygodnia (skumulowana sześciu-okresowa różnica dla modelu dla danych dziennych oraz pierwsza różnica dla modelu dla danych tygodniowych),
2. zmiany ilości po upływie 3 tygodni (skumulowana osiemnasto-okresowa różnica dla modelu dla danych dziennych oraz skumulowana trzy okresowa różnica dla modelu dla danych tygodniowych).

¹ Można także badać reakcję konsumentów na promocje przesunięte w czasie, jednak w niniejszym badaniu takie scenariusze nie były brane pod uwagę. Wyniki takich symulacji można znaleźć w Błażejowski, Kufel (2007) oraz Kufel, Błażejowski (2007).



Wykres 2. Zmiany ilości sprzedanych średnich opakowań proszku E w zależności od występujących na rynku równoległych 2 promocji sprzedaży
Źródło: opracowanie własne.



Wykres 3. Zmiany ilości sprzedanych średnich opakowań proszku E w zależności od występujących na rynku równoległych 3 promocji sprzedaży
Źródło: opracowanie własne.

Tabela 1. Różnice po 1 oraz 3 tygodniach w ilościach sprzedanych średnich opakowań proszku marki E w zależności od założonego scenariusza promocji

Marki podlegające promocji w tym samym okresie	Okres trwania promocji	Wyniki uzyskane na podstawie model dla danych tygodniowych	Wyniki uzyskane na podstawie model dla danych dziennych
Vizir_Dosia	3 tyg.	0.028538645	0.325942885
	1 tyd.	-0.933447515	-1.076876348
Vizir_Persil	3 tyg.	-0.757238589	-0.322858974
	1 tyd.	-0.840536881	-0.870381249
E_Dosia	3 tyg.	0.762861158	1.45603844
	1 tyd.	-0.939534223	-0.37578493
E_Persil	3 tyg.	-0.05059613	0.174359505
	1 tyd.	-0.847021724	-0.451511433
Vizir_E	3 tyg.	0.507394114	3.237584173
	1 tyd.	-0.60211164	0.407564609

Źródło: obliczenia własne.

Tabela (1) przedstawia porównanie wartości różnic po 1 tygodniu oraz po 3 tygodniach dla średnich opakowań marki E dla równoległych 2 promocji, natomiast tabela (2) przedstawia porównanie wartości różnic po tygodniu oraz po trzech tygodniach dla średnich opakowań marki E dla równoległych 3 promocji.

Tabela 2. Różnice po 1 oraz 3 tygodniach w ilościach sprzedanych średnich opakowań proszku marki E w zależności od założonego scenariusza promocji

Marki podlegające promocji w tym samym okresie	Okres trwania promocji	Wyniki uzyskane na podstawie modelu dla danych tygodniowych	Wyniki uzyskane na podstawie modelu dla danych dziennych
Vizir_Dosia_E	3 tyg.	0.420489022	1.735521604
	1 tyd.	-0.947843753	-0.704110451
Vizir_Dosia_Persil	3 tyg.	-0.837937223	-0.710731554
	1 tyd.	-1.139358953	-1.802573527
Dosia_E_Persil	3 tyg.	-0.173967986	0.495826574
	1 tyd.	-1.144567868	-1.080239052
Vizir_E_Persil	3 tyg.	-0.295810713	1.350562113
	1 tyd.	-0.855872712	-0.658048947

Źródło: obliczenia własne.

Warto zwrócić uwagę na to, że wartości skumulowanych różnic dla modelu dla danych dziennych oraz tygodniowych są w zdecydowanej większości bardzo podobne. W 3 przypadkach występuje niezgodność znaków tych różnic, choć w przypadku wyników dla scenariusza zakładającego równoległą promocję marki E oraz Persil moduły różnic ilości sprzedanych opakowań po 3 tygodniach nieznacznie różnią się od 0. Występują natomiast wartości, które dla obu modeli są bardzo podobne lub wręcz niemal identyczne.

5. Podsumowanie

Analizując uzyskane wyniki należy zadać pytanie: w oparciu o który model należy dokonywać symulacji wyników projektowanych akcji promocyjnych? Odpowiedź będzie tutaj dwojaka. Jeżeli symulacje mają dać odpowiedź jedynie na pytanie o globalne efekty promocji (w większości przypadków okres cenowych promocji sprzedaży stanowi wielokrotność tygodnia), to wystarczy skorzystać z modelu dla danych tygodniowych. Przemawiają za tym przede wszystkim względy praktyczne związane z koniecznością oszacowania znacznie mniejszej liczby parametrów. Jeśli natomiast symulacja ma dać odpowiedź na pytanie nie tylko o globalne efekty promocji, ale także o dokładny przebieg reakcji konsumentów (co może być ważne przy planowaniu stanów magazynowych czy liczby pracowników obsługujących poszczególne stoiska), to należy wykorzystać model oszacowany na podstawie danych dziennych. Agregacja danych dziennych powoduje bowiem zatarcie części informacji, zatem modele dla danych zagregowanych nie pozwolą na dokładny opis reakcji konsumentów na impulsy cenowe.

Literatura

- Błażejowski M., (2006a), Econometric Model of „Promotion Bubble”: Identification, Analysis and Application, w: Zieliński Z. (red.), *Dynamic Econometric Models*, Wydawnictwo UMK, Toruń 2006, vol. 7.
- Błażejowski M., (2006b), Wykorzystanie ekonometrycznych modeli sprzedaży do opisu „bąbli promocyjnych” i budowy scenariuszy cenowych akcji promocyjnych, w: Dittman, P., Krupowicz, J. (red.), *Prognozowanie w zarządzaniu firmą*, Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, nr 1112, Wydawnictwo AE im. Oskara Langego we Wrocławiu, Wrocław 2006.
- Błażejowski M., Kufel T. (2007), Projektowanie scenariuszy promocji cenowych z wykorzystaniem wielorównaniowych ekonometrycznych modeli popytu konsumpcyjnego, *Zeszyt Taksonomia* 14, Wydawnictwo AE we Wrocławiu (w druku).
- Kufel T., Błażejowski M., (2006), Ekonometryczne modelowanie procesów sprzedaży dla danych o wysokiej częstotliwości obserwowania, w: Pociecha J. (red.), *Prze-strzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*, Wydawnictwo AE w Krakowie, Kraków 2007, 195–210.