METODY ILOŚCIOWE
W BADANIACH EKONOMICZNYCH

QUANTITATIVE METHODS IN ECONOMICS

Vol. XIII, No. 3

Szkoła Główna Gospodarstwa Wiejskiego
Wydział Zastosowań Informatyki i Matematyki
Katedra Ekonometrii i Statystyki

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH

Tom XIII/3

Redaktor Naczelny
Bolesław Borkowski

Warszawa 2012

KOMITET REDAKCYJNY

Prof. Zbigniew Binderman – przewodniczący, Prof. Bolesław Borkowski, Prof. Leszek Kuchar, Prof. Wojciech Zieliński, Prof. Stanisław Gędek, Dr hab. Hanna Dudek, Dr Agata Binderman – sekretarz.

RADA NAUKOWA

Prof. Bolesław Borkowski – przewodniczący (SGGW w Warszawie),
Prof. Zbigniew Binderman (SGGW w Warszawie),
Prof. Paolo Gajo (University of Florence),
Prof. Evgeny Grebenikov (Computing Centre of Russia Academy of Sciences, Moscow),
Prof. Yuriy Kondratenko (Black Sea State University, Ukraine),
Prof. Vassilis Kostoglou (Alexander Technological Educational Institute of Thessaloniki),
Prof. Robert Kragler (University of Applied Sciences, Weingarten, Germany),
Prof. Yochanan Shachmurove (The City College of The City University of New York);
Prof. Alexander N. Prokopenya (Brest University, Belarus),
Dr Monika Krawiec – sekretarz (SGGW w Warszawie).

SKŁAD I ŁAMANIE: Dr Jolanta Kotlarska, Dr Elżbieta Saganowska

RECENZENCI WSPÓŁPRACUJĄCY Z REDAKCJĄ

Prof. Iacopo Bernetti (University of Florence)
Prof. Agata Boratyńska (Szkoła Główna Handlowa w Warszawie)
Prof. Paolo Gajo (University of Florence)
Prof. Yuiry Kondratenko (Black Sea State University, Ukraine)
Prof. Vassilis Kostoglou (Alexander Technological Educational Institute of Thessaloniki, Greece),
Prof. Karol Kukuła (Uniwersytet Rolniczy w Krakowie)
Prof. Wanda Marcinkowska – Lewandowska (Szkoła Główna Handlowa w Warszawie)
Prof. Yochanan Shachmurove (The City College of the City University of New York)
Prof. Ewa Marta Syczewska (Szkoła Główna Handlowa w Warszawie)
Prof. Dorota Witkowska (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie)
Prof. Wojciech Zieliński (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie)
Dr Lucyna Błażejczyk – Majka (Uniwersytet Adama Mickiewicza w Poznaniu)
Dr Michaela Chocolata (University of Economics in Bratislava, Slovakia)

Redaktor językowy: Agata Kropiwiec
Redaktor statystyczny: Prof. Wojciech Zieliński
Native speaker: Prof. Yochanan Shachmurove
Redaktorzy tematyczni:
ekonometria i statystyka – Prof. Bolesław Borkowski
inżynieria finansowa – Prof. Dorota Witkowska
wielowymiarowa analiza danych – Prof. Wiesław Szczesny
ekonomia matematyczna – Prof. Zbigniew Binderman

ISSN 2082 – 792X

© Copyright by Katedra Ekonometrii i Statystyki SGGW, Warszawa 2012

Wersją pierwotną jest wersja papierowa

Druk: Agencja Reklamowo-Wydawnicza A. Grzegorczyk, www.grzeg.com.pl

SPIS TREŚCI

Iwona Bąk – Wykorzystanie drzew klasyfikacyjnych i regresyjnych do analizy wyjazdów turystycznych gospodarstw domowych seniorów w Polsce 7

Aneta Becker – Zastosowanie metody ELECTRE TRI do oceny zróżnicowania regionalnego Polski pod względem wyposażenia w nowoczesne technologie ICT w 2011 roku 18

Jarosław Becker – Integracja metody AHP i ELECTRE TRI – rozwiązania decyzyjne w systemie DSS (część II) 27

Katarzyna Bień-Barkowska – Model sekwencyjnego zawierania transakcji – zastosowanie do analizy procesu transakcyjnego na kasowym rynku złotego 42

Agata Binderman – Rozwój polskiego rolnictwa w kontekście regionalnego zróżnicowania w latach 1998-2010 52

Krzysztof Gajowniczek, Tomasz Ząbkowski – Problemy modelowania rezygnacji klientów w telefonii komórkowej 65

Stanisław Gędek – Kształtowanie się kursu korony szwedzkiej wobec podstawowych walut światowych 79

Ewa Małgorzata Gnatowska – O własnościach szeregów czasowych i płynności akcyjnych funduszy ETF notowanych na europejskich giełdach 89

Rumiana Górska, Dorota Narojczyk – Badanie związków między zmiennymi realnymi a zmiennymi nominalnymi w polskiej gospodarce w latach 1997-2011 97

Andrzej Karpio – Kilka uwag dotyczących stopy zwrotu w terminie do wykupu 107

Maria Kaźmierska-Zatoń, Wojciech Zatoń – Zastosowanie kaskad informacyjnych w badaniu zachowań konsumentów 117

Małgorzata Kobylińska – Wykorzystanie zanurzania obserwacji w próbie do konstrukcji kart kontrolnych 127

Joanna Landmesser – Analiza karier równoległych z wykorzystaniem nieparametrycznego modelu ryzyk konkurencyjnych 137

Lidia Luty – Zróżnicowanie województw Polski pod względem struktury obszarowej gospodarstw ekologicznych 149

Małgorzata Machowska-Szewczyk, Agnieszka Sompolska-Rzechuła – Ocena stabilności wyników klasyfikacji województw Polski pod względem poziomu życia ludności 159

Mieczysław Połoński – Prognozowanie czasu zakończenia inwestycji na podstawie jej bieżącego zaawansowania 169

Agnieszka Sompolska-Rzechuła – Wpływ metody doboru cech na efektywność klasyfikacji na przykładzie analizy jakości życia w świetle zrównoważonego rozwoju..... 180

Elżbieta Soszyńska – Konwergencja technologiczna a potencjał społeczno-technologiczny krajów Unii Europejskiej..... 191

Maria Szmuksta-Zawadzka, Jan Zawadzki – O metodzie prognozowania brakujących danych w dziennych szeregach czasowych z lukami systematycznymi 202

Ryszard Szupiluk – Dywergencje Bosego-Einsteina w analizie podobieństw finansowych szeregów czasowych 213

Marek Szymański – Stabilność parametrów modelu rynkowego szacowanego w oparciu o stopy zwrotu WIG 222

Stanisław Urban, Anna Kowalska – Ocena statystyczna rynku usług gastronomicznych w Polsce 232

Tomasz Warowny – O pewnych kryteriach inwestowania w opcje na akcje 243

Adam Waszkowski – Estymacja krzywej dochodowości stóp procentowych dla Polski 253

Dorota Witkowska – Budowa mierników syntetycznych do oceny efektywności europejskich funduszy inwestycyjnych 262

Tomasz Ząbkowski, Wiesław Szczesny – Badanie atrakcyjności oferty dostępu do internetu za pomocą analizy gradacyjnej 276

Monika Zielińska-Sitkiewicz – Zastosowanie metod wielowymiarowej analizy dyskryminacyjnej do oceny kondycji firm deweloperskich 288

Paweł Ziemba, Ryszard Budziński – Selekcja kryteriów oceny jakości serwisów internetowych z wykorzystaniem pojemności informacyjnej Hellwiga 302

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 7 – 17

WYKORZYSTANIE DRZEW KLASYFIKACYJNYCH I REGRESYJNYCH DO ANALIZY WYJAZDÓW TURYSTYCZNYCH GOSPODARSTW DOMOWYCH SENIORÓW W POLSCE

Iwona Bąk
Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: iwona.bak@zut.edu.pl

Streszczenie: W artykule przedstawiono wyniki badań dotyczące klasyfikacji wyjazdów turystycznych seniorów ze względu na rodzaj wyjazdu oraz segmentacji gospodarstw domowych seniorów w Polsce ze względu na ich uczestnictwo w ruchu turystycznym. W badaniu uwzględniono indywidualne wyjazdy zrealizowane przez gospodarstwa domowe seniorów w okresie 1.10.2008-30.09.2009. Do klasyfikacji wyjazdów turystycznych ze względu na rodzaj wyjazdu wykorzystano drzewa klasyfikacyjne, natomiast do segmentacji gospodarstw domowych wykorzystano drzewo regresyjne.

Słowa kluczowe: gospodarstwa domowe seniorów, wyjazdy turystyczne, drzewa klasyfikacyjne i regresyjne

WSTĘP

W społeczeństwach wielu krajów świata, w tym Polski, coraz większą i znaczną część stanowi pokolenie ludzi starszych będących w wieku emerytalnym lub do niego się zbliżających. Zmiany zachodzące w strukturze demograficznej społeczeństw i stale rosnąca liczba osób starszych spowodowała, że producenci towarów i usług zaczęli dostosowywać ich rodzaje i asortymenty do potrzeb osób tej grupy wiekowej [Bąk 2011]. Współcześni seniorzy różnią się od swoich poprzedników tym, że żyją dłużej, są bogatsi, lepiej wykształceni i są zdrowsi. Także inaczej patrzą na wiek emerytalny – chcą realizować swoje marzenia i rozwijać pasje, a rynek towarów, a zwłaszcza usług w tym im pomaga. Rosnąca liczba osób starszych, a szczególnie przebywających na emeryturze, może stać się ważnym czynnikiem rozwoju szeroko pojętego przemysłu turystycznego. Profil demograficzny poszczególnych segmentów rynku turystycznego będzie

w przyszłości wyraźnie różnicować ofertę usługową przedsiębiorstw turystycznych. Zdaniem wielu autorów, nabycie przez osoby w wieku emerytalnym umiejętności zagospodarowania czasu wolnego i prowadzenie przez nie aktywnego (na miarę ich sił i potrzeb) stylu życia umożliwi zachowanie dobrej sprawności psychofizycznej w starości i czerpanie radości także i z tej fazy ich życia [Trafiałek 2006]. Pozytywny wpływ aktywności turystycznej na zdrowie i samopoczucie starszych osób jest niezaprzeczalny i potwierdzony wieloma wynikami badań.

Polscy seniorzy, na tle seniorów zachodnioeuropejskich czy amerykańskich, a także innych grup wiekowych własnego kraju wykazują stosunkowo niski poziom konsumpcji turystycznej. Główną przyczyną tej sytuacji jest nie tylko znacznie niższa siła nabywcza współczesnego, polskiego seniora, ale również wyznawany przez niego system wartości, w którym turystyka oraz inne formy aktywności spędzania czasu wolnego nie zajmują poczesnego miejsca. Biorąc pod uwagę rosnące zasoby czasu wolnego oraz zwiększającą się liczbę osób w tym wieku należy sądzić, że będą one stymulująco oddziaływać na kształtowanie koniunktury na rynku usług turystycznych. W tym celu niezbędne jest podejmowanie badań w sferze konsumpcji turystycznej, które dostarczą istotnych informacji o motywach i zachowaniach konsumpcyjnych turystów-seniorów oraz szacunkowych wielkościach środków, które mogą oni wydatkować na wypoczynek.

W artykule sformułowano dwa cele badawcze. Pierwszy z nich dotyczy klasyfikacji wyjazdów turystycznych seniorów[1] ze względu na rodzaj wyjazdu (krajowy, zagraniczny), a tym samym wskazania tych zmiennych niezależnych (predyktorów), które dzielą próbę na najbardziej homogeniczne klasy pod względem wyjazdów. Natomiast cel drugi to segmentacja gospodarstw domowych seniorów (głowa gospodarstwa była w wieku 60 lat i więcej) w Polsce ze względu na ich uczestnictwo w ruchu turystycznym. Jako narzędzia badawcze wykorzystano drzewa klasyfikacyjne (do klasyfikacji wyjazdów turystycznych) i regresyjne (do segmentacji gospodarstw domowych).

Dane statystyczne na temat turystyki wyjazdowej seniorów zaczerpnięto z badań ankietowych *„Turystyka i wypoczynek w gospodarstwach domowych”* przeprowadzonych przez Główny Urząd Statystyczny w 2009 roku. Dane mają charakter reprezentacyjny i pochodzą z badań cyklicznych przeprowadzanych co cztery lata. W badaniu uwzględniono indywidualne wyjazdy zrealizowane przez gospodarstwa domowe seniorów w okresie 1.10.2008-30.09.2009.

[1] W pracy zamiennie używane są pojęcia takie jak: „osoby starsze” i „seniorzy” w odniesieniu do osób powyżej 60 roku życia.

ISTOTA DRZEW KLASYFIKACYJNYCH I REGRESYJNYCH

Drzewa klasyfikacyjne i regresyjne zaliczane są do metod statystycznej analizy wielowymiarowej. Znajdują zastosowanie do klasyfikacji obiektów wówczas, gdy w zbiorze badanych zmiennych można wyróżnić zmienną zależną, a badane zmienne (zależna i niezależne) mogą być mierzone zarówno na skalach słabych (nominalna, porządkowa), jak i na skalach mocnych (przedziałowa, ilorazowa) [Gatnar, Walesiak 2004, s. 56-59].

Drzewa klasyfikacyjne i regresyjne są graficzną reprezentacją modelu postaci [Gatnar 2008, s. 37-39]:

$$Y = f(\mathbf{x}_i) = \sum_{k=1}^{K} \alpha_k \boldsymbol{I}(\mathbf{x}_i \in R_k), \tag{1}$$

gdzie:
Y – zmienna zależna,
R_k ($k = 1,\ldots,K$, K – liczba segmentów) to podprzestrzenie (segmenty) przestrzeni zmiennych objaśniających $\mathbf{X}^L$ ($X_1,X_2,\ldots,$
X_L, L – liczba zmiennych objaśniających),
$\mathbf{x}_i = [x_{i1},x_{i2},\ldots,x_{iL}]$ – obserwacje ze zbioru rozpoznawalnego,
α_k – parametry modelu,
$\boldsymbol{I}$ – funkcja wskaźnikowa.

Sposób definiowania funkcji wskaźnikowej $\boldsymbol{I}$ zależy od charakteru zmiennych objaśniających ($X_1,\ldots,X_L$). Gdy zmienne te mają charakter metryczny, to każdy z segmentów R_k jest definiowany przez jego granice w przestrzeni $\mathbf{X}^L$ w następujący sposób:

$$\boldsymbol{I}(\mathbf{x}_i \in R_k) = \prod_{l=1}^{L} \boldsymbol{I}(v_{kl}^{(d)} \le x_{il} \le v_{kl}^{(g)}), \tag{2}$$

gdzie wartości $v_{kl}^{(d)}$ i $v_{kl}^{(g)}$ oznaczają odpowiednio górną oraz dolną granicę odcinka w l-tym wymiarze przestrzeni.

Jeżeli zmienne $X_1,\ldots,X_L$ mają charakter niemetryczny, to podprzestrzeń R_k można zdefiniować jako

$$\boldsymbol{I}(\mathbf{x}_i \in R_k) = \prod_{l=1}^{L} \boldsymbol{I}(x_{il} \in B_{kl}), \tag{3}$$

gdzie B_{kl} to podzbiór zbioru kategorii zmiennej X_l, tj. $B_{kl} \subseteq V_l$.

Jeżeli zmienna zależna Y w modelu (1) jest zmienną nominalną, to taki model nazywamy dyskryminacyjnym i reprezentuje go drzewo klasyfikacyjne. Parametry α_k dla tego modelu wyznaczamy jako

$$\alpha_k = \arg\max_j p(C_j / \mathbf{x}_i \in R_k), \tag{4}$$

gdzie $p(C_j/\mathbf{x}_i \in R_k)$ oznacza prawdopodobieństwo *a posteriori*, że obserwacja z segmentu R_k należy do klasy C_j.

Jeżeli zmienna zależna Y w modelu (1) jest mierzona na skalach mocnych, to ten model jest modelem regresji, a jego graficzną postacią jest drzewo regresyjne. Parametry modelu regresji obliczamy według wzoru:

$$\alpha_k = \frac{1}{N(k)} \sum_{\mathbf{x}_i \in R_k} y_i \,, \tag{5}$$

gdzie: N(k) – liczba obserwacji znajdujących się w segmencie R_k, y_i – wartości przyjmowane przez zmienną zależną w segmencie R_k.

Do oceny jakości podziału przestrzeni zmiennych objaśniających $\mathbf{X}^L$ wykorzystuje się następujące miary:[2]

1. dla zmiennej zależnej niemetrycznej: błąd klasyfikacji, wskaźnik Giniego, miarę entropii, statystykę χ^2,
2. dla zmiennej zależnej metrycznej – wariancję zmiennej zależnej.

EMPIRYCZNE MODELE DRZEW KLASYFIKACYJNYCH I REGRESYJNYCH

W artykule do wyznaczenia drzew klasyfikacyjnych i regresyjnych wykorzystano procedurę CART. Obliczeń dokonano w programie Statistica 9.0 przy założeniach przedstawionych w tab. 1.

Tabela 1. Założenia przyjęte w procedurze CART

Wyszczególnienie	Modele ogólne drzew	
	klasyfikacyjnych	regresyjnych
Koszty błędnej klasyfikacji	równe	-
Miary dopasowania (reguła podziału)	wskaźnik Giniego	-
Kryterium stopu	przy błędnej klasyfikacji	przytnij według wariancji
Minimalna liczność	30	30
Maksymalna liczba węzłów	1000	1000

Źródło: opracowanie własne

[2] Sposoby wyznaczania i własności miar wykorzystywanych do oceny jakości podziału przestrzeni zmiennych są szeroko omówione w pracach [Gatnar 2001], [Gatnar, Walesiak 2004], [Gatnar 2008].

Do klasyfikacji wyjazdów turystycznych seniorów ze względu na rodzaj wyjazdu wykorzystano drzewa klasyfikacyjne. Jako zmienną zależną przyjęto rodzaj wyjazdu (krajowy, zagraniczny), natomiast w zbiorze zmiennych niezależnych uwzględniono[3]:

1. predyktory jakościowe:
- forma wyjazdu: wczasy, wycieczki (impreza objazdowa, pielgrzymka), inna (rodzina, działka);
- pośrednictwo w zakupie usług turystycznych: korzystał, nie korzystał;
- główny środek transportu wykorzystywany na dojazd: kolej, PKS lub inna autobusowa linia przewozowa, autokar, samochód osobowy, inny (samolot, prom);
- charakter odwiedzanego obszaru: obszar miejski (stolica, aglomeracje miejskie), miejscowość turystyczna, obszary górskie i wyżynne, obszary położone nad wodą (morze, akwen śródlądowy lub ciek wodny), uzdrowisko, obszar wiejski;
- cel wyjazdu: wypoczynek (rekreacja, wakacje), zwiedzanie (architektura,, kultura, przyroda), odwiedziny u krewnych lub znajomych oraz uroczystości rodzinne, zdrowotny, inny;

2. predyktor ilościowy: roczne wydatki ogółem poniesione w związku z wyjazdem.

Do segmentacji gospodarstw domowych wykorzystano drzewa regresyjne. Zmienną zależną zdefiniowano jako łączne roczne wydatki poniesione przez gospodarstwo domowe na wyjazdy turystyczne (wartości od 37 zł do 14860 zł), natomiast zbiór zmiennych niezależnych tworzyły:
- predyktory jakościowe: płeć (kobieta, mężczyzna), wykształcenie (bez wykształcenia, podstawowe, zasadnicze zawodowe, średnie, wyższe), miejsce zamieszkania (wieś, miasto poniżej 20 tys. mieszkańców, miasto od 20 do 99 tys. mieszkańców, miasto od 100 do 199 tys. mieszkańców, miasto od 200 do 499 tys. mieszkańców, miasto 500 tys. mieszkańców i więcej), typ biologiczny gospodarstwa domowego (małżeństwo bez dzieci, małżeństwo z dziećmi, gospodarstwo jednoosobowe, pozostałe), rodzaj wyjazdu (krajowy, zagraniczny);
- predyktory ilościowe: przeciętny miesięczny dochód gospodarstwa, wiek, liczba osób pracujących, liczba wyjazdów w ciągu roku.

W tabeli 2 przedstawiono procedurę wyboru drzewa regresyjnego oraz klasyfikacyjnego, które następnie wykorzystano do interpretacji wyników.

[3] Wszystkie zmienne niezależne zaproponowane do budowy drzew klasyfikacyjnych wykazują statystycznie istotne zależności z rodzajem wyjazdu. Porównaj badanie [Bąk, Wawrzyniak 2009].

W wyniku zastosowanej procedury CART otrzymano sekwencję 6 drzew klasyfikacyjnych i 19 drzew regresyjnych. Następnie na podstawie analizy wykresu przedstawiającego poziom kosztów sprawdzianu krzyżowego oraz kosztów resubstytucji na tle złożoności drzewa, wybrano dla każdego rodzaju drzew po trzy drzewa optymalne (kryterium – najmniejsza różnica między kosztem sprawdzianu krzyżowego a kosztem resubstytucji)[4]. W kolejnym etapie drzewa optymalne oceniono pod względem złożoności oraz liczby i ważności wykorzystanych przy podziale predyktorów [Batóg, Mojsiewicz, Wawrzyniak 2011, s. 169]. Za najlepsze uznano drzewo klasyfikacyjne nr 3 oraz drzewo regresyjne nr 14.

Interpretując wybrane drzewo klasyfikacyjne (rys. 1) sformułowano następujące wnioski wykorzystując w tym celu regułę zdań warunkowych typu *„jeżeli…, to…”*:

- jeżeli formą wyjazdu była wycieczka (impreza objazdowa, pielgrzymka) mająca na celu zwiedzanie (architektury, kultury, przyrody), to wyjazd był wyjazdem zagranicznym; natomiast, gdy wyjeżdżano w innym celu, to wyjazd był wyjazdem krajowym;
- jeżeli formą wyjazdu były wczasy lub inna forma (rodzina, działka), a środkiem transportu była kolej, PKS lub inna autobusowa linia przewozowa, to wyjazd był wyjazdem krajowym; natomiast gdy korzystano z innego środka transportu (samolot, prom), to wyjazd był wyjazdem zagranicznym.

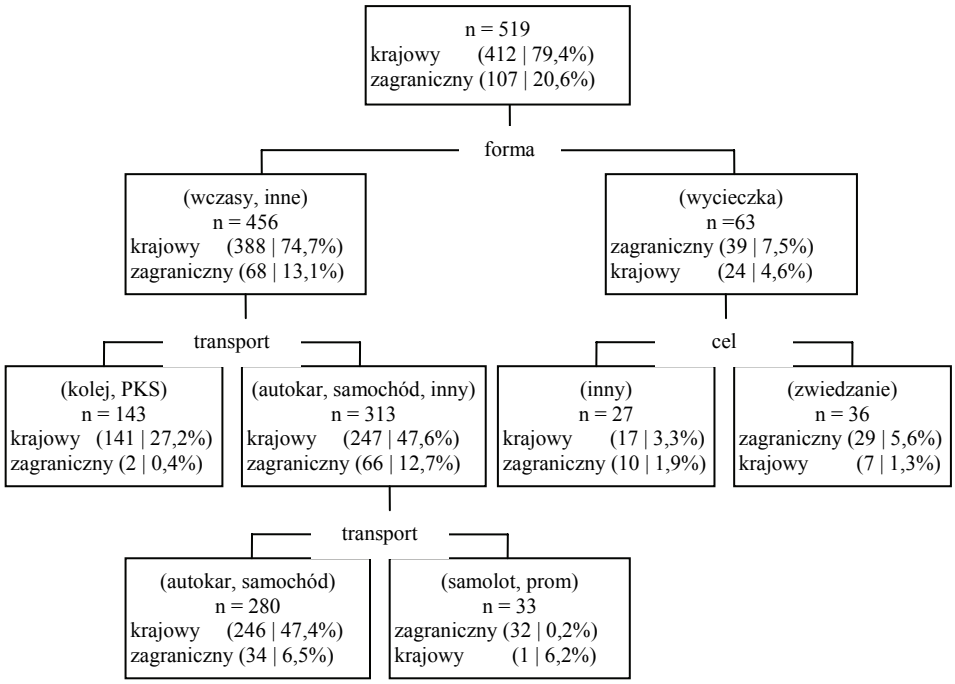
Tabela 2. Drzewa optymalne wybrane ze względu na zbieżność kosztów sprawdzianu krzyżowego oraz kosztów resubstytucji

Wyszczególnienie	Modele ogólne drzew	
	klasyfikacyjnych	regresyjnych
Liczba drzew w sekwencji drzew	6	19
Numery drzew optymalnych	5, 3,2	18, 16, 14
Numer drzewa wybranego	3	14
– liczba węzłów dzielonych		3
– liczba węzłów końcowych		6
– ważność predyktorów	1. forma 2. transport 3. cel 4. pośrednictwo 5. wydatki 6. charakter odwiedzanego obszaru	1. rodzaj wyjazdu 2. dochód 3. liczba wyjazdów 4. miejsce zamieszkania 5. wykształcenie 6. wiek 7. typ rodziny 8 liczba pracujących 9. płeć

Źródło: opracowanie własne

[4] Program *Statistica* za najlepsze drzewa uznał drzewo klasyfikacyjne o nr 5 i drzewo regresyjne o nr 18, czyli drzewa o jednym węźle dzielonym i o dwóch węzłach końcowych.

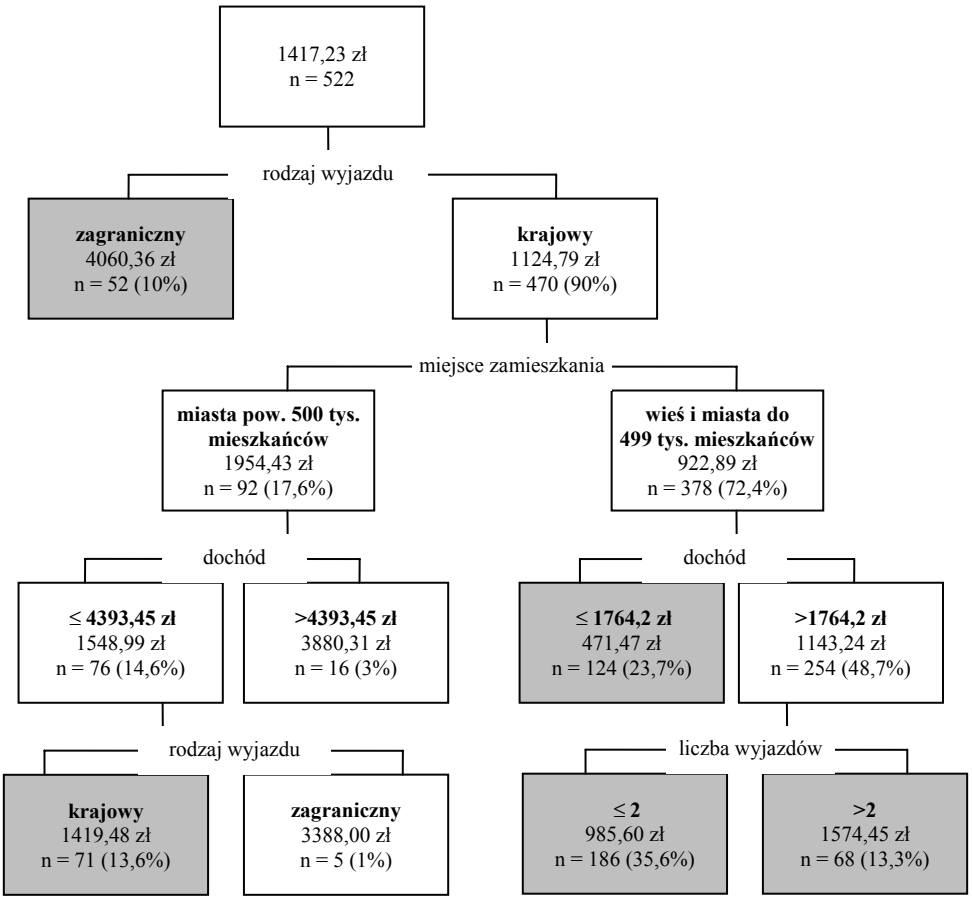
Rysunek 1. Drzewo klasyfikacyjne dla rodzaju wyjazdu (w nawiasach podano odsetki obliczone w stosunku do liczebności całej próby)



Źródło: opracowanie własne

Na podstawie drzewa regresyjnego nr 14 wydzielonych zostało 7 segmentów gospodarstw domowych ze względu na ich roczne wydatki na wyjazdy turystyczne. Przy wyborze segmentu do interpretacji przyjęto założenie, że jego liczebność powinna stanowić przynajmniej 10% liczebności próby. Warunek ten spełnia tylko pięć węzłów końcowych, które na rys. 2 zostały wyróżnione szarym kolorem. Wyniki segmentacji zamieszczono w tab. 3.

Rysunek 2. Drzewo regresyjne dla rocznych wydatków ogółem poniesionych w związku z wyjazdem turystycznym w ciągu roku (w nawiasach podano odsetki obliczone w stosunku do liczebności całej próby)



Źródło: opracowanie własne

Tabela 3. Charakterystyka segmentów gospodarstw domowych seniorów ze względu na ich wydatki na wyjazdy turystyczne

Nr	Charakterystyka segmentu	Przeciętne roczne wydatki na wyjazdy turystyczne (zł)	Liczebność segmentu (% badanej próby)
1	Gospodarstwa domowe biorące udział w wyjazdach zagranicznych	4060,37	52 (10,0%)
2	Gospodarstwa domowe o miesięcznych dochodach nie wyższych niż 4393,45 zł., zlokalizowane w miastach powyżej 500 tys. mieszkańców biorące udział w wyjazdach krajowych	1419,48	71 (13,6%)
3	Gospodarstwa domowe o miesięcznych dochodach nie wyższych niż 1764,47 zł., zlokalizowane na wsi lub w miastach do 499 tys. mieszkańców biorące udział w wyjazdach krajowych	471,47	124 (23,8%)
4	Gospodarstwa domowe o miesięcznych dochodach powyżej 1764,47 zł., zlokalizowane na wsi lub w miastach do 499 tys. mieszkańców biorące udział w co najwyżej dwóch wyjazdach krajowych	985,60	186 (35,63%)
5	Gospodarstwa domowe o miesięcznych dochodach powyżej 1764,47 zł., zlokalizowane na wsi lub w miastach do 499 tys. mieszkańców biorące udział w co najmniej dwóch wyjazdach krajowych	1574,46	68 (13,0%)

Źródło: opracowanie własne

PODSUMOWANIE

Zastosowane w artykule drzewa klasyfikacyjne umożliwiły wykrycie tych predyktorów, które w sposób istotny dzielą próbę na jednorodne klasy ze względu na rodzaj wyjazdu. Najistotniejszymi predykatorami w tym przypadku okazały się następujące zmienne niezależne: forma wyjazdu, cel wyjazdu i środek transportu. Wśród wyjazdów turystycznych realizowanych przez gospodarstwa domowe seniorów w okresie 1.10.2008-30.09.2009 dominowały wyjazdy krajowe na wczasy, do rodziny lub działkę, a głównym środkiem transportu była kolej, PKS lub samochód. Natomiast wyjazd zagraniczny był organizowany w formie wycieczki mającej na celu zwiedzanie architektury, kultury, przyrody a środkiem transportu był samolot.

Wykorzystanie drzew regresyjnych pozwoliło na wydzielenie segmentów gospodarstw domowych, które znacząco różniły się pod względem poziomu przeciętnych rocznych wydatków na wyjazdy turystyczne. Okazało się, że najliczniejszą grupę gospodarstw domowych seniorów (ponad 35%) stanowiły gospodarstwa wydające na turystykę przeciętnie w ciągu roku około 986 zł i są to gospodarstwa, których członkowie biorą udział w co najwyżej dwóch wyjazdach krajowych. Dogłębna analiza tego segmentu wykazała, że znajdujące się w nim gospodarstwa najczęściej korzystają z wyjazdów krajowych krótkoterminowych (2-4 dni) mających na celu odwiedziny krewnych lub znajomych. Najwyższe wydatki na wyjazdy ponoszą gospodarstwa biorące udział w wyjazdach zagranicznych, z tym, że liczebność tego segmentu jest nieznaczna i wynosi tylko 52 gospodarstwa, co stanowi ok. 10% wszystkich gospodarstw aktywnych turystycznie.

Uzyskane wyniki badań pozwalają poznać preferencje gospodarstw domowych seniorów w zakresie rodzaju wyjazdu oraz przeciętnych miesięcznych wydatków na wyjazdy turystyczne. Powyższe informacje mogą pozwolić odpowiednim organizacjom (np. Klubom Seniora), stowarzyszeniom i biurom turystycznym na przygotowanie oferty odpowiedniej do oczekiwań badanych zbiorowości.

BIBLIOGRAFIA

Batóg B., Mojsiewicz M., Wawrzyniak K. (2011) Segmentacja gospodarstw domowych ze względu na popyt potencjalny i zrealizowany na rynku ubezpieczeń życiowych w Polsce, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 176, Taksonomia 18. Klasyfikacja i analiza danych – teoria i zastosowania, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.

Bąk I., Wawrzyniak K. (2009) Zastosowanie analizy korespondencji w badaniach związanych z motywami wyboru rodzajów wyjazdów turystycznych przez emerytów i rencistów w 2005 roku, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 47, Taksonomia 16. Klasyfikacja i analiza danych – teoria i zastosowania, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.

Bąk I. (2011) Analiza wyjazdów turystycznych emerytów, Wiadomości Statystyczne nr 12, GUS, Warszawa.

Gatnar E. (2001) Nieparametryczna metoda dyskryminacji i regresji, Wydawnictwo Naukowe PWN, Warszawa.

Gatnar E. (2008) Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji. Wydawnictwo Naukowe PWN, Warszawa.

Gatnar E., Walesiak M. (2004) Metody statystycznej analizy wielowymiarowej w badaniach marketingowych, Wydawnictwo AE we Wrocławiu, Wrocław.

Trafiałek E. (2006) Starzenie się i starość. Wybór tekstów z gerontologii społecznej, Wszechnica Świętokrzyska, Kielce.

USING OF CLASSIFICATION AND REGRESSION TREES FOR ANALYSIS OF HOUSEHOLDS TOURIST TRIPS OF SENIORS IN POLAND

Summary: In the article the present results of the classification of touristic travels of seniors according to the kind of the travel and the segmentation of households of seniors according to their participation in tourist traffic. In the research were used the individual trips realized by seniors households in Poland in period 1.10.2008-30.09.2009. The classification of touristic travels of seniors was conducted by means of classification trees and the segmentation of households – by means of regression trees.

Keywords: households of seniors, tourist departures, classification and regression trees

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 18 – 26

ZASTOSOWANIE METODY ELECTRE TRI DO OCENY ZRÓŻNICOWANIA REGIONALNEGO POLSKI POD WZGLĘDEM WYPOSAŻENIA W NOWOCZESNE TECHNOLOGIE ICT W 2011 ROKU

Aneta Becker
Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: aneta.becker@zut.edu.pl

Streszczenie: Celem artykułu jest ocena zróżnicowania regionalnego Polski pod względem wyposażenia przedsiębiorstw w nowoczesne technologie ICT, w 2011 roku. W badaniach zastosowano metodę ELECTRE TRI opartą na modelu preferencji w postaci relacji przewyższania. Relacje te są określane na podstawie posiadanych informacji i reprezentują ściśle określone preferencje decydenta. Poza tym model ten nie narzuca przechodniości preferencji i dopuszcza sytuację nieporównywalności.

Słowa kluczowe: technologie teleinformatyczne (ICT), klasyfikacja, relacja przewyższania, metoda ELECTRE TRI

WSTĘP

Na przestrzeni ostatnich kilkudziesięciu lat rozwinęły się w nauce o podejmowaniu decyzji kierunki, które przyczyniły się do utworzenia europejskiej (francuskiej) szkoły podejmowania decyzji. Jej prekursorem jest prof. Bernard Roy [Dias, Mousseau 2003]. Badania prowadzone przez uczonych tej szkoły doprowadziły do opracowania nowej metodologii podejmowania decyzji oraz skonstruowania szeregu metod wielokryterialnych, wśród nich rodzinę metod ELECTRE, znajdujących szerokie zastosowanie w różnorodnych problemach decyzyjnych [Merad i in. 2004].

Można wyróżnić następujące problemy wielokryterialnego wspomagania decyzji, w których mają zastosowanie metody ELECTRE:

- wybór – wybór najlepszych (wyróżniających się) wariantów decyzyjnych, z punktu widzenia zbioru kryteriów,
- porządkowanie (rankingowanie) – uszeregowanie wszystkich wariantów od najlepszego do najgorszego, z punktu widzenia zbioru kryteriów,
- klasyfikację (sortowanie) – podział zbioru wariantów na klasy (kategorie), uporządkowane między sobą pod względem preferencji.[Merad i in. 2004; Trzaskalik 2006]

Wśród szerokiej gamy wielokryterialnych metod wspomagania decyzji metodą, która została specjalnie zaprojektowana do rozwiązywania zadań klasyfikacyjnych jest ELECTRE TRI. Podobnie jak cała rodzina metod ELECTRE opiera się ona na relacji przewyższania S. Wspomniana relacja jest binarna i mówi, że wariant a przewyższa wariant b (aSb), jeżeli biorąc pod uwagę dostępne informacje dotyczące preferencji decydenta, istnieją wyraźne przesłanki, żeby stwierdzić, iż wariant a jest co najmniej tak dobry jak wariant b i nie istnieją żadne istotne powody, aby stwierdzenie to odrzucić [Roy 1991; La Gauffre i in. 2007; Roy i Słowiński 2008; Figueira i in. 2009]. Należy nadmienić, że na temat preferencji, które jest podstawowym pojęciem w teorii ekonomii, szczególnie mikroekonomii (teorii wyboru konsumenta) odzwierciedlającym gusty konsumenta i pozwalającym dokonywać wyborów spośród rozmaitych alternatyw, obszernie pisze w swojej pracy między innymi E. Panek [2003] .

Celem artykułu jest klasyfikacja województw Polski ze względu na poziom wykorzystania technologii informacyjnych i telekomunikacyjnych (ICT – ang. *Information and Communication Technologies*) w przedsiębiorstwach. Informacje statystyczne wykorzystane w artykule pochodzą z GUS (Głównego Urzędu Statystycznego) i obejmują materiał z 2011 r. w ujęciu regionalnym. W badaniach klasyfikacyjnych zastosowano metodę wielokryterialnego wspomagania decyzji ELECTRE TRI. Wyniki grupowania mają na celu wyodrębnienie obszarów charakteryzujących się zbliżoną intensywnością korzystania z nowoczesnych rozwiązań teleinformatycznych, czyli technologii przetwarzających, gromadzących i przesyłających informacje w formie elektronicznej. W artykule weryfikuje się pogląd, że województwa wyróżniające się potencjałem rozwojowym w dużym stopniu wykorzystują technologie informacyjne i komunikacyjne.

METODA WIELOKRYTERIALNEGO WSPOMAGANIA DECYZJI OPARTA NA RELACJI PRZEWYŻSZANIA

Przykładem metody wspomagania decyzji, którą można stosować do zagadnień klasyfikacyjnych jest metoda ELECTRE TRI [Doumpos, Zopounidis 2002; La Gauffre i in. 2007]. Procedura koncentruje się na tzw. profilach separujących od siebie klasy. Każdy wariant decyzyjny (obiekt) można opisać ze względu na jego wartości na kryteriach (cechach). Danymi wejściowymi w metodzie ELECTRE TRI są wagi kryteriów oraz progi: nierozróżnialności, preferencji i veta. Należy

również podać liczbę klas i określić ich granice, czyli profile separujące. Działanie procedury obliczeniowej opiera się na wykonywaniu szeregu testów (zgodności i niezgodności) [La Gauffre i in. 2007; Roy i Słowiński 2008]. W metodzie tej dokonuje się porównań każdego wariantu decyzyjnego ze wszystkimi profilami separującymi klasy. W wyniku tych działań otrzymuje się szereg hipotez, na podstawie których podejmowana jest decyzja czy wariant przewyższa dane profile czy też nie. [Doumpos, Zopounidis 2002; Dias, Mousseau 2003; Merad i in. 2004].

W metodzie ELECTRE TRI umieszczanie obiektów w poszczególnych klasach następuje według dwóch komplementarnych procedur: optymistycznej i pesymistycznej. Klasyfikacja dokonywana jest na podstawie znajomości relacji przewyższania S dla każdej uporządkowanej pary (a, b_h), gdzie: a $(a = a_1, a_2, ..., a_n)$ jest wariantem decyzyjnym, a b_h profilem będącym górną granicą klasy (kategorii) C_h i dolną granicą klasy C_{h+1}, $h = 1,2,...,p$. Procedura optymistyczna polega na porównaniu wariantu a kolejno z profilami b_h $(h = 1,2,...,p-1,p)$, rozpoczynając od profilu najniższego (b_1). Jeśli b_h jest pierwszym napotkanym profilem takim, że $b_h Pa$ (co oznacza, że b_h jest preferowany nad a), to a zostaje przydzielone do klasy C_h. Z kolei w procedurze pesymistycznej następuje porównanie a kolejno z profilami b_h $(h = p, p-1,...,1,0)$, począwszy od profilu najwyższego (b_p). Jeśli b_h jest pierwszym napotkanym profilem takim, że aSb_h, to a zostaje przydzielone do klasy C_{h+1} ($h = 1,2,...,p$) [La Gauffre i in. 2007; Doumpos, Zopounidis 2002].

MATERIAŁ EMPIRYCZNY

W marcu 2012 r. Główny Urząd Statystyczny (GUS) opublikował wyniki badań przeprowadzonych w ramach statystyki społeczeństwa informacyjnego w Polsce. Badania te zostały zrealizowane w kwietniu 2011 r. i obejmowały główne obszary wykorzystania technologii informacyjnych i (tele)komunikacyjnych w przedsiębiorstwach i gospodarstwach domowych. W przypadku pomiarów dotyczących podmiotów gospodarczych, reprezentatywna próba liczyła około 14 tysięcy przedsiębiorstw, w których liczba pracujących wynosiła co najmniej 10, a prowadzona działalność gospodarcza zaklasyfikowana była według Polskiej Klasyfikacji Działalności (PKD) do następujących sekcji: C – przetwórstwo przemysłowe, F – budownictwo, G – handel hurtowy i detaliczny, naprawa pojazdów samochodowych włączając motocykle, H – transport i gospodarka magazynowa, I – działalność związana z zakwaterowaniem i usługami gastronomicznymi, J – informacja i komunikacja, L – działalność związana z obsługą rynku nieruchomości, M – działalność profesjonalna, naukowa i techniczna. [Wykorzystanie technologii 2012]

W analizowanym roku wśród podmiotów biorących udział w badaniu dominowały małe firmy (10-49 pracujących), które stanowiły nieco ponad 80% ogółu. Do grupy średnich przedsiębiorstw o liczbie pracujących od 50 do 249 osób należało 16,51% podmiotów. Natomiast najniższy odsetek firm reprezentowały duże przedsiębiorstwa, w których liczba pracujących wynosiła powyżej 250 osób – 3,32%. Najwięcej przedsiębiorstw, w stosunku do wszystkich badanych, pochodziło z województw: mazowieckiego (17%), śląskiego (12,69%) i wielkopolskiego (11,10%). Natomiast najmniej firm reprezentowało województwa: podlaskie (2,23%), opolskie (2,29%) i świętokrzyskie (2,56%). W układzie regionalnym NUTS-1 największy udział w badaniach miały firmy pochodzące z regionu centralnego (województwa: łódzkie i mazowieckie) – 23,76%, natomiast najmniejszy podmioty regionu południowo-zachodniego (województwa: dolnośląskie i opolskie), które stanowiły prawie 10% ogółu.

Według prezentowanych przez GUS informacji nieco ponad 97% firm, które brały udział w badaniu korzystało z komputerów. Dostęp do Internetu deklarowało prawie 94%, przy czym 77,5% spośród nich poprzez łącza szerokopasmowe, natomiast prawie 21% za pomocą wąskopasmowych łączy bezprzewodowych. Dominowały podmioty (94%) posiadające dostęp do Internetu łączami o prędkości od 2 do10 Mb/s. Z bezpłatnego oprogramowania *open source* korzystało 78% badanych podmiotów. W tym z bezpłatnych aplikacji ERP lub CRM nieco ponad 7%, natomiast prawie 21% z systemów operacyjnych (np. Linux). Najpopularniejsze były przeglądarki internetowe oraz aplikacje biurowe.

Spośród firm z dostępem do Internetu, prawie 65% posiadało własną stronę internetową. Podmioty gospodarcze najczęściej korzystały ze strony internetowej, aby zaprezentować katalogi wyrobów lub cenniki swoich towarów bądź usług (46,9%) i zapewnić ochronę danych osobowych lub do homologacji bezpieczeństwa (30,6%). Najrzadszą funkcją strony była personalizacja jej zawartości dla stałych użytkowników (8,1%) oraz sprawdzanie stanu realizacji zamówienia online (9,3%).

Nieznacznie ponad 66% firm korzystało z automatycznej wymiany danych z podmiotami zewnętrznymi. Na ogół w celu wymiany informacji z organami administracji publicznej (64,3%), rzadziej po to, aby wysyłać lub otrzymywać dokumenty transportowe (np. listy przewozowe) – 17%. Natomiast z automatycznej wymiany informacji wewnątrz przedsiębiorstwa korzystało 34,5% podmiotów, głównie w księgowości, sporadycznie w zarządzaniu dystrybucją. Z kolei 30,5% firm otrzymywało lub wysyłało e-faktury. Najbardziej rozpowszechnioną formą były e-faktury nienadające się do automatycznego przetwarzania np. e-mail z załącznikiem w formacie PDF. Według informacji GUS w 2011 r. blisko jedna trzecia przedsiębiorstw stosowała procedury, których celem było ograniczenie negatywnego wpływu na środowisko. Działania te dotyczyły zwłaszcza redukcji zużycia energii pochłanianej przez wyposażenie ICT – 24,6% oraz zużycia papieru – 23% podmiotów.

WYNIKI BADAŃ

W badaniach klasyfikacyjnych posłużono się materiałem empirycznym, zebranym i opublikowany przez GUS [Wykorzystanie technologii 2012]. W doświadczeniu brała udział odmienna liczba firm charakteryzująca poszczególne województwa, dlatego zgromadzone dane unormowano do wartości względnych. Spośród znacznej liczby zmiennych opisujących teleinformatykę wyłoniono następujący dziewięcioelementowy zbiór kryteriów, który obejmował odsetek przedsiębiorstw:

- X_1 – korzystających w swojej działalności z komputerów,
- X_2 – mających dostęp do Internetu,
- X_3 – posiadających własne strony WWW,
- X_4 – korzystających z automatycznej wymiany danych z podmiotami zewnętrznymi,
- X_5 – wykorzystujących wewnętrzną automatyczną wymianę danych,
- X_6 – korzystających z bezpłatnego oprogramowania *open source*,
- X_7 – otrzymujących lub wysyłających e-faktury,
- X_8 – z dostępem do Internetu poprzez łącza bezprzewodowe (wąsko lub szerokopasmowe),
- X_9 – posiadających procedury, których stosowanie ogranicza negatywny wpływ na środowisko.

Obliczenia wykonano z pomocą programu komputerowego ELECTRE TRI 2.0a. Za podział wynikowy uznano grupowanie na 3 klasy. Ze względu na liczbę grup (k=3), w omawianym problemie decyzyjnym, zdefiniowano dwa profile określające granice klas. Dla poszczególnych profili podano wartości progów: nierozróżnialności, preferencji i veta na każdym z dziewięciu kryteriów. Z uwagi na subiektywny charakter metody oraz jej założenia przeprowadzono szereg eksperymentów wykorzystując odmienny zestaw profili i progów. Przyjęto także następujące wartości wag dla poszczególnych kryteriów: w_1=3, w_2=4, w_3=5, w_4=8, w_5=5, w_6=6, w_7=7, w_8=8, w_8=9. Najniższą wagę nadano pierwszemu kryterium, ponieważ korzystanie z komputera w firmach jest współcześnie zjawiskiem powszechnym. Podobna sytuacja dotyczy dostępu do Internetu, który (w analizowanym okresie) znajdował się na wysokim poziomie i dla podmiotów, które chcą zajmować znaczącą pozycję na rynku konkurencji jest zjawiskiem oczywistym. Najwyższą wagę przypisano kryterium świadczącym o proekologicznym podejściu przedsiębiorstw w swojej działalności. Przykładem wspomnianego działania może być między innymi redukcja zużycia papieru (podczas kopiowania, drukowania) lub energii elektrycznej (np. wyłączanie komputerów o określonej porze dnia) oraz ograniczenia podróży służbowych przez wykorzystanie możliwości jakie dają wideokonferencje.

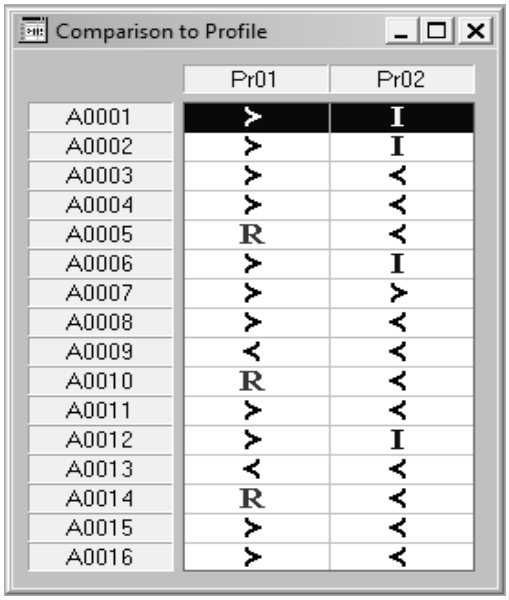
Zgodnie z relacją przewyższania aSb_h uznano, że jest ono autorytatywne, gdy współczynnik wiarygodności σ przekracza przyjęty w badaniu poziom cięcia

λ=0,76. Szczegółowy opis współczynników wykorzystywanych przy konstruowaniu procedury klasyfikacyjnej metody ELECTRE TRI zawierają między innymi prace: Dias i Mousseau (2003) oraz La Gauffre i in. (2007). Należy wyjaśnić, że współczynnik $\sigma(a,b_h)\in[0,1]$ określany jest mianem stopnia wiarygodności twierdzenia, że aSb_h – odpowiednio b_hSa. Przy czym λ jest to wcześniej określony poziom cięcia (próg odcięcia), który przyjmuje wartości z przedziału $[0{,}5;1]$. Przewyższanie w stopniu 1 jest w pełni poparte przez wszystkie kryteria. Porównanie σ i λ ujawnia występowanie czterech sytuacji, które zachodzą pomiędzy wariantami decyzyjnymi i profilami separującymi klasy.

- aIb_h; a jest nierozróżnialne z b_h,
- $a \succ b_h$; a jest preferowane wobec b_h,
- $a \prec b_h$: b_h jest preferowane wobec a,
- aRb_h; a jest nieporównywalne z b_h.

Rysunek 1 przestawia relacje jakie zachodziły pomiędzy szesnastoma wariantami decyzyjnymi (obiektami – województwami) a ustalonymi profilami oddzielającymi klasy dla eksperymentu, którego wyniki prezentuje tabela 1.

Rysunek 1. Relacje pomiędzy obiektami i profilami oddzielającymi klasy



Źródło: opracowanie własne na podstawie danych GUS z wykorzystaniem programu ELECTRE TRI 2.0a

Przydział województw do klas nastąpił na podstawie znajomości relacji pomiędzy wariantami decyzyjnymi i profilami według procedur: optymistycznej i pesymistycznej. Zaprezentowany w tabeli 1 wynik to wybrane, przykładowe rozwiązanie spośród wielu uzyskanych podziałów.

Tabela 1. Wyniki klasyfikacji województw według stopnia wykorzystania technologii teleinformatycznych uzyskane metodą ELECTRE TRI

Procedura optymistyczna			Procedura pesymistyczna		
klasa 1	klasa 2	klasa 3	klasa 1	klasa 2	klasa 3
podkarpackie	lubelskie	dolnośląskie	łódzkie	lubelskie	dolnośląskie
świętokrzyskie	lubuskie	kujawsko-pomorskie	podkarpackie	lubuskie	kujawsko-pomorskie
	łódzkie	małopolskie	podlaskie	opolskie	małopolskie
	opolskie	mazowieckie	świętokrzyskie	pomorskie	mazowieckie
	podlaskie	śląskie	warmińsko-mazurskie	wielkopolskie	śląskie
	pomorskie			zachodniopomorskie	
	warmińsko-mazurskie				
	wielkopolskie				
	zachodniopomorskie				

Źródło: opracowanie własne na podstawie danych GUS z wykorzystaniem programu ELECTRE TRI 2.0a

Numeracja grup związana była z hierarchią ważności skupień. Do skupienia pierwszego zaklasyfikowano województwa, w których wykorzystanie technologii teleinformatycznych było na najniższym poziomie, drugiej – województwa przeciętne, natomiast klasę trzecią reprezentowały obiekty, w których znajdują się firmy będące liderami wykorzystania technologii ICT. Analizując wyniki uzyskane w wyniku zastosowania procedury optymistycznej, województwa zostały rozdzielone pomiędzy trzy klasy. Przy czym większość obiektów zaklasyfikowano do grupy drugiej (przeciętnej), natomiast w grupie najsłabszej znalazły się tylko dwa województwa, a w trzeciej (najlepszej) uplasowało się pięć województw. Z kolei rozpatrując rezultaty uzyskane z pomocą podejścia pesymistycznego, które jest bardziej rygorystyczne, można zauważyć, że zbiór obiektów także został rozdzielony pomiędzy trzy klasy. Najwięcej, czyli sześć województw zaklasyfikowano do grupy przeciętnej (drugiej) i po pięć zawierały klasy: pierwsza i trzecia. Porównując wyniki uzyskane w konsekwencji zastosowania dwóch różnych procedur rozdziału obiektów do grup, można stwierdzić, że struktura poszczególnych klas jest zbliżona. W przypadku klasy trzeciej, zarówno w ujęciu optymistycznym jak i pesymistycznym występują te same województwa. Natomiast województwa: łódzkie, podlaskie i warmińsko-mazurskie, które według procedury optymistycznej znalazły się w drugim skupieniu, w ujęciu pesymistycznym pojawiły się w grupie najsłabszej. Oprócz wymienionych województw, które w zależności od procedury zmieniały swoje położenie, pozostałe regiony można określić mianem stabilnych. Oznacza to, że zdecydowanie przynależą do swoich grup, niezależnie od zastosowanego podejścia.

Decydent może wybrać pomiędzy podejściem pesymistycznym a optymistycznym. Na decyzję mają wpływ jego preferencje. Odmienności idei optymistycznej

i pesymistycznej jest zauważalna w końcowym przydziale obiektów do grup. Przykładem może być obiekt, który znajduje się pomiędzy dwoma profilami separującymi klasy. Gdy jest on nieporównywalny z profilem podejście pesymistyczne przypisuje go do niższej grupy niż optymistyczne.

PODSUMOWANIE

Metoda ELECTRE TRI pozwala zobrazować dysproporcje występujące pomiędzy województwami Polski ze względu na stopień wykorzystania teleinformatyki. Procedura ta wykorzystywana jest do klasyfikacji zbioru wariantów decyzyjnych (obiektów) na podzbiory (grupy). Zakłada się, że klasy są porównywalne w sensie preferencji. Można zestawić dwa skupienia i stwierdzić, że jedno jest lepsze od drugiego lub odwrotnie. Metoda ELECTRE wymaga od analityka znacznej wiedzy na temat analizowanego problemu decyzyjnego i umiejętności potrzebnych do określenia profili separujących klasy, nadania wag kryteriom oraz wyznaczenia progów. Ma charakter subiektywny, gdyż posługując się nią wiele zależy od preferencji decydenta, proponowanych przez niego ustawień parametrów wstępnych badania i ostatecznego wyboru propozycji rozwiązania.

Wyniki badań wskazują, że województwa charakteryzujące się wysokim poziomem: infrastruktury komunikacyjnej, nasycenia sprzętem informatycznym, potrafią odpowiednio wykorzystać technologie ICT w procesie zarządzania przedsiębiorstwem i umiejętnie przekazują informacje zwłaszcza w kontaktach z otoczeniem zostały zaklasyfikowane do najlepszej grupy (trzeciej). W klasie pierwszej znalazły się województwa o niskim potencjale rozwojowym co przekłada się na intensywność wykorzystywania technologii informacyjnych i komunikacyjnych.

BIBLIOGRAFIA

Dias L., Mousseau V. (2003) IRIS: A DSS for Multiple Criteria Sorting Problems, Journal of Multi-Criteria Decision Analysis, Nr 12, str. 286-287.

Doumpos M., Zopounidis C. (2002) Multi-criteria classification methods in financial and banking decision, International Transactions in Operational Research, Nr 9, str. 568, 570-571.

Figueira J. R., Greco S., Roy B. (2009) ELECTRE methods with interaction between criteria: An extension of the concordance index, European Journal of Operational Research, Nr 199, str. 481.

La Gauffre P., Haidar H., Poinard D, Laffréchine K., Baur R., Schiatti M.. (2007) A multicriteria decision support methodology for annual rehabilitation programs for water networks, Computer-Aided Civil and Infrastructure Engineering, Nr 22, str. 479, 488.

Merad M. M., Verdel T., Roy B., Kouniali S. (2004) Use of multi-criteria decision-aids for risk zoning and management of large area subjected to mining-induced hazards, Tunnelling and Underground Space Technology, Nr 19, str. 167, 172-173.

Panek E. (2003) Ekonomia matematyczna, Wyd. Akademii Ekonomicznej w Poznaniu, Poznań.

Roy B., Słowiński R (2008) Handing effects of reinforced preference and counter-veto in credibility of outranking, European Journal of Operational Research, Nr 188, str. 185, 186-187.

Roy B. (1991) The outranking approach and the foundations of ELECTRE methods, Theory and Decision Nr 31 (1), str. 49-73.

Trzaskalik T. (red.) (2006) Metody wielokryterialne na polskim rynku finansowym, PWE, Warszawa, str. 43.

Wykorzystanie technologii informacyjno-(tele)komunikacyjnych w przedsiębiorstwach i gospodarstwach domowych w 2011 r., GUS. Zasoby Internetu z dnia 30. 03. 2012 r.
http://www.stat.gov.pl/gus/5840_wykorzystanie_ict_PLK_HTML.htm

APPLICATION OF THE ELECTRE TRI METHOD FOR EVALUATION OF REGIONAL DIVERSITY OF POLAND IN TERMS OF NEW ICT TECHNOLOGIES APPLICATION IN 2011

Abstract: The aim of the article is the evaluation of regional diversity of Poland in terms of new ITC technology application in companies in 2011. Conducted research includes the ELECTRE TRI method that is based on a model of preferences in a form of outranking relations. Relations are defined according to information and represent strictly defined preferences of a decision maker. What is more, the model does not impose the transitivity of preferences and enables a situation of comparison impossibility.

Keywords: information and communication technology (ICT), classification, outranking relation, ELECTRE TRI method

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 27 – 41

INTEGRACJA METODY AHP I ELECTRE TRI – ROZWIĄZANIA DECYZYJNE W SYSTEMIE DSS (CZĘŚĆ II)

Jarosław Becker
Katedra Inżynierii Systemów Informacyjnych
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: jbecker@wi.zut.edu.pl

Streszczenie: W artykule skoncentrowano się na ukazaniu praktycznych aspektów prowadzonych badań naukowo-inżynierskich związanych z integracją metody AHP i ELECTRE TRI w informatycznym systemie wspomagania decyzji. W architekturze systemu DSS ujęto algorytmy obu metod. Działają one na wspólnym zbiorze obiektów (wariantów decyzyjnych), które można uporządkować (AHP) oraz przydzielić do zdefiniowanych grup (ELECTRE TRI). W części końcowej opracowania zilustrowano funkcjonalność metod w formie studium przypadku, generując rozwiązania decyzyjne przy użyciu opracowanego prototypu DSS 2.0.

Słowa kluczowe: metody ilościowe we wspomaganiu decyzji, wielokryterialna analiza decyzyjna, systemy wspomagania decyzji

WSTĘP

Artykuł jest kontynuacją badań obejmujących przegląd i analizę funkcjonalną metod ilościowych z obszaru WAD (Wielokryterialnej Analizy Decyzyjnej) na potrzeby budowy informatycznego systemu wspomagania decyzji.[1] Dyskusja dotyczyła funkcjonalności metod w systemie informatycznym, w aspekcie rozwiązywania złożonych (wielokryterialnych, wieloetapowych, grupowych, iteracyjnych i masowych) problemów decyzyjnych związanych z wyborem, uporządkowaniem i grupowaniem jednorodnych wariantów decyzyjnych (np. ujętych w formie wniosków, ofert). Autor uargumentował wybór metody WPL jako metody bazowej (charakteryzującej się dużą nośnością informacyjną), która posłużyła za platformę do definiowania problemów

[1] Artykuł opublikowano w zeszłorocznym wydaniu Czasopisma jako część I.

decyzyjnych oraz organizacji danych wejściowych, niezbędnych do zintegrowania i zastosowania komplementarnych metod (m. in. AHP i ELECTRE TRI) w systemie informatycznym.

Celem artykułu (część II) jest ukazanie inżynierskiej strony realizowanych badań nad integracją metody AHP i ELECTRE TRI w informatycznym systemie klasy DSS (wersja DSS 2.0). Obszerne omówienie wspomnianych metod zawierają pozycje: [Saaty 1980], [Saaty 2004], [La Gauffre i in. 2007], [Roy, Słowiński 2008]. W pracy skupiono się na modelu integracji danych wejściowych oraz procedurach ich preparacji. Funkcjonalność uzyskanych rozwiązań prototypowania zilustrowano na praktycznym przykładzie analizy decyzyjnej wniosków, które składają rolnicy starający się o pomoc finansową na działalność związaną z produkcją mleka. Badanie zrealizowano w formie studium przypadku na danych teoretycznych. Zachowano założenia formalnej procedury podziału środków finansowych, które ustanowiono w ramach działania „Modernizacja gospodarstw rolnych" objętego Programem Rozwoju Obszarów Wiejskich na lata 2007-2013.

IDEA INTEGRACJI METOD AHP I ELECTRE TRI W SYSTEMIE INFORMATYCZNYM WSPOMAGANIA DECYZJI

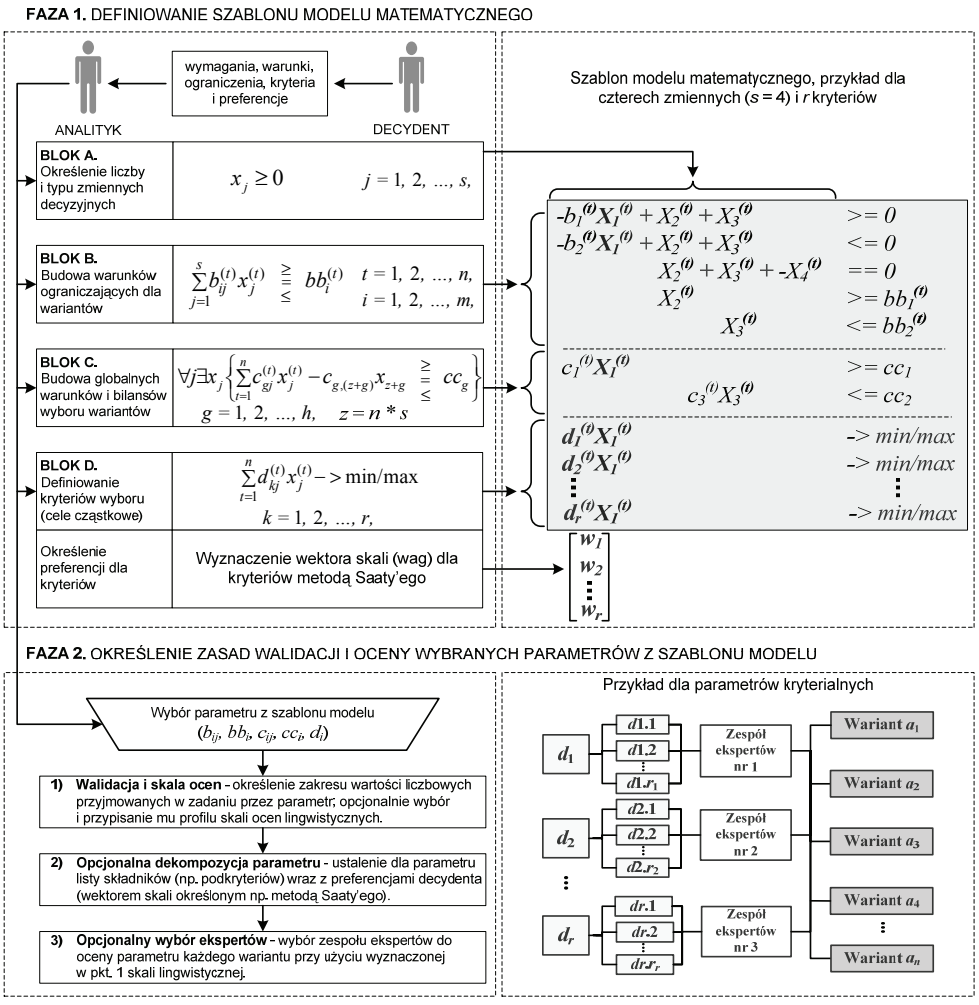
W literaturze można zauważyć wiele odwołań do pojęcia „procesy integracyjne". F. B. Vernadat (1996) rozumie ten termin jako, cyt.: *„połączenie niejednorodnych składników w całość tak, że współdziałając w ramach tej całości, wzmagają swoją skuteczność"*. Integracja metod w informatycznym systemie wspomagania decyzji polega najczęściej na wykorzystaniu ich funkcjonalności w ramach spójnego, logicznego i kompleksowego procesu. Uzyskany w ten sposób efekt synergii objawia się zmniejszeniem obszaru niepewności podejmowanych decyzji oraz wzrostem użytecznych informacji (podpowiedzi, sugestii działania).

W systemie DSS założono, że z każdym zadaniem decyzyjnym związany jest zbiór obiektów (wariantów decyzyjnych). Wspomaganie koncentruje się na możliwości wyboru tych obiektów (optymalizacja WPL), ich uporządkowaniu (ranking AHP) oraz przyporządkowaniu do zdefiniowanych klas (grupowanie ELECTRE TRI). Trzonem integracji metod w systemie DSS jest zunifikowana, blokowa struktura, stosowana do definiowania różnej postaci szablonów modeli matematycznych i pozwalająca na wyczerpujący opis (zbiór metadanych) rozpatrywanych, jednorodnych obiektów oraz kryteria, preferencje i reguły ich analizy. W wymiarze inżynierskim jest to platforma służąca do określenia struktur informacyjnych bazy modeli WPL (nazwana umownie platformą WPL, rys. 1), która narzuca pewien schemat postępowania podczas budowy modelu oraz dostarcza mechanizmy kontroli jego spójności i sprawności działania (badanie na danych testujących, czy projektowany układ dla zadania decyzyjnego nie jest sprzeczny). Powstały w efekcie projektowania na platformie WPL szablon modelu

matematycznego pozwala na dostarczenie na wejścia integrowanych metod (WPL, AHP i ELECTRE TRI; rys. 1, faza 1):

- zbioru warunków ograniczających bloku B, które charakteryzują wewnętrzną strukturę każdego *t*-tego wariantu za pomocą określonej w bloku A liczbą *s* zmiennych decyzyjnych x_j (typu: rzeczywistego, całkowitego lub binarnego), parametrów $b_{ij}^{(t)}$ oraz wartości ograniczających $bb_i^{(t)}$,

Rysunek 1. Formalizacja zadania decyzyjnego w systemie DSS (platforma WPL)



Źródło: opracowanie własne

- zbioru warunków ograniczających i bilansów bloku C, które wyznaczają zasady wyboru wariantów, np. łączną liczbę wybieranych wariantów lub inną graniczną wartość cc_g, dla sumy iloczynu wybranej zmiennej x_j i parametru $c_{gj}^{(t)}$,
- wektora wartości kryteriów (reprezentowany w bloku D jako zbiór r preferowanych celów cząstkowych) oraz odpowiadający mu wektor skali (wag w_k dla każdego k-tego celu cząstkowego, $k = 1, 2, \ldots, r$) wyznaczany przy użyciu metody Saaty'ego.

Wprowadzona w systemie DSS procedura walidacji parametrów modelu [Becker 2010] pozwala na kontrolę i wycenę każdego z nich, uwzględniając określenie przedziału dopuszczalnych wartości dla parametru (np. $\langle d_{k,min}, d_{k,max}\rangle$, $k = 1, 2, \ldots, r$) oraz możliwość wybór profilu skali ocen dla parametru (rys. 1, faza 2). Skali bazującej na określonej funkcji przynależności, o wybranej z zakresu od 2 do 11 liczbie stopni i indywidualnie określonej kafeterii (zbiorze konceptów lingwistycznych, np. mało, średnio, dużo). Istnieje również możliwość dekompozycji parametru na składniki oraz wybór zespołu ekspertów oceniających dany parametr (lub jego składowe) w każdym wariancie przy użyciu skali punktowej lub lingwistycznej.

Opracowany system DSS to rozwiązanie hybrydowe, które łączy techniki komputerowego przetwarzania danych z algorytmami różnych metod wspomagania decyzji. Na podstawie przygotowanej struktury informacyjnej (szablonu modelu matematycznego) dla zadania decyzyjnego wprowadzany jest do systemu zbiór jednorodnych obiektów (wariantów decyzyjnych: $a_1, a_2, \ldots, a_n$). Każdy wariant to zapis (rekord) w tabeli relacyjnej bazy danych i jednocześnie autonomiczny model matematyczny, inaczej sformalizowana postać zadania programowania linowego, które posiada rozwiązanie (rys. 2).

Optymalizacyjna procedura wyboru najlepszych, z punktu widzenia określonych preferencji, obiektów zakłada ulokowanie wskazanych rekordów jako submodeli w zadaniu głównym (rys. 2, krok A1). Z budową zadania związany jest zapis cząstkowych funkcji celu w postaci równań (blok D), ich normalizacja i ujęcie w funkcji użyteczności [Becker 2008]. Rywalizacja wariantów decyzyjnych w zadaniu głównym zmierza w kierunku osiągnięcia ustalonego przez decydenta wzorca preferencji, rozumianego jako optimum funkcji użyteczności. Niech $G(d^*)$ jest nadrzędną funkcją celu. W zbiorze decyzji dopuszczalnych D^* określony jest zbiór funkcji cząstkowych $f_1, f_2, \ldots, f_r$, które przyjmują wartości dodatnie. W zadaniu maksymalizuje się sumę wskaźników jakości, które przyjmują wartości ujemne w przypadku, gdy cząstkowa funkcja celu f_k dąży do minimum i nieujemne w sytuacji przeciwnej ($f_k \rightarrow max$). Wynikają one ze sprowadzenia syntezy preferowanych równań celów cząstkowych (rys. 2, krok A2)

$$\sum_{t=1}^{n} d_{kj}^{(t)} x_j^{(t)} - x_{(z+h+k)} = 0 \tag{1}$$

do postaci funkcji celu (rys. 2, krok A4)

$$G(d^*) = \sum_{k=1}^{r} f_k(d^*) = \sum_{k=1}^{r} w_k u_k x_{(z+h+k)} \rightarrow maksimum\,. \quad (2)$$

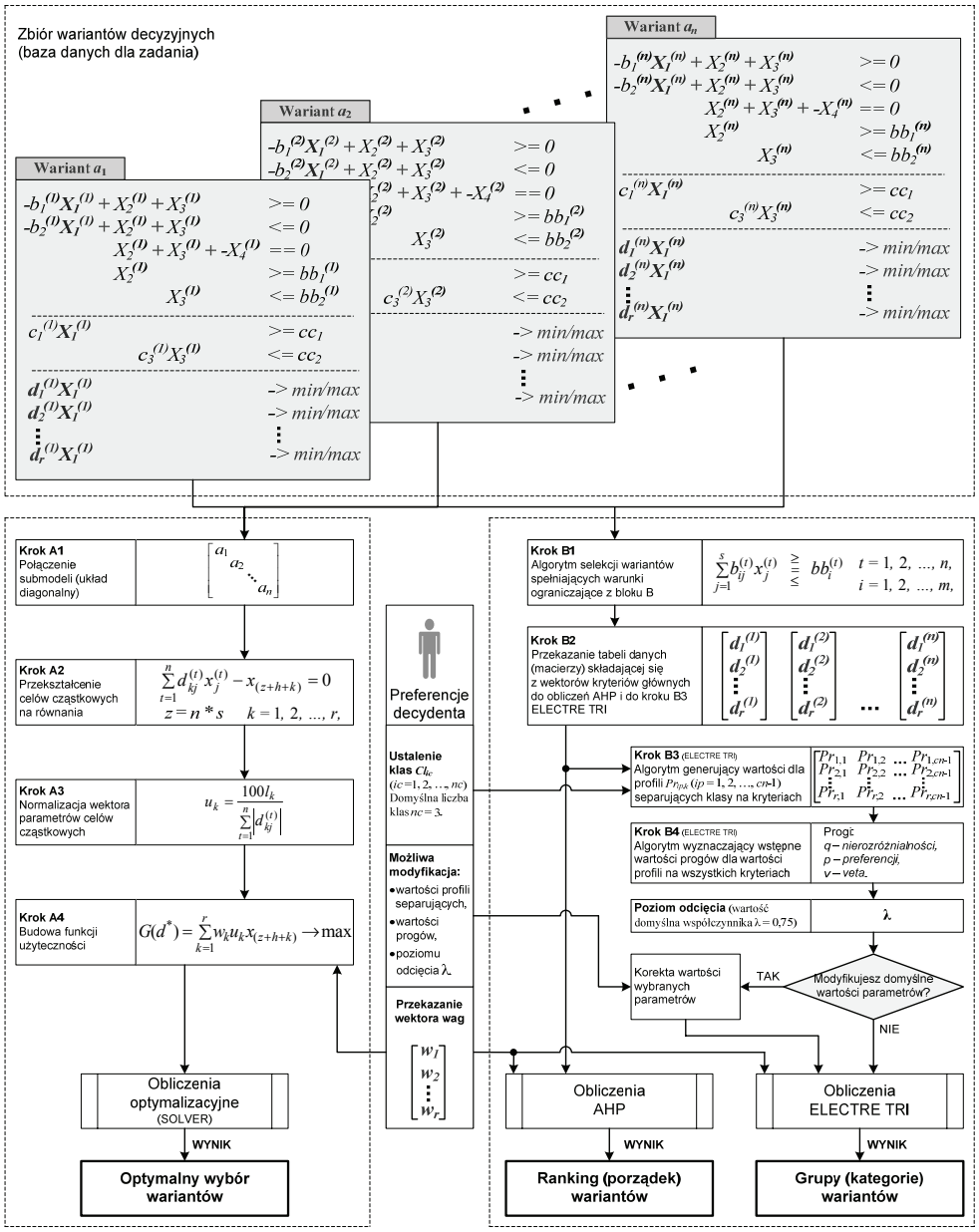
Przy czym w_1, w_2, ..., w_r są współczynnikami wagowymi (preferencjami decydenta) przykładanymi do osiągnięcia różnych celów. Natomiast u_k są parametrami technicznymi [Budziński 2001] sprowadzającymi k-te cele cząstkowe do ich jednakowej rangi w obliczeniach optymalizacyjnych (rys. 2, krok A3). Suma wartości składników funkcji użyteczności (globalnego kryterium celu) osiąga ekstremum na zbiorze rozwiązań dopuszczalnych (wariantów decyzyjnych), co umożliwia znalezienie rozwiązania najlepszego z możliwych.

W konstrukcji systemu DSS założono, że integrowane metody AHP i ELECTRE TRI będą funkcjonować na wspólnym zbiorze danych, wariantów decyzyjnych (rys. 2, baza danych dla zadania). Spełnienie tego wymogu skłoniło do opracowania odpowiedniej procedury (w tym algorytmów; rys. 2, krok B1-B4) preparacji danych wejściowych dla obu metod. Korzystnym efektem tego podejścia (integracji metod na platformie WPL) jest poszerzenie zakresu funkcjonalności obu procedur, szeregowania (AHP) i grupowania (ELECTRE TRI), o algorytm selekcji wariantów spełniających warunki ograniczające z bloku B. Warto w tym miejscu przypomnieć, że w optymalizacyjnej procedurze wyboru (rys 2., krok A1-A4) warunki te, jako składnik submodeli matematycznych, zawsze są lokowane w modelu głównym i rozpatrywane łącznie jako zadanie programowania liniowego. Formalnie metody AHP i ELECTRE TRI tego typu warunków nie uwzględniają. Dlatego postanowiono, żeby na etapie preselekcji wariantów (rys. 2., krok B1) użytkownik systemu DSS mógł sam o tym zadecydować.

Kolejny algorytm systemu (rys. 2, krok B2) dla wyznaczonego zbioru wariantów, z poszczególnych rekordów pobiera wektory kryteriów, tworzy tabelę (macierz), która jest źródłem danych dla obliczeń AHP i kolejnego etapu preparacji danych wejściowych ELECTRE TRI (krok B3). Do uzyskania rankingu wariantów system pobiera jeszcze wektor współczynników wagowych $\{w_1, w_2, ..., w_r\}$, które zostały określone przez decydenta na wstępie, podczas definiowania szablonu modelu matematycznego (rys. 1, blok D). Natomiast do obliczeń ELECTRE TRI preferencje trafiają dopiero po wstępnym wyznaczeniu przez system niezbędnych wartości parametrów: klas Cl_{ic} (ic =1, 2, …, nc), profili Pr_{ip} (ip = 1, 2, …, nc-1), progów (q – nierozróżnialności, p – preferencji, v – veta) i wskaźnika odcięcia $\lambda \in (0{,}5;\ 1\rangle$; który wyznacza poziom wiarygodności dla twierdzenia dotyczącego relacji przewyższania.

Przy pierwszym uruchomieniu procedury grupowania generowany jest automatycznie (bez ingerencji użytkownika) domyślny podział wariantów dla ustalonych z góry wartości parametrów. Założono, że będzie to podział na 3 klasy (nc = 3), przy współczynniku odcięcia $\lambda = 0{,}76$.

Rysunek 2. Idea integracji metod w systemie DSS



Źródło: opracowanie własne

W kroku B3 (rys. 2) algorytm generuje wartości dla profili Pr_{ip} separujących klasy na kryteriach d_k ($k = 1, 2, \ldots, r$). W działaniu tym wykorzystuje się podsystem profili skal punktowych i lingwistycznych stosowany do oceny

parametrów modelu matematycznego (rys. 1, faza 2). System dobiera skalę, której liczba stopni (τ) odpowiada wyznaczonej liczbie klas (nc). Następnie wykorzystuje jej wartości do podziału zakresu $\langle d_{k,min}, d_{k,max}\rangle$ na odcinki, których długości zależą od wybranej charakterystyki. W ustawieniach domyślnych systemu DSS wybrano funkcję liniową, która daje podział na odcinki o tej samej długości. Krańce odcinków to wstępne wartości profili $Pr_{k,ip}$ na k-tym kryterium. Wartości te można modyfikować, np. wygenerować na podstawie charakterystyki nieliniowej i jeśli zachodzi taka potrzeba skorygować każdą z osobna. Rezultatem generowania profili dla każdego kryterium jest macierz.

W kolejnym etapie (rys. 2, krok B4) dla każdej wartości granicznej $Pr_{k,ip}$ system wyznacza wstępne wielkości trzech progów według prostej zależności liniowej $\{q_{k,ip} = \alpha_q\, Pr_{k,ip},\ p_{k,ip} = \alpha_p\, Pr_{k,ip},\ v_{k,ip} = \alpha_v\, Pr_{k,ip}\}$, gdzie $0 < \alpha_q < \alpha_p < \alpha_v < 1$. Przy pierwszym uruchomieniu procedury grupowania progi są generowane dla ustalonych wartości umownych: $\alpha_q = 0{,}03$; $\alpha_p = 0{,}20$; $\alpha_v = 0{,}40$. Następnie, podobnie jak w przypadku profili, system pozwala użytkownikowi skorygować wartości parametrów $\{\alpha_q, \alpha_p, \alpha_v\}$ oraz poszczególnych progów $\{q_{k,ip}, p_{k,ip}, v_{k,ip}\}$.

STUDIUM ROZWIĄZAŃ DECYZYJNYCH W SYSTEMIE DSS

Funkcjonalność przedstawionych rozwiązań modelowych integracji metod w informatycznym systemie DSS można zilustrować na praktycznym przykładzie. Wariantami decyzyjnymi w zadaniu są wnioski, które składają rolnicy starający się o pomoc finansową na działalność związaną z produkcją mleka w ramach działania „Modernizacja gospodarstw rolnych" objętego Programem Rozwoju Obszarów Wiejskich na lata 2007-2013. Problem decyzyjny sprowadza się do oceny wniosków złożonych o dofinansowanie i dokonanie wyboru tych, które z punktu widzenia określonych kryteriów i preferencji dotacjodawcy są najlepsze. W badaniu zachowano założenia formalnej procedury podziału środków finansowych. Na podstawie oryginalnego szablonu dokumentu określono dane teoretyczne dla 10 wniosków.

Osoby ubiegające się o dotację na produkcję mleka muszą spełniać szereg warunków, np.: należy być właścicielem gospodarstwa rolnego zarejestrowanym w ewidencji producentów gospodarstw rolnych, nie przekraczać wieku emerytalnego, w przypadku zajmowania się produkcją mleka należy taką działalność prowadzić minimum 1 rok, z kolei powierzchnia gospodarstwa nie może być mniejsza niż 1 ha i nie większa niż 300 ha. Poza tym beneficjent musi mieć już przyznaną minimalną kwotę mleczną w wysokości 20 tys. kg. Wielkość ekonomiczna gospodarstwa powinna wynosić min 4 ESU (europejska jednostka wielkości – określa dochodowość (wielkość, siłę) gospodarstw rolnych; 1 ESU = 1200 EUR). W 2011 roku graniczna kwota dofinansowania produkcji mleka przypadająca na jednego beneficjenta (wniosek) wynosiła 300 000 zł.

Tabela 1. Warunki i kryteria wyboru wniosków

Parametr	Opis
BLOK B. WARUNKI BINARNE DLA WNIOSKU (0 - nie, 1 - tak)	
bb1	Powierzchnia gospodarstwa (grunty rolne) mieści się w przedziale od 1 do 300 ha?
bb2	Beneficjent posiada wykształcenie rolnicze lub staż rolniczy?
bb3	Kwota mleczna wynosi co najmniej 20 tyś kg?
bb4	Wielkość ekonomiczna gospodarstwa wynosi co najmniej 4 ESU?
bb5	Wiek beneficjenta (kobiety od 18 do 60 lat; mężczyźni od 18 do 65 lat)?
bb6	Beneficjent posiada numeru ident. w ewidencji producentów lub gospodarstw rolnych?
bb7	Beneficjent jest obywatelem kraju Uni Europejskiej?
bb8	Produkcji mleka trwa więcej niż 1 rok?
bb9	Kwota zapotrzebowania nie przekracza 300 tyś zł?
bb10	Wniosek złożony w terminie?
bb11	Wniosek wypełniony zgodnie z wymaganiami formalnymi (druk, rubryki, podpisy)?
bb12	Załączono komplet dokumentów?
bb13	Biznes plan poprawny?
bb14	Operacja realizowana na terenie Polski?
bb15	Wnioskodawca mieszka na terenie Polski?
BLOK C. WARUNKI OGRANICZAJĄCE DLA ZADANIA	
c1	Kwota dotacji celowej (zapotrzebowanie w tys. zł)
cc1	Pula środków finansowych przeznaczonych na dotacje (tys. zł)
cc2	Ograniczenie liczby wniosków (szt.)
BLOK D. KRYTERIA WYBORU WNIOSKÓW	
d01	Innowacyjność operacji (0 pkt - mała, 1 pkt - przeciętna, 2 pkt - znacząca)
d02	Powierzchania gospodarstwa nie przekracza średniej wartości powierzchni gospodarstwa w danym regionie (0 pkt - przekracza, 1 pkt - nie przekracza)
d03	Wpływ operacji na środowisko naturalne (0 pkt - duży wpływ, 1 pkt - umiarkowany wpływ, 2 pkt - bark wpływu)
d04	Wysokość wnioskowanej kwoty (250 tyś zł do 300 tyś zł = 0 pkt, 200 tyś zł do 249 tyś zł = 3 pkt, mniej niż 200 tyś zł = 5 pkt)
d05	Operacja popawia komfort życia zwierząt (0 pkt = nie, 1 pkt = tak)
d06	Operacja podniesie jakość mleka (0 pkt - nie, 1 pkt - umiarkowanie, 2 pkt - znacznie)
d07	Beneficjent ma mniej niż 40 lat (0 pkt - nie, 1 pkt - tak)
d08	Staż pracy w rolnictwie przy produkcji mleka (3 pkt - 1 rok, 2 pkt - 2 do 4 lat, 1 pkt - 5 do 7 lat, 0 pkt - powyżej 8 lat)
d09	Ocena punktowa wiekości ekonomicznej gospodarstwa ESU (0-10 pkt. - rys. 5a)
d10	Ocena punktowa wielkości kwoty mlecznej (0-10 pkt. - rys. 5b)

Źródło: opracowanie własne na podstawie Programu Rozwoju Obszarów Wiejskich na lata 2007-2013, działanie „Modernizacja gospodarstw rolnych" (http://www.minrol.gov.pl).

Wnioski oceniane są w dwóch etapach. W pierwszej kolejności poddawane są kontroli formalnej, która polega sprawdzeniu m. in. czy: wniosek został wypełniony prawidłowo i na właściwym formularzu, podane wartości parametrów mieszczą się w wyznaczonych przedziałach, został złożony na czas i jest załączony komplet dokumentów. Pełną listę warunków binarnych zawiera tabela 1 (blok B), niespełnienie któregokolwiek skutkuje odrzuceniem wniosku. W takim przypadku w formalnej procedurze przewidziano możliwość poprawienia dokumentu w określonym czasie. Przygotowany w systemie DSS szablon modelu

matematycznego dla rozpatrywanego zadania wyboru wniosków jest nieskomplikowany. Założono, że każdy wniosek będzie reprezentowany przez jedną zmienną typu binarnego (oznaczenie '*X01*', rys. 3). Warunki ograniczające bloku B to zbiór prostych nierówności $\{X01 \leq bb1,\ X01 \leq bb2,\ ...,\ X01 \leq bb15\}$, w których parametry z prawej strony znaku relacji przyjmują wartości binarne. Część z nich, jak: *bb1, bb3, bb4, bb5, bb8* i *bb9*, wyliczana jest automatycznie przez systemy DSS na podstawie podanej w formularzu wartości bezwzględnej (*ocn*) oraz odpowiednio przygotowanej reguły (np. $if(ocn_1 \geq 1\ .and.\ ocn_1 \leq 300,\ bb1 := 1,\ bb1 := 0)$. Pozostał wprowadzane są do systemu przy użyciu binarnej skali ocen (*'tak'* $\rightarrow$ $bb1 := 1$; *'nie'* $\rightarrow$ $bb1 := 0$). Jeżeli wszystkie warunki ograniczające z bloku B są spełnione to wniosek przechodzi do kolejnego etapu, w którym jest poddawany ocenie merytorycznej.

Rysunek 3. Podgląd szablonu modelu matematycznego zdefiniowanego dla zadania wyboru wniosków

SZABLON [DSS008_000000]

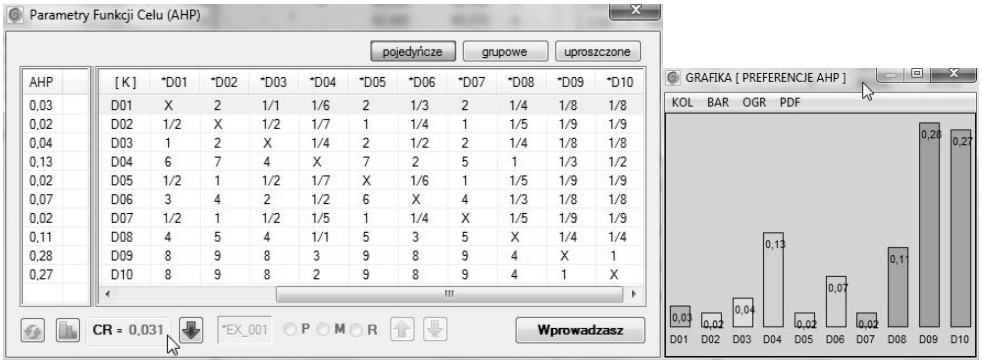
IND	BILANSE	X01_Wybór	F01...	F02...	F03...	F04...	F05...	F06...	F07...	F08...	F09...	F10...	REL	OGRAN
B01	¤Powierzchnia gospodarstwa (gr...	1,000											<=	bb1
B02	¤Wykształcenie rolnicze (lub sta	1,000											<=	bb2
B03	¤Minimalna kwota mleczna (20 tys	1,000											<=	bb3
B04	¤Minimalna wielkość ekonomicz...	1,000											<=	bb4
B05	¤Wiek beneficjenta (M: 18-65; K:	1,000											<=	bb5
B06	¤Posiadanie nr id. w ew. produce	1,000											<=	bb6
B07	¤Beneficjent jest obywatelem kra	1,000											<=	bb7
B08	¤Prowadzi gosp. rolne (produkcję	1,000											<=	bb8
B09	¤Kwota zapotrzebowania nie prz...	1,000											<=	bb9
B10	¤Wniosek złożony w terminie	1,000											<=	bb10
B11	¤Wniosek wypełniony zgodnie z...	1,000											<=	bb11
B12	¤Załączono komplet dokumentó...	1,000											<=	bb12
B13	¤Biznes plan poprawny	1,000											<=	bb13
B14	¤Operacja realizowana na terenie	1,000											<=	bb14
B15	¤Wnioskodawca mieszka na ter...	1,000											<=	bb15
C01	¤Kwota zapotrzebowania	c1											<=	cc1
C02	Liczba wniosków	1,000											<=	cc2
D01	Innowacyjność operacji	d1	-n1 ...										==	0
D02	Powierzchania gosp. nie przekrac	d2		-n2 ...									==	0
D03	Wpływy operacji na środowisko ...	d3			-n3 ...								==	0
D04	Wysokość wnioskowanej kwoty ...	d4				-n4 ...							==	0
D05	Operacja popawia komfort życia z	d5					-n5 ...						==	0
D06	Operacja podniesie jakość mleka	d6						-n6 ...					==	0
D07	Beneficjent ma mniej niż 40 lat	d7							-n7 ...				==	0
D08	Staż pracy w rolnictwie przy pro	d8								-n8 ...			==	0
D09	Ocena Wiekości ekonomicznej ...	d9									-n9 ...		==	0
D10	Ocena pkt. na podst. wysokości k	d10										-n1...	==	0
CEL	FUNKCJA WPL		a01...	a02...	a03...	a04...	a05 ...	a06...	a07 ...	a08...	a09...	a10...	=>	MAX

Źródło: opracowanie własne z wykorzystaniem systemu DSS 2.0.

W oryginalnej procedurze ocena merytoryczna wniosku wyrażona jest sumą punktów uzyskanych na poszczególnych kryteriach. Ilość zdobytych punktów decyduje o pozycji wniosku w rankingu. Do oceny wniosków przyjęto 10 kryteriów $\{d01, d02, ..., d10\}$. W tabeli 1 (blok D) zamieszczono ich opis oraz sugerowane w dokumentacji różne skale punktowe. W systemie DSS zrezygnowano z niejednorodnych rozpiętości, dla wszystkich kryteriów przyjęto

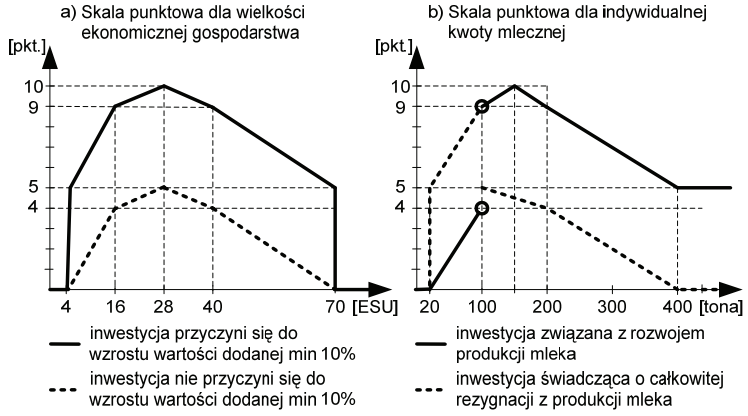
jednolitą skalę od 0 do 10 punktów a w zamian określono preferencje (wektor wag) odzwierciedlające tę różnorodność. Obserwując rys. 4 (kolumny AHP i [K], lub wykres) wyraźnie widać, że najsilniej preferowanymi kryteriami są *d09* oraz *d10,* które miały pierwotnie przydzieloną skalę o największym zakresie od 0 do 10pkt. Natomiast najniższe współczynniki wagowe nadano: *d02, d05 i d07,* ponieważ miały skalę binarną {0 pkt.; 1 pkt.}. Podczas konstrukcji wektora skali system obliczał na bieżąco wartość wskaźnika *CR*. Uzyskana na koniec wartość $CR = 0{,}031$ świadczy o zachowaniu spójności ocen porównywanych parami kryteriów.

Rysunek 4. Okno wyznaczania wektora skali dla kryteriów



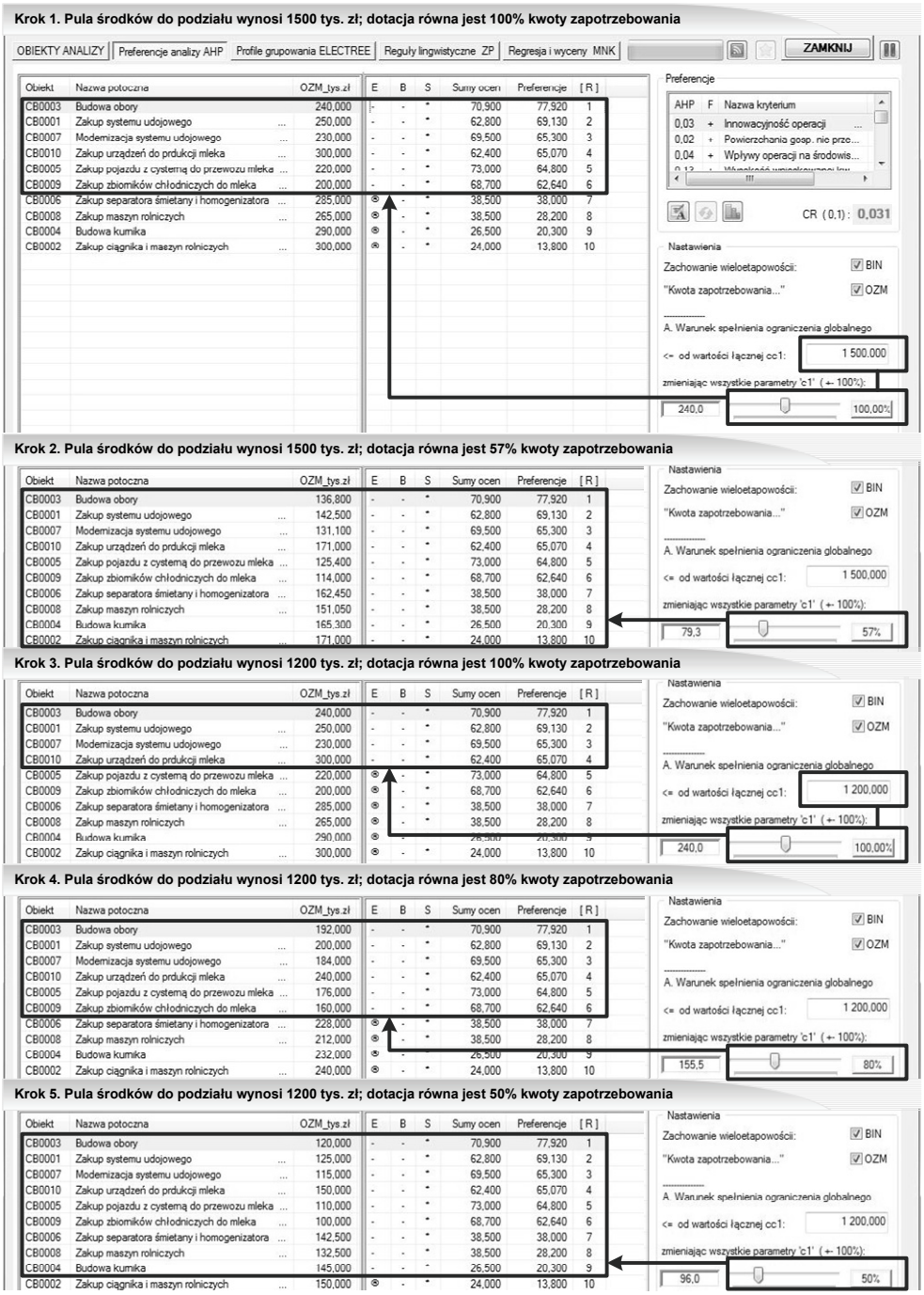
Źródło: opracowanie własne z wykorzystaniem systemu DSS 2.0.

Rysunek 5. Formalne oceny punktowe dla kryterium ESU i indywidualnej kwoty mlecznej



Źródło: opracowanie własne na podstawie Programu Rozwoju Obszarów Wiejskich na lata 2007-2013, działanie „Modernizacja gospodarstw rolnych" (http://www.minrol.gov.pl).

Rysunek 6. Ranking wniosków metodą AHP i symulacja podziału środków finansowych



Źródło: opracowanie własne z wykorzystaniem systemu DSS 2.0.

Podobnie jak to miało miejsce w bloku B, na podstawie podanej w formularzu wartości bezwzględnej (*ocn*) oraz odpowiednio przygotowanej w programatorze systemu DSS reguły, w bloku D wyliczane są wartości punktowe dla kryteriów: *d02, d04, d7, d8, d9, d10* (patrz: tabela 1). Na szczególną uwagę zasługują dwa kryteria (*d9 i d10*), które mają największy wpływ na pozycję wniosków w rankingu. Prawodawcy dotacji unijnych opracowali dla nich złożone reguły przyznawania punktów (rys. 5 a i b). Pozostałe kryteria podlegają ocenie eksperckiej, dlatego przyporządkowano im w systemie odpowiednie skale lingwistyczne (np. skala dla *d03*:*'duży wpływ'* → 0 pkt., *'umiarkowany wpływ'* → 5 pkt.; *'brak wpływu'* → 10 pkt.).

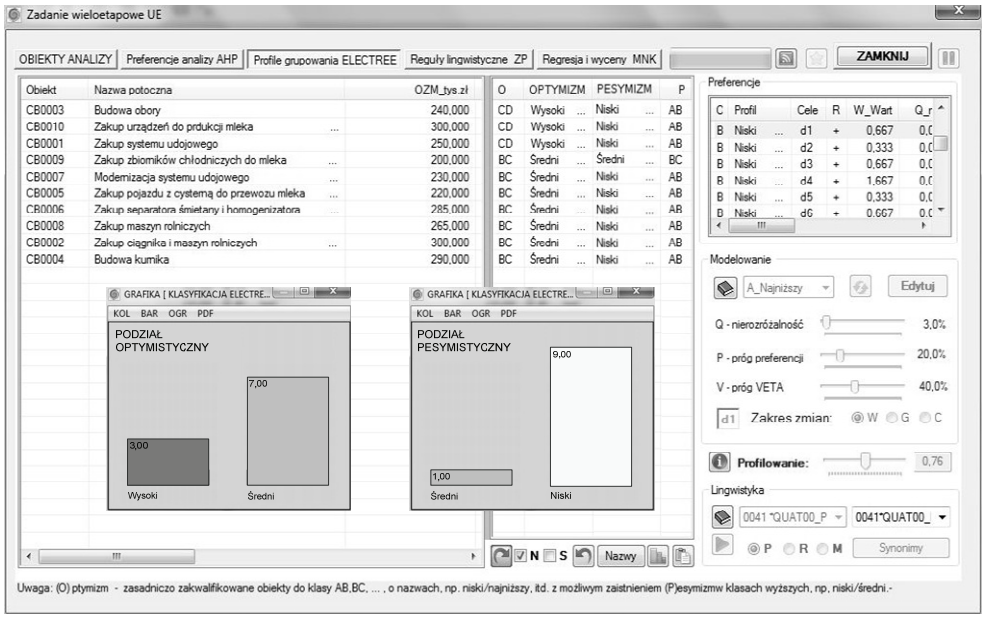
Po zdefiniowaniu struktury informacyjnej zadania decyzyjnego – szablonu modelu matematycznego, zakresu i reguł transformacji wartości jego parametrów oraz skal lingwistycznych do ich eksperckiej oceny – wprowadzono do systemu wartość kwoty do podziału (*cc1* = 1500 tys. zł) oraz dane dla dziesięciu wniosków. Następnie na podstawie rankingu AHP przeprowadzono badanie symulacyjne rozdziału środków finansowych. W ostatniej fazie, przy pomocy ELECTRE TRI, przydzielono wnioski do trzech zdefiniowanych klas jakości: niskiej, średniej, wysokiej. Działanie to pozwoliło odpowiedzieć na szereg interesujących pytań. Na przykład: jaka jest struktura podziału wniosków, czy jest zgodna z porządkiem uzyskanym w rankingu, czy wszystkie wnioski są porównywalne (założenie AHP), jaką kategorię stanowią faworyci rankingu, czy są to wysoko, przeciętnie, a może nisko preferowane wnioski?

Po zatwierdzeniu w menu systemu DSS pozycji o nazwie *„Analiza, rankingi i klasyfikacje"* pojawia się okno z pięcioma zakładkami: 1) Obiekty analizy, 2) Preferencje analizy AHP, 3) Profile grupowania ELECTRE TRI, 4) Reguły lingwistyczne ZP, 5) Regresja i wyceny MNK. Dwie opcje, przedostatnia wykorzystująca metody teorii zbiorów przybliżonych do poszukiwania reguł decyzyjnych oraz ostatnia oferująca analizę regresji i wyceny wariantów decyzyjnych nie zostały omówione, ponieważ wykraczają poza zakres opracowania (określony w tytule). Po uruchomieniu okna system ustawia się na pierwszej zakładce. Użytkownik może w niej dokonać selekcji wariantów według dowolnych warunków filtrowania i przekazać wybrany podzbiór do kolejnych zakładek (w omawianym przykładzie do analizy przekazano wszystkie wnioski).

Funkcjonalność opcji *„Preferencje analizy AHP"* zilustrowano w pięciu krokach (rys. 6). W pierwszym kroku, po uaktywnieniu zakładki, system wyświetlił ranking wniosków, od najbardziej do najmniej preferowanych, według kolumny *Preferencje*. W kolumnie *Sumy ocen* widoczne są punkty obliczone bez uwzględnienia wektora wag, natomiast w *OZM_tys zł* ukazana jest możliwa do uzyskania kwota dotacji. Dotację otrzymają beneficjenci, których wnioski zajmują najwyższą lokatę i jednocześnie suma kwot ich zapotrzebowania nie przekracza puli środków do podziału. Dla ustalonej na wstępie puli do rozdysponowania *cc1* = 1,5 mln zł i przy zachowaniu wielkości dotacji na poziomie 100% kwoty

zapotrzebowania zakwalifikowało się sześć najlepszych wniosków (rys. 6, krok 1). System informuje o tym fakcie wyróżniając odpowiednio wiersze w kolumnie *'E'* ('-/⊗' → weszły/nie weszły do rozwiązania). W drugim kroku symulacji poszukiwano odsetka kwoty zapotrzebowania, przy którym dotację mogliby otrzymać wszyscy beneficjenci. W tym celu zmniejszano stopniowo wartość odsetka i kontrolowano pojawiające się w kolumnie *'E'* oznaczenia '⊗'. Okazało się, że warunek ten może zostać spełniony przy zmniejszeniu dotacji do poziomu 57% kwot zapotrzebowania (rys 2, krok 2). W kolejnych krokach symulacji zmniejszono o trzy miliony pulę dofinansowania (*cc1* = 1,2 mln zł) i ponownie rozważano stosunek liczby wniosków otrzymujących dotację do wielkości ich kwot, stanowiących odsetek zgłoszonego zapotrzebowania. W kroku 3 przy 100% dotacjach zakwalifikowały się cztery wnioski, w kroku 4 przy 80% było to sześć wniosków, a przy 50% dziewięć.

Rysunek 7. Analiza podziału wniosków na kategorie metodą ELECTRE TRI



Źródło: opracowanie własne przy zastosowaniu systemu DSS 2.0.

W ostatnim etapie badania, wybierając zakładkę *„Profile grupowania ELECTRE TRI",* system pogrupował wnioski z punktu widzenia przyjętych preferencji do trzech klas jakości (*nc* = 3): AB – niska, BC – średnia, CD – wysoka. Klasy te są definiowane niezależnie od wariantów oraz są porównywalne z punktu widzenia preferencji decydenta. W rozważanym przykładzie (w obliczeniach) system zastosował umownie wygenerowaną macierz dla dwóch profili separujących (kolumny) i dziesięciu kryteriów (wiersze), domyślne wartości parametrów: $\alpha_q = 0{,}03$; $\alpha_p = 0{,}20$ i $\alpha_v = 0{,}40$ używane do wyznaczenia wartości

progów na poszczególnych profilach oraz współczynnik odcięcia (poziom wiarygodności) $\lambda = 0{,}76$. Zgodnie z założeniami metody ELECTRE TRI przydziału wariantów do wyznaczonych klas dokonano za pomocą procedury optymistycznej i pesymistycznej (rys 7). Uzyskany podział nie jest jednoznaczny, w obydwu podejściach struktury podziału wniosków znacząco się różnią. W ujęciu optymistycznym są 3 wnioski wysoko preferowane i 7 średnio, natomiast w pesymistycznym tylko 1 przeciętnie preferowany, a pozostałe nisko. Niezgodność ta sygnalizuje występowanie pewnych niejednoznaczności. W praktyce, gdy wynikowe przydziały się różnią, ostatecznego przyporządkowania dokonuje decydent. Trzy wnioski o identyfikatorze: CB0003, CB0010 i CB0007, zostały zaklasyfikowane do skrajnych grup. Oznacza to, że są one nieporównywalne z wyznaczonymi profilami separującymi oraz, wnioskując dalej, z wariantami, dla których taka porównywalność zachodzi. Metoda AHP zakłada pełną porównywalność obiektów (nie powinno się w niej zestawiać np. Rolls-Royce z Fiatem 126p). W rankingu wnioski te znalazły się w pierwszej czwórce, więc tym bardziej jest to asumpt dla decydenta do zastanowienia i wnikliwej analizy danych. Tylko jeden wniosek z numerem CB0009 uzyskał jednoznaczny przydział do średniej klasy w obu podejściach. Pozostałe sześć wniosków na pewno nie należy do kategorii wysoko preferowanych ich przydział waha się w zależności od podejścia na granicy klasy średniej (optymizm) i niskiej (pesymizm). Generalnie struktura przydziału wniosków do klas w kolejności od wysokiej do niskiej w dobrym stopniu odpowiada porządkowi uzyskanemu w rankingu. Wynika to z faktu, że większość wniosków (7 z 10) była porównywalna z profilem separującym kategorię słabą od średniej.

ZAKOŃCZENIE

Opracowany system DSS spełnia wymagania rozwiązania zorientowanego na modele, komunikację i grupowe wspomaganie decyzji. Modele są dostarczane użytkownikowi w postaci gotowych propozycji (szablonów) wraz z narzędziami umożliwiającymi ich rozwijanie w kierunku własnych potrzeb. Elastyczne struktury informacyjne systemu DSS (repozytoria baz danych i szablonów modeli matematycznych) sprawiają, że jest narzędziem przystosowanym do szybko i często zmieniających się potrzeby użytkowników. Stanowi on spójną platformę informatyczną o otwartej architekturze. Platformę, która pozwala na różne konfiguracje składników tego układu, mianowicie: *dane* – różne typy numeryczne lub formy lingwistyczne, *modele* – proste (jednowektorowe), złożone (hierarchiczne), uwzględniające zbiór warunków ograniczających oraz wartości parametrów będące wynikiem zastosowania reguł, *metody* – oparte na optymalizacji WPL, wieloatrybutowej teorii użyteczności (AHP) lub wykorzystujące relację przewyższania (ELECTRE TRI), *wyniki* – dowolne graficzne formy prezentacji uzyskanych rezultatów analizy decyzyjnej.

BIBLIOGRAFIA

Becker J. (2008) Architektura informatycznego systemu generowania wielokryterialnych rozwiązań decyzyjnych: (cz. 1) Koncepcja budowy modelu WPL oparta na niestandardowych zadaniach decyzyjnych, Seria IBS PAN: Badania Systemowe, Tom 64, Wyd. Instytut Badań Systemowych PAN & Polskie Towarzystwo Badań Operacyjnych i Systemowych, Warszawa.

Becker J. (2010) Integracja źródeł wiedzy w modelowaniu zadań WPL na potrzeby systemów klasy DSS, Seria: Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą nr 27, Wydawnictwo: BEL Studio Sp. z o. o., Bydgoszcz, s. 49-60.

Budziński R. (2001) Metodologiczne aspekty systemowego przetwarzania danych ekonomiczno-finansowych w przedsiębiorstwie, Monografia, Rozprawy i Studia T. (CDXLVI)372. Wydawnictwo Naukowe US w Szczecinie, Szczecin.

La Gauffre P., Haidar H., Poinard D. (2007), A multicriteria decision support methodology for annual rehabilitation programs for water networks. "Computer-Aided Civil and Infrastructure Engineering" 22, s. 478-488.

Roy B., Słowiński R. (2008), Handing effects of reinforced preference and counter-veto in credibility of outranking. "European Journal of Operational Research", 188, s.186-187.

Saaty T.L. (1980), The analytic hierarchy process: Planning, priority setting, resource allocation, McGraw-Hill International Book Co., New York.

Saaty T.L. (2004), The Analytic Hierarchy and Analytic Network Processes for the Measurement of Intangibles Criteria and for Decision-Making, in: Figueira J., Greco S., Ehrgott M. (edit.), Multiple Criteria Decision Analysis: State of the Art Surveys, 67 page chapter, Kluwer Academic Publishers, Dordrecht.

Vernadat, F.B. (1996) Enterprise Modeling and Integration: Principles and Applications. London: Chapman & Hall.

Program Rozwoju Obszarów Wiejskich na lata 2007-2013 (http://www.minrol.gov.pl)

AHP AND ELECTRE TRI METHOD INTEGRATION – STUDIES ABOUT DECISION SOLUTIONS IN DSS SYSTEM (PART II)

Abstract: The article is a continuation of research that includes the review and functional analysis of quantitative methods within MCDA for designing needs of an information system supporting decisions (the article was published in a previous edition of the Magazine as part I). The aim of the part II is to show practical aspects of conducted research connected with integration of AHP and ELECTRE TRI methods with their application. Research was conducted in a form of studies over decision solutions that were conducted with the use of prepared system of DSS class (prototype version DSS 2.0). Methods work in the common set of objects that may be put in order (AHP) and devote to defined groups (ELECTRE TRI).

Keywords: quantitative methods in decision support, multiple criteria decision analysis, decision support system

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 42 – 51

MODEL SEKWENCYJNEGO ZAWIERANIA TRANSAKCJI – ZASTOSOWANIE DO ANALIZY PROCESU TRANSAKCYJNEGO NA KASOWYM RYNKU ZŁOTEGO

Katarzyna Bień-Barkowska[1]
Instytut Ekonometrii, Szkoła Główna Handlowa
e-mail: katarzyna.bien@sgh.waw.pl
Departament Systemu Finansowego, Narodowy Bank Polski

Streszczenie: W artykule dokonano próby aproksymacji stopnia zróżnicowania intencji i oczekiwań dealerów międzybankowego kasowego rynku złotego na podstawie strumienia składanych zleceń zakupu lub sprzedaży euro. Artykuł wpisuje się w obszar badań mikrostruktury rynku walutowego. Zaprezentowano w nim uogólnioną w pracy [Wünsche 2007] specyfikację ekonometrycznego modelu sekwencyjnego zawierania transakcji pierwotnie zaproponowanego w pracy [Easley i in. 1996]. Model ten umożliwia pomiar „zawartości informacyjnej" poszczególnych transakcji poprzez oszacowanie czterech miar: (1) prawdopodobieństwa napływu nowej informacji na rynek, (2) prawdopodobieństwa, że nowa informacja okaże się zła dla waluty bazowej, (3) intensywności transakcji wywołanych napływem nowej informacji (*informed trading*), (4) intensywności transakcji, które nie są umotywowane napływem nowej informacji na rynek, a wynikają na przykład z zarządzania płynnością (*uninformed trading*). W badaniu empirycznym dokonano oszacowania zróżnicowania poziomu zmienności kursu EUR/PLN w okresach charakteryzujących się napływem dobrej informacji, złej informacji lub zakwalifikowanych jako okresy bez napływu nowych sygnałów informacyjnych.

Słowa kluczowe: mikrostruktura rynku, modele sekwencyjnego zawierania transakcji, kurs walutowy

[1] Autorka składa podziękowania firmie Thomson Reuters za udostępnienie danych z systemu Reuters Dealing 3000 Spot Matching oraz Piotrowi Odrzywołkowi, Pawłowi Sobolewskiemu i dr Dobiesławowi Tymoczko za uwagi merytoryczne. Opinie prezentowane w artykule są prywatnymi opiniami Autorki, a nie stanowiskiem Narodowego Banku Polskiego.

WPROWADZENIE

Badania mikrostruktury rynku definiuje się jako studia nad procesem wymiany dóbr przy uwzględnieniu sformalizowanych reguł obrotu handlowego (por. [O'Hara 1995]). W wielu teoretycznych modelach formułowanych w tym obszarze przyjmuje się założenie, że podmioty zawierające transakcje na rynku finansowym zdobywają informację i „uczą się" na podstawie zachowań innych uczestników rynku. W mikrostrukturalnych modelach informacji (ang. *information models*) scenariusz składanych zleceń zakupu i sprzedaży stanowi czytelną przesłankę na temat oczekiwań dotyczących przyszłej wartości instrumentu finansowego (por. Glosten i Milegrom 1985], [Easley i O'Hara 1987]). Zgodnie z wynikami badań prezentowanych w literaturze poświęconej mikrostrukturze rynków walutowych, przepływ zleceń (zakup netto waluty bazowej) wywiera istotny wpływ na kształtowanie się krótkookresowych (wewnątrzdziennych, dziennych) zmian kursu walutowego (por. [Evans i Lyons 2002a, 2002b], [Berger I in. 2007], [Rime i in. 2007], [Scalia 2008] i.in.). W odniesieniu do kasowego rynku złotego istotną statystycznie zależność pomiędzy zakupem netto waluty bazowej a stopami zwrotu dla kursu USD/PLN (w I połowie 2004 r.) oraz EUR/PLN
(w II połowie 2004 oraz w 2007 r.) wykazano w pracy [Bień 2010].

Celem artykułu jest prezentacja modelu sekwencyjnego zawierania transakcji (ang. *sequential trade models*) EKOP, należącego do obszaru mikrostrukturalnych modeli informacji. Model EKOP jest popularnym narzędziem ekonometrycznym umożliwiającym pomiar „zawartości informacyjnej" procesu transakcyjnego, czyli intencji i oczekiwań inwestorów na rynku kierowanym zleceniami. Na takim rynku strony transakcji (dealerzy walutowi w przypadku rynku walutowego) mogą składać dwa podstawowe typy zleceń: zlecenia rynkowe i zlecenia z limitem ceny. Zlecenia rynkowe zakupu (sprzedaży) waluty bazowej są realizowane po najbardziej konkurencyjnych cenach ask (bid) dostępnych na rynku i skutkują natychmiastowym zawarciem transakcji. Z uwagi na pewność realizacji takich zleceń, są one w literaturze tradycyjnie postrzegane jako wynikające z napływu nowej informacji na rynek, gdyż wartość takiej informacji z reguły podlega bardzo szybkiej deprecjacji w czasie (por. [Glosten 1994], [Sepii 1997]). Zlecenia z limitem ceny zakupu (sprzedaży) mogą być natomiast wprowadzane do arkusza zleceń po cenach nie odpowiadających najbardziej konkurencyjnym cenom ask (bid) i mogą oczekiwać na realizację przez zadany przez dealera okres.

Rozróżnienie pomiędzy zleceniami rynkowymi i zleceniami z limitem ceny pozwala na wyróżnienie w badaniach mikrostruktury rynku tzw. transakcji zakupu i transakcji sprzedaży. Transakcja zakupu waluty bazowej następuje w wyniku realizacji rynkowego zlecenia zakupu waluty bazowej względem oczekującego na realizację najbardziej konkurencyjnego zlecenia sprzedaży waluty bazowej z limitem ceny. Dealer kupujący walutę bazową w takiej transakcji jest określany

jako agresor – podmiot, którego działanie doprowadziło bezpośrednio do zawarcia transakcji. Analogicznie, do transakcji sprzedaży waluty bazowej dochodzi w wyniku realizacji rynkowego zlecenia sprzedaży waluty bazowej względem oczekującego na realizację, najbardziej konkurencyjnego, zlecenia zakupu waluty bazowej z limitem ceny. Agresorem w takiej transakcji jest zatem dealer sprzedający walutę bazową.

Zastosowanie modelu EKOP do analizy danych dotyczących poszczególnych typów zleceń zakupu i sprzedaży umożliwia wyodrębnienie dwóch rodzajów transakcji zakupu lub sprzedaży: (1) transakcji wynikających z napływu nowej informacji na rynek (ang. *informed trades*) oraz (2) transakcji nie mających związku z napływem nowych sygnałów informacyjnych, a wynikających np. z zarządzania płynnością (ang. *uninformed trades, liquidity trades, noise trades*). Model umożliwia również oszacowanie tzw. prawdopodobieństwa zawierania transakcji na podstawie napływu nowej informacji (ang. *probability of informed trading*), czyli oszacowanie liczby transakcji zainicjowanych poprzez napływ nowych sygnałów informacyjnych wśród wszystkich zawartych na rynku transakcji. W literaturze poświęconej zastosowaniom modeli informacji, model EKOP był dotychczas wykorzystywany w badaniach migracji nowej informacji w okresach, w których dokonywano podziału akcji (ang. *stock splits*) (por. [Easley, O'Hara i Saar 2001]), publikacji nowej informacji dotyczącej standingu finansowego spółek (por. [Benos i Jochec 2007] i [Reza i Wilson 2007]), wyjaśnienia wahań płynności na rynku (por. [Easley et al. 1996]; [Brockman i Chung 2000]; [Easley et al. 2008]). Model ten stanowił również podstawę oceny wewnątrzdziennej sezonowości zawierania transakcji na podstawie napływu nowych sygnałów informacyjnych a zatem fluktuacji „intencji i oczekiwań" uczestników rynku (por. [Gençay, Gradojevic i Selcuk 2007] i [Gençay i Gradojevic 2008]).

W artykule zaprezentowano przykład zastosowania uogólnionej przez Wünsche (2007) wersji modelu EKOP do danych z systemu Reuters Dealing 3000 Spot Matching obejmujących międzybankowe transakcje wymiany euro (waluta bazowa) za złotego w okresie styczeń-lipiec 2008 r. Wyniki estymacji umożliwiły także pomiar zróżnicowania wewnątrzdziennej zmienności kursu EUR/PLN w okresach charakteryzujących się napływem dobrej lub złej informacji dla złotego.

TEORETYCZNA KONSTRUKCJA MODELU EKOP I JEJ ROZWINIĘCIA

W modelu EKOP zakłada się, że na rynku występują dwie grupy inwestorów: (1) mający dostęp do informacji (*informed*) oraz (2) niemający dostępu do informacji (*uninformed, liquidity*). W danej jednostce czasu prawdopodobieństwo napływu na rynek sygnału informacyjnego jest stałe i wynosi

α. Prawdopodobieństwo tego, że dany sygnał informacyjny powiązany jest z napływem „złej" informacji dla waluty bazowej wynosi δ, natomiast tego, że informacja jest „dobra" wynosi odpowiednio $1-\delta$. Transakcje zakupu i sprzedaży waluty bazowej odbywają się zgodnie z niezależnymi procesami Poissona o wartościach oczekiwanych (intensywnościach) $\lambda_{Z,t}$ oraz $\lambda_{S,t}$, odpowiednio. W każdym z analizowanych okresów, inwestorzy dysponujący informacją mogą wkraczać na rynek dokonując transakcji ze stałą intensywnością μ, a inwestorzy, którzy nie odebrali żadnych sygnałów informacyjnych z intensywnością ε. Na początku każdego z wyróżnionych interwałów czasowych na rynku może pojawić się sygnał informacyjny z prawdopodobieństwem α. Zostaje on wówczas „odebrany" przez część dealerów walutowych, którzy – postępując racjonalnie – postanawiają wykorzystać swoją przewagę informacyjną na rynku. Jeżeli informacja jest „dobra" dla waluty bazowej (świadczy o chwilowym niedowartościowaniu waluty bazowej), inwestorzy ci włączają się do procesu transakcyjnego dokonując transakcji zakupu waluty bazowej z intensywnością (średnią liczbą transakcji w ciągu wyróżnionego okresu) μ. Oprócz podmiotów mających dostęp do sygnałów informacyjnych na rynku obecni są również dealerzy, którzy zawsze z taką samą intensywnością ε zawierają transakcje zakupu i sprzedaży – dokonują ich w celach „nieinformacyjnych", czyli niezależnych od napływu sygnałów informacyjnych. W związku z tym, w następstwie zaobserwowania „dobrej" informacji (dla waluty bazowej, czyli EUR), oczekiwana liczba transakcji zakupu w piętnastominutowym interwale czasowym wynosi: $\lambda_Z = \mu + \varepsilon$, natomiast średnia liczba transakcji sprzedaży: $\lambda_S = \varepsilon$. Analogicznie, jeżeli informacja okazała się „zła" (świadczyła o chwilowym przewartościowaniu kursu), ci dealerzy, którzy „odebrali" sygnał informacyjny dokonują agresywnych transakcji sprzedaży waluty bazowej. W takim przypadku, oczekiwana liczba transakcji zakupu w ciągu piętnastu minut wynosi: $\lambda_Z = \varepsilon$, natomiast oczekiwana liczba transakcji sprzedaży: $\lambda_S = \mu + \varepsilon$. W okresach, w których na rynek nie napłynęły żadne sygnały informacyjne, aktywnymi uczestnikami procesu transakcyjnego są tylko banki, które nie dysponują pełną informacją, a zatem zarówno oczekiwana liczba transakcji zakupu jak i sprzedaży waluty bazowej wynosi: $\lambda_Z = \lambda_S = \varepsilon$.

Model EKOP ma cztery parametry α, δ, ε, μ, a jego klasyczna konstrukcja ekonometryczna wykorzystuje mieszankę trzech dwuwymiarowych rozkładów Poissona o różnych wartościach oczekiwanych. Prawdopodobieństwo jednoczesnego zaobserwowania z_t transakcji zakupu oraz s_t transakcji sprzedaży między momentami $t-1$ oraz t ma postać:

$$P(Z_t = z_t, S_t = s_t) = \alpha(1-\delta)P_{POI}(z_t, s_t; \varepsilon + \mu, \varepsilon) + \alpha\delta\, P_{POI}(z_t, s_t; \varepsilon, \varepsilon + \mu) + (1-\alpha)P_{POI}(z_t, s_t; \varepsilon, \varepsilon) \quad (1)$$

gdzie $P_{POI}(z_t, s_t, \lambda_Z, \lambda_S)$ oznacza łączny dwuwymiarowy rozkład Poissona dla transakcji zakupu i sprzedaży waluty bazowej o intensywnościach: $\lambda_B \in \{\varepsilon + \mu, \varepsilon\}$ oraz $\lambda_S \in \{\varepsilon, \varepsilon + \mu\}$.

Z przyjętego założenia, że transakcje zakupu i sprzedaży zawierane są zgodnie z procesem Poissona wynika, że liczby transakcji zakupu i sprzedaży waluty bazowej w poszczególnych trzech rodzajach okresów (charakteryzującym się napływem złej informacji lub dobrej informacji, lub bez napływu sygnałów informacyjnych) są od siebie całkowicie niezależne. Zgodnie z klasyczną specyfikacją modelu, omawiane procesy nie mogą podlegać bowiem ani autokorelacji, ani korelacji krzyżowej. Dodatkowo, na podstawie wzoru (1), kowariancja pomiędzy procesami zakupu i sprzedaży (wynikającymi z mieszanki trzech procesów Poissona) ma postać $Cov(Z_t, S_t) = -\alpha^2\delta(1-\delta)\mu^2$, czyli implikuje ujemną korelację pomiędzy zmiennymi Z_t i S_t. Empiryczne wyniki badań dowodzą jednak, że okresowe fluktuacje liczebności transakcji zakupu i sprzedaży charakteryzują się dodatnią korelacją – choćby ze względu na występowanie wewnątrzdziennej lub wewnątrztygodniowej sezonowości (por. [Venter i de Jongh 2004]). Inną wadą, wynikającą z przyjęcia założenia o rozkładzie Poissona, stanowi zaakceptowanie *implicite* restrykcji o równości pomiędzy wartością oczekiwaną liczby transakcji zakupu lub sprzedaży a wariancją liczby zakupu i sprzedaży. Empiryczne badania wskazują jednak, że dyspersja (wariancja) rozkładu badanych zmiennych jest dużo większa niż wynikałoby to z rozkładu Poissona (por. [Venter i de Jongh 2004]).

Zaproponowane w pracy [Wünsche 2007] rozwiązanie polega na wykorzystaniu innych niż rozkład Poissona rozkładów prawdopodobieństwa dla liczby transakcji zakupu i sprzedaży. [Wünsche 2007] proponuje wykorzystanie dwuwymiarowego rozkładu ujemnego dwumianowego (*Negative Binomial,* NegBin)[2]. Dwuwymiarowy rozkład liczby transakcji zakupu i sprzedaży waluty bazowej ma wówczas postać:

$$P_{NEG}(Z_t = z_t, S_t = s_t) = \frac{\Gamma(z_t + s_t + \kappa)}{\Gamma(z_t + 1)\Gamma(s_t + 1)\Gamma(\kappa)}\left(\frac{\kappa}{\zeta}\right)^{\kappa}\left(\frac{\lambda_z}{\zeta}\right)^{z_t}\left(\frac{\lambda_s}{\zeta}\right)^{s_t} \quad (2)$$

[2] Rozkład ujemny dwumianowy można traktować jako mieszankę rozkładu Poissona z rozkładem gamma. Zarówno intensywność transakcji zakupu jak i sprzedaży zależą jednocześnie od tego samego ukrytego czynnika ω o rozkładzie gamma. Implikuje to dodatnią korelację zmiennych (zob. [Wünsche 2007]).

gdzie $\zeta = \kappa + \lambda_z + \lambda_s$, natomiast λ_z, oraz λ_s oznaczają intensywności odpowiednio transakcji zakupu i sprzedaży waluty bazowej.

Podobnie jak w klasycznym modelu EKOP, bazującym na rozkładzie Poissona, oczekiwane liczby transakcji zakupu i sprzedaży są równe odpowiednio $E(Z) = \lambda_z$ i $E(S) = \lambda_s$, natomiast wariancje procesów brzegowych mają postać: $V(Z) = \lambda_z + \kappa^{-1}\lambda_z^2$ oraz $V(S) = \lambda_s + \kappa^{-1}\lambda_s^2$. Parametr $E(Z) = \kappa$ odpowiada zatem za miarę dyspersji rozkładu (wariancja może być większa od wartości oczekiwanej). Model umożliwia większą dyspersję w porównaniu z rozkładem Poissona. Dwuwymiarowy rozkład NegBin zbiega do dwuwymiarowego rozkładu Poissona jeśli $\kappa \to \infty$. Przy wykorzystaniu mieszanki rozkładów ujemnych dwumianowych do opisu łącznego rozkładu liczby transakcji zakupu i sprzedaży waluty bazowej, odpowiednikiem rozkładu prawdopodobieństwa zadanego wzorem (1) jest rozkład:

$$\begin{aligned} P(Z_t = z_t, S_t = s_t) &= \alpha\delta P_{NEG}(z_t, s_t; \varepsilon, \varepsilon + \mu) \\ &+ \alpha(1-\delta) P_{NEG}(z_t, s_t; \varepsilon + \mu, \varepsilon) \\ &+ (1-\alpha) P_{NEG}(z_t, s_t; \varepsilon, \varepsilon) \end{aligned} \tag{3}$$

na podstawie którego możemy otrzymać funkcję wiarygodności:

$$\begin{aligned} L_{EKOP}(\Theta) = \prod_{t=1}^{T} & \frac{\Gamma(z_t + s_t + \kappa)}{\Gamma(z_t + 1)\Gamma(s_t + 1)\Gamma(\kappa)} \varepsilon^{z_t + s_t} \kappa^{-\kappa} (2\varepsilon + \kappa)^{z_t + s_t + \kappa} \\ \cdot & \left[(1-\alpha) + \left[\alpha\delta(1+a)^{s_t} + \alpha(1-\delta)(1+a)^{z_t} \right] \left(\frac{1}{1+b} \right)^{z_t + s_t + \kappa} \right] \end{aligned} \tag{4}$$

gdzie $a = \dfrac{\mu}{\varepsilon}$, $b = \dfrac{\mu}{2\varepsilon + \kappa}$.

Na podstawie ocen parametrów modelu można dokonać jednoznacznego podziału rozważanych okresów, w których dokonuje się obrotu na: (1) okresy, w których zaobserwowano napływ dobrej informacji, (2) okresy, w których zarejestrowano napływ złej informacji oraz (3) okresy, w których nie było napływu jakichkolwiek sygnałów informacyjnych. Identyfikacji takiej można dokonać na podstawie prawdopodobieństw *a posteriori* wyznaczonych na podstawie twierdzenia Bayesa:

$$P(NI_t \mid z_t, s_t) = (1-\alpha) P_{Neg}(Z_t, S_t; \varepsilon, \varepsilon) / L_{EKOP} \tag{5}$$

$$P(ZI_t \mid z_t, s_t) = \alpha\delta P_{Neg}(Z_t, S_t; \varepsilon, \varepsilon + \mu) / L_{EKOP} \tag{6}$$

$$P(DI_t \mid z_t, s_t) = \alpha(1-\delta) P_{Neg}(Z_t, S_t; \varepsilon + \mu, \varepsilon) / L_{EKOP} \tag{7}$$

gdzie $P(NI_t \mid z_t, s_t)$ oznacza prawdopodobieństwo braku napływu sygnałów informacyjnych, $P(ZI_t \mid z_t, s_t)$ – prawdopodobieństwo napływu złej informacji dla waluty bazowej, a $P(GI_t \mid z_t, s_t)$ – prawdopodobieństwo napływu dobrej informacji (prawdopodobieństwa pod warunkiem zaistnienia z_t transakcji zakupu oraz s_t transakcji sprzedaży w okresie t). Końcowej identyfikacji okresów dokonuje się na podstawie tego, które z oszacowanych prawdopodobieństw jest dla danego momentu t największe.

BADANIE PROCESU TRANSAKCYJNEGO NA RYNKU EUR/PLN

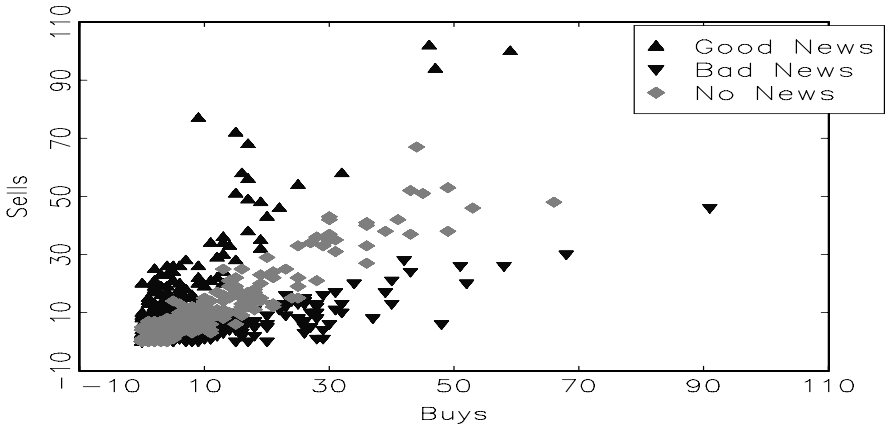
W przykładzie empirycznym wykorzystano zbiory danych dotyczące transakcji na międzybankowym kasowym rynku złotego, zarejestrowanych w systemie Reuters 3000 Spot Matching System, w okresie styczeń-lipiec 2008 r.[3] Z uwagi na małą aktywność rynku w nocy i dni świąteczne, okresem badania objęto dni robocze w godz. 9.00-17.00 CET. Szeregi danych tikowych poddano agregacji do piętnastominutowej częstotliwości, wyodrębniając zmienne: (1) logarytmiczną stopę zwrotu ze średniego kursu (kurs mid) EUR/PLN (wyrażoną w punktach bazowych), (2) liczbę transakcji sprzedaży i (3) liczbę transakcji zakupu.

Model EKOP zakłada, że parametry są niezmienne w czasie. Tymczasem sygnały informacyjne napływają na rynek finansowy „seriami", ponieważ źródła takich informacji są ze sobą często powiązane. W celu wychwycenia potencjalnych trendów oraz sezonowości w kształtowaniu się poszczególnych współczynników modelu, zastosowano procedurę zaproponowaną w pracach [Gencay, Gradojevic i Selcuk 2007] oraz [Gencay i Gradojevic 2008]. Model EKOP oszacowano[4] dla każdego z 145 dni wchodzących w zakres próby oddzielnie, każdorazowo na podstawie czterdziestu wewnątrzdziennych obserwacji odpowiadających piętnastominutowym interwałom czasowym. Na podstawie oszacowań parametrów modelu każdy z piętnastominutowych interwałów zakwalifikowano do jednego z trzech reżimów informacyjnych (por. wzory 5-7). Wyodrębnienie różnych okresów przedstawiono na rysunku 1. Można zauważyć, że istotna nadwyżka transakcji zakupu euro nad transakcjami sprzedaży pozwala wyodrębnić okres o napływie dobrej informacji dla waluty bazowej (euro). Analogicznie, istotna nadwyżka transakcji sprzedaży nad transakcjami zakupu umożliwia wyodrębnienie okresu o napływie złej informacji dla euro, a zatem dobrej dla złotego.

[3] Więcej informacji na temat samej platformy transakcyjnej, jej funkcjonowania oraz struktury danych znaleźć można w [Bień 2010].

[4] Model oszacowano na podstawie samodzielnie przygotowanych kodów do estymacji w programie ekonometrycznym Gauss (wersja 8.0). W procedurze maksymalizacji funkcji wiarygodności wykorzystano bibliotekę Maxlik.

Rysunek 1. Podział okresów transakcyjnych na podstawie dysproporcji transakcji zakupu i sprzedaży otrzymanej za pomocą modelu EKOP.



Źródło: obliczenia własne

Wyodrębnienie trzech typów okresów dla procesu zawierania transakcji pozwala odpowiedzieć na pytanie, w jakim stopniu aktywność dealerów polegająca na składniu agresywnych zleceń zakupu lub sprzedaży (z różną intensywnością) wpływa na wewnątrzdzienną fluktuację kursu walutowego. W celu zbadania takiego efektu oszacowano regresje:

$$r_t = \omega_{NI}\hat{P}(NI_t \mid z_t, s_t) + \omega_{ZI}\hat{P}(ZI_t \mid z_t, s_t) + \omega_{GI}\hat{P}(GI_t \mid z_t, s_t) + \nu_{t,r} \tag{8}$$

$$r_t^2 = \alpha_{NI}\hat{P}(NI_t \mid z_t, s_t) + \alpha_{ZI}\hat{P}(ZI_t \mid z_t, s_t) + \alpha_{GI}\hat{P}(GI_t \mid z_t, s_t) + \nu_{t,r^2} \tag{9}$$

Wyniki oszacowań przedstawiono w tabeli 1. Można zauważyć, że poszczególne reżimy informacyjne generują całkowicie inne zachowanie stóp zwrotu. W przypadku napływu dobrej informacji dla euro (lub złej dla złotego) kurs EUR/PLN rośnie średnio o około 4,35 punktu bazowego (deprecjacja złotego), a w przypadku napływu złej informacji dla euro (lub dobrej dla złotego) kurs EUR/PLN maleje średnio o około 4,4 punktu bazowego (aprecjacja złotego). Wynik taki jest zgodny z oczekiwaniami. Należy zwrócić uwagę, że podział na poszczególne stany informacyjne przeprowadzo na podstawie wiedzy o liczbie transakcji zakupu i sprzedaży w poszczególnych okresach. Procedura taka wykorzystuje zatem informację o wartości przepływu zleceń i potwierdza rezultaty otrzymane w pracy [Bień 2010]. Ciekawe wyniki otrzymano w przypadku regresji dla kwadratów stóp zwrotu. Można zauważyć, że zmienność kursu EUR/PLN jest większa w przypadku napływu złej informacji dla złotego (lub dobrej dla euro), niż w okresach charakteryzujących się napływem dobrej informacji dla złotego (lub dobrej dla euro). Wniosek ten potwierdza, że kurs EUR/PLN w nieco większym stopniu zależy od informacji (ogłoszeń, oczekiwań, wyników badań itp.) dotyczących wartości walut krajów *emerging markets* (w tym złotego) niż tych

dotyczących wartości głównych walut. Dyspersja kursu wzrasta bardziej w następstwie negatywnych informacji dla złotego. Tym samym otrzymane wyniki wskazują na występowanie dźwigni finansowej jako jednej ze stylizowanych cech finansowych szeregów czasowych o bardzo wysokiej częstotliwości.

Tabela 1. Wpływ reżimów informacyjnych na stopę zwrotu i kwadrat stopy zwrotu z kursu EUR/PLN. Wartości p odpowiadają statystykom t-studenta obliczonym na podstawie średnich błędów szacunku odpornych na autokorelację i heteroskedastyczność składnika losowego (*autocorrelation and heteroskedasticity robust standard errors*).

parametry	oszacowanie	wart. p	parametry	oszacowanie	wart. p
	równanie dla r:			równanie dla r^2:	
ω_{NI}	-0,216	0,017	α_{NI}	23,513	0,000
ω_{GI}	4,357	0,000	α_{GI}	51,547	0,000
ω_{ZI}	-4,403	0,000	α_{ZI}	48,651	0,000

Źródło: obliczenia własne

BIBLIOGRAFIA

Benos, E., Jochec M. (2007) Testing the PIN Variable Working Paper, University of Illinois.

Berger D., Chaboud A., Chernenko S., Howorka E., Wright J. (2008) Order Flow and Exchange Rate Dynamics in Electronic Brokerage System Data, Journal of International Economics, No. 75, s. 31–62.

Bień K. (2010) Przepływ zleceń a kurs walutowy. Badanie mikrostruktury międzybankowego kasowego rynku złotego, Bank i Kredyt, 5, 5-39.

Brockman P., Chung D. Y. (2000) Informed and Uniformed Trading in an Electronic Market-Driven Environment, The Financial Review, 35, 125 –146.

Danielsson, J., Luo J., Payne R. (2002) Exchange Rate Determination and Inter-Market Flow Effects, Mimeo, London School of Economics.

Easley D., O'Hara M. (1987) Price, Trade Size, and Information in Securities Markets, Journal of Financial Economics, No. 19, s. 69–90.

Easley D., O'Hara, M., Saar G. (2001) How Stock Splits Affect Trading: A Microstructure Approach, Journal of Financial and Quantitative Analysis, 36, 25–51.

Easley D., Kiefer N., O'Hara M., Paperman J. (1996) Liquidity, Information and Infrequently Traded Stocks, Journal of Finance, 51, 1405–1436.

Evans M. D., Lyons R. K. (2002a) Informational Integration and FX Trading, Journal of International Money and Finance, 21, 807–831.

Evans M. D., Lyons R. K. (2002b) Order Flow and Exchange Rate Dynamics, Journal of Political Economy, 110, 170–180.

Gençay, R., Gradojevic, N. (2008) Informed Trading in Electronic Foreign Market, Working paper, Department of Economics, Simon Fraser University, Canada.

Gençay, R., Gradojevic, N, Selçuk, F. (2007) When Do Informed Traders Arrive in Foreign Exchange Markets, Working paper, Department of Economics, Simon Fraser University, Canada.

Glosten, L. R. (1994) Is the Electronic Open Limit Order Book Inevitable?, The Journal of Finance, 49, 1127–1161.

Glosten L. R., Milgrom P. R. (1985) Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, Journal of Financial Economics, 14, 71–100.

O'Hara M. (1995), Market Microstructure, Basil Blackwell, Oxford.

Reza, S., Wilson C. (2007) Does Corporate Ownership Impact the Probability of Informed Trading?, International Journal of Business Research, 7, 188–210.

Rime D., Sarno L., Sojli E. (2010), Exchange Rate Forecasting, Order Flow and Macroeconomic Information, Journal of International Economics, No. 80, s. 72-88.

Scalia A. (2008) Is Foreign Exchange Intervention Effective? Some Microanalytical Evidence from the Czech Republic, Journal International Money and Finance, 27, 529–546.

Seppi, D. J. (1997) Liquidity Provision with Limit Orders and a Strategic Specialist, Review of Financial Studies, 10, 103–150.

Wünsche O. (2007) Using Mixed Poisson Distributions in Sequential Trade Models, Working Paper, Group Quantitative Risk Methodology.

Venter J., de Jongh D. (2004) Extending the EKOP Model to Estimate the Probability of Informed Trading", Studies in Economics and Econometrics 30(2), 25-39.

THE SEQUENTIAL TRADE MODEL – APPLICATION TO THE ANALYSIS OF THE TRADING PROCESS IN THE POLISH ZLOTY MARKET

Abstract: In the paper we estimate the degree of expectation heterogeneity among currency dealers on the interbank spot market of the EUR/PLN currency pair. We use the flow of buy and sell orders submitted to the market (orders to buy or to sell euro). In a market microstructure study we present the generalized version (see [Wünsche 2007]) of the sequential trade model proposed in [Easley et al. 1996]. The aim of the model is to estimate the informational content of trades upon four distinct measures: (1) probability of the news arrival, (2) probability that the news is bad, (3) arrival rate of informed trades and (4) arrival rate of uninformed trades (i.e. liquidity trades). In the empirical part of the paper we estimate the impact of the news regime (periods classified as good, bad and neutral for the base currency) on the volatility of the EUR/PLN exchange rate.

Keywords: market microstructure, sequential trade models, exchange rate

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 52 – 64

ROZWÓJ POLSKIEGO ROLNICTWA W KONTEKŚCIE REGIONALNEGO ZRÓŻNICOWANIA W LATACH 1998-2010[1]

Agata Binderman
Katedra Zastosowań Informatyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: agata_binderman@sggw.pl

Streszczenie: W pracy przy pomocy wielowymiarowej analizy porównawczej oceniono regionalne zróżnicowanie rolnictwa w Polsce, w latach 1998-2010. Do analizy użyto cztery różne mierniki syntetyczne. Na ich podstawie dokonano uporządkowania i klasyfikacji województw. Wyniki pokazują, że pomimo szybkiego w ostatnich latach wzrostu poziomu rolnictwa dysproporcje miedzy województwami nie maleją. W niniejszej pracy zbadano zróżnicowanie poziomu rozwoju rolnictwa, przyjmując jeden sposób normalizacji zmiennych, jeden wzór określający odległość między rozważanymi obiektami i cztery różne sposoby klasyfikacji tych obiektów – zarówno wzorcowe jak i bezwzorcowe.

Słowa kluczowe: normalizacja zmiennych, miernik syntetyczny, poziom rozwoju rolnictwa, klasyfikacja, grupowanie

WSTĘP

Do analizy zjawisk złożonych, takich jak poziom rozwoju czy potencjał rolnictwa oraz oceny województw pod tym kątem, konieczne jest rozważenie wielu czynników [Harasim 2006, 2009 Krasowicz, Kukuła 2006]. Zmienne objaśniające dane zjawisko są zarówno mierzalne jak i niemierzalne. Analiza na podstawie tych danych może mieć charakter niejednoznaczny. Do sumarycznego charakteryzowania zjawisk złożonych stosuje się więc zmienne syntetyczne (agregatowe). Zastąpienie ciągu wielu cech objaśniających przez zmienną syntetyczną daje pewną ocenę badanego zjawiska. Zmienne syntetyczne poza swą

[1] Praca wykonana w ramach zadania badawczego pt. „Dynamika regionalnego zróżnicowania rolnictwa w Polsce w latach 1998 – 2011”.

niejednoznacznością mają jednak taką wadę, że nie zawsze można im nadać interpretację merytoryczną. Istnieje wiele metod tworzenia zmiennych syntetycznych. Metody wzorcowe zakładają istnienie pewnego hipotetycznego obiektu wzorcowego. Uporządkowania badanych obiektów dokonuje się w zależności od osiągniętych przez nie odległości od obiektu wzorcowego [Binderman 2011]. Metody te wykorzystują odpowiednio wybrane zmienne diagnostyczne (objaśniające) charakteryzujące badane zjawisko i różnią się miedzy sobą, co do sposobu normalizacji zmiennych oraz postaci funkcji agregujących [Kukuła 2000, Malina 2004, Młodak 2006, Nowak 1990, Zeliaś 2000].

Badania autorki dotyczące polskiego rolnictwa [Binderman 2005a,b 2006a,b, 2007, 2008, 2009, 2009a, 2010, 2010a, 2011] wykazały, że klasyfikacje i grupowania województw, uzyskiwane za pomocą obiektów wzorcowych w istotny sposób mogą zależeć zarówno od wybranej metryki, sposobu normalizacji zmiennych jak i od wyboru wzorca. W niniejszej pracy autorka podjęła się oceny zróżnicowania poziomu rolnictwa polskich województw w latach 1998-2010, przyjmując jeden sposób normalizacji zmiennych, jeden wzór określający odległość między rozważanymi obiektami i cztery różne sposoby klasyfikacji tych obiektów – trzy wzorcowe i jeden bezwzorcowy.

W badaniach regionalnych, przy konstrukcji mierników syntetycznych duże znaczenie odgrywa dobór odpowiedniego zestawu cech, które jako zmienne diagnostyczne będą charakteryzowały analizowane obiekty. W literaturze jest wiele propozycji tych cech (mierników), opisujących regionalne zróżnicowanie rolnictwa. Do najważniejszych przyczyn regionalnego zróżnicowania rolnictwa w Polsce zalicza się: warunki naturalne, zaszłości historyczne związane z systemami organizacji i zarządzania, gęstość zaludnienia, stopień uprzemysłowienia i urbanizacji oraz rozwój infrastruktury, przemiany agrarne, stopień wyposażenia rolnictwa w czynniki wytwórcze, politykę przestrzennego zagospodarowania kraju, tradycje i poziom kultury rolnej [Dąbkowski 2000; Gołaś, Wysocki 1997; Harasim 2001; Heady, Dillon 1961, Heller, Zawisza (red.) 2000, Poczta, Bartkowiak 2012].

Analizując regionalne zróżnicowanie rolnictwa warto wspomnieć o pracy [Krasowicz, Kukuła 2006], gdzie przedstawiono zarys badań prowadzonych przez Instytut Uprawy, Nawożenia i Gleboznawstwa (IUNG) w Puławach nad regionalnym zróżnicowaniem polskiego rolnictwa oraz monografie wydane przez IUNG [Harasim red. 2006, 2009], w których wielu autorów przedstawiło wyniki dotyczące regionalnego zróżnicowania produkcji rolniczej w Polsce. Według autorów Krasowicza i S. Kukuły zróżnicowanie rolnictwa i obszarów wiejskich w Polsce jest ważną przesłanką polityki regionalnej, wyznacznikiem problemów, wymagających wsparcia ze strony nauki. Zakres badań IUNG obejmował między innymi regionalne zróżnicowanie: wykorzystania potencjału produkcyjnego, intensywności organizacji, produkcji zbóż, możliwości i rozwoju działalności alternatywnych. Polska w stosunku do rozwiniętych gospodarczo krajów UE ma

znaczny potencjał produkcyjny rolnictwa, który jest słabo wykorzystany, przy dużym, regionalnym jego zróżnicowaniu.

METODYKA BADAŃ

W celu przeprowadzenia badań zebrano dane empiryczne dotyczące polskiego rolnictwa zarówno w ujęciu przestrzennym (wojewódzkim), jak i dynamicznym - za lata 1998-2010. Źródłem danych były publikacje, roczniki statystyczne województw, roczniki statystyczne GUS z lat 1998-2012 oraz dane z Powszechnych Spisów Rolnych 2002, 2010 i Narodowych Spisów Powszechnych 2002, 2010. Wcześniejsze materiały zostały pozyskane bezpośrednio z Departamentu Statystyki Rolnictwa i Środowiska GUS.

W prowadzonych badaniach, do wyboru zmiennych diagnostycznych wykorzystano wstępnie 76 cech dotyczących rolnictwa oraz 4 cechy makroekonomiczne, takie jak np. powierzchnia gruntów ogółem czy produkt krajowy brutto, opisujące poszczególne województwa. Czynników określających poziom i rozwój rolnictwa jest oczywiście wiele. W pracy brane były pod uwagę zarówno cechy pokazujące jakość rolniczej przestrzeni produkcyjnej danego województwa, strukturę agrarną, wielkość nakładów na produkcję rolną czy stopień wyposażenia rolnictwa w czynniki wytwórcze i poziom produktywności rolnictwa, jak również gęstość zaludnienia, stan i strukturę zatrudnienia, rozwój infrastruktury w regionie oraz cechy uwzględniające czynniki społeczne (np. dotyczące edukacji). Ważnym aspektem w badaniach nad zróżnicowaniem regionalnym rolnictwa jest powiązanie gospodarstw rolnych z rynkiem, mierzone towarowością produkcji, takimi jak produkcja towarowa czy wielkość skupu produktów rolnych.

Listę rozpatrywanych 80 cech zamieszczono w pracy doktorskiej autorki [Binderman 2007]. Ponieważ zbyt liczny zbiór cech mógłby utrudniać badania [zob. np. Zeliaś 2000], do dalszej analizy przyjęte zostały te wskaźniki, które umożliwiały merytoryczną ocenę analizowanego zjawiska i posiadały wysoką zdolność różnicującą badane obiekty. Ze względu na to, że nie wszystkie cechy występowały w formie wskaźników, dokonano odpowiednich przeliczeń, w wyniku których otrzymano następujące 43 wskaźniki: udział UR w % powierzchni gruntów ogółem; udział gruntów ornych w % powierzchni UR; udział sadów w % powierzchni UR; udział łąk i pastwisk w % powierzchni UR; udział odłogów i ugorów w % powierzchni UR; wskaźnik waloryzacji rolniczej przestrzeni produkcyjnej w pkt; plony zbóż w t z 1 ha; plony ziemniaków w t z 1 ha; plony buraków cukrowych w t z 1 ha; plony rzepaku i rzepiku w t z 1 ha; obsada bydła na 100 ha UR; obsada trzody chlewnej na 100 ha UR; obsada owiec na 100 ha UR; obsada drobiu na 100 ha UR; skup zbóż na 1 ha UR w kg; skup ziemniaków na 1 ha UR w kg; skup żywca rzeźnego w przelicz na mięso w kg na 1 ha UR; skup mleka w l na 1 ha UR; skup zbóż w kg na 1 ha powierzchni upraw; skup ziemniaków w kg na 1 ha powierzchni upraw; skup

buraków cukrowych w kg na 1 ha powierzchni upraw; skup rzepaku i rzepiku w kg na 1 ha powierzchni upraw; skup warzyw w kg na 1 ha powierzchni upraw; skup owoców z drzew w kg na 1 ha powierzchni upraw; skup żywca rzeźnego w przeliczeniu na mięso w kg na szt. pogłowia; skup mleka w litrach na szt. pogłowia; wartość skupu w zł produktów rolnych ogółem na 1 ha UR; wartość skupu w zł produktów roślinnych na 1 ha UR; wartość skupu w zł produktów zwierzęcych na 1 ha UR; towarowa produkcja rolnicza (zł/ha); towarowa produkcja roślinna (zł/ha); towarowa produkcja zwierzęca (zł/ha); zużycie nawozów mineralnych w kg na 1 ha UR; liczba ciągników w sztukach na 100 ha UR; zasoby siły pociągowej na 100 ha UR w jednostkach pociągowych; wartość dodana brutto w rolnictwie w zł na 1 ha UR; wartość dodana brutto w rolnictwie w zł na 1 pracującego w rolnictwie; nakłady inwestycyjne w rolnictwie w zł na 1 ha UR; wskaźnik zatrudnienia w rolnictwie w %; udział wartości brutto środków trwałych w rolnictwie w % ogółem; średnia powierzchnia gospodarstwa rolnego; produkt krajowy brutto w zł na 1 mieszkańca; drogi publiczne w km na 100 km^2. Spośród 43 cech wybrano takie, które spełniały następujące kryteria formalno – statystyczne: kompletność i porównywalność danych dla wszystkich województw, kompletność i porównywalność danych w latach 1998-2010, dostateczna zmienność przestrzenna, brak nadmiernego skorelowania zmiennych.

W pierwszej kolejności status diagnostyczności otrzymały cechy, które posiadały zdolności dyskryminacyjne badanych obiektów, tzn. te, które odznaczały się dostatecznie dużą zmiennością po obiektach. Wartość krytyczna współczynnika zmienności V* została wyznaczona na poziomie 0,1. Eliminacji podlegały cechy uznawane za quasi-stałe, dla których współczynnik zmienności V był mniejszy od wartości krytycznej, tj. $V<0,1$. O dalszej selekcji decydowała względna amplituda wahań. Ze zbioru zmiennych wyeliminowano te cechy, dla których współczynnik względnej amplitudy wahań A przybierał wartości z przedziału [1;1,2] [Kukuła 1989]. W ostatnim kroku doboru zmiennych, ze względu na kryterium korelacyjne, zbadane zostały właściwości informacyjne cech oraz ich wzajemne powiązania. Celem tego etapu badań był dobór zmiennych diagnostycznych, które wniosą do analizy jak najwięcej niepowielających się informacji na temat badanego zjawiska. Do badań posłużono się metodą parametryczną Hellwiga [Hellwig 1981]. Analizy dokonano na podstawie macierzy R, zawierającej współczynniki korelacji liniowej Pearsona między badanymi zmiennymi. Kryterium klasyfikacji zmiennych była wartość krytyczna r*. Przyjmując arbitralnie wartość krytyczną $r^*=0,7$, cechy, dla których współczynniki korelacji były mniejsze od wartości krytycznej, utworzyły grupę zmiennych podobnych. W każdej grupie wyróżniono jedną cechę centralną oraz zmienne satelitarne, które były silnie skorelowane ze zmienną centralną. Współczynnik korelacji pomiędzy cechą centralną i zmienną satelitarną był większy niż wartość krytyczna r*, natomiast cechy nienależące do grup potraktowano, jako zmienne izolowane. Z zestawu potencjalnych cech diagnostycznych wyeliminowano wszystkie te zmienne, których współczynnik korelacji z wyróżnioną cechą był, co do wartości bezwzględnej większy od r*.

Z pozostałych cech wyłoniono kolejną o największej sumie współczynników korelacji i wyeliminowano zmienne z nią najsilniej skorelowane. W kolejnych etapach procedurę tą powtarzano aż do wyczerpania zasobu cech.

Dla porównania dokonano również analiz rozważanych cech przy użyciu innych metod selekcji zmiennych. Przeprowadzono analizę czynnikową oraz analizę Biplot. Wszystkie obliczenia numeryczne wykonano przy użyciu pakietów programów statystycznych Statistica (analiza czynnikowa), SAS (analiza Biplot) oraz arkusza kalkulacyjnego Excel.

Zastosowanie analizy czynnikowej i analizy Biplot pozwoliło na zredukowanie dużej liczby zmiennych do kilku wzajemnie nieskorelowanych czynników. Do opisu regionalnego zróżnicowania rolnictwa wyłoniono ostatecznie zestaw 10 cech. Wszystkie wybrane zmienne miały charakter stymulant, co oznacza, że większe wartości tych cech informowały o wyższym poziomie rozwoju badanego zjawiska. Wybrane zmienne określają całościowo poziom rolnictwa w danych latach, w Polsce, umożliwiając jednocześnie pokazanie różnic, jakie występują pomiędzy województwami. Poniżej została zamieszczona ostateczna lista wybranych 10 zmiennych diagnostycznych.

X_1 Udział użytków rolnych w % powierzchni ogółem.
X_2 Wskaźnik waloryzacji rolniczej przestrzeni produkcyjnej (pkt).
X_3 Plony buraków cukrowych w tonach z 1 hektara.
X_4 Obsada bydła w sztukach na 100 hektarów użytków rolnych.
X_5 Skup ziemniaków w kilogramach na 1 hektar użytków rolnych.
X_6 Skup rzepaku i rzepiku w kg na 1 hektar powierzchni upraw.
X_7 Skup owoców z drzew w kg na 1 hektar powierzchni upraw.
X_8 Wartość skupu produktów rolnych ogółem w złotych na 1 ha UR.
X_9 Liczba ciągników w sztukach na 100 ha UR.
X_{10} Nakłady inwestycyjne w rolnictwie w złotych na 1 ha UR.

Dane można przedstawić za pomocą macierzy trójwymiarowej (województwo*wartość cechy*rok) bądź za pomocą macierzy, w której każdy wiersz przedstawia obiekt, reprezentowany przez cechy danego województwa w danym roku. W pracy wybrano ten drugi sposób. Otrzymano w ten sposób m=16*13 obiektów $\mathbf{Q}_1,\mathbf{Q}_2,...,\mathbf{Q}_{208}$, z których każdy opisany został przez n=10 cech $X_1,X_2,...,X_{10}$. Wartości przyjętych zmiennych diagnostycznych dla 16 województw oraz dla 13 lat (1998-2010), utworzyły macierz $\mathbf{X}$ o wymiarach 10×208. W macierzy tej wiersz np. 1, 17, 33, 49, 65, 81, 97, 113, 129, 145, 161, 177 oraz 193, reprezentuje województwo dolnośląskie, w poszczególnych latach okresu 1998-2010.

Na podstawie wartości przyjętych przez zmienne diagnostycznych dla 16 województw w poszczególnych 13 latach badanego okresu, stworzono dwa stałe (statyczne), hipotetyczne województwa, minimalne (najgorsze) $\mathbf{Q}_0$ oraz maksymalne (najlepsze) $\mathbf{Q}_{209}$, opisane odpowiednio za pomocą najmniej i najbardziej korzystnego zestawu wartości cech, województwa te w niniejszej pracy będą reprezentowane przez wektory $\mathbf{x}_0$ oraz $\mathbf{x}_{209}$ o 10 składowych każdy.

Wektory te są wzorcami dla całego badanego okresu 1998-2010, wyznaczając w n-wymiarowej przestrzeni Euklidesa $\Re_+^n$ kostkę $[\mathbf{x}_0,\mathbf{x}_{209}]$, oznacza to, że dla każdego $i\in[1,208]$: $x_i\in[\mathbf{x}_0,\mathbf{x}_{209}]$. W ten sposób do dalszej analizy uzyskano macierz

$$\mathbf{X}=\left[x_{ij}\right]_{210\times 10}$$

danych mającą 210 wierszy i 10 kolumn.

Ponieważ wybrane zmienne diagnostyczne miały różne miana oraz różne rzędy wielkości, zmienne te poddano normalizacji. W celu sprowadzenia zmiennych do porównywalności, spośród wielu rodzajów normowania wybrano i zastosowano unitaryzację zerowaną. Wybór metody normowania zmiennych był konsekwencją rezultatów, otrzymanych przez autorkę (Binderman 2010). Unormowane wartości dla poszczególnych zmiennych przy m=208, n=10, obliczono według wzorów [Kukuła 2000]:

$$z_{ij}=\frac{x_{ij}-x_{0j}}{x_{m+1j}-x_{0j}},\ 0\le i\le 209,\ 1\le j\le 10 \tag{1}$$

Przekształcone w ten sposób cechy, poprzez wyeliminowanie jednostek miar stały się wzajemnie porównywalne. Zmienne z_{ij} przekształcone za pomocą metody unitaryzacji zerowanej (MUZ), przyjmują wartości z przedziału [0,1]. Dokonane przekształcenia symbolicznie można zapisać: $\mathbf{Z}=\varphi(\mathbf{X})$, gdzie $\mathbf{X}$ – macierz obserwacji, $\mathbf{Z}=[z_{ij}]_{210\times 10}$. Po transformacji zmiennych, statyczne wektory wzorcowe przedstawiają się następująco: $\mathbf{z}_0=\mathbf{0}=[0,0,...,0]$, $\mathbf{z}_{209}=\mathbf{1}=[1,1,...,1]$, tj. w macierzy $\mathbf{Z}$ pierwszy wiersz składa się z samych zer, natomiast ostatni z samych jedynek. W pracy do mierzenia odległości między wektorami (obiektami) $\mathbf{x}=(x_1,x_2,...,x_n),\mathbf{y}=(y_1,y_2,...,y_n)\in\Re_+^n$ wykorzystano prostą do interpretacji metrykę Minkowskiego:

$$d(\mathbf{x},\mathbf{y})=\left(\sum_{i=1}^{n}(x_i-y_i)^p\right)^{\frac{1}{p}}\ \text{przy p=4.}$$

W pracach autorki [Binderman 2004, 2005b, 2006] pokazano, że stosowanie metod opartych tylko na jednym wzorcu, w wielu przypadkach prowadzi do otrzymania błędnych wyników, które nie spełniają warunków poprawności [Jackson 1971]. Z powyższego powodu w niniejszej pracy, do analizy regionalnego zróżnicowania polskiego rolnictwa wykorzystano dwa mierniki, które są oparte na dwóch wzorcach oraz bezwzorcowy, izotoniczny miernik rozwoju typu M. Pluty [Cieślak 1974, Kukuła 2000, Malina 2004, Zeliaś 2000].

Użyty miernik bezwzorcowy został określony za pomocą wzoru:

$$m_1(\mathbf{x})=\frac{1}{n}\sum_{j=1}^{n}x_j\,, \tag{2}$$

natomiast z metod wykorzystujących dwa wzorce $\mathbf{x}_0$ i $\mathbf{x}_{m+1}$ użyto miernik TOPSIS [Hwang, Yoon 1981] i miernik autorki [Binderman 2006]. Mierniki te określone są dla danego wektora $\mathbf{x} \in [\mathbf{x}_0, \mathbf{x}_{m+1}]$ za pomocą wzorów [Binderman 2011]:

$$\mathsf{m}_2(\mathbf{x}) = \frac{d(\mathbf{x}_0, \mathbf{x})}{d(\mathbf{x}_0, \mathbf{x}) + d(\mathbf{x}_{m+1}, \mathbf{x})} \tag{3}$$

$$\mathsf{m}_3(\mathbf{x}) = \frac{1}{2} + \frac{d(\mathbf{x}_0, \mathbf{x}) - d(\mathbf{x}_{m+1}, \mathbf{x})}{2d(\mathbf{x}_0, \mathbf{x}_{m+1})}, \tag{4}$$

gdzie w naszym przypadku m=208.

Podane wyżej mierniki po normalizacji zmiennych przy pomocy unitaryzacji zerowanej – wzór (1) są znormalizowane tj.

$$0 \le \mathsf{m}_k(\mathbf{z}) \le 1 \quad \text{dla} \quad \mathbf{z} \in < \mathbf{z}_0, \mathbf{z}_{m+1} >, \; k = 1,2,3.$$

Zauważmy ponadto, że: $\mathsf{m}_k(\mathbf{z}_0) = 0, \; \mathsf{m}_k(\mathbf{z}_{m+1}) = 1, \; \mathsf{m}_k(\mathbf{z}_a) = 0{,}5;$ dla k = 1,2,3, gdzie $\mathbf{z}_a := (\mathbf{z}_0 + \mathbf{z}_{m+1}) / 2 = \mathbf{1} / \mathbf{2}$ jest obiektem „pośrednim", być może fikcyjnym.

Do „wypadkowej" analizy zróżnicowania poziomu rozwoju polskiego rolnictwa w latach 1998-2010 wybrano klasyfikator, będący średnią geometryczną trzech powyższych mierników m_1, m_2, m_3, który został obliczony według wzoru:

$$\mathrm{m}(\mathbf{z}) = \sqrt[3]{\mathrm{m}_1(\mathbf{z})\,\mathrm{m}_2(\mathbf{z})\,\mathrm{m}_3(\mathbf{z})}\;, \; \mathbf{z} \in [\mathbf{z}_0, \mathbf{z}_{\mathrm{m}+1}], \tag{5}$$

gdzie mierniki m_1, m_2, m_3 są określone są za pomocą wzorów (2), (3), (4), odpowiednio. Warto zauważyć, że miernik ten, jako średnia geometryczna funkcji znormalizowanych jest również znormalizowany.

WYNIKI BADAŃ

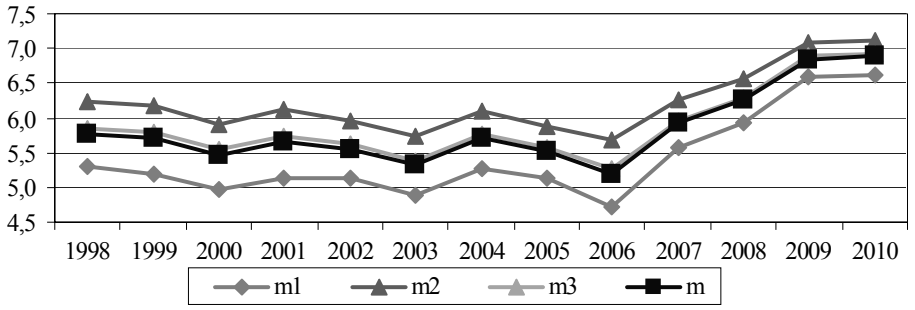
Niech $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{16} \in \Re_+^{13}$ oznaczają wektory, przyporządkowane alfabetycznie województwom ($\mathbf{w}_1$ – woj. dolnośląskie,..., $\mathbf{w}_{16}$ – woj. zachodnio-pomorskie), których składowymi są wartości miernika $\mathsf{m}(\mathbf{z})$ w poszczególnych latach okresu 1998-2010, gdzie miernik $\mathsf{m}(\mathbf{z})$ jest określony za pomocą wzoru (5). W tabeli 1 podano wyniki przeprowadzonych badań, w której wiersze są współrzędnymi wektorów $\mathbf{w_1}$, $\mathbf{w_2}$,...,$\mathbf{w_{16}}$. Natomiast na rysunku 1 zilustrowano wartości mierników $\mathsf{m}(\mathbf{z})$, $\mathsf{m}_1(\mathbf{z})$, $\mathsf{m}_2(\mathbf{z})$, $\mathsf{m}_3(\mathbf{z})$, dla Polski w latach 1998-2010.

Tabela 1. Wartości mierników m(z) dla województw w latach 1998-2010

	'98	**'99**	**'00**	**'01**	**'02**	**'03**	**'04**	**'05**	**'06**	**'07**	**'08**	**'09**	**'10**
Dolnośląskie	0,38	0,38	0,37	0,36	0,38	0,34	0,38	0,37	0,34	0,39	0,40	0,43	0,43
Kujaw.-pom.	0,42	0,41	0,38	0,40	0,39	0,38	0,41	0,39	0,38	0,41	0,44	0,48	0,48
Lubelskie	0,46	0,42	0,43	0,44	0,42	0,39	0,41	0,40	0,34	0,39	0,42	0,49	0,47
Lubuskie	0,23	0,25	0,23	0,24	0,22	**0,19**	0,28	0,24	0,22	0,27	0,27	0,30	0,34
Łódzkie	0,38	0,37	0,34	0,38	0,35	0,33	0,35	0,35	0,33	0,37	0,39	0,46	0,47
Małopolskie	0,33	0,32	0,32	0,32	0,33	0,31	0,31	0,30	0,29	0,33	0,33	0,37	0,39
Mazowieckie	0,40	0,39	0,37	0,39	0,37	0,35	0,38	0,39	0,37	0,42	0,46	0,53	0,51
Opolskie	0,45	0,43	0,44	0,44	0,43	0,41	0,45	0,44	0,42	0,47	0,49	0,49	0,51
Podlaskie	0,29	0,29	0,28	0,29	0,29	0,29	0,28	0,26	0,25	0,28	0,28	0,31	0,31
Podkarpackie	0,36	0,37	0,36	0,37	0,36	0,36	0,36	0,38	0,34	0,39	0,42	0,43	0,45
Pomorskie	0,34	0,34	0,32	0,32	0,33	0,31	0,33	0,33	0,31	0,35	0,38	0,41	0,40
Śląskie	0,27	0,26	0,24	0,25	0,25	0,24	0,27	0,25	0,24	0,32	0,36	0,38	0,36
Świętokrzyskie	0,35	0,35	0,32	0,35	0,33	0,31	0,33	0,31	0,27	0,30	0,32	0,38	0,38
Warm.-mazur.	0,36	0,36	0,34	0,35	0,35	0,37	0,36	0,36	0,35	0,39	0,41	0,43	0,43
Wielkopolskie	0,40	0,39	0,38	0,41	0,41	0,38	0,43	0,41	0,39	0,46	0,51	**0,57**	0,52
Zach.-pomors.	0,36	0,37	0,35	0,34	0,35	0,35	0,37	0,34	0,35	0,37	0,37	0,39	0,42

Źródło: obliczenia własne

Rysunek 1. Wartości mierników dla Polski w latach 1998-2010



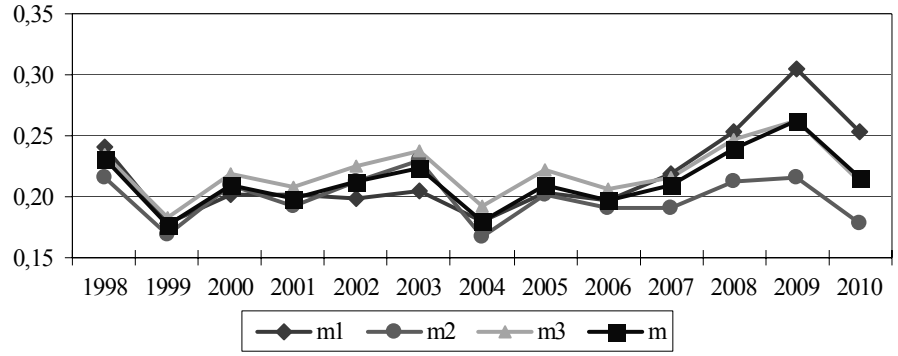
Źródło: opracowanie własne

Na podstawie przedstawionych wyników można zauważyć, że w Polsce w latach 1998-2010 występowały duże dysproporcje między województwami. Syntetyczne mierniki w badanym okresie przyjmowały wartości od 0,19 dla województwa lubuskiego w 2003 roku do 0,57 dla województwa wielkopolskiego w 2009 roku. Przeciętnie użyteczność $\bar{\mu}$ wszystkich województw w latach 1998-2010 wynosiła 0,36 $\left(\text{tj. } \bar{\mu} = \frac{1}{208}\sum_{i=1}^{208} f\left(\mathbf{z}_i\right) = 0,36\right)$. Wyniki pokazują, że w badanym okresie, według przyjętej metody najniżej oceniany jest poziom rozwoju rolnictwa w województwie lubuskim. Wartość syntetycznego miernika

tego województwa w badanym okresie wahała się w granicach 0,19 - 0,34 i była znacznie niższa od pozostałych województw. Na podstawie wyników zamieszczonych w tabeli 1 oraz rysunku 1 można zauważyć, że w latach 1998-2010 absolutny miernik rozwoju Polski, gdy w jednym roku rósł, to w następnych dwóch latach malał, natomiast w ostatnich czterech latach analizowanego okresu, tj. od roku 2006 w większości województw użyteczności znacząco wzrosły. Tendencję tę wykazują wszystkie mierniki bez względu na metodę obliczeń. Wzrost wskaźników w ostatnich latach oznacza większy poziom rozwoju rolnictwa, przeciętnie w Polsce.

Podany niżej rysunek 2 przedstawia wartości rozstępów tj. różnice między najwyższą a najniższą wartością mierników $m(\mathbf{z})$, $m_1(\mathbf{z})$, $m_2(\mathbf{z})$, $m_3(\mathbf{z})$ województw dla poszczególnych lat okresu 1998-2010.

Rysunek 2. Zróżnicowanie rozwoju rolnictwa w Polsce w latach 1998-2010



Źródło: opracowanie własne

Wyznaczone mierniki pozwoliły na dokonanie rankingu rozważanych obiektów. W tabeli 2 podano wyniki uporządkowania województw według wartości mierników syntetycznych. Analizując uporządkowanie województw według poziomu użyteczności można zauważyć, że na początku listy w badanym okresie, w rankingu znajdowało się województwo opolskie, następnie wielkopolskie i lubelskie. Końcowe lokaty zajęły natomiast województwa podkarpackie, śląskie i na ostatnim miejscu lubuskie.

Tabela 2. Uporządkowanie województw według $m(\mathbf{z})$

	'98	**'99**	**'00**	**'01**	**'02**	**'03**	**'04**	**'05**	**'06**	**'07**	**'08**	**'09**	**'10**
Dolnośląskie	6	6	6	8	5	9	5	7	8	7	8	9	9
Kujaw.-pom.	3	3	4	4	4	4	4	4	3	4	4	5	4
Lubelskie	1	2	2	1	2	2	3	3	9	6	5	3	5
Lubuskie	16	16	16	16	16	16	14	16	16	16	16	16	15
Łódzkie	7	8	9	6	9	10	10	9	10	10	9	6	6
Małopolskie	13	13	13	13	12	13	13	13	12	12	13	14	12

	‘98	‘99	‘00	‘01	‘02	‘03	‘04	‘05	‘06	‘07	‘08	‘09	‘10
Mazowieckie	5	5	5	5	6	7	6	5	4	3	3	2	3
Opolskie	2	1	1	2	1	1	1	1	1	1	2	4	2
Podlaskie	14	14	14	14	14	14	15	14	14	15	15	15	16
Podkarpackie	8	9	7	7	7	6	9	6	7	8	6	8	7
Pomorskie	12	12	12	12	11	12	12	11	11	11	10	10	11
Śląskie	15	15	15	15	15	15	16	15	15	13	12	12	14
Świętokrzyskie	11	11	11	9	13	11	11	12	13	14	14	13	13
Warm.-mazur.	9	10	10	10	8	5	8	8	5	5	7	7	8
Wielkopolskie	4	4	3	3	3	3	2	2	2	2	1	1	1
Zach.-pomors.	10	7	8	11	10	8	7	10	6	9	11	11	10

Źródło: obliczenia własne

Wykorzystując wyniki obliczeń zamieszczone w tabeli 1 dokonano podziału województw na 4 grupy, charakteryzujące się zbliżonym poziomem rozwoju rolnictwa. Podstawą podziału województw na klasy za cały badany okres były wartości średniego miernika rozwoju. Za średni miernik rozwoju $\bar{m}_i$, i-tego województwa, przyjęto: $\bar{m}_i = \frac{1}{13}\sum_{j=1}^{13} w_{ij}$, $i = 1,2,...,16$; gdzie w_{ij} oznacza j-tą składową wektora $\mathbf{w}_i$, np. w_{22} oznacza wartość miernika $m(\mathbf{z})$ dla województwa kujawsko-pomorskiego w roku 1999. Wykorzystano do tego szeregi przedziałowe rozdzielcze, w których rozpiętości przedziałów klasowych były równe w przybliżeniu jednej czwartej rozstępu za cały okres. Wyniki klasyfikacji województw dla badanego okresu przedstawiono w tabeli 3. Analizując zróżnicowanie regionalne rolnictwa w Polsce wyodrębniono cztery typologiczne grupy województw. Syntetyczne zestawienie wyników wskazuje, że Polska jest krajem wyraźnie zróżnicowanym ze względu na poziom rozwoju rolnictwa.

Tabela 3. Podział województw na grupy wg $m(\mathbf{z})$ w badanym okresie

Grupa	Poz.	Województwo	$\bar{m}_i$
I	1	Opolskie	0,45
	2	Wielkopolskie	0,44
	3	Lubelskie	0,42
II	4	Kujawsko-pomorskie	0,41
	5	Mazowieckie	0,41
	6	Dolnośląskie	0,38
	7	Podlaskie	0,38
	8	Łódzkie	0,37
	9	Warmińsko-mazurskie	0,37
III	10	Zachodnio-pomorskie	0,36
	11	Pomorskie	0,34
	12	Świętokrzyskie	0,33
	13	Małopolskie	0,33
IV	14	Podkarpackie	0,29
	15	Śląskie	0,28
	16	Lubuskie	0,25

Źródło: obliczenia własne

Do I grupy zaliczono 3 województwa (opolskie, wielkopolskie, lubelskie) odznaczające się najwyższym poziomem rozwoju rolnictwa. Trzeba jednak zauważyć, że województwo lubelskie, pomimo, że zostało zakwalifikowane ze względu na średnią wartość miernika syntetycznego w badanym okresie do grupy I, to znajduje się niejako na granicy I i II grupy (od 2005 roku znajdowało się w II grupie). Grupa II obejmuje 6 województw: kujawsko-pomorskie, mazowieckie, dolnośląskie, podlaskie, łódzkie, warmińsko-mazurskie. Charakteryzują się one dość wysokim poziomem rozwoju rolnictwa (użyteczności powyżej wartości przeciętnej, która wynosi 0,36). Grupa III obejmuje 4 województwa: zachodnio-pomorskie, pomorskie, świętokrzyskie i małopolskie. Województwo zachodnio-pomorskie można zaliczyć do regionów o średnim poziomie rozwoju rolnictwa, (mierniki syntetyczne dla tego województwa kształtują się na poziomie zbliżonym do przeciętnego), natomiast pozostałe regiony o niskim poziomie rozwoju rolnictwa, wartości mierników syntetycznych kształtują się poniżej poziomu przeciętnego. W IV grupie znalazły się trzy województwa: podkarpackie, śląskie i lubuskie, co wskazuje na to, że są to regiony o najniższym poziomie rozwoju rolnictwa w Polsce. W prawie każdym roku rozważanego okresu najniższy poziom rozwoju rolnictwa wykazało województwo lubuskie.

WNIOSKI

Przeprowadzone badania wykazały, że wskaźniki rozwoju rolnictwa Polski wyraźnie wzrosły w ostatnich czterech latach badanego okresu 1998 – 2010. Przyczyn takich zmian wybranego miernika syntetycznego można upatrywać w przystąpieniu Polski do Unii Europejskiej. Mimo tego znacznego wzrostu, poziom zróżnicowania rozwoju rolnictwa polskich województw nie maleje, co więcej wykazuje tendencje wzrostową – przy wykluczeniu wyników z roku 2010. Dla potwierdzenia tego zjawiska wskazana jest dalsza, wielowymiarowa analiza występujących zależności.

BIBLIOGRAFIA

Binderman A (2004) Przestrzenne zróżnicowanie potencjału rolnictwa w Polsce w latach 1989-1998, RNR, Seria G., T.91, Z.1, Wydawnictwo „Wieś Jutra", Warszawa, 51-60.

Binderman A. (2005a) Klasyfikacja polskich województw według poziomu rozwoju rolnictwa, RNR, Seria G., T.92, Z.1, Warszawa, 42-51.

Binderman A. (2005b) O problemie wyboru wzorca przy badaniu przestrzennego zróżnicowania potencjału rolnictwa w Polsce, MIBE V, Warszawa, str. 46-57.

Binderman A. (2006a) Wykorzystanie funkcji użyteczności do badania przestrzennego zróżnicowania rolnictwa, RNR SERiA, T.VIII, Z.5, Warszawa, s.5-11.

Binderman A. (2006b) Klasyfikacja obiektów oparta na dwóch wzorcach, EiOGŻ, Zeszyty Naukowe SGGW, nr 60, Warszawa, 25-34.

Binderman A. (2007) Wielowymiarowa analiza regionalnego zróżnicowania rolnictwa w Polsce, praca doktorska, SGGW, Warszawa.

Binderman A. (2008) Zastosowanie liniowej i nieliniowej funkcji użyteczności do badania poziomu rolnictwa w Polsce, MIBE IX, wyd. SGGW, Warszawa, str. 29-38.

Binderman A. (2009) Dynamika rozwoju rolnictwa w Polsce po akcesji do Unii Europejskiej, Roczniki Nauk Rolniczych, SERiA, Tom XI, Zeszyt 3.

Binderman A. (2009a) Zależność oceny zróżnicowania rolnictwa w Polsce od wybranych mierników syntetycznych, MIBE X, wyd. SGGW, Warszawa, 30-41.

Binderman A. (2010) Wpływ sposobu normalizacji zmiennych na ocenę regionalnego zróżnicowania rolnictwa, MIBE XI, wyd. SGGW, Warszawa, 25-38.

Binderman A. (2010a) Porównanie poziomu rozwoju rolnictwa województw w Polsce, RN SERIA, T XII, z.2, 29-34.

Binderman A. (2011) Wielokryterialne metody analizy zróżnicowania polskiego rolnictwa w 2009 roku, MIBE XII, 2, 58-68.

Borkowski B, Dudek H., Szczesny W. (2006) O pewnych problemach przekształcania wartości cech, ACTA AGRARIA ET SILVESTRIA, vol. XLVII

Cieślak M. (2004) Taksonomiczna procedura programowania rozwoju gospodarczego i określania zapotrzebowania na kadry kwalifikowane, Przegląd Statystyczny, z. 1.

Dąbkowski W. (2000) Regionalne struktury rolnictwa polskiego, IRWiR PAN, Warszawa.

Harasim A (2001) Wskaźniki oceny regionalnego zróżnicowania rolnictwa, Pamiętnik Puławski z. 124, Puławy.

Harasim A. (red.) (2006) Regionalne Zróżnicowanie Produkcji Rolniczej w Polsce, Studia I Raporty IUNG – PIB, 3.

Harasim A. (red.) (2009) Wybrane elementy regionalnego zróżnicowania rolnictwa w Polsce, Studia i Raporty IUNG – PIB, 15.

Heady E.O., Dillon J.L. (1961) Agricultural Production Functions, Iowa State University.

Heller J., Zawisza S. (red.) (2000) Problemy regionalizacji w procesie kształtowania obszarów wiejskich, Wyd. AT-R, Bydgoszcz.

Hellwig Z. (1981) Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych, [w:] Welfe W. (red.) Metody i modele ekon.-matemat. w doskonaleniu zarządzania gosp. socjalistyczną, PWE, Warszawa.

Hwang C. L., Yoon K. (1981) Multiple Attribute Decision Making, Methods and Applications, Springer – Verlag, New York.

Jacson D. M. (1971) The stability of classifications of binary attribute data, Technical Report 70-65, Cornell University, 1-13.

Krasowicz S., Kukuła S. (2006) Regionalne zróżnicowanie polskiego rolnictwa w świetle badań IUNG – PIB, Acta Agr. Silv., Sekcja Ekon., 46 (2): 73-89.

Kukuła K. (1989) Statystyczna analiza strukturalna i jej zastosowanie w sferze usług produkcyjnych dla rolnictwa, ZN, seria spec.: Monografie, nr 89, AE, Kraków 1989.

Kukuła K. (2000) Metoda unitaryzacji zerowanej, PWN, Warszawa.

Malina A. (2004) Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, AE, Seria Monografie nr 162, Kraków.

Młodak A. (2006) Analiza taksonomiczna w analizie regionalnej, Difin, Warszawa.

Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno - gospodarczych, PWE, Warszawa.

Poczta W., Bartkowiak N. (2012) Regionalne zróżnicowanie rolnictwa w Polsce, Journal of Agribusiness and Rural Development 1(23), 95-109.

Zeliaś A. (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, Akademia Ekonomiczna, Kraków.

REGIONAL DIFFERENTIATION OF POLISH AGRICULTURE IN THE YEARS 1998-2010

Abstract: The paper contains the multidimensional analysis of the regional differentiation of Polish agriculture in the years 1998-2010. The study is based on several synthetic measures - without patterns and which use two patterns of objects. On the basis of synthetic measures there were made arrangement and classification of voivodeships. The results show that despite the rapid growth in recent years of the level of agricultural development, the disparities between voivodships in Poland did not decrease.

Keywords: synthetic measures, utility functions, agriculture development level, classification, class division

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 65 – 78

PROBLEMY MODELOWANIA REZYGNACJI KLIENTÓW W TELEFONII KOMÓRKOWEJ

Krzysztof Gajowniczek, Tomasz Ząbkowski
Katedra Informatyki, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: krzysztof_gajowniczek@sggw.pl, tomasz_zabkowski@sggw.pl

Streszczenie: Przewidywanie i zarządzanie rezygnacjami klientów jest problemem wielu działalności gospodarczych, ale jest ono szczególnie dotkliwe w silnie konkurencyjnym sektorze telefonii komórkowej. Ze względu na wysokie koszty pozyskania nowych klientów i korzyści wynikające z utrzymania istniejących, istotną rolę w tego typu problemach odgrywają modele przewidujące rezygnację klientów. W tym kontekście autorzy artykułu, na podstawie danych empirycznych, zwracają uwagę na takie kwestie jak: (1) przygotowanie danych do analizy (2) problem doboru cech, (3) dobór odpowiednich technik modelowania wraz z oceną ich przydatności w kampaniach utrzymaniowych.

Słowa kluczowe: modelowanie rezygnacji, telefonia komórkowa, modele klasyfikacyjne

WPROWADZENIE

Działania prowadzone dotychczas przez operatorów telekomunikacyjnych, koncentrowały się, w głównej mierze, na dążeniu do zwiększenia udziału w rozwijającym się rynku przez przejmowanie klientów od konkurencji lub pozyskiwaniu zupełnie nowych klientów. Wraz ze wzrostem stopnia dojrzałości rynku przed przedsiębiorstwami telekomunikacyjnymi pojawiały się nowe wyzwania. Na rynkach dojrzałych podstawowym celem działania staje się utrzymanie istniejących klientów, zbudowanie z nimi trwałych relacji oraz pozyskanie ich lojalności i zaufania. W warunkach rosnącej konkurencji, stworzenie grupy stałych klientów wydaje się być podstawowym warunkiem poprawy zyskowności firmy i zachowania dotychczasowej pozycji rynkowej. Na obecnym etapie rozwoju polskich firm telekomunikacyjnych niezbędne staje się więc zachowanie właściwych proporcji między zdobywaniem nowych

a utrzymywaniem istniejących klientów. Zdobywanie nowych klientów jest niewątpliwie ważnym zadaniem. Istnieje jednak potrzeba przewidywania zmian zachodzących na rynku i planowania działań nakierowanych na utrzymanie wartościowych i lojalnych klientów.

Przewidywanie i zarządzanie rezygnacjami klientów jest problemem wielu branż, ale jest szczególnie dotkliwe w silnie konkurencyjnym i obecnie w znacznie zliberalizowanym sektorze telefonii komórkowej. Rezygnacja abonenta w branży telekomunikacyjnej odnosi się do przejścia klienta od jednego operatora telekomunikacyjnego do drugiego. Zjawisko takie w nomenklaturze branżowej określa się z języka angielskiego – churn. Wielu abonentów często rezygnuje z jednego usługodawcy na rzecz innego w poszukiwaniu lepszych stawek, usług lub korzyści wynikających z podpisania nowej umowy (np. otrzymywanie najnowszych aparatów). Ocenia się, że średni wskaźnik rezygnacji dla polskiego sektora telefonii komórkowej wynosi 2,2% miesięcznie, co oznacza blisko 26% utraconych klientów rocznie. Koszty marketingowe potrzebne do pozyskania nowego abonenta (SAC ang. subscriber acquisition cost) w amerykańskim sektorze telefonii komórkowej szacuje się na 300 do 600 dolarów [Wei i Chiu 2002], natomiast w przypadku polskiego sektora te koszty oscylują wokół 132 zł [Grupa Telekomunikacja Polska 2011]. Warto nadmienić, że koszt utrzymania istniejących abonentów jest blisko pięć razy niższy niż koszt ich pozyskania, oraz co istotne mają oni tendencję do generowania większych przychodów, ponieważ są mniej wrażliwi na cenę i często korzystają z dodatkowych usług.

Ze względu na wysokie koszty pozyskania nowych klientów i znacznie korzyści wynikające z utrzymania istniejących, istotną rolę odgrywają modele przewidujące rezygnację klientów. Dają one możliwość wcześniejszego wskazania klientów chcących zrezygnować z usług operatora, dzięki czemu realne staje się wdrożenie w odpowiednim momencie działań, mających na celu ograniczenie tego zjawiska, poprzez różnego rodzaju akcje utrzymaniowe i promocyjne.

W tym kontekście, autorzy na podstawie danych empirycznych [Neslin 2002], zwrócą uwagę na takie kwestie związane z modelowaniem zjawiska rezygnacji klientów jak: (1) problem doboru cech mogących determinować rezygnację klientów w branży telekomunikacyjnej, (2) problematykę odpowiedniego przygotowania danych do analizy, (3) dobór odpowiednich technik modelowania oraz budowę modeli wraz z oceną ich przydatności w kampaniach utrzymaniowych klientów operatora.

PROBLEM REZYGNACJI KLIENTÓW W LITERATURZE

W przypadku rynku usług, na którym znajduje się wielu klientów i wielu dostawców, zachodzi swoboda zawierania umów, w związku z czym, pojawia się możliwość migracji i nabywania usług gdzie indziej.

W praktyce, w usługach abonamentowych można wyróżnić dwa najczęściej stosowane typy rezygnacji [Lu 2002]. Pierwszym z nich jest rezygnacja

niezamierzona (ang. involuntary), wynikająca z przyczyn finansowych. W takim przypadku usługodawca rozwiązuje kontrakt z klientem zalegającym w płatnościach, w celu ograniczenia potencjalnych przyszłych start. W ramach rezygnacji finansowej, wyróżnia się również sytuację, w której konsument nie jest w stanie regulować dalszych zobowiązań pieniężnych [Burez i Van den Poel 2007], [Coussement i Van den Poel 2008]. Drugi typ rezygnacji związany jest z obniżoną satysfakcją z otrzymywanych usług oraz ze skutecznymi działaniami marketingowymi firm konkurencyjnych. Nad drugim typem rezygnacji, nazywanym również rezygnacją komercyjną, skupia się największą uwagę podczas badań nad tym zjawiskiem. Badania prowadzone nad rezygnacja finansową, są bardzo podobne do badań prowadzonych nad skoringiem kredytowym. W powyższej sytuacji, rezygnację finansową jest dużo łatwiej przewidzieć niż rezygnacje komercyjną, lecz jest znacznie trudniej jej przeciwdziałać [Burez i Van den Poel 2008].

Istnieją także alternatywne definicje zakładające, że rezygnacja klienta związana jest ze względnym obniżeniem poziomu usług, dzięki czemu, można przedstawić ją wykorzystując nachylenie krzywej obrazującej popyt klienta na dana usługę lub produkt. [Glady i in. 2009a, 2009b] zauważają, że powyższe terminy odejścia są mocno zależne od definicji i ceny pojedynczego produktu lub usługi, dlatego też proponują zdefiniować konsumenta zagrożonego rezygnacją jako tego, dla którego obniżeniu ulega poziom przyszłych wpływów CLV[1], a dokładniej rzecz ujmując, przez ujemne nachylenie krzywej CLV. [Hwang i in. 2004] idą o krok dalej sugerując, że model CLV powinien brać pod uwagę nie tylko potencjalną wartość przyszłych wpływów, lecz zarówno wartość przeszłych wpływów pieniężnych oraz prawdopodobieństwo rezygnacji z usług.

Poziom rezygnacji mierzony jest zazwyczaj jako odsetek bazy klientów, którzy opuścili danego operatora i podawany jest w odniesieniu do ustalonego okresu czasu, np. w skali miesięcznej lub rocznej. W segmencie telefonii komórkowej odsetek odejść, waha się w zależności od kraju i usługodawcy. W skali rocznej może on sięgać 25-40% w usługach abonamentowych oraz 70-80% w przypadku usług przedpłaconych [Kohs 2006].

Skala problemu odchodzenia klientów uzasadnia potrzebę jego precyzyjnej identyfikacji oraz działania z wyprzedzeniem. Badacze zjawiska podkreślają, że istotną rolę w całym procesie analizy odejść, odgrywa rodzaj danych i ich jakość. Wyróżnia się dane na temat klienta informujące o: demografii, finansach, czasie trwania kontraktu, dodatkowych usług czy wreszcie dane o połączeniach wychodzących i przychodzących [Ahn i in. 2006], [Kim i Yoon 2004], [Sulikowski 2008], [Van den Poel i Larivière 2004].

[1] CLV - (ang. Customer Lifetime Value), analiza mająca na celu ocenę aktualnej wartości przyszłych przepływów pieniężnych, związanych z relacjami z nabywcą, a więc ocenę wartości klienta w czasie.

Najczęściej w praktyce do modelowania rezygnacji klientów stosuje się techniki uczenia nadzorowanego jak np.: logit i probit [Ahn i in. 2006], [Burez i Van den Poel 2007], [Coussement i Van den Poel 2008], [Keramati i Ardabili 2011], [Madden i in. 1999], [Seo i in. 2008], które są rozwinięciem klasycznych metod regresyjnych, zaadaptowanych specjalnie na potrzeby klasyfikacji obiektów, do z góry ustalonej liczy klas. Z kolei [Huang i in. 2012], [Hung i in. 2006], [Karus i Dumas 2011], [Neslin i in. 2006], [Wei i Chiu 2002], stosowali drzewa decyzyjne, będące graficzną metodą wspomagania procesu decyzyjnego, stosowanego w teorii decyzji. Inni, m.in. [Daskalaki i in. 2006], [Huang i in. 2012], [Hung i in. 2006], [Karus i Dumas 2011], [Neslin i in. 2006], [Tsai i Lu 2009], [Waal i Toit 2008], wykorzystują sztuczne sieci neuronowe.

Do modelowania rezygnacji stosowano także metody grupowania modeli, takie jak random forest, bagging i boosting [Burez i Van den Poel 2007], [Burez i Van den Poel 2008], [Coussement i Van den Poel 2008]. Są to techniki, polegające na opracowaniu sekwencji prostych modeli, przy czym każdy kolejny model, przykłada większą wagę do tych obserwacji, które zostały błędnie zaklasyfikowane przez poprzednie modele.

Na koniec należy nadmienić, że skuteczność podawanych przez autorów modeli jest różna, zależy bowiem od przyjętych założeń, jakości posiadanych danych oraz definicji samej rezygnacji (inna dla klientów abonamentowych a inna dla klientów usług przedpłaconych).

OPIS DANYCH

Dane wykorzystane w pracy do modelowania rezygnacji stanowi zbiór „Cell2Cell: The Churn Game" [Neslin 2002], pozyskany z Centrum Zarządzania Relacjami z Klientami Uniwersytetu Duke'a, mieszczącego się w Karolinie Północnej w Stanach Zjednoczonych. Stanowią one reprezentatywny wycinek całej bazy danych, należącej do anonimowej firmy działającej w sektorze telefonii komórkowej w USA.

Dane zawierają 71047 obserwacji, gdzie każda obserwacja odpowiada indywidualnemu klientowi. Każdej obserwacji przypisanych jest 78 zmiennych, z czego 75 potencjalnych zmiennych objaśniających posłuży do późniejszej budowy modeli. Zestaw zmiennych objaśniających zawiera 34 zmienne ilościowe znajdujące się na skali ilorazowej, 38 zmienne jakościowe wyrażone liczbowo znajdujące się na skali dychotomicznej, 2 zmienne jakościowe wyrażone liczbowo znajdujące się na skali nominalnej oraz jedną zmienną jakościową opisaną za pomocą etykiet tekstowych. Wszystkie zmienne objaśniające pochodzą z tego samego okresu czasu, lecz modelowana binarna zmienna objaśniana (przyjmująca wartości 0 oraz 1) oznaczona jako „CHURN", została zaobserwowana w okresie od 31 do 60 dni, od momentu obserwacji pozostałych zmiennych. W zbiorze znajduje się dodatkowo zmienna „CALIBRAT", służąca do identyfikacji próby

uczącej i próby testowej, zawierających odpowiednio 40000 oraz 31047 obserwacji (por. Tab. 1).

Tablica 1. Liczebności próby uczącej i próby testowej

Dane Churn	Liczba obserwacji	Odsetek rezygnacji
Próba ucząca	40000	50%
Próba testowa	31047	1,96%

Źródło: opracowanie własne

Próba ucząca zawiera 20000 przypadków zaklasyfikowanych jako osoby rezygnujące (oznaczone jako 1) oraz 20000 przypadków zaklasyfikowanych jako osoby nierezygnujące (oznaczone jako 0). Wyżej opisany podział zbioru uczącego, w którym odsetek klasy wyróżnionej (osoby rezygnujące) jest taki sam jak odsetek klasy niewyróżnionej, jest mało realistyczny, natomiast prowadzi do ominięcia problemu niezrównoważenia klas [Napierała i Stefanowski 2011], prowadzącego do błędnej specyfikacji modelu oraz słabej generalizacji wiedzy.

W próbie testowej, na której będzie sprawdzana jakość zbudowanego modelu, znajduje się jedynie 1,96% osób rezygnujących. Z tak małym odsetkiem klasy wyróżnionej (oznaczony cyfrą 1), bardzo często można spotkać się w praktyce biznesowej podczas budowy modeli klasyfikacyjnych. Ponadto, wymieniona wartość odsetka, odpowiada rzeczywistej wartości wskaźnika rezygnacji w firmie, z której pochodzą dane.

PRZYGOTOWANIE DANYCH DO ANALIZY

Większość surowych dostępnych w bazach czy hurtowaniach danych jest w postaci danych transakcyjnych, które należy przygotować, aby móc je wykorzystać do analizy. Częstym problem w danych może być również niekompletność (braki danych). Baza danych może również zawierać obserwacje, które są przestarzałe lub zbędne, dane znajdujące się w nieodpowiednim formacie dla modeli czy wreszcie wartości, które są niezgodne z zasadami czy zdrowym rozsądkiem. Aby dane były przydatne do celów eksploracji danych muszą przejść przez wstępną analizę. Nadrzędnym celem staje się minimalizacja GIGO[2] czyli minimalizacja tzw. śmieciowych danych dostających się do późniejszej budowy modelu. Szacuje się [Larose 2006], że sam etap wstępnej obróbki danych może zajmować nawet 60% czasu poświęconego na cały proces eksploracji danych.

W przypadku dostępnego zbioru danych problem dotyczył jedynie braków danych. W bazie danych znajduje się 10 zmiennych, w których takowe braki występują. Tab. 2 przedstawia liczebność braków danych dla poszczególnych zmiennych. Zmienne te dotyczą połączeń głosowych oraz wieku klientów.

[2] GIGO (ang. garbage in – garbage out) – wprowadzisz błędne dane a uzyskasz błędne wyniki.

Tablica 2. Liczebność braków danych dla poszczególnych zmiennych

Cecha	Opis	Braki danych
REVENUE	Średni miesięczny przychód	216
MOU	Średnia liczba minut w miesiącu	216
RECCHRGE	Średnia wartość miesięcznych zobowiązań	216
DIRECTAS	Średnia liczba połączeń bezpośrednich w miesiącu	216
OVERAGE	Średnia liczba minut ponad plan w miesiącu	216
ROAM	Średnia liczba połączeń roamingowych	216
CHANGEM	Procentowa zmiana wykorzystania minut w danych okresach	502
CHANGER	Procentowa zmiana w przychodach w danych okresach	502
AGE1	Wiek pierwszego członka rodziny	1244
AGE2	Wiek kolejnego członka rodziny	1244

Źródło: opracowanie własne

W rekordach z brakami danych ze zbioru uczącego, znajduje się aż 70% klasy wyróżnionej (klientów rezygnujących), podczas gdy stosunek osób rezygnujących i nierezygnujących w całym zbiorze uczącym wynosi 50%. Usunięcie obserwacji z brakami danych w takiej sytuacji mogłoby doprowadzić do zaburzenia wzorców znajdujących się w danych. W związku z tym, w dalszej części zaproponowana została technika uzupełniania tychże braków danych.

Do najczęstszych sposobów radzenia sobie z brakującymi danymi należą: zastąpienie wartości brakującej pewną z góry ustaloną wartością stałą lub zastąpienie ich wartością średnią (dla zmiennych liczbowych) bądź dominantą (dla zmiennych jakościowych). Takie postępowanie nie zawsze jest właściwe, ponieważ w ten sposób mocno ingerujemy w rozkład rzeczywisty danej cechy. Panaceum w takiej sytuacji może okazać się proces generowania liczb losowych z zaobserwowanego rozkładu prawdopodobieństwa danej zmiennej.

Rozkład prawdopodobieństwa dla tej samej cechy w obrębie obydwu klas (rezygnujący vs lojalni klienci) może się diametralnie różnić, w związku z czym, przeprowadzono test do weryfikacji hipotezy o równości wartości średnich badanej cechy w dwóch klasach, który ma postać:

$$u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{1}$$

gdzie $\bar{x}_1, \bar{x}_2$ są wartościami średnimi obydwu klas, s_1^2, s_2^2 są wariancjami badanych klas oraz n_1, n_2 są licznościami poszczególnych klas. Dla dużej próby statystyka u ma standaryzowany rozkład normalny $N(0,1)$. Wyniki przeprowadzonego testu, w którym zbiór danych na dwie rozłączne podpróby rozdzielała zmienna objaśniana „CHURN", zostały przedstawione w poniższej tabeli (Tab. 3).

Tablica 3. Wyniki testu na równość średnich w dwóch grupach

	Grupująca: CHURN, Grupa 1: 0, Grupa 2: 1			
Zmienna	Średnia 0	Średnia 1	u	p-value
REVENUE	59,2187	57,9548	3,44835	0,000564
MOU	543,2069	482,8304	13,76491	0,000000
RECCHRGE	47,7824	44,6525	15,82461	0,000000
DIRECTAS	0,9180	0,8351	4,55609	0,000005
OVERAGE	39,1729	42,3594	-3,99238	0,000065
ROAM	1,1506	1,3940	-3,23494	0,001217
CHANGEM	-5,3443	-24,4463	9,00406	0,000000
CHANGER	-1,4102	-0,7000	-2,20327	0,027579
AGE1	43,6244	43,4101	1,348094	0,089932
AGE2	44,3017	44,0885	1,692853	0,090786

Źródło: opracowanie własne

Na poziomie istotności $\alpha = 0{,}05$ we wszystkich przypadkach odrzucamy hipotezę zerową mówiącą o tym, że wartości średnie w obydwu klasach są sobie równe. Jedynie w przypadku zmiennych „AGE1" i „AGE2" brak jest podstaw do odrzucenia hipotezy zerowej mówiącej, że wartości średnie w obrębie obydwu klas nie różnią się od siebie.

Proces przygotowania danych do analizy zakończony został etapem generowania danych z zaobserwowanych wcześniej rozkładów i ich uzupełnieniem do zbioru danych. Generowanie liczb losowych w tych przypadkach zostało przeprowadzone za pomocą modułu „Metody symulacji" w pakiecie Statistica. Symulacja danych dla zmiennych „AGE1" i „AGE2", została wykonana za pomocą metody Monte Carlo [Brandt 1998], [Drew i Homem 2012], która jest stosowana do modelowania matematycznego złożonych procesów, tak by można było przewidzieć ich wyniki za pomocą podejścia analitycznego.

W przypadku zmiennych informujących o wskaźnikach na temat połączeń głosowych, należy uwzględnić fakt, że występuje pomiędzy nimi ścisła współzależność funkcyjna. Związek ten został zmierzony za pomocą współczynnika korelacji rang Spearmana. W świetle tych wyników do symulacji braków danych w tych zmiennych została wykorzystana metoda uwzględniająca zależności funkcyjne, gdyż w przeciwny razie uzyskane wyniki mogłyby stać ze sobą w sprzeczności. W tym celu do losowania liczb z zachowaniem współzależności funkcyjnej , została wykorzystana Metoda Imana-Conovera [Haas 1999], wykorzystująca fakt istnienia związku między zmiennymi, mierzonego za pomocą współczynnika korelacji rang Spearmana.

DOBÓR ZMIENNYCH DO MODELU

W modelach opracowywanych w ramach koncepcji data mining jednym z zadań jest dobór zmiennych objaśniających spośród listy kandydatów. Dla

przykładu, budowanie zmiennych może być procesem zautomatyzowanym, w efekcie czego powstaje duża liczba cech sięgająca tysięcy lub nawet setek tysięcy zmiennych. Niestety, wykorzystywane metody analityczne, w obliczu tak dużych wolumenów danych mogą stać się bezużyteczne.

Dobór zmiennych objaśniających za pomocą testu χ^2 Pearsona

W niniejszej pracy dobór zmiennych został oparty m.in. o nieparametryczną statystykę zgodności χ^2 Pearsona oraz wartość p-value dla każdej zmiennej niezależnej [Lu 2002]:

$$\chi_d^2 = \sum_1^k \frac{(n_i - np_i)^2}{np_i} \tag{2}$$

gdzie n_i są doświadczalnymi liczebnościami poszczególnych przedziałów, natomiast p_i są prawdopodobieństwami przynależności do tych przedziałów. W przypadku zmiennych ciągłych zakres wartości zmiennej dzielony jest na k przedziałów, natomiast zmienne jakościowe nie są przekształcane. Trzeba ponadto zaznaczyć, iż wyników tej metody należy traktować, jako pewne rozwiązane heurystyczne będące pomocnym do wyodrębnienia tych zmiennych, które można uwzględnić na etapie budowy modeli.

W kolejnym kroku, za pomocą analizy macierzy współczynników korelacji dokonano wyboru takich zmiennych objaśniających, które są silnie skorelowane ze zmienną objaśnianą i jednocześnie słabo skorelowane między sobą. Pozwala to ominąć często występujący w analizach problem współliniowości zmiennych objaśniających [Maddala 2006]. Wobec tego, ze zbioru danych objaśniających zostały wybrane wszystkie te pary zmiennych, dla których współczynnik korelacji wynosił powyżej 0.7, aby ostatecznie wyeliminować tę zmienną, która charakteryzowała się mniejszą wartością statystyki χ^2 obliczoną powyżej. Ostatecznie z całego zbioru danych pozostało 38 zmiennych objaśniających, w których 22 zmienne ma charakter ilościowy i 16 zmiennych ma charakter jakościowy.

Dobór zmiennych objaśniających za pomocą analizy składowych głównych

Na tym etapie wykorzystano również analizę składowych głównych (PCA), która jest techniką redukcji wymiaru zaliczaną do technik uczących się bez nadzoru. Celem analizy jest poszukiwanie takiego zbioru zmiennych, mniej licznego od zbioru zmiennych oryginalnych, dla którego można z pewnym, ale możliwie najmniejszym błędem, odtworzyć wartości zmiennych oryginalnych. Aby taka redukcja była możliwa, między oryginalnymi zmiennymi muszą zachodzić zależności statystyczne. Metoda ta przekształca oryginalne, skorelowane zmienne na nowe nieskorelowane zmienne, tzw. składowe główne, które

wyjaśniają w maksymalnym stopniu całkowitą wariancję z próby zmiennych pierwotnych.

Za pomocą tej analizy, do dalszego etapu budowy modeli wybrano 6 zmiennych, utworzonych poprzez poszczególne składowe główne, które wyjaśniały całkowitą zmienność w 97,38%.

Podejście eksperckie do doboru zmiennych objaśniających

Jako trzecią strategię doboru zmiennych objaśniających, zastosowano subiektywny ekspercki wybór zmiennych do modelu. W tym celu, spośród wszystkich zmiennych wybrano dziesięć najlepszych predyktorów ilościowych, charakteryzujących się największą wartością statystyki χ^2.

MODELOWANIE REZYGNACJI

Do modelowania rezygnacji klientów wykorzystanych został szereg technik. W szczególności, celem rozwiązania postawionego problemu klasyfikacyjnego zastosowano sztuczne sieci neuronowe, drzewa klasyfikacyjne, drzewa klasyfikacyjne ze wzmacnianiem, regresję logistyczną oraz analizę dyskryminacyjną.

Do oceny modeli wykorzystywane były dwie miary: wykres przyrostu oraz pole pod krzywą ROC.

Wykres przyrostu (ang. lift chart) jest graficznym sposobem podsumowania użyteczności modeli do przewidywania wartości zmiennej objaśnianej przyjmującej dwie wartości, natomiast w przypadku, gdy modelowana zmienna przyjmuje więcej wartości, można stworzyć wykres przyrostu i wykres zysku oddzielnie dla każdej z klas. Jak pokazuje [Larose 2006], przyrost można zdefiniować następująco:

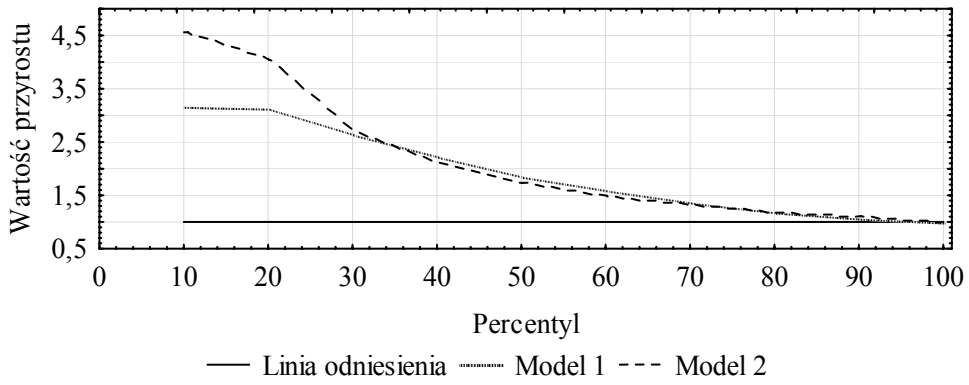
$$przyrost = \frac{procent\ \ pozytywnych\ trafień\ w\ zbiorze\ \ pozytywnych\ klasyfikacji}{procent\ \ pozytywnych\ trafień\ w\ całym\ zbiorze\ danych} \quad (3)$$

Podczas obliczania przyrostu najpierw sortuje się przypadki według prawdopodobieństwa przynależności do danej klasy. Następnie wartości przyrostu oblicza się dla kolejnych percentyli całej zbiorowości tzn. dla 10%, 20% itd. przypadków z największym przewidywanym prawdopodobieństwem przynależności do danej klasy. Uzyskane w ten sposób punkty łączy się linią, która zazwyczaj wolno spada, aż do osiągnięcia wartości 1,0 przy wyborze wszystkich przypadków. Dodatkowo na wykres nanosi się linię odniesienia odpowiadającą losowemu wyborowi 10%, 20% itd. przypadków ze zbioru danych.

Końcowa postać wykresu przedstawiona jest na Rys. 1, na którym widać, że największy przyrost osiągnięty jest dla najmniejszych percentyli. Z wykresu tego typu można odczytać, iż biorąc 10% przypadków (oś x) najpewniej zaklasyfikowanych do odpowiedniej klasy, otrzymamy próbkę, w której co najmniej 4,5 razy więcej przypadków (oś y) należy do wybranej klasy. Innymi

słowy, zastosowanie modelu zwiększa prawdopodobieństwo dotarcia do pożądanej klasy w stosunku do losowego wyboru obserwacji.

Rysunek 1. Graficzne porównanie wartości przyrostu wybranych modeli



Źródło: opracowanie własne

W niniejszej pracy, do oceny modeli klasyfikacyjnych wykorzystywana jest wartość przyrostu dla pierwszego percentyla bazy danych klientów. Jest to podyktowane względami ekonomicznymi, gdyż operator telekomunikacyjny nie kieruje kampanii utrzymaniowej do szerokiej bazy klientów, lecz skupia się na niewielkim odsetku np. 1 – 2% bazy klientów w skali miesiąca, o największym prawdopodobieństwie rezygnacji z usług. Przykładowo, wobec całkowitej liczby klientów u danego operatora sięgającej ok. 10 mln, 1% stanowi grupę ok. 100 tyś. klientów w skali miesiąca, którym zostanie przedstawiona oferta utrzymaniowa.

Jeśli chodzi o ogólną jakość klasyfikatora, konstruuje się miarę (ściśle związaną z indeksem Giniego), będącą polem powierzchni pomiędzy krzywą modelu a osią poziomą (AUC ang. Area Under Curve), którą można w przybliżeniu obliczyć poprzez całkowanie numeryczne metodą trapezów [Fortuna i in. 2002]:

$$AUC \approx \frac{2 - \sum_{i=1}^{n} \left[(Fx_{i+1} + Fx_i)(Lx_{i+1} - Lx_i) \right]}{2} \tag{4}$$

gdzie Fx_i są wartościami prostej utworzonej poprzez dystrybuantę linii odniesienia, natomiast Lx_i są wartościami krzywej utworzonej z wartości zysku obliczonego dla każdego percentyla. Ponadto $AUC \in \langle 0.5,1 \rangle$ oraz $indeks\ Giniego = 2AUC - 1$.

W wyniku przeprowadzonych eksperymentów numerycznych okazało się, że najlepszą metodą doboru zmiennych objaśniających do modelu opisującego

rezygnacje klientów, okazała się metodą χ^2 połączona z późniejszą eliminacją skorelowanych zmiennych w celu uniknięcia współliniowości. Modele charakteryzowały się nie tylko dużym przyrostem dla pierwszego percentyla, ale w szczególności znacznie większą jakością klasyfikacji na całym zbiorze testowym mierzoną polem pod krzywą ROC. Zachowanie jak największej ilości odpowiednich zmiennych pozwala zawrzeć w modelu nawet te najdrobniejsze wzorce i nie traci się cennych informacji zawartych w danych. Poza tym nie są one w żaden sposób przekształcone tak jak w przypadku metody składowych głównych, która okazała się metodą najgorszą.

Najlepszymi modelami w grupach zmiennych w $3/3$ przypadków (przez przypadki rozumiane są trzy różne metody selekcji zmiennych), okazały się algorytmy wzmacniania drzew klasyfikacyjnych, gdzie dla najlepszego modelu przyrost wyniósł 3,93 oraz pole $AUC_{Boosting} = 0{,}6600$. Na kolejnym miejscu uplasowały się modele drzew klasyfikacyjnych C&RT. Następnie, bardzo podobne wyniki uzyskały modele sieci neuronowych i modele dyskryminacyjne, natomiast najgorsze wyniki dawał model regresji logistycznej (por. Tab. 4).

Tablica 4. Zbiorcze wyniki dla opracowanych modeli

Lp.	Model (nazwa)	Przyrost w 1	Pole AUC
Dobór zmiennych za pomocą testu χ^2 i eliminacji skorelowanych zmiennych			
1	Sieci neuronowe (SANN)	2,29	0,6217
2	Drzewa klasyfikacyjne (C&RT)	3,11	0,6204
3	Wzmacniane drzewa klasyfikacyjne (Boosting)	3,93	0,6600
4	Regresja logistyczna (Logit)	2,13	0,6165
5	Analiza dyskryminacyjna (GDA)	2,78	0,6166
Dobór 10 zmiennych ilościowych za pomocą podejścia eksperckiego			
6	Sieci neuronowe (SANN)	2,62	0,6140
7	Drzewa klasyfikacyjne (C&RT)	2,78	0,6192
8	Wzmacniane drzewa klasyfikacyjne (Boosting)	3,93	0,6352
9	Regresja logistyczna (Logit)	1,96	0,5877
10	Analiza dyskryminacyjna (GDA)	2,13	0,5879
Dobór zmiennych za pomocą Analizy Składowych Głównych			
11	Sieci neuronowe (SANN)	2,45	0,6078
12	Drzewa klasyfikacyjne (C&RT)	1,63	0,5930
13	Wzmacniane drzewa klasyfikacyjne (Boosting)	3,44	0,6083
14	Regresja logistyczna (Logit)	1,63	0,5868
15	Analiza dyskryminacyjna (GDA)	1,63	0,5866

Źródło: opracowanie własne

Na podstawie przeprowadzonych analiz zaobserwowano, że na rezygnacje klienta najbardziej wpływają następujące zmienne: „EQPDAYS" – im dłuższy czas posiadania bieżącej taryfy tym większe ryzyko odejścia, „MONTHS" – im dłuższy czas pobytu w sieci tym większe prawdopodobieństwo rezygnacji. „RETCALS" – większa ilość połączeń do sekcji utrzymania klienta wpływa na rezygnację, „MOU" i "OVERAGE" – mniejsze średnie i całościowe miesięczne wykorzystywanie minut sprzyja rezygnacji oraz zmienna „UNIQSUBS" mówiąca o tym, że klient posiadający mniej unikalnych usług częściej rezygnuje.

PODSUMOWANIE

Celem pracy było przedstawienie możliwości modeli klasyfikacyjnych w problemie rezygnacji klientów telefonii komórkowej.

W tym kontekście autorzy artykułu, na podstawie danych empirycznych zwrócili uwagę na takie kwestie jak: (1) problem doboru cech mogących determinować rezygnację klientów w branży telekomunikacyjnej, (2) problematykę odpowiedniego przygotowania danych do analizy (3) dobór odpowiednich technik modelowania oraz budowę modeli wraz z oceną ich przydatności w kampaniach utrzymaniowych klientów operatora.

W szczególności badania wykazały, że:

(i) Najlepszą techniką doboru zmiennych skutkującą największą wartością przyrostu dla pierwszego percentyla oraz polem powierzchni AUC, była metoda oparta o statystykę χ^2 i eliminację skorelowanych zmiennych;

(ii) Algorytm wzmacniania drzew klasyfikacyjnych wykazywał największą trafność klasyfikacji (wyrażoną przyrostem i polem pod krzywą ROC) spośród testowanych metod;

(iii) Modele przewidujące rezygnację klientów, zbudowane za pomocą przedstawionych technik są dobrym narzędziem do osiągnięcia celów strategicznych firmy, nakierowanych na przeciwdziałanie zjawisku rezygnacji.

BIBLIOGRAFIA

Ahn J-H., Han S-P., Lee Y-S. (2006) Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry, Telecommunications policy 30, str. 552-568.

Brandt S. (1998) Analiza danych. Metody statystyczne i obliczeniowe, Wydawnictwo Naukowe PWN, Warszawa.

Burez J., Van den Poel D. (2007) CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services, Expert Systems with Applications 32(2), str. 277-288.

Burez J., Van den Poel D. (2008) Separating financial from commercial customer churn: A modeling step towards resolving the conflict between the sales and credit department, Expert Systems with Applications 35 str. 497-514.

Coussement K., Van den Poel D. (2008) Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques, Expert Systems with Applications 34(1), str. 313-327.

Daskalaki S., Kopanas I., Avouris N. (2006) Evaluation of classifiers for an uneven class distribution problem, Applied Artificial Intelligence Vol. 20 (5), str. 381-417.

Drew S., Homem T. (2012) Some Large Deviations Results for Latin Hypercube Sampling, Methodology and Computing in Applied Probability 14 (2), str. 203-232.

Fortuna Z., Macukow B., Wąski J. (2002) Metody numeryczne, Wydawnictwo Naukowo-Techniczne, Warszawa.

Glady N., Baesens B., Croux C. (2009a) Modeling churn using customer lifetime value, European Journal of Operational Research 197, str. 402-411.

Glady N., Baesens B., Croux C. (2009b) A modified Pareto/NBD approach for predicting customer lifetime value, Expert Systems with Applications 36, str. 2062-2071.

Grupa Telekomunikacja Polska. (2011) Sprawozdanie Zarządu z działalności Grupy Kapitałowej Telekomunikacja Polska w pierwszym półroczu 2011 roku,. str. 19-21, Pozyskano z : http://www.telix.pl/images/sprawozdania/TP_Grupa_2q2011.pdf.

Haas C. (1999) On Modeling Correlated Random Variables in Risk Assessment, Risk Analysis 19 (6), str. 1205-1214.

Huang B., Kechadi M., Buckley B. (2012) Customer churn prediction in telecommunications, Expert Systems with Applications 39, str. 1414-1425.

Hung S.Y., Yen D.C., Wang H.Y. (2006) Applying data mining to telecom churn management, Expert Systems with Applications 31, str. 515–524.

Hwang H., Jung T., Suh E. (2004) An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, Expert System with Applications 26, str. 181-188.

Karus S., Dumas M. (2011) Predicting the maintainability of XSL transformations, Science of Computer Programming 76, str. 1161-1176.

Keramati A., Ardabili S. (2011) Churn analysis for an Iranian mobile operator, Telecommunications Policy 35, str. 344-356.

Kim H., Yoon C. (2004) Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market, Telecommunications Policy 28, str. 751–765.

Kohs G. (2006) Comparison of Churn Rates, Inside Market Research, June 2006, Pozyskano z : http://insidemr.blogspot.com/2006/06/comparison-of-churn-rates.html.

Larose D.T. (2006) Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych, Wydawnictwo Naukowe PWN, Warszawa.

Lu J. (2002) Predicting customer churn in the telecommunications industry – An applications of survival analysis modeling using SAS, Proceedings of SUGI 27, Orlando, Florida, Paper 114.

Maddala G. S. (2006) Ekonometria, Wydawnictwo Naukowe PWN, Warszawa.

Madden G., Savage S., Coble-Neal G. (1999) Subscriber churn in the Australian ISP market, Information Economics and Policy 11, str. 195-207.

Napierała K., Stefanowski J. (2011) BRACID: a comprehensive approach to learning rules from imbalanced data, Journal of Intelligent Information Systems, Pozyskano z : http:// www.springerlink.com/content/d484131415313k17/fulltext.pdf .

Neslin S. (2002) Cell2Cell: The churn game. Cell2Cell Case Notes. Hanover, NH: Tuck School of Business, Dartmoth College.

Neslin S., Gupta S., Kamakura W., Lu J., Mason C. (2006) Defection detection: Measuring and understanding the predictive accuracy of customer churn models, Journal of Marketing Research 43(2), str. 204-211.

Seo D., Ranganathan C., Babad Y. (2008) Two-level model of customer retention in the US mobile telecommunications service market, Telecommunications Policy, Volume 32, Issue 3-4, str. 182-196.

Sulikowski P. (2008) Mobile Operator Customer Classification in Churn Analysis, Proceedings of the SAS Global Forum Conference, San Antonio, Texas, paper 344.

Tsai C-F., Lu Y-H. (2009) Customer churn prediction by hybrid neural networks, Expert Systems with Applications 36, str. 12547-12553.

Waal de D., Toit du J. (2008) Gaining Insight into Customer Churn Prediction using Generalized Additive Neural Networks, Proceedings of SATNAC - South Africa Telecommunication Networks and Applications.

Wei C., Chiu I-T. (2002) Turning telecommunications call details to churn prediction: a data mining approach, Expert Systems with Applications 23, str.103-112.

PROBLEMS OF CHURN MODELLING AT CELLULAR TELECOMMUNICATION

Abstract: Managing of customer churn is a serious problem at many businesses but is particularly important in the highly competitive and liberalized the cellular telecommunication sector. Due to the high costs of acquiring new customers and significant benefits of keeping existing ones, the predictive models for churn classification play an important role in this business. In this context, based on empirical data, the authors will tackle such issues as: (1) the problem of variable selection to determine the churn, (2) the problem of data preparation for the analysis; (3) selection of appropriate modeling techniques and the construction of models with their evaluation to support the retention campaigns.

Keywords: churn modelling, cellular telecommunication, classification models

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 79 – 88

KSZTAŁTOWANIE SIĘ KURSU KORONY SZWEDZKIEJ WOBEC PODSTAWOWYCH WALUT ŚWIATOWYCH

Stanisław Gędek
Katedra Ekonomii, Politechnika Rzeszowska
e-mail: gedeks@prz.edu.pl

Streszczenie: Celem pracy była analiza zachowania kursu korony szwedzkiej wobec walut światowych. Jako narzędzie analizy wykorzystany został model VAR. Otrzymane wyniki wskazują, że w badanym okresie korona szwedzka pozostawała pod wpływem euro, jednakże wpływ ten stawał się coraz słabszy. Zmienność kursów dolara w znacznie mniejszym stopniu wpływała na kursy korony szwedzkiej.

Słowa kluczowe: kursy walutowe, korona szwedzka, kointegracja, model VAR, funkcja odpowiedzi na impuls

WPROWADZENIE

Światowy rynek walutowy jest największym i najszybciej rosnącym rynkiem. Ocenia się, że dzienne obroty na tym rynku osiągały w roku 2010 niemal 4 bilony dolarów, podczas gdy w roku 2001 obroty na tym rynku wynosiły jedynie około 1,2 biliona dolarów[1]. Dla porównania produkt krajowy brutto Stanów Zjednoczonych w roku 2010 wynosił 14,7 biliona dolarów. Rynek ten funkcjonuje praktycznie 24 godziny na dobę , przez 7 dni w tygodniu.

W czasie obowiązywania systemu z Bretton Woods, od drugiej wojny światowej do wczesnych lat 70 XX wieku, wahania kursów walutowych były bardzo małe. Wprowadzenie systemu płynnych kursów walutowych spowodowało znaczne zwiększenie zakresu wahań, przekraczające bardzo często 10% w ciągu dnia i 100% w ciągu roku [Karras i in. 2005, Stockman 1988].

Zmienność kursów walutowych jest przedmiotem intensywnych badań. Zazwyczaj jednak przyjmuje się, że jest to czynnik zewnętrzny determinujący

[1] *Report on global foreign exchange market activity in 2010*. Monetary and Economic Department, Bank for International Settlements 2011.

zachowanie innych zmiennych makroekonomicznych [por. np. Benita i Lauterbach 2007, Froot i Stein 1991, McKenzie 1999, Nucci i Pozzolo 2001]. W podejmowanych rzadziej badaniach mających wyjaśnić przyczyny zmienności kursów walutowych przyjmuje się najczęściej, w zasadzie *implicite*, założenie Friedmana [Friedman 1953], iż zmienność kursu walutowego w systemie kursów płynnych jest odzwierciedleniem zmienności parametrów charakteryzujących gospodarkę danego kraju [Edwards 2006, Lanne i Vesala 2006, Tarashev 2007], lub stosowanej polityki gospodarczej [Beine i in. 2006, Devereux i Engel 2003, Galí, i Monacelli 2005, Reinhart i Rogoff 2004]. Zagadnienie wzajemnego wpływu kursów walutowych podejmowane było jeszcze rzadziej [Bednarz i Gędek 2007, Witkowska 2011]. Zgodnie z teorią optymalnych obszarów walutowych [Mussa 2002] natężenie sił stabilizujących walutę danego kraju jest odwrotnie proporcjonalne do wielkości jego gospodarki. Stąd należy przypuszczać, że kursy walut tych krajów Unii Europejskiej, które nie wprowadziły wspólnej waluty, będą podlegały procesowi konwergencji względem euro. Procesowi takiemu podlegał na przykład kursy walut Polski, Węgier i Czech [Witkowska 2011], ale nie Wielkiej Brytanii [Bednarz i Gędek 2007].

Przedmiotem niniejszego opracowania jest analiza zachowania się kursu korony szwedzkiej wobec podstawowych walut światowych. Jak dowodzą badania [Bask i de Luna 2005] utrzymanie przez Szwecję i Danię walut narodowych i powstrzymanie się od bliższej współpracy z Europejskim Bankiem Centralnym nie spowodowało zasadniczych negatywnych skutków dla gospodarek tych krajów. Kryzys z przełomu lat 2008/9 wstrząsnął gospodarką szwedzką znacznie mocniej niż innymi krajami UE, jednakże gospodarka szwedzka wyszła z kryzysu znacznie szybciej niż inne kraje [2]. Gospodarka szwedzka jest dosyć silnie powiązana z gospodarką światową (export wynosi ok. 30% PKB), zarówno z UE, jak i pozostałymi obszarami gospodarczymi [Ketels 2012]. Czyni to Szwecję i jej walutę interesującym przedmiotem badań. Punktem wyjścia w badaniach przeprowadzonych w niniejszej pracy jest hipoteza, iż na kurs korony szwedzkiej do walut światowych wpływa kurs tych walut do euro i dolara. Celem opracowania jest weryfikacja tej hipotezy.

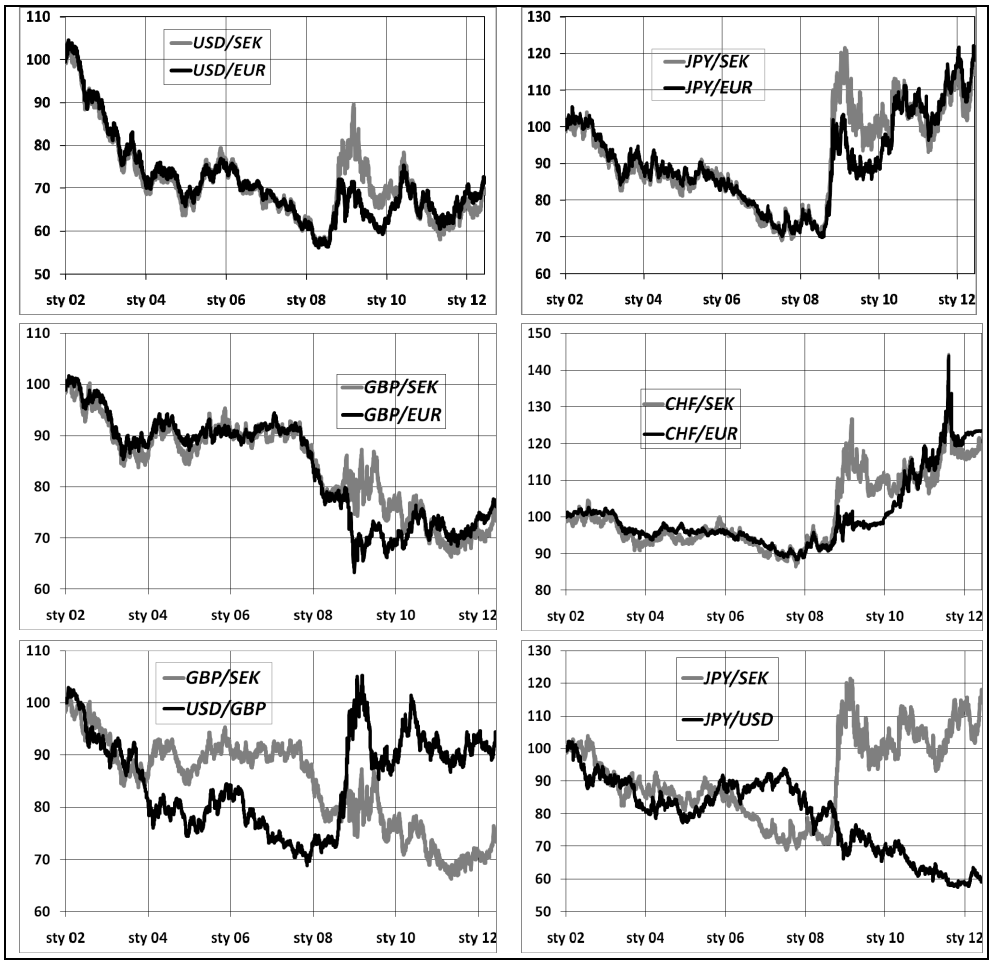
DANE EMPIRYCZNE I METODYKA BADAŃ

Analiza współzależności pomiędzy badanymi kursami przeprowadzona została w oparciu o podawane przez Reuters[3] zamknięcia dziennych notowań tych kursów. Szeregi czasowe tych zmiennych obejmowały okres od07.01.2001 do 08.06.2012 i liczyły po 2720 obserwacji, 544 pełne tygodnie 5 dniowe. Pojedyncze braki danych w poszczególnych szeregach czasowych były

[2] *Swedish national reform program 2011*. Government Office of Sweden. Stockholm 2011.
[3] Dostępne na stronie stooq.com.

uzupełniane przy pomocy średnich z obserwacji sąsiednich. Wykres indeksów dynamiki tych szeregów czasowych przedstawiony został na rysunku 1. Wykorzystanie indeksów kursów zamiast poziomów zmiennych umożliwiło porównanie ich przebiegu.

Rysunek 1. Przebieg indeksów badanych kursów walutowych



Źródło: obliczenia własne

Dosyć powszechnie przyjmuje się, że procesem losowym generującym szeregi czasowe kursów walutowych jest proces błądzenia losowego [por. na przykład: Belaire-Franch i Opong 2005, Choi 1999, Kilian i Taylor 2003]. Powoduje to, że szeregi czasowe kursów walutowych są niestacjonarne. Analiza przebiegu zmienności badanych kursów walutowych przedstawionych na rysunku 1 tę hipotezę potwierdza. Potwierdzają tę hipotezę w odniesieniu do

badanych kursów walutowych również wyniki testów stacjonarności zawarte w tabeli 1[4].

Tabela 1. Wyniki testów stacjonarności szeregów czasowych badanych kursów walutowych

Zmienna (kurs)	Test ADF				Test KPSS		
	Poziomy zmiennych		Pierwsze różnice		Statystyka testu		Wartość krytyczna ($\alpha = 0{,}05$)
	Statystyka testu	p	Statystyka testu	p	Poziomy zmiennych	Pierwsze różnice	
USD/SEK	-1,7015	0,4304	-54,4252	0,0000	17,4142	0,0994	0,461
EUR/SEK	-2,5190	0,2111	-55,4360	0,0000	4,70839	0,0951	
GBP/SEK	-1,4316	0,5682	-53,6752	0,0000	24,4839	0,0334	
CHF/SEK	-1,7737	0,3940	-53,9870	0,0000	15,1022	0,0593	
JPY/SEK	-1,8392	0,3617	-55,0625	0,0000	5,8670	0,1928	
EUR/USD	-1,8157	0,3732	-55,7351	0,0000	21,1198	0,1811	
GBP/EUR	-1,2573	0,6514	-52,9016	0,0000	25,3576	0,0786	
CHF/EUR	-0,4202	0,9034	-52,8880	0,0000	11,6808	0,2840	
JPY/EUR	-1,3147	0,6249	-58,0013	0,0000	6,3557	0,4344	
GBP/USD	1,5406	0,5130	-53,4901	0,0000	5,6452	0,2094	
CHF/USD	-1,2656	0,6476	-55,5742	0,0000	25,1997	0,03013	
JPY/USD	-0,4249	0,9026	-57,1793	0,0000	22,7165	0,1619	

Źródło: obliczenia własne

Niestacjonarność szeregów czasowych kursów walutowych wymaga zastosowania do analizy ich współzależności metodyki opracowanej przez Engla i Grangera [Engle i Granger 1987], rozwiniętej następnie przez Johansena [Johansen 1988] oraz Johansena i Juselius [Johansen i Juselius 1990]. Idea tej procedury polega na przekształceniu modelu *VAR* (*Vector Auto Regression*) dla poziomów zmiennych o postaci:

$$\mathbf{x}_t = \mathbf{A}_0\mathbf{d}_t + \sum_{i=1}^{r} \mathbf{A}_i\mathbf{x}_{t-i} + \mathbf{e}_t \tag{1}$$

gdzie: $\mathbf{x}_t = [x_{t1} \ \ldots \ x_{tk}]^{\mathrm{T}}$ – wektor obserwacji na bieżących wartościach pierwszych różnic zmiennych objaśnianych, $\mathbf{d}_t = [d_0 \ d_t \ d_{1\,t\text{-}1,} \ d_{1\,t\text{-}r} \ d_{l\,t\text{-}r,}]^{\mathrm{T}}$ – wektor egzogenicznych składników równań, którego składowymi są odpowiednio: stała równania oraz bieżące i opóźnione wartości zmiennych egzogenicznych, $\mathbf{A}_0$ – macierz parametrów przy zmiennych wektora $\mathbf{d}_t$, $\mathbf{A}_i$ – macierz parametrów przy

[4] Na niestacjonarność badanych szeregów czasowych wskazuje zarówno test Grangera, jak i test KPSS. Taka sytuacja „silnie wskazuje na niestacjonarność zmiennych" [Welfe 2009, str. 368]. Więcej na temat testów ADF i KPSS por. Welfe [2009], str. 360 i dalsze.

opóźnionych zmiennych wektora $\mathbf{x}_t$, $\mathbf{e}_t = [\mathbf{e}_{1t}$ **...** $\mathbf{e}_{kt}]^T$ wektory reszt równań modelu, *r* - rząd opóźnienia, do postaci *VECM* (*Vector Error Correction Model*) [5]:

$$\Delta \mathbf{x}_t = \mathbf{\Psi}_0 \mathbf{d}_t + \mathbf{\Pi} \mathbf{x}_{t-1} + \sum_{i=1}^{r} \mathbf{\Pi}_i \Delta \mathbf{x}_{t-i} + \boldsymbol{\varepsilon}_t \tag{2}$$

gdzie: $\mathbf{\Psi}_0$ – macierz parametrów przy zmiennych wektora $\mathbf{d}_t$, $\mathbf{\Pi} = \sum_{j=1}^{k} \mathbf{A}_i - \mathbf{I}$; $\boldsymbol{\varepsilon}_t$.- reszty modelu. Macierz $\mathbf{\Pi}$ modelu (2) jest wykorzystywana w teście Johansena. Końcowym efektem testu Johannesa jest możliwość określenia modelu, w oparciu o który przeprowadzona zostanie analiza współzależności szeregów czasowych[6]. Możliwe są bowiem trzy sytuacje: macierz $\mathbf{\Pi}$ jest pełnego rzędu i wówczas model (2) jest modelem *VAR* dla poziomów zmiennych, rząd macierzy $\mathbf{\Pi}$ większy od 0 i mniejszy od *k* wskazuje na liczbę wektorów kointegracyjnych i wymaga zastosowania modelu (2), rząd macierzy $\mathbf{\Pi}$ jest 0 wówczas model (3) jest modelem *VAR* dla przyrostów zmiennych.

Badanie współzależności kursów korony szwedzkiej przeprowadzone zostało w ten sposób, że test Johansena przeprowadzony został dla par szeregów czasowych: kurs SEK do danej waluty i kurs tej waluty do EUR, kurs SEK do danej waluty i kurs tej waluty do USD. Następnym krokiem była estymacja odpowiednich modeli dla wymienionych par kursów. W dalszej kolejności, w oparciu o wyniki estymacji modeli przeprowadzona została analizę przyczynowości[7]. Do analizy przyczynowości wykorzystany został wariant Walda testu Grangera[8]. Podsumowaniem badań będzie analiza funkcji odpowiedzi na impuls (IRF), która pozwala na opis przebiegu w czasie zależności pomiędzy badanymi zmiennymi[9].

WYNIKI BADAŃ

Przeprowadzone testy kointegracji Johansena dla par kursów walutowych (kurs korony szwedzkiej do danej waluty i kurs tej waluty do euro bądź dolara) dla każdego z wydzielonych okresów wykazały, że macierz $\mathbf{\Pi}$ jest pełnego rzędu, za

[5] Sposób otrzymywania modelu *VECM* można znaleźć w pracach Millsa [2002] i Osińskiej [2006].

[6] Szczegóły konstrukcji testu Johansena można znaleźć w pracy: Johansen S. (2000) Modelling of cointegration in the vector autoregressive model, *Economic Modelling*, Tom 17 (2000), str. 359-373.

[7] Chodzi tu o tak zwaną przyczynowość w sensie Grangera. Zmienna *x* jest przyczyną *y* w sensie Grangera, jeśli bieżące wartości *y* są lepiej objaśniane przy użyciu opóźnionych i bieżących wartości *x* niż bez ich wykorzystania. Więcej na ten temat por. Osińska [2006], str.212.

[8] Więcej na temat funkcji *IRF* por.: Kusideł E. (2000) *Modele wektorowo- autoregresyjne VAR. Metodologia i zastosowania*. ABSOLWENT, Łódź.

[9] Więcej na temat testu Grangera i jego wariantów por. Osińska [2006], str.212.

wyjątkiem: pary kursów JPY/SEK – JPY/USD w okresach: *kryzys* i *po kryzysie*, : pary kursów CHF/SEK – CHF/USD w okresie *po kryzysie*. Oznacza to, że poza wymienionymi przypadkami, do analizy współzależności badanych kursów może być wykorzystany model VAR dla poziomów zmiennych dany zależnością (1).

W oparciu o wyniki estymacji modeli przeprowadzona została analiza przyczynowości, w której weryfikowana była hipoteza, że na kurs korony szwedzkiej do danej waluty wpływał kurs danej waluty do euro i do dolara. Wyniki testu Grangera weryfikującego tę hipotezę zmieszczone zostały w tabeli 2. Przypadki, w których statystyka testu przyczynowości Grangera była istotnie różna od zera zaznaczone zostały pogrubieniem, a w przypadku, gdy prawdopodobieństwo błędu I rodzaju było zawarte w przedziale (0,05; 0,10), dodatkowo kursywą.

Tabela 2. Wyniki testu przyczynowości Grangera badającego współzależność badanych kursów

Okres	Kurs SEK do:	Przyczyna – kurs:			
		EUR do danej waluty		USD do danej waluty	
		τ	p	τ	p
Przed kryzysem	USD (EUR)	**16,5502**	**0,0054**	***9,5149***	***0,0902***
	GBP	***5,1475***	***0,0762***	**12,6751**	**0,0266**
	CHF	**11,1055**	**0,0493**	7,6492	0,1767
	JPY	**12,7797**	**0,0467**	5,0143	0,5420
Kryzys	USD (EUR)	6,9287	0,1397	9,2046	0,5128
	GBP	***9,4625***	***0,0920***	**12,1057**	**0,0334**
	CHF	**11,3450**	**0,0229**	2,7537	0,7379
	JPY	3,5257	0,4740	brak kointegracji	
Po kryzysie	USD (EUR)	8,7554	0,1192	3,1685	0,6740
	GBP	***10,3434***	***0,0661***	2,4442	0,7849
	CHF	***4,8394***	***0,0889***	brak kointegracji	
	JPY	0,7811	0,9782	brak kointegracji	

Źródło: obliczenia własne

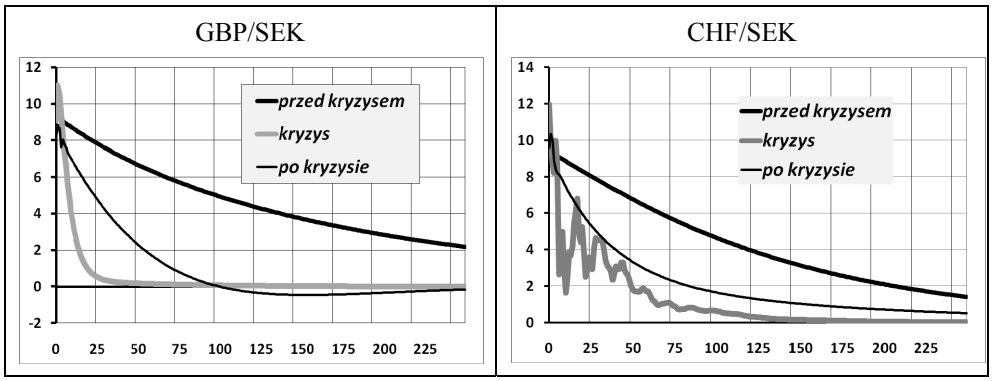
Dane zawarte w tabeli 2 wskazują, że dla kursu korony szwedzkiej do walut światowych w znacznej części przypadków przyczyną (w sensie Grangera) był kurs danej waluty do Euro. Kursy dolarowe tych walut oddziaływały w tak rozumianym sensie na kursy korony szwedzkiej znacznie słabiej. Dane zawarte w tabeli 2 wskazują też, że oddziaływanie to zmieniało się w zależności od okresu.

W okresie „przed kryzysem" na wszystkie rozpatrywane kursy korony szwedzkiej do euro wpływał kurs tych walut do euro. W pozostałych dwu okresach jedynie kurs korony do funta brytyjskiego i franka szwajcarskiego pozostawał pod wpływem kursów tych walut do euro. Kursy korony szwedzkiej do dolara i jena kształtowały się natomiast w tych dwu okresach niezależnie od kursów EUR/USD

i JPY/EUR. Oddziaływania kursów dolarowych walut światowych na kursy korony szwedzkiej do dolara widoczne jest natomiast jedynie w okresie „przed kryzysem" i tylko na kurs EUR/SEK i GBP/SEK oraz w okresie „kryzys", gdzie oddziaływanie to dotyczy wyłącznie kursu GBP/SEK. W okresie „po kryzysie" przeprowadzony test przyczynowości Grangera nie pozwolił na stwierdzenie, iż kursy korony szwedzkiej do wybranych walut światowych kształtowały się pod wpływem kursów tych walut do dolara.

Więcej o kształtowaniu się reakcji badanych zależności pomiędzy kursami walutowymi może powiedzieć funkcja odpowiedzi na impuls (*IRF*). Na rysunku 2 przedstawiony został wykres przebiegu funkcji *IRF*, opisujący rozkład w czasie reakcji kursów GBP/SEK i CHF/SEK na impuls ze strony kursów odpowiednio: GBP/EUR i CHF/EUR w każdym z wydzielonych okresów. Wykres funkcji odpowiedzi na impuls skonstruowany jest w ten sposób, że na osi rzędnych wykresu odłożona jest wielkość reakcji odpowiedniego kursu korony szwedzkiej na zmianę kursu GBP/EUR lub CHF/EUR o 1, a na osi odciętych wyrażony w dniach horyzont czasowy oddziaływania tego impulsu. Analiza funkcji odpowiedzi na impuls dotyczyła tylko kursów GBP/SEK i CHF/SEK, ponieważ tylko w przypadku tych dwu kursów stwierdzony został wpływ kursów odpowiednio GBP/EUR i CHF/EUR w każdym z wydzielonych okresów.

Rysunek 2. Przebieg funkcji odpowiedzi na impuls kursów korony szwedzkiej ze strony kursów odpowiednio GBP/EUR i CHF/EUR



Źródło: obliczenia własne

Analiza funkcji odpowiedzi na impuls dotyczy trzech elementów: kierunku oddziaływania impulsu, siły tego impulsu, oraz jego rozkładu w czasie, w tym zwłaszcza szybkości wygasania. Wykresy przebiegu funkcji reakcji na impuls zamieszczone na rysunku 2 ujawniają w tym zakresie kilka elementów wspólnych dla obydwu kursów i wszystkich wydzielonych okresów. Przede wszystkim pokazują, że w każdym okresie i w przypadku obydwu kursów zmiany kursów GBP/EUR i GBP/SEK oraz kursów CHF/EUR i CHF/SEK zmieniały się w tym samym kierunku w każdym z wyróżnionych okresów. Oznacza to, że osłabienie

euro wobec funta brytyjskiego (wzrost wartości kursu GBP/EUR) skutkowało osłabieniem korony szwedzkiej wobec funta (wzrost wartości kursu GBP/SEK) i na odwrót. Analogiczny wniosek dotyczy relacji korony i franka szwajcarskiego.

Przebieg funkcji odpowiedzi na impuls przedstawiony na rysunku 2 pozwala też zauważyć, że reakcja kursów korony szwedzkiej do funta brytyjskiego i franka szwajcarskiego na zmiany kursów tych walut do euro nie jest wyłącznie natychmiastowa. Jej maksimum nie zawsze przypadało w pierwszym dniu po wystąpieniu impulsu w postaci kursu zmian kursu, tak GBP/EUR, jak i CHF/EUR. Wygasanie impulsów w każdym z analizowanych okresów i w przypadku obydwu badanych kursów korony odbywało się bardzo powoli.

Przebieg funkcji *IRF* jest podobny w przypadku obydwu kursów korony szwedzkiej. Widoczne są natomiast różnice dotyczące przebiegu tej funkcji pomiędzy poszczególnymi okresami. Dotyczą one zarówno siły oddziaływania impulsów, jak i szybkości wygasania reakcji na impuls. W obydwu przypadkach, zarówno w przypadku kursu GBP/SEK, jak i CHF/SEK njsilniejsze było oddziaływanie na impuls ze strony kursu odpowiednie waluty w okresie „kryzys", wtedy też reakcja na impuls wygasała najszybciej. W okresie „po kryzysie" reakcje kursów GBP/SEK i CHF/SEK były bardzo podobne. Reakcje na impulsy ze strony kursów GBP/EUR, jak i CHF/EUR były najsłabsze i wygasały wyraźnie szybciej niż w przypadku okresu „przed kryzysem" i nieznacznie wolniej niż w okresie „kryzys".

Analiza przebiegu funkcji reakcji na impuls stanowi efekt końcowy całości przeprowadzonej analizy ekonometrycznej, której przedmiotem było zachowanie kursów korony szwedzkiej wobec walut światowych. Analiza ta uzupełniła wnioski wynikające z badania kointegracji. W szczególności pozwoliła stwierdzić, że reakcje kursów korony szwedzkiej na zmiany kursów innych walut zmieniły się w czasie.

PODSUMOWANIE

Przeprowadzona analiza zachowania się korony szwedzkiej do podstawowych walut światowych pozwoliła na częściowe potwierdzenie postawionej hipotezy. Stwierdzone zostało, iż korona szwedzka pozostawała pod wpływem euro. Wpływ ten był jednakże zmienny w czasie. Jedynie w pierwszym z wydzielonych okresów, „ przed kryzysem", zmienność kursów euro do dolara i jena oddziaływała na kursy korony szwedzkiej względem tych walut. W dwu pozostałych okresach nie stwierdzono takiego wpływu. We wszystkich wydzielonych okresach na kurs korony szwedzkiej do funta brytyjskiego i franka szwajcarskiego wpływała zmienność kursów tych walut do euro. Jednakże i w tym przypadku zależność ta zmieniała się w czasie i można mówić o pewnym osłabianiu tej zależności co wykazała analiza funkcji *IRF*.

Kursy walut światowych: euro, jena, funt brytyjskiego, frank szwajcarskiego względem dolara wykazywały dużo słabszy związek z kursami korony szwedzkiej

do tych walut. Tutaj również, podobnie jak w poprzednim przypadku, wpływ ten zmieniał się w czasie i słabł.

Stosunkowo dobre wyniki gospodarki szwedzkiej w ostatnich latach tłumaczone są niekiedy tym, że Szwecja znajduje się poza strefą euro. Wyniki badań wskazują, że w ostatnich latach osłabiają się również związki ze strefą euro odbywające się poprzez rynek walutowy. Kurs korony szwedzkiej względem walut światowych kształtuje się w dużym stopniu autonomicznie. Otwarte pozostaje pytanie, czy jest to jedna z przyczyn względnego sukcesu Szwecji w wychodzeniu z kryzysu, czy też skutek zmian w gospodarce tego kraju.

BIBLIOGRAFIA

Bask M., Luna de X. (2005) EMU and the stability and volatility of foreign exchange: Some empirical evidence. Chaos, Solitions and Fractals, Tom 25, str. 737–750.

Bednarz J. Gędek S. (2007) Współzależności kursów euro i funta brytyjskiego, Metody Ilościowe w Badaniach Ekonomicznych, Tom VIII, str. 25-34.

Beine M., Lahaye J., Laurent S., Neely C.J., Palm F.C. (2006) Central Bank Intervention and Exchange Rate Volatility, Its Continuous and Jump Components, Federal Reserve Bank of St. Louis Working Paper 2006-031C.

Belaire-Franch J., Opong K. K. (2005) Some evidence of random walk behavior of Euro exchange rates using ranks and signs, Journal of Banking & Finance, Tom 29, str. 1631–1643.

Benita G., Lauterbach B. (2007) Policy Factors and Exchange Rate Volatility, Panel Data versus a Specific Country Analysis, International Research Journal of Finance and Economics, Tom 7, str. 7-22.

Choi I. (1999) Testing the Random Walk Hypothesis for Real Exchange Rates, Journal of Applied Econometrics, Tom 14, str. 293-308.

Devereux M.B., Engel C. (2003) Monetary Policy in the Open Economy Revisited. Price Setting and Exchange-Rate Flexibility, Review of Economic Studies, 70, str. 765-783.

Edwards E. (2006) The Relationship between Exchange Rates and Inflation Targeting Revisited. Central Bank of Chile Working Papers, No 409.

Engle R.F., Granger C.W.J. (1987) Co-integration and Error Correction, Econometrica, Tom 55 str. 251-276

Friedman, M. (1953) Essays in Positive Economics, University of Chicago Press, Chicago.

Froot K., Stein J. (1991) Exchange Rates and Foreign Direct Investment. An Imperfect Capital Markets Approach, Quarterly Journal of Economics, Tom 196, str. 1191-1218.

Galí, J., Monacelli, T. (2005) Monetary policy and exchange rate volatility in a small open economy, Review of Economic Studies, Tom 72, str. 707–734.

Johansen, S. (1988) Statistical Analysis of Cointegration Vectors' Journal of Economic Dynamics and Control, Tom 12, str. 231–254.

Johansen S. (2000) Modelling of cointegration in the vector autoregressive model, Economic Modelling, Tom 17 (2000), str. 359-373.

Johansen, S., Juselius, K. (1990) Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money, Oxford Bulletin of Economics and Statistics, Tom 52, str. 169–210.

Karras G., Lee J.M., Stokes H. (2005) Sources of Exchange-Rate Volatility. Impulses or Propagation? , International Review of Economics and Finance 14, str. 213–226.

Ketels C. (2012) Sweden's Position in the global economy. Globaliseringsforum Rapport #2, Stockholm.

Kilian L., Taylor M.P. (2005) Why is it so difficult to beat the random walk forecast of exchange rates?, Journal of International Economics, Tom 60, str. 85–107.

Kusideł E. (2000) Modele wektorowo- autoregresyjne VAR. Metodologia i zastosowania. ABSOLWENT, Łódź.

Lanne M., Vesala T. (2006) The effect of a transaction tax on exchange rate volatility, Bank of Finland Research Discussion Papers, 11/2006.

McKenzie M. (1999) The Impact of Exchange Rate Volatility on International Trade Flows, Journal of Economic Surveys, Tom 13, str. 71-106.

Mills T.C. (2002) The Econometric Modeling of Financial Time Series, Cambridge University Press, Cambridge.

Mussa M. (2002) The euro versus the dollar: not a zero sum game, Journal of Policy Modeling, Tom 24, str. 361-372.

Nucci F. Pozzolo A.F. (2001) Investment and the Exchange Rate: An Analysis with Firm-Level Panel Data, European Economic Review, Tom 45, str. 259-283.

Osińska M. (2006) Ekonometria finansowa, PWE, Warszawa.

Reinhart C., Rogoff K. (2004) The Modern History of Exchange Rate Arrangements: A Reinterpretation, Quarterly Journal of Economics, Tom 119, str.1-48.

Report on global foreign exchange market activity in 2010. Monetary and Economic Department, Bank for International Settlements, 2011.

Stockman A. (1988) Real Exchange Rate Variability under Pegged and Floating Exchange Rate System: An Equilibrium Theory, University of Rochester Working Paper No 2565.

Swedish national reform program 2011. Government Office of Sweden. Stockholm 2011.

Tarashev N.A. (2007) Speculative Attacks and the Information Role of Interest Rate, Journal of the European Economic Association, Tom 5, str. 1-36.

Welfe A. (2009) Ekonometria. Metody i ich zastosowanie, PWE, Warszawa

Witkowska D. (2011) Kointegracja kursów walutowych Polski, Węgier i Czech, Metody Ilościowe w Badaniach Ekonomicznych, Tom XII, str. 399-408.

THE BEHAVIOUR OF SWEDISH KRONA EXCHANGE RATES

Abstract, The aim of the paper was the analysis of the behaviour of Swedish krona (SEK) exchange rates. The tool of analysis was a VAR model. The analysis has shown that SEK exchange rates were and determined to a large extend by the variability of euro exchange rates. The influence of US dollar exchange rates was much less significant.

Keywords: exchange rates, Swedish krona, VAR model, impulse response function

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 89 – 96

O WŁASNOŚCIACH SZEREGÓW CZASOWYCH I PŁYNNOŚCI AKCYJNYCH FUNDUSZY ETF NOTOWANYCH NA EUROPEJSKICH GIEŁDACH

Ewa Małgorzata Gnatowska
Dom Inwestycyjny Nehrebetius S. A. w Krakowie
e-mail: ewa.gnatowska@nehrebetius.pl

Streszczenie: Praca jest poświęcona ETF-om, czyli funduszom notowanym na giełdzie. Na początku przedstawione zostają krótko trzy z nich, które pojawiły się niedawno na GPW w Warszawie. Następnie badane są własności szeregów czasowych dla funduszu opartego na indeksie WIG 20 w porównaniu z innymi europejskimi ETF-ami firmy Lyxor opartymi o indeksy giełdowe. Dyskutowana jest słaba efektywność rynku oraz płynność tych instrumentów w kontekście kryzysu wypłacalności państw europejskich.

Słowa kluczowe: exchange traded fund, płynność, słaba efektywność rynku

WSTĘP: NOWE INSTRUMENTY NA WARSZAWSKIEJ GIEŁDZIE

Exchange Traded Funds to stosunkowo nowe produkty inwestycyjne, powstałe na początku lat dziewięćdziesiątych, łączące w sobie cechy tradycyjnych funduszy inwestycyjnych oraz notowanych na giełdzie akcji. Duża część z nich naśladuje zachowanie indeksów, które mogą być zarówno dobrze znanymi indeksami giełdowymi, jak i syntetycznymi wskaźnikami. Postrzegane jako sposób zdobycia ekspozycji na dany rynek bez konieczności nabycia koszyka instrumentów tworzących indeks ani instrumentów pochodnych na nim opartych. Jako zalety tej inwestycji podkreśla się relatywnie niższe koszty zarządzania niż dla tradycyjnych funduszy, korzyści podatkowe oraz dostępność instrumentów w obrocie giełdowym, często połączoną z dużą płynnością. Na GPW w Warszawie mamy już trzy takie instrumenty, wyemitowane przez firmę Lyxor Asset Management, jednego z największych dostarczycieli ETF-ów w Europie. Pierwszy pojawił się we wrześniu 2010 roku dając możliwość inwestycji w indeks WIG 20. Dwa kolejne naśladują DAX i SP 500 i notowane są od czerwca 2011.

CELE PRACY I WYKORZYSTANE DANE

Wraz z rozwojem tych produktów w latach dziewięćdziesiątych XX wieku pojawiły się pierwsze opracowania naukowe, starające się opisać własności szeregów czasowych stóp zwrotu osiągniętych z inwestycji w ETF-y. Kolejnym poruszanym zagadnieniem była kwestia płynności tych instrumentów, notowanych zresztą nieraz jednocześnie na różnych giełdach.

W pracy tej podejmę próbę analizy pierwszego polskiego funduszu ETF i porównaniu go z innymi reprezentantami szerokiej rodziny europejskich funduszy akcyjnych firmy Lyxor. Zakres notowań to na razie (do sierpnia 2012) niecałe dwa lata, ponieważ jednak rozważam tylko dzienne stopy zwrotu, ilość danych zdaje się być wystarczające do wykonania pierwszych analiz ilościowych. Otrzymane wyniki przedyskutuję w kontekście obecnego kryzysu wypłacalności państw strefy euro. Notowania ETF-ów oraz inne niezbędne informacje zaczerpnęłam ze stron internetowych poszczególnych giełd oraz twórcy ETF – firmy Lyxor, posiłkowałam się też serwisem www.stooq.pl oraz www.bloomberg.com.

Z ETF-em bazującym nas WIG 20 warto porównać na przykład ETF na hiszpański indeks IBEX lub francuski CAC 40 albo, w obrębie tak zwanych rynków wschodzących, ETF bazujący na indeksie 20 największych spółek notowanych w Istambule. Dla wszystkich rozważanych funduszy przyjęłam zakres danych odpowiadający polskiemu ETF, oczekując jednocześnie, że te instrumenty, które istniały już od kilku lat, mają dzisiaj inne własności, niż „nowe" ETF-y. Pełny zakres badanych funduszy zawiera poniższa tabela.

Tabela 1. Charakterystyka badanych funduszy ETF

Nazwa	Indeks	Giełda	Notowany od	Waluta
Lyxor ETF WIG 20	WIG 20	GPW	2010-09-21	złoty
Lyxor ETF DAX	DAX	Xetra ETF	2006-06-28	euro
Lyxor ETF IBEX 35	IBEX 35	System ciągły	2006-10-03	euro
Lyxor ETF FTSE 100	FTSE 100	LSE	2009-11-06	funt
Lyxor ETF CAC 40	CAC 40	Euronext Paris	2001-01-22	euro
Lyxor ETF FTSE MIB	FTSE MIB	Borsa Italiana	2003-11-12	euro
Lyxor ETF TURKEY	DJ Turkey Titans 20	Euronext Paris	2006-08-29	euro

Źródło: www.lyxoretf.pl

METODOLOGIA I WYNIKI BADAŃ

Słaba efektywność rynku

W literaturze pojawiło się kilka prac dotyczących efektywności rynku w odniesieniu do ETF-ów, omawiających szczegółowo różne kategorie efektywności oraz zestaw przydatnych w badaniach testów. Zdecydowałam się,

opierając się na pracy G. Rompolisa[1] (dotyczącej amerykańskich ETF-ów emitowanych przez firmę iShares) skupić na efektywności słabej. Testowałam hipotezę zerową: stopy zwrotu z ETF podlegają błądzeniu losowemu, przeciw hipotezie alternatywnej: stopy zwrotu z ETF nie podlegają błądzeniu losowemu, z wykorzystaniem dwóch testów statystycznych parametrycznych. Przeprowadziłam test autokorelacji, stosując model autoregresji rzędu K opisany przez równanie:

$$R(t) = a_0 + \sum_{k=1}^{K} a_k R(t-k) + e\ (t) \tag{1}$$

R(t) jest tutaj stopą zwrotu z ETF w dniu t, R(t-k) dla k=1,2,..,K – analogiczną stopą zwrotu w dniu t-k, a_k dla k=0,1,..K współczynnikami liczbowymi, a e(t) oznacza błąd losowy. Przyjęłam w trzech kolejnych badaniach: K=1,2,3. Jeżeli obliczona wartość Q-statystyk okazałaby się nieistotna statystycznie, uznalibyśmy, że rynek ETF jest efektywny w formie słabej. Wartość Q dla równania AR(K) zdefiniowałam wzorem Ljunga- Boxa:

$$Q = n(n+2)\sum_{k=1}^{K} \frac{\rho_k}{n-k} \tag{2}$$

gdzie n jest ilością obserwacji, ρ_k oznacza autokorelację rzędu k, a k=1,2,...K. Statystyka ta ma rozkład Chi-kwadrat o K stopniach swobody.

Dodatkowo dla rozważanego wyżej równania (1) obliczyłam tak zwaną statystykę LM (Breush-Godfrey Lagrange Multiplier Test) o rozkładzie Chi-kwadrat o K (odpowiednio 1, 2 i 3) stopniach swobody, równą iloczynowi liczby obserwacji i współczynnika determinacji w równaniu regresji wielorakiej (1).

Tabela 2. Wyniki badania słabej efektywności z wykorzystaniem testu Ljunga-Boxa

Instrument	AR(1)		AR(2)		AR(3)	
	statystyka Q	wartość p	statystyka Q	wartość p	statystyka Q	wartość p
ETF WIG	1,42	0,23	9,51	0,01	9,84	0,02
ETF DAX	8,30	0,00	9,69	0,01	14,62	0,00
ETF IBEX	6,97	0,01	10,45	0,01	15,58	0,00
ETF FTSE	4,56	0,03	7,62	0,02	10,55	0,01
ETF CAC	1,94	0,16	3,16	0,21	7,38	0,06
ETF MIB	3,38	0,07	5,07	0,08	7,75	0,05
ETF TUR	1,74	0,19	1,74	0,42	2,16	0,54

Źródło: obliczenia własne

1 Rompotis G. (2011) Testing Weak-Form Efficiency of Exchange Traded Funds Market, AESTIMATIO, The EIB International Journal of Finance; Tom 2, str.2-33.

Tabela 3. Wyniki badania słabej efektywności z wykorzystaniem statystyki LM

Instrument	AR(1)		AR(2)		AR(3)	
	statystyka LM	wartość p	statystyka LM	wartość p	statystyka LM	wartość p
ETF WIG	1,45	0,23	9,89	0,01	9,94	0,02
ETF DAX	8,43	0,00	10,87	0,00	14,62	0,00
ETF IBEX	7,19	0,01	12,00	0,00	15,13	0,00
ETF FTSE	4,59	0,03	8,43	0,01	10,20	0,02
ETF CAC	2,03	0,15	3,46	0,18	7,07	0,07
ETF MIB	3,36	0,07	5,51	0,06	7,46	0,06
ETF TUR	1,79	0,18	1,79	0,41	2,19	0,53

Źródło: obliczenia własne

Wyniki zawarte są w tabelach 2. i 3. I dla obu testów są ze sobą zgodne. Wskazują, że na poziomie ufności 95%, jedynie rynki ETF-ów: francuskiego, włoskiego i tureckiego można uznać za efektywne w formie słabej. O tyle nie jest to zaskoczeniem, że dwa pierwsze istnieją stosunkowo długo, odpowiednio od 2001 i od 2003 roku, a turecki ETF notowany jest na tej samej dużej giełdzie co francuski (Euronext Paris). Silna zależność AR(3) stwierdzona została natomiast dla instrumentu niemieckiego, hiszpańskiego, brytyjskiego i pierwszego polskiego funduszu ETF. Dwa ostatnie notowane są od niedawna (odpowiednio od listopada 2009 i września 2010), nie dziwi więc tutaj odstępstwo od słabej efektywności.

Moim przypuszczeniem było także, że na efektywność rynku może wpływać w istotny sposób zaangażowanie państwa w kryzys płynnościowy w Europie, zarówno jako beneficjenta ewentualnej pomocy (Hiszpania, Włochy) jak i głównego decydenta w tej sprawie (Niemcy, Francja). Dla ETF CAC i ETF MIB jednak hipoteza zerowa nie została odrzucona, oczekiwania te potwierdziły się natomiast w przypadku Niemiec i Hiszpanii.

Zagadnienie płynności

Dalsza część pracy będzie dotyczyła płynności rynku. Cena EFT zmienia się w ciągu dnia, na jego koniec publikowana jest cena zamknięcia oraz NAV, czyli wartość aktywów funduszu przypadająca na jedną akcję. Praca Chacko, Dasa i Fana [2] wprowadziła miarę płynności opartą na różnicy między tymi wielkościami, daną wyrażeniem:

$$BILLIQ = -10000 \circ \log \frac{NAV}{NAV + |ETF - NAV|} \tag{3}$$

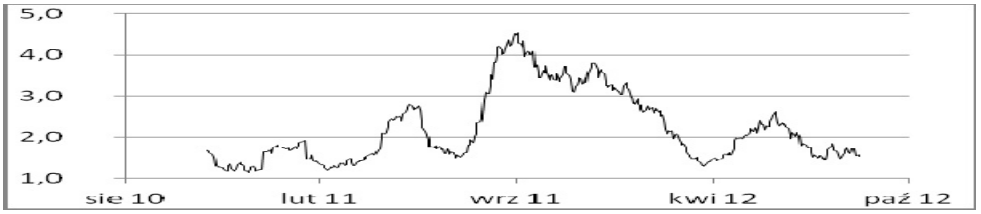
[2] Chacko G., Das S., Fan R. (2011) An Index-Based Measure of Liquidity, Santa Clara University Leavel School of Business Research Paper Series No 11-10.

udowadniając, że jest ona wysoko skorelowana z czynnikami opisującymi płynność szerokiego rynku.[3] Dla ETF z rozważanej grupy otrzymałam następujący historyczny przebieg wielkości BILLIQ:

Rysunek 1. Historyczny przebieg płynności dla ETF WIG

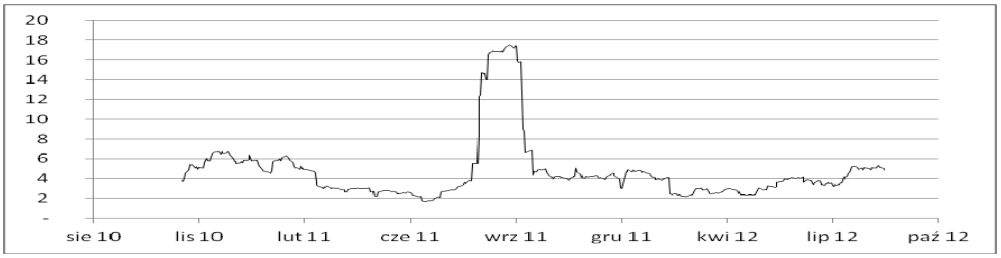
Źródło: obliczenia własne

Rysunek 2. Historyczny przebieg płynności dla ETF DAX



Źródło: obliczenia własne

Rysunek 3. Historyczny przebieg płynności dla ETF IBEX



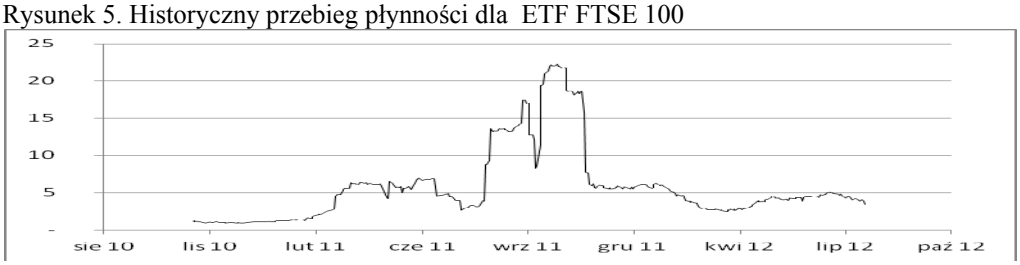
Źródło: obliczenia własne

[3] Zależność ma interpretację ciągłej stopy zwrotu wyrażonej w punktach bazowych. Dla zachowania przejrzystości wykresów przyjęłam średnią 30 -okresową.

Rysunek 4. Historyczny przebieg płynności dla ETF CAC

Źródło: obliczenia własne

Rysunek 5. Historyczny przebieg płynności dla ETF FTSE 100



Źródło: obliczenia własne

Rysunek 6. Historyczny przebieg płynności dla ETF MIB

Źródło: obliczenia własne

Rysunek 7. Historyczny przebieg płynności dla ETF TUR

Źródło: obliczenia własne

Niektóre parametry statystyczne charakteryzujące płynność rynków zawiera poniższa tabela.

Tabela 3. Wyniki badania płynności wyrażonej wskaźnikiem BILLIQ

Instrument	Średnia	Odchylenie standardowe	Maksimum
ETF WIG 20	16,5	5,9	36,3
ETF DAX	2,2	0,9	4,5
ETF IBEX	4,7	3,3	17,5
ETF FTSE	5,8	5,0	22,3
FTSE CAC	2,2	0,7	4,0
FTSE MIB	6,6	4,2	19,8
ETF TUR	17,0	6,2	39,5

Źródło: obliczenia własne

Implikowana przez ceny ETF-ów płynność rynku maleje wraz ze wzrostem wartości BILLIQ. Zależy ona od kilku czynników, po pierwsze aktywności mających ją zapewnić animatorów danego rynku, po drugie, obecności dużych inwestorów, mogących doprowadzić do wyrównania się ceny funduszu ETF i jego NAV w wyniku transakcji arbitrażowych. W tym kontekście, to, jak długo dany instrument jest na danej giełdzie notowany, powinno wyjaśniać różnice w średnim poziomie płynności (Tabela 3.). Zostało to potwierdzone przez powyższe badania poza przypadkiem funduszu ETF TUR opartego na spółkach tureckich, wynik którego zbliżony jest do wartości odnotowanej dla ETF-a na indeks WIG 20 i wyraźnie wyższy od pozostałych. Nie jest tu zapewne bez znaczenia, że oba indeksy należą do wskaźników giełdowych grupy rynków wschodzących.

Jednak tym, co najbardziej rzuca się w oczy to przebieg historyczny tych zależności. Wszystkie one wykazują maksimum na jesieni 2011 roku, jedynie ETF CAC zdaje się mieć drugie maksimum na początku 2012 (Rysunek 4.), a na rynku niemieckim obniżona płynność utrzymuje się przez najdłuższy czas (Rysunek 2.) Jak już wspominałam, rozważany okres, to czas trwającego do dzisiaj kryzysu wypłacalności. Francja i Niemcy to najsilniejsze kraje strefy euro, zmuszone ponieść największe koszty ratowania państw bardziej zadłużonych. Państwami z grupy tak zwanych krajów peryferyjnych są tu Hiszpania i Włochy. Rysunek 3. zdaje się sugerować, że Hiszpania przeżywała problemy z płynnością stosunkowo krótko – obserwujemy wąskie maksimum w okolicach września i października 2011 roku. Polska, Wielka Brytania, a tym bardziej Turcja, uwikłane są w kryzys bardziej pośrednio, jako, że do strefy euro nie należą, mimo to i one zaraziły się nim, czego dowodem są wysokie maksima na wykresach. Rysunek 5 nieco różni się kształtem od pozostałych, ale pamiętać musimy o, podobnej do amerykańskiego QE2, operacji skupu obligacji skarbowych przeprowadzonej przez rząd brytyjski, nie pozostającej bez wpływu również na rynek akcji.

PODSUMOWANIE

Stwierdziliśmy, że rynki europejskich funduszy notowanych na giełdzie nie są często efektywne w formie słabej. Ponieważ znanym faktem jest, że dla samych indeksów ta własność zachodzi na większości giełd, wniosek stąd, że szeregi czasowe ETF mogą mieć inne własności niż instrumenty na których bazują, co warto wziąć pod uwagę opracowując ilościowe strategie inwestycyjne. Na płynność poszczególnych ETF-ów silnie wpływa sytuacja gospodarcza na świecie, a nie tylko charakterystyka samej giełdy, na której są notowane.

Ze względu na rozmiar pracy nie przeprowadziłam tutaj innych, pojawiających się często w literaturze w kontekście funduszy ETF, testów słabej efektywności rynku, w szczególności testów na istnienie w równaniu (1) pierwiastka jednostkowego (rozszerzony test Dicka-Fullera lub nieparametryczny test Phillipsa-Perona). Może to stanowić przedmiot dalszych badań.

BIBLIOGRAFIA

Nawrot W. (2007) Exchange-Traded Funds (ETF). Nowe produkty na rynku funduszy inwestycyjnych , CeDeWu.PL, Warszawa.

Chacko G., Das S., Fan R. (2011) An Index-Based Measure of Liquidity, Santa Clara University Leavel School of Business Research Paper Series No 11-10.

Chacko G., Jurek J., Stafford E. (2008) The Price of Immediacy, Journal of Finance 63(3), str. 1253-1290.

Rompotis G. (2011) Testing Weak-Form Efficiency of Exchange Traded Funds Market, AESTIMATIO, The EIB International Journal of Finance; Tom 2, str.2-33.

TIME SERIES PROPERTIES AND LIQUIDITY OF SOME EQUITY ETF-S TRADED ON EUROPEAN STOCK EXCHANGES

Abstract: This paper is devoted to ETFs, introducing briefly three of them, which are traded on the Warsaw Stock Exchange. Next time series properties are investigated for the fund following WIG 20 Index in comparison with other European ETFs provided by Lyxor Asset Management based on stock market indices. Weak-form efficiency of the market is discussed and liquidity of these instruments described, in context of the solvency crisis in Europe.

Keywords: exchange traded fund, liquidity, weak-form efficiency of the market

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 97 – 106

BADANIE ZWIĄZKÓW MIĘDZY ZMIENNYMI REALNYMI A ZMIENNYMI NOMINALNYMI W POLSKIEJ GOSPODARCE W LATACH 1997-2011

Rumiana Górska, Dorota Narojczyk
Katedra Finansów
Wyższa Szkoła Finansów i Zarządzania w Warszawie
rumiana.gorska@op.pl, dorota.narojczyk@wp.pl

Streszczenie: Artykuł przedstawia wyniki analizy związków między wybranymi zmiennymi realnymi a nominalnymi w polskiej gospodarce w latach 1997-2011. Większość zidentyfikowanych zależności przejawia się z opóźnieniami w czasie. Wzrost gospodarczy zależy w głównej mierze od czynników należących do realnej sfery gospodarki. Na spadek bezrobocia wpływa wzrost PKB oraz inwestycje. Wzrost stopy referencyjnej NBP powoduje spadek inflacji, lecz ma ujemny wpływ na wzrost PKB. Inflacja jest zjawiskiem przede wszystkim monetarnym, zależy głównie od wzrostu podaży pieniądza. Wzrost zadłużenia zagranicznego Polski skutkuje wzrostem podaży pieniądza.

Słowa kluczowe: zmienne realne, zmienne nominalne, korelogram wzajemny, modele ADL, stopa procentowa, Produkt Krajowy Brutto

WSTĘP

Dyskusja na temat mechanizmu powiązań między zmiennymi opisującymi realną stronę gospodarki, takimi jak Produkt Krajowy Brutto (PKB), konsumpcja, inwestycje czy stopa bezrobocia a zmiennymi nominalnymi, takimi jak podaż pieniądza, poziom cen oraz nominalna stopa procentowa jest tematem szeroko omawianym w literaturze ekonomicznej. Można przy tym wyróżnić dwie główne szkoły, które cieszyły się najszerszym poparciem w środowisku naukowym i na przemian wpływały na politykę gospodarczą różnych krajów w XX wieku.

Przypomnijmy, że według teorii Johna Maynarda Keynesa (Keynes, 1936), PKB jest wyznaczany przez ogólnogospodarczy popyt, a prowadzona przez rząd polityka kierowania popytem powinna odgrywać główną rolę w polityce

makroekonomicznej. Ta teoria stała się popularna wśród ekonomistów, którzy uważają, że rynki są niedoskonale, stąd potrzebna jest aktywna polityka interwencyjna państwa. W czasie kryzysu rząd powinien wspierać popyt poprzez rozmaite działania, w tym również kreować popyt poprzez wzrost podaży pieniądza. Według szkoły keynesowskiej przyczyny zarówno inflacji, jak i deflacji leżą w wahaniach składników zagregowanego popytu, czyli w realnej gospodarce. Sektor monetarny jedynie dostosowuje się do poziomu tego popytu.

Monetaryzm jest drugą popularną szkołą w teorii makroekonomii XX wieku. Współczesny monetaryzm wywodzi się z tzw. ilościowej teorii pieniądza Miltona Friedmana (Friedman, 1956), która podkreśla znaczenie podaży pieniądza przy analizie polityki makroekonomicznej. Według szkoły monetarystów nominalna ilość pieniądza jest główną przyczyną zmiany ogólnego poziomu cen w długim okresie i wahań poziomu realnej aktywności gospodarczej w krótkim okresie.

Założenie, że zmienne nominalne nie mają wpływu na zmienne realne nazywane jest w literaturze ekonomicznej klasyczną dychotomią. Zgodnie z tą koncepcją zmienne realne zależą od postępu technologicznego oraz gustów konsumenckich, natomiast podaż pieniądza nie wpływa na realną stronę gospodarki. Stwierdzenie, że wzrost podaży pieniądza powoduje proporcjonalny wzrost cen jest znane jako hipoteza neutralności pieniądza.

Powiązania między realną stroną gospodarki a zmiennymi nominalnymi są często przedmiotem badań w kontekście różnych modeli makroekonomicznych - modelu realnego cyklu gospodarczego, modelu sztywnych cen i modelu efektów płynnościowych. King i Watson (1996) badają związki między stopą procentową a realną stroną gospodarki na podstawie danych z powojennej gospodarki Stanów Zjednoczonych. Ich badania nie potwierdzają, że realna lub nominalna stopa procentowa może być indykatorem wyprzedzającym realną aktywność gospodarczą.

Mertens (2010) bada związki kowariancyjne między zagregowanym produktem a realną stopą procentową w warunkach występowania strukturalnych szoków monetarnych i technologicznych. Związki między produktem i realną stopą procentową są różne dla okresów 1959-1979 (okres nazywany *The Great Inflation*), gdy większe znaczenie miały szoki monetarne, i 1982-2006 (*The Great Moderation*), gdy większe znaczenie miały szoki technologiczne. W pierwszym okresie realna stopa procentowa jest antycykliczna, a w drugim okresie jest procykliczna.

W Polsce badanie neutralności pieniądza przeprowadzili Brzoza-Brzezina, Kłos, Kot i Łyziak (2002). Badania te nie prowadzą do jednoznacznych wniosków na temat długookresowych efektów szoków monetarnych na realną gospodarkę. W swoim artykule autorzy przytaczają wyniki podobnych badań dla gospodarek USA, Australii, Wielkiej Brytanii, Kanady, Niemiec, Japonii i innych krajów, część z nich potwierdza hipotezę, że pieniądz jest neutralny względem dochodu. Według innych badań (Bruzda, 2002) reakcja realnej gospodarki na decyzje Rady Polityki Pieniężnej NBP pojawia się po czterech kwartałach. Badania te

przeprowadzono dla danych z lat 1993-2001, kiedy polska gospodarka była jeszcze we wczesnym okresie transformacji. Obszerną analizę mechanizmu transmisji polityki pieniężnej w ostatnich latach w Polsce przeprowadzili Demczuk, Łyziak, Przystupa, Sznajderska i Wróbel (2012).

Funkcjonowanie współczesnej gospodarki Polski jest uzależnione nie tylko od decyzji krajowych agentów gospodarczych (przedsiębiorstwa i gospodarstwa domowe), rządu oraz władzy monetarnej (NBP), ale również od międzynarodowej pozycji Polski uwzględniającej zadłużenie zagraniczne kraju. Zmienne opisujące gospodarkę oddziałują na siebie wzajemnie, a część tych wpływów ujawnia się z opóźnieniem w czasie. W związku z tym, celem niniejszego artykułu jest identyfikacja dynamicznej struktury powiązań i zależności zachodzących między wybranymi zmiennymi realnymi a nominalnymi opisującymi polską gospodarkę, przy uwzględnieniu wpływu zadłużenia zagranicznego. Badanie przeprowadzono na podstawie danych kwartalnych z okresu 1997-2011.

DANE

Zmienne, które są uwzględnione w badaniu, podzielono na dwie grupy. Pierwsza grupa to zmienne realne, w skład której wchodzą: PKB, inwestycje i konsumpcja, stopa bezrobocia, wynagrodzenie realne oraz realna stopa procentowa. Druga grupa są to zmienne nominalne, w tym: podaż pieniądza (agregat M3), stopa inflacji (względna zmiana cen towarów i usług konsumpcyjnych), stopa referencyjna[1] Narodowego Banku Polskiego (NBP), która w badaniu pełni rolę nominalnej stopy procentowej, zadłużenie zagraniczne kraju oraz kurs walutowy (EUR)[2]. Wszystkie zmienne są przedstawione w postaci wskaźników dynamiki, liczone jako iloraz poziomu zjawiska w danym kwartale do analogicznego kwartału w roku poprzednim.

Dane o PKB, konsumpcji, inwestycji, bezrobociu oraz wynagrodzeniu realnym pochodzą ze strony internetowej Głównego Urzędu Statystycznego.

Dane o podaży pieniądza, stopie referencyjnej NBP, zadłużeniu zagranicznym oraz kursie walutowym pochodzą ze strony internetowej NBP. Ponadto, do celów niniejszego badania została obliczona realna stopa procentowa jako różnica stopy referencyjnej i stopy inflacji.

[1] Z uwagi na brak danych o stopie referencyjnej w latach 1997-1999 stopa nominalna jest liczona jako stopa lombardowa NBP pomniejszona o 4 punkty procentowe. (Różnica między stopą lombardową a stopą referencyjną w późniejszym okresie wynosiła 4 punkty procentowe).

[2] Od 1999r.

METODOLOGIA

Badanie wzajemnych powiązań między wybranymi zmiennymi przeprowadzono w dwóch etapach. W pierwszym etapie zbadano zależności między wszystkimi możliwymi parami zmiennych za pomocą wzajemnych korelogramów, nie zważając na przesłanki teorii ekonomicznej o istnieniu lub nieistnieniu danej zależności. Wzajemne korelogramy umożliwiają badanie zależności między wybraną zmienną a przeszłymi, obecnymi oraz przyszłymi wartościami drugiej zmiennej, co pozwala także na uchwycenie zależności występujących z opóźnieniem lub wyprzedzeniem. Dla wszystkich zmiennych obliczono również współczynniki autokorelacji.

W drugim etapie, dla każdej zmiennej oszacowano odpowiedni model z klasy autoregresyjnych modeli z rozkładem opóźnień ADL (*autoregressive distributed lags*) na podstawie zależności otrzymanych w pierwszym etapie. Modele ADL uwzględniają możliwość wystąpienia różnych dostosowań dynamicznych, dotyczących nie tylko zmiennej objaśnianej, ale również zmiennych objaśniających:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + ... + \beta_0 x_t + \beta_1 x_{t-1} + ... + \gamma_0 z_t + \gamma_1 z_{t-1} + ... \quad (1)$$

Wybór specyfikacji modelu dokonano za pomocą kryteriów informacyjnych. Aby uniknąć badania zależności zmiennych niestacjonarnych względem stacjonarnych, badanie przeprowadzono dla przyrostów zmiennych, które są stacjonarne[3].

WYNIKI

Badanie wzajemnych zależności między zmiennymi

Celem pierwszego etapu badania jest zdiagnozowanie siły i kierunku zależności między wszystkimi możliwymi parami zmiennych oraz opóźnień z jakimi występują te zależności.

W tabeli 1 podano wartości istotnych współczynników korelacji między poszczególnymi zmiennymi oraz odpowiadające im opóźnienia. W nawiasach podano rząd opóźnienia zmiennej z danej kolumny. Główna przekątna tabeli zawiera wartości współczynników autokorelacji.

[3]Stacjonarność szeregów zbadano za pomocą testu ADF.

Tabela 1. Współczynniki korelacji i autokorelacji

	PKB	I	C	REAL	WR	U	CPC	REF	M	ER	ZADL
PKB	0,36(1) -0,29(4)	0,5(0)	0,51(0)	-	0,3(0)	-0,24(0)	-0,48(3)	-0,34(2) -0,58(3) -0,36(4)	-	0,36(2)	-
I	0,5(0)	-0,3(8)	0,29(0)	-0,27(2)	0,27(0) 0,37(3)	-	-	-	-	-	-
C	0,51(0)	0,29(0)	-0,27(5)	-	0,3(0)	-	-	-	-	-	-
REAL	-	-	-	-	0,34(0)	0,33(5)	-0,26(0) 0,29(4)	0,5(0)	-	-0,31(1)	-0,3(1)
WR	0,3(0)	0,27(0)	0,3(0)	0,34(0)	-	-	-0,25(0)	-	-	-	-
U	-0,48(1) -0,48(2) -0,48(3)	-0,4(2) -0,4(3)	-	-0,33(1)	-	0,6(1)	-0,29(0) -0,32(1)	-0,39(0) -0,49(1)	0,27(2)	0,41(0)	-
CPC	0,47(1)	-	0,29(2)	-0,26(0) -0,37(5)	-0,25(0)	-0,29(0)	0,39(1)	0,69(0) -0,42(4)	0,26(4)	-0,31(0)	-
REF	0,4(2)	-	-	0,5(0) -0,46(5)	0,27(2)	-0,39(0)	0,69(0) 0,51(1)	0,55(1)	-	-0,4(0) 0,27(3)	-
M	-	-	-	-	-	-0,32(1) -0,31(2)	0,35(4)	0,32(2)	-0,5(4)	0,48(0) -0,38(1)	0,46(0) 0,32(5)
ER	-0,35(1)	-	-	-	-	0,41(0)	-0,31(0) 0,32(4)	-0,4(0) -0,57(1)	0,48(0)	-0,5(4)	0,81(0) -0,46(4)
ZADL	-	-	-	-	-	-	-	-	0,46(0) -0,38(4)	0,81(0) -0,5(4)	-0,52(4)

Źródło: badania własne

Oznaczenia zmiennych są następujące:

- PKB – Produkt Krajowy Brutto
- I – Inwestycje (Nakłady brutto na środki trwałe)
- C – Konsumpcja (Spożycie indywidualne)
- REAL – Realna stopa procentowa (REAL=REF–CPC)
- WR – Przeciętne miesięczne wynagrodzenia realne brutto w gospodarce narodowej ogółem
- U – Stopa rejestrowanego bezrobocia
- CPC – Stopa inflacji (względna zmiana cen towarów i usług konsumpcyjnych)
- REF – Stopa referencyjna NBP (stan na koniec okresu)
- M – Podaż pieniądza (agregat M3)
- ER – Kurs walutowy EUR/PLN (dane od roku 1999)
- ZADL – Zadłużenie zagraniczne (dane od roku 1999)

Produkt Krajowy Brutto jest dodatnio skorelowany z konsumpcją, inwestycjami[4], realnym wynagrodzeniem i kursem walutowym[5] oraz ujemnie skorelowany ze stopą referencyjną NBP, stopą inflacji oraz stopą bezrobocia. Inwestycje i konsumpcja są skorelowane jedynie ze zmiennymi realnymi. Korelacja między realną stopą procentową a referencyjną stopą oraz stopą inflacji wynika z faktu, iż stopa realna jest równa różnicy między stopą nominalną a stopą inflacji. Realna stopa procentowa jest dodatnio skorelowana również

[4]Wynika to z faktu, że konsumpcja i inwestycje są składnikami PKB.
[5]Wzrost kursu jest równoznaczny z deprecjacją waluty krajowej.

z wynagrodzeniem realnym oraz ujemnie skorelowana z kursem walutowym i zadłużeniem zagranicznym sprzed kwartału. Wynagrodzenie realne jest dodatnio skorelowane ze wszystkimi zmiennymi realnymi i ujemnie skorelowane ze stopą inflacji. Stopa bezrobocia wzrasta po spadku zmiennych realnych (PKB, inwestycje, realna stopa procentowa). Zidentyfikowano słabą ujemną zależność między stopą bezrobocia a stopą inflacji.

Inflacja jest dodatnio skorelowana z opóźnionymi wartościami PKB i konsumpcji, co może być potwierdzeniem tezy, iż ożywienie gospodarcze, w tym popyt konsumpcyjny są jednym z czynników generujących inflację. Kolejnym czynnikiem wpływającym na wzrost inflacji jest wzrost podaży pieniądza sprzed 4 kwartałów. Bieżąca stopa referencyjna NBP jest dodatnio skorelowana ze stopą inflacji. Im wyższa jest bieżąca inflacja, tym większe jest prawdopodobieństwo podwyżki stopy referencyjnej NBP. Wzrost stopy referencyjnej sprzed 4 kwartałów z kolei wpływa na spadek inflacji, co jest efektem celu polityki monetarnej NBP.

Wzrostowi stopy referencyjnej towarzyszy krótkoterminowy spadek kursu walutowego (aprecjacja waluty krajowej), co prawdopodobnie wiąże się z przypływem kapitału zagranicznego, przyciągniętego wyższymi stopami procentowymi.

Podaż pieniądza jest najsilniej skorelowana z zadłużeniem zagranicznym, kursem walutowym oraz stopą inflacji sprzed 4 kwartałów.

Kurs walutowy jest skorelowany ze stopą inflacji, stopą referencyjną, podażą pieniądza oraz zadłużeniem zagranicznym. Zidentyfikowano również zależność między kursem walutowym a PKB oraz bezrobociem.

Zadłużenie zagraniczne jest skorelowane z kursem walutowym oraz podażą pieniądza.

Przedstawione zależności obejmują jedynie związki dwustronne między zmiennymi, natomiast nie uwzględniają powiązań wielostronnych.

Badanie zależności za pomocą modeli z klasy ADL

W kolejnym etapie badania dla każdej ze zmiennych zbudowano i oszacowano model ADL, który uwzględnia dynamiczny i wielowymiarowy charakter zależności między zmiennymi. Po odpowiedniej weryfikacji wyeliminowano zmienne objaśniające, które są nieistotne. Zbiorcze wyniki oszacowań parametrów, opóźnień zmiennych oraz wartości *p* są podane w Tabeli 2. W nawiasach ( ) podano rząd opóźnienia zmiennej z danej kolumny. W nawiasach [ ] podano wartość *p*.

Tabela 2. Oszacowania parametrów modeli ADL

	PKB	I	C	REAL	WR	U	CPC	REF	M	ER	ZADL
PKB	-0,16(4) [0,09]	0,08(0) [0,00]	0,23(0) [0,00]			-3,76(0) [0,03]		-0,32(3) [0,00]			
I	1,67(0) [0,00]	-0,32(8) [0,01]			1,17(3) [0,00]						
C	0,51(0) [0,00]		-0,31(5) [0,01]		0,21(0) [0,06]						
REAL							-1(0) [0,00]	1(0) [0,00]			
WR	0,5(0) [0,00]	0,08(1,2) [0,03]			-0,43(1) [0,06]		-0,41(0) [0,00]				
U	-0,02(1) [0,00]	-0,004(2) [0,00]				0,38(1) [0,00]					
CPC			0,1(1) [0,04]		-0,2(0) [0,00]			0,55(0) [0,00] -0,24(4) [0,00]	7,3(1) [0,00] 5,53(2) [0,02] 4,94(4) [0,03]		
REF	0,21(0) [0,00]						0,77(0) [0,00] 0,37(1) [0,00] 0,26(4) [0,00]				
M	0,007(0) [0,03]						0,007(1) [0,06] 0,011(2) [0,00]		-0,34(4) [0,00]		0,17(0) [0,00]
ER	-0,03(1) [0,00]										1,16(0) [0,00]
ZADL										0,6(0) [0,00] 0,23(4) [0,03]	-0,43(4) [0,00]

Źródło: badania własne

Produkt Krajowy Brutto

$$\hat{PKB}_t = -0{,}16PKB_{t-4} + 0{,}08I_t + 0{,}23C_t - 3{,}76U_t - 0{,}32REF_{t-3}, \qquad R^2=0{,}64$$

Na wzrost Produktu Krajowego Brutto wpływa wzrost konsumpcji, wzrost inwestycji oraz spadek bezrobocia. Na zmianę PKB wpływają również decyzje NBP o zmianie stopy referencyjnej. Podwyżka tej stopy hamuje wzrostu PKB po 3 kwartałach.

Polityka monetarna NBP

$$\hat{REF}_t = 0{,}21PKB_t + 0{,}77CPC_t + 0{,}37CPC_{t-1} + 0{,}26CPC_{t-4}, \qquad R^2=0{,}77$$

Decyzje o zmianie stopy referencyjnej NBP są uzależnione w głównej mierze od zmian inflacji (największy jest wpływ inflacji bieżącej). Dynamika PKB jest drugim czynnikiem wpływającym na stopę referencyjną – model potwierdza, iż argumentem branym pod uwagę przy zmianie stopy referencyjnej jest również ożywienie lub spowolnienie gospodarcze.

Bezrobocie

$$\hat{U}_t = -0{,}02PKB_{t-1} - 0{,}004I_{t-2} + 0{,}38U_{t-1}, \qquad R^2=0{,}56$$

Wzrost PKB wpływa na spadek bezrobocia po kwartale, natomiast wzrost inwestycji – po 2 kwartałach.

Stopa inflacji

$$\hat{CPC}_t = 0{,}1C_{t-1} - 0{,}2WR_t + 0{,}55REF_t - 0{,}24REF_{t-4} + 7{,}3M_{t-1} + 5{,}5M_{t-2} + 4{,}9M_{t-4}$$

$$R^2=0{,}80$$

Na wzrost inflacji główny wpływ ma wzrost podaży pieniądza, co wydaje się potwierdzać tezę, że inflacja jest zjawiskiem przede wszystkim monetarnym. Czynnikiem hamującym inflację jest decyzja NBP o podniesieniu stopy procentowej sprzed 4 kwartałów. Pozostałe zmienne objaśniające o mniejszym wpływie to konsumpcja oraz wynagrodzenie realne.

Kurs walutowy, zadłużenie zagraniczne a podaż pieniądza

$$\hat{ER}_t = -0{,}03PKB_{t-1} + 1{,}16ZADL_t, \qquad R^2=0{,}83$$

$$\hat{ZADL}_t = 0{,}6ER_t + 0{,}23ER_{t-4} - 0{,}43ZADL_{t-4}, \qquad R^2=0{,}81$$

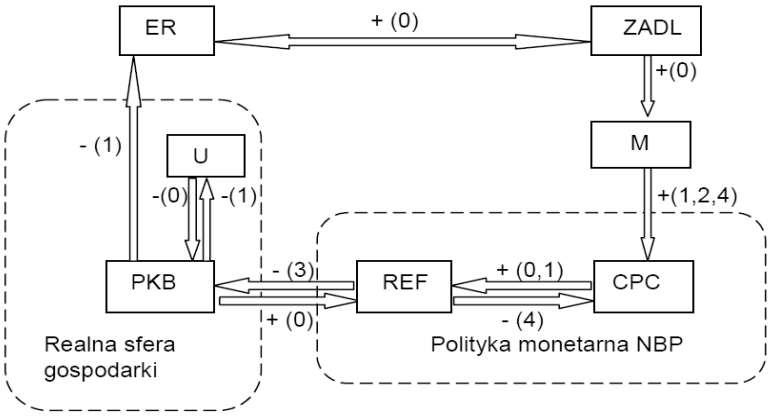
$$\hat{M}_t = 0{,}007PKB_t + 0{,}007CPC_{t-1} + 0{,}011CPC_{t-2} - 0{,}34M_{t-1} + 0{,}17ZADL_t,$$

$$R^2=0{,}59$$

Istnieje dwustronna dodatnia zależność między wzrostem zadłużenia a deprecjacją kursu walutowego. Wynikać to może m. in. z faktu, iż zadłużenie zagraniczne jest podawane w walucie obcej (euro). Czynnikiem wpływającym na aprecjację kursu walutowego jest wzrost PKB i spadek zadłużenia. Wzrost zadłużenia natomiast skutkuje wzrostem podaży pieniądza. W niewielkim stopniu na wzrost podaży pieniądza wpływa również PKB oraz inflacja z poprzednich okresach.

Podsumowanie głównych zidentyfikowanych zależności z uwzględnieniem ich dynamicznego charakteru przedstawiono za pomocą schematu 1. W nawiasie podano liczbę opóźnień.

Schemat 1. Główne wzajemne powiązania między zmiennymi



Źródło: badania własne

Podsumowanie

Wzrost PKB w badanym okresie zależał w głównej mierze od inwestycji, konsumpcji oraz zatrudnienia. Wzrost stopy referencyjnej NBP miał ujemny wpływ na wzrost PKB, który ujawniał się po 3 kwartałach. Na spadek bezrobocia wpływał wzrost PKB oraz inwestycji. Stopa inflacji zależała od podaży pieniądza oraz od stopy referencyjnej NBP i w mniejszym stopniu od popytu konsumpcyjnego. Polityka monetarna NBP, skierowana na bezpośredni cel inflacyjny była skuteczna. Średnio po 4 kwartałach od decyzji Rady Polityki Pieniężnej o podwyżce referencyjnej stopy procentowej następował spadek inflacji, wiązało się to jednak ze spadkiem PKB. Wzrost zadłużenia zagranicznego Polski powodował wzrost podaży pieniądza, co z kolei przekładało się na wzrost inflacji. Według oszacowań parametrów modeli wzrost zadłużenia zagranicznego Polski nie przyczyniał się do wzrostu PKB.

ZAKOŃCZENIE

Przeprowadzona analiza miała na celu zidentyfikowanie dynamicznych powiązań między zmiennymi realnymi (PKB, inwestycje, konsumpcja, stopa bezrobocia) a zmiennymi nominalnymi (podaż pieniądza, stopa inflacji, nominalna stopa procentowa, kurs walutowy i zadłużenie zagraniczne) w polskiej gospodarce na podstawie danych kwartalnych z lat 1997-2011. Z przeprowadzonych badań wynikają następujące wnioski. Większość zidentyfikowanych zależności przejawia się z opóźnieniami w czasie. Wzrost gospodarczy Polski w głównej mierze zależy od czynników należących do realnej strony gospodarki. Na procesy inflacyjne

największy wpływ ma podaż pieniądza. Wzrost zadłużenia zagranicznego kraju powoduje wzrost podaży pieniądza. Decyzje NBP o wzroście stopy referencyjnej hamują inflację, ale i wzrost gospodarczy.

BIBLIOGRAFIA

Bruzda J. (2002) Szybkość reakcji gospodarki na bodźcie monetarne, Wiadomości statystyczne, GUS.

Brzoza-Brzezina M., Kot A., Kłos B., Łyziak T. (2002) Hipoteza neutralności pieniądza, Materiały i Studia NBP nr 142.

Demczuk O., Łyziak T., Przystupa J., Sznajderska A., Wróbel E. (2012) Mechanizm transmisji polityki pieniężnej w Polsce. Co wiemy w 2011 roku? Materiały i Studia NBP nr 270.

Friedman M. (1956) The Quantity Theory of Money - a Restatement, Studies in the Quantity Theory of Money, Chicago Press.

Keynes J.M. (1936) The General Theory of Employment, Interest and Money, Macmillan Cambridge University Press.

King R.G., Watson M.W. (1996) Money, Prices, Interest Rates and the Business Cycle, The Review of Economics and Statistics, vol. 78, No. 1, pp. 35-53.

Maddala G. S. (20008) Ekonometria, Wydawnictwo Naukowe PWN, Warszawa.

Mertens E. (2010) Structural Shocks and the Comovements between Output and Interest Rate, Journal of Economic Dynamics and Control, vol 34, Issue 6, pp 1171-1186.

STUDY ON THE LINKS BETWEEN REAL AND NOMINAL VARIABLES IN THE POLISH ECONOMY IN YEARS 1997-2011

Abstract: The article presents results of the analysis of links between selected real and nominal variables in the Polish economy in years 1997-2011. Most of the identified dependences appear with delays in time. Economic growth depends mainly on real economy factors. GDP growth and investments impact on the decrease of unemployment rate. The rise of central bank reference rate causes a decrease of inflation rate, but has a negative effect on GDP growth. The inflation is a monetary phenomenon, depends mainly on money supply. The increase of polish foreign debt causes money supply growth.

Keywords: real variables, nominal variables, cross-correlogram, ADL models, interest rate, Gross Domestic Product

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 107 – 116

KILKA UWAG DOTYCZĄCYCH STOPY ZWROTU W TERMINIE DO WYKUPU

Andrzej Karpio
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
andrzej_karpio@sggw.pl

Streszczenie: Stopa zwrotu w terminie do wykupu (YTM) jest podstawową i powszechnie stosowaną miarą efektywności inwestycji w papiery dłużne. Wykorzystywana przez praktyków definicja zakłada stałe płatności kuponowe i jednakowe okresy odsetkowe. Są to ograniczenia nierealistyczne. W prezentowanej pracy autor zajmuje się przypadkiem ogólnym, podaje wyrażenia na wartość wewnętrzną obligacji bez wymienionych założeń. W konsekwencji wyprowadza wzór na stopę zwrotu w terminie do wykupu w postaci uwzględniającej zmienne okresy odsetkowe, posiadający własności asymptotyczne zgodne z metodologią wyceny papierów dłużnych i dający możliwość dalszych przybliżeń.

Słowa kluczowe: obligacja, wartość wewnętrzna, krzywa dochodowości, stopa zwrotu w terminie do wykupu, bieżąca stopa zwrotu

WSTĘP

Problem określenia efektywności inwestycji w papiery dłużne jest bardzo istotny z punktu widzenia rynku finansowego, przede wszystkim ze względu na rolę jaką odgrywają na nim instrumenty emitowane przez skarb państwa. Możliwa do osiągnięcia stopa zwrotu, wielkość płatności kuponowych, stopy dyskontowe są podstawą konstrukcji instrumentów dłużnych innych emitentów, którzy opierają się na parametrach charakteryzujących instrumenty skarbowe. Jednym z podstawowych jest stopa zwrotu w terminie do wykupu YTM (ang. *yield to maturity*) definiowana przy założeniu płaskiej struktury terminowej. Na ogół nie da się podać jej dokładnej wartości, wymagane są przybliżenia, a zatem założenia upraszczające definicję. W pracy zrezygnowano z wielu uproszczeń, chociażby z jednakowych okresów odsetkowych, czy też założenia, że w momencie

wyznaczania stopy zwrotu w terminie do wykupu mamy przed sobą pełne okresy odsetkowe. Intencją autora jest uporządkowanie wiedzy na temat YTM i podanie wzorów, zarówno ścisłych, gdy jest to możliwe, jak i przybliżonych, ale w ogólniejszej postaci niż są stosowane w praktyce. Dodatkowo przeprowadzona zostanie dyskusja własności stopy zwrotu w terminie do wykupu. Rozważania zostaną ograniczone do obligacji bez dodatkowych opcji w rodzaju prawa do wcześniejszego wykupu przez emitenta lub prawa do przedstawienia do wcześniejszego wykupu przez inwestora. Również nie będą rozważane obligacje zamienne, te bowiem bardziej przypominają opcje niż instrumenty dłużne.

ROZWAŻANIA WSTĘPNE

W tej części pracy zostaną podane założenia wstępne oraz przeprowadzona dyskusja ogólnego wzoru na wartość wewnętrzną obligacji. Załóżmy, że mamy do czynienia z obligacją wypłacającą odsetki w chwilach przyszłych $t_1, t_2, \ldots, t_N$ (daty). Ponadto, chwilę dzisiejszą, czyli datę wyceny obligacji, oznaczać będziemy symbolem t_0. W tym momencie odsetki mogą być wypłacane, ale nie muszą. Dodatkowo, dla wygody, chwilę końcową t_N oznaczać będziemy symbolem T. Wówczas następuje wykup obligacji oraz wypłata ostatniego kuponu. Przyjęte oznaczenia pozwalają wyrazić w postaci różnic: termin do wykupu $T - t_0$ oraz okresy odsetkowe $\Delta t_{i-1} = t_i - t_{i-1}$, gdzie $i = 1,2,\ldots,N$. Zachodzi wówczas oczywista równość: $\sum_{i=1}^{N} \Delta t_{i-1} = T - t_0$. Przyjmujemy umowę, że wszelkie okresy czasu będą wyrażone w latach. Niech $h_{i-1,i}$ oznacza strumień odsetek (kwotę odsetek odniesioną do roku). Wówczas, w okresie odsetkowym Δt_{i-1} wypłata będzie równa: $h_{i-1,i}\Delta t_{i-1}$. Strumień odsetek wiąże się z nominalną stopą kuponową w okresie od t_{i-1} do t_i związkiem: $h_{i-1,i} = s_{i-1,i}P_M$, gdzie P_M jest wartością nominalną obligacji.

Do wyceny obligacji konieczne jest przyjęcie wymaganych stóp procentowych, związanych z momentami wypłaty odsetek. Będą one grały rolę stóp dyskontowych. Załóżmy zatem, że dysponujemy strukturą terminową zadawaną funkcją $r(t)$. Z definicji $r(t)$ jest nominalną stopą natychmiastową (spot), czyli stopą zwrotu z inwestycji zaczynającej się w chwili obecnej t_0 i kończącej się w chwili przyszłej *t*. Na użytek dalszych rozważań wprowadzimy oznaczenie: $r(t_i) = r_{0,i}$. Nie ograniczając ogólności rozważań możemy założyć, że struktura terminowa jest funkcją ciągłą i ma skończoną wartość graniczną: $\lim_{t\to\infty} = r_{0,\infty}$. Założenie to wynika z praktyki rynkowej. Inwestycje długoterminowe,

na przykład dwudziesto- lub trzydziestoletnie charakteryzują się skończonymi nominalnymi stopami procentowymi, niewiele różniącymi się od stóp kilku lub kilkunastoletnich. Nie czynimy żadnych założeń dotyczących monotoniczności $r(t)$, ale warto przypomnieć standardową terminologię: normalna struktura terminowa występuje wówczas, gdy funkcja jest rosnąca; odwrócona, gdy jest malejąca i płaska, gdy jest to funkcja stała. Pomijamy inne przypadki, gdy $r(t)$ nie jest monotoniczna, chociaż takie sytuacje mają miejsce na rzeczywistych rynkach finansowych. Większość wyników zaprezentowanych w pracy nie zależy od monotoniczności struktury terminowej. Dużo ważniejsze jest przyjęte wyżej założenie o skończonej granicy przy *t* dążącym do nieskończoności.

Standardowe wyrażenie na wartość wewnętrzną obligacji, jako suma zdyskontowanych przepływów finansowych, przybiera postać [Karpio, 2008, 3]:

$$P_{t_0T} = \sum_{i=1}^{N} \frac{h_{i-1,i}\Delta t_{i-1}}{\prod_{m=1}^{i}\left(1 + r_{0,i}\Delta t_{m-1}\right)} + \frac{P_M}{\prod_{m=1}^{N}\left(1 + r_{0,N}\Delta t_{m-1}\right)} \tag{1}$$

gdzie symbol P_{t_0T} w jawny sposób uwzględnia chwilę dzisiejszą t_0 (data wyceny) i końcową *T* (wykup obligacji). Wyrażenie to jest mało wygodne do dalszej dyskusji poza dwoma wnioskami, które nasuwają się natychmiast: wartość wewnętrzna jest malejącą funkcją stóp natychmiastowych $r_{0,i}$ oraz jej wartość, przy zerowych stopach procentowych jest równa: $P_{t_0T} = \sum_{i=1}^{N} h_{i-1,i}\Delta t_{i-1} + P_M$. Zatem wyraża się w postaci sumy płatności kuponowych i wartości nominalnej. Dalsza dyskusja wymaga przekształcenia powyższego wzoru. W tym celu korzystamy z definicji jednookresowych stóp terminowych $f_{i-1,i}$, związanych z okresami odsetkowymi Δt_{i-1} [Karpio, 2008, 4], [Luenberger, 2003]:

$$1 + f_{i-1,i}\Delta t_{i-1} = \frac{\prod_{m=1}^{i}\left(1 + r_{0,i}\Delta t_{m-1}\right)}{\prod_{m=1}^{i-1}\left(1 + r_{0,i-1}\Delta t_{m-1}\right)} \tag{2}$$

Dodatkowo, zastępując strumienie odsetek $h_{i-1,i}$ stopami kuponowymi $s_{i-1,i}$ możemy otrzymać następującą zależność:

$$\frac{P_{t_0T}}{P_M} - 1 = \sum_{i=1}^{N} \frac{\left(s_{i-1,i} - f_{i-1,i}\right)\Delta t_{i-1}}{\prod_{m=1}^{i}\left(1 + r_{0,i}\Delta t_{m-1}\right)} \tag{3}$$

Dyskusję tego wzoru wraz z wnioskami ważnymi z punktu widzenia inwestorów, można znaleźć we wcześniejszej pracy autora [Karpio, 2008, 3]. Warto w tym miejscu zwrócić uwagę na to, że wartość wewnętrzna istotnie zależy od relacji pomiędzy stopami kuponowymi $s_{i-1,i}$ i stopami terminowymi $f_{i-1,i}$. Te ostatnie, w zależności od konstrukcji obligacji, mogą być stopami WIBOR lub rentownością

bonów skarbowych, jeśli rozważamy obligacje o zmiennym oprocentowaniu. Skarb Państwa najczęściej oferuje oprocentowanie powiązane z tymi dwoma parametrami rynkowymi. Oczywiście, nie znamy przyszłych stóp, ani rynkowych, ani kuponowych, ale emitent zawsze deklaruje określoną relację pomiędzy nimi. Zaglądając do listów emisyjnych obligacji skarbowych[1] zauważymy, że obligacje detaliczne mają stopy kuponowe związane ze stopą WIBOR (utożsamianą w tym przypadku z $f_{i-1,i}$) relacją: $s_{i-1,i} = pf_{i-1,i}$, gdzie p jest dla danej obligacji stałym czynnikiem, najczęściej przyjmującym wartości: 0,98, 0,95 itd.. W obecnej sytuacji rynkowej współczynniki p są mniejsze od jedności, z wyjątkiem serii wyemitowanej w lutym 2012 roku, dla której $p = 1$. Warto w tym momencie wspomnieć, że ten przypadek prowadzi do anomalnego zachowania się ceny obligacji w funkcji stóp dyskontowych [Karpio, 2012, 5]. Wspomniana seria TZ0215 jest ostatnią notowaną na Giełdzie Papierów Wartościowych w Warszawie. Kolejne emisje mają identyczną konstrukcję, ale nie są notowane na wolnym rynku. Z kolei oprocentowanie obligacji hurtowych powiązane jest z rentownością bonów skarbowych. Ich związek ze stopami terminowymi opisany jest równością: $s_{i-1,i} = s + f_{i-1,i}$, w której s jest rzędu jednego procenta. W pracy [Karpio, 2008, 3] zamieszczono dyskusję powyższej zależności zakładając ogólną relację: $s_{i-1,i} = s + pf_{i-1,i}$, zatem obejmującą obie grupy wspomnianych obligacji skarbowych. W szczególnym przypadku otrzymujemy obligacje o stałym oprocentowaniu ($p = 0$) lub obligacje zerokuponowe ($s = 0,\ p = 0$).

STOPA ZWROTU W TERMINIE DO WYKUPU

Definicja stopy zwrotu w terminie do wykupu r_{YTM} jest następująca: Jest to stopa spełniająca równanie:

$$\frac{P_{t_0T}}{P_M} - 1 = \sum_{i=1}^{N} \frac{\left(s_{i-1,i} - r_{YTM}\right)\Delta t_{i-1}}{\prod_{m=1}^{i}\left(1 + r_{YTM}\Delta t_{m-1}\right)} \tag{4}$$

w którym P_{t_0T} jest ceną rynkową obligacji lub wartością wewnętrzną, jeśli r_{YTM} potraktujemy jako stopę wymaganą przez inwestora. Rozwiązanie tego równania nastręcza wiele kłopotów i może być otrzymane jedynie w postaci przybliżonej, z wyjątkiem pewnych szczególnych przypadków. Cena jest funkcją ciągłą i monotonicznie malejącą do zera przy stopie r_{YTM} dążącej do nieskończoności, zatem istnieje jednoznaczne rozwiązanie powyższego równania przy zadanej wartości P_{t_0T}. Obserwując notowania obligacji na Giełdzie Papierów Wartościowych w Warszawie można zauważyć, że ich cena niewiele różni się

[1] Listy emisyjne są dostępne na stronach Ministerstwa Finansów: www.mf.gov.pl

od wartości nominalnej. Poniższa tabela podaje notowania obligacji trzyletnich o zmiennym oprocentowaniu. W oczy rzuca się mała płynność, niestety jest to cecha wszystkich obligacji występujących na polskim rynku kapitałowym, w tym również skarbowych.

Tabela 1. Tabela notowań obligacji trzyletnich na GPW w dniu 7 09 2012

Nazwa obligacji	Data ostatniej transakcji	Kurs zamknięcia
TZ0214	2012-09-06	100,19
TZ0215	2012-09-05	100,21
TZ0513	2012-09-03	100,00
TZ0514	2012-09-06	100,30
TZ0813	2012-09-05	99,91
TZ0814	2012-09-05	100,22
TZ1112	2012-08-17	99,90
TZ1113	2012-09-07	100,05
TZ1114	2012-09-07	100,20

Źródło: Giełda Papierów Wartościowych w Warszawie

W konsekwencji nasuwa się przypuszczenie, że rozwiązania równania (4) można poszukiwać przybliżając funkcję $P_{t_0T}(r_{YTM})$ styczną w punkcie, w którym zachodzi równość: $P_{t_0T} = P_M$. Z ciągłości i monotoniczności wynika, że taka wartość argumentu istnieje dokładnie jedna, bowiem wartości rozważanej funkcji zawierają się w przedziale $\left\langle P_M \sum_{i=1}^{N} s_{i-1,i}\Delta t_{i-1} + P_M, 0\right)$. Oznaczając szukaną stopę symbolem r_{YTM}^M (gdzie wskaźnik M nawiązuje do wartości nominalnej P_M) musimy rozwiązać równanie:

$$\sum_{i=1}^{N} \frac{\left(s_{i-1,i} - r_{YTM}\right)\Delta t_{i-1}}{\prod_{m=1}^{i}\left(1 + r_{YTM}\Delta t_{m-1}\right)} = 0 \tag{5}$$

Jeśli stopa kuponowa jest stała ($s_{i-1,i} = s$), to ścisłym rozwiązaniem jest $r_{YTM}^M = s$, w pozostałych przypadkach ponownie musimy zadowolić się rozwiązaniem przybliżonym. Jeśli iloczyny w mianownikach zastąpimy jedynkami, to rozwiązaniem będzie średnia ważona czasem stóp kuponowych:

$$r_{YTM}^M \approx \sum_{i=1}^{N} s_{i-1,i} \frac{\Delta t_{i-1}}{T - t_0} \tag{6}$$

Pochodna ceny względem stopy r_{YTM}^{M} jest równa:

$$P'(r_{YTM}) = -\frac{P_M}{r_{YTM}}\left(1-\frac{1}{\prod_{i=1}^{N}(1+r_{YTM}\Delta t_{m-1})}\right) - \\ -P_M\sum_{i=1}^{N}\left(\frac{(s_{i-1,i}-r_{YTM})\Delta t_{i-1}}{\prod_{m=1}^{i}(1+r_{YTM}\Delta t_{m-1})}\sum_{m=1}^{i}\frac{\Delta t_{m-1}}{1+r_{YTM}\Delta t_{m-1}}\right) \tag{7}$$

Suma po prawej stronie zawiera różnice $s_{i-1,i}-r_{YTM}$, które są bliskie zeru gdy $r_{YTM}=r_{YTM}^{M}$. Średnia stopa kuponowa niewiele różni się od swoich składników, przede wszystkim wówczas, gdy mamy do czynienia ze stabilną sytuacją rynkową. Wówczas stopy kuponowe zmieniają się w niewielkim zakresie, szczególnie w krótkim lub średnim okresie czasu. W konsekwencji, przybliżonemu współczynnikowi kierunkowemu szukanej stycznej można nadać wartość:

$$P'(r_{YTM}^{M}) \approx -\frac{P_M}{r_{YTM}^{M}}\left(1-\frac{1}{\prod_{i=1}^{N}(1+r_{YTM}^{M}\Delta t_{m-1})}\right) \tag{8}$$

Pisząc równanie stycznej i wyznaczając z niej r_{YTM} znajdujemy przybliżoną wartość stopy zwrotu w terminie do wykupu:

$$r_{YTM} \approx r_{YTM}^{M}\left(1-\frac{P_{t_0T}-P_M}{P_M}\left(1-\frac{1}{\prod_{m=1}^{N}(1+r_{YTM}^{M}\Delta t_{m-1})}\right)^{-1}\right) \tag{9}$$

W przypadku obligacji o stałym kuponie $s_{i-1,i}=s=const$ stopa r_{YTM}^{M} jest równa stopie kuponowej s i jest to ścisłe rozwiązanie równania (4). Ponadto, pominięte wyżej wyrazy zawierające różnice $s_{i-1,i}-r_{YTM}$ są dokładnie równe zeru. Zatem w tym przypadku otrzymane przybliżenie stopy zwrotu w terminie do wykupu korzysta jedynie z przybliżenia zależności $P_{t_0T}(r_{YTM})$ styczną.

Przyjrzyjmy się równaniu na r_{YTM}^{M}. Zamiast przybliżonego rozwiązania w postaci średniej ważonej czasem stóp kuponowych możemy skorzystać z innego oszacowania. Załóżmy, że ciąg stóp kuponowych $s_{i-1,i}$ dla $i=1,2,\ldots,N$ jest ograniczony, co jest naturalnym założeniem z punktu widzenia rynku finansowego. Górne i dolne ograniczenia oznaczmy odpowiednio: $s_{\max}=\sup\{s_{i-1,i}\}$, $s_{\min}=\inf\{s_{i-1,i}\}$.

Po skorzystaniu z tożsamości:

$$\sum_{i=1}^{N} \frac{\Delta t_{i-1}}{\prod_{m=1}^{i}\left(1+r_{YTM}\Delta t_{m-1}\right)} = \frac{1}{r_{YTM}}\left(1-\frac{1}{\prod_{m=1}^{N}\left(1+r_{YTM}\Delta t_{m-1}\right)}\right) \tag{10}$$

równanie (4) prowadzi do oszacowania:

$$s_{\min} \le r_{YTM}^{M} \le s_{\max} \tag{11}$$

Przyjęte rozwiązanie w postaci średniej ważonej czasem stóp kuponowych spełnia powyższą nierówność, ale to oszacowanie pozwala we wzorze (9) wstawić za r_{YTM}^{M} jakąkolwiek wartość spełniającą warunek (11), na przykład średnią arytmetyczną kresów ciągu $s_{i-1,i}$ lub obecną wartość stopy kuponowej. Ten drugi wariant jest szczególnie użyteczny z praktycznego punktu widzenia, bowiem nie wymaga znajomości wszystkich stóp kuponowych, w tym prognozowanych, a jedynie tej, która jest właśnie obowiązująca. Ponadto widać dlaczego obligacje o stałym kuponie prowadzą do ścisłego rozwiązania równania (4), wówczas oba kresy są jednakowe. Z otrzymanej zależności wynika dodatkowy ważny wniosek mający zastosowanie dla obligacji bezterminowych lub długoterminowych. Nawet w tym pierwszym przypadku, gdy ciąg stóp kuponowych jest nieskończony, założenie o jego ograniczoności pozostaje w mocy. W konsekwencji, w granicy $N \to \infty$ stopa zwrotu w terminie do wykupu, otrzymana z równania (9), przyjmuje wartość:

$$r_{YTM} = r_{YTM}^{M}\frac{P_{t_0T}}{P_M} \tag{12}$$

Jej jawna postać zależy od przyjętego oszacowania stopy r_{YTM}^{M}. Jeśli mamy do czynienia z obligacjami o stałym kuponie *s*, to otrzymane rozwiązanie prowadzi do stopy zwrotu zadanej wzorem:

$$r_{YTM} = s\frac{P_{t_0T}}{P_M} \tag{13}$$

W przypadku obligacji bezterminowych jest to ścisła wartość, którą można otrzymać bezpośrednio ze wzoru (4) na cenę obligacji, bez jakichkolwiek przybliżeń. Dla obligacji długoterminowych powyższy wzór jest przybliżony, ale często stosowany w praktyce, przede wszystkim ze względu na swoją prostotę.

Na zakończenie warto zwrócić uwagę na następujący aspekt zaprezentowanego rozwiązania równania (5), gdzie czynniki w mianownikach zastąpiono najmniejszą ich wartością, a mianowicie jedynkami. Jeśli poszukujemy lepszego przybliżenia, to możemy uwzględnić następny wyraz rozwinięcia i skorzystać z równości:

$$\frac{1}{\prod_{m=1}^{i}\left(1+r_{YTM}\Delta t_{m-1}\right)} \approx 1-r_{YTM}\left(t_i-t_0\right) \tag{14}$$

Wówczas rozwiązanie będzie miało postać:

$$r_{YTM}^{M} = \frac{\sum_{i=1}^{N} s_{i-1,i}\Delta t_{i-1}}{T - t_0 + \sum_{i=1}^{N} s_{i-1,i}\Delta t_{i-1}\left(t_{i-1} - t_0\right)} \tag{15}$$

Pozostałe wyrażenia na stopę zwrotu w terminie do wykupu pozostaną bez zmian, ale występująca w nich stopa r_{YTM}^{M} będzie miała wartość zadaną powyższym wzorem. Jednak to przybliżenie nie prowadzi do ścisłego rozwiązania równania (5), gdy mamy do czynienia z obligacjami o stałym kuponie. Jednak przykłady liczbowe wykorzystujące dane empiryczne pokazują, że lepsze wyniki daje przybliżenie wzorem (6) lub oszacowanie (11).

INNE PODEJŚCIE DO STOPY YTM

Najczęściej wykorzystywana do obliczeń przybliżona wartość stopy zwrotu w terminie do wykupu wykorzystuje przybliżenie zależności $P_{t_0T}\left(r_{YTM}\right)$ funkcją wymierną postaci:

$$P_{t_0T} = \frac{\alpha + \beta \cdot r_{YTM}}{1 + \gamma \cdot r_{YTM}} \tag{15}$$

Współczynniki α, β, γ dla obligacji o stałym kuponie wyznaczamy z warunków na wartości funkcji w punktach $r_{YTM} = 0$ (wtedy $P_{t_0T} = \sum_{i=1}^{N} h_{i-1,i}\Delta t_{i-1} + P_M$), $r_{YTM} = s$ (wówczas $P_{t_0T} = P_M$) oraz na wartość asymptotyczną przy $r_{YTM} \to +\infty$, przyjmując, że wartość wewnętrzna dąży do $-P_M$. Dwa pierwsze punkty leżą na krzywej, natomiast trzeci jest nierealistyczny, gdyż w rzeczywistości graniczną wartością P_{t_0T} jest zero. Takie podejście daje wzór dobrze sprawdzający się dla rzeczywistych wartości stopy zwrotu w terminie do wykupu niewiele różniącej się od zera i od $r_{YTM} = s$. Ma on postać [Fabozzi i in. 2000], [Luenberger, 2003]:

$$r_{YTM} \approx 2\frac{sP_M + \dfrac{P_M - P_{t_0T}}{T - t_0}}{P_M + P_{t_0T}} \tag{16}$$

Jednak w porównaniu z wyrażeniem zaproponowanym w poprzednim rozdziale, nie prowadzi on do właściwej wartości asymptotycznej charakteryzującej obligacje bezterminowe. Ponadto nie uwzględnia zmiennych okresów odsetkowych, które występują na rzeczywistym rynku. W szczególności, stopa obliczana z powyższego wzoru będzie jednakowa bez względu na to czy odsetki są wypłacane raz w roku, dwa razy, czy w innych okresach jeśli tylko obligacje mają jednakową cenę i taki sam termin do wykupu. Jest to oczywiste uproszczenie, inna jest stopa zwrotu gdy reinwestujemy kapitał częściej, a inna gdy rzadziej. Zatem wzór ten nie uwzględnia

realiów rynku obligacji, niemniej jednak daje dosyć dobre przybliżenie i dlatego jest powszechnie stosowany. Można w tym miejscu zaproponować jego wariant mający zastosowanie do obligacji o zmiennym oprocentowaniu Mianowicie, jeśli drugi wspomniany wyżej warunek zastąpimy nieco ogólniejszym, ale przybliżonym: $r_{YTM} \approx \sum_{i=1}^{N} s_{i-1,i} \frac{\Delta t_{i-1}}{T - t_0}$, dla $P_{t_0 T} = P_M$, to otrzymamy następującą wartość stopy zwrotu w terminie do wykupu:

$$r_{YTM} \approx 2 \frac{P_M \sum_{i=1}^{N} s_{i-1,i} \frac{\Delta t_{i-1}}{T - t_0} + \frac{P_M - P_{t_0 T}}{T - t_0}}{P_M + P_{t_0 T}} \tag{17}$$

Posiada ona te same mankamenty co stopa wyznaczona ze wzoru (16), ale ma ogólniejszą postać uwzględniającą zmienne stopy kuponowe i niejednakowe okresy odsetkowe i również daje dobre przybliżenie. Gdy rozważamy obligacje o stałym kuponie wzór (17) staje się poprzednim wzorem.

UWAGI KOŃCOWE

W podsumowaniu warto wspomnieć, że często stosowany w praktyce wzór przybliżony (16) określający stopę zwrotu w terminie do wykupu [Fabozzi, 2000] ma bardzo ograniczone zastosowanie. Przede wszystkim dlatego, że nie uwzględnia zmiennych okresów odsetkowych, istotnych z inwestycyjnego punktu widzenia. Ponadto, otrzymuje się go przy założeniu, które nie ma nic wspólnego z rzeczywistym asymptotycznym zachowaniem się wartości wewnętrznej: $P_{t_0 T} \to -P_M$, gdy $r_{YTM} \to +\infty$. Zaproponowany przez autora wzór (9) nie ma tych mankamentów. Co więcej, uwzględnia zmienne okresy odsetkowe, ma własności asymptotyczne zgodne z wyceną obligacji, między innymi prowadzi do bieżącej stopy zwrotu (13), będącej ścisłym rozwiązaniem równania (5). Dodatkowo, niejako przy okazji, pojawiła się modyfikacja wyrażenia na stopę YTM uwzględniająca oszacowania zmiennych stóp kuponowych (11). Niewielka ilość miejsca nie pozwala przytoczyć przykładów wykorzystania podanych przez autora wzorów dla obligacji notowanych na rynku finansowym, ale otrzymane przybliżenia okazują się bardzo dobre. W przykładowych oszacowaniach, w zależności od terminu do wykupu, otrzymane wartości różnią się od rzeczywistych znacznie mniej niż jeden procent, do obliczeń przyjęto duży zakres zmienności kuponów i okresów wykupu obligacji. Warto również wspomnieć, że podobne wyrażenia do (9) można otrzymać przy szacowaniu wewnętrznej stopy zwrotu wykorzystywanej przy ocenie projektów inwestycyjnych oraz rzeczywistej rocznej stopy oprocentowania kredytów.

BIBLIOGRAFIA

Fabozzi F., J., Fong G. (2000) Zarządzanie portfelem inwestycji finansowych przynoszących stały dochód, PWN, Warszawa.

Fabozzi F., J. (2000) Rynki obligacji. Analiza i strategie, WIG – PRESS, Warszawa.

Karpio A. (2008) Some Aspects of Debt Securities Valuation In Descrete Time, Optimum. Studia Ekonomiczne, Nr 3 (39), str. 189 – 198.

Karpio A. (2008) Stopy terminowe przy różnych procesach akumulacji kapitału, Matematyczne Aspekty Ekonomii (red. W. Kulpa), Wydawnictwo Uniwersytetu Kardynała Stefana Wyszyńskiego, Warszawa, str. 49 -60.

Karpio A. (2012) Anomalies in bond pricing, 73 Konferencja International Atlantic Economical Society, Istambuł, 28-31 marca 2012.

Luenberger D., G. (2003) Teoria inwestycji finansowych, PWN, Warszawa.

A FEW REMARKS ON YIELD TO MATURITY

Abstract: Yield to maturity is the fundamental and common used measure of investment effectiveness in debt securities. Almost always it is defined under non realistic assumptions, namely that coupon payments are constant and coupon periods are the same. In the presented work author concerns the general case. He gives the expressions for intrinsic value of bonds without simplifying assumptions. The implication of the resignation with simplifying assumption is deriving the formulas for yield to maturity in general case, taking into account various coupon periods.

Key words: bond, intrinsic value, yield curve, yield to maturity, current yield

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 117 – 126

ZASTOSOWANIE KASKAD INFORMACYJNYCH W BADANIU ZACHOWAŃ KONSUMENTÓW

Maria Kaźmierska-Zatoń
Instytut Finansów i Rachunkowości, PWSZ w Skierniewicach
e-mail: zaton@uni.lodz.pl
Wojciech Zatoń
Zakład Bankowości Centralnej i Rynku Pieniężnego, Uniwersytet Łódzki
e-mail: zaton@uni.lodz.pl

Streszczenie: Podejmowanie decyzji wymaga przetworzenia i oceny różnego rodzaju informacji, zarówno informacji prywatnych decydenta, jak również tych wynikających z obserwacji zachowań innych podmiotów. Sekwencyjny proces decyzyjny, w którym decydenci ignorują informacje prywatne a kierują się zachowaniami innych decydentów, prowadzi do powstania kaskady informacyjnej. Jest to proces racjonalnego podejmowania decyzji w oparciu o ograniczoną informację. W artykule obok podstawowego modelu kaskady informacyjnej zaprezentowano też przykład zastosowania tego modelu do symulowania zachowań konsumenckich.

Słowa kluczowe: kaskady informacyjne, proces decyzyjny, zachowania konsumenta

WSTĘP

Dokonywanie wyborów przez decydentów odbywa się zazwyczaj z wykorzystaniem informacji dwojakiego rodzaju: informacji (sygnałów) prywatnych decydenta i jego osobistych przekonań oraz obserwacji zachowań innych decydentów.

Kaskada informacyjna powstaje wówczas, gdy w procesie decyzyjnym informacje prywatne są ignorowane, a decyzje podejmowane na podstawie obserwacji zachowań innych decydentów. Takie zachowanie jest racjonalnie uzasadnione, jeśli decydenci postępują zgodnie z regułą Bayesa.

Eksperymenty laboratoryjne wskazują na dużą skłonność decydentów do tworzenia kaskad informacyjnych (por. [Anderson i Holt 1996, 1997]) Formalna

identyfikacja i analiza kaskad informacyjnych powstających w rzeczywistości jest jednak bardzo trudna ze względu na ograniczone możliwości pomiaru różnych składników realnych procesów decyzyjnych. Mimo to występowanie kaskad informacyjnych można obserwować zarówno w dziedzinie zachowań społecznych (por. np. [Kuran 1997], [Shiller 2005]), wyborów konsumenckich ([Ge i Messinger 2009], [Duan i in. 2009]), czy też w podejmowaniu decyzji na rynkach finansowych i ubezpieczeniowych ([D'Arcy i Oh 1997], [Lee 1998], [Hirshleifer i Teoh 2003], [Heog 2008]).

Celem artykułu jest zaprezentowanie podstawowego modelu kaskady informacyjnej oraz przedstawienie zastosowania modelu kaskady informacyjnej w symulowaniu zachowań konsumentów.

MODEL KASKADY INFORMACYJNEJ

Funkcjonowanie kaskad informacyjnych zostało pierwotnie opisane w celu wyjaśnienia tworzenia się tendencji zachowań w grupach społecznych, zwłaszcza w dziedzinie mody (dwa pionierskie artykuły w tej dziedzinie to: [Banerjee 1992] oraz [Bikhchandani, Hirshleifer i Welch 1992]. Zwrócono uwagę na fakt, że tendencja do zachowań stadnych może wynikać z prostego naśladownictwa i presji społecznej, ale też może być uzasadniona na gruncie racjonalnego wnioskowania z wykorzystaniem ograniczonej informacji.

Eksperymenty laboratoryjne (por. [Anderson i Holt 1996, 1997]) wykazały, że kaskady informacyjne mogą powstać łatwo i szybko. Jeden z nich [Anderson i Holt 1997], w którym uczestniczyła grupa studentów, polegał na losowaniu kul z dwóch urn. Prowadzący eksperyment przygotował dwie urny: UC – zawierającą dwie kule czerwone i jedną niebieską, UN – zawierającą dwie kule niebieskie i jedną czerwoną. Początkowo, w sposób losowy, wybrana została jedna z tych urn (uczestnicy eksperymentu mogli obserwować losowanie, ale nie byli w stanie odróżnić, która z urn została wylosowana), Kolejni studenci losowali po jednej kuli z urny (kolor wylosowanej kuli był niewidoczny dla reszty uczestników) i ogłaszali swoją decyzję, czy sądzą, że jest to urna UC, czy UN. Następnie kula była zwracana do urny.

Wykorzystanie reguły Bayesa do formułowania sądów o prawdopodobnym typie urny, z której odbywa się losowanie jest racjonalnym postępowaniem prowadzącym do powstawania kaskad informacyjnych.

Jeśli pierwsza osoba wylosowała kulę czerwoną to wskaże ona, że losowanie odbywa się z urny UC. Jeśli również druga osoba wylosowała kulę czerwoną to i ona wskaże, że losowanie odbywa się z urny UC. Każda następna osoba wskaże na urnę UC niezależnie od koloru kuli, którą wylosuje (zignoruje sygnał prywatny). Powstanie zatem kaskada informacyjna.

Jeśli druga osoba wylosowała kulę niebieską to, wiedząc że pierwsza osoba wylosowała kulę czerwoną[1], znajdzie się w sytuacji, w której oba sygnały się równoważą. W takim przypadku druga osoba podejmie decyzję o wskazaniu rodzaju urny w sposób losowy. Z prawdopodobieństwem równym 0,5 wskaże wtedy urnę UC tym samym zapoczątkowując kaskadę informacyjną[2]. Z kolei wskazując urnę UN, druga osoba zrównoważy decyzję osoby pierwszej, stawiając trzecią losującą osobę w sytuacji podejmowania decyzji wyłącznie na podstawie koloru wylosowanej przez nią kuli (analogicznie, jak w przypadku pierwszego decydenta). W tym przypadku po wyborach dwóch decydentów kaskada informacyjna nie powstanie.

Analogiczne wnioskowanie prowadzące do wskazywania urny UN będzie uzasadnione w sytuacji, jeśli pierwsza osoba wylosowałaby kulę niebieską.

W podstawowym modelu kaskady informacyjnej (por. [Banerjee 1992], [Bikhchandani, Hirshleifer i Welch 1992], a także [Kaźmierska-Zatoń i Zatoń 2012]) zakłada się, że kolejni decydenci sekwencyjnie podejmują decyzje dotyczące wyboru jednej z dwóch opcji: A lub R w warunkach dwóch możliwych stanów natury: D i Z. Znane są prawdopodobieństwa a priori wystąpienia tych stanów. Wybór opcji A oznacza podjęcie działania korzystnego w sytuacji, gdy rzeczywistym stanem jest stan D, wybór opcji R jest lepszy w warunkach stanu Z. Każdy decydent podejmuje decyzję na podstawie prywatnego sygnału sugerującego bieżący stanu natury oraz obserwacji decyzji podjętych wcześniej przez innych decydentów. Prywatny sygnał może pojawić się w jednym z dwóch wariantów: H i L. Sygnał H wskazuje, że lepszą opcją jest A (jeśli stan natury jest D), sygnał L wskazuje, że lepszym wyborem jest R (jeśli stanem natury jest Z). Zakładając że, $p > 0,5$ prawdopodobieństwa warunkowe otrzymania sygnałów H i L w zależności od stanu rzeczywistości przedstawia tabela 1.

Tabela 1. Prawdopodobieństwa pojawienia się sygnałów H i L w zależności od stanu natury ($p > 0,5$)

Sygnał	Stan	
	D	Z
H	*p*	*1-p*
L	*1-p*	*p*

Źródło: opracowanie własne

[1] Formalnie druga osoba wie jedynie, że pierwsza wskazała urnę UC, ale jeśli z założenia każdy z uczestników podejmuje decyzje zgodnie z regułą Bayesa to jest pewne, że pierwsza osoba wylosowała kulę czerwoną.

[2] Warto zwrócić uwagę, że w tej sytuacji powstanie i trwanie kaskady informacyjnej polegającej na wskazywaniu urny UC przez kolejnych uczestników może odbywać się w warunkach, kiedy tylko pierwszy uczestnik wylosował kulę czerwoną, pozostali zaś kule niebieskie.

Wartość p można zatem interpretować jako siłę sygnału – wartość bliska 1 oznacza dużą wartość informacyjną i wiarygodność sygnału.

Zakładając, że prawdopodobieństwa a priori wystąpienia stanów D i Z są równe ($P(D) = P(Z) = 0{,}5$), warunkowe prawdopodobieństwo występowania stanu D przy otrzymaniu sygnału H wyniesie:

$$P(D \mid H) = \frac{P(D) * P(H \mid D)}{P(D) * P(H \mid D) + P(Z) * P(H \mid Z)} = \frac{0{,}5 * p}{0{,}5 * p + 0{,}5 * (1-p)} = p \qquad (1)$$

Ze względu na typ wybieranej opcji można wyróżnić dwa rodzaje kaskad: wstępującą KA oraz zstępującą KR. KA powstaje, kiedy kolejni decydenci wybierają opcję A, zaś KR – kiedy kolejni decydenci wybierają opcje R.

Prawdopodobieństwo powstania kaskady typu KA (podobnie jak kaskady KR) po decyzjach dwóch osób (n=2, przy założeniu, że oba stany D i Z są jednakowo prawdopodobne) wyniesie[3]:

$$P(KA, n = 2) = P(KR, n = 2) = \frac{1 - p + p^2}{2} \qquad (2)$$

Prawdopodobieństwo powstania kaskady typu KA lub KR po decyzjach n osób można obliczyć ze wzoru (3)[4]:

$$P(KA, n) = P(KR, n) = \frac{1 - \left(p - p^2\right)^{\frac{n}{2}}}{2} \qquad (3)$$

Tabela 1 zawiera wartości prawdopodobieństw powstania kaskad informacyjnych (łącznie KA i KR) w zależności od liczby decydentów n i wartości prawdopodobieństwa p (siły informacyjnej prywatnego sygnału).

[3] Kaskada KA powstanie po decyzjach dwóch decydentów, jeśli zajdzie jeden z niżej wymienionych przypadków (w nawiasach podano prawdopodobieństwa wystąpienia zdarzeń):

- przy stanie D obaj decydenci otrzymali sygnały H ($p * p$);
- przy stanie D pierwszy decydent otrzymał sygnał H, drugi sygnał L i losowo (z prawdopodobieństwem 0,5) wybrał opcję A ($p * (1-p) * 0{,}5$);
- przy stanie Z obaj decydenci otrzymali sygnały H ($(1-p) * (1-p)$);
- przy stanie Z pierwszy decydent otrzymał sygnał H, drugi sygnał L i losowo (z prawdopodobieństwem 0,5) wybrał opcję A ($(1-p) * p * 0{,}5$).

[4] por. Bikhchandani, S., Hirshleifer, D., Welch, I. (1992) A theory of fads, fashion, custom, and cultural change as informational cascades, The Journal of Political Economy, vol. 100, no. 5, str. 998.

Tabela 1. Prawdopodobieństwa powstania kaskady informacyjnej (łącznie KA i KR) w zależności od *n* i *p*

Liczba decydentów *n*	*p* = 0,55	*p* = 0,65	*p* = 0,75	*p* = 0,85	*p* = 0,95
2	0,753	0,773	0,813	0,873	0,953
3	0,877	0,891	0,919	0,954	0,990
4	0,939	0,948	0,965	0,984	0,998
5	0,970	0,975	0,985	0,994	1,000
6	0,985	0,988	0,993	0,998	1,000
7	0,992	0,994	0,997	0,999	1,000
8	0,996	0,997	0,999	1,000	1,000
9	0,998	0,999	0,999	1,000	1,000
10	0,999	0,999	1,000	1,000	1,000

Źródło: obliczenia własne

Z rezultatów w tabeli 1 wynika, że prawdopodobieństwo powstania kaskady informacyjnej jest bardzo duże i już przy kilku decydentach rośnie do 1, nawet przy stosunkowo słabej sile informacyjnej sygnału (niskiej wartości *p*).

Drugim istotnym aspektem kaskad informacyjnych (obok łatwości ich powstawania) jest ich poprawność. Jeśli prawdziwym stanem natury jest np. stan D to kaskadą poprawną (KP) będzie kaskada typu KA, błędną (KB) - kaskada KR. Prawdopodobieństwo powstania kaskady poprawnej po decyzjach dwóch osób wyniesie[5]:

$$P(KP \mid D, n=2) = P(KA \mid D, n=2) = p * p + p * (1-p) * 0{,}5 = \frac{p * (p+1)}{2} \quad (4)$$

Prawdopodobieństwo powstania kaskady błędnej będzie natomiast określone wzorem (5):

$$P(KB \mid D, n=2) = P(KR \mid D, n=2) = (1-p) * (1-p) + (1-p) * p * 0{,}5 = \frac{(p-2) * (p-1)}{2} \quad (5)$$

Po decyzjach n osób omawiane powyżej prawdopodobieństwa będą określone wzorami (6) i (7)[6]:

$$P(KP \mid D, n) = \frac{p * (p+1) * [1 - (p - p^2)^{\frac{n}{2}}]}{2 * (1 - p + p^2)} \quad (6)$$

[5] Wzory (4) i (5) wynikają wprost z wcześniej analizy dotyczącej powstawania kaskad KA.
[6] por. Bikhchandani, S., Hirshleifer, D., Welch, I. (1992) A theory of fads, fashion, custom, and cultural change as informational cascades, The Journal of Political Economy, vol. 100, no. 5, str. 998.

$$P(KB \mid D,n) = \frac{(p-2)*(p-1)*[1-(p-p^2)^{\frac{n}{2}}]}{2*(1-p+p^2)} \quad (7)$$

W tabeli 2 przedstawiono prawdopodobieństwa ukształtowania się poprawnych i błędnych kaskad informacyjnych w zależności od siły informacyjnej sygnału p oraz liczby decydentów n.

Tabela 2. Prawdopodobieństwa powstawania poprawnych i błędnych kaskad informacyjnych w zależności od siły sygnału p i liczby decydentów n

Liczba decydentów n	p = 0,55		p = 0,75		p = 0,95	
	KP	KB	KP	KB	KP	KB
2	0,426	0,326	0,656	0,156	0,926	0,026
3	0,497	0,380	0,742	0,177	0,962	0,027
4	0,532	0,407	0,779	0,186	0,970	0,027
5	0,549	0,420	0,795	0,189	0,972	0,028
6	0,558	0,427	0,802	0,191	0,972	0,028
7	0,562	0,430	0,805	0,192	0,972	0,028
8	0,564	0,432	0,807	0,192	0,972	0,028
9	0,565	0,433	0,807	0,192	0,972	0,028
10	0,566	0,433	0,808	0,192	0,972	0,028

Źródło: obliczenia własne

Wyniki zawarte w tabeli 2 wskazują, że zgodnie z oczekiwaniami, prawdopodobieństwo powstania poprawnej kaskady rośnie wraz ze wzrostem wiarygodności sygnału p i liczby decydentów n. Stosunkowo wysokie jest jednak prawdopodobieństwo powstania błędnej kaskady, które maleje wyraźnie dopiero wtedy, gdy siła informacyjna sygnału jest bardzo wysoka.

MODEL KASKADY INFORMACYJNEJ W SYMULOWANIU ZACHOWAŃ KONSUMENTÓW

Przedstawiony podstawowy model kaskady informacyjnej jest prosty i jest mało prawdopodobne, żeby w takiej formie mógł wystąpić w rzeczywistości. W tej części artykułu przedstawiono propozycję bardziej skomplikowanego modelu będącego bardziej realistyczną ilustracją kaskad informacyjnych[7]. Może on

[7] Model o zbliżonej konstrukcji (uwzględniający dodatkowo przekazaywanie informacji o jakości produktu (obejrzanego filmu) między dwoma kolejnymi decydentami) został przedstawiony w:De Vany, A., Lee, C. (2001) Quality signals in information cascades and the dynamics of the distribution of motion picture box office revenues, Journal of Economic Dynamics and Control, vol. 25, str. 593-614. Model przedstawiony w artykule można zaklasyfikować, jako pośredni wariant między modelem podstawowym i przedstawionym we wspomnianej pracy.

posłużyć do opisu niektórych wyborów konsumenckich, zwłaszcza, gdy publicznie dostępne są oceny popularności i udziałów w rynku różnych produktów, jak i opinie pojedynczych konsumentów (np. dostępne w internecie).

Założenia modelu są następujące:

1. Konsumenci na rynku określonej grupy produktów podejmują sekwencyjnie decyzje o wyborze jednego z dostępnych produktów. Podjęta decyzja może polegać również na rezygnacji z wyboru jakiegokolwiek produktu.
2. Na rynku dostępne są cztery produkty, z których dwa są dobrej jakości (D1 i D2) i dwa złej (Z1 i Z2). Prawdopodobieństwa a priori wyboru dobrego i złego produktu są zatem równe i wynoszą 0,5.
3. Każdy z konsumentów otrzymuje również prywatny sygnał wskazujący na prawdopodobną jakość ostatnio wybranego produktu. Prawdopodobieństwo p określa siłę informacyjną sygnału zgodnie z założeniami modelu przedstawionego w poprzedniej części artykułu[8].
4. Prowadzona jest lista rankingowa produktów będąca miarą popularności każdego z produktów i jego udziałów w rynku. Jest ona uaktualniana po wyborze każdego produktu i stanowi informację publiczną dostępną wszystkim konsumentom.
5. Przed wyborem pierwszego konsumenta wybierany jest losowo jeden z czterech produktów. Zgodnie z jakością tego produktu i przyjętym prawdopodobieństwem p, pierwszy konsument otrzymuje prywatną informację sugerującą jakość tego produktu. W zależności od otrzymanego sygnału pierwszy konsument podejmuje decyzje o wyborze produktu lub nie.
6. Kolejny konsument $j+1$ może podjąć następujące decyzje:
   a. Jeśli konsument j wybrał produkt x, zaś konsument $j+1$ otrzymał informację prywatną wskazującą na dobrą jakość tego produktu (sygnał H) to również wybierze ten produkt.
   b. Jeśli konsument j wybrał produkt x, zaś konsument $j+1$ otrzymał informację prywatną wskazującą na słabą jakość tego produktu (sygnał L) to w sposób losowy (z prawdopodobieństwem 0,5) podejmie jedną z dwóch decyzji: zdecyduje się na produkt x, który wybrał konsument j lub dokona wyboru produktu na podstawie listy rankingowej[9].
   c. Jeśli konsument j nie wybrał produktu, zaś konsument $j+1$ otrzymał informację prywatną wskazującą na dobrą jakość ostatnio wybieranego produktu (sygnał H) to w sposób losowy (z prawdopodobieństwem 0,5) podejmie jedną z dwóch decyzji: wybierze produkt na podstawie listy rankingowej lub nie wybierze żadnego produktu.

[8] Por. tabela 1. Stan D oznaczać będzie w tym przypadku dobry produkt (D1 lub D2), Z zły (Z1 lub Z2).

[9] Wybór z listy rankingowej jest dokonywany w sposób losowy zgodnie z popularnością produktów (liczbą dotychczasowych wyborów danego produktu).

d. Jeśli konsument *j* nie wybrał produktu, zaś konsument *j+1* otrzymał informację prywatną wskazującą na słabą jakość ostatnio wybieranego produktu (sygnał L) to nie wybierze żadnego produktu.

W celu analizy przedstawionego modelu przeprowadzono symulacje na próbie 1000 fikcyjnych konsumentów i czterech fikcyjnych produktów. Dla każdej z wybranych wartości siły sygnału *p* wykonano 1000 symulacji. Wyniki dotyczące wyborów konsumentów zawiera tabela 3.

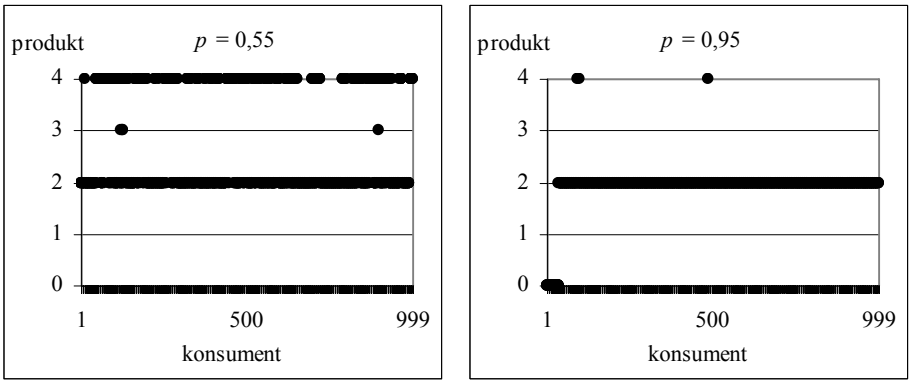
Tabela 3. Wybór produktów w modelu kaskady informacyjnej zachowań konsumenckich (w % ogółem)

Siła sygnału *p*	Odsetek wyborów produktów					
	D1	D2	Z1	Z2	D1+D2	Z1+Z2
0,55	30,3	35,4	16,0	18,3	65,7	34,3
0,60	36,5	36,3	14,2	13,0	72,8	27,2
0,65	42,4	43,5	6,6	7,5	85,9	14,1
0,70	45,7	46,5	4,5	3,3	92,2	7,8
0,75	50,0	45,8	2,4	1,8	95,8	4,2
0,80	49,4	47,7	1,1	1,8	97,1	2,9
0,85	48,9	49,9	0,5	0,7	98,8	1,2
0,90	49,7	49,7	0,3	0,3	99,4	0,6
0,95	49,4	50,0	0,4	0,2	99,4	0,6

Źródło: obliczenia własne

Wyniki zawarte w tabeli 3 potwierdzają, że pomimo zwiększenia możliwości decyzyjnych w stosunku do podstawowego modelu kaskady informacyjnej nadal istnieją silne tendencje do formowania się kaskad informacyjnych wraz ze wzrostem siły informacyjnej prywatnych sygnałów.

Rysunek 1. Przykładowe wybory konsumentów dla dwóch różnych wartości sygnału *p*.



Źródło: obliczenia własne.

Uwagi: oznaczenia osi pionowej:
0 – brak wyboru, 1 – wybór D1, 2 –wybór D2, 3 – wybór D3, 4 – wybór D4.

Na rysunku 1 przedstawione zostały przykładowe wybory konsumentów w dwóch symulacjach dla różnych wartości p. W przypadku słabej wartości informacyjnej sygnału widać duże zróżnicowanie wyborów konsumentów, w drugim przypadku kaskada informacyjna zostaje silnie utrwalona i naruszana jest bardzo sporadycznie.

Przykładowa weryfikacja przedstawionego modelu wykorzystuje dane fikcyjne, wobec tego trudno formułować jednoznaczne wnioski na temat wykorzystania tego modelu w praktyce. Planowane są dalsze prace w kierunku wykorzystania danych rzeczywistych.

PODSUMOWANIE

Przedstawiony model jest, mimo wprowadzonych rozszerzeń, stosunkowo dużym uproszczeniem rzeczywistości. Uzyskane wyniki potwierdzają jednak silne tendencje do tworzenia się kaskad informacyjnych w sytuacjach, gdy wybory są podejmowane w dużej mierze na podstawie obserwacji zachowań innych decydentów. Jednym z kroków w kierunku rozbudowy modelu, mogłoby być większe zróżnicowanie jakości produktów (nie tylko podział dychotomiczny na produkty dobre i złe). Innym udoskonaleniem byłoby zróżnicowanie korzyści z podejmowania różnych decyzji przez decydentów w zależności od tego ilu decydentów wcześniej podjęło dane decyzje (por. np. [Gill 2008]). Korzyści z nabycia np. nowoczesnego produktu mogą w oczywisty sposób wynikać z faktu, ilu konsumentów wcześniej nabyło ten produkt[10].

Przedstawione propozycje modeli opisują procesy decyzyjne, a dostępne dane empiryczne dotyczą zazwyczaj efektów tych procesów. Uzyskanie danych z pomiarów różnych składników realnych procesów decyzyjnych jest bardzo trudne, czasami wręcz niemożliwe. Na sposób przetwarzania informacji może też w dużym stopniu wpływać struktura procesu decyzyjnego. Stanowić to może istotne ograniczenie zastosowań proponowanych modeli kaskad informacyjnych.

BIBLIOGRAFIA

Anderson, L., Holt, C. (1996) Classroom Games: Information Cascades, Journal of Economic Perspectives, vol.10, str.187-193.

Anderson, L., Holt, C. (1997) Information Cascades in the Laboratory, American Economic Review, vol. 87, str.847-862.

Banerjee, A. (1992) A simple model of herd behavior, Quarterly Journal of Economics, vol. 107, no 3, str.797-817.

[10] Klasycznymi przykładami w tym względzie są faksy i telefony komórkowe.

Bikhchandani, S., Hirshleifer, D., Welch, I. (1992) A theory of fads, fashion, custom, and cultural change as informational cascades, The Journal of Political Economy, vol. 100, no. 5, s. 992-1026.

Bikhchandani, S., Hirshleifer, D., Welch, I. (1998) Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades, Journal of Economic Perspectives, vol. 12, no. 3, str. 151-70.

De Vany, A., Lee, C. (2001) Quality signals in information cascades and the dynamics of the distribution of motion picture box office revenues, Journal of Economic Dynamics and Control, vol. 25, str. 593-614.

Duan W., Gu B., Whinston A. (2009) Informational cascades and software adoption on the internet: an empirical investigation, MIS Quarterly, vol. 33, no. 1, str 23-48.

D'Arcy, S., Oh, P. (1997) The Cascade Effect in Insurance Pricing, Journal of Risk and Insurance, vol. 64, no. 3, str. 465-480.

Ge, X., Messinger, P., Li, J. (2009) Influence of Soldout Products on Consumer Choice, Journal of Retailing, vol. 85, no. 3, str. 274–287.

Gill, G. (2008) Resonance within the Multi-Client System: Criticality, Cascades and Tipping Points. Informing Science, vol. 11, str. 311-348

Heog S. (2008) Informational cascade in the insurance market, The Journal of Risk and Insurance, vol. 75, no. 1, str.145-165.

Hirshleifer, D., Teoh, S. (2003) Herd Behaviour and Cascading in Capital Markets: a Review and Synthesis, European Financial Management, vol. 9, str. 25-66.

Kaźmierska-Zatoń M., Zatoń W. (2012) Model kaskad informacyjnych w procesie podejmowania decyzji, Zeszyty Naukowe UE w Poznaniu, tom: Matematyka i informatyka na usługach ekonomii. Informatyka ekonomiczna, w druku.

Kuran T. (1997) Private truths and public lies: The social consequences of preference falsification, Cambridge, Harvard University Press.

Lee, I. (1998) Market crashes and informational avalanches, Review of Economic Studies, vol. 65, str. 741-759.

Shiller, R. (2005) Irrational exuberance, Princeton University Press, Princeton, New Jersey.

INFORMATIONAL CASCADES APPLIED TO ANALYSIS OF CONSUMERS BEHAVIOUR

Abstract: If in the decision-making process people ignore their private information and make decisions mostly based upon what others did then informational cascade may quickly arise. The article presents the basic model of informational cascades which can be easily confirmed in experimental studies. The more realistic model that can be applied to analysis of consumers' behaviour is also presented.

Keywords: informational cascades, decisions making, consumers behaviour

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 127 – 136

WYKORZYSTANIE ZANURZANIA OBSERWACJI W PRÓBIE DO KONSTRUKCJI KART KONTROLNYCH

Małgorzata Kobylińska
Katedra Metod Ilościowych
Uniwersytet Warmińsko-Mazurski w Olsztynie
angosiak@poczta.onet.pl

Streszczenie: Sterowanie procesem produkcyjnym wymaga monitorowania obserwowanych odchyleń zmiennych diagnostycznych od wartości celowych. Podstawową rolę w działaniach sterowania jakością odgrywają karty kontrolne, które są jednym z podstawowych narzędzi statystycznej kontroli jakości. Należą do bieżących metod kontroli, jakości, a jednocześnie przy właściwym stosowaniu mają duże znaczenie przy poprawie jakości produkcji. W 1975 roku zostało wprowadzone, przez Tukey'a pojęcie zanurzania obserwacji w próbie wielowymiarowej. Miara zanurzania może być wykorzystywana do klasyfikacji obserwacji wielowymiarowych względem ich oddalenia od centralnego skupienia oraz do wykrywania i eliminacji obserwacji nietypowych. W pracy zaprezentowano wykorzystanie miary zanurzania obserwacji w próbie do konstrukcji kart kontrolnych.

Słowa kluczowe: statystyczne sterowanie jakością procesów, karty kontrolne, zanurzania obserwacji w próbie

WPROWADZENIE

Nowe technologie umożliwiają wytwarzanie produktów charakteryzujących się coraz wyższym poziomem użyteczności. Ważną rolę w funkcjonowaniu nowoczesnego przedsiębiorstwa odgrywa sprawnie działający system zarządzania jakością usług i wyrobów. Przedsiębiorstwo produkcyjne dąży do tego, żeby jak najlepszej jakości były produkty, które wytwarza. Jakość produktu decyduje o jego przydatności dla klienta oraz satysfakcji klienta z produktu, który został przez niego nabyty. Dobra jakość wytwarzanych przez przedsiębiorstwo produktów lub świadczonych przez nie usług pozwala na utrzymanie się na rynku, na którym konkurencja jest bardzo duża. Na proces produkcyjny oddziałują pewne czynniki,

które są przyczyną jego zmienności. Na podstawie kontroli przebiegu charakterystyk wartości zaobserwowanych parametrów danego procesu można dokonać analizy jego zmienności.

Chcąc skutecznie zarządzać jakością przedsiębiorstwo powinno wyznaczyć stan ilościowy jakości produktu. W tym przypadku konieczne jest ustalenie zbioru cech (właściwości), według których będzie przeprowadzana ocena produktu lub usługi. Właściwości te mogą być wyrażone za pomocą wartości liczbowych, lub mogą odpowiadać im pewne kategorie obrazujące dany stan rzeczy [Hamrol 2008].

Sterowanie jakością procesów produkcyjnych obejmuje miedzy innymi: opracowanie planów jakości procesów i wyrobów, zarządzanie konfiguracją i ryzykiem w procesach produkcyjnych, nadzorowanie jakości oraz ocenę zdolności procesów produkcyjnych [Szkoda 2004].

Ważną rolę w sterowaniu jakością procesów produkcyjnych odgrywają metody statystyczne, wśród nich metody statystyki opisowej i matematycznej. Statystyczne sterowanie procesem produkcyjnym (SPC) ma na celu poprawę jego jakości poprzez realizowaną w rzeczywistym czasie kontrolę jego przebiegu, w celu wykrycia ewentualnych rozregulowań. Zmienność procesu może być wywołana przez dwa rodzaje przyczyn: losowe lub specjalne. Pierwsze z nich są stale obecne w procesie, nieuniknione w danych warunkach oraz nie są łatwo rozpoznawalne, kolejne – pojawiają się nieregularnie i trudno przewidzieć ich skutki, mogą być stałym elementem procesu lub pojawiać się w nim przypadkowo [Wiśniewska, Malinowska 2011].

Celem statystycznego testowania procesem produkcyjnym jest doprowadzenie go do stabilnego poziomu, utrzymanie go na tym poziomie oraz zapewnienie spełnienia założonych wymagań przez wyroby i usługi. Proces produkcyjny nazywamy stabilnym, jeżeli czynniki wpływające na jego przebieg maja charakter wyłącznie losowy [Sałaciński 2009].

KARTY KONTROLNE JEDNOWYMIAROWE I WIELOWYMIAROWE

Karty kontrolne zostały wprowadzone przez W. A. Shewharta w 1924 roku i stały się niezwykle użyteczne zarówna w produkcji jak i usługach. Ich wykorzystanie do monitorowania procesów produkcyjnych umożliwia znaczną redukcję zmienności oraz prowadzi do wzrostu jakości produkcji. Zadaniem kart kontrolnych jest dostarczenie informacji, wykorzystując metody graficzne, o stabilności procesu, czyli czy wybrane do kontroli danych cech parametry statystyczne „nie zmieniają się” w czasie [Hamrol 2008].

W kartach kontrolnych Shewharta wykorzystywane są prawa rachunku prawdopodobieństwa i statystyki matematycznej. Jeżeli jakiekolwiek sygnały świadczą o rozregulowaniu procesu produkcji, można podjąć decyzje o jego przerwaniu i skorygowaniu jego przebiegu.

Karty kontrolne konstruowane dla zmiennych ilościowych mogą monitorować informacje o pojedynczych właściwościach procesu (karty kontrolne

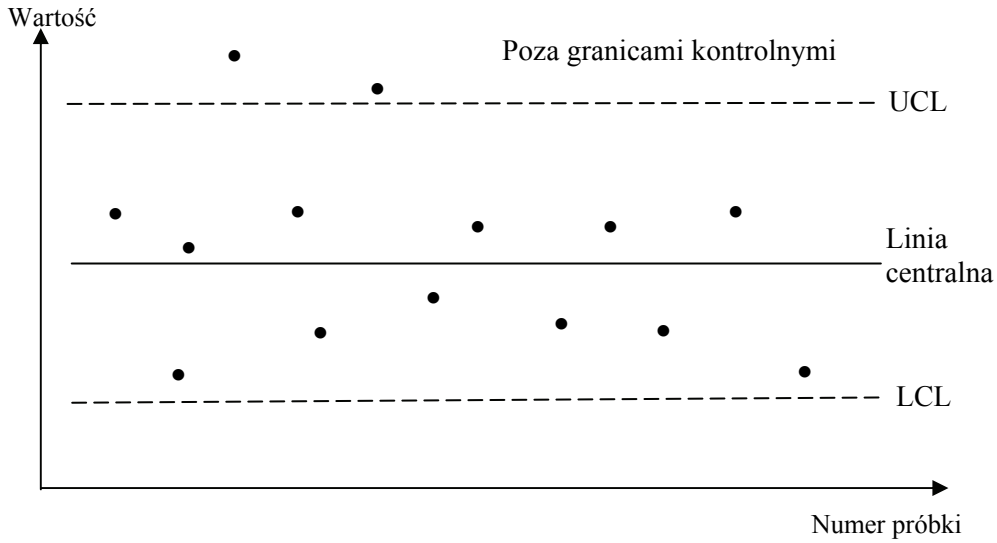
jednowymiarowe) lub wielu jego właściwościach (karty kontrolne wielowymiarowe). Umożliwiają monitorowanie zmian charakterystyk procesu dotyczących położenia, rozproszenia lub jego stabilności [Płaska 2000].

Klasyfikacji kart kontrolnych dokonać można ze względu na liczbę rozpatrywanych cech, rodzaj rozkładu analizowanej cechy lub jej charakter. Wśród klasycznych kart kontrolnych wyróżniamy karty kontrolne dla cech dyskretnych (zliczeniowych) oraz cech ciągłych (mierzalnych). Karty sekwencyjne (karta MA, EWMA, IX-MR, CUSUM) wykorzystywane są do wykrywania pojawiających się trendów oraz niewielkich przesunięć wartości średniej procesu produkcyjnego. Szczegółowo typy kart kontrolnych zostały omówione w pracy Sałacińskiego [Sałaciński 2009].

Karta kontrolna jako narzędzie wspomagające regulację procesów produkcji jest wykresem przedstawiającym statystyki np. wartość średnią, odchylenie standardowe w czasie lub według numeru kolejnej próbki pobranej do badania. Głównym elementem każdej karty jest diagram kontrolny, na którym nanoszone są:

- Linia centralna (CL) – jest zwykle średnią z rozważanych danych procesu,
- Linie kontrolne - określane jako górna granica kontrolna (UCL) oraz dolna granica kontrolna (LCL) (w zależności od potrzeb badania stosuje się jednostronny lub dwustronny obszar krytyczny). W wielu przypadkach wyznacza się także granice ostrzegawcze.

Rysunek 1. Karta kontrolna



Źródło: Aczel [2000]

Linie kontrolne wyznaczają zakres zmienności monitorowanej statystyki procesu, tzn. zakres, której przyczyną są tylko czynniki losowe. Jeżeli badana

statystyka przekracza granice kontrolne lub gdy jej wykres wykazuje pewne stałe zmiany, proces może być nieustabilizowany i powinien być skorygowany [Aczel 2000].

W rzeczywistości często mamy do czynienia z procesami produkcji, do opisu których wykorzystuje się kilka właściwości, które mogą być od siebie niezależne lub powiązane ze sobą liniowo. W przypadku monitorowania dwóch właściwości procesu wykorzystać można elipsy kontrolne. Jednak na ich wykresie nie ma możliwości zaobserwowania trendów, wzorców, tendencji procesu oraz jego zmienności w czasie.

W celu monitorowania w czasie parametrów statystycznych wybranych do kontroli wielu cech procesu można wykorzystać kilka kart kontrolnych. Niesie to za sobą kilka wad np. konieczność konstrukcji i analizowania znacznej liczby kart kontrolnych lub pomijanie zależności pomiędzy cechami. Ocena procesu ze względu na każdą z właściwości oddzielnie może być myląca, w związku z wzajemnymi powiązania między nimi [Demski 2010].

Pomocnym narzędzie w celu monitorowania procesów wielowymiarowych jest karta T^2 Hotellinga, na której zamiast zaobserwowanych w czasie wartości przedstawiana jest przekształcona wartość, tzw. odległość T^2 Holettinga. Jest ona wyznaczona, jako różnica pomiędzy średnią wartością w próbce i średnią w populacji.

$$T^2(p) = n\left[(\bar{x}(p) - \bar{\bar{x}})^T s^{-1} (\bar{x}(p) - \bar{\bar{x}})\right] \tag{1}$$

gdzie

n - liczba jednostek w próbce,

$\bar{x}(p)$ - wektor średnich w próbce,

$\bar{\bar{x}}$ - wektor średnich uregulowanego procesu,

s^{-1} - macierz odwrotna do macierzy kowariancji s [patrz np. Dembski 2010].

Wartości tej odległości są nanoszone na karcie kontrolnej dla poszczególnych grup pomiarowych odnotowanych na osi poziomej karty. Należy zaznaczyć, że na kartach kontrolnych Hotellinga dolna granica kontrolna nie występuje. W przypadku zaobserwowania rozregulowania procesu konieczne jest zidentyfikowanie cechy, która była tego przyczyną. W tym celu sporządza się wykresy jednowymiarowych kart kontrolnych np. $\bar{x} - R$. Jak już wcześniej wspomniano jest to działanie dość pracochłonne, szczególnie gdy badanych jest kilkanaście lub kilkadziesiąt właściwości procesu.

Użytecznym narzędziem kontrolowania procesu o wielu właściwościach może być analiza głównych składowych. Umożliwia ona ze zbioru wszystkich zmiennych charakteryzujących dany proces uzyskanie zestawu nowych zmiennych, nazywanych składowymi głównymi. Stosunkowo niewielka liczba składowych głównych przenosi większość informacji zawartych w dużej liczbie zmiennych

wejściowych. W przypadku nadzorowania procesu z wykorzystaniem tej metody w pierwszym kroku przeprowadza się analizę składowych głównych, a w kolejnym konstruuje się karty kontrolne Hotellinga dla wyodrębnionych najważniejszych składowych.

W większości wykorzystywanych kart kontrolnych zakłada się, że zebrane wyniki pomiarów mają rozkład normalny lub zbliżony do normalnego. Założenie to jest wymagane ze względu na postać wzorów, zgodnie z którymi wyznacza się granice kontrolne. W przypadku niespełnienia założenia dotyczącego zgodności z rozkładem normalnym na karcie mogą pojawiać się fałszywe sygnały o rozregulowaniu procesu produkcji lub nie będą odnotowane wartości świadczące o jego niestabilności [Greber 2002].

Jednym z głównych założeń w metodach konstrukcji kart kontrolnych Hotellinga oraz tych, w których wykorzystuje się składowe główne jest, żeby zależność pomiędzy właściwościami procesu była liniowa [Demski 2010].

ZANURZANIE OBSERWACJI W PRÓBIE W KARTACH KONTROLNYCH

Dokonując analizy statystycznej zjawisk empirycznych często napotyka się na utrudnienia wynikające z niespełnienia założeń stawianych wykorzystywanym metodom statystycznym. Wiele z tych metod wymaga spełnienia założeń dotyczących zgodności z rozkładem normalnym lub co najmniej przybliżonego rozkładu normalnego. W związku z powyższym zachodzi konieczność stosowania metod statystycznych, które pomijają klasyczne założenia stochastyczne.

W 1975 roku zostało wprowadzone przez Tukey'a pojecie zanurzania obserwacji w zbiorze danych. Pojęcie to było rozwijane przez wielu badaczy z punktu widzenia jego przydatności w statystycznej analizie danych, szczególnie w analizie danych z obserwacjami nietypowymi. W literaturze przedmiotu szeroko omówione zostały różne kryteria wyznaczania miary zanurzenia obserwacji w próbie m. in. kryterium Mahalanobisa, kryterium zanurzania sympleksowego Liu lub kryterium Barnett'a [Liu i in. 1999, Rousseeuw i Ruts 1996].

Niech $P_n^p = \{x_1, x_2, ..., x_n\}$ będzie układem wektorów wyrażających próbę p-wymiarową o liczebności n pochodzącą z pewnego p-wymiarowego rozkładu określonego dystrybuantą F_p oraz niech $\theta \in R^p$ będzie pewnym punktem z przestrzeni rzeczywistej R^p. Punkt θ może być punktem z próby P_n^p. Koncepcja zanurzania obserwacji w próbie Tukey'a oparta jest na następującej definicji.

Definicja 1. Zanurzaniem punktu θ w próbie P_n^p nazywamy funkcję próby $zan_p(\theta, P_n^p)$ o następującej postaci

$$zan_p(\theta, P_n^p) = \min_{\|u\|=1} \#\left\{ i : u^T \theta \geq u^T x_i \right\}, \tag{2}$$

gdzie u jest wektorem kolumnowym z przestrzeni R^p o długości 1, czyli o normie euklidesowej $\|u\| = 1$, natomiast $\#\{\cdot\}$ oznacza liczebność rozważanego zbioru.

W pracy do konstrukcji kart kontrolnych wykorzystane zostało kryterium zanurzania Mahalanobisa. Jest ono określone według następującej definicji.

Definicja 2. Miarą zanurzania Mahalanobisa ($Mzan_p$) punktu θ w próbie P_n^p nazywamy funkcję

$$Mzan_p(\theta; P_n^p) = \left[1 + Q(\theta, P_n^p)\right]^{-1}, \tag{3}$$

gdzie $Q(\theta, P_n^p) = (\theta - \bar{x})^T S^{-1} (\theta - \bar{x})$ *jest odległością Mahalanobisa wektora* θ *od wektora średnich* $\bar{x}$, przy czym $\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_p \end{bmatrix}$, $\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_p \end{bmatrix}$,

$\bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j$, S *jest macierzą kowariancji między rozważanymi* p *wektorami, natomiast* S^{-1} *jest jej macierzą odwrotną.*

W sensie matematycznym miara zanurzania $Mzan_p$ wyznaczona na podstawie wzoru (3) bazuje na odległości między punktem przestrzeni p wymiarowej i wektorem średnich $\bar{x}$, według metryki wyznaczonej odwrotnością macierzy S. Własności miary zanurzania $Mzan_p$ wynikają bezpośrednio z podstawowych własności odległości Mahalanobisa [patrz np. Mahalanobis1936, Bartkowiak 1976].

Załóżmy, że proces produkcyjny charakteryzowany jest za pomocą p właściwości (p>1). Niech G będzie pewnym p-wymiarowym rozkładem teoretycznym określonym dystrybuantą F_G oraz niech $y_1, y_2, ..., y_m$ będzie układem p-wymiarowych wektorów pochodzących z rozkładu G. $P_{m(G)}^p = \{y_1, y_2, ..., y_m\}$ o liczebności m jest próbą odniesienia

w kontekście monitorowanego procesu produkcji. Niech $P_{n(F)}^{p} = \{x_1, x_2, ..., x_n\}$ będzie układem obserwowalnych wektorów procesu produkcji wyrażającym próbę *p*-wymiarową o liczebności *n* pochodzącą w pewnego *p*-wymiarowego rozkładu określonego dystrybuantą F_F. Zbadanie, czy proces produkcji uległ rozregulowaniu sprowadza się do sprawdzenia, czy obserwacje próby $P_{n(F)}^{p}$ należą do rozkładu G. Karty kontrolne, które są wykorzystywane w tym celu oparte są na mierze zanurzania obserwacji w próbie. Miary zanurzania wyznaczone dla obserwacji próby umożliwiają ich uporządkowanie względem oddalenia od centralnego skupienia, które jest wyznaczone przez obserwacje z najwyższymi wartościami miary zanurzania. Punkty, którym odpowiadają najniższe wartości tej miary leżą na zewnątrz „chmury danych" i mogą być traktowane jako odstające, czyli nie należące do danego rozkładu.

Procedura sprawdzania czy dany proces uległ rozregulowaniu sprowadza się do konstrukcji kart kontrolnych z wykorzystaniem miary zanurzania obserwacji oraz próbie [Liu 1995]. W tym celu wyznaczane są miary zanurzania obserwacji prób $P_{m(G)}^{p}$ i $P_{n(F)}^{p}$. Karta kontrolna oparta na miarach zanurzania jest wykresem wartości $r_{G_m}(x_i)$ w czasie z wartością CL=0,5 i linią kontrolną α. Jeżeli wartości $r_{G_m}(x_i)$ są mniejsze od zadanej wartości α, to wystąpiło rozregulowanie procesu. Wartości statystyki $r_{G_m}(x_i)$ wyznaczamy jako

$$r_{G_m}(x_i) = \#\left\{y_j / zan_G(y_j) \le zan_G(x_i), i = 1,2,...,n, j = 1,2,...,m\right\}/m \tag{4}$$

gdzie $\#\{\cdot\}$ oznacza liczebność rozważanego zbioru, $zan_G(\cdot)$ jest wartością miary zanurzania obserwacji w próbie wyznaczoną na podstawie wzoru 3.

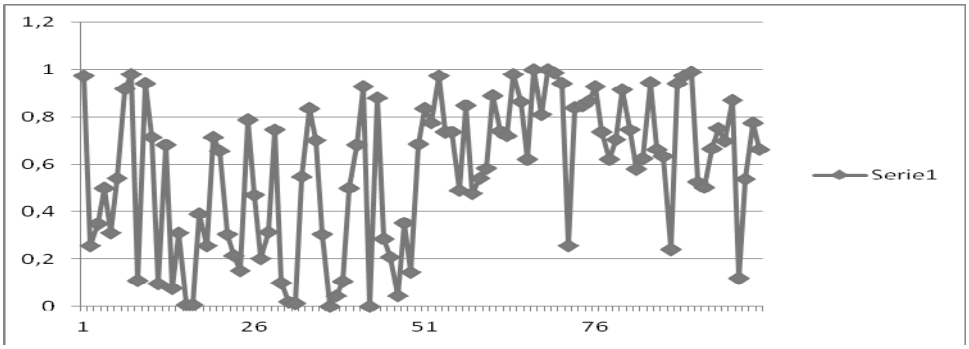
W celu zilustrowania konstrukcji kart kontrolnych posłużono się próbami dwuwymiarowymi, przyjmując za każdym razem inną konfiguracje badanych prób:

- Próba PD1 pochodzi ze standardowego dwuwymiarowego rozkładu normalnego,
- Próba PD2 – pierwszych 250 obserwacji pochodzi ze standardowego dwuwymiarowego rozkładu normalnego, kolejnych 100 z dwuwymiarowego rozkładu normalnego z wektorem wartości oczekiwanych $[3,3]$ i macierzą kowariancji $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,
- Próba PD8 – pierwszych 50 obserwacji pochodzi ze standardowego dwuwymiarowego rozkładu normalnego, kolejnych 50 pochodzi

z dwuwymiarowego rozkładu normalnego z zerowym wektorem wartości oczekiwanych oraz macierzą kowariancji $\begin{bmatrix} 0{,}5 & 0 \\ 0 & 0{,}5 \end{bmatrix}$.

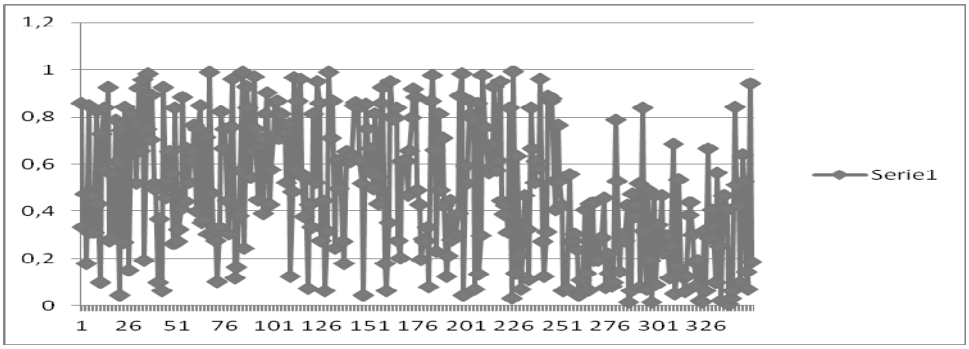
PD1 jest próbą odniesienia w kontekście monitorowanego procesu.

Rysunek 1. Karta kontrolna PD8



Źródło: opracowanie własne

Rysunek 2. Karta kontrolna PD2



Źródło: opracowanie własne

Na rysunkach przedstawione zostały procesy, w których widoczne są rozregulowania. Na pierwszym nich (rys.1) od 51 próbki wzrosła wartość statystyki $r_{G_m}(x_i)$. Wysoka jej wartość świadczy o tym, że większej części obserwacji próby PD1 odpowiadają niższe wartości miary zanurzania w porównaniu z obserwacjami próby PD8. Oznacza to, że w próbie PD8 występuje większe skupienie obserwacji wokół wektora wartości średnich w porównaniu z próbą PD1. Na rysunku 2 zauważyć można, że od 251 próbki zmalały wartości statystyki $r_{G_m}(x_i)$. Może to być uwarunkowane przesunięciem rozkładu lub zwiększeniem jego zróżnicowania.

PODSUMOWANIE

W pracy zaprezentowano wykorzystanie miary zanurzania obserwacji w próbie do konstrukcji kart kontrolnych. Klasyczne wielowymiarowe karty kontrolne oparte na statystyce χ^2 lub na odległości T^2 Hotellinga zakładają żeby dane pochodzące z pomiarów dokonywanych przy sterowaniu procesem miały rozkład zgodny z rozkładem normalnym. Założenie to jest również wymagane przy konstrukcji elipsy kontrolnej dla przypadku dwuwymiarowego.

Konstrukcja kart kontrolnych opartych na miarach zanurzania obserwacji w próbie dla przypadku wielowymiarowego nie wymaga złożonych przekształceń danych oraz zawansowanych metod analizy statystycznej. Nie jest wymagane również spełnienia założeń dotyczących zgodności z rozkładem normalnym.

W literaturze dotyczącej zarządzania jakością procesów produkcji szeroko omówione zostały różne rodzaje kart kontrolnych. Ich wybór zależy od specyfiki procesu, który ma być monitorowany. W pracy omówiona została karta kontrolna oparta na statystyce r, która podobnie jak karta $\bar{x}$ analizuje zmiany miar położenia oraz zmienności. W literaturze przedmiotu omówione zostały inne karty kontrolne, w których wykorzystuje się zanurzanie obserwacji w próbie [Liu 1995, Liu i in. 2004]. Będą one przedmiotem dalszych rozważań.

BIBLIOGRAFIA

Aczel A.D (2000) Statystyka w zarządzaniu, PWN, Warszawa.

Bartkowiak A., (1976), An Abbreviated Method of Calculating the Mahalanobis Distance or Residua Sum of Squares in a Linear Regression Model, Applicationes Mathematicae, 15: 215-222.

Demski T. (2010) Sterowanie jakością procesu o wielu właściwościach: wielowymiarowe karty kontrolne i inne narzędzia, www.statsoft.pl/czytelnia.html

Greber T. (2002) Karty kontrolne przy ocenie liczbowej sterowania procesami- założenia i analiza, www.statsoft.pl/czytelnia.html

Hamrol A. (2008) Zarządzanie jakością z przykładami, PWN, Warszawa.

Liu R.Y, Parelius J.M., Singh K. (1999) Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference, The Annals of Statistics, 27, 783-858.

Liu R. (1995), Control Chart for Multivariate Processes, Journal of the American Statistical Association December, Vol. 90, No. 432.

Liu, R.Y., Singh, K. and Teng, J.H. (2004) DDMA-charts: nonparametric multivariate moving average controlcharts based on data depth. Allgemeines Statisches Archiv, 88, 235-258.

Mahalanobis P. C., (1936), On the Generalized Distance in Statistics, Proc. Nat. Inst. Sci. India, 2, 49-55.

Płaska S. (2000) Wprowadzenie do statystycznego sterowania procesami technologicznymi, Wydawnictwo Politechniki Lubelskiej, Lublin.

Malinowska M., Wiśniewska E. (2011) Zarządzanie jakością żywności. Systemy, koncepcje, instrumenty, Difin.

Rousseeuw P.J. , Ruts I. (1996), Bivariate Location Depth, Applied Statistics , 45, 516-526.
Sałaciński T. (2009) SPC-statystyczne sterowanie procesami produkcji, Oficyna Wydawnicza Politechniki Warszawskiej.
Szkoda J. (2004) Sterowanie jakością procesów produkcyjnych. Teoria i praktyka, Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie.
Tukey J.W., (1975) Mathematics and the Picturing of Data, Proceedings of the International Congress of Mathematicians, 523-531.

USING OBSERVATION DEPTH IN A SAMPLE TO CONSTRUCT CONTROL CARDS

Abstract: Controlling a production process requires monitoring of the observed deviations of diagnostic variables from target values. The fundamental role in quality control is played by control cards, which are one of the main tools of statistical quality control. They are one of ongoing quality control methods and – when used properly –are very important in improving production quality. In 1975, Tukey introduced the concept of observation depth in a multi-dimensional sample. The depth measure can be used to classify multi-dimensional observations in regard to the distance from the central cluster and to detect and eliminate atypical observations. This article presents the application of a measure of observation depth in a sample in the construction of control cards.

Keywords: observation depth measure in a sample, control cards

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 137 – 148

ANALIZA KARIER RÓWNOLEGŁYCH Z WYKORZYSTANIEM NIEPARAMETRYCZNEGO MODELU RYZYK KONKURENCYJNYCH

Joanna Landmesser
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: joanna_landmesser@sggw.pl

Streszczenie: Przedmiotem badania są dwie wpływające na siebie nawzajem kariery równoległe: rodzinna i zawodowa. Analizie poddano dane zgromadzone w wyniku badania retrospektywnego „Biografie Zawodowe, Rodzinne i Edukacyjne", przeprowadzonego przez Instytut Statystyki i Demografii SGH w 2006 roku. Dotyczyły one zdarzeń zanotowanych w biografiach aktywnych zawodowo młodych mężatek. Do modelowania wykorzystano nieparametryczny model zależnych ryzyk konkurencyjnych.

Słowa kluczowe: nieparametryczny model ryzyk konkurencyjnych, kariery równoległe, ryzyka zależne

WPROWADZENIE

W pracy podjęta została próba analizy zależności występujących przy realizacji dwóch karier: rodzinnej i zawodowej. W analizie historii zdarzeń kariery rozumiane są jako sekwencje wydarzeń określonego typu, realizujące się w sposób równoległy lub szeregowy. Kariery rodzinna i zawodowa są przykładem karier równoległych, ponieważ zazwyczaj odgrywają się w tym samym czasie. Losy kobiet będących matkami, często z trudem łączących kariery rodzinną i zawodową, stanowią przykład karier konkurujących ze sobą. Przy karierach konkurujących, inaczej niż w wypadku karier komplementarnych, ma miejsce rywalizacja o wykorzystanie ograniczonych zasobów jednostki.

Celem badania jest przeprowadzenie modelowania zależności pomiędzy realizowanymi przez jednostki karierami, a dokładniej: modelowanie powiązań pomiędzy karierami rodzinną i zawodową realizowanymi przez zamężne kobiety.

Analiza związków między decyzjami prokreacyjnymi kobiet a ich aktywnością zawodową pomogłaby na znalezienie odpowiedzi na pytania typu:

- Jak wystąpienie zdarzenia w karierze rodzinnej, polegające na urodzeniu dziecka, wpływa na przerwanie zatrudnienia?
- Jak przerwanie zatrudnienia wpływa na karierę rodzinną kobiety?

KONCEPCJE TEORETYCZNE WYJAŚNIAJĄCE KSZTAŁTOWANIE SIĘ KARIER RODZINNYCH I ZAWODOWYCH

To, w jaki sposób kształtują się postawy prokreacyjne wobec kariery zawodowej, stanowi często przedmiot zainteresowań socjologów. W okresie powojennym aktywność zawodowa kobiet w większości krajów europejskich była niska. Wynikało to z tradycyjnego modelu kulturowego, w którym pracę kobiet podporządkowywano funkcjom rodzinnym. Zazwyczaj po zakończeniu edukacji kobiety podejmowały pracę, ale trwale ją przerywały po wyjściu za mąż lub po urodzeniu dziecka. Dopiero m.in. wzrost świadomości kobiet w latach 60-tych doprowadził do tego, że kobiety masowo rozpoczęły wchodzić na rynek pracy.

Spośród koncepcji teoretycznych, które można wykorzystać w badaniu karier równoległych, wymienić należy: ekonomiczną teorię funkcjonowania rodzin i gospodarstw domowych (*new home economics*), teorię inwestowania w kapitał ludzki (*human capital theory*), teorię inwestowania w kapitał społeczny (*social capital theory*), teorie płodności, koncepcję drugiego przejścia demograficznego (*second demographic transition*), czy teorię społecznych interakcji (por. szczegółowe rozważania w pracy [Ptak-Chmielewska 2005]).

Teoria funkcjonowania rodzin i gospodarstw [Becker 1993] odnosi się do podziału czasu poświęcanego przez członków rodziny dla pracy zawodowej oraz dla pracy w gospodarstwie domowym. Na rynku pracy przeważa ten z partnerów, który jest bardziej predysponowany do pracy zawodowej. Zazwyczaj uwarunkowania biologiczne powodują, że to kobiety przeważają w pracy na rzecz gospodarstw. Współcześnie jednak kobiety coraz częściej decydują się na potomstwo, nie rezygnując z pracy zawodowej. Zgodnie z teorią inwestowania w kapitał ludzki [Becker 1993] poświęcanie przez kobietę czasu na edukację przynosi wzrost dochodów z pracy, ale w konsekwencji wzrost aktywności zawodowej wywołuje spadek płodności (kariera zawodowa konkuruje z karierą rodzinną). Teoria inwestowania w kapitał społeczny określa kapitał społeczny jako potencjał wynikający ze spoistości relacji międzyludzkich, siłę więzi społecznych, które służą interesom całego społeczeństwa, a nie wyłącznie interesom grupowym lub indywidualnym [Loury 1977], [Coleman 1990], [Stone, Hughes 2002]. Ponieważ przyszłe życie młodych zależy od rodzinnego kapitału społecznego, stąd podkreśla się znaczenie rodziny jako głównej formy inwestowania w kapitał społeczny. Na przykład Parcel i Menaghan zwracają uwagę na wpływ pracy zawodowej matek na problemy z zachowaniem dziecka [Parcel, Menaghan 1993].

Niektóre teorie płodności skupiają się na określeniu modelu, który byłby w stanie odzwierciedlić zmiany w poziomie płodności. Hipoteza Easterlina mówi o cykliczności zmian płodności. Polega ona na tym, że osoby urodzone w rocznikach o dużej liczebności decydują się na mniej liczne potomstwo (i na odwrót) [Easterlin 1978]. Zgodnie z teorią drugiego przejścia demograficznego formuła rodziny ewaluuje w kierunku indywidualizmu i samorealizacji jej członków [Van De Kaa 1987]. Tradycyjna rodzina straciła na popularności i doszło do spadku liczby zawieranych małżeństw, opóźniania urodzenia pierwszego dziecka w małżeństwie, spadku płodności kobiet, obniżania się współczynnika urodzeń niższej kolejności. Dzięki rozwojowi środków kontroli urodzeń stało się możliwe długie i nieprzerywane prokreacją uczestnictwo kobiet na rynku pracy. Większość kobiet stara się łączyć rodzenie dzieci z pracą zawodową. Dobrze zarabiające kobiety wolą pozostać aktywne zawodowo i tym samym świadomie opóźniają prokreację. Również Polacy coraz później zakładają rodziny oraz coraz później decydują się na dziecko. Decyzję o założeniu rodziny, a także jej późniejszy kształt, determinują czynniki związane z funkcjonowaniem rynku pracy i sytuacją ekonomiczną gospodarstw domowych. Na odkładanie w czasie decyzji o potomstwie może wpływać lęk przed utratą pracy, chęć spełniania aspiracji zawodowych, poświęcenia się karierze oraz świadomość, że czas spędzany w pracy uniemożliwi wychowanie dziecka oraz poświęcanie mu niezbędnej uwagi. Począwszy od roku 1990 notuje się więc i w Polsce spadek dzietności oraz odraczanie decyzji macierzyńskich.

TEORIA ZALEŻNYCH RYZYK KONKURENCYJNYCH

Modelowanie zależnych ryzyk konkurencyjnych stosowano początkowo w trakcie studiów nad umieralnością z różnych przyczyn, następnie poszerzono zakres zastosowań tego typu modeli o analizę innych problemów demograficznych i społecznych [Chiang 1968], [Elandt-Johnson 1976], [Hakulinen 1977], [Courgeau 1985], [Willekens 1989], [Blossfeld, Hamerle, Mayer 1989]. Polskimi przykładami wykorzystania teorii ryzyk konkurencyjnych w demografii są prace [Frątczak, Jóźwiak, Paszek 1996], [Frątczak 1999], [Ptak-Chmielewska 2003]. Są one poświęcone wynikom badań retrospektywnych zrealizowanych w Polsce w latach 1988 i 1991, a dotyczących kariery rodzinnej, zawodowej i migracyjnej jednostek. Przeprowadzenie podobnego badania w roku 2006 skłania do analizy zaktualizowanego materiału empirycznego, czego starano się dokonać w niniejszej pracy.

W wypadku modelowania dwóch karier, np. kariery rodzinnej i kariery zawodowej dla kobiet, konkurencyjność polega na tym, że realizujące się w ich trakcie zdarzenia mogą występować w różnej kolejności (przykładowo narodziny dziecka mogą wystąpić przed przerwaniem zatrudnienia). Każdej zamężnej pracującej kobiecie przypisano dwuwymiarową zmienną losową (T_1, T_2), gdzie

T_1 jest czasem upływającym od momentu zamążpójścia do chwili narodzin pierwszego dziecka, natomiast T_2 oznacza długość czasu zatrudnienia, mierzonego od momentu zamążpójścia do chwili zaprzestania wykonywania pracy. Dla tak zdefiniowanej dwuwymiarowej zmiennej określono funkcję przeżycia $S(t_1,t_2)$, wyrażającą prawdopodobieństwo, że do chwil (t_1,t_2) nie wystąpiło żadne ze zdarzeń:

$$S(t_1,t_2)=\Pr[T_1>t_1,T_2>t_2]=\begin{cases}S_1(t_1,t_2), & t_1\le t_2,\\ S_2(t_1,t_2), & t_2\le t_1.\end{cases} \tag{1}$$

Wariant $S_1(t_1,t_2)$ dotyczy sytuacji, w której narodziny dziecka wyprzedzają zaprzestanie zatrudnienia, a $S_2(t_1,t_2)$ - gdy zaprzestanie pracy wyprzedza narodziny. Odpowiednie funkcje hazardu są postaci (dla $t_1=t_2=t$):

$$\begin{aligned} h_1^1(t)&=-\frac{\partial S_1(t_1,t_2)/\partial t_1}{S_1(t_1,t_2)}, \quad h_1^2(t)=-\frac{\partial S_2(t_1,t_2)/\partial t_1}{S_2(t_1,t_2)},\\ h_2^1(t)&=-\frac{\partial S_2(t_1,t_2)/\partial t_2}{S_2(t_1,t_2)}, \quad h_2^2(t)=-\frac{\partial S_1(t_1,t_2)/\partial t_2}{S_1(t_1,t_2)}, \end{aligned} \tag{2}$$

gdzie: $h_1^i(t)$ - warunkowe prawdopodobieństwo wystąpienia narodzin jako zdarzenia *i*-tej kolejności w chwili *t*, gdy do chwili *t* nie wystąpiło żadne ze zdarzeń, $h_2^i(t)$ - warunkowe prawdopodobieństwo wystąpienia zaprzestania zatrudnienia jako zdarzenia *i*-tej kolejności w chwili *t*, gdy do chwili *t* nie wystąpiło żadne ze zdarzeń, *i*=1,2.

Dla zmiennej losowej (T_1,T_2) wyznacza się rozkłady brzegowe:

$$S(t_1,0)=G_1(t_1)=\Pr[T_1>t_1], \quad S(0,t_2)=G_2(t_2)=\Pr[T_2>t_2], \tag{3}$$

z którymi powiązane są funkcje hazardu

$$h_1(t_1)=-\frac{dG_1(t_1)/dt_1}{G_1(t_1)} \quad \text{oraz} \quad h_2(t_2)=-\frac{dG_2(t_2)/dt_2}{G_2(t_2)}. \tag{4}$$

W sytuacji, gdy zmienne T_1 oraz T_2 są niezależne, następujące funkcje hazardu są sobie równe:

$$h_1^1(t) = h_1^2(t) = h_1(t_1) \text{ dla } t_1 = t,$$
$$h_2^1(t) = h_2^2(t) = h_2(t_2) \text{ dla } t_2 = t. \tag{5}$$

Brak spełnienia powyższych równości oznacza zależność zmiennych. W pracy założono liniową zależność funkcyjną między czasami trwania, czyli $T_1 = aT_2$ (jest to jeden z wielu możliwych typów zależności). Taka liniowa zależność między czasami trwania oznacza proporcjonalność odpowiednich funkcji hazardu, czyli $h_2(t_2)/h_1(t_1) = a$ [Frątczak, Jóźwiak, Paszek 1996].

Celem przeprowadzenia analizy porównawczej dla zdarzeń z dwóch karier dodatkowo rozważono grupy kobiet doświadczających zdarzeń tylko jednej z karier. Dla kobiet, które jedynie urodziły dzieci, zdefiniowano zmienną losową U_1 z funkcjami przeżycia i hazardu

$$G_1(u_1) = \Pr[U_1 > u_1], \qquad \bar{h}_1(u_1) = -\frac{dG_1(u_1)/du_1}{G_1(u_1)}, \tag{6}$$

natomiast dla kobiet, które doświadczyły jedynie zdarzenia z kariery zawodowej (przestały pracować), zmienną losową U_2 z funkcjami przeżycia i hazardu

$$G_2(u_2) = \Pr[U_2 > u_2], \qquad \bar{h}_2(u_2) = -\frac{dG_2(u_2)/du_2}{G_2(u_2)}. \tag{7}$$

Jeśli po przyrównaniu do siebie funkcji hazardu według kolejności występowania zdarzeń z brzegowymi funkcjami hazardu (wzory (5)) dla kobiet doświadczających obu zdarzeń okaże się, że ma miejsce zależność czasów trwania, to można przystąpić do jej dokładniejszej analizy. W tym celu porównuje się wartości brzegowych funkcji hazardu dla grup jednostek doświadczających dwóch zdarzeń (wzory (4)) z wartościami funkcji hazardu w grupach jednostek, dla których wystąpiło tylko jedno z badanych zdarzeń (wzory (6) i (7)). Wyznaczone ilorazy funkcji hazardu

$$\frac{\bar{h}_1(u_1)}{h_1(t_1)} = b \qquad \text{oraz} \qquad \frac{\bar{h}_2(u_2)}{h_2(t_2)} = c, \tag{8}$$

przy założeniu, że są one stałe w przedziałach czasu, oznaczają zależności odpowiednich czasów oczekiwania: $T_1 = bU_1$ oraz $T_2 = cU_2$. Jeśli $a, b > 1$, to czas oczekiwania na dane zdarzenie przy jego współwystępowaniu z drugim jest

dłuższy niż czas oczekiwania na to zdarzenie, gdy realizuje się ono w pojedynkę. Rozkład statystyk określonych wzorami (8) (czyli rozkład ilorazu dwóch funkcji hazardu) zbadano w pracy [Frątczak, Jóźwiak, Paszek 1996].

Zastosowany w niniejszej pracy model zależnych ryzyk konkurencyjnych jest modelem nieparametrycznym. Estymacja odpowiednich funkcji hazardu jest dokonywana zgodnie z metodologią tablic trwania życia i wymaga przyjęcia założeń dotyczących stałości wskaźników hazardu w podokresach czasu (w pracy przyjęto 12-miesięczne podokresy) oraz równomiernego rozłożenia zdarzeń w przedziale czasu. Dane są grupowane w przedziałach o krańcach równej rozpiętości $[\tau_j, \tau_{j+1})$, $j = 0,1,...,J$. Przyjmuje się oznaczenia: d_j – liczba jednostek doznających wydarzenia w *j*-tym przedziale, n_j – liczba jednostek dożywających początku *j*-tego przedziału (populacja wystawiona na ryzyko doznania wydarzenia w tym przedziale), przy czym $n_j = n_{j-1} - d_{j-1}$. Wartości funkcji przeżycia $\hat{G}_j$ oraz funkcji hazardu $\hat{h}_j$ dla *j*-tego przedziału ustala się zgodnie ze wzorami [StataCorp. 2007]:

$$\hat{G}_j = \prod_{k=0}^{j-1}\left(\frac{n_k - d_k}{n_k}\right) \quad \text{oraz} \quad \hat{h}_j = \frac{d_j}{(n_j - d_j/2)(\tau_{j+1} - \tau_j)}. \tag{9}$$

WYNIKI ESTYMACJI MODELU ZALEŻNYCH RYZYK KONKURENCYJNYCH

Wykorzystane dane empiryczne pochodziły z badania retrospektywnego „Biografie Zawodowe, Rodzinne i Edukacyjne" przeprowadzonego przez Instytut Statystyki i Demografii SGH w 2006 roku. W badaniu zbierano szczegółowe informacje na temat historii edukacyjnych, rodzinnych, migracyjnych oraz zawodowych respondentów z dokładnością co do miesiąca i roku wydarzenia. Badaniem objęto 3005 kobiet urodzonych w latach 1966-1981, czyli kobiet, które w roku 2006 miały ukończone od 25 do 40 lat. Wyłoniona na potrzeby analiz prezentowanych w niniejszej pracy podpróba liczyła 1166 zamężnych i pracujących kobiet. Zanotowały one w swojej małżeńskiej biografii wydarzenia albo tylko z kariery rodzinnej (417 kobiet urodziło pierwsze dziecko), albo tylko z zawodowej (22 kobiety przerwały ciągłość zatrudnienia), albo z obu karier naraz (zarówno narodziny potomstwa, jak i przerwanie zatrudnienia doświadczyło 727 kobiet). Dostępne w bazie informacje dotyczyły m.in. terminów takich zdarzeń, jak zawarcie związku małżeńskiego, narodziny pierwszego dziecka, utrata pracy. Wszystkie spośród wybranych kobiet zawarły związek małżeński i na początku trwania małżeństwa były zatrudnione. Przeciętny czas zatrudnienia

(niekoniecznie u jednego pracodawcy) wyniósł 30,62 miesiące, zaś przeciętny czas oczekiwania na potomstwo to 18,66 miesięcy.

W pierwszej kolejności w oparciu o brzegowe funkcje przeżycia wyznaczono wartości funkcji hazardu $h_1(t_j)$ dla urodzeń dziecka, a następnie wartości funkcji hazardu $h_2(t_j)$ dla przerwania zatrudnienia wśród kobiet realizujących zarówno karierę rodzinną, jak i zawodową (Tab. 1).

Tabela 1. Wartości funkcji przeżycia oraz hazardu dla urodzeń oraz przerwania zatrudnienia wśród kobiet realizujących karierę rodzinną i zawodową

Czas trwania małżeństwa	Urodzenia				Przerwania zatrudnienia			
	Wejście	Wyjście	Funkcja przeżycia	Funkcja hazardu	Wejście	Wyjście	Funkcja przeżycia	Funkcja hazardu
$[\tau_j, \tau_{j+1})$	n_j	d_j	$G_1(t_j)$	$h_1(t_j)$	n_j	d_j	$G_2(t_j)$	$h_2(t_j)$
[0,12)	727	356	0,5103	0,0540	727	225	0,6905	0,0305
[12,24)	371	205	0,2283	0,0636	502	223	0,3838	0,0476
[24,36)	166	72	0,1293	0,0462	279	85	0,2669	0,0300
[36,48)	94	44	0,0688	0,0509	194	57	0,1884	0,0287
[48,60)	50	22	0,0385	0,0470	137	38	0,1362	0,0268
[60,72)	28	10	0,0248	0,0362	99	19	0,1100	0,0177
[72,84)	18	4	0,0193	0,0208	80	15	0,0894	0,0172
[84,96)	14	7	0,0096	0,0556	65	21	0,0605	0,0321
[96,108)	7	5	0,0028	0,0926	44	13	0,0426	0,0289
[108,120)	2	1	0,0014	0,0556	31	8	0,0316	0,0247

Źródło: obliczenia własne

Podane w tabeli wielkości n_j oznaczają liczbę jednostek wchodzących do danego przedziału, natomiast wielkości d_j – liczbę jednostek doznających wydarzenia w przedziale (odpowiednio liczba kobiet, które rodzą dziecko lub liczba kobiet, które przerywają zatrudnienie). Wartości funkcji przeżycia oraz funkcji hazardu ustalono zgodnie ze wzorami (9). Przykładowo: $G_1(t_1) = (727 - 356)/356 = 0{,}5103$,

$h_1(t_1) = 356/[(727 - 356/2) \cdot 12] = 0{,}054$.

Wskaźniki hazardu dla narodzin osiągają wyższe wartości niż odpowiadające im w kolejnych podokresach wskaźniki dla zaprzestania zatrudnienia. Ponadto hazard dla narodzin dziecka wykazuje podwyższony poziom w pierwszych 2 latach oraz po 7 latach od zamążpójścia. Skłonność do przerwania pracy zawodowej była najwyższa w drugim roku po ślubie.

Następnie wyznaczono intensywności urodzeń jako zdarzeń pierwszej oraz drugiej kolejności ($h_1^i(t)$ dla i=1,2) i przerwania zatrudnienia również jako zdarzeń pierwszej oraz drugiej kolejności ($h_2^i(t)$ dla i=1,2) (Tab. 2).

Tabela 2. Wartości funkcji hazardu dla urodzeń i przerwania zatrudnienia jako zdarzeń pierwszej oraz drugiej kolejności wśród kobiet realizujących karierę rodzinną i zawodową

Czas trwania małżeństwa	Zdarzenia pierwszej kolejności					Zdarzenia drugiej kolejności				
	Wejście	Wyjście		Funkcja hazardu		Wejście	Wyjście		Funkcja hazardu	
$[\tau_j, \tau_{j+1})$	n_j^1	d_{1j}^1	d_{2j}^1	$h_1^1(t)$	$h_2^1(t)$	n_j^2	d_{1j}^2	d_{2j}^2	$h_1^2(t)$	$h_2^2(t)$
[0,12)	727	334	54	0,0497	0,0064	727	22	171	0,0026	0,0222
[12,24)	339	168	24	0,0549	0,0061	534	37	199	0,0060	0,0382
[24,36)	147	52	18	0,0358	0,0109	298	20	67	0,0058	0,0211
[36,48)	77	36	4	0,0508	0,0044	211	8	53	0,0032	0,0239
[48,60)	37	15	2	0,0424	0,0046	150	7	36	0,0040	0,0227
[60,72)	20	7	1	0,0354	0,0043	107	3	18	0,0024	0,0153
[72,84)	12	3	1	0,0238	0,0072	86	1	14	0,0010	0,0148
[84,96)	8	3	0	0,0385	0,0000	71	4	21	0,0048	0,0289
[96,108)	5	4	1	0,1111	0,0185	46	1	12	0,0018	0,0250
[108,120)	0	0	0	×	×	33	1	8	0,0026	0,0230

Źródło: obliczenia własne

Tym razem wielkości n_j^i oznaczają liczbę kobiet wchodzących do przedziału narażonych na ryzyko zdarzenia *i*-tej kolejności, natomiast wielkości d_{1j}^i oraz d_{2j}^i – liczby kobiet rodzących oraz przerywających zatrudnienie w danym przedziale czasowym jako zdarzenie *i*-tej kolejności. Przykładowy rachunek dla drugiej wartości funkcji hazardu dla urodzeń jako zdarzeń 1-ej kolejności wygląda następująco: $h_1^1(t) = 168/[(339 - 168/2) \cdot 12] = 0{,}0549$.

Z zestawienia wskaźników hazardu dla pierwszych urodzeń i pierwszych przerwań zatrudnienia wynika, że występują pomiędzy tymi wskaźnikami znaczące różnice. Dla wydarzeń pierwszej kolejności intensywność urodzeń jest we wszystkich analizowanych podokresach większa niż intensywność zaprzestania zatrudnienia ($h_1^1(t) > h_2^1(t)$ dla każdego *t*). Dla wydarzeń drugiej kolejności sytuacja jest odmienna: intensywność urodzeń zaistniałych po przerwaniu zatrudnienia jest mniejsza niż intensywność zaprzestania zatrudnienia następującego po narodzinach dziecka ($h_1^2(t) < h_2^2(t)$ dla każdego *t*). Porównując ze sobą wyniki przedstawione w Tab. 1 i Tab. 2 można wysunąć wniosek o występowaniu współzależności zdarzeń z kariery rodzinnej i zawodowej (ponieważ $h_1^1(t) \neq h_1^2(t) \neq h_1(t_1)$ oraz $h_2^1(t) \neq h_2^2(t) \neq h_2(t_2)$).

Aby porównać intensywności występowania zdarzeń w grupach kobiet doświadczających tylko jednego zdarzenia lub obu zdarzeń łącznie, wyznaczono

dodatkowo wartości funkcji hazardu dla urodzeń $\overline{h}_1(u_j)$ oraz dla przerwania zatrudnienia $\overline{h}_2(u_j)$ dla kobiet realizujących tylko jedną karierę (Tab. 3).

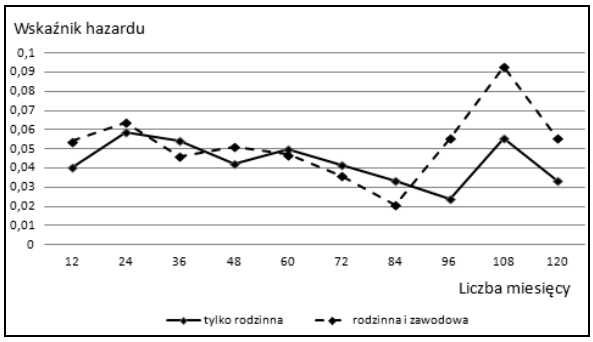
Tabela 3. Wartości funkcji przeżycia oraz hazardu dla urodzeń oraz przerwania zatrudnienia wśród kobiet realizujących tylko jedną karierę (rodzinną lub zawodową)

Czas trwania małżeństwa	Urodzenia				Przerwania zatrudnienia			
	Wejście	Wyjście	Funkcja przeżycia	Funkcja hazardu	Wejście	Wyjście	Funkcja przeżycia	Funkcja hazardu
$[\tau_j, \tau_{j+1})$	n_{1j}	d_{1j}	$G_1(u_j)$	$\overline{h}_1(u_j)$	n_{1j}	d_{1j}	$G_2(u_j)$	$\overline{h}_2(u_j)$
[0,12)	417	163	0,6091	0,0405	22	7	0,6818	0,0315
[12,0)	254	132	0,2926	0,0585	15	7	0,3636	0,0507
[24,36)	122	60	0,1487	0,0543	8	1	0,3182	0,0111
[36,48)	62	25	0,0887	0,0421	7	3	0,1818	0,0455
[48,60)	37	17	0,0480	0,0497	4	1	0,1364	0,0238
[60,72)	20	8	0,0288	0,0417	3	0	1,0000	0,0000
[72,84)	12	4	0,0192	0,0333	3	1	0,0909	0,0333
[84,96)	8	2	0,0144	0,0238	2	0	1,0000	0,0000
[96,108)	6	3	0,0072	0,0556	2	1	0,0455	0,0556
[108,120)	3	1	0,0048	0,0333	1	0	1,0000	0,0000

Źródło: obliczenia własne

Dla kariery rodzinnej otrzymane wyniki w zestawieniu z uzyskanymi poprzednio (Tab. 1) zaprezentowano graficznie (Rys. 1).

Rysunek 1. Intensywność urodzeń dla kobiet realizujących kariery: tylko rodzinną, rodzinną i zawodową



Źródło: obliczenia własne

Można zauważyć, że intensywność urodzeń jest wyższa w dwóch pierwszych oraz po siódmym roku trwania małżeństwa dla kobiet, które doświadczają obu karier, w porównaniu z intensywnością urodzeń u kobiet realizujących tylko karierę rodzinną. Ze względu na małą liczbę obserwacji dla kobiet realizujących tylko karierę zawodową podobna analiza jest utrudniona.

Wyniki zamieszczone w Tab. 4 dają informację o poziomie stałej w przedziałach czasu zależności pomiędzy czasami oczekiwania.

Tabela 4. Relacje wartości wybranych funkcji hazardu dla grup kobiet z jedną i dwiema karierami

Czas trwania małżeństwa	$\frac{T_1}{T_2}=\frac{h_2(t_j)}{h_1(t_j)}$	$\frac{h_2^1(t_j)}{h_1^1(t_j)}$	$\frac{h_2^2(t_j)}{h_1^2(t_j)}$	$\frac{\overline{h}_1(u_j)}{h_1(t_j)}$	$\frac{\overline{h}_2(u_j)}{h_2(t_j)}$
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
[0,12)	0,57	0,13	8,68	0,75	1,03
[12,24)	0,75	0,11	6,38	0,92	1,07
[24,36)	0,65	0,30	3,65	1,18	0,37
[36,48)	0,56	0,09	7,43	0,83	1,58
[48,60)	0,57	0,11	5,71	1,06	0,89
[60,72)	0,49	0,12	6,46	1,15	0,00
[72,84)	0,83	0,30	15,15	1,60	1,93
[84,96)	0,58	0,00	5,99	0,43	0,00
[96,108)	0,31	0,17	13,65	0,60	1,92
[108,120)	0,44	×	8,97	0,60	0,00

Źródło: obliczenia własne

Wartości w kolumnie *(2)* stanowią obliczenia dla współczynników liniowej zależności funkcyjnej pomiędzy obserwowanymi czasami trwania T_1 oraz T_2 ($T_1 = aT_2$) (np. $0{,}57 = h_2(t_1)/h_1(t_1) = 0{,}0305/0{,}054$). Czas upływający do chwili narodzin dziecka jest wyraźnie krótszy od czasu zatrudnienia w późniejszych latach trwania małżeństwa (po 7 roku); dla młodych mężatek podobieństwo analizowanych czasów trwania jest większe. W kolumnach *(3)* i *(4)* podano wartości stosunków odpowiednich funkcji hazardu dla zdarzeń 1-ej i 2-ej kolejności.

Wartości współczynników zawartych w kolumnach *(5)* i *(6)* niższe od 1 dowodzą wzajemnego przyspieszającego wpływu współwystępujących zdarzeń. Przykładowo: w pierwszym roku trwania małżeństwa czas oczekiwania na dziecko jest krótszy dla kobiet, które realizują obie kariery (czyli rodzą dzieci i przerywają pracę zawodową) niż dla kobiet realizujących tylko karierę rodzinną (ponieważ $T_1 = 0{,}75U_1$). Podobnie przedstawia się sytuacja w 2, 4, 8, 9 i 10 roku. W latach 3, 5, 6 i 7 oczekiwanie na urodzenie dziecka wydłuża się w grupie kobiet doświadczających zdarzeń z kariery rodzinnej i zawodowej. Niestety analiza dla zdarzeń w ramach kariery zawodowej jest utrudniona ze względu na małą liczbę obserwacji dla kobiet realizujących tylko karierę zawodową.

PODSUMOWANIE

Analiza dowiodła występowania w badanym okresie współzależności zdarzeń z kariery rodzinnej i kariery zawodowej kobiet. Intensywność dla narodzin dziecka okazała się podwyższona w dwóch pierwszych latach oraz po siedmiu

latach od zamążpójścia, natomiast skłonność do przerwania pracy zawodowej była najwyższa w 2 roku po ślubie. Ponadto intensywność urodzeń była wyższa w dwóch pierwszych oraz po siódmym roku trwania małżeństwa dla kobiet, które doświadczały obu karier, w porównaniu z intensywnością urodzeń u kobiet realizujących tylko karierę rodzinną. Intensywność urodzeń jako zdarzeń pierwszej kolejności była większa niż intensywność zaprzestania zatrudnienia, natomiast dla zdarzeń drugiej kolejności miała miejsce odwrotna relacja. Wśród starszych mężatek zanotowano mniejsze podobieństwo obu analizowanych czasów trwania.

Warto podkreślić, że obserwowany wzrost popularności badań retrospektywnych i panelowych skłania do prowadzenia analiz na poziomie mikro, czyli na poziomie jednostek, przy wykorzystaniu specjalistycznych narzędzi analizy. Wyniki badań zaprezentowane w niniejszym artykule wskazują na walory technik z zakresu analizy historii zdarzeń, dzięki którym staje się możliwe uchwycenie wzajemnych zależności między różnymi procesami.

BIBLIOGRAFIA

Becker G. S. (1993) Human capital. A theoretical and empirical analysis with special reference to education. Third edition, The University of Chicago Press, Chicago.

Blossfeld H. P., Hamerle A., Mayer K. U. (1989) Event history analysis. Statistical theory and application in the social sciences, Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey.

Chiang C. D. (1968) Introduction to stochastic process in biostatistics, Willey, New York.

Coleman J. S. (1990) Foundations of social theory, Cambridge Harvard University Press, Cambridge.

Courgeau D. (1985) Interaction between spatial mobility, family and career life-cycle: A French survey, European Sociological Review, Vol. 1, nr 2, str. 139 - 162.

Easterlin R. A. (1978) The economics and sociology of fertility: a synthesis, w: Lilly Ch. (red.) Historical studies of changing fertility, Princeton University Press, Princeton.

Elandt-Johnson, R.C. (1976) Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates, Scandinavian Actuarial Journal, nr 1, str. 37-51.

Frątczak E. (1999) Modelowanie cyklu życia jednostki i rodziny. Teoria i praktyka, Oficyna Wydawnicza SGH, Warszawa.

Frątczak E., Jóźwiak J., Paszek B. (1996) Zastosowania analizy historii zdarzeń w demografii, Oficyna Wydawnicza SGH, Warszawa.

Hakulinen T. (1977) Net probabilities in the theory of competing risks, Scandinavian Actuarial Journal, str. 65 - 80.

Loury G. (1977) A dynamic theory of racial income differences, w: Wallace P. A., Le Mund A. (red.) Women, minorities and employment discrimination, Lexington Books, Lexington, Mass.

Parcel T. L., Menaghen E. G. (1993) Family social capital and children's behavior problems, Social Psychology Quarterly, Vol. 56, nr 2.

Ptak-Chmielewska A. (2003) Studia trzech karier równoległych w demografii z wykorzystaniem modelu ryzyk konkurencyjnych, Studia Demograficzne, nr 2(144), str. 79-97.

Ptak-Chmielewska A. (2005) Przebieg równoległych karier kobiet w Polsce w okresie poprzedzającym transformację systemową w świetle badań retrospektywnych z 1988 i 1991 roku. Wykorzystanie modeli parametrycznych i semiparametrycznych, Studia Demograficzne, nr 1(147), str. 72-103.

StataCorp. (2007) Stata statistical software: release 10, College Station.

Stone W., Hughes J. (2002) Social capital. Empirical meaning and measurement validity, Research Paper No. 27, June 2002, Australian Institute of Family Studies, Melbourne.

Van de Kaa D. J. (1987) Europe's second demographic transition, Population Bulletin, nr 42(1), str. 3-24.

Willekens F. (1989) Understanding the interdependence between parallel careers, Paper presented at the Workshop "Female Labor Market Behaviour and Fertility: Preferences, Restrictions, Behaviour", 20-22 April, NIDI, The Hague.

PARALLEL CAREERS ANALYSIS USING A NONPARAMETRIC COMPETING RISKS MODEL

Abstract: The subjects of the study are two parallel careers influencing each other: family and occupational careers. We analyzed data collected from a retrospective survey "Family, Occupational and Educational Biographies", conducted by the Institute of Statistics and Demography, Warsaw School of Economics in 2006. The data related to the events recorded in the biographies of economically active young married women. The nonparametric model for dependent competing risks was used.

Keywords: nonparametric competing risks model, parallel careers, dependent risks

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 149 – 158

ZRÓŻNICOWANIE WOJEWÓDZTW POLSKI POD WZGLĘDEM STRUKTURY OBSZAROWEJ GOSPODARSTW EKOLOGICZNYCH

Lidia Luty
Katedra Statystyki Matematycznej
Uniwersytet Rolniczy w Krakowie
e-mail: rrdutka@cyf-kr.edu.pl

Streszczenie: Rolnictwo ekologiczne jest systemem gospodarowania wykorzystującym naturalne procesy zachodzące w gospodarstwie rolnym. Celem rolnictwa ekologicznego jest produkcja żywności o wysokiej jakości zdrowotnej oraz utrzymanie obszarów środowiska przyrodniczego w naturalnym stanie. Dynamiczny rozwój rolnictwa ekologicznego w Polsce obserwujemy od roku 2004. Wprowadzono wówczas dotacje do ekologicznej produkcji rolnej dla gospodarstw posiadających certyfikat zgodności i będących w trakcie przestawiania. Celem opracowania będzie przedstawienie statystycznej analizy struktury obszarowej gospodarstw ekologicznych w województwach Polski w latach 2007-2010. Do podziału zbioru obiektów na podzbiory, które charakteryzuje relatywnie niski stopień zróżnicowań badanej struktury zostanie wykorzystana metoda eliminacji wektorów. Łańcuchowe miary zmienności struktur pozwolą ocenić przebieg zmian strukturalnych.

Słowa kluczowe: gospodarstwo ekologiczne, taksonomia struktur, dynamika struktur

WSTĘP

Produkcja w gospodarstwach ekologicznych stymulowana jest naturalnymi środkami i metodami produkcji. Tradycje rolnictwa ekologicznego w Polsce sięgają czasów przedwojennych, jednak dopiero w 1990 roku przeprowadziło pierwszą inspekcję gospodarstw i nadano atesty 27 gospodarstwom. W roku 1998 wprowadzono dotacje do kosztów kontroli gospodarstw, a w kolejnym roku dopłaty do powierzchni upraw ekologicznych. W 2001 roku ustawowo

uregulowano status rolnictwa ekologicznego. Wprowadzenie dopłat do produkcji ekologicznej po akcesji Polski do UE w ramach realizacji programu rolnośrodowiskowego przyczyniło się także do znacznego przyrostu gospodarstw ekologicznych.

W latach 2004 – 2010 w Polsce liczba gospodarstw ekologicznych średnio z roku na rok wzrastała o 33% (tabela 1), a ich całkowita powierzchnia o 57%. Na podobnym poziomie kształtowało się średnie tempo zmian liczby gospodarstw ekologicznych w przyjętych grupach obszarowych (Tabela 1). W roku 2010 liczba gospodarstw o powierzchni od 50 do 100ha była siedmiokrotnie większa niż w roku 2004, gdy w pozostałych grupach obszarowych wzrost były na niewiele mniejszym poziomie.

Tabela 1. Przyrosty oraz indeksy łańcuchowe, średnie tempo zmian liczby gospodarstw ekologicznych wg grup obszarowych w Polsce w latach 2004-2010

Rok		Grupy obszarowe [ha]						Ogółem
		do 5	5-10	10-20	20-50	50-100	powyżej 100	
2004	liczba	714	940	978	677	263	188	3760
2005	liczba	2002	1879	1516	1017	496	272	7182
	przyrost	1288	939	538	340	233	84	3422
	indeks	2,80	2,00	1,55	1,50	1,88	1,45	1,91
2006	liczba	2434	2374	1939	1333	698	409	9187
	przyrost	432	495	423	316	202	137	2005
	indeks	1,22	1,26	1,28	1,31	1,41	1,50	1,28
2007	liczba	3323	2995	2281	1776	943	552	11870
	przyrost	889	621	342	443	245	143	2683
	indeks	1,37	1,26	1,18	1,33	1,35	1,35	1,29
2008	liczba	5437	3510	2666	1923	939	421	14896
	przyrost	2114	515	385	147	-4	-131	3026
	indeks	1,64	1,17	1,17	1,08	1,00	0,76	1,25
2009	liczba	5739	3794	3301	2404	1226	627	17091
	przyrost	302	284	635	481	287	206	2195
	indeks	1,06	1,08	1,24	1,25	1,31	1,49	1,15
2010	liczba	4876	4996	4532	3391	1850	934	20579
	przyrost	-863	1202	1231	987	624	307	3488
	indeks	0,85	1,32	1,37	1,41	1,51	1,49	1,20
średnie tempo zmian		1,38	1,32	1,29	1,31	1,38	1,31	1,33

Źródło: opracowanie własne

Najwyższy procentowy wzrost w grupach obszarowych do 100ha miał miejsce w roku 2005 w porównaniu do roku 2004, wówczas liczba gospodarstw o powierzchni do 5ha wzrosła o 188%, a o powierzchni od 50 do 100ha o 88%. Najniższy wzrost liczby gospodarstw ekologicznych o powierzchni powyżej 100 ha w stosunku do roku go poprzedzającego zanotowano w roku 2010. W tym

samym roku, w stosunku do roku 2009 spadła liczba gospodarstw, których powierzchnia była do 5ha.

Tabela 2. Maksymalne i minimalne liczby gospodarstw ekologicznych w grupach obszarowych w województwach Polski w latach 2007-2010

Wyszczególnienie			Rok			
			2007	2008	2009	2010
Grupy obszarowe [ha]	do 5	max	885	1233	1461	1095
			Małopolskie			
		min	14	20	17	13
			Opolskie			
	5-10	max	490	579	498	666
			Małopolskie	Podkarpackie	Lubelskie	Małopolskie
		min	11	15	13	20
			Opolskie			
	10-20	max	311	348	567	793
			Mazowieckie		Podlaskie	
		min	13	12	11	18
			Opolskie			
	20-50	max	271	283	403	608
			Zachodniopomorskie		Warmińsko-mazurskie	
		min	11	11	14	13
			Opolskie			
	50-100	max	213	186	236	382
			Zachodniopomorskie			
		min	4	5	5	10
			Opolskie			
	powyżej 100	max	165	106	161	242
			Zachodniopomorskie			
		min	0	0	3	5
			Opolskie			

Źródło: opracowanie własne

Polskie rolnictwo oparte na ekologicznych metodach wytwarzania jest i było regionalnie zróżnicowane. W latach 2007-2010 najwięcej gospodarstw ekologicznych o powierzchni do 5ha było w województwie małopolskim (tabela 2), z kolei w grupach obszarowych powyżej 50 ha - w województwie zachodniopomorskim. Województwa południowo-wschodniej Polski (małopolskie, podkarpackie, lubelskie) w badanych latach miały najwięcej ze wszystkich województw Polski gospodarstw ekologicznych o powierzchni 5-10 ha. W dwóch ostatnich latach województwa podlaskie i warmińsko-mazurskie miały najwięcej gospodarstw ekologicznych z grup obszarowych odpowiednio 10-20 ha i 20-50 ha. W województwie opolskim liczba gospodarstw ekologicznych we wszystkich grupach obszarowych w analizowanych latach była najniższa.

Średnie tempa zmian liczby gospodarstw ekologicznych w wyróżnionych grupach obszarowych w województwach Polski w latach 2007-2010 (tabela 3) były zróżnicowane. Największe średnioroczne wzrosty odnotowano w województwach północnej Polski (warmińsko-mazurskim, zachodniopomorskim i pomorskim). Natomiast najmniejsze w województwie podkarpackim, gdzie w grupach obszarowych powyżej 10 ha liczba gospodarstw średnio z roku na rok spadała. Zmniejszenie liczby gospodarstw odnotowano także w województwie opolskim, dotyczyło ono gospodarstw o powierzchni do 5 ha.

Tabela 3. Średnie tempa zmian liczby gospodarstw ekologicznych wg grup obszarowych w województwach Polski, w latach 2007-2010

Województwo	Grupy obszarowe [ha]					
	do 5	5-10	10-20	20-50	50-100	powyżej 100
Dolnośląskie	1,261	1,251	1,210	1,119	1,453	1,618
Kujawsko-pomorskie	1,210	1,185	1,165	1,049	1,112	1,276
Lubelskie	1,095	1,115	1,145	1,132	1,150	1,119
Lubuskie	1,230	1,487	1,332	1,241	1,411	1,300
Łódzkie	1,259	1,045	1,181	1,172	1,613	1,266
Małopolskie	1,074	1,108	1,113	1,158	1,333	1,491
Mazowieckie	1,090	1,119	1,157	1,303	1,438	1,337
Opolskie	0,977	1,219	1,115	1,057	1,362	-
Podkarpackie	1,250	1,081	0,990	0,979	0,992	0,900
Podlaskie	1,232	1,331	1,400	1,313	1,291	1,362
Pomorskie	1,411	1,344	1,359	1,332	1,266	1,280
Śląskie	1,153	1,195	1,165	1,041	1,279	1,358
Świętokrzyskie	1,057	1,073	1,087	1,170	1,137	1,160
Warmińsko-mazurskie	1,484	1,548	1,483	1,337	1,434	1,372
Wielkopolskie	1,188	1,223	1,244	1,254	1,219	1,121
Zachodniopomorskie	1,438	1,377	1,432	1,293	1,215	1,136

Źródło: opracowanie własne

Celem niniejszego opracowania było przedstawienie analiza struktury obszarowej gospodarstw ekologicznych w województwach Polski w latach 2007-2010 oraz wyznaczenie grup województw o podobnej strukturze obszarowej w 2010 roku.

MATERIAŁ I METODA ANALIZY

Analizę obejmującą głównie lata 2007-2010 prowadzono w oparciu o dane dostępne z raportów Głównego Inspektoratu Jakości Handlowej Artykułów Rolno-Spożywczych. Rozważane szeregi czasowe opisują liczbę gospodarstw ekologicznych wg grup obszarowych określonych następująco: do 5 ha, 5-10 ha, 10-20 ha, 20-50 ha, 50-100 ha, powyżej 100 ha.

W celu określenia tempa zmian struktury utworzonej przez *m* składowych opisanej odpowiednio w okresie *t* oraz $t+\tau$ wektorami:

$$S_t = [\alpha_{1t} \quad \alpha_{2t} \quad ..... \quad \alpha_{mt}] \tag{1}$$

$$S_{t+\tau} = [\alpha_{1(t+\tau)} \quad \alpha_{2(t+\tau)} \quad ..... \quad \alpha_{m(t+\tau)}] \tag{2}$$

gdzie α_{jt}, $\alpha_{j(t+\tau)}$ oznacza odpowiednio udział *j* składowej struktury w okresie *t* i $t+\tau$.

zastosowano miernik określony następująco [Kukuła, 1989]:

$$v_{t,t+\tau} = \frac{\sum_{j=1}^{m} |\alpha_{j(t+\tau)} - \alpha_{jt}|}{2} \tag{3}$$

Wartości $v_{t,t+\tau}$ określa stopień zmian struktury w okresie od $t+\tau$ do *t*. Miernik ten przybiera wartości z przedziału <0,1>, przy czym większa jego wartość oznacza większe zmiany w strukturze.

Przeciętne wartości miernika $v_{t,t+\tau}$ wyznaczono zgodnie ze wzorem:

$$\bar{v} = \frac{\sum_{t=1}^{n} v_{t,t+1}}{n-1} \tag{4}$$

Strukturę obszarową gospodarstw wg przyjętych grup obszarowych w jednym badanym roku, w województwach Polski można przedstawić w postaci macierzy $[a_{ij}]$:

$$[a_{ij}]_{\substack{i=1,2,...,n \\ j=1,2,...,m}} = \begin{bmatrix} a_{11} & a_{12} & ..... & a_{1m} \\ a_{21} & a_{22} & ..... & a_{2m} \\ .... & .... & ..... & .... \\ .... & .... & ..... & .... \\ a_{n1} & a_{n2} & ..... & a_{nm} \end{bmatrix} \tag{5}$$

gdzie a_{ij} opisuje udział liczby gospodarstw z *j*-tej grupy obszarowej w liczbie wszystkich gospodarstw w *i*-tym województwie, przy czym spełnione są następujące warunki: $0 \le a_{ij} \le 1$ oraz $\sum_{j=1}^{m} a_{ij} = 1$, $i = 1, \ 2,..., \ n$.

Dla każdej pary województw wyznaczono miarę zróżnicowania według wzoru [Chomątowski i Sokołowski, 1978]:

$$d_{ik} = 1 - \sum_{j=1}^{m} \min\{\alpha_{ij}, \alpha_{kj}\} \tag{6}$$

Elementy d_{ik} tworzą symetryczną macierz $[d_{ik}]_{(i,k=1,...,n)}$.

Wartości krytyczną odległości między województwami oszacowano w oparciu o formułę:

$$d^* = \min_i \max_k \{d_{ik}\} \tag{7}$$

Podstawą podziału województw na grupy podobne ze względu na rozpatrywaną strukturę było przekształcenie macierzy $[d_{ik}]_{(i,k=1,\ldots,n)}$ w binarną macierz podobieństw $[p_{ik}]_{(i,k=1,\ldots,n)}$, której elementy p_{ik} są zdefiniowane następująco:

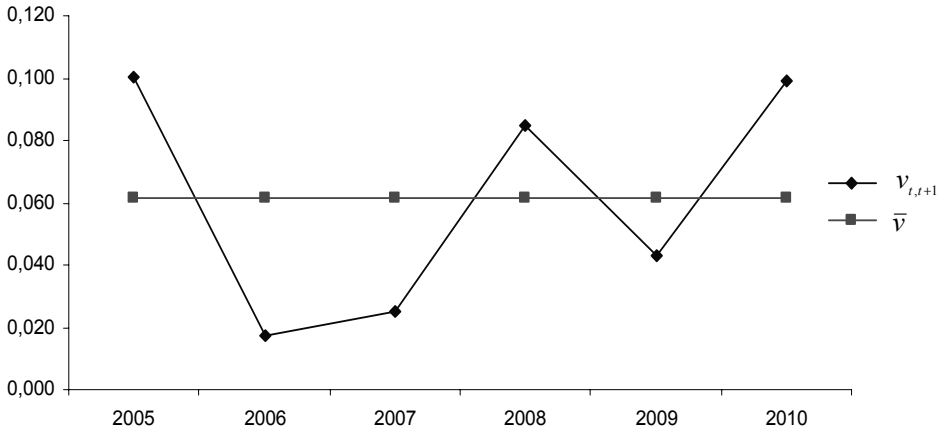
$$p_{ik} = \begin{cases} 0, & gdy \quad d_{ik} < d^* \\ 1, & gdy \quad d_{ik} \geq d^* \end{cases} \tag{8}$$

Macierz $[p_{ik}]_{(i,k=1,\ldots,n)}$ stanowi punkt wyjścia do grupowania województw zgodnie z procedurą eliminacji wektorów zaproponowaną przez S. Chomątowskiego i A. Sokołowskiego [1978].

WYNIKI BADAŃ

Przeciętne zmiany w strukturze obszarowej gospodarstw ekologicznych w całej Polsce, w latach 2004-2010 określa wartość $\bar{v} = 0{,}062$. Na rysunku 1 wartościom przeciętnych zmian strukturalnych odpowiada pozioma linia, która pozwala zwrócić uwagę na okresy ponadprzeciętnych zmian strukturalnych.

Rysunek 1. Przebieg zmian wartości miernika dynamiki struktury obszarowej ($v_{t,t+1}$) liczby gospodarstw ekologicznych w Polsce w latach 2004-2010



Źródło: opracowanie własne

Silne zmiany badanej struktury miały miejsce w latach 2004-2005, 2007-2008 oraz 209-2010. Brak jednak stabilizacji struktury o czym świadczą zróżnicowane wartości $v_{t,t+1}$.

Województwa Polski różnią się znacznie między sobą strukturą obszarową gospodarstw ekologicznych. Charakteryzuje je także różne tempo zmian analizowanej struktury (tabela 4). W latach 2007-2010 najbardziej zmieniła się struktura obszarowa gospodarstw w województwach podkarpackim, warmińsko-mazurskim, zachodniopomorskim oraz pomorskim. W województwach tych kierunek zmian struktury nie pokrywał się z tendencjami ogólnokrajowymi. Najmniejszym z kolei zmianom podlegała analizowana struktura w województwie lubelskim.

Tabela 4. Zmiany struktury liczby gospodarstw ekologicznych wg grup obszarowych w województwach Polski, w latach 2007-2010

Województwo	$v_{2007,2008}$	$v_{2008,2009}$	$v_{2009,2010}$	$\bar{v}$
Dolnośląskie	0,141	0,057	0,103	0,100
Kujawsko-pomorskie	0,164	0,097	0,066	0,109
Lubelskie	0,033	0,045	0,012	0,030
Lubuskie	0,128	0,045	0,126	0,100
Łódzkie	0,193	0,085	0,070	0,116
Małopolskie	0,075	0,079	0,158	0,104
Mazowieckie	0,072	0,066	0,037	0,058
Opolskie	0,088	0,095	0,169	0,117
Podkarpackie	0,024	0,283	0,181	0,163
Podlaskie	0,099	0,109	0,042	0,083
Pomorskie	0,187	0,038	0,166	0,131
Śląskie	0,084	0,074	0,116	0,091
Świętokrzyskie	0,086	0,032	0,139	0,086
Warmińsko-mazurskie	0,197	0,130	0,076	0,134
Wielkopolskie	0,139	0,086	0,092	0,106
Zachodniopomorskie	0,201	0,043	0,157	0,134
Polska	0,072	0,022	0,113	0,069

Źródło: opracowanie własne

W roku 2010 struktura obszarowa gospodarstw ekologicznych województwa małopolskiego i warmińsko-mazurskiego znacząco odróżniała się od omawianej struktury z pozostałych województw Polski (tabela 5). W województwie małopolskim w roku 2010 aż 50,8% gospodarstw ekologicznych miało powierzchnię do 5ha, a 30,9% powierzchnię od 5 do 10 ha. Przeciwnie w województwie warmińsko-mazurskim 84,7 % gospodarstw miało powierzchnię powyżej 10ha, z tego 56,1% powierzchnię od 10 do 50 ha.

W roku 2010 przyjmując progową wartość zróżnicowania na poziomie $d^* = 0{,}331$, wykorzystując metodę eliminacji wektorów wyodrębniono 3 grupy

województw podobnych (rysunek 2). Grupa I obejmuje województwa: dolnośląskie, kujawsko-pomorskie, lubelskie, lubuskie, mazowieckie, opolskie, podlaskie, pomorskie oraz zachodniopomorskie. W województwach tej grupy dominują gospodarstwa ekologiczne o powierzchni 5-50 ha stanowiąc co najmniej 60% wszystkich gospodarstw.

Tabela 5. Województwo (*k*-te), dla którego wartości miernika d_{ik} dla *i*-tego województwa Polski w roku 2010 osiągały odpowiednio wartość maksymalną oraz minimalną

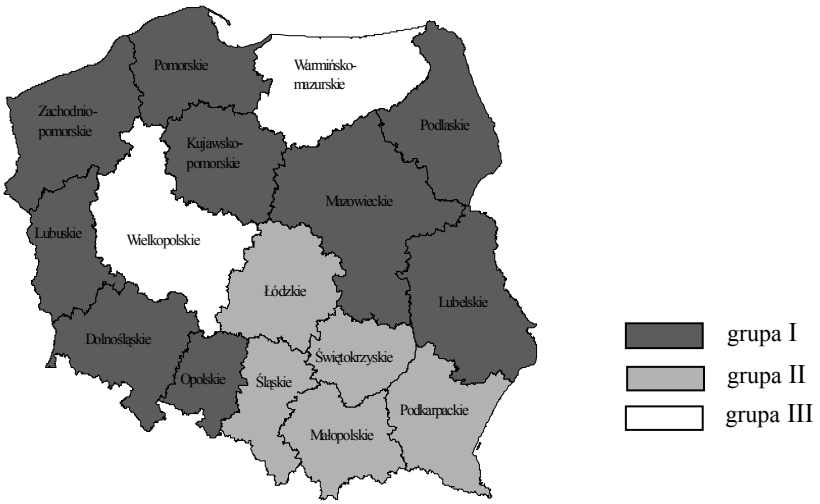
i	Województwo	Województwo, dla którego	
		$\max_k d_{ik}$	$\min_k d_{ik}$
1	Dolnośląskie	Małopolskie	Pomorskie
2	Kujawsko-pomorskie	Małopolskie	Mazowieckie
3	Lubelskie	Wielkopolskie	Łódzkie
4	Lubuskie	Małopolskie	Zachodniopomorskie
5	Łódzkie	Warmińsko-mazurskie	Lubelskie
6	Małopolskie	Warmińsko-mazurskie	Podkarpackie
7	Mazowieckie	Małopolskie	Opolskie
8	Opolskie	Małopolskie	Mazowieckie
9	Podkarpackie	Warmińsko-mazurskie	Łódzkie
10	Podlaskie	Małopolskie	Lubuskie
11	Pomorskie	Małopolskie	Dolnośląskie
12	Śląskie	Podlaskie	Małopolskie
13	Świętokrzyskie	Warmińsko-mazurskie	Łódzkie
14	Warmińsko-mazurskie	Małopolskie	Pomorskie
15	Wielkopolskie	Małopolskie	Lubuskie
16	Zachodniopomorskie	Małopolskie	Lubuskie

Źródło: opracowanie własne

Grupę II reprezentuje województwo: łódzkie, małopolskie, podkarpackie, śląskie i świętokrzyskie. We wszystkich tych województwach liczba gospodarstw ekologicznych o powierzchni do 5 ha stanowiła najwyższy odsetek, od 35,7% w województwie łódzkim do 50,8% w województwie małopolskim. Wiele gospodarstw ekologicznych w województwach tej grupy miało powierzchnię od 5 do 10 ha. Natomiast nieliczne powierzchnię powyżej 50 ha.

Województwa warmińsko-mazurskie i wielkopolskie tworzą grupę III. Dominują tu gospodarstwa ekologiczne o powierzchni co najmniej 10ha. Ich udział w ogólnej liczbie gospodarstw był powyżej 75%. Grupę tę odróżnia od poprzednich największy odsetek gospodarstw o powierzchni powyżej 50ha: w warmińsko-mazurskim - 18,3%, w wielkopolskim - 31%.

Rysunek 2. Grupy województw Polski podobnych ze względu na strukturę obszarową gospodarstw ekologicznych w roku 2010



Źródło: opracowanie własne

WNIOSKI

Województwa Polski są zróżnicowane pod względem struktury obszarowej gospodarstw ekologicznych. Można wydzielić trzy grupy województw podobnych do siebie pod względem badanej struktury. Oznacza to, że w Polsce występują trzy różne typy struktury obszarowej gospodarstw ekologicznych. Różne są także kierunki i tempo zmian badanej struktury. W większości wyróżnionych województw zmiany są niestabilne, co jest także obserwowane w odniesieniu do wszystkich gospodarstw ekologicznych w kraju. W latach 2007-2010 średnioroczne wzrosty liczby gospodarstw ekologicznych o powierzchni od 50 do 100 ha w większości województw były znacznie większe niż w grupie obszarowej do 5 ha. Wyjątek stanowią województwa północnej Polski (kujawsko-pomorskie, pomorskie, warmińsko-mazurskie, zachodniopomorskie), gdzie liczba tych gospodarstw już w roku 2007 była duża oraz województwo podkarpackie, w którym to przyrosty liczby gospodarstw w badanym okresie odnotowano jedynie w odniesieniu do gospodarstw o powierzchni do 10ha. Zmiany struktury obszarowej gospodarstw ekologicznych są więc jednym ze wskaźników dostosowywania się tego rodzaju rolnictwa do warunków ekonomicznych Duże gospodarstwa są w znacznie korzystniejszej sytuacji, ze względu chociażby na ograniczoną skalę wytwórczości małych gospodarstw.

BIBLIOGRAFIA

Chomątowski S., Sokołowski A. (1978) Taksonomia struktur. Przegląd Statystyczny, z. 2, s.217-226.

Golinowska M., Janowska-Biernat J. (2006) Rozwój rolnictwa ekologicznego na Dolnym Śląsku. Prace Naukowe AE we Wrocławiu, nr 1118, 244-240.

Grabiński T., Wydymus S., Zeliaś A. (1989) Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych. PWN, Warszawa.

Kukuła K. (1989) Statystyczna analiza strukturalna i jej zastosowanie w sferze usług produkcyjnych dla rolnictwa. Zeszyty Naukowe AE w Krakowie, Seria specjalna: Monografie, 89, Kraków.

Łuczka-Bakuła W. (2007) Rynek żywności ekologicznej. Wyznaczniki i uwarunkowania rozwoju, PWN, Warszawa, 33-43.

Małyszem M, Szmidt K. (2003) Rolnictwo ekologiczne w Polsce i krajach Unii Europejskiej. Wieś i Rolnictwo, nr 4, 59-75.

Raport o stanie rolnictwa ekologicznego w Polsce w latach 2004-2010, Główny Inspektorat Jakości Handlowej Artykułów Rolno-Spożywczych, Warszawa.

THE DIVERSITY OF POLISH VOIVODESHIPS IN TERMS OF THE STRUCTURE OF AREA OF THE ORGANIC FARMS

Abstract: Organic farming is a system of farming that uses natural processes used in the farm. The aim of organic farming is the production of food of healthy high quality and the keeping of environmental areas in the natural state. Dynamic development of organic farming in Poland has been observed since 2004. At that time there were introduced the subsidies to organic agricultural production for holdings with the certificate of conformity and in the course of transformation. Purpose of this paper is to present a statistical analysis of the structure of the area of the organic farms in Polish voivodeships in the years 2007-2010. To divide the set of objects into subsets, which is characterised by a relatively low degree of variations of the tested structure is used the method of elimination of vectors. Chain measure of variation of structures will enable to evaluate the course of structural changes.

Keywords: organic farm, taxonomy of structures, dynamics of structures

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 159 – 168

OCENA STABILNOŚCI WYNIKÓW KLASYFIKACJI WOJEWÓDZTW POLSKI POD WZGLĘDEM POZIOMU ŻYCIA LUDNOŚCI

Małgorzata Machowska-Szewczyk
Katedra Metod Sztucznej Inteligencji i Matematyki Stosowanej
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: mmachowska@wi.zut.edu.pl
Agnieszka Sompolska-Rzechuła
Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: asompolska@zut.edu.pl

Streszczenie: W pracy przedstawiono ocenę stabilności klasyfikacji województw Polski pod względem poziomu życia w latach 2002-2009. Dobór cech diagnostycznych został przeprowadzony w sposób dynamiczny, dzięki czemu otrzymany zestaw cech końcowych jest aktualny nie tylko w okresie badanym, ale także w przyszłości. Do klasyfikacji województw zastosowano metodę Warda oraz sprawdzono dopasowanie dendrogramów do macierzy odległości **D** za pomocą współczynnika korelacji kofenetycznej, wskaźnika całościowego dopasowania i miary STRESS. Wykorzystując wskaźniki podobieństwa wyników klasyfikacji stwierdzono, że klasyfikacje charakteryzują się dość dobrą stabilnością w czasie.

Słowa kluczowe: stabilność wyników klasyfikacji, metoda Warda, miary dopasowania

WPROWADZENIE

W ostatnich latach można zaobserwować wzrost zainteresowania takimi pojęciami jak: dobrobyt ekonomiczny, poziom życia, jakość życia, warunki życia. Po okresie fascynacji tempem postępu technicznego i ekonomicznego oraz intensywnych badań nad rozwojem gospodarczym, nastąpił etap analizy korzyści i zagrożeń, wynikających z postępu cywilizacyjnego dla człowieka. Zatem, poza badaniem poziomu życia pod względem materialnym, pojawiły się propozycje

badania poziomu życia pod względem zaspokajania potrzeb duchowych, moralnych, emocjonalnych, psychicznych. Jednak w pełni obiektywna ocena tej kategorii jest zadaniem niezmiernie trudnym, gdyż z poziomem życia związane są takie cechy jak: ukryte bezrobocie, liczba samobójstw, ilość oraz sposób wykorzystania czasu wolnego czy możliwości korzystania z dóbr kultury, o których trudno jest pozyskać informacje i dokonać ich prawidłowej kwantyfikacji [Zeliaś 2004]. Z uwagi na wieloaspektowość do badania poziomu życia można wykorzystać metody wielowymiarowej analizy danych.

W artykule przedstawiono wyniki taksonomicznego badania poziomu życia w województwach Polski w ujęciu dynamicznym. Celem pracy jest ocena stabilności wyników klasyfikacji województw Polski ze względu na wyróżnioną kategorię w latach 2002-2009. W okresie trwającego w Polsce rozwarstwienia społecznego i pauperyzacji znacznej części społeczeństwa skierowanie uwagi na problem oceny poziomu życia ma szczególne znaczenie.

DOBÓR CECH DIAGNOSTYCZNYCH W UJĘCIU DYNAMICZNYM

Do doboru zmiennych, które charakteryzują się relatywnie wysokimi wahaniami wartości oraz gdy celem jest przewidywanie przyszłego rozwoju badanego zjawiska, wykorzystać można współczynniki zmienności oraz korelacji liniowej, obliczane na podstawie danych dotyczących nie tylko ostatniego okresu, ale również kilku poprzednich okresów [Zeliaś 2004].

Ocenę stabilności klasyfikacji przeprowadzono na podstawie podziałów województw Polski w poszczególnych latach okresu 2002-2009 pod względem poziomu życia mieszkańców. Wykorzystano dane statystyczne udostępnione przez Główny Urząd Statystyczny w Banku Danych Lokalnych na stronie internetowej http://www.stat.gov.pl/bdl/app/portret.dims do wyznaczenia wartości wskaźników, które charakteryzują analizowane zjawisko. W badaniu przyjęto następujący zestaw potencjalnych cech diagnostycznych:

X_1 – liczba ludności na 1 km^2,
X_2 – udział ludności w wieku przedprodukcyjnym w ogólnej liczbie ludności,
X_3 – udział ludności w wieku produkcyjnym w ogólnej liczbie ludności,
X_4 – udział ludności w wieku poprodukcyjnym w ogólnej liczbie ludności,
X_5 – ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym,
X_6 – ludność w wieku poprodukcyjnym na 100 osób w wieku przedprodukcyjnym,
X_7 – ludność w wieku poprodukcyjnym na 100 osób w wieku produkcyjnym,
X_8 – liczba kobiet na 100 mężczyzn,
X_9 – zgony na 1000 ludności,
X_{10} – przyrost naturalny na 1000 ludności,
X_{11} – urodzenia żywe na 1000 ludności,
X_{12} – liczba małżeństw zawartych w ciągu roku na 1000 ludności,
X_{13} – liczba rozwodów na 1000 ludności,
X_{14} – zgony niemowląt na 1000 urodzeń żywych,

X_{15} – przeciętne miesięczne wydatki na 1 osobę w zł,
X_{16} – stopa bezrobocia w %,
X_{17} – liczba ofert pracy ogółem na 1 bezrobotnego,
X_{18} – wskaźnik zatrudnienia ogółem w %,
X_{19} – przeciętne miesięczne wynagrodzenie brutto w relacji do średniej krajowej (Polska=100),
X_{20} – przeciętna powierzchnia użytkowa mieszkania w m^2 na 1 osobę,
X_{21} – liczba mieszkań na 10 tys. ludności,
X_{22} – liczba studentów na 10 tys. ludności,
X_{23} – liczba praktyk lekarskich w miastach na 10 tys. ludności,
X_{24} – liczba praktyk lekarskich na wsi na 10 tys. ludności,
X_{25} – liczba ludności na 1 aptekę ogólnodostępną,
X_{26} – liczba osób korzystających ze świadczeń pomocy społecznej na 10 tys. ludności,
X_{27} – stopień wykorzystania miejsc noclegowych w %,
X_{28} – czytelnicy bibliotek publicznych na 1000 ludności,
X_{29} – wypożyczenia księgozbioru na 1 czytelnika,
X_{30} – liczba ludności na 1 miejsce w kinach stałych,
X_{31} – liczba widzów i słuchaczy w teatrach i instytucjach muzycznych na 1000 ludności,
X_{32} – PKB na 1 mieszkańca,
X_{33} – nakłady inwestycyjne na 1 mieszkańca w zł,
X_{34} – drogi publiczne o twardej nawierzchni w km na 100 km^2 powierzchni,
X_{35} – liczba samochodów osobowych zarejestrowanych na 1000 ludności,
X_{36} – liczba ofiar śmiertelnych na 100 wypadków drogowych,
X_{37} – emisja przemysłowych zanieczyszczeń pyłowych powietrza w t na 100 km^2,
X_{38} – emisja przemysłowych zanieczyszczeń gazowych powietrza w t na 100 km^2,
X_{39} – udział parków narodowych w ogólnej powierzchni w %,
X_{40} – plony z 1 ha zbóż ogółem w dt,
X_{41} – dochody budżetu województwa ogółem na 1 mieszkańca w zł,
X_{42} – wydatki z budżetu województwa ogółem na 1 mieszkańca w zł,
X_{43} – liczba podmiotów gospodarczych na tys. mieszkańców.

Metoda doboru cech diagnostycznych w ujęciu dynamicznym oraz procedura doboru zostały przedstawione w pracy [Machowska-Szewczyk, Sompolska-Rzechuła 2012]. Po selekcji cechy: X_1, X_{22}, X_{26}, X_{32}, X_{36}, X_{39} tworzą nową listę zmiennych, które są słabo skorelowane między sobą i równocześnie silnie skorelowane z pozostałymi cechami, wyeliminowanymi ze zbioru cech diagnostycznych.

METODA WARDA I OCENA JAKOŚCI DENDROGRAMU

Spośród bardzo wielu metod hierarchicznych do badań wybrano metodę Warda. Na każdym etapie tej metody optymalizuje się otrzymany podział, przez

połączenie dwóch elementów, stosując kryterium minimalnego wzrostu łącznej wewnątrzgrupowej sumy kwadratów odchyleń wszystkich wartości zmiennych dla każdego obiektu od ich średnich grupowych. Zapewnia ona zatem homogeniczność wewnątrz skupień i heterogeniczność między skupieniami, przez co uznawana jest za bardzo efektywną [Ward 1963]. Na podstawie przeprowadzonych badań można stwierdzić, że metoda Warda jest prawie o 40% bardziej efektywna[1] w porównaniu z drugą z kolei pod względem efektywności – metodą najdalszego sąsiedztwa [Malina 2004]. Ponadto metoda Warda może być wykorzystana do określenia wstępnej liczby klas na podstawie wzrokowej analizy dendrogramu, jednak jest to sposób subiektywny, ponieważ trudno jednoznacznie odpowiedzieć na pytanie, w którym miejscu należy przerwać proces aglomeracji, tzn. na ile klas podzielić daną zbiorowość.

Zaletą metod hierarchicznych (w tym także metody Warda) jest graficzna prezentacja klasyfikacji za pomocą dendrogramu, który ilustruje kolejność łączenia się obiektów w klasy i poziomy, na których obiekty po raz pierwszy się połączyły. W literaturze przedmiotu można znaleźć propozycje mierzenia dopasowania dendrogramu do macierzy odległości **D**, takie jak np.: współczynnik korelacji kofenetycznej, wskaźnik całościowego dopasowania, statystyka Gowera, miara STRESS [Gordon 1999]). Porównuje się wyjściową macierz odległości **D** z macierzą odległości **C**, otrzymaną na podstawie dendrogramu - zwaną macierzą dendrogramu lub macierzą kofenetyczną. Elementami macierzy **C** są odległości odpowiadające poziomom łączenia, na których pary obiektów łączą się w tym samym skupieniu pierwszy raz.

Dopasowanie dendrogramu można ocenić za pomocą tzw. współczynnika korelacji kofenetycznej (cophenetic correlation coefficient) [Balicki 2009]:

$$r_{kof.} = \frac{\left[\frac{n(n-1)}{2}\sum_{r<s} d_{rs}\cdot c_{rs} - \sum_{r<s} d_{rs}\cdot\sum_{r<s} c_{rs}\right]}{\sqrt{\left[n(n-1)/2\sum_{r<s} d_{rs}^2 - \left(\sum_{r<s} d_{rs}\right)^2\right]\left[n(n-1)/2\sum_{r<s} d_{rs}^2 - \left(\sum_{r<s} d_{rs}\right)^2\right]}}, \quad (1)$$

gdzie: n - liczba obiektów, $r,s \in \{1,...,n\}$, d_{rs} - elementy macierzy odległości między obiektami **D**, c_{rs} - elementy macierzy kofenetycznej **C**, n –liczba obiektów.

Dendrogram dobrze odzwierciedla różnice między obiektami lub podobieństwa obiektów, jeżeli wartość współczynnika korelacji kofenetycznej jest bliska 1. Wartość współczynnika przeciętnie maleje wraz ze wzrostem liczby obiektów i prawie nie zależy od liczby zmiennych.

[1] Przez efektywność metod taksonomicznych należy rozumieć zdolność prawidłowego rozpoznawania rzeczywistej struktury obiektów w wielowymiarowych przestrzeniach cech, którą bada się na przykładach empirycznych lub na podstawie zbiorów sztucznie wygenerowanych [Malina 2004].

Inne sposoby oceny grupowania oparte są na różnicach odległości w dwóch macierzach: wyjściowej macierzy odległości i macierzy kofenetycznej. Ocenia się zatem poziom niezgodności między d_{rs} a c_{rs} w macierzach **D** i **C**. Wykorzystują one niektóre relacje, jakie zachodzą między sumami odległości w tych macierzach: $\sum d_{rs}$ oraz $\sum c_{rs}$, różnicami między odpowiadającymi sobie odległościami: $d_{rs} - c_{rs}$ oraz ich sumami: $\sum (d_{rs} - c_{rs})$. Dodatnia suma różnic będzie świadczyła o łącznym skróceniu odległości na dendrogramie w stosunku do odległości wyjściowych, i odwrotnie – ujemna suma różnic będzie informowała o rozciągnięciu odległości. Im większa jest bezwzględna wartość tej wielkości, tym gorzej dendrogram oddaje rzeczywiste odległości. Badanie różnic stanowi podstawę wyznaczania wskaźnika całościowego dopasowania [Balicki 2009]:

$$\delta = \sum_{\substack{r,s \\ r<s}} \left| d_{rs} - c_{rs} \right| . \tag{2}$$

Inną miarą, wprowadzoną w 1964 roku przez Kruskala, stosowaną w niemetrycznym skalowaniu wielowymiarowym jest miara o nazwie STRESS (STandardized Residual Sum of Squares) [Balicki 2009]:

$$\text{STRESS} = \left[\left(\sum_{r<s}^{r,s} (d_{rs} - c_{rs})^2 \right) / \left(\sum_{r<s}^{r,s} d_{rs}^2 \right) \right]^{\frac{1}{2}}, \tag{3}$$

gdzie c_{rs} są elementami macierzy kofenetycznej, $r, s \in \{1, ..., n\}$.

Tabela 1. Wybrane odległości d_{rs} z wyjściowej macierzy odległości **D** i odległości dendrogramu c_{rs}

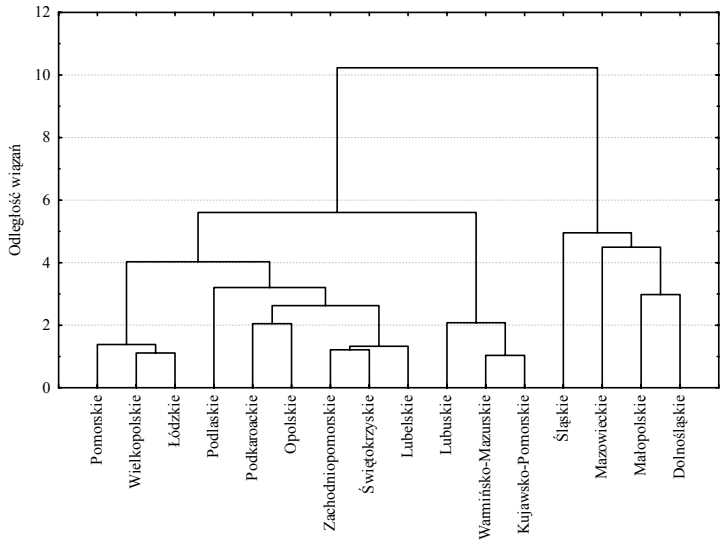
Pary obiektów	2002		2003		…	2008		2009	
	d_{rs}	c_{rs}	d_{rs}	c_{rs}	…	d_{rs}	c_{rs}	d_{rs}	c_{rs}
1-2	4,53	8,58	4,43	7,81	…	3,18	10,95	3,08	10,59
1-3	4,33	8,58	4,01	7,81	…	2,65	10,95	2,78	10,59
1-4	5,33	8,58	5,05	7,81	…	4,25	10,95	3,91	10,59
1-5	4,04	6,53	3,80	6,36	…	1,35	1,33	1,70	1,76
…	…	…	…	…	…	…	…	…	…
13-15	1,94	8,58	1,77	2,31	…	2,20	10,95	2,17	10,59
13-16	1,77	2,60	2,58	7,81	…	0,99	0,99	1,08	1,08
14-15	3,27	8,58	2,92	7,81	…	3,61	10,95	3,62	10,59
14-16	2,62	2,60	2,07	1,95	…	2,38	3,34	2,33	2,46
15-16	1,98	8,58	1,55	7,81	…	1,70	10,95	1,63	10,59
$r_{kof.}$	0,5259		0,4762		…	0,5402		0,5220	
δ	1,1016		1,0019		…	1,5450		1,4708	
STRESS	361,9405		345,0522		…	509,4228		475,2159	

Źródło: opracowanie własne

W każdym roku badanego okresu 2002-2009 przeprowadzono klasyfikację województw metodą Warda i sporządzono dendrogramy. Na tej podstawie wyznaczono macierze kofenetyczne i zbadano dopasowanie dendrogramów do macierzy odległości **D** oraz policzono współczynniki korelacji kofenetycznej, wskaźniki całościowego dopasowania i miary STRESS. Ponieważ macierze kwadratów odległości euklidesowych oraz macierze kofenetyczne mają wymiar 16×16, ponadto wszystkich kombinacji pomiędzy różnymi województwami jest $n(n-1)/2=120$, to w tabeli 3 przedstawiono jedynie fragmenty tych macierzy oraz wartości miar dopasowania.

Najwyższa wartość współczynnika korelacji kofenetycznej została zaobserwowana dla roku 2005, co świadczy o dobrym dopasowaniu dendrogramu do wyjściowej macierzy odległości. Współczynniki korelacji kofenetycznej dla dendrogramów w pozostałych latach przyjmują wartości bliskie 0,5, co oznacza umiarkowane dopasowanie macierzy **C** i **D**. Najmniejsze różnice między odległościami w macierzach **C** i **D** a tym samym najniższe wartości miar: całościowego dopasowania oraz STRESS, występują dla roku 2003. Zatem jedynie w roku 2005 województwa wykazują wyraźną hierarchiczną strukturę grupową (rys. 1).

Rysunek 1. Dendrogram podziału województw Polski w roku 2005 za pomocą metody Warda



Źródło: opracowanie własne

W celu oceny stabilności klasyfikacji województw Polski ustalono jednakową liczbę skupień oraz wcześniej wyodrębniony taki sam zbiór zmiennych diagnostycznych w każdym roku badanego okresu 2002-2009. Po analizie

dendrogramów przyjęto trzy klasy typologiczne i otrzymano wyniki klasyfikacji, które prezentuje tabela 2.

Tabela 2. Wyniki dynamicznej klasyfikacji województw Polski w latach 2002-2009

Rok	Skupienie	Województwa
2002	G1	lubuskie, podlaskie, zachodniopomorskie, pomorskie, podkarpackie, warmińsko-mazurskie, świętokrzyskie, lubelskie, kujawsko-pomorskie
	G2	śląskie, małopolskie, opolskie, wielkopolskie, łódzkie
	G3	mazowieckie, dolnośląskie
2003	G1	lubuskie, warmińsko-mazurskie, zachodniopomorskie, lubelskie, podkarpackie, pomorskie, kujawsko-pomorskie
	G2	śląskie, małopolskie, świętokrzyskie, opolskie, wielkopolskie, łódzkie
	G3	mazowieckie, podlaskie, dolnośląskie
2004	G1	lubuskie, warmińsko-mazurskie, zachodniopomorskie, lubelskie, podkarpackie, kujawsko-pomorskie, podlaskie
	G2	śląskie, małopolskie, świętokrzyskie, pomorskie, opolskie, wielkopolskie, łódzkie
	G3	mazowieckie, dolnośląskie
2005	G1	podlaskie, zachodniopomorskie, świętokrzyskie, pomorskie, opolskie, wielkopolskie, łódzkie, lubelskie, podkarpackie
	G2	lubuskie, warmińsko-mazurskie, kujawsko-pomorskie
	G3	śląskie, mazowieckie, małopolskie, dolnośląskie
2006	G1	podlaskie, podkarpackie, opolskie, zachodniopomorskie, świętokrzyskie, lubelskie, lubuskie, warmińsko-mazurskie, kujawsko-pomorskie
	G2	śląskie
	G3	mazowieckie, małopolskie, dolnośląskie, pomorskie, łódzkie, wielkopolskie
2007	G1	podlaskie, podkarpackie, opolskie, zachodniopomorskie, świętokrzyskie, lubelskie, lubuskie, warmińsko-mazurskie, kujawsko-pomorskie
	G2	śląskie
	G3	mazowieckie, małopolskie, dolnośląskie, pomorskie, łódzkie, wielkopolskie
2008	G1	podlaskie, podkarpackie
	G2	opolskie, zachodniopomorskie, świętokrzyskie, lubelskie, lubuskie, warmińsko-mazurskie, kujawsko-pomorskie
	G3	śląskie, mazowieckie, małopolskie, pomorskie, łódzkie, dolnośląskie, wielkopolskie
2009	G1	podlaskie
	G2	opolskie, lubelskie, lubuskie, podkarpackie, warmińsko-mazurskie, zachodniopomorskie, świętokrzyskie, kujawsko-pomorskie
	G3	śląskie, mazowieckie, małopolskie, pomorskie, łódzkie, dolnośląskie, wielkopolskie

Źródło: opracowanie własne

OCENA STABILNOŚCI KLASYFIKACJI

Do oceny stopnia podobieństwa wyników klasyfikacji obiektów w dwóch momentach można wykorzystać wskaźniki zgodności wyników podziałów, oparte na tablicy kontyngencji. Rozpatruje się wszystkie pary obiektów $O_i, O_j, i, j \in \{1,...,n\}, i \neq j$ ze względu na ich przynależność do grup typologicznych z obu podziałów. Każdej parze obiektów badania przyporządkowuje się wartość 1, jeśli obiekty te należą do tej samej grupy typologicznej lub wartość 0, jeśli obiekty te należą do dwóch różnych grup typologicznych. W ten sposób otrzymuje się dwie macierze przyporządkowania: $\mathbf{P}_t$ dla klasyfikacji w momencie *t* oraz $\mathbf{P}_{t'}$ dla klasyfikacji w momencie t'. Na tej podstawie buduje się czteropolową tablicę kontyngencji (tabela 5).

Tabela 3. Czteropolowa tablica kontyngencji wyników klasyfikacji w dwóch momentach

	Podział w momencie t'	
Podział w momencie *t*	1	0
1	*a*	*b*
0	*c*	*d*

Źródło: opracowanie własne na podstawie [Nowak 1990, s. 139].

Wskaźnik podobieństwa wyników klasyfikacji, oparty na tablicy kontyngencji w wymiarach 2×2 ma następującą postać [Nowak 1990]:

$$w_{tt'} = \frac{a+d}{a+b+c+d}. \tag{4}$$

Miara (4) przyjmuje wartości z przedziału $\langle 0,1 \rangle$, wartości bliskie 1 wskazują na dużą zgodność wyników obu klasyfikacji. Jeżeli w jednej klasyfikacji otrzyma się jedną grupę *n* elementową a w drugiej *n* grup jednoelementowych, to wskaźnik (4) przyjmie wartość 0. Natomiast $w_{tt'} = 1$, gdy oba podziały dają identyczne wyniki. Jeżeli porównuje się wyniki klasyfikacji z wielu momentów, to wskaźniki zgodności podziałów można przedstawić w macierzy zgodności o postaci [Nowak 1990]:

$$\mathbf{W} = \begin{bmatrix} 1 & w_{12} & \dots & w_{1T} \\ w_{21} & 1 & \dots & w_{2T} \\ \vdots & \vdots & \vdots & \vdots \\ w_{T1} & w_{T2} & \dots & 1 \end{bmatrix}. \tag{5}$$

Dla klasyfikacji otrzymanych metodą Warda w kolejnych latach okresu 2002-2009 wyznaczono macierz wskaźników podobieństwa wyników klasyfikacji (tabela 6).

Tabela 4. Macierz wskaźników podobieństwa wyników klasyfikacji w latach 2002-2009

Lata	2002	2003	2004	2005	2006	2007	2008	2009
2002	1,0000	0,8167	0,8000	0,6167	0,7167	0,7167	0,6333	0,6667
2003	0,8167	1,0000	0,8333	0,5833	0,5833	0,5833	0,6167	0,6500
2004	0,8000	0,8333	1,0000	0,6167	0,7000	0,7000	0,6667	0,6833
2005	0,6167	0,5833	0,6167	1,0000	0,6000	0,6000	0,5833	0,5833
2006	0,7167	0,5833	0,7000	0,6000	1,0000	1,0000	0,8333	0,8833
2007	0,7167	0,5833	0,7000	0,6000	1,0000	1,0000	0,8333	0,8833
2008	0,6333	0,6167	0,6667	0,5833	0,8333	0,8333	1,0000	0,9333
2009	0,6667	0,6500	0,6833	0,5833	0,8833	0,8833	0,9333	1,0000

Źródło: opracowanie własne

Analizując elementy macierzy wskaźników podobieństwa wyników klasyfikacji województw Polski pod względem poziomu życia mieszkańców w latach 2002-2009 można uznać, że podziały te są dość stabilne w czasie. Największą zgodność wykazują podziały w sąsiednich latach, np. klasyfikacje w latach 2006 i 2007 dają jednakowe wyniki ($w_{2006;2007}$=1). Klasyfikacja otrzymana w roku 2003 wykazuje najsłabszą zgodność z klasyfikacjami w latach 2005, 2006 i 2007 oraz podział w roku 2005 z podziałami w latach 2008 i 2009 (wskaźniki zgodności podziałów w tych przypadkach wynoszą 0,5833).

Najsłabsza zgodność klasyfikacji w roku 2003 w porównaniu z latami 2005, 2006 i 2007 wynika z dość dużego tempa wzrostu PKB na 1 mieszkańca, liczby studentów na 10 tys. ludności, liczby osób korzystających ze świadczeń pomocy społecznej na 10 tys. ludności.

PODSUMOWANIE

W pracy przedstawiono ocenę stabilności klasyfikacji województw Polski pod względem poziomu życia mieszkańców. Badany okres obejmował lata 2002-2009. Dynamiczne ujęcie zastosowano na etapie doboru cech diagnostycznych, wykorzystując funkcje trendów liniowych dla współczynników zmienności przyjętych cech oraz funkcje trendów liniowych współczynników korelacji. Dynamiczny dobór zmiennych diagnostycznych ma tę zaletę, że otrzymany zestaw cech końcowych jest nie tylko aktualny w okresie badanym, ale także w przyszłości. Na podstawie wyłonionego zbioru cech diagnostycznych dokonano, metodą Warda, klasyfikacji województw Polski w każdym roku badanego okresu oraz zbadano dopasowanie dendrogramów do macierzy odległości **D** za pomocą współczynnika korelacji kofenetycznej, wskaźnika całościowego dopasowania i miary STRESS. Wartości miar wskazują na dość dobrą lub umiarkowaną zgodność macierzy odległości i dendrogramów. W kolejnym kroku badania dokonano oceny stabilności otrzymanych podziałów, za pomocą wskaźników podobieństwa wyników klasyfikacji i otrzymano, że klasyfikacje charakteryzują się dość dobrą stabilnością w czasie.

Przedstawione podejście pozwoliło na wykorzystanie danych przekrojowo-czasowych w dokonaniu klasyfikacji województw Polski pod względem poziomu życia, uwzględniając dynamikę zmian zachodzących w wartościach cech ilustrujących badane obiekty w całym okresie badania. Województwa Polski są zróżnicowane pod względem poziomu życia. Badania o charakterze dynamicznym powinny być prowadzone w sposób ciągły, gdyż systematyczna analiza i ocena przestrzennego zróżnicowania poziomu życia pozwala na bieżącą kontrolę jego stanu oraz stosowanie odpowiedniej polityki społeczno-gospodarczej, w przypadku zaobserwowania negatywnych zjawisk.

BIBLIOGRAFIA

Balicki A. (2009) Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne, Wydawnictwo Uniwersytetu Gdańskiego, str. 283-291.

Gordon A. D. (1999) Classification, Chapman & Hall/CRC, Boca Raton.

Machowska-Szewczyk M., Sompolska-Rzechuła A. (2012), Dynamiczny dobór cech w taksonomicznej analizie obiektów, Wiadomości Statystyczne, nr 9, str. 34-49.

Malina A. (2004) Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, Wydawnictwo Akademii Ekonomicznej w Krakowie.

Nowak E. (1990) Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych, PWE, Warszawa.

Zeliaś A. (red) (2004) Poziom życia w Polsce i krajach Unii Europejskiej, PWE, Warszawa.

Ward J. H. (1963) Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association 58.

EVALUATION OF THE RESULTS STABILITY OF THE POLISH VOIVODSHIPS' CLASSIFICATION

Abstract: The work presents evaluation of the results stability of the Polish regions' classification in terms of a living standard. The research covered the period 2002-2009. Selection of diagnostic features was carried out in a dynamic way, so that the final set of features is valid not only in the period considered, but also in the future. For the regions' classification Ward's method was used and the fit of dendrograms distance to matrix D was checked by using the cophenetic correlation coefficient, the indicator of the overall fit and the STRESS measure. Using the similarity indices of the classification results, it was found that the classifications have a fairly good stability in time.

Keywords: stability of results of classification, Ward's method, measures of fit

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 169 – 179

PROGNOZOWANIE CZASU ZAKOŃCZENIA INWESTYCJI NA PODSTAWIE JEJ BIEŻĄCEGO ZAAWANSOWANIA

Mieczysław Połoński
Katedra Geoinżynierii
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: mieczyslaw_polonski@sggw.pl

Streszczenie: Jednym z najważniejszych obowiązków menadżera budowy na etapie prowadzenia robót budowlanych, jest kontrola jej zawansowania w stosunku do planowanych terminów i budżetu. Metodą pozwalającą na efektywne śledzenie tych parametrów na tle zaawansowania rzeczowego jest metoda wartości wypracowanej (EVM). W artykule zwrócono szczególną uwagę na prognozowanie czasu zakończenia inwestycji. W literaturze można spotkać kilka metod zastosowania wybranych wskaźników EVM do prognozowania czasu dokończenia robót budowlanych. Celem artykułu było pokazanie najważniejszych z nich oraz sprawdzenie na przykładzie dużego obiektu budowlanego realizowanego w ostatnich latach w Warszawie uzyskiwanych tymi metodami wyników oraz wskazanie rozwiązania, które w analizowanych warunkach wykazało największą zgodność prognozy z rzeczywistym czasem trwania robót.

Słowa kluczowe: wartość wypracowana, earned schedule, prognozowany czas, obiekt budowlany

WSTĘP

Jednym z najważniejszych obowiązków menadżera budowy na etapie prowadzenia robót budowlanych, oprócz dbania o zgodność realizowanych robót z projektem i jakość wykonywanych elementów i konstrukcji, jest kontrola jej zawansowania w stosunku do planowanych terminów. Na obecnym etapie złożoności robót budowlanych, przy występowaniu różnego rodzaju ryzyk [Skorupka 2009] oraz zaangażowania znacznych środków finansowych, poszukuje się metod ilościowych, które pozwoliłyby skutecznie powiązać ocenę stanu zaawansowania rzeczowego robót z przewidywanymi kosztami oraz planowanymi

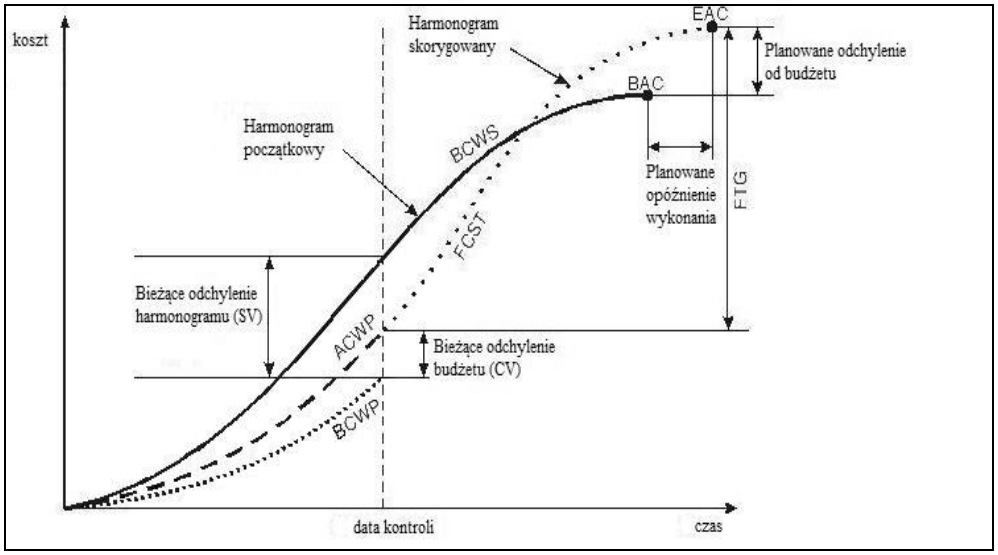
terminami realizacji poszczególnych etapów robót i całego przedsięwzięcia. Jedną z takich technik jest metoda wartości wypracowanej nazywana w literaturze Earned Value Management (EVM).

Założenia samej metody EVM, szczególnie ukierunkowane na kontrole planowanych i ponoszonych kosztów są dobrze opisane w literaturze, więc w artykule podane zostaną podstawowe informacje pozwalające na wprowadzenie w nieco odmienną tematykę, związaną z prognozowaniem czasu trwania inwestycji, wykonywaną jednak z zastosowaniem podstawowych parametrów używanych w metodzie EVM.

PODSTAWOWE ZAŁOŻENIA METODY EARNED VALUE

Celem opracowania metody EVM było powiązanie rzeczowego zaawansowania stanu robót i poniesionych kosztów na tle wartości planowanych. Istotę założeń metody Earned Value przedstawiono na rysunku 1 (więcej informacji o metodzie EVM można znaleźć w [Webb 2008, Wilkens 1999]).

Rysunek 1. Elementy metody Earned Value



Źródło: Wilkens, 1999

Krzywa ilustrująca przebieg *BCWS*, czyli planowany koszt planowanej pracy wyznaczana jest podczas fazy planowania projektu, a jej zakończeniem jest *BAC,* czyli planowany budżet projektu. Krzywe ilustrujące przebieg *BCWP,* czyli planowany koszt wykonanej pracy oraz *ACWP*, czyli rzeczywisty koszt wykonanej pracy wyznaczane są podczas realizacji projektu i mogą zostać ustalone jedynie do dnia kontroli. Dane z krzywych *ACWP, BCWS i BCWP* stanowią podstawę do

obliczania dalszych wskaźników. Z punktu widzenia prognozowania czasu inwestycji najważniejszymi są dwa:

- *SV* (Scheduled Variance - odchylenie od harmonogramu) - jest to wskaźnik w funkcji kosztów, którego wartość ujemna oznacza opóźnienie wyrażone przez kwotę niewykorzystanego budżetu,

$$SV = BCWP - BCWS \tag{1}$$

- *SPI* (Schedule Performance Index - wskaźnik wykonania harmonogramu) - jest to wskaźnik wykonania harmonogramu w funkcji czasu (mniejszy od 100 % oznacza opóźnienie).

$$SPI = \frac{BCWP}{BCWS} \tag{2}$$

Na podstawie ww wykresu można również zapoznać się z odchyleniami od budżetu oraz harmonogramu zarówno w dniu kontroli, jak również wartościami prognozowanymi na zakończenie projektu. To, co jest bardzo istotne dla procesu prognozowania dalszego przebiegu budowy to fakt, że badania na ponad 700 dużych amerykańskich projektach wykazały, że przy niespełna 15% - 20% zaawansowaniu projektu użycie metody EVM daje możliwość przewidzenia wyniku końcowego z dużą dokładnością i to niezależnie od typu monitorowanego kontraktu, programu czy usługi [Szczurowski 2008].

METODY WYZNACZANIA CZASU ZAKOŃCZENIA INWESTYCJI

W opisywanych w artykule metodach przewidywany czas, jaki zajmie realizacja całej inwestycji (*ETTC* - Estimated Time At Completion) składa się z dwóch składników: czasu, jaki upłynął od rozpoczęcia prac do dnia aktualizacji (*AD* - Actual Duration) oraz przewidywanego czasu jaki zajmie dokończenie projektu (*PDWR* - Planned Duration of Work Remaining).

$$ETTC = AD + PDWR \tag{3}$$

Składnik *AD* nie budzi wątpliwości i zawsze jest wartością znaną i pewną. Element *PDWR* jest natomiast wartością przewidywaną, obarczoną niepewnością i podlegającą różnym metodom estymacji. Sposób jego wyznaczania zależy od przyjętej metody oraz w każdej metodzie dodatkowo od rodzaju wskaźnika, za pomocą którego chcemy odwzorować tempo przyszłych robót. Otrzymany wynik w dużej mierze zależy właśnie od wyboru tego wskaźnika, który powinien być ustalony na podstawie dotychczasowego przebiegu robót, prognozy tempa przyszłych prac , doświadczenia i wiedzy menadżera projektu.

W klasycznej metodzie EVM czas potrzebny do zakończenia robót prognozowano na podstawie następującej formuły [Webb 2008]:

$$PDWR = \frac{PD - (AD * SPI)}{SPI} \tag{4}$$

gdzie *PD* (Planned Duration) - planowany czas realizacji inwestycji, *SPI* (Schedule Performance Index) - wskaźnik wykonania harmonogramu (2). Jednak podczas licznych badań stwierdzono, że ta formuła niezbyt efektywnie wyznacza poszukiwany czas zakończenia robót i w literaturze podano kilka innych metod jego obliczania.

W artykule skupiono się na trzech podstawowych metodach prognozowania: PVM (Planned Value Method) , EDM (Earned Duration Method) i ESM (Earned Schedule Method). Poniżej podano wzory do wyznaczania prognozowanego czasu wszystkimi trzema metodami. We wzorach starano się zachować oryginalne oznaczenia, używane przez autorów tych metod, a odwołania do synonimów używanych oznaczeń podano w objaśnieniach do wzorów.

Planned Value Method (PVM)

Autorem tej metody jest Anbari [2003]. Podobnie, jak w pozostałych opisanych poniżej metodach, rozróżnił on trzy możliwe scenariusze przebiegu niewykonanych dotychczas robót:

- Jeżeli pozostały czas do zakończenia ma być zgodny z planem

$$ETTC_{PV1} = PD - TV \tag{5}$$

gdzie *TV (*Time Variance*)*

$$TV = \frac{SV}{PV_{Rate}} = \frac{SV * PD}{BAC} = \frac{\left(EV - PV\right)}{BAC} * PD \tag{6}$$

gdzie *SV (*Schedule Variance) - odchylenie od harmonogramu, czyli różnica między *BCWP* i *BCWS (1*), PV_{Rate} (Planned Value Rate) - średni, planowany przerób przypadający na przyjętą jednostkę czasu np. miesiąc, *EV* (Earned Value) - wartość wypracowana oznaczana też *BCWP, PV* (Planned Value) - planowany koszt planowanej pracy czyli końcowa wartość *BCWS, BAC* (Budget At Completion) - planowany koszt wykonania obiektu.

- Jeżeli pozostały czas do zakończenia robót ma być zgodny z aktualnym tempem prowadzenia robót odniesionym do planowanych kosztów, należy odwołać się do wskaźnika *SPI* (2)

$$ETTC_{PV2} = \frac{PD}{SPI} \tag{7}$$

- Jeżeli pozostały czas do zakończenia robót ma być zgodny nie tylko z dotychczasowym tempem odniesionym do planowanych kosztów, lecz również uwzględniać faktyczne przeroby finansowe na budowie, należy zastosować wskaźnik *SCI*

$$ETTC_{PV3} = \frac{PD}{SCI} \tag{8}$$

gdzie *SCI* (Schedule Cost Index lub *CR* - Critical Ratio) - wskaźnik krytyczny

$$SCI = SPI * CPI \tag{9}$$

gdzie *CPI* (Cost Performance Index) - wskaźnik wykonania budżetu

$$CPI = \frac{BCWP}{ACWP} \tag{10}$$

Earned Duration Method (EDM)

Autorem tej metody jest Jacob [2003]. Sumaryczny czas całej inwestycji składa się z dwóch elementów: czasu jaki już upłynął *AD* oraz czasu jaki jest potrzebny na zakończenie robót *PDWR* z indeksem wskazującym na metodę Earned Duration:

$$ETTC_{ED} = AD + PDWR_{ED} \text{ gdzie} \tag{11}$$

$$PDWR_{ED} = \frac{(PD - ED)}{PF} \tag{12}$$

gdzie *ED* (Earned Duration) to wypracowany czas, czyli czas, jaki pierwotnie był planowany na zrealizowanie tego, co zostało wykonane

$$ED = AD * SPI \tag{13}$$

a *PF* (Performance Factor) to wskaźnik wykonania zależny od charakteru obiektu, pozwalający ocenić czy dotychczasowe tempo prac zostanie utrzymane.

Jeżeli zakładamy, że tempo planowanych prac zostanie utrzymane wówczas możemy założyć *PF*=1 i wówczas:

$$PDWR_{ED1} = (PD - ED) \tag{14}$$

$$ETTC_{ED1} = AD + PDWR_{ED1} = AD + PD - AD * SPI = PD + AD * (1 - SPI) \tag{15}$$

Jeżeli zakładamy, że tempo planowanych prac będzie takie jak dotychczas, wówczas możemy założyć *PF*=*SPI* i wówczas:

$$PDWR_{ED2} = \frac{(PD - ED)}{SPI} \tag{16}$$

$$ETTC_{ED2} = AD + PDWR_{ED2} = AD + \frac{(PD - AD * SPI)}{SPI} = \frac{PD}{SPI} \tag{17}$$

Jeżeli natomiast zakładamy, że tempo planowanych prac będzie wypadkową nie tylko dotychczas utrzymywanego tempa prac, ale również dotychczasowego przerobu finansowego, wówczas możemy założyć *PF*=*SCI* i wówczas:

$$PDWR_{ED3} = \frac{(PD - ED)}{SCI} \tag{18}$$

$$ETTC_{ED3} = AD + PDWR_{ED3} = AD + \frac{(PD - ED)}{SCI} = \frac{PD}{SCI} + AD * \left(1 - \frac{1}{CPI}\right) \tag{19}$$

Jeżeli obliczenia należy wykonać w okresie, gdy termin aktualizacji *AD* przekroczył termin planowany zakończenia robót *PD* (czyli realizacja trwa już dłużej niż planowano) we wzorach (14), (16) i (18) wartość *PD* należy zastąpić *AD* czyli otrzymamy formuły:

$$PDWR_{ED1} = (AD - ED);\ PDWR_{ED2} = \frac{(AD - ED)}{SPI};\ PDWR_{ED3} = \frac{(AD - ED)}{SCI} \quad (20)$$

W metodzie EDM wprowadzony został dodatkowo jeszcze jeden wskaźnik nazwany *TCSPI* (To Complete Schedule Performance Index). Wskazuje on dodatkowy wysiłek, jaki należy podjąć, aby ukończyć projekt w założonym terminie.

$$TCSPI = \frac{(PD - ED)}{(PD - AD)} \quad (21)$$

Jeżeli termin zakończenia robót został przesunięty, za wartość *PD* w mianowniku można wstawić nowy, dyrektywny termin zakończenia robót, nazywany *LRS* (Letest Revised Schedule). Wówczas dodatkowy wysiłek można obliczyć formułą:

$$TCSPI(LRS) = \frac{(PD - ED)}{(LRS - AD)} \quad (22)$$

Earned Schedule Method (ESM)

W metodzie ESM (której autorem jest Lipke [2009]) kluczowym parametrem jest tzw *ES* (Earned Schedule) czyli czas, jaki pierwotnie był planowany na zrealizowanie tego, co zostało wykonane (podobnie jak parametr *ED* w metodzie EDM, jednak jego sposób wyznaczania jest nieco inny). W klasycznej metodzie Earned Value odchylenie od harmonogramu *SV* (1) było mierzone „w pionie", wzdłuż osi rzędnych, jako różnica *BCWP* - *BCWS* i wyrażane w jednostce walutowej. W ten sposób poprzez jednostki wyrażone w pieniądzach starano się pokazać opóźnienie (przyśpieszenie) wykonywanych robót. Taki sposób wyznaczania opóźnienie często był podnoszony jako jeden z głównych zarzutów w stosunku do metody Earned Value. W metodzie ESM parametr *ES* wyznaczany jest ze wzoru:

$$ES = C + \frac{(EV - PV_C)}{(PV_{C+1} - PV_C)} \quad (23)$$

gdzie *C* to liczba okresów aktualizacji od początku robót do dnia bieżącej aktualizacji, w których skumulowana wartość *PV* (czyli planowany koszt planowanej pracy, *BCWS*) w kolejnych okresach aktualizacji od początku robót jest mniejsza od bieżącej, skumulowanej wartości *EV* (czyli wartości wypracowanej *BCWP*), *EV* - skumulowana wartość wypracowana (*BCWP*) w okresie bieżącej aktualizacji a PV_C, PV_{C+1} to skumulowana wartość *PV* (*BCWS*) w okresie *C, C+1* itd.

Znając wartość ES wyznaczamy *PDWR* i *ETTC*:

$$ETTC_{ES} = AD + PDWR_{ES} \text{ gdzie} \quad (24)$$

$$PDWR_{ES} = \frac{(PD - ES)}{PF} \quad (25)$$

gdzie *PF* (Performance Factor) podobnie jak w metodzie EDM to wskaźnik wykonania zależny od charakteru obiektu, pozwalający ocenić czy dotychczasowe tempo prac zostanie utrzymane. W zależności od oceny projektu stosujemy:

- jeżeli zakładamy, że tempo planowanych prac zostanie utrzymane wówczas możemy założyć *PF*=1 i wówczas:

$$PDWR_{ES1} = (PD - ES);\ ETTC_{ES1} = AD + PDWR_{ES1} = AD + PD - ES \tag{26}$$

- jeżeli zakładamy, że tempo planowanych prac będzie takie jak dotychczas, wówczas możemy założyć *PF*=*SPI(t)* i wówczas:

$$PDWR_{ES2} = \frac{(PD - ES)}{SPI(t)} \tag{27}$$

$$\text{gdzie } SPI(t) = \frac{ES}{AD} \tag{28}$$

- jeżeli natomiast zakładamy, że tempo planowanych prac będzie wypadkową nie tylko dotychczas utrzymywanego tempa prac, ale również dotychczasowego przerobu finansowego, wówczas możemy założyć *PF*=*SCI(t)* i wówczas: [Vandevoorde, Vanhoucke 2006]

$$PDWR_{ES3} = \frac{(PD - ES)}{SCI(t)} \text{ gdzie} \tag{29}$$

$$SCI(t) = SPI(t) * CPI \tag{30}$$

W metodzie tej, podobnie jak w metodzie EDM, również można wyznaczyć indeksy *TCSPI* (To Complete Schedule Performance Index):

$$TCSPI(t) = \frac{(PD - ES)}{(PD - AD)} \tag{31}$$

$$TCSPI(t) - LRS = \frac{(PD - ES)}{(LRS - AD)} \tag{32}$$

gdzie *LRS* (Letest Revised Schedule) to nowy, dyrektywny termin zakończenia robot.

BADANIA EMPIRYCZNE

Obliczenia, które przeprowadzono miały dwa podstawowe cele:

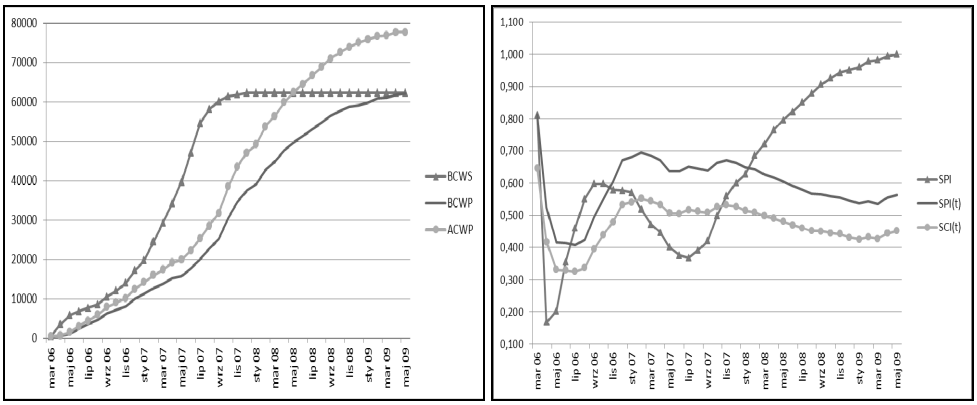
- porównanie uzyskanych opisanymi powyżej metodami wartości prognozowanych czasów zakończenia robót w zmieniających się warunkach, w zależności od zaawansowanie robót na budowie,
- sprawdzenie, który z trzech używanych we wszystkich analizowanych metodach wskaźników *PF* najlepiej nadaje się do wyznaczenia prognozy.

Obiekt, dla którego dysponowano niezbędnymi danymi i na podstawie którego przeprowadzono wszystkie opisane obliczenia [Komendarek 2010,

Połoński, Komendarek 2011], to samodzielny, duży, kubaturowy budynek użyteczności publicznej wyposażony we wszystkie niezbędne instalacje. Budynek posiada dwie kondygnacje podziemne i dziewięć kondygnacji nadziemnych. Powierzchnia całkowita budynku to ok. 16500 m^2, powierzchnia użytkowa - 9600m^2 a kubatura 75000m^3. Realizacja inwestycji zaplanowana została przez Generalnego Wykonawcę na 23 miesiące (rozpoczęcie - marzec 2006, zakończenie - grudzień 2007) a koszt realizacji na 62,267 mln. zł. Faktyczne zakończenie robót nastąpiło w maju 2009 roku po 39 miesiącach od rozpoczęcia robót a całkowity koszt robót wyniósł 77,627 mln zł. Śledzenie miesięcznych raportów na temat rzeczowego postępu robót oraz sporządzanych krzywych *BCWP* i *ACWP* na tle znanej od początku krzywej BCWS, pozwoliło na bieżąco dokonywać oceny ponoszonych wydatków na tle wartości planowanych, a w efekcie prognozować m.in. czas, jaki zajmie realizacja całej inwestycji (*ETTC*).

Wszystkie obliczenia wykonano samodzielnie w programie Excel, na podstawie przedstawionych powyżej wzorów, stosując miesięczne raporty zawierające dane na temat planowanych kosztów *BCWS*, zaawansowania rzeczowego, wartości wypracowanej *BCWP* oraz faktycznie poniesionych kosztów za wykonane prace *ACWP*. Przy wyznaczaniu wartości *ES* i *SPI(t)* w metodzie ESM skorzystano z arkusza obliczeniowego do wyznaczania tych parametrów udostępnionego przez Like [2006], przy czym arkusz ten został przez autora nieco zmodyfikowany. Poniżej w postaci wykresów przedstawiono skondensowane, najważniejsze rezultaty badań.

Rysunek 2. Wartości podstawowych parametrów charakteryzujących przebieg robót w czasie budowy

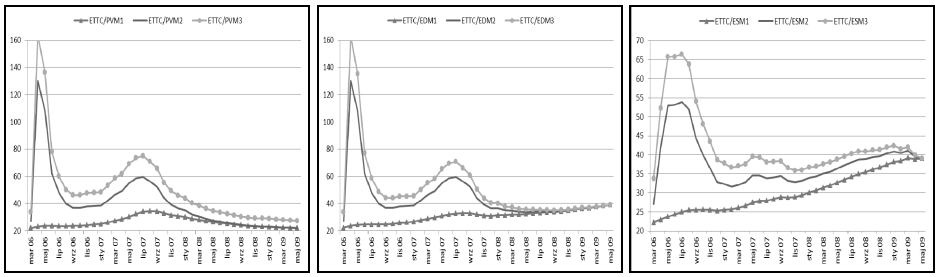


Źródło: obliczenia własne

Na rys. 2 z lewej strony przedstawiono przebieg wartości podstawowych monitorowanych parametrów a więc planowanych kosztów *BCWS*, wartości wypracowanej *BCWP* i realnie ponoszonych kosztów *ACWP*. Jak można zauważyć wartość *BCWP* od planowanego terminu zakończenia robót (12.2007) utrzymuje

się na stałym poziomie, natomiast krzywe *BCWP* i *ACWP* rosną do czasu faktycznego zakończenia prac (05.2009), przy czym ponoszone koszty wyraźnie przekraczają wartości planowane. Na tym samym rysunku z prawej strony przedstawiono przebieg wartości głównych wyliczanych wskaźników, na podstawie których, oblicza się całkowity planowany czas trwania inwestycji *ETTC*. Przebieg wartości wskaźników *SPI(t)* i *SCI(t)* w czasie jest dosyć podobny, tyko nieco przesunięty co do wartości (wskaźnik *SPI(t)* jest nieco większy od *SCI(t)*), natomiast wskaźnik SPI przyjmuje zupełnie inne wartości, szczególnie od momentu planowanego zakończenia robót.

Rysunek 3. Prognozowany całkowity czas trwania inwestycji *ETTC* wyznaczony metodami PVM, EDM i ESM przy różnych scenariuszach wyznaczania parametru *PF*



Źródło: obliczenia własne

Na rys. 3 przedstawiono zestawienie zmieniającego się przebiegu prognozowanego całkowitego czasu trwania inwestycji *ETTC* wyznaczonego metodami PVM, EDM i ESM, przy czym w każdej z tych metod rozpatrzono trzy możliwe scenariusze: pierwszy, gdy wskaźnik wykonania PF zakłada przebieg dalszych robót zgodny z planem; drugi, gdy wskaźnik *PF* ma być zgodny z dotychczasowym tempem prowadzenia robót oraz trzeci, gdy wskaźnik *PF* ma być zgodny nie tylko z dotychczasowym tempem prac, lecz również uwzględniać faktyczne uzyskiwane przeroby finansowe na budowie. We wszystkich trzech metodach łatwo zauważyć na samym początku robót charakterystyczny skok planowanych wartości *ETTC* nawet do ok. 160 miesięcy (dla drugiego i trzeciego scenariusza wskaźnika PF), przy czym w metodzie ESM jest on znacznie mniejszy, bo sięga maksymalnie 66 miesięcy. Skok ten jest spowodowany bardzo wolnym tempem prowadzonych robót w kilku pierwszych miesiącach budowy, co natychmiast przekłada się na prognozowany czas trwania robót. Jednak w dalszych miesiącach we wszystkich metodach wartości *ETTC* znacznie spadają i wahają się w granicach ok. 20 do 70 miesięcy. Rozpatrując każdą metodę oddzielnie należy zauważyć, że w przypadku metody PVM po przekroczeniu planowanego terminu zakończenia robót, prognozowane czasy *ETTC* cały czas spadają (i to przy wszystkich rodzajach wskaźników *PF*) wbrew wyraźnej tendencji faktycznego wydłużania robót. Taki wynik w dużej mierze podważa wiarygodność

uzyskiwanych wyników tą metodą. W metodach EDM i ESM tej tendencji już się nie zauważa.

W metodzie EDM po przekroczeniu planowanego terminu zakończenia prognozowane wartości *ETTC* są bardzo zbieżne dla wszystkich scenariuszy parametru *PF*, jednak biorąc pod uwagę prognozę wyznaczoną na pół roku przed planowanym zakończeniem robót (07.2007), odbiega ona dosyć znacznie od ostatecznej wartości.

Wyraźnie najlepsze rezultaty osiągnięto metodą ESM przy zastosowaniu drugiego i trzeciego scenariusza wskaźnika *PF*. Szacowane wartości prognozowanego czasu *ETTC* są najbardziej stabilne i już po roku prowadzenia robót (to znaczy po ok. 30% faktycznego czasu realizacji prac) ustalają się w zakresie ok. 35-40 miesięcy, co potwierdziło się w praktyce.

PODSUMOWANIE

Dysponując wiarygodnym harmonogramem planowanych robót budowlanych oraz prowadząc właściwie opracowaną metodycznie kontrolę postępu prac można wiarygodnie prognozować całkowity czas trwania robót a tym samym termin zakończenia budowy. Do wyznaczenia takiej prognozy należy dysponować regularnymi, wiarygodnymi raportami na temat rzeczowego zaawansowania planowanych robót, kosztów planowanych *BCWS*, wartości wypracowanej *BCWP* oraz faktycznych ponoszonych kosztów *ACWP*. Tak zebrane dane stanowią podstawę do wyliczenia dalszych wskaźników obrazujących w syntetycznej formie przebieg prac na budowie i wyznaczenia prognozy terminu zakończenia robót *ETTC*.

Rozpatrując uzyskane wyniki obliczeń na szczególną uwagę zasługuje metoda Earned Schedule (ESM). Pozwala ona najbardziej stabilnie i wiarygodnie prognozować całkowity czas trwania inwestycji, przy czym wskazane jest równoczesne rozpatrywanie drugiego i trzeciego scenariusza przebiegu dalszych robót (wskaźnika *PF*) a więc opartego na założeniu utrzymywania dotychczasowego tempa prac (wyznaczonego wskaźnikiem *SPI(t)*) oraz równoczesnego uwzględnienia dotychczasowego tempa robót i przerobu finansowego (wyznaczanego wskaźnikiem *SCI(t)*). Przy wykonywaniu obliczeń tą metodą pomocny jest arkusz obliczeniowy udostępniony przez autora tej metody [Lipke 2006] do obliczania wartości *ES* i *SPI(t)*, a dalsze obliczenia można wykonać samemu np. w programie Excel w oparciu o podane w artykule wzory.

Wskazane jest prowadzenie dalszych badań nad możliwością prognozowania czasu trwania robót budowlanych, szczególnie z uwzględnieniem ich specyficznych warunkach realizacji, zwłaszcza metodą ESM.

BIBLIOGRAFIA

Anbari F. (2003) Earned value method and extensions, Project Manage Journal, Dec. 34(4), p.12–23.

Jacob D. (2003) Forecasting project schedule completion with earned value metrics, The Measurable News (March), 1. p. 7–9.

Lipke W. (2009) Project Duration Forecasting ...a comparison of Earned Value Management methods to Earned Schedule, The Measurable News (May) 2, p. 24-31.

Lipke W. (2006) http://www.earnedschedule.com/Calculator.shtml

Komendarek P (2010) Bieżąca kontrola kosztów wykonania robót budowlanych na przykładzie budowy Naczelnego Sądu Administracyjnego przy ul. Boduena w Warszawie, maszynopis, SGGW.

Połoński M., Komendarek P. (2011) Bieżąca kontrola kosztów realizacji obiektu budowlanego metodą earned value, Metody ilościowe w badaniach ekonomicznych, Tom XII/2 - s. 279-290.

Skorupka D. (2009) Method of planning construction projects taking into account risk factors, Operations Research and Decision, Wroclaw, p. 119-128.

Szczurowski L. (2008) Metoda Earned Value - materiał wykładowy http://www.ioz.pwr.wroc.pl/ pracownicy/szczurowski/piwz/EV.pdf

Webb A. (2008) Wartość wypracowana w praktyce, PROED, Warszawa.

Wilkens T. (1999) Earned Value, Clear and Simple, Primavera Systems, materiały w formacie PDF.

Vandevoorde S., Vanhoucke M. (2006) A comparison of di□erent project duration forecasting methods using earned value metrics. International Journal of Project Management 24 (2006) s. 289–302.

FORECASTING CIVIL STRUCTURE DURATION ON THE BASIS OF PROGRESS OF WORKS

Abstract: During construction works one of the most important duties of a project manager is to control progress of construction works with relation to scheduled deadlines and budget. A method allowing following both of aforementioned indicators with relation to progress of works is the EVM. This article focuses on forecasting civil structure duration. In literature one can find various methods of how to use chosen indicators to forecast civil structure duration. The aim of this article is to list few most important of this methods and show results of each, using the example of a large enclosed structure constructed in Warsaw in last years. The article reveals one method which, in given situation, proved to be the most accurate against real time of duration of construction works.

Keywords: earned value, earned schedule, forecasting duration, civil structure

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 180 – 190

WPŁYW METODY DOBORU CECH NA EFEKTYWNOŚĆ KLASYFIKACJI NA PRZYKŁADZIE ANALIZY JAKOŚCI ŻYCIA W ŚWIETLE ZRÓWNOWAŻONEGO ROZWOJU

Agnieszka Sompolska-Rzechuła
Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: asompolska@zut.edu.pl

Streszczenie: W artykule podjęto próbę odpowiedzi na pytanie: Czy wyniki otrzymane za pomocą różnych metod doboru cech mają wpływ na efektywność klasyfikacji? Do badania wykorzystano dwie metody doboru cech: parametryczną metodę coraz metodę odwróconej macierzy współczynników korelacji. Skuteczność grupowań zweryfikowano za pomocą wskaźników homogeniczności, heterogeniczności i poprawności skupień. W ocenie efektywności grupowań wykorzystano podejście z medianą Webera. Badanie dotyczyło powiatów województwa zachodniopomorskiego pod względem jakości życia w świetle zrównoważonego rozwoju.

Słowa kluczowe: metoda doboru cech, efektywność klasyfikacji, jakość życia, rozwój zrównoważony

WPROWADZENIE

Celem artykułu jest próba odpowiedzi na pytanie: Czy wyniki otrzymane za pomocą różnych metod doboru cech mają wpływ na efektywność klasyfikacji? Dobór cech jest jednym z najważniejszych, a jednocześnie najtrudniejszych zagadnień. Niezbędna jest kompleksowa znajomość analizowanego zagadnienia oraz specyfiki powiązań pomiędzy zjawiskami społeczno-gospodarczymi. Od jakości zestawu cech zależy wiarygodność ostatecznych wyników i trafność podejmowanych decyzji [Gatnar, Walesiak 2004]. Próbę odpowiedzi na postawione pytanie podjęto na podstawie badania taksonomicznego powiatów ziemskich województwa zachodniopomorskiego, dokonując klasyfikacji obiektów

na podstawie zbioru cech otrzymanych metodami: parametryczną oraz metodą odwróconej macierzy współczynników korelacji. Podziału powiatów dokonano pod względem obiektywnej jakości życia mieszkańców w świetle zrównoważonego rozwoju w roku 2010.

Istnieje bardzo wiele określeń i klasyfikacji jakości życia, a problemami z jej zakresu zajmują się przedstawiciele wielu dyscyplin naukowych (filozofii, socjologii, psychologii, ekonomii oraz statystyki). Jako kategoria wyrażająca stopień samorealizacji człowieka jakość życia powinna być podstawowym przedmiotem zainteresowania społeczeństwa. Jak podaje Tadeusz Borys [Borys 2008] trzy kategorie: jakość życia, rozwój społeczny, gospodarczy i środowiskowy oraz instrumentarium tego rozwoju tworzą hierarchiczny układ pojęć ściśle ze sobą związanych i powinny być przedmiotem zintegrowanego pomiaru wskaźnikowego. Wspólne cechy zmian rozwojowych jakości życia i rozwoju zrównoważonego[1] znajdują odzwierciedlenie w powiązaniu opisu wskaźnikowego. Duże znaczenie mają wskaźniki rozwoju zrównoważonego w opisie pośredniej jakości życia oraz przy tworzeniu pośrednich wskaźników jakości życia. Większość wskaźników zrównoważonego rozwoju tworzy pośredni obraz obiektywnej jakości życia. Trwały i zrównoważony rozwój, w większości definicji, postrzegany jest jako taki sposób gospodarowania, który prowadzi do poprawy jakości życia.

Obliczenia zostały wykonane w arkuszu kalkulacyjnym Excel oraz programach: Statistica i **R**.

OPIS METODY

Zastosowanie metod wielowymiarowej analizy porównawczej wymaga wyboru obiektów oraz zbioru cech diagnostycznych charakteryzujących poszczególne obiekty.

Po określeniu i zgromadzeniu danych dotyczących wstępnego zestawu cech należy podjąć odpowiednie działania weryfikacyjne według dwóch najistotniejszych kryteriów [Młodak 2006]:

1. Zmienność – cechy powinny wykazywać odpowiednią zmienność, czyli skutecznie dyskryminować obiekty. Do oceny zmienności wartości cech wykorzystuje się współczynnik zmienności:

$$v_j = \frac{s_j}{\bar{x}_j} \tag{1}$$

[1] W literaturze można znaleźć wiele określeń zrównoważonego rozwoju, przykładem ujęcia ogólnego jest definicja trwałego i zrównoważonego rozwoju, według której: „istotą rozwoju zrównoważonego i trwałego jest zapewnienie trwałej poprawy jakości życia współczesnych i przyszłych pokoleń poprzez kształtowanie właściwych proporcji między trzema rodzajami kapitału: ekonomicznym, ludzkim i przyrodniczym" Piontek F. (2001) Ekonomia a rozwój zrównoważony, Ekonomia i środowisko, str. 19.

gdzie: $\bar{x}_j$ to średnia arytmetyczna wartości cechy X_j, zaś s_j jest odchyleniem standardowym j-tej cechy, $j = 1, \ldots, m$, m – liczba cech.

2. Korelacja – dwie cechy silnie ze sobą skorelowane są nośnikami podobnej informacji, zatem jedna z nich jest zbędna. Do oceny siły związku między cechami stosuje się współczynnik korelacji. Punktem wyjścia jest macierz współczynników korelacji między wszystkimi parami cech:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdots & 1 \end{bmatrix} \tag{2}$$

gdzie: r_{jk} to współczynniki korelacji liniowej Pearsona j-tej i k-tej cechy.

Metodą wykorzystywaną do dyskryminacji cech bazującą na macierzy współczynników korelacji jest metoda parametryczna, która jest wygodna w użyciu, ponieważ jest prosta rachunkowo.

Metoda parametryczna posiada jednak dwie zasadnicze wady [Młodak 2006, Panek 2009]:

- jest wrażliwa na wartości odstające, co oznacza, że na wysoką wartość współczynnika korelacji może, w dużym stopniu, wpływać jej wysokie skorelowanie nawet z jedną z cech,
- uwzględnia wyłącznie bezpośrednie powiązania cechy z innymi cechami, nie uwzględniając powiązań pośrednich.

Skutecznym sposobem zniwelowania pierwszej niedogodności jest zastąpienie w pierwszym kroku sumy elementów kolumny (wiersza) macierzy **R** przez ich medianę. Pozwala to uodpornić analizę na zaburzenia spowodowane przez obserwacje odstające.

Druga wada może być wyeliminowana poprzez zastosowanie metody odwróconej macierzy współczynników korelacji [Panek 2009, Malina, Zeliaś 1997]. Procedura eliminacji jest następująca: korzystając z macierzy współczynników korelacji **R**, wyznacza się macierz $\mathbf{R}^{-1} = [r^{ij}]$, gdzie wartości r^{ij} są elementami macierzy odwrotnej $\mathbf{R}^{-1}$. Element diagonalny r^{ii} macierzy $\mathbf{R}^{-1}$ jest równy jedności, jeśli zmienna X_j jest ortogonalna względem pozostałych zmiennych. W przypadku nieortogonalności $r^{ii} \in (1, +\infty)$, gdy zmienna jest nadmiernie skorelowana z pozostałymi, elementy diagonalne macierzy odwrotnej $\mathbf{R}^{-1}$ są znacznie większe od jedności, co jest symptomem złego uwarunkowania macierzy **R**.

Cechy nadmiernie skorelowane, którym odpowiadają elementy diagonalne r^{ii} o wartościach większych niż 10, są eliminowane z pierwotnego zbioru cech. Jeżeli takie elementy nie występują, to procedurę uznaje się za zakończoną. Ponownie wyznacza się macierz odwrotną $\mathbf{R}^{-1}$ dla zredukowanego zbioru cech i analizuje jej

elementy diagonalne. Procedurę powtarza się do momentu osiągnięcia stabilności macierzy $\mathbf{R}^{-1}$, czyli pojawienia się elementów diagonalnych, których wartości nie przekraczają znacząco 10.

Otrzymany zbiór cech diagnostycznych stanowił podstawę klasyfikacji obiektów. Spośród wielu metod hierarchicznych do badania wybrano metodę Warda, która różni się od wszystkich pozostałych metod tym, że do oszacowania odległości między skupieniami wykorzystuje się podejście analizy wariancji. Metoda ta zmierza do minimalizacji sumy kwadratów odchyleń dowolnych dwóch hipotetycznych skupień, które mogą zostać uformowane na każdym etapie analizy. Ważną cechą tej metody jest zapewnienie minimalizacji kryterium wariancyjnego, które głosi, że wariancja wewnątrz skupień jest minimalna. Metoda Warda zapewnia zatem homogeniczność wewnątrz skupień i heterogeniczność między skupieniami, przez co uznawana jest za bardzo efektywną [Ward 1963].

Ostatnim etapem analizy taksonomicznej obiektów jest sprawdzenie jakości uzyskanych podziałów. Do oceny jakości klasyfikacji stosuje się mierniki homogeniczności oraz heterogeniczności skupień, wykorzystując koncepcję środka ciężkości grupy i odległości od niego. W badaniu wykorzystano podejście, w którym środek ciężkości danej grupy zastąpiony został medianą Webera jej elementów. Mediana Webera stanowi wielowymiarowe uogólnienie klasycznego pojęcia mediany. Chodzi o wektor, który minimalizuje sumę euklidesowych odległości od danych punktów reprezentujących rozpatrywane obiekty, a więc znajduje się niejako „pośrodku" nich, ale jest jednocześnie uodporniony na występowanie obserwacji odstających [Młodak 2006].

W ocenie homogeniczności otrzymanych grup wykorzystano miernik o następującej postaci [Młodak 2006]:

$$hm_{6\,\max}^{*} = \max_{k=1,\ldots,p} hm_6^{*}\left(P_k\right) \tag{3}$$

gdzie:

$$hm_6^{*}\left(P_k\right) = \operatorname*{med}_{i:O_i \in P_k} \delta\left(O_i, \Gamma_{\theta k}\right) \tag{4}$$

jest medianą odległości obiektów grupy P_k od jej wektora medianowego Webera,

$$\Gamma_{\theta k} = \left(\theta_{1P_k}, \theta_{2P_k}, \ldots, \theta_{mP_k}\right) \tag{5}$$

jest wektorem medianowym Webera, δ - odległość obiektów grupy P_k od jej wektora medianowego Webera, O_i - obiekty, θ_{mP_k} - mediana Webera rozpatrywanego układu m cech diagnostycznych, k – liczba klas, $k = 1,2,\ldots,p$, p – liczba skupień otrzymanych na danym poziomie grupowania.

Natomiast w ocenie heterogeniczności zastosowano miernik:

$$ht_{6\,\mathrm{mn}}^{*} = \min_{k=1,\ldots,p} ht_{6}^{*}(P_k) \tag{6}$$

gdzie:

$$ht_{6}^{*}(P_k) = \operatorname*{med}_{\substack{i=1,\ldots,p \\ i\neq k}} \delta(\Gamma_{\theta i}, \Gamma_{\theta k}) \tag{7}$$

jest medianą odległości pomiędzy medianą Webera danej grupy z analogicznymi wektorami dla pozostałych grup.

W ocenie poprawności grupowania wykorzystano kompleksowy miernik o postaci:

$$ct_{6}^{*} = \frac{hm_{6\,\mathrm{mx}}^{*}}{ht_{6\,\mathrm{mn}}^{*}} \tag{8}$$

CHARAKTERYSTYKA MATERIAŁU BADAWCZEGO

Doboru cech diagnostycznych dokonano za pomocą dwóch metod: parametrycznej oraz odwróconej macierzy współczynników korelacji. Wstępna lista cech diagnostycznych obejmowała wskaźniki ujęte w grupy i kategorie wskaźników zrównoważonego rozwoju, które zostały przedstawione w tabeli 1.

Tabela 1. Grupy i kategorie wskaźników zrównoważonego rozwoju

Ład środowiskowy	Ład ekonomiczny	Ład społeczny
• Zmiany klimatu • Energia • Ochrona powietrza • Ekosystemy morskie • Zasoby słodkiej wody • Użytkowanie gruntów • Bioróżnorodność • Gospodarka odpadami	• Rozwój gospodarczy • Zatrudnienie • Innowacyjność • Transport • Zrównoważone wzorce produkcji	• Zmiany demograficzne • Zdrowie publiczne • Integracja społeczna • Edukacja • Dostęp do rynku pracy • Bezpieczeństwo publiczne • Zrównoważone wzorce konsumpcji

Źródło: opracowanie własne na podstawie [Wskaźniki zrównoważonego rozwoju Polski, GUS 2011]

Do opisu jakości życia przyjęto następujący zestaw cech [Województwo zachodniopomorskie, podregiony, powiaty, gminy 2011],:

X_1 - ludność w wieku nieprodukcyjnym na 100 osób w wieku produkcyjnym,
X_2 - małżeństwa zawarte na 1000 ludności,
X_3 - urodzenia żywe na 1000 ludności,
X_4 - zgony niemowląt na 1000 ludności,
X_5 - przyrost naturalny na 1000 ludności,
X_6 - rozwody na 1000 ludności,

X_7 - separacje na 100 tys. ludności,
X_8 - saldo migracji na 1000 ludności,
X_9 - liczba ludności na 1 placówkę biblioteczną,
X_{10} - liczba ludności na 1 instytucję kultury,
X_{11} - liczba ludności na 1 lekarza,
X_{12} - liczba ludności na 1 aptekę i punkt apteczny,
X_{13} - beneficjenci pomocy społecznej w % ogółu ludności,
X_{14} - osoby niepełnosprawne poniżej 16 roku życia na 1000 ludności poniżej 16 roku życia,
X_{15} - osoby niepełnosprawne powyżej 16 roku życia na 1000 ludności powyżej 16 roku życia,
X_{16} - korzystający z noclegów na 1000 ludności,
X_{17}- udzielone noclegi na 1000 ludności,
X_{18} - lesistość w %,
X_{19} - ludność korzystająca z oczyszczalni ścieków w % ludności ogółem,
X_{20} - emisja zanieczyszczeń pyłowych w tonach na km^2,
X_{21} - emisja zanieczyszczeń gazowych w tonach na km^2,
X_{22} - odpady wytworzone w ciągu roku w tys. t na km^2,
X_{23} - stopień redukcji wytworzonych zanieczyszczeń w %,
X_{24} - udział powierzchni o szczególnych walorach przyrodniczych prawnie chronionej w powierzchni powiatu (w%),
X_{25} - udział rezerwatów w powierzchni o szczególnych walorach przyrodniczych prawnie chronionej (w %),
X_{26} - pomniki przyrody na km^2,
X_{27} - nakłady na ochronę środowiska w tys. zł na km^2,
X_{28} - podmioty gospodarki narodowej w sektorze prywatnym na 1000 ludności,
X_{29} - stopa bezrobocia rejestrowanego (w %),
X_{30} - bezrobotne kobiety w liczbie bezrobotnych ogółem w %,
X_{31} - bezrobotni trwale bezrobotni w liczbie bezrobotnych ogółem w %,
X_{32} - długość sieci wodociągowej w km na 1 km^2,
X_{33} - długość sieci kanalizacyjnej w km na 1 km^2,
X_{34} - zasoby mieszkaniowe na 1000 ludności,
X_{35} - mieszkania w miastach wyposażone w łazienkę w % ogółu mieszkań,
X_{36} - mieszkania w miastach wyposażone w gaz z sieci w % ogółu mieszkań,
X_{37} - drogi publiczne powiatowe o twardej nawierzchni w km na 1 km^2,
X_{38} - wypadki drogowe na 10 tys. ludności,
X_{39} - śmiertelne ofiary wypadków drogowych na 10 tys. ludności.

Metoda parametryczna pozwoliła na wyodrębnienie następującego zbioru cech diagnostycznych: X_4, X_{10}, X_{19}, X_{22}, X_{24}, X_{27}, X_{30}, X_{34}, X_{37}.

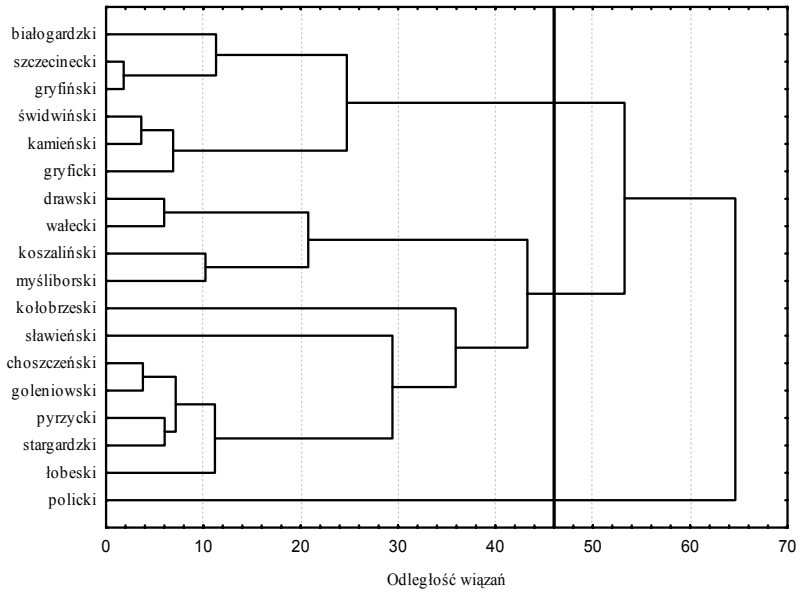
Natomiast wykorzystanie metody odwróconej macierzy współczynników korelacji doprowadziło do uzyskania zbioru cech: X_{10}, X_{23}, X_{24}, X_{32}, X_{33}, X_{34}, X_{37}, X_{38}, X_{39}.

WYNIKI BADANIA

Wykorzystując zbiór cech diagnostycznych uzyskany metodami: parametryczną i odwróconej macierzy współczynników korelacji dokonano klasyfikacji powiatów ziemskich województwa zachodniopomorskiego. Analizując dendrogramy otrzymano po trzy skupienia powiatów.

Dendrogram uzyskany metodą Warda na podstawie zbioru cech otrzymanych drugą metodą został przedstawiony na rysunku 1.

Rysunek 1. Dendrogram powiatów ziemskich województwa zachodniopomorskiego



Źródło: opracowanie własne

W tabeli 2 przedstawiono wyniki grupowania powiatów ziemskich metodą Warda na podstawie zbiorów cech uzyskanych metodami: parametryczną i odwróconej macierzy współczynników korelacji.

Tabela 2. Wyniki grupowania powiatów ziemskich metodą Warda na podstawie zbiorów cech uzyskanych metodami: parametryczną i odwróconej macierzy współczynników korelacji

Grupowanie powiatów metodą Warda na podstawie zbioru cech uzyskanych metodą					
parametryczną			odwróconej macierzy współczynników korelacji		
grupa I	grupa II	grupa III	grupa I	grupa II	grupa III
wałecki, drawski, goleniowski, choszczeński, kołobrzeski stargardzki, szczecinecki, myśliborski, pyrzycki, sławieński, policki, gryficki	świdwiński, kamieński, gryfiński,	białogardzki, koszaliński, łobeski	policki	łobeski, stargardzki, pyrzycki, goleniowski, choszczeński, sławieński, kołobrzeski, myśliborski, koszaliński, wałecki, drawski	gryficki, kamieński, świdwiński, gryfiński, szczecinecki, białogardzki

Źródło: opracowanie własne

Skuteczność grupowań zweryfikowano wyznaczając wartości wskaźników homogeniczności, heterogeniczności i poprawności skupień (tabela 3) [Młodak 2006].

Tabela 3. Wartości wskaźników homogeniczności, heterogeniczności i poprawności skupień

Wskaźniki	Wariant oparty na zbiorze cech uzyskanych metodą odwróconej macierzy współczynników korelacji	Wariant oparty na zbiorze cech uzyskanych metodą parametryczną
homogeniczności skupień	390,81	733,80
heterogeniczności skupień	1846,72	732,30
poprawności skupień	0,21	1,00

Źródło: obliczenia własne

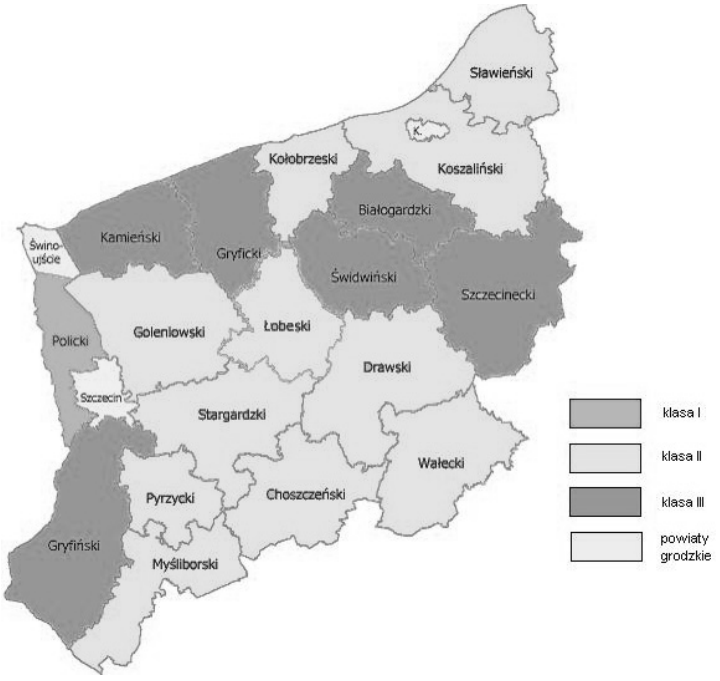
Analizując wyniki dotyczące efektywności grupowań, przedstawione w tabeli 3, można stwierdzić, że klasyfikacja otrzymana metodą Warda na podstawie zbioru cech uzyskanych metodą odwróconej macierzy współczynników korelacji dała lepsze rezultaty pod każdym względem, czyli zarówno homogeniczności i heterogeniczności, jak i poprawności grupowania w porównaniu z klasyfikacją otrzymaną z wykorzystaniem metody parametrycznej.

Klasyfikacja oparta na metodzie odwróconej macierzy współczynników korelacji wyodrębniła trzy skupienia, wśród których jest skupienie jednoelementowe – powiat policki. Powiat ten charakteryzuje się najwyższym

stopniem redukcji wytworzonych zanieczyszczeń (w %) i wysokim udziałem powierzchni o szczególnych walorach przyrodniczych prawnie chronionej w powierzchni powiatu, jak również korzystnymi wartościami wskaźników dotyczących infrastruktury technicznej (np. długość sieci kanalizacyjnej w km na 1 km^2 oraz drogi publiczne powiatowe o twardej nawierzchni w km na 1 km^2). Klasa druga z największą liczbą powiatów (11) wyróżnia się najniższym stopniem redukcji wytworzonych zanieczyszczeń (ponad siedmiokrotnie niższym w porównaniu ze średnią ogólną) ale wysokim udziałem powierzchni o szczególnych walorach przyrodniczych prawnie chronionej w powierzchni powiatu. Pozostałe wskaźniki oscylują wokół średnich ogólnych. Trzecia klasa wyróżnia się niekorzystnie pod względem bardzo niskiego stopnia redukcji wytworzonych zanieczyszczeń (w %) oraz najniższą wartością długości dróg publicznych powiatowe o twardej nawierzchni w km na 1 km^2.

Podział powiatów województwa zachodniopomorskiego pokazuje rys. 2.

Rysunek 2. Podział powiatów województwa zachodniopomorskiego metodą Warda na podstawie zbioru cech otrzymanych metodą odwróconej macierzy współczynników korelacji



Źródło: opracowanie własne

PODSUMOWANIE

W pracy przedstawiono próbę odpowiedzi na pytanie dotyczące wpływu zbiorów cech diagnostycznych otrzymanych różnymi metodami doboru cech na efektywność klasyfikacji. W badaniu wykorzystano dwie metody doboru cech: parametryczną i odwróconej macierzy współczynników korelacji. Parametryczna procedura doboru cech posiada dwie niedogodności, które są niwelowane w metodzie odwróconej macierzy współczynników korelacji. Otrzymane zbiory posłużyły do klasyfikacji ziemskich powiatów województwa zachodniopomorskiego pod względem obiektywnej jakości życia w świetle zrównoważonego rozwoju. Badanie dotyczyło 2010 roku. Efektywność klasyfikacji zbadano wykorzystując wskaźniki homogeniczności, heterogeniczności oraz poprawności grupowań, w których role środków ciężkości odgrywała mediana Webera. Klasyfikacja wykorzystująca zbiór cech uzyskany metodą odwróconej macierzy współczynników korelacji dała lepsze rezultaty pod względem wszystkich trzech kryteriów.

Badanie wykazało, iż metody klasyfikacji mogą być skutecznym narzędziem w ocenie jakości życia mieszkańców, a wyniki metod doboru cech do badania taksonomicznego mają wpływ na jakość i na rezultaty klasyfikacji.

BIBLIOGRAFIA

Borys T., Rogala P. (red.) (2008) Jakość życia na poziomie lokalnym – ujęcie wskaźnikowe, Program Narodów Zjednoczonych ds. Rozwoju, Warszawa, str. 9-10

Gatnar E., Walesiak M. (2004) Metody statystycznej analizy wielowymiarowej w badaniach marketingowych, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, str. 320

Młodak A. (2006) Analiza taksonomiczna w statystyce regionalnej, Difin, Warszawa, str. 31

Panek T. (2009) Statystyczne metody wielowymiarowej analizy porównawczej, Szkoła Główna Handlowa w Warszawie, str. 22

Malina A., Zeliaś A. (1997) O budowie taksonomicznej miary jakości życia, Taksonomia 4, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, str. 238-263

Ward J. H. (1963), Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association, No. 58

Piontek F. (2001) Ekonomia a rozwój zrównoważony, Ekonomia i środowisko, Białystok, str. 19

Województwo zachodniopomorskie, podregiony, powiaty, gminy (2011), Urząd Statystyczny w Szczecinie

Wskaźniki zrównoważonego rozwoju Polski (2011), Główny Urząd Statystyczny, Urząd Statystyczny w Katowicach

THE INFLUENCE OF THE METHOD OF THE FEATURE SELECTION ON THE CLASSIFICATION'S EFFICIENCY BASED ON THE QUALITY OF LIFE IN LIGHT ON THE SUSTAINABLE DEVELOPMENT

Abstract: In the article attempts to answer the question: Do the results, obtained by means of the various feature selection method, have any influence on the classification's efficiency? For the analysis two methods were used: parametric method and the matrix inverse method of the correlation coefficients. The effectiveness of classifications was checked by use of homogeneity, heterogeneity and correctness of clustering coefficients. The approach was used in the assessment of the classification's efficiency, with the center of gravity replaced with the Weber's median. The analysis was local and concerned the districts in zachodniopomorskie province in terms of the quality of life in the light of sustainable development.

Keywords: the feature selection method, the classification's efficiency, quality of life, sustainable development

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 191 – 201

KONWERGENCJA TECHNOLOGICZNA A POTENCJAŁ SPOŁECZNO-TECHNOLOGICZNY KRAJÓW UNII EUROPEJSKIEJ

Elżbieta Soszyńska
Instytut Stosowanych Nauk Społecznych, Uniwersytet Warszawski
e-mail: esosz@uw.edu.pl

Streszczenie: W artykule przedstawione są rezultaty estymacji szeregu modeli luki technologicznej dla weryfikacji następujących hipotez w UE27 w latach 1995-2009. Po pierwsze, wzrost TFP jest określany przez siłę procesu krajowych innowacji oraz szybkość imitacji. Po drugie, społeczno-technologiczny potencjał gospodarki określa rozmiary jej innowacji oraz możliwości eksploatacji imitacji. Po trzecie, trzy potencjalne determinanty wzrostu TFP, a mianowicie ICT, pracownicy z przygotowaniem uniwersyteckim oraz bariery regulacyjne mogą stanowić ważne stymulatory bądź bariery wzrostu produktywności.

Słowa kluczowe: ogólna produktywność czynników produkcji, luka technologiczna, innowacje, zdolność absorpcji

WPROWADZENIE

Timmer, Inklaar, O'Mahony i Van Ark [2010] przeprowadzając badania wzrostu gospodarczego w Europie w latach 1950 – 2006 doszli między innymi do konkluzji, że zróżnicowanie wzrostu produktywności pracy wśród krajów Unii Europejskiej w ostatnich dekadach jest głównie skutkiem różnic w efektywności, z jaką nakłady podstawowych czynników produkcji są wykorzystywane. Zjawisko to mierzone jest zwyczajowo wzrostem ogólnej produktywności czynników produkcji (total factor productivity- TFP).

W pracy na podstawie danych empirycznych weryfikowana jest jako główna hipoteza badawcza, iż na przełomie XX i XXI wieku w UE27 występowała konwergencja w technologii, zaś głównym jej mechanizmem były imitacje a nie innowacje. Hipotezami pomocniczymi, jest twierdzenie, że dla wprowadzenia gospodarki na ścieżkę wzrostu stymulowanego innowacjami, niezbędne jest

zbudowanie odpowiedniego potencjału społeczno-technologicznego kraju, a także wraz ze zbliżaniem się do światowej granicy czy też światowego obszaru najnowszych technologii rośnie rola szkolnictwa wyższego i przyjaznego klimatu dla zakładania i zamykania firm, czyli tak zwanego procesu deregulacji, natomiast maleje znaczenie w procesie generowania wzrostu TFP technologii ogólnego zastosowania, czyli nowych technologii komunikowania się.

W celu rozpoznania przyczyn występujących między krajami różnic stóp wzrostu TFP na ogół wykorzystuje się modele wzrostu TFP, których podstawy teoretyczne bazują na teorii luki technologicznej. Według tej teorii zacofanie w technologii krajów rozwijających się w odniesieniu do krajów liderów światowej technologii stanowi potencjał dla tych pierwszych, który względnie tanio można uruchomić w procesie doganiania [Gerschenkron, 1962]. Jednak potencjał ten można uaktywnić jedynie po spełnieniu pewnych warunków, które w literaturze przedmiotu nazywane są między innymi zdolnością do absorpcji innowacji. W modelu luki technologicznej wzrost TFP jest określany przez siłę krajowej działalności innowacyjnej oraz zakres i szybkość imitacji najlepszych technologii światowych, niezależnie gdzie zostały stworzone[1].

Porównanie poziomu i stopy wzrostu TFP po drugiej wojnie światowej w krajach Unii Europejskiej (UE) i USA prowadzi do wniosku, że od 1995 roku wzrost TFP znacznie został spowolniony w UE i odwrotnie przyspieszony w USA [Timmer i in. 2010]. Analizując powyższe tendencje Aghion i Howitt [2004] argumentują fakt zaistnienia tych rozbieżności między krajami UE i USA zdolnością czy też gotowością społeczno-technologiczną tych krajów do tworzenia innowacji i eksploatowania możliwości imitacyjnych. Między innymi w ich opinii tradycyjnie europejskie instytucje – w sensie reguł gry społeczno-ekonomicznej – w wyższym stopniu sprzyjają procesowi zmniejszania luki technologicznej niż generowania innowacji. Wśród argumentów wyróżniają poniższe. Systemy edukacji w Europie są bardziej nakierowane na przygotowanie zawodowe niż doskonalenia jakości kształcenia w szkołach wyższych. Rynki kapitałowe są w znacznie wyższym stopniu nastawione na współpracę z dużymi firmami inkubatorami niż start-up'ami. Regulacje rynku pracy raczej promują doskonalenie zawodowe, a wstrzymują proces realokacji siły roboczej między firmami. Zaś budowane systemy innowacji, w tym prawo w sferze ochrony patentowej oraz publiczne instytuty badawczo-rozwojowe stymulują tworzenie innowacji przyrostowych aniżeli przełomowych.

W celu zweryfikowania powyższych idei w pierwszych rozdziale tego przyczynku badawczego w standardowym modelu doganiania (catch-up model)

[1] Innowacje na ogół są definiowane jako nowe idee, pomysły w procesie tworzenia produktów, procesów i innej działalności firm, które prowadzą do zwiększenia wartości dodanej. W dalszej analizie przyjmuje się za innowacje nowość w stosunku do krajów, liderów w zakresie światowej technologii. W pozostałych przypadkach mamy do czynienia z imitacjami światowych technologii, ich kopiowaniem.

w technologii uwzględnione będą dodatkowe zmienne objaśniające, ilustrujące przede wszystkim różnice instytucjonalne, występujące w krajach UE27. Tego typu zmienne mogą wpływać na tempo procesu innowacji i szybkość imitacji technologii. W rozdziale drugim do modeli wprowadzone będą syntetyczne zmienne objaśniające, reprezentujące z jednej strony działalność innowacyjną i imitacyjną zaś z drugiej potencjał społeczno-technologiczny krajów jako podstawowe uwarunkowanie procesu generowania innowacji oraz absorpcji i adaptacji innowacji niezależnie od miejsca ich kreowania. Zarówno działalność innowacyjna jak imitacyjna oraz potencjał społeczno-technologiczny czy też gotowość społeczno-technologiczna są pojęciami wieloaspektowymi. Stąd też dla oszacowania tych potencjalnych zmiennych objaśniających wykorzystano analizę czynnikową. Ostatnia część opracowania zawiera wnioski z badań.

Badanie przeprowadzono na danych panelowych dla 27 krajów UE za okres 1995 – 2009 wyszczególniając trzy panele, a mianowicie dla lat: 1995-1999, 2000-2004, 2005-2009. W celu zweryfikowania poszczególnych hipotez badawczych oszacowano kilkadziesiąt modeli ekonometrycznych, z których wyniki estymacji wybranych kilkunastu, w pełni poprawnie zweryfikowanych pod względem formalno-statystycznym, zamieszczono w poniższych tabelach.

MODEL LUKI TECHNOLOGICZNEJ W PROCESIE GENEROWANIA WZROSTU TFP

Podstawowym, wyjściowym modelem dla zweryfikowania wyżej wyszczególnionych hipotez badawczych, jest model o poniższej specyfikacji:

$$growthTFP = \beta(gaptech) + \gamma X + \delta X * gaptech + \varepsilon \quad (1)$$

gdzie: growth TFP – stopa wzrostu ogólnej produktywności czynników produkcji; X – oznacza jedną spośród szeregu potencjalnych determinant wzrostu TFP, będących w kręgu zainteresowań makropolityków; β – podstawowy parametr modelu, za pomocą którego szacuje się wagę imitacji technologii w procesie wyjaśniania zmian w TFP, która zależy od rozmiarów luki technologicznej; γ – reprezentuje bezpośrednie efekty czynnika X we wzroście TFP; δ – wskazuje, czy zmienna objaśniająca X skutkuje większym efektem we wzroście TFP w krajach, które znajdują się bliżej światowej granicy technologicznej (ocena ujemna parametru) czy też dalej (ocena dodatnia). ε - zmienna losowa modelu.

Zgodnie z teorią wzrostu endogenicznego oraz teorią Schumpetera przyjmuje się, że podstawową determinantą wzrostu TFP na przełomie XX i XXI wieku są innowacje. Są one wynikiem kompleksowego łańcucha działań, a zatem nie są bezpośrednio mierzalne [Greenhalgh i in. 2010]. Mierzy się ich zmiennymi symptomatycznymi w postaci nakładów bądź efektów, ewentualnie zmiennymi syntetycznymi. Jako często wykorzystywanymi w badaniach zmiennymi symptomatycznymi innowacji są: udział wydatków na B+R w PKB bądź patenty.

Dosyć powszechnie spotykanym w literaturze poglądem badaczy - zajmujących się oddziaływaniem jakościowych czynników na wzrost gospodarczy - jest hipoteza, że informacje oraz nowe technologie komunikowania się (information and technology communication - ICT) są jednym z podstawowych stymulatorów wzrostu TFP, czyli pośredniego kanału kształtowania wzrostu gospodarczego. Zakłada się, że największy udział we wzroście produktywności pracy ma ICT. Teza te bazuje na założeniu, że ICT generują pozytywne efekty zewnętrzne, w postaci komplementarnych inwestycji, zmian organizacyjnych czy też efektów sieciowych. Jest to teza wysoce dyskusyjna, gdyż dowody na występowanie tych efektów mają charakter niejednoznaczny. W wielu modelach ekonometrycznych zakłada się również, że źródłem wzrostu TFP są efekty zewnętrzne uzyskane z tytułu wysokich kwalifikacji siły roboczej. Vandenbusseche i inni [2006] przedstawili model, w którym gospodarki z większym udziałem pracowników z edukacją uniwersytecką wykazywały wyższą stopę innowacji, którą to zależność wyjaśniali faktem, iż wysoko kwalifikowana siła robocza stwarza relatywnie większe korzyści w działalności innowacyjnej niż imitacyjnej.

Inną zmienną, na którą w kontekście stymulowania wzrostu TFP zwracane jest zainteresowanie, jest czynnik konkurencyjności. Pewna grupa badaczy, na podstawie wyników badań empirycznych, dochodzi do konkluzji, że zależność między stopą wzrostu TFP a konkurencyjnością nie jest liniowa [Timmer i in. 2010, str. 248].

Tabela 1. Innowacje, imitacje a wzrost TFP w Unii Europejskiej w latach 1995-2009. Zmienna objaśniana: stopa wzrostu TFP. Modele panelowe z efektami stałymi.

Zmienne objaśniające, parametry struktury stochastycznej	M O D E L E				
	I	II	III	IV	V
gaptech	0,536***	0,664***	0,574***	0,515***	0,432***
R&D		-0,014			
Patents			0,027	-0,308	0,239*
Gaptechpatents				0,346	
Inw					0,091
Govexp					-0,410***
N	81	81	81	81	81
R^2	0,453	0,682	0,496	0,487	0,517
SkorR^2	0,432	0,665	0,470	0,452	0,478
D-W	1,920	1,926	1,928	1,911	2,008

Źródło: Obliczenia własne na podstawie – gaptech, inw z bazy Penn World Tables 7.0; govexp, patents z Eurostatu; R&D według World Development Indicators 2010, wzrost TFP według The Conference Board Total Economy Database January 2011, http://www.conference-board.org/data/economydatabase/07.05.2011.

Współczynniki regresji istotne statystycznie odpowiednio na poziomie: * - 0,1; **-0,05; ***- 0,01. Przeciętna stopa zmian TFP – odpowiednio z lat: 1995-1999; 2000-2004;2005-2009. R&D – wskaźnik udziału wydatków ogółem na badania i rozwój w produkcie krajowym brutto, przeciętna wartość odpowiednio w latach: 1966-1999, 2000-2003, 2004-2007. Govexp – przeciętna wartość udziału wydatków rządowych w produkcie krajowym brutto w latach: 1997-1999, 2000-2003, 2004-2007. Inw – przeciętna wartość udziału wydatków inwestycyjnych ogółem w produkcie krajowym brutto w latach: 1995-1999, 2000-2004, 2005-2009. Gaptech – odwrotność realnego produktu krajowego brutto per capita w przeliczeniu do produktu USA jako kraju lidera światowej technologii, odpowiednio dla lat: 1996, 2001, 2006. Patents – przeciętna liczba patentów, przyznawana przez USPTO, w przeliczeniu na milion mieszkańców, odpowiednio za lata: 1993-1996, 1997-2000, 2001-2004. Wzrost TFP – przeciętna stopa wzrostu ogólnej produktywności czynników produkcji za lata: 1995-1999, 2000-2004, 2005-2009. Gaptechpatents – zmienna interakcyjna między luką technologiczną a patentami.

Modele ze względu na występowanie heteroskedastyczności estymowane są ważoną metodą najmniejszych kwadratów. Testem Kołmogorowa-Smirnowa sprawdzano, czy rozkład zmiennej losowej jest rozkładem normalnym. Weryfikowano jednorodność wariancji zmiennej losowej testem Goldfelda-Quandta. Liniowość modeli była sprawdzana testem serii. Dopasowanie modelu do danych empirycznych sprawdzano testem Fishera-Snedecora, zaś autokorelację testem Durbina-Watsona. Badanie istotności wpływu zmiennych objaśniających na zmienną objaśnianą przeprowadzono testem t-Studenta. Odporność współczynników regresji na zmiany specyfikacji modelu oraz zmiany próby statystycznej sprawdzono testem Levine'a. W tabelach zamieszczono wyniki estymacji modeli, których rozkład zmiennej losowej jest rozkładem normalnym, nie występuje autokorelacja reszt modelu pierwszego rzędu, model jest wystarczająco dobrze dopasowany do danych empirycznych oraz potwierdzony został wybór postaci funkcyjnej modelu testem serii, zaś współliniowość między zmiennymi objaśniającymi, mierzona czynnikiem inflacji wariancji (VIF) nie przekraczała 10. W celu zmniejszenia skutków problemu endogeniczności obserwacje wybranych zmiennych objaśniających zgromadzono dla początkowych lat panelu bądź z przesunięciem czasowym. Zastosowano pośrednią miarę luki technologicznej jako odwrotności relacji agregatowego PKB per capita w stosunku do analogicznego wskaźnika kraju z granicy technologicznej [Kubielas 2009, str.137]. Zamieszczone w tabelach oceny współczynników regresji są standaryzowane (wersja beta), a zatem są bezpośrednio porównywalne.

Wyniki estymacji modeli z tabeli 1 potwierdzają główną hipotezę badawczą postawioną we wprowadzeniu, że w latach 1995 – 2009 przeciętnie rzecz biorąc przy innych warunkach niezmienionych w krajach UE występowała konwergencja technologiczna, zaś jej podstawowym mechanizmem były imitacje. Oszacowane współczynniki regresji przy zmiennej gaptech są dodatnie i istotne statystycznie, co oznacza, że im większa jest odległość kraju od światowej granicy technologicznej, tym w wyższym stopniu za pomocą dyfuzji wiedzy zagranicznej stymulowany jest

wzrost TFP. Zmienną gaptech można nazwać zmienną konwergencji w technologii. W badanym okresie uchwycono pozytywny, ale skromniejszy rozmiarowo w porównaniu do imitacji, wpływ innowacji na wzrost TFP jedynie w przypadku V modelu (tabela 1), gdy kontrolowano rozmiary inwestycji jako symptomatyczną miarę potencjału technologicznego. Ocena współczynnika regresji przy zmiennej patents jest w modelu V dodatnia oraz istotna statystycznie. Przy braku kontroli potencjału społeczno-technologicznego nie odnotowano wpływu innowacji (zmienne B&R oraz patents, modele I-IV tabela1) na wzrost TFP w krajach UE w badanym okresie. Ponadto współczynniki regresji przy tych zmiennych są nieodporne na zmiany specyfikacji modelu bazowego.

Tabela 2. Efekty nowych technologii komunikacji, edukacji uniwersyteckiej oraz barier regulacyjnych we wzroście TFP UE27 w latach 1995-2009. Zmienna objaśniana: stopa wzrostu TFP. Modele panelowe z efektami stałymi.

Zmienne objaśniające	M O D E L E					
	I	II	III	IV	V	VI
gaptech	0,565***	0,231	0,544***	0,866***	0,713***	0,717*
internet	0,056	-0,277				
gaptechinternet		0,548***				
tertiary			-0,067	0,173		
gaptechtertiary				-0,387		
regulation					0,346***	-0,120
gaptechregulation						1,256***
N	81	81	81	81	81	81
R^2	0,453	0,485	0,480	0,433	0,471	0,505
SkorR^2	0,424	0,451	0,453	0,396	0,443	0,472
D-W	1,928	1,879	1,999	1,992	2,004	1,909

Źródło: obliczenia własne na podstawie: Internet – dane z Eurostatu, tertiary - z bazy Barro&Lee, regulation – 2010 Economic Freedom Dataset Fraser Institute.
Internet – dostęp gospodarstw domowych do Internetu, w % odpowiednio w latach: 1999, 2004, 2009. Tertiary – udział ludności z wykształceniem odpowiadającym poziomowi trzeciemu w populacji 15 lat i więcej w %, odpowiednio w latach: 1995, 2000, 2005. Regulation – zakres regulacji rynku pracy, rynku kredytowego oraz zakładania i prowadzenia działalności gospodarczej w latach: 1995, 2001, 2006. Gaptechinternet – zmienna interakcyjna między odległością od światowej granicy technologicznej a zmienną symptomatyczną nowoczesnych technologii komunikowania się. Gaptechtertiary – zmienna interakcyjna między odległością od światowej granicy technologicznej a edukacją poziomu trzeciego. Gaptechregulation – zmienna interakcyjna między odległością od światowej granicy technologicznej a stopniem regulacji działalności gospodarczej. Inne bazy oraz oznaczenia jak powyżej.

W celu uchwycenia idei, że pewne instytucje – w rozumieniu gry społeczno-ekonomicznej – w sposób istotny wpływają na proces innowacyjny bądź na szybkość transferu wiedzy w postaci imitacji do bazowego modelu doganiania wprowadzone zostały kolejno zmienne symptomatyczne ICT, edukacji na

poziomie trzecim oraz regulacji, których wybrane wyniki estymacji są prezentowane w tabeli 2. Za pomocą zmiennej symptomatycznej internet na poziomie makro nie uchwycono, w latach 1995-2009 w UE27 pozytywnych efektów zewnętrznych ICT we wzroście TFP (modele I i II tabela 2). Istotny statystycznie okazał się oszacowany współczynnik regresji przy zmiennej gaptechinternet (model 2 tabela 2). Ocena współczynnika jest dodatnia, a zatem można wnioskować, że prawdopodobnie rozprzestrzenianie informacji i rozwój nowoczesnych technologii komunikowania się stymuluje wzrost TFP przede wszystkim w krajach oddalonych od światowej przestrzeni technologicznej, czyli w wyższym stopniu sprzyja transferowi wiedzy w postaci imitacji niż procesowi innowacji. Nie odnotowano wpływu edukacji uniwersyteckiej na stopę wzrostu TFP w UE27 (zmienne: tertiary oraz gaptechtertiary w modelach III i IV tabela 2). Ten wynik nie jest zaskoczeniem, gdyż ta zmienna symptomatyczna jest pewnym elementem kapitału ludzkiego w ilościowym ujęciu, kształtowanego na poziomie trzecim. Z badań empirycznych nad oddziaływaniem kapitału ludzkiego na wzrost gospodarczy oraz wzrost TFP – prowadzonych przez różne ośrodki – stwierdza się na ogół występowanie pozytywnych efektów zewnętrznych kapitału ludzkiego na poziomie makro, jeżeli spełnionych jest szereg dodatkowych warunków. Wśród nich wyróżnia się: zapewnienie odpowiedniej jakości tego kapitału, jego wykorzystania, symetrycznego rozkładu w dostępie do niego, a także odpowiedniej struktury w zależności od odległości kraju od światowego obszaru najnowszych technologii [Soszyńska 2012, str. 13-47]. Zmienna regulacji jest faktycznie zmienną syntetyczną, za pomocą której mierzony jest stopień swobody działania na rynku pracy, kredytowym oraz rynku produktów. Jej wyższy pomiar oznacza wyższy poziom deregulacji rynków. Zgodnie z teorią ekonomii głównego nurtu wyższy poziom deregulacji sprzyja działaniu mechanizmów rynkowych poprzez zwiększenie stopnia konkurencyjności w gospodarce. Jednak pojęcie konkurencyjności jest terminem wieloaspektowym i ponadto w kontekście budowy gospodarki stymulowanej innowacjami pewna grupa badaczy udowadnia, że zależność miedzy innowacjami a konkurencyjnością nie jest liniowa. Na podstawie wyników estymacji modelu V (tabela 2 zmienna regulation) można wnioskować, że deregulacja przeciętnie rzecz ujmując – przy innych warunkach constans – stymuluje wzrost TFP. Jednak zestawienie wyników oszacowań współczynników regresji modelu V i VI (tabela 2), prowadzi do konkluzji, iż procesy deregulacyjne rynków pracy, kredytowego i produktów w latach 1995-2009 w UE27 pozytywnie oddziaływały na wzrost TFP w krajach, które były oddalone od światowej granicy technologicznej.

Podsumowując można stwierdzić, że wyniki estymacji modeli luki technologicznej przedstawione w powyższych tabelach i oszacowane na danych zagregowanych dla UE 27 są zbliżone do rezultatów, które otrzymali Timmer i inni (2010) dla UE za pomocą modeli panelowych konstruowanych na skorygowanych pomiarach TFP o przeciętna liczbę godzin pracy, a także o strukturę siły roboczej oraz strukturę gałęziową gospodarki. Nie uchwycono wpływu edukacji

uniwersyteckiej w procesie stymulowania wzrostu TFP oraz zanotowano pozytywne efekty ICT na wzrost TFP jedynie w krajach oddalonych od światowej granicy technologicznej. Zaś rola deregulacji w kształtowaniu wzrostu TFP jest istotna statystycznie przede wszystkim w krajach UE27 oddalonych od światowej granicy technologicznej.

INNOWACJE, IMITACJE, POTENCJAŁ SPOŁECZNO - TECHNOLOGICZNY A WZROST TFP

Twierdzenie, że różnice w rozwoju gospodarczym między krajami są przede wszystkim skutkiem zróżnicowania w technologii, znalazło mocne poparcie nie tylko w teorii luki technologicznej według Gerschenkrona [1962], ale również w teorii innowacji Schumpetera [1934]. Wśród badaczy istnieje konsensus, że zarówno zmniejszenie luki technologicznej, jak również luki w produktywności, a w następstwie wprowadzenie gospodarki na ścieżkę wzrostu gospodarczego stymulowanego innowacjami nie jest bezwarunkowe. Najpierw należy stworzyć zdolność, potencjał społeczno-technologiczny gospodarki na rzecz generowania innowacji oraz absorpcji i adaptacji imitacji. Pierwotnie pojęcia te zostały wprowadzone do literatury pod nazwą potencjału społecznego, gotowości technologicznej, zdolności absorpcyjnej, czy też systemu innowacji [Soszyńska 2010, str.13-47]. Są to pojęcia wieloaspektowe. Mimo iż zbiory czynników, za pomocą których definiuje się te pojęcia nie pokrywają się, to na ogół występują w nich: szeroko pojęta infrastruktura techniczna, w tym infrastruktura informatyczna; kapitał wiedzy w postaci skumulowanych nakładów i efektów działalności sektora badawczo-rozwojowego, kapitał ludzki, zwyczajowo definiowany jako edukacja i kompetencje; rozwój rynku finansowego; jakość instytucji w rozumieniu reguł gry ekonomicznej; odpowiedni poziom otwartości i stabilności gospodarki.

Ze względu na wieloaspektowość analizowanego pojęcia zgromadzono pomiary 13 zmiennych indywidualnych, na bazie których metodą analizy czynnikowej wyodrębniono trzy zmienne syntetyczne. W zbiorze zmiennych indywidualnych znalazły się: Internet, R&D, patents, inw, inwpryw – stopa inwestycji sektora prywatnego jako udział w PKB (Eurostat), govinw – stopa inwestycji rządowych jako udział w PKB (Eurostat), tertiary, edu – przeciętna liczba lat edukacji formalnej (Barro&Lee), scientists&technologists – udział pracowników nauki i specjalistów nauk ścisłych w sile roboczej (Eurtostat), FDI – udział napływu netto bezpośrednich inwestycji zagranicznych w PKB (WDI 2010), Real GDPpc – realny PKB per capita (Eurostat) według parytetu siły nabywczej, GE - miara jakości pracy administracji państwowej (Baza Governance Matters).

Po wstępnej obróbce danych za lata 1995-2009 dla krajów UE27 przeprowadzono ich analizę z punktu widzenia przydatności do przeprowadzenia analizy czynnikowej. Wyznacznik macierzy korelacji zmiennych pierwotnych

równał się 0,00, a miara KMO adekwatności doboru próby – 0,784, zaś oszacowana na bazie danych empirycznych istotność dla testu sferyczności Bartletta – 0,00. Te informacje wskazują, iż dane nadają się do przeprowadzenia analizy czynnikowej. Na podstawie bazy danych za pomocą pakietu statystycznego wygenerowano trzy czynniki, których wartość własna była większa od 1. Za pomocą tych czynników odtworzono ponad 65% wariancji pierwotnego zbioru zmiennych. Wyodrębnianie czynników przeprowadzono regresją z zastosowaniem metody największej wiarygodności bez rotacji.

Na podstawie analizy macierzy składowych odnotowano, iż największe udziały w kształtowaniu pierwszego czynnika miały odpowiednio zmienne: patenty, GE, realGDPpc, B&R, Scientists&technoligists Internet. Czynnik ten określono jako system innowacji – innovsyst. Największe udziały w tworzeniu drugiego czynnika miały zmienne: edu, tertiary, stąd przypisano temu czynnikowi nazwę kapitał ludzki – humcapital. Trzeci syntetyczny czynnik otrzymano przede wszystkim na podstawie zmiennych: FDI, govinw, inw – są to podstawowe elementy gotowości technologicznej, zatem zmienną captech – nazwano potencjałem czy też gotowością technologiczną .

Tabela 3. Innowacje, konwergencja technologiczna, potencjał społeczno-technologiczny a wzrost TFP w UE27 w latach 1995-2009. Zmienna objaśniana – stopa wzrostu TFP. Modele panelowe z efektami stałymi.

Zmienne objaśniające	M O D E L E		
	I	II	III
Gaptech	0,781***	1,028***	0,545***
Innovsyst	0,254*	-0,240	-0,103
humcapital	0,303***	0,108	0,130*
captech	0,220***	0,198**	-0,024
gaptechinnovsyst		0,689**	
innovsysthumcapital		-0,168**	-0,217***
freetrade			0,297***
N	69	69	69
R^2	0,649	0,760	0,790
SkorR^2	0,616	0,728	0,762
D-W	1,784	1,805	1,947

Źródło: Obliczenia własne. Freetrade – z 2010Economic Freedom Dataset Fraser Institute. Freetrade – miara otwartości w handlu zagranicznym, szacowana na bazie takich składowych, jak: wysokość taryf celnych, ocena barier pozataryfowych pochodząca z badan ankietowych, proporcja faktycznej wymiany handlowej w stosunku do estymowanej wymiany na podstawie modeli regresji, nadwyżka czarnorynkowego kursu walutowego nad kursem oficjalnym, stopień kontroli przepływów kapitałowych. Zmienne interakcyjne: gaptechinnovsyst między odległością od światowej granicy technologicznej a zmienną system innowacyjny, nazywana w literaturze zmienną imitacyjną; innovsysthumcapital między systemem innowacyjnym a kapitałem ludzkim. Pozostałe bazy i oznaczenia jak powyżej.

Wprowadzenie do modeli luki technologicznej obok zmiennej konwergencji w technologii (gaptech) oraz systemu innowacyjnego (innovsyst) zmiennych objaśniających reprezentujących potencjał społeczno-technologicznego prowadzi do zwiększenia stopnia wyjaśnienia zmienności stopy wzrostu TFP. Wynika to z porównania współczynników determinacji (R^2 i skorR^2) modeli z tabeli 1 oraz modeli z tabeli 3. Zaobserwowano występowanie w UE27 w latach 1995-2009 konwergencji w technologii. Ocena współczynnika regresji przy zmiennej gaptech jest dodatnia i istotna statystycznie (tabela 3). Zmienna innovsyst obejmuje wiele aspektów systemu innowacyjnego, które nie tylko są kojarzone z działalnością innowacyjną ale również imitacyjną. Celem wyodrębnienia działalności imitacyjnej wprowadzono do zbioru potencjalnych zmiennych zmienną interakcyjną gaptechinnovsyst (model II tabela 3). Na podstawie wyników estymacji modeli z tabeli 3 stwierdza się, że nie odnotowano w sposób wyraźny wpływu innowacji na stopę wzrostu TFP w UE27 w latach 1995-2009. Głównym mechanizmem kształtowania dynamiki TFP była w tym okresie – przeciętnie rzecz ujmując - w krajach UE27 imitacja technologii i była ona prawdopodobnie obok stworzenia odpowiedniego potencjału społeczno-technologicznego podstawową stymulantą konwergencji w technologii. Ocena współczynnika regresji przy zmiennej gaptechinnovsystem jest dodatnia i istotna statystycznie. Zarówno potencjał społeczny w analizowanym okresie w UE27 (zmienna humcapital) jak i potencjał technologiczny (zmienna captech) stymulowały, przy innych warunkach constans, wzrost TFP. Jednak przy danym kapitale ludzkim poziom rozwoju systemu innowacyjnego był prawdopodobnie niewystarczający, gdyż wyhamowywał wzrost TFP (zmienna innovsysthumcapital, modele II i III w tabeli 3).

PODSUMOWANIE

Badania ekonometryczne potwierdziły, iż w krajach UE27 w latach 1995-2009 występowała konwergencja w technologii, ale głównym jej mechanizmem sprawczym – przeciętnie rzecz ujmując - były imitacje a nie innowacje. Potencjał społeczny, jak i technologiczny stymulował w badanym okresie wzrost TFP. Wyniki badań interakcyjności między systemem innowacyjnym a potencjałem społeczno-technologicznym wskazują, iż występowało niedopasowanie między systemem innowacyjnym w krajach UE27 a potencjałem społeczno-technologicznym, czego przykładem jest ujemne oddziaływanie na wzrost TFP przez zmienną interakcyjną między systemem innowacyjnym a kapitałem ludzkim. Nie zanotowano wpływu na wzrost TFP takich elementów potencjału społeczno-technologicznego, jak edukacja na poziomie trzecim i jej zmieniająca się rola w procesie zbliżania się do światowej granicy technologicznej. Pozytywne efekty we wzroście TFP UE27 w latach 1995-2009 technologii ogólnego zastosowania (ICT) oraz procesów deregulacji

gospodarki uchwycono, jako istotne statystycznie, w przypadku krajów rozwijających się, które są oddalone od światowej granicy technologicznej.

BIBLIOGRAFIA

Aghion P., Howitt P. (2006) Joseph Schumpeter Lecture. Appropriate Growth Policy: A Unifying Framework, Journal of the European Economic Association, 4(2/3), pp.269-314.

Gerschenkron A. (1962) Economic Backwardness in Historical Perspective, Harvard University Press, Cambridge M. A.

Greenhalgh Ch., Rogers M. (2010) Innovation, Intellectual Property, and Economic Growth, Princeton University Press, New Jersey.

Kubielas S. (2009) Innowacje i luka technologiczna w gospodarce globalnej opartej na wiedzy. Strukturalne i makroekonomiczne uwarunkowania, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.

Schumpeter J. (1934) The Theory of Economic Development, Harvard University Press, Cambridge Mass.

Soszyńska E. (2012) Modernizacja technologiczna, potencjał społeczny a wzrost gospodarczy – wnioski dla Polski, [w:] Woźniak M. G. (red.) Gospodarka Polski 1990-2011. Modernizacja, Wydawnictwo Naukowe PWN, Tom 2, str. 13-47.

Timmer M. P., Inklaar R., O'Mahony M., Van Ark B. (2010) Economic Growth in Europe, Cambridge University Press, New York.

Vandenbussche J., Aghion P., Meghir C. (2006) Growth, Distance to the Frontier and Composition of Human Capital, Journal of Economic Growth, 11(2), pp. 97-127.

TECHNOLOGICAL CONVERGENCE AND SOCIO - TECHNOLOGICAL CAPABILITIES IN THE EUROPEAN UNION COUNTRIES

Abstract: There are presented results of several macro technology-gap models estimation to capture following ideas in the EU27 countries in years 1995-2009. First, TFP growth is determined by the strength of the domestic innovation process and the speed of imitation. Second, the socio-technological capabilities of an economy determine to what extent it innovates and exploits imitation opportunities. Third, the three possible determinants of TFP growth, namely ICT, university educated-workers and regulatory barriers have been analyses as important drivers of productivity growth.

Keywords: TFP, technological gap, innovation, absorptive capacity

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 202 – 212

O METODZIE PROGNOZOWANIA BRAKUJĄCYCH DANYCH W DZIENNYCH SZEREGACH CZASOWYCH Z LUKAMI SYSTEMATYCZNYMI

Maria Szmuksta-Zawadzka
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
Jan Zawadzki
Katedra Zastosowań Matematyki w Ekonomii
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: jan.zawadzki@zut.edu.pl

Streszczenie: W pracy przedstawiono metodę modelowania a następnie prognozowania w sytuacji, gdy w szeregu czasowym dla danych dziennych występują luki systematyczne. Podstawą budowy prognoz były regularne hierarchiczne modele szeregu czasowego opisujące wahania o rocznym. Wahania o cyklu tygodniowym były opisywane za pomocą zmiennej grupującej, w skład której wchodziły dni podobne oraz tego rodzaju zmiennych dla pozostałych dni. W modelach wystąpiły także zmienne o charakterze migawkowym oznaczające występowanie świąt oraz dni około świątecznych. Rozważania o charakterze teoretyczne zostały zilustrowane przykładem empirycznym dla założonego wariantu luk w danych. Przeprowadzona została analiza dokładności błędów prognoz inter- i ekstrapolacyjnych ogółem oraz w dezagregacji na dni tygodnia, miesiące i święta oraz dni około świąteczne.

Słowa kluczowe: szeregi czasowe, dane dzienne, prognozowanie brakujących danych, luki systematyczne

WPROWADZENIE TEORETYCZNE

W modelowaniu i prognozowaniu na podstawie kompletnych szeregów czasowych dla danych dziennych można wykorzystywać m.in. modele szeregu czasowego z wahaniami okresowymi zawierające dwa zespoły zmiennych zerojedynkowych opisujących wahania o cyklu 7-dniowym oraz 12-miesięcznym.

Zapis analityczny takiego modelu z liniowym trendem jest następujący [por.Kufel 2010; Szmuksta-Zawadzka, Zawadzki 2011]:

$$Y_t = \alpha_1 t + \alpha_0 + \sum_{k=1}^{7} d_{0k} Q_{kt} + \sum_{s=1}^{12} d_{0s} Q_{sMt} + \sum_{j=1}^{r} a_{0j} D_{jt} + U_t \,, \tag{1}$$

gdzie: $\sum_{k=1}^{7} d_{ok} = \sum_{s=1}^{12} d_{0s} = 0$.

W modelu tym uwzględniono także zmienne D_{jt} oznaczające występowanie dni świątecznych oraz około świątecznych. Na parametry d_{ok} oraz d_{0i} nałożony został warunek sumowalności do zera. Następstwem użycia do opisu wahań dwóch zespołów zmiennych zerojedynkowych będzie to, że dla każdego rodzaju wahań szacowanych będzie o jeden parametr mniej niż wynosi długość każdego z cykli. Oznacza to, że w podmacierzach zawierających zmienne Q_{kt} oraz Q_{sMt} w jednym z wierszy (zazwyczaj ostatnim w ramach cyklu) wystąpią elementy równe minus jeden. Parametry d_{ok} oraz d_{0i} interpretuje się wówczas jako odchylenia od wyrazu wolnego, będącego średnią z wyrazów wolnych dla poszczególnych dni i miesięcy.

W innym, równoważnym zapisie, nie nakłada się na parametry warunku sumowalności do zera, zastępując elementy równe minus jeden zerami. Wyraz wolny odnosić się będzie wtedy na ogół do ostatnich okresów cykli na przykład niedzieli i grudnia. Parametry dla poszczególnych dni i miesięcy będą odchyleniami od tak zdefiniowanego wyrazu wolnego.

Z literatury przedmiotu wiadomo, że w procesie modelowania i prognozowania dla danych w postaci szeregów czasowych z wahaniami sezonowymi dla okresów jednostkowych nie krótszych niż dekada użyteczne są także regularne modele hierarchiczne. W modelach tych podzielniki p_i cyklu wahań m spełniają jednocześnie dwa warunki [Szmuksta-Zawadzka, Zawadzki 2002]:

$$2 \le p_i \le 0\frac{m}{2} \tag{2}$$

$$\prod_{i=1}^{\frac{m}{2}} p_i = m \tag{3}$$

Dla danych miesięcznych wyróżnia się siedem modeli hierarchicznych (cztery dwustopniowe i trzy trójstopniowe). Natomiast dla danych o cyklu 36 dekadowym modeli tych 25, w tym 5 dwu-, 14 trój-, 6 czterostopniowych**.**

W przypadku danych dziennych, ze względu na nieparzystą liczbę dni w tygodniu, nie jest możliwa budowa regularnych modeli hierarchicznych. Dlatego, podobnie jak w modelach klasycznych, osobno należy modelować wahania o cyklu tygodniowym oraz o cyklu rocznym.

Wahania o cyklu rocznym można opisać za pomocą regularnych modeli hierarchicznych, w których parametry dla poszczególnych stopni hierarchii są średnimi z ocen parametrów dla odpowiednich miesięcy.

Sprawa opisu wahań o cyklu tygodniowym jest bardziej skomplikowana ze względu na wspomnianą wyżej, nieparzystą liczbę dni w tygodniu. Dlatego też zaproponować można procedurę polegającą na grupowania dni podobnych. Dokonując wyboru dni podobnych należy brać pod uwagę przede wszystkim kryteria oparte na przesłankach merytorycznych. Przykładem takiego kryterium może być podział na dni robocze i dni świąteczne. Z kolei dni robocze można podzielić na dni z początku tygodnia oraz jego końca itp. Wyodrębnienie dni podobnych można przeprowadzić także kierując się zbliżonymi wielkościami średnich wartości zmiennej prognozowanej lub też zbliżonymi ocenami parametrów w modelach dla pełnych danych. Ocena parametru przy zmiennej grupującej będzie równa średniej z ocen dla dni podobnych. Natomiast pozostałe dni będą uwzględnione w modelu bez zmian. Ponadto w modelach mogą występować także zerojedynkowe zmienne migawkowe, charakteryzujące występowanie dni świątecznych i około świątecznych.

Sposób budowy oraz liczba dni wchodzących w skład zmiennej grupującej odgrywa istotną rolę w przypadku modelowania i prognozowania w warunkach występowania systematycznych luk w danych. Występowanie luk tego rodzaju oznacza, że przynajmniej o jednym dniu tygodnia lub (i) przynajmniej o jednym miesiącu nie posiadamy żadnej informacji.

Do modelowania i prognozowania, w omawianej sytuacji, nie mogą być wykorzystane klasyczne modele szeregu czasowego, ze względu na zerowe wariancje zmiennych odnoszących się do tych podokresów (miesięcy i dni), w których występują luki. Natomiast do tego celu, mogą być wykorzystanie regularne modele hierarchiczne opisujące wahania o cyklu rocznym. Wahania o cyklu tygodniowym będą opisywane za pomocą zmiennej grupującej oraz zmiennych zerojedynkowych nie wchodzących w jej skład. Zmienna grupująca powinna zawierać przynajmniej jeden dzień, w którym nie występują luki. Prognozy zarówno interpolacyjne jak i ekstrapolacyjne dla okresów, w których występują z luki, będą sumami: iloczynów parametru kierunkowego trendu, wyrazu wolnego, parametrów modelu hierarchicznego dla poszczególnych stopni hierarchii oraz parametru dla dni podobnych, wyznaczonego dla dnia, o których dostępne były informacje. Dla pozostałych dni dodawany będzie, zamiast parametru przy zmiennej grupującej, parametr dla określonego dnia.

Występowania luk systematycznych może pociągnąć za sobą daleko idące następstwa związane z wyborem istotnych składowych w modelach hierarchicznych. W zależności od liczby i rozmieszczenia luk systematycznych następstwa te mogą być następujące [Zawadzki(red.) 2003]:

1. Przynajmniej część składowych należących do różnych stopni hierarchii może być ze sobą silnie skorelowana, a niektóre z nich mogą być nawet dokładnie współliniowe.
2. Część składowych może przyjmować stałe wartości, a więc charakteryzować się zerowymi wariancjami - składowe te należy wyeliminować jeszcze przed rozpoczęciem procesu modelowania.
3. Ponadto niektóre składowe mogą być liniowymi kombinacjami innych składowych. Oznaczać to będzie konieczność szacowania wielu wersji modeli różniących się zestawami składowych nie tworzących takich kombinacji. Wszystkie te wersje charakteryzować się będą identycznymi ocenami parametrów struktury stochastycznej. Tak więc będą one nierozróżnialne z punktu widzenia własności prognostycznych. Różnić się będą natomiast dokładnością prognoz *ex post*. Dlatego też dla celów prognozowania *ex ante* wybierane będą modele o minimalnych ocenach błędów prognoz *ex post.*

Egzemplifikacją przeprowadzonych wyżej rozważań o charakterze teoretycznym będzie przykład empiryczny dotyczący modelowania i prognozowania dziennej sprzedaży paliw płynnych w jednej ze stacji benzynowych.

WYNIKI BADAŃ EMPIRYCZNYCH.

Rozpatrywany będzie jeden wariant występowania systematycznych luk w danych. Luki będą występować w dwóch spośród siedmiu dni tygodnia (w środy i w piątki) oraz w dwóch spośród 12 miesięcy (w sierpniu i w październiku). Oznacza to, że dostępnych będzie 430 spośród 724 obserwacji dla całego szeregu, a więc obejmować będą około 40 procent jego długości. Luki otrzymano przez „wymazanie" 294 obserwacji z pełnego szeregu.

Szacowane będą hierarchiczne modele szeregu czasowego z liniowym trendem oraz składowymi opisującymi wahania o cyklu rocznym i tygodniowym. Do opisu wahań o cyklu 12-miesięcznym zostaną wykorzystane klasyczne regularne dwu- i trójstopniowe modele hierarchiczne. Natomiast wahania o cyklu tygodniowym opisywane będą za pomocą zmiennej grupującej obejmującej cztery spośród pięciu dni roboczych (oprócz czwartku). Oznacza to, że prognozy dla tych dni, w tym dla środy i piątku, będą średnią z ocen dla poniedziałku i wtorku. Dokonując wyboru zmiennej grupującej (GR) kierowano się zbliżonymi ocenami parametrów modelu dla tych dni dla pełnych danych [por. Szmuksta-Zawadzka, Zawadzki 2011].W modelach wystąpią także zmienne zerojedynkowe dla czwartku i soboty oraz tego rodzaju zmienne dla dni świątecznych i około świątecznych. W celach porównawczych, odnoszących się do: stopnia opisu oraz dokładności *ex post* prognoz ekstrapolacyjnych, został oszacowany model (1) a następnie na jego zostały wyznaczone prognozy.

W tabeli 1 zestawione zostały oceny parametrów struktury stochastycznej: współczynników determinacji (R^2), odchyleń standardowych składników losowych (S_e), współczynników zmienności losowej(V_{SE}) oraz statystyk empirycznych testu Durbina-Watsona (DW).

Tabela 1 Oceny parametrów struktury stochastycznej modeli hierarchicznych

Model	R^2	SE(l)	VS(%)	DW
PEL	0,527	682,8	13,50	1,94
H26	0,402	764,8	15,50	1,86
H34	0,468	720,6	14,61	2,10
H43	0,481	711,4	14,42	2,14
H62	0,470	719,9	14,60	2,11
H223	0,401	763,3	15,47	1,86
H232	0,389	771,4	15,64	1,83
H322	0,444	735,8	14,91	2,02

Źródło: opracowanie własne

Z informacji zawartych w tabeli wynika, że najwyższą ocenę współczynnika determinacji wynoszącą 52,65% oraz najniższą oceną współczynnika zmienności losowej (13,50%) charakteryzuje się model klasyczny (PEL). Spośród modeli hierarchicznych zbliżone do niego własności predyktywne posiadają trzy modele dwustopniowe; H43, H34 i H62 oraz jeden model trójstopniowy - H322. Oceny współczynników determinacji dla tych modeli mieszczą się w przedziale od 44,4 do 48,1% a współczynniki zmienności losowej od 14,42 do 14,91%.

W przypadku pozostałych trzech modeli oceny współczynników zmienności losowej kształtują się w granicach 15,47-15,64%.

Dokonując oceny kształtowania się własności predyktywnych oszacowanych równań możemy uznać je za zadowalające. Na zjawisko malejącej dokładności opisu wraz ze skracaniem długości okresu jednostkowego z jednocześnie dopuszczalnymi ocenami współczynników zmienności losowej, zwracają uwagę Hozer, Zawadzki [1990]. Statystyki empiryczne testu Durbina-Watsona zawarte w przedziale od 1,830 do 2,289 informują o bardzo słabym dodatnim lub ujemnym skorelowaniu reszt rzędu pierwszego.

Na podstawie oszacowanych modeli zostały wyznaczone dwa rodzaje prognoz: interpolacyjne i ekstrapolacyjne. Prognozy interpolacyjne dotyczą tych okresów (dni, miesięcy i świąt oraz dni przed- i poświątecznych), w których

wystąpiły luki w danych. Natomiast prognozy ekstrapolacyjne zostały wyznaczone dla kolejnego roku, bezpośrednio następującego po okresie estymacyjnym.

W tabeli 2 zestawione zostały oceny błędów prognoz interpolacyjnych. Odnosić się one będą do tych dni i miesięcy, w których wystąpiły luki w danych, a także do jedynego święta (tj. Święta Wniebowzięcia Najświętszej Marii Panny (WNMP)) przypadające w dniu 15 sierpnia oraz dnia je poprzedzającego i po nim następującego. Zostaną także przeanalizowane błędy prognoz dla zmiennej grupującej (GR).

Tabela 2 Oceny średnich względnych błędów prognoz interpolacyjnych(w %)

	Liczba prognoz	H26	H34	H43	H62	H223	H232	H322
OG	294	10,14	9,35	9,01	10,97	10,20	9,62	8,86
Q3	104	12,16	11,79	11,46	14,74	12,42	12,29	11,57
Q5	103	12,41	11,40	11,04	13,86	12,56	12,33	11,26
GR	240	12,10	11,28	10,80	13,53	12,22	11,87	10,99
Q8M	62	13,07	10,15	8,79	10,70	12,96	9,93	8,79
Q10M	62	10,33	10,95	11,25	10,44	10,06	9,90	9,74
WNMP	2	5,21	12,60	4,04	2,13	5,05	2,81	7,48
P_WNMP	4	22,17	26,25	23,56	22,68	22,20	23,01	24,14

Źródło: opracowanie własne

Z informacji zawartych w tabeli wynika, że błędy względne prognoz interpolacyjnych ogółem, otrzymanych na podstawie modeli hierarchicznych, zawierają się w przedziale od 8,86 procent dla modelu H322 do 10,97 procent dla modelu H62.

Z analizy kształtowania się ocen błędów prognoz dla środy wynika, że dla sześciu spośród siedmiu modeli hierarchicznych (z wyłączeniem modelu H62), oceny te różnią się mniej niż 1,5 punktu procentowego od oceny maksymalnej prognoz ogółem. Zdecydowanie wyższą ocenę błędu, wynoszącą 14,74%, otrzymano dla modelu H62. Oceny błędów dla piątku kształtują się na poziomie niższym o około 0,5 punktu procentowego w porównaniu ze środą. Minimalna jego ocena jest niższa o 0,42 punktu procentowego. Niższe jest także zróżnicowanie ocen – ich rozstępy przyjęły następujące wartości: 3,28 punktu dla środy i 2,82 punktu dla piątku. Minimalną ocenę łączną dla dni wchodzących w skład zmiennej grupującej (obejmującej cztery dni robocze (oprócz czwartku)), wynoszącą 10,80% otrzymano dla modelu H43. Jest ona nieznacznie niższa od minimalnych ocen dla środy i piątku (odpowiednio o 0,66 i 0,24 punktu). Oznacza to, że błędy prognoz dla poniedziałku i wtorku, nie dotyczące dni, w których wystąpiły luki, powinny być niższe od otrzymanych dla środy i piątku. Tak jest w istocie - przyjęły one wartości wynoszące one odpowiednio: 7,33 i 8,49%.

Z porównania minimalnych ocen błędów prognoz dla miesięcy wynika, że dla sierpnia jest ona bardzo nieznacznie niższa (o 0,07 punktu), a dla października mniej niż o 1 punkt procentowy wyższa od oceny minimalnej prognoz ogółem. Wprawdzie dla października otrzymano ocenę minimalna nieco wyższą, ale błędy prognoz dla tego miesiąca charakteryzują się niższym zróżnicowaniem. Ich rozstęp dla października wynosi 1,51 punktu procentowego a dla sierpnia 4,28 punktu.

Dla święta WNMP oceny błędów prognoz otrzymanych na podstawie modeli hierarchicznych są znacznie niższe od ocen otrzymanych na podstawie tych modeli dla dnia je poprzedzającego i po nim następującego. Różnica ocen minimalnych błędów prognoz otrzymanych odpowiednio dla predyktorów H62 i H26 przekracza 20 punktów procentowych.

W tabeli 3 zestawione zostały błędy prognoz ekstrapolacyjnych. Zmienne zerojedynkowe Q_{kt} oraz Q_{sMt} oznaczają w tej tabeli odpowiednio: dni tygodnia oraz miesiące w roku, dni świąteczne zastąpione zostały skrótami nazw tych świąt a nazwy dni przed- i po świątecznych są skrótami nazw świąt poprzedzonych symbolem P_:

- NROK - zmienna przyjmująca wartość jeden w dniu Nowego Roku,
- P_NROK - zmienna przyjmująca wartość jeden w dniu poprzedzającym Nowy Rok i dzień po nim następujący,
- WIELK - zmienna przyjmująca wartość jeden w pierwszym i drugim dniu Wielkanocy,
- P_WIELK - zmienna przyjmująca wartość jeden w dniu poprzedzającym Wielkanoc i dzień następujący po jej zakończeniu,
- M1_3 - zmienna przyjmująca wartość jeden w dniach 1-3 maja,
- P_M1_3 - zmienna przyjmująca wartość jeden w dniu 30 kwietnia i 4 maja,
- BC - zmienna przyjmująca wartość jeden w dniu święta Bożego Ciała,
- P_BC - zmienna przyjmująca wartość jeden w dniu poprzedzającym święto Bożego Ciała i dzień po nim następujący,
- WNMP - zmienna przyjmująca wartość jeden w dniu Święta Wniebowzięcia Najświętszej Panny Marii (15 sierpnia),
- P_WNMP - zmienna przyjmująca wartość jeden w dniach 14 i 16 sierpnia,
- WSW - zmienna przyjmująca wartość jeden w dniu święta Wszystkich Świętych (1 listopada),
- P_WSW - zmienna przyjmująca wartość jeden w dniu 31 października oraz 2 listopada,
- Li11 - zmienna przyjmująca wartość jeden w dniu Święta Niepodległości (11 listopada),
- P_Li11 - zmienna przyjmująca wartość jeden w dniach: 10 i 12 listopada,
- BN - zmienna przyjmująca wartość jeden w pierwszy i drugi dzień Świąt Bożego Narodzenia,

- P_BN - zmienna przyjmująca wartość jeden w dniach 24 i 27 grudnia.

Tabela 3. Oceny błędów prognoz ekstrapolacyjnych

	PEL	H34	H43	H26	H62	H223	H232	H322
Og.	13,46	16,01	15,19	14,30	16,28	**11,34**	11,39	13,18
Q1	16,83	15,08	14,36	14,52	15,02	**12,83**	12,98	13,63
Q2	10,95	14,80	14,17	13,59	16,38	**11,26**	12,27	13,07
Q3	11,92	12,63	12,18	10,78	14,27	**9,01**	9,07	11,25
Q4	12,46	19,96	19,86	20,01	21,63	**16,72**	17,11	18,46
Q5	13,32	12,78	11,72	10,23	10,90	**7,04**	7,30	8,50
Q6	17,85	26,53	24,27	22,02	25,46	15,23	**14,80**	19,63
Q7	12,23	14,28	**14,19**	15,26	16,91	15,39	14,99	19,10
GR	13,32	18,53	17,66	16,65	19,48	16,59	**13,34**	15,64
Q1M	14,77	17,31	**17,09**	18,88	17,33	19,09	17,94	17,85
Q2M	11,46	10,13	10,38	9,93	15,04	**9,92**	10,10	10,35
Q3M	6,58	11,46	11,22	13,23	14,18	8,57	8,95	**7,24**
Q4M	10,29	**13,53**	15,73	14,88	16,05	14,40	13,88	13,54
Q5M	18,11	25,49	22,01	15,02	25,76	**14,95**	16,06	26,56
Q6M	16,35	18,97	18,99	13,77	15,63	14,35	**12,36**	22,22
Q7M	10,19	13,66	10,87	8,73	17,90	8,82	**8,63**	12,12
Q8M	11,67	20,21	11,86	10,86	10,62	10,80	11,40	15,61
Q9M	14,57	**13,96**	19,79	15,78	17,35	15,57	16,19	14,83
Q10M	16,52	14,87	14,90	17,30	14,74	16,63	16,49	**14,64**
Q11M	17,08	16,46	**15,04**	18,01	15,32	17,88	19,21	17,17
Q12M	15,57	16,46	**14,02**	17,26	15,31	17,44	16,93	15,78
NROK	14,62							
P_NROK								
WIELK	16,77	21,87	22,32	21,82	22,65	21,78	**21,59**	22,31
P_WIELK								
M1_3	**17,35**							
P_M1_3	3,84	17,62	18,95	**16,52**	23,38	16,82	17,85	17,29
BC	7,84	**8,45**	10,26	11,34	9,12	11,79	9,69	10,30
P_BC								
WNMP	14,74	6,91	3,20	14,57	6,80	14,41	8,14	**1,69**
P_WNMP	18,48	10,92	1,41	11,52	3,55	11,34	4,77	5,56
WSW	1,29	15,45	15,52	15,63	**12,82**	15,65	15,53	14,46
P_WSW	19,74	6,31	6,70	6,67	7,17	6,67	6,63	**5,75**
Li11								
P_Li11	1,67	5,49	5,40	5,51	**5,33**	5,52	5,58	5,60

Źródło: opracowanie własne

W tabeli zamieszczone zostały oceny błędów prognoz dla tych dni świątecznych i około świątecznych, które nie przekraczają, przyjętego kryterium dopuszczalności prognoz na poziomie 25%. Dla dni i miesięcy przyjęto kryterium o 5 punktów procentowych niższe. Z informacji zawartych w tabeli wynika,

że oceny błędów prognoz ekstrapolacyjnych ogółem otrzymane dla modeli hierarchicznych mieszczą się w przedziale od 11,34 dla modelu H223 do 16,28 % dla modelu H62. W ich kształtowaniu się widoczna jest prawidłowość polegająca na tym, że oceny błędów prognoz otrzymanych na podstawie predyktorów dwustopniowych (H26, H34, H46 i H62) są wyższe od ocen błędów prognoz dla predyktorów trójstopniowych (H223, H232, H322). Maksymalna ocena błędu prognoz otrzymana dla predyktora H322, wynosząca 13,18% jest o 1,12 punktu procentowego niższa od minimalnej oceny błędu dla predyktora dwustopniowego. Oceny błędów prognoz dla predyktorów trójstopniowych są też niższe od oceny błędu prognoz otrzymanych na podstawie predyktora, którego parametry były szacowane z wykorzystaniem pełnych danych.

W przypadku prognoz ekstrapolacyjnych dla dni tygodnia otrzymanych na podstawie modeli hierarchicznych, także widoczna jest różnica w ich efektywności między modelami dwu- i trójstopniowymi. Prognozy otrzymane na podstawie predyktorów trójstopniowych charakteryzują się ocenami błędów niższymi od 1 do 3 punktów procentowych od ocen otrzymanych dla predyktorów dwustopniowych. Dla dni roboczych najbardziej efektywny, tzn. charakteryzujący się najniższą oceną względnego błędu prognoz, okazał się predyktor H223, a dla soboty był to predyktor H232. Jedynie dla niedzieli nieco wyższą dokładnością charakteryzowały się dwa predyktory dwustopniowe (H34 i H43). Z informacji zawartych w tabeli wynika, że minimalne oceny błędów prognoz otrzymanych na podstawie modeli hierarchicznych dla trzech dni tygodnia (poniedziałku, środy i piątku) oraz błąd prognoz ogółem, są niższe od odpowiadających im błędów prognoz otrzymanych na podstawie predyktora szacowanego na podstawie pełnych danych.

Z informacji dotyczących kształtowania się ocen błędów prognoz ekstrapolacyjnych dla miesięcy otrzymanych na podstawie predyktora bez luk wynika, że wahają się one w przedziale od 6,58 procent dla marca do 20,39 procent dla maja. Ocenę poniżej 10 procent otrzymano jeszcze dla kwietnia. Dokładność prognoz otrzymanych na podstawie predyktorów hierarchicznych była bardziej zróżnicowana niż w przypadku prognoz ogółem i niż dla dni tygodnia

Z porównania minimalnych ocen błędów otrzymanych na podstawie modeli hierarchicznych i modelu dla pełnych danych wynika, że dla pierwszych siedmiu miesięcy oraz września i października bardziej dokładne okazały się predyktory oparte na równaniach hierarchicznych. Różnice te mieściły się w przedziale od 0,18 punktu procentowego dla lutego do 7,95 punktu dla lipca. Dla pozostałych trzech miesięcy nieznacznie lepszy okazał się model dla pełnych danych z różnicą w dokładności od 0,11 punktu dla grudnia do 0,41 dla listopada. Jednocześnie zwraca uwagę fakt, że w odróżnieniu od dni tygodnia, dokładność dwu- i trójstopniowych predyktorów hierarchicznych dla miesięcy była zbliżona. Minimalne oceny błędów prognoz w poszczególnych miesiącach zawarte są w przedziale od 7,24% dla lutego do 17,09% procent dla stycznia, otrzymano dla sześciu z siedmiu predyktorów (poza H62) - zostały one zaznaczone „tłustym

drukiem". Dla każdego z tych modeli oceny minimalne ocen błędów otrzymano dla dwóch miesięcy. Oznacza to, że należy wykorzystywać różne modele hierarchiczne dla celów prognozowania. Oceny niższe od 10 procent otrzymano jeszcze dla lipca i lutego. Ocenami powyżej 15 procent charakteryzują się ostatnie miesiące roku a bliską tej wielkości jest także ocena dla maja.

Z analizy błędów prognoz dla dni świątecznych oraz przed- i poświątecznych wynika, że prognozy dopuszczalne otrzymano jedynie dla Święta Wniebowzięcia Najświętszej Marii Panny oraz Wszystkich Świętych.

W pozostałych przypadkach dopuszczalne były prognozy bądź tylko dla samych Świąt (Nowego Roku, Wielkanocy, Bożego Ciała) bądź dni poprzedzających i następujących po świętach: 1_3 Maja i Niepodległości oraz dni po nich następujących. Niemożność otrzymania prognoz dopuszczalnych dla pozostałych świąt i dni około świątecznych wynikała z dużej skali wahań o charakterze losowym zakłócającym wyrazistość prawidłowości.

PODSUMOWANIE

Z analizy błędów prognoz inter- i ekstrapolacyjnych ogółem oraz w dezagregacji na dni tygodnia, miesiące oraz część dni świątecznych oraz około świątecznych wynika praktyczna użyteczność predyktorów hierarchicznych w prognozowaniu brakujących danych, w którym luki systematyczne obejmowały około 40 procent długości szeregu. W pełni zasadna okazała się także procedura wykorzystania zmiennej grupującej obejmującej cztery dni robocze (bez czwartku), z których w dwóch wystąpiły braki w danych.

BIBLIOGRAFIA

Hozer J.,Zawadzki J. (1990) Zmienna czasowa i jej rola w badaniach ekonometrycznych, PWN,Warszawa.

Kufel T. (2010) Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwośći prognozowania,Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń.

Szmuksta – Zawadzka M., Zawadzki J. (2002) Hierarchiczne modele szeregów czasowych z wahaniami sezonowymi. Budowa. Estymacja. Prognozowanie., [w:] Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych, Wyd. AE, Kraków, str. 193-204.

Szmuksta-Zawadzka, M., Zawadzki, J. (2011) The ex post analysis of the forecast accuracy of liquid fuel daily sales, Research Papers of Wrocław University of Economics, Econometrics nr 32, s. 194-204.

Zawadzki J. (red) (2003) Zastosowanie hierarchicznych modeli szeregów czasowych w prognozowaniu zmiennych ekonomicznych z wahaniami sezonowymi, Akademia Rolnicza, Szczecin.

ABOUT METHOD OF FORECASTING THE MISSING DATA IN DAILY TIME SERIES WITH SYSTEMATIC GAPS

Abstract: This paper presents a method for modeling and then forecasting in situation, when in time series for daily data contain systematic gaps. Base of construction were regular hierarchical time series models describing annual fluctuations. Weekly fluctuations were described as a grouping variable, which contains similar days and this type variables for other days. In models were used also dummy variables describing holidays and days pre- and post- holidays. Theoretical considerations were illustrated by empirical example for selected variant of gaps. Based on the same estimated equations, inter- and extrapolation predictions ware built. For both types of prediction – in general and in disaggregation to weekdays and months and holidays and days pre- and post holidays.

Keywords: data gaps, daily data, forecasting

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 213 – 221

DYWERGENCJE BOSEGO-EINSTEINA W ANALIZIE PODOBIEŃSTW FINANSOWYCH SZEREGÓW CZASOWYCH[1]

Ryszard Szupiluk
Katedra Informatyki Gospodarczej
Szkoła Główna Handlowa w Warszawie
e-mail: rszupi@sgh.waw.pl

Streszczenie: Ocena wzajemnego podobieństwa finansowych szeregów czasowych jest jednym z problemów, w którym kwestia właściwego doboru metody analitycznej zaznacza się bardzo wyraźnie. Z reguły problem ten sprowadzany jest do analizy korelacji - co nie zawsze prowadzi do właściwych rezultatów. Często są to oceny wręcz sprzeczne ze wizualną obserwacją lub wiedzą ekspercką. Powodów takiego stanu rzeczy można upatrywać zarówno we właściwościach samej miary korelacyjnej i jej adekwatności do analizowanych danych, jak również w aspekcie metodologicznym przeprowadzanego badania. W niniejszym artykule zaproponujemy alternatywne rozwiązanie oparte na miarach dywergencji, w szczególności dywergencji Bosego-Einsteina. Przeprowadzone eksperymenty na poglądowych danych symulowanych potwierdzają użyteczność zaproponowanych rozwiązań

Słowa kluczowe: podobieństwo szeregów czasowych, miary dywergencji, dywergencja Bosego-Einsteina

WSTĘP

Podobieństwo między zmiennymi, wektorami czy funkcjami może być różnie definiowane. W przypadku analizy i modelowania danych empirycznych, w rzeczywistych problemach ekonomicznych, często oczekuje się by matematyczne ilościowe oceny podobieństwa odpowiadały potocznej intuicji z tym pojęciem wiązanej. W takim rozumieniu problematyka oceny podobieństwa

[1] Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/03/B/HS4/05092

instrumentów finansowych, reprezentowanych przez odpowiadające im szeregi czasowe, zasadniczo ujmowana jest w dwojaki sposób.

Najpopularniejszym, najszerzej zbadanym i opisanym podejściem są metody korelacyjne. Podobieństwo jest tu interpretowane w kategoriach statystycznych zależności drugiego rzędu. Podejście to ma długą tradycję, doprowadziło także do rozwiązania wielu fundamentalnych zagadnień, lub stanowiło podstawę tychże rozwiązań. Wiąże się to, z ogólnym znaczeniem statystyk drugiego rządu i rozkładu normalnego w analizie danych. Rozkład normalny, którego znaczenie statystyczne ugruntowane jest w centralnym twierdzeniu statystycznym, jest w pełni zdefiniowany przez statystyki drugiego rzędu. Stosunkowo prosta estymacja statystyk drugiego rzędu, w przypadku rozkładu gaussowskiego, pozwalana na uzyskanie pełnej informacji statystycznej o badanym zjawisku. Dodatkowo, na bazie rozkładu gaussowskiego, oraz dominującego paradygmatu estymacji, w postaci metody największej wiarygodności, można otrzymać „wygodne", liniowe modele statystyczne [10].

Istnieje jednak szereg zagadnień, gdzie wspomniany aparat korelacyjny nie jest tak adekwatny. Takie uwarunkowania jak: zależności nieliniowe, dane o klastrowej strukturze, niestacjonarność zmiennych, w przypadku mechanicznego zastosowania aparatu korelacyjnego, mogą prowadzić do niewłaściwych wniosków [2,9]. Transformacje zmiennych z kolei, mogą doprowadzić do zniekształcenia pierwotnych związków między zmiennymi. W efekcie, badanie zależności korelacyjnych, wymaga w dużej mierze indywidualnego nadzoru, co utrudnia wykorzystanie w automatycznych systemach rozpoznawania wzorców [7,8]. Także na gruncie samych metod statystycznych znajdziemy zagadnienia, kiedy wariancja jest z założenia niejednoznaczna. Jednym z takich przykładów są zastosowania analizy składowych niezależnych. Jest to wielowymiarowa metoda z założenia adresowana do zmiennych niegaussowskich (poza jedną dopuszczalną w zbiorze) oraz eksplorująca statystyki wyższych rzędów. Przy czym uzyskiwane komponenty w tej metodzie cechują się, niejednoznacznością względem wariancji. Oznacza to, że może być ona dowolnie skalowalna, co powoduje że wiele algorytmów realizujących ICA z założenia standaryzuje wariancje do jedności [3]. Przy czym, „wizualna" charakterystyka zmienności tych komponentów może być wysoce zróżnicowana.

Alternatywne podejście w ocenie podobieństwa, polega na postawieniu zagadnienia jako problemu segmentacji (grupowania). Prowadzi to szerokiego spektrum różnych technik w których podobieństwo oceniane jest z reguły jako odległość euklidesowa między badanymi obiektami, lub w ogólnym przypadku jako odległość mierzona określoną *p*-normą.

$$D_p\|\mathbf{x}-\mathbf{y}\|_p = \left(\sum_{i=1}^{L}|x_i - y_i|^p\right)^{1/p} \tag{1}$$

Jest to także popularne podejście, którego jednak głównym mankamentem jest, ponownie niejednoznaczność otrzymywanych wyników. W przypadku dwóch zmiennych (szeregów czasowych, sygnałów) z samej informacji o odległości między nimi, trudno jest ocenić stopień ich podobieństwa. Pewnym rozwiązaniem może być tu przyjęcie określonej zmiennej referencyjnej, względem której obliczane są odległości, jednak całokształt tak uzyskanej informacji o podobieństwie jest dalece niejednoznaczny i względny.

Powyższe ograniczenia, były i są motywacją do poszukiwania nowych miar podobieństwa. W ostatnim czasie szczególnie intensywnie badane są miary dywergencji. Rozwój metod opartych na dywergencjach, wiąże się przede wszystkim ze spektakularnymi sukcesami ich zastosowań, w zagadnieniu nieujemnej faktoryzacji macierzy, dzięki której, można z kolei dokonać niezwykłych operacji na obrazach i wzorcach, np. wyodrębnić pojedynczy jego fragment. Otworzyło to nowe możliwości budowy systemów rozpoznawania wzorców, m.in. dla komunikacji ze sparaliżowanymi, za pomocą analizy obrazów fal mózgowych. Jedną z cech charakterystycznych dywergencji, jest ich ogólny brak symetrii, co może zostać wykorzystane do oceny podobieństwa sygnałów (szeregów czasowych). W niniejszym artykule, zostanie zaproponowana metod oceny podobieństwa szeregów czasowych (sygnałów), oparta na dywergencji Bosego-Einsteina. Prezentacja tej metody zostanie dokona w kontekście filtracji szeregów czasowych przy wykorzystaniu analizy składowych niezależnych. Pozwoli to na wykazanie naturalnych ograniczeń związanych z analizą korelacyjną. Przeprowadzone eksperymenty praktyczne, zarówno na poglądowych danych symulowanych, jak i na rzeczywistych instrumentach finansowych potwierdzają użyteczność zaproponowanych rozwiązań.

DYWERGENCJE BOSEGO-EINSTEINA I PODOBIEŃSTWO

Dywergencją $D(y\,||\,z)$, nazywana jest funkcja dwuargumentowa określona na nieujemnych zmiennych *z* i *y,* która spełnia warunek $D(y\,||\,z)\geq 0$, gdzie $D(y\,||\,z)=0$ wtedy i tylko wtedy, gdy $y=z$ [1,5,6]. Dywergencja nie musi natomiast spełniać nierówności trójkąta $D(y\,||\,z)\leq D(y\,||\,x)+D(x\,||\,z)$ oraz nie musi być spełniony warunek symetryczności $D(y\,||\,x)=D(x\,||\,y)$. Dla części dywergencji konieczny jest warunek sumowania się wartości zmiennych z i y do jedności. Dywergencje mogą być zdefiniowane dla wielkości ciągłych jak i dyskretnych. Obecnie funkcje dywergencji stosowane są do oceny podobieństwa (lub jego braku) między nieujemnymi zmiennymi, wektorami, macierzami lub funkcjami. Do najpopularniejszych i najszerszych klas dywergencji zalicza się dywergencję Bregmana oraz dywergencję Csiszar'a [5].

Jedną z popularnych dywergencji jest dywergencji Bosego-Einsteina, która dla wektorów $\mathbf{z}=[z_1,z_2,...,z_L]$ oraz $\mathbf{y}=[y_1,y_2,...,y_L]$, gdzie $y_i,z_i\in[0,1]$ zdefiniowana jest jako

$$D_{BE}^{\alpha}(\mathbf{y}\,\|\,\mathbf{z})=\sum_{i=1}^{L}\left(y_i\ln\frac{(1+\alpha)y_i}{y_i+\alpha z_i}+\alpha z_i\ln\frac{(1+\alpha)z_i}{y_i+\alpha z_i}\right). \tag{2}$$

Dywergencja ta posiada szereg interesujących właściwości, m.in. $D_{BE}^{\alpha}(\mathbf{y}\,\|\,\mathbf{z})=D_{BE}^{1/\alpha}\left(\mathbf{y}\|\mathbf{z}\right)$ oraz $D_{BE}^{\alpha\to\infty}(\mathbf{y}\,\|\,\mathbf{z})=D_{KL}\left(\mathbf{y}\|\mathbf{z}\right)$, gdzie D_{KL} oznacza dywergencję Kullback'a-Leiblera.

Istotą cechą dywergencji (2) jest brak symetrii, co może zostać wykorzystane przy ocenie podobieństwa sygnałów. Dla sygnałów, w pewnym ogólnym sensie „statystycznie podobnych" można oczekiwać, że kolejność argumentów nie będzie miała znaczenia. W szczególności, dla przypadku sygnałów losowych nie posiadających z założenia wzorców czy regularności, można założyć, że kolejność argumentów, w niesymetrycznej dywergencji (2) nie ogrywa roli. Oznacza to, że dla sygnałów szumowych v_1,v_2 o tym samym rozkładzie miara dywergencji (2), powinna być symetryczna $D_{BE}(v_1\,\|\,v_2)=D_{BE}(v_2\,\|\,v_1)$. Efekt symetrii, standaryzowanych do przedziału [0,1] sygnałów, może być mierzony jako

$$q=\mathrm{abs}\left(\log\frac{D_{BE}(z\,\|\,v)}{D_{BE}(v\,\|\,z)}\right). \tag{3}$$

Badając symetrię, z wykorzystaniem dywergencji Bosego-Einsteina, należy mieć na uwadze wpływ parametru α na jej wartości. Obecnie zostanie to szerzej przedstawione.

Zwróćmy uwagę, że wyrażenie pod znakiem sumy w (2) można rozpisać jako

$$\begin{aligned}f(y,z)&=y\ln\frac{(1+\alpha)y}{y+\alpha z}+\alpha z\ln\frac{(1+\alpha)z}{y+\alpha z}=\\&=y\ln\big((1+\alpha)y\big)-y\ln\big(y+\alpha z\big)+\alpha z\ln\big((1+\alpha)z\big)-\alpha z\ln\big(y+\alpha z\big)\end{aligned}. \tag{4}$$

Dla $\alpha \in (0,1)$, wszystkie wielkości logarytmowane przyjmują wartości z zakresu $(0,2)$. Dla takiego przypadku, możliwe jest następujące rozwinięcie $\ln(x)$ w szereg Taylora

$$\ln x = (x-1) - \frac{(x-1)^2}{2} + \ldots + (-1)^{n+1} \frac{(x-1)^n}{n} + \ldots \approx x - 1 . \tag{5}$$

W efekcie

$$\begin{aligned} & f(y,z) \approx y\big((1+\alpha)y - 1\big) - y\big(y + \alpha z - 1\big) + \\ & + \alpha z\big((1+\alpha)z - 1\big) - \alpha z\big(y + \alpha z - 1\big) = \alpha (y-z)^2 . \end{aligned} \tag{6}$$

Uwzględniając (2) i (6) dla $\alpha \in (0,1)$ zachodzi

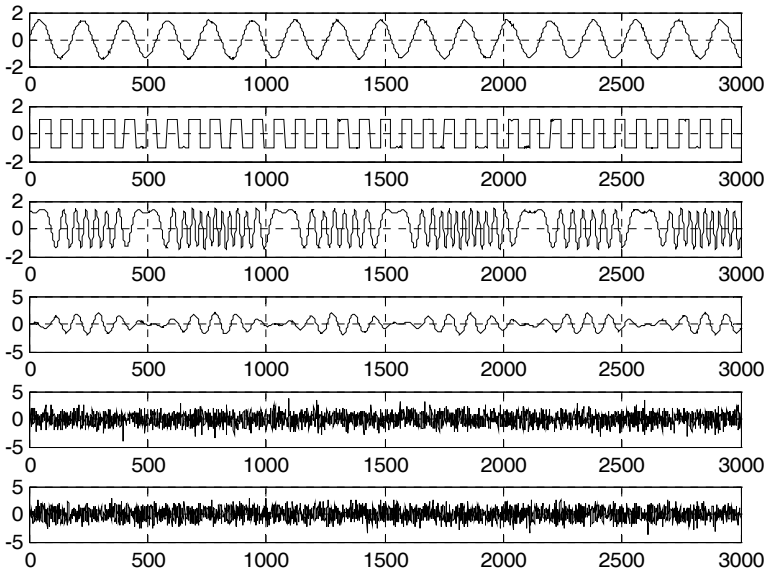
$$D_{BE}^{\alpha}(\mathbf{y} \,||\, \mathbf{z}) = \alpha \left\| \mathbf{y} - \mathbf{x} \right\|_2^2 + R , \tag{7}$$

gdzie *R* oznacza resztę aproksymacji. Przy czym, z (7) wynika, że to *R* odpowiada za efekt asymetrii $D_{BE}^{\alpha}(\mathbf{y} \,||\, \mathbf{z})$. Oznacza to, że od doboru parametru α zależy czułość metody w zakresie oceny podobieństwa. Jednak jak zauważmy w części eksperymentalnej, nawet przyjęcie symetryzującego $\alpha \in (0,1)$, pozwala na osiągnięcie satysfakcjonujących wyników.

EKSPERYMENT PRAKTYCZNY

Powyższe rozważania zostaną obecnie zaprezentowane w kontekście symulowanych badań komputerowych. Postawiony wyżej problem, poszukiwania dobrej, „intuicyjnej", ale jednocześnie ilościowej miary podobieństwa, obrazuje rysunek 1. Widoczne na nim sygnały są wzajemnie niezależne (a więc i zdekorelowane), mają także jednostkowe wariancje.

Rysunek 1. Podobieństwo a korelacja



Źródło: opracowanie własne

Odpowiadająca tym sygnałom macierz współczynników korelacji **C** ma postać

$$\mathbf{C} = \begin{bmatrix} 1.0000 & 0.0594 & -0.0526 & -0.0030 & -0.0502 & -0.0033 \\ 0.0594 & 1.0000 & -0.0098 & 0.0098 & 0.0420 & 0.0142 \\ -0.0526 & -0.0098 & 1.0000 & 0.0226 & -0.0460 & -0.0087 \\ -0.0030 & 0.0098 & 0.0226 & 1.0000 & -0.0108 & -0.0204 \\ -0.0502 & 0.0420 & -0.0460 & -0.0108 & 1.0000 & -0.0084 \\ -0.0033 & 0.0142 & -0.0087 & -0.0204 & -0.0084 & 1.0000 \end{bmatrix}$$

Widać, że korelacyjna ocena podobieństwa, nie odpowiada wizualnej ocenie, w szczególności w odniesieniu do dwóch ostatnich szumowych sygnałów. Z kolei, zastosowanie odległości opartej na p-normie przy p=2, prowadzi do wyników prezentowanych w Tabeli 1.

Tabela 1. Odległości między sygnałami mierzone p-normą, przy p=2

	S1	S2	S3	S4	S5	S6
S1	0	0,0333	0,0409	0,0316	0,0287	0,0287
S2	0,0333	0	0,0379	0,0267	0,0226	0,0239
S3	0,0409	0,0379	0	0,0351	0,0331	0,0335
S4	0,0316	0,0267	0,0351	0	0,0184	0,0197
S5	0,0287	0,0226	0,0331	0,0184	0	0,0153
S6	0,0287	0,0239	0,0335	0,0197	0,0153	0

Źródło: opracowanie własne

Jak widać, choć podobieństwo między sygnałami jest stosunkowo właściwie oceniane, to jednocześnie symetryczność tej oceny, istotnie ogranicza dalszą interpretację. Nie sposób ocenić także, jaki charakter mają analizowane sygnały. Pomiar odległości za pomocą dywergencji Bosego-Einsteina nawet przy symetryzującym parametrze $\alpha = 0{,}5$ pozwala na wyraźne odróżnienie sygnałów podobnych, co prezentuje Tabela 2.

Tabela 2. Odległości między sygnałami mierzone symetryzowaną dywergencją Bosego-Einsteina przy $\alpha = 0{,}5$

	S1	S2	S3	S4	S5	S6
S1	0	0,2354	0,4102	0,2038	0,1661	0,1674
S2	0,2415	0	0,3855	0,1525	0,1113	0,1241
S3	0,375	0,3406	0	0,2978	0,2756	0,2809
S4	0,2197	0,1641	0,3541	0	0,0663	0,0764
S5	0,1876	0,1264	0,3398	0,071	0	0,0435
S6	0,1881	0,1392	0,3437	0,0812	0,0428	0

Źródło: opracowanie własne

Dodatkowo stopień symetryczności oceny podobieństwa za pomocą dywergencji Bosego-Einsteina, pozwala na wnioskowanie o statystycznym podobieństwie typowym dla sygnałów losowych. Wyniki oceny symetrii, mierzonej parametrem (3) prezentuje Tabela 3.

Tabela 3. Stopień symetryczności dla odległości między sygnałami mierzony symetryzowaną dywergencją Bosego-Einsteina przy $\alpha = 0{,}5$

	S1	S2	S3	S4	S5	S6
S1	0	0,0255	0,0898	0,0751	0,1216	0,1165
S2	0	0	0,1237	0,0738	0,1272	0,1147
S3	0	0	0	0,1731	0,2095	0,202
S4	0	0	0	0	0,0684	0,0606
S5	0	0	0	0	0	0,0158
S6	0	0	0	0	0	0

Źródło: opracowanie własne

Otrzymane wyniki potwierdzają zasadność wykorzystania opisanej metody oceny podobieństwa opartej na dywergencjach Bosego-Einsteina. Mogą mięć one szerokie zastosowanie w automatycznych systemach transakcyjnych, w których jednym z podstawowych problemów jest poszukiwanie podobieństw między historycznymi wzorcami. Ma to szczególne znaczenie, gdy do systemów automatycznych chcemy przekazać ludzką wiedzę i doświadczenie. Jak widać z przedstawionego przykładu brak korelacji między danymi nie musi oznaczać braku podobieństw w potocznym tego słowa rozumieniu.

ZAKOŃCZENIE

W artykule przedstawiono koncepcję oceny podobieństwa wykorzystującą niesymetryczne właściwości miar dywergencji. Zastosowana dywergencja Bosego-Einsteina, jako typowy przykład niesymetrycznej dywergencji, może być zastąpiona innym typem dywergencji. Dywergencja Bosego-Einsteina ma interesujące właściwości zmiany stopnia symetryczności w zależności od parametru aplha. Jego zmiany pozwalają na różnicowanie stopnia asymetrii miary, co pozwala dostosować czułość metody do posiadanych danych. Należy także zauważyć, że fakt zakładanej nieujemności sygnałów, nie jest tu zasadniczym ograniczeniem, gdyż oceniany i porównywany jest kształt rzeczywistych sekwencji danych. Dlatego dla danego sygnału jak również szumu referencyjnego, zawsze możliwym jest dokonanie przesunięcia lub standaryzacji do wartości nieujemnych.

BIBLIOGRAFIA

Amari, S. (1985) Diferential-Geometrical Methods in Statistics. Springer Verlag.

Anscombe, F. J. (1973) Graphs in statistical analysis. The American Statistician 27: 17–21.

Cardoso J.-F. and Comon P. (1996): Independent component analysis, a survey of some algebraic methods. In Proc. ISCAS Conference, volume 2, pages 93–96, Atlanta.

Cichocki, A., Zdunek, R., Amari, S.: (2006) Csiszar's Divergences for Non-Negative Matrix Factorization: Family of New Algorithms, Lecture Notes in Computer Science, Vol. 3889, pp. 32--39, Springer Verlag, Heidelberg.

Cichocki, A., Zdunek, R., Phan, A.-H., Amari, S. (2009) Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis. John Wiley.

Csiszar, I. (1974) Information measures: A critical survey. In: Prague Conference on Information Theory. Volume A, pp. 73--86. Academia Prague.

Krutsinger J. (1997) Trading Systems: Secrets of the Masters, McGraw-Hill.

Luo, Y., Davis, D., Liu, K.: (2002) A Multi-Agent Decision Support System for Stock Trading. In: The IEEE Network Magazine Special Issue on Enterprise Networking and Services, vol.16, No. 1.

Rodgers J. L. and Nicewander W. A. (1988) Thirteen ways to look at the correlation coefficient. The American Statistician, 42(1):59–66.

Therrien, C.W. (1992) Discrete Random Signals and Statistical Signal Processing. Prentice Hall, New Jersey.

BOSE-EINSTEIN DIVERGENCES FOR SIMILARITY ANALYSIS IN FINANCIAL TIME SERIES

Abstract: The similarity assessment of the financial time series is the one of problems where the proper methodological choice is very important. The typical correlation approach can lead to misleading results. Often the similarity score is contrary to the visual observations, expert's knowledge and even a common sense. The reasons of such situations can be associated with the properties of the correlation measure and its adequateness for analyzed data, as well as in terms of methodology aspects. In this article, we point these disadvantages associated with the use of correlation to assess the similarity of financial time series as well as we propose the alternative solution based on divergence measures. In particular, we focus on the Bose-Einstein divergence. The practical experiments with simulated data confirm the validity of our concept.

Keywords: time series similarity, divergence measures, Bose-Einstein divergence

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 222 – 231

STABILNOŚĆ PARAMETRÓW MODELU RYNKOWEGO SZACOWANEGO W OPARCIU O STOPY ZWROTU WIG

Marek Szymański
Katedra Organizacji i Zarządzania
Wyższa Szkoła Administracji i Biznesu im. Eugeniusza Kwiatkowskiego w Gdyni
e-mail: m.szymanski@wsaib.pl

Streszczenie: Celem artykułu jest przedstawienie problemów, z jakimi można się spotkać stosując metodę skumulowanej nadzwyczajnej stopy zwrotu, gdy pomiędzy oknem estymacji i oknem obserwacji dojdzie do zmiany koniunktury na giełdzie. Wyniki badań sugerują, że zachodzi silny związek między średnią stopą zwrotu z WIG-u w oknie estymacji i parametrami modelu rynkowego. W wyniku zmiany średniej stopy zwrotu z indeksu między oknem estymacji i oknem obserwacji oszacowany model rynkowy może okazać się nieprzydatny dla określenia wpływu badanego zdarzenia na notowania określonych akcji.

Słowa kluczowe: skumulowana nadzwyczajna stopa zwrotu, Giełda Papierów Wartościowych w Warszawie, WIG

WSTĘP

Skumulowana nadzwyczajna stopa zwrotu (CAR – *cumulative abnormal return*) jest metodą powszechnie stosowaną do zbadania wpływu na notowania akcji określonej spółki zdarzeń takich jak: ogłoszenie planu rozwoju przez fuzje i przejęcia [Schipper, Thompson 1983], zapowiedź, akceptacja warunków i przeprowadzenie fuzji (dla wszystkich zaangażowanych spółek) [Asquith 1983, Asquith, i in. 1983, Malatesta 1983, Huang, Walkling 1987, Hamrol, Tarczyńska 2002, Magenheim, Mueller 1988], ogłoszenie wezwania do sprzedaży akcji [Bradley i in. 1983], ogłoszenie wyników finansowych spółek [Szyszka 2001a], prognozy finansowe [Chen i in. 1997, Nogalski, Szymański 2003] lub ich korekta [Szyszka 2001b], decyzje legislacyjne, administracyjne lub sądowe dotyczące fuzji [Stillman 1983, Eckbo 1983, Wier 1983, Kamma i in. 1988, Bradley, Schipani 1983], planowanie lub wprowadzenie zmian w statutach lub rozporządzeniach

spółek [Jarrell, Paulsen 1987, DeAngelo, Rice 1983, Linn, McConnell 1983, Malatesta, Walkling 1988; Ryngaert 1988], *proxy contests* [Dodd, Warner 1983], zmiany w strukturze własności [Jarrell, Paulsen 1988, Holderness, Sheehan 1988, Dann, DeAngelo 1983, Bradley, Wakeman 1983, Klein, Rosenfeld 1988, Thosar 1996], zmiany na kluczowych stanowiskach w firmie [Warner i in. 1988], stosowanie rozmaitych technik w sytuacji próby wrogiego przejęcia [Loh i in. 1995, Thosar 1996, Loh, Rathinasamy 1997, Davidson i in. 1998], zapowiedzi dotyczące wysokości dywidendy [Impson 1997], wydanie rekomendacji giełdowych [Buzała 2012] i innych.

Podstawowym założeniem metody jest istnienie zależności między stopą zwrotu z akcji określonej spółki i stopą zwrotu z określonego portfela akcji (najczęściej indeksu giełdowego). Odkrycie tej zależności pozwala sprawdzić czy po zaistnieniu badanego zdarzenia zależność jest wciąż obowiązująca, czy też zdarzenie przyniosło dodatnie lub ujemne dodatkowe stopy zwrotu. W oparciu o dane uzyskane dla różnych spółek charakteryzujących się podobnymi zdarzeniami możliwe jest określenie średniej skumulowanej nadzwyczajnej stopy zwrotu (ACAR), która może służyć jako podstawa do wyciągnięcia wniosków dotyczących rodzaju zdarzenia, nie zaś konkretnego przypadku.

Choć idea CAR wydaje się być dość prosta, a założenia dopracowane, wyniki jakie uzyskuje się badając wpływ określonego zdarzenia na notowania akcji konkretnej spółki mogą być wrażliwe na zmiany koniunktury giełdowej pomiędzy tzw. oknem estymacji i oknem obserwacji.

PROCEDURA CAR

Stosowanie CAR obejmuje cztery główne etapy:

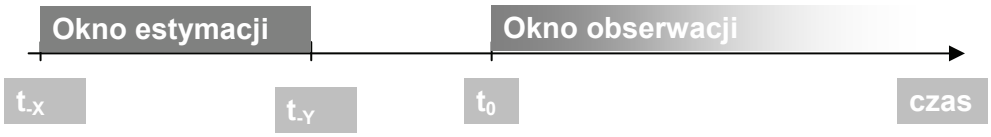
1. Wybór kluczowych dat.
2. Oszacowanie parametrów modelu rynkowego.
3. Obliczenie nadzwyczajnych stóp zwrotu.
4. Skumulowanie nadzwyczajnych stóp zwrotu.

Stosowanie metody CAR wymaga określenia trzech kluczowych dat: t_0, t_{-X} i t_{-Y} (rys.1). Pierwsza z wymienionych dat jest dniem zaistnienia zdarzenia, którego wpływ na notowania spółki jest badany. Dzień t_0 rozpoczyna tzw. okno obserwacji, czyli okres, z którego dane powinny pokazywać wpływ badanego zdarzenia na notowania spółki. Nierzadko początek okna obserwacji przesuwany jest przed dzień t_0, dzięki czemu można określić na przykład czy doszło do wykorzystania informacji poufnych.

Metodyka CAR nie wyznacza żadnych stałych dat określających granice okna estymacji. W praktyce stosuje się dowolne dwie daty sprzed dnia t_0.: -12 i -7 [Asquith 1983], -130 i -31 [Asquith i in. 1983], -308 i -109 [Ruback 1983], -210 i -61 [Kamiński, Lewandowski 2000], -60 i -16 [Buzała 2012] lub nawet

dwie daty „otaczające" dzień t_0 (np. Shipper i Thompson [1983] zastosowali okno estymacji obejmujące od -24 i +23 miesiące wokół dnia zdarzenia).

Rysunek 1. Kluczowe daty w metodzie CAR



Źródło: opracowanie własne

Na podstawie rzeczywistych obserwacji pochodzących z okna estymacji szacuje się parametry modelu rynkowego, który prezentuje związek między stopami zwrotu z akcji badanej spółki i określonego indeksu:

$$R_{it} = \alpha_i + \beta_i \cdot R_{mt} + e_{it} \tag{1}$$

gdzie:

R_{it} – stopa zwrotu dla waloru *i* w dniu *t*,
α_i, β_i – parametry równania regresji oszacowane na podstawie danych empirycznych,
R_{mt} – stopa zwrotu z indeksu rynkowego *m* w dniu *t*,
e_{it} – składnik resztowy.

Okno estymacji w założeniu ma obejmować dane obrazujące „normalne" zmiany cen akcji badanej spółki. W idealnych warunkach nadzwyczajne stopy zwrotu dostrzegalne w oknie obserwacji mają wskazywać dodatni lub ujemny efekt zdarzenia w porównaniu z warunkami „normalnymi"

Rzeczywista stopa zwrotu obliczana jest według wzoru:

$$R_{it} = \frac{P_{it} + D_{it}}{P_{it-1}} \tag{2}$$

gdzie:

R_{it} – stopa zwrotu dla waloru *i* w dniu *t*,
P_{it} – cena waloru i w dniu *t*,
P_{it-1} – cena waloru i w dniu *t-1*,
D_{it} – dywidenda z akcji *i*, do której prawo przyznano w dniu *t*.

Stopę zwrotu z indeksu rynkowego oblicza się w zbliżony sposób:

$$R_{mt} = \frac{P_{mt}}{P_{mt-1}} \tag{3}$$

gdzie:

R_{mt} – stopa zwrotu dla indeksu *m* w dniu *t*,
P_{it} – poziom indeksu *m* w dniu *t*,
P_{it-1} – poziom indeksu *m* w dniu *t-1*,

Różnica pomiędzy rzeczywistą stopą zwrotu z akcji *i*, a stopą zwrotu szacowaną na podstawie modelu regresji stanowi nadzwyczajną stopę zwrotu (*abnormal return* – AR):

$$AR_{it} = R_{it} - (\alpha_i + \beta_i \cdot R_{mt}) \tag{4}$$

gdzie:

Ar_{it} – nadzwyczajna stopa zwrotu z akcji *i* w dniu *t*,

R_{it} – stopa zwrotu dla waloru *i* w dniu *t,*

α_i, β_i – parametry równania regresji oszacowania na podstawie danych empirycznych z okna estymacji,

R_{mt} – stopa zwrotu z indeksu *m* w dniu *t*.

Suma nadzwyczajnych stóp zwrotu w kolejnych dniach stanowi skumulowaną nadzwyczajną stopę zwrotu:

$$CAR_{i,k-n} = \sum_{t=k}^{t=n} AR_{it} \tag{5}$$

gdzie:

$CAR_{i,k-n}$ – skumulowana nadzwyczajna stopa zwrotu z akcji *i* w okresie od dnia *k* do dnia *n*,

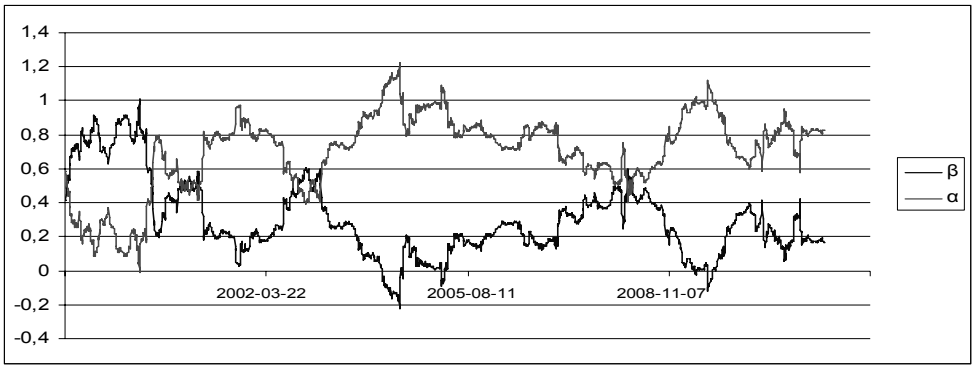
AR_{it} – nadzwyczajna stopa zwrotu z akcji *i* w dniu *t*.

Do oszacowania poziomu istotności parametrów modelu wykorzystuje się statystykę t-Studenta.

NIESTABILNOŚĆ PARAMETRÓW MODELU RYNKOWEGO – ŹRÓDŁA I SKUTKI

Parametry modelu rynkowego oszacowane w oparciu o indeks giełdowy z reguły nie są stabilne. Zmienność zarówno parametru α, jak i β może być znaczna, co widać chociażby na przykładzie akcji spółki Żywiec SA (rys. 2.).

Rysunek 2. Wartość parametrów modeli rynkowych dla stopy zwrotu z akcji Żywca



Źródło: obliczenia własne

Przedstawione na wykresie wielkości parametrów modeli obliczono na podstawie stóp zwrotu z akcji Żywca i Warszawskiego Indeksu Giełdowego (WIG) w ciągu 150 kolejnych sesji giełdowych poprzedzających dzień ustalenia modelu. Na wykresie zaprezentowano parametry 3018 modeli rynkowych ustalonych w dniach kolejnych sesji giełdowych.

Parametr α przyjmuje wartości od 0 do 1,2, podczas gdy parametr β – od -0,2 do 1. Dodatkowo na wykresie można dostrzec przeciwzależność zmian obu parametrów. W ciągu 30 sesji giełdowych (czyli relatywnie krótkiego okresu) między 10 maja 2004 r. i 22 czerwca 2004 r. oszacowany parametr β spadł z 0,1569 do 0,0878, zaś parametr α wzrósł z 0,8429 do 0,9127. Modele rynkowe Żywca oszacowane w obu zdefiniowanych wyżej dniach o oparciu o dane pochodzące ze 150 sesji poprzedzających obie daty byłyby zatem wyraźnie różne.

Zmienność parametrów modelu nie musi wpływać negatywnie na ocenę CAR jako metody badawczej. Istotna w tym kontekście jest przyczyna zmienności. Można wyodrębnić trzy główne rodzaje źródeł zmian parametrów modelu w czasie.

1. Zdarzenie dotyczące spółki. CAR jest metodą stworzoną dla zbadania wpływu określonego zdarzenia na wycenę akcji spółki. Jeśli do takiego zdarzenia dochodzi, pojawia się nadzwyczajna stopa zwrotu w oknie obserwacji. Takie nadzwyczajne stopy zwrotu mogą w przyszłości wpłynąć na kształt modelu rynkowego, jeśli znajdą się w oknie estymacji, przy okazji zwiększając wariancję składnika resztowego „nowego" modelu. W rezultacie statystyczna istotność „nowych" nadzwyczajnych stóp zwrotu może zostać niedoszacowana.
2. Zdarzenia niezwiązane ze spółką. Na notowania spółek giełdowych mogą mieć wpływ nie tylko czynniki związane z nimi bezpośrednio, ale także dotyczące innych spółek publicznych. Wzrost popytu na akcje spółki, która ogłosiła optymistyczne prognozy może współwystępować ze zmniejszonym popytem na akcje innych spółek. Zdarzenia takiego rodzaju z reguły także zwiększają wariancję składnika resztowego i wpływają na ocenę statystycznej istotności skumulowanej nadzwyczajnej stopy zwrotu.
3. Zdarzenia modelowo-fundamentalne. Parametry strukturalne modelu mogą ulegać zmianie pod wpływem wahań koniunktury giełdowej. Jeśli zmiana koniunktury będzie miała miejsce pomiędzy oknem estymacji i obserwacji efektem może być fundamentalne zaburzenie uzyskanych wyników. Model rynkowy oszacowany na podstawie danych pochodzących z okna estymacji może charakteryzować się niską wariancją składnika resztowego, ale całkowicie nie przystawać do zmienionych warunków obowiązujących w oknie obserwacji.

Trzecią przyczynę niestabilności parametrów modelu rynkowego można uznać za szczególnie nieprzyjemną dla badacza: nie zwiększa wariancji składnika resztowego i sugeruje dobre dopasowanie modelu do rzeczywistości, choć rzeczywistość w oknie obserwacji jest już inna, niż opisana modelem.

Przykładowo, jeśli w oknie estymacji średnia stopa zwrotu z WIG wynosiła 1,02, a oszacowany model rynkowy przyjął postać:

$$R_{it} = 0 + 1 \cdot R_{WIGt} + e_{it} \tag{6}$$

zaś w oknie obserwacji średnia stopa zwrotu z WIG wynosiła 0,98, to po 30 sesjach giełdowych należałoby się spodziewać spadku poziomu notować spółki o 46%. Jeśli w tym samym czasie doszło do spadku o 20%, to obliczony CAR wyniósłby 26%. Zdarzenie, jakie miało miejsce w dniu t_0 (inne niż nastanie bessy) zostałoby uznane za pozytywne dla spółki.

Jednakże możliwe jest, że warunkach bessy (spadek WIG w tempie 2% na sesję) oszacowany wcześniej model rynkowy jest nieaktualny, zaś „obowiązujący" w nowych warunkach przyjmuje postać:

$$R_{it} = 0{,}01 + 1{,}01 \cdot R_{WIGt} + e_{it} \tag{7}$$

W takiej sytuacji zamiast oczekiwać spadku o 46% w ciągu 30 kolejnych sesji, należałoby spodziewać się niewielkiej obniżki o 0,6%. W rezultacie po rzeczywistym spadku o 20%, obliczony CAR wyniósłby -19,4%. Na takiej podstawie trudno uznać zdarzenie z dnia t_0 za pozytywne.

Problemem otwartym pozostaje jednak czy parametry modelu rynkowego są czułe na zmiany stopy zwrotu z WIG.

KONIUNKTURA GIEŁDOWA A PARAMETRY MODELI RYNKOWYCH – WYNIKI BADAŃ

Badaniu poddano 10 polskich spółek publicznych reprezentujących różne sektory i należących do najdłużej notowanych na Giełdzie Papierów Wartościowych w Warszawie. Dla każdej z nich obliczono dzienne stopy zwrotu od 3 stycznia 1999 r. do 31 grudnia 2011, a następnie oszacowano parametry serii modeli rynkowych. Modele szacowane były w oparciu o stopę zwrotu z indeksu WIG i z akcji spółki w 150-sesyjnych oknach estymacji. Liczba oszacowanych modeli dla każdej spółki przekraczała 3000[1]. Każdy oszacowany model rynkowy charakteryzuje się wartościami parametrów strukturalnych (α i β) i średnim poziomem stopy zwrotu z WIG-u. Następnie oszacowano funkcję regresji, w której za zmienną niezależną przyjęto średni poziom stopy zwrotu z WIG w oknie estymacji, zaś za zmienną zależną – parametr β modelu rynkowego. Wyniki przedstawiono w tab. 1.

[1] W 13-letnim okresie badawczym odbyło się około 3250 sesji giełdowych. Od 150. sesji możliwe było szacowanie modeli rynkowych, każda następna sesja pozwalała obliczyć kolejną parę stóp zwrotu (z akcji spółki i WIG-u) i skonstruować nowy model (obejmujący sesje od 2 do 151, od 3 do 152 itd.). Różnice w liczbie oszacowanych modeli rynkowych dla poszczególnych spółek wynikają np. z zawieszeń notowań i innych zdarzeń uniemożliwiających obliczenie stopy zwrotu.

Tabela 1. Parametry funkcji regresji parametru β modelu rynkowego względem średniego poziomu WIG w oknie estymacji

Spółka	Liczba oszacowanych modeli rynkowych	α (t-Studenta)	β (t-Studenta)
Próchnik	3082	11,08 (1,48)	-10,24 (1,37)
Żywiec	3018	-40,63 (15,61)***	40,95 (15,73)***
Mostalexp	3104	-26,61 (3,99)***	27,89 (4,18)***
Millennium	3094	-60,52 (17,45)***	61,73 (17,79)***
Efekt	3061	-9,22 (2,06)**	9,84 (2,19)**
Vistula	3092	-50,69 (14,81)***	53,31 (14,98)***
Rafako	3103	57,45 (14,28)***	-56,70 (14,09)***
Stalexport	3104	-48,65 (7,08)***	49,91 (7,25)***
Dębica	3102	4,07 (1,5)	-3,49 (1,29)
Echo	3081	-45,38 (11,50)***	46,13 (11,69)***

Źródło: obliczenia własne
** istotne na 5-procentowym poziomie istotności i dwustronnym obszarze krytycznym
*** istotne na 1-procentowym poziomie istotności i dwustronnym obszarze krytycznym

Większość badanych spółek charakteryzuje się silną zależnością między giełdową koniunkturą i wartościami parametrów modelu rynkowego. Spośród 10 badanych spółek jedynie w Próchnik i Dębica charakteryzowały się modelami rynkowymi „nieczułymi" na zmiany średniego poziomu WIG-u. W pozostałych przypadkach zagrożenie związane z zmianą koniunktury giełdowej pomiędzy oknem estymacji i oknem obserwacji można uznać za duże. Jeśli te proporcje są charakterystyczne dla wszystkich spółek giełdowych, to w 80% przypadków obliczenia CAR dokonywane najprostszą metodą jest ryzykowne.

PODSUMOWANIE I REMEDIA

Zmiana koniunktury giełdowej między oknem estymacji i obserwacji może spowodować, że skumulowana nadzwyczajna stopa zwrotu wprowadzi badacza w błąd. Można przypisać pozytywny wpływ na wycenę spółki zdarzeniu, które w rzeczywistości zmniejszyło jej wartość; można zdarzenie w rzeczywistości pozytywne oceniać negatywnie; można także przypisać jakiekolwiek znaczenie

zdarzeniu nieistotnemu. Aby uniknąć takiego ryzyka, można zamiast najczęściej stosowanego podejścia zastosować jakąś jego modyfikację.

1. Skrócić okresu między oknem estymacji i oknem obserwacji. Dzięki takiemu podejściu zmniejsza się ryzyko wyraźnej zmiany średniego poziomu stopy zwrotu z WIG. Zamiast tego zagrożenia pojawia się wówczas ryzyko pojawienia się w oknie estymacji efektów badanego zdarzenia, jeśli rynek przewiduje jego zaistnienie.
2. Zastosować do estymacji parametrów modelu takiego okna estymacji, w którym warunki rynkowe byłyby zbliżone do warunków występujących w oknie obserwacji. Z podejściem takim wiąże się z kolei ryzyko zbytniego oddalenia obu okien. Może się wówczas okazać, że wprawdzie tempo zmian WIG-u jest w obu oknach zbliżone, ale model rynkowy oszacowany na podstawie danych z okna estymacji sprzed wielu miesięcy już się zdezaktualizował.
3. Zastosować do szacowania modelu rynkowego „sztuczny" portfel akcji zamiast WIG-u. Ciekawe może okazać się skonstruowanie portfela spółek o parametrze β zbliżonym do tegoż parametru spółki, której CAR zamierza się obliczyć. Takie podejście można spotkać w literaturze [Asquith 1983], zaś jego zalety wydają się być oczywiste: spółki o zbliżonym parametrze β mogą mieć zbliżoną „czułość" obu parametrów na zmiany koniunktury rynkowej. W rezultacie na ewentualne zmiany zareagują podobnie, co z kolei pozwoli na ujawnienie rzeczywistej skumulowanej nadzwyczajnej stopy zwrotu bez względu na skalę zmiany WIG między oknem estymacji i oknem obserwacji.

BIBLIOGRAFIA

Asquith P. (1983) Merger bids, uncertainty, and stockholder returns, Journal of Financial Economics, nr 11, str. 51-83.

Asquith P., Bruner R. F., Mullins D. W., Jr. (1983) The gains to bidding firms from merger, Journal of Financial Economics, nr 11, str. 121-139.

Bradley M., Desai A., Han Kim E. (1983) The rationale behind interfirm tender offers, Journal of Financial Economics, nr 11 str. 183-206.

Bradley M., Schipani C. A. (1991) The Economic Importance of the Business Judgment Rule: An Empirical Analysis of the Trans Union Decision and Subsequent Delaware Legislation [w:] Sametz A. W. (red.) The Battle for Corporate Control: Shareholder Rights, Stakeholder Interest, and Managerial Responsibility, Salomon Brothers Center for the Study of Financial Institutions, Leonard N. Stern School of Business, New York University, New York, str. 105-129.

Bradley M., Wakeman L. M. (1983) The wealth effects of targeted share repurchases, Journal of Financial Economics, nr 11 str. 301-328.

Buzała P. (2012) reakcja inwestorów na rekomendacje giełdowe. Implikacje dla efektywności rynku akcji [w:] D. Zarzecki (red.) Zarządzanie finansami. Upowszechnianie i transfer wyników badań, Uniwersytet Szczeciński, Szczecin, str. 11-20.

Chen C., Lin J. W., Sauer D. (1997) Earning announcements, quality and quantity of information, and stock price changes, The Journal of Financial Research Vol. XX, 4, str. 483-502.

Dann L. W., DeAngelo H. (1983) Standstill agreements, privately negotiated stock repurchases, and the market for corporate control, Journal of Financial Economics, nr 11, str. 275-300.

Davidson W. N. III, Pilger Th., Szakmany A. (1998) Golden Parachutes, Board and Committee Composition, and Shareholder Wealth, The Financial Review, nr 33, 4, str. 17-32.

DeAngelo H., Rice E. M. (1983) Antitakeover charters amendments and stockholder wealth, Journal of Financial Economics, nr 11, str. 329-360.

Dodd P., Warner J. B. (1983) On corporate governance: A study of proxy contest, Journal of Financial Economics, nr 11, str. 401-438.

Eckbo B E. (1983) Horizontal mergers, collusion, and stockholder wealth, Journal of Financial Economics, nr 11, str. 241-273.

Hamrol M., Tarczyńska J. (2002) Wpływ fuzji i przejęć na wartość rynkową przediębiorstw w Polsce [w:] Duraj J. (red.) Przedsiębiorstwo na rynku kapitałowym, KAiSP Uniwersytetu Łódzkiego, Łódź, str. 131-140.

Holderness C. G., Sheehan D. P. (1988) The Role of Majority Shareholders in Publicly Held Corporations: An Exploratory Analysis, Journal of Financial Economics, nr 20, str. 317-346.

Huang Y-S., Walkling R. A. (1987) Target abnormal returns associated with acquisition announcements. Payment, acquisition form, and managerial resistance, Journal of Financial Economics, nr 19, str. 329-349.

Impson M. 1997 Market reaction to dividend-decrease announcement: public utilities vs unregulated industrial firms, „The Journal of Financial Research" nr XX/3, str. 407-422.

Jarrell G. R., Paulsen A. B. (1987) Shark repellents and stock prices: The effect of antitakeover amendments since 1980, Journal of Financial Economics, nr 19, str. 127-168.

Jarrel G. A., Paulsen A. B. (1988) Dual-Class Recapitalization as Antitakeover Mechanism: The Recent Evidence, Journal of Financial Economics, nr 20, str. 129-152.

Kamma S., Weintrop J., Wier P. (1988) Investors' Perception of the Delaware Supreme Court Decision in Unocal v. Mesa, Journal of Financial Economics, nr 20, str. 419-430.

Klein A., Rosenfeld J. (1988) Targeted Share Repurchase and Top Management Changes, Journal of Financial Economics, nr 20, str. 493-506.

Linn S. C., McConnell J. J. (1983) An empirical investigation of the impact of 'antitakeover' amendments on common stock prices, Journal of Financial Economics, nr 11, str. 361-399.

Loh Ch., Bezjak J. R., Toms H. (1995) Voluntary Corporate Divestitures as Antitakeover Mechnisms, The Financial Review, nr 30, 1, str. 41-60.

Loh Ch., Rathinasamy R. S. (1997) The Impact of Antitakeover Devices on the Valuations Consequences of Voluntary Corporate Selloffs, The Financial Review, nr 32, 3, str. 691-708.

Magenheim E. B., Mueller D. C. (1988) Are Acquiring-Firm Shareholders Better Off After an Acquisition? [w:] Coffee J. C., Lowenstein L., Rose-Ackerman S. (red.) Knights, Raiders, and Targets, Oxford University Press, New York, Oxford, str. 171-193.

Malatesta P. H. (1983) The wealth effect of merger activity and the objective functions of merging firms, Journal of Financial Economics, nr 11, str. 155-181.

Malatesta P. H., Walkling R. A. (1988) Poison pill securities: Stockholder wealth, profitability, and ownership structure, Journal of Financial Economics, nr 20, str. 347-376.

Nogalski B., Szymański M. (2003) Sterowanie wartością spółki giełdowej za pomocą prognozy zysku [w:] Duraj J. (red.) Wartość przedsiębiorstwa – z teorii i praktyki zarządzania, Szkoła Wyższa im. Pawła Włodkowica w Płocku, Uniwersytet Łódzki, Wydawnictwo Naukowe Novum, Płock-Łódź, str. 65-73.

Ryngaert M. (1988) The Effect of Poisson Pill Securities on Shareholders Wealth, Journal of Financial Economics, nr 20, str. 377-417.

Schipper K., Thompson R. (1983) Evidence on the capitalized value of merger activity for acquiring firms, Journal of Financial Economics, nr 11, str. 85-119.

Stillman R. (1983) Examining antitrust policy towards horizontal mergers, Journal of Financial Economics, nr 11, str. 225-240.

Szyszka A. (2001a) Reakcje inwestorów. Kwartalne raporty finansowe, Nasz rynek kapitałowy, nr 1(133), str. 30-38.

Szyszka A. (2001b) Reakcje inwestorów. Korekty prognoz finansowych, Nasz rynek kapitałowy, nr 2(134), str. 89-92.

Thosar S. (1996) Tenders Offers and Target Management Responses: Managerial Entrenchment Versus Stockholder Interest Revisited, The Financial Review, nr 31, 1, February, str. 87-104.

Warner J. B., Watts R. L., Wruck K. H. (1988) Stock Prices and Top Management Changes, Journal of Financial Economics, nr 20, str. 461-492.

Wier P. (1983) The costs of antimerger lawsuits: Evidence from the stock market, Journal of Financial Economics, nr 11, str. 207-224.

STABILITY OF MARKET MODEL PARAMETERS ESTIMATED USING DAILY RETURNS OF WARSAW STOCK EXCHANGE INDEX

Abstract: The aim of this paper is describing problems one can face using the CAR method, when the market direction change occurs between the estimation window and the event window. Such situation can cause the market model is inappriopriate to market conditions in the event window. Consequently calculated cumulative abnormal return becomes useless as a measure of market reaction to the event.

Keywords: cumulative abnormal return, Warsaw Stock Exchange, Warsaw Stock Exchange Index

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 232 – 242

OCENA STATYSTYCZNA RYNKU USŁUG GASTRONOMICZNYCH W POLSCE

Stanisław Urban, Anna Kowalska
Katedra Ekonomiki i Organizacji Gospodarki Żywnościowej
Uniwersytet Ekonomiczny we Wrocławiu
e-mail: stanislaw.urban@ue.wroc.pl; anna.kowalska@ue.wroc.pl

Streszczenie: Celem artykułu jest ocena rozwoju usług gastronomicznych w Polsce. Zaprezentowano w nim zróżnicowanie poziomu rozwoju w poszczególnych regionach i województwach. Wykorzystano do tego analizę skupień, taksonomię, liniową funkcję trendu i dynamikę zmian.

Słowa kluczowe: analiza skupień, gastronomia, taksonomia

WSTĘP

Gastronomia to dział gospodarki zaliczany do sektora usługowego jako część składowa handlu. Gastronomia zajmuje w kompleksie żywieniowym miejsce pomiędzy przemysłem spożywczym i handlem. Przedsiębiorstwa gastronomiczne prowadzą działalność produkcyjną, handlową i usługową. Działalność produkcyjna polega na wytwarzaniu wyrobów kulinarnych, zaopatrywaniu ludności w posiłki i napoje gotowe do spożycia. Działalność usługowa polega na stworzeniu warunków umożliwiających konsumpcję, a także świadczenie różnego rodzaju usług zaspokajających potrzeby konsumentów w zakresie rozrywki, spędzania wolnego czasu, odpoczynku i spotkań towarzyskich.

Placówki gastronomiczne mogą prowadzić dodatkową działalność np. sprzedaż produktów kulinarnych przeznaczonych do spożycia w domu. Gastronomia prowadzi też działalność handlową polegającą na sprzedaży wyprodukowanych posiłków oraz towarów handlowych.

Gastronomia w Polsce należy do najszybciej rozwijających się działów gospodarki narodowej, przy czym ma ona szanse dalszego rozwoju. Jednocześnie występuje silne zróżnicowanie regionalne w zakresie wyposażenia w zakłady gastronomiczne, które nadal rośnie.

Rozwój gastronomii w dużym stopniu wiąże się z rozwojem turystyki. Gastronomia ze względu na dużą pracochłonność może dać zatrudnienie wielu pracownikom, co sprzyja zmniejszaniu bezrobocia. Stąd jej rozwój jest bardzo ważny i potrzebny.

Celem artykułu jest dokonanie statystycznej oceny rozwoju usług gastronomicznych w Polsce w latach 1999-2010.

MATERIAŁ ŹRÓDŁOWY I METODY BADAWCZE

Badaniami objęto placówki gastronomiczne w Polsce w latach 1999-2010 zatrudniające powyżej 9 osób. Wyszczególniono podział rynku gastronomicznego ze względu na typy placówek, ich liczbę oraz przychody uzyskiwane przez nie w poszczególnych regionach i województwach. Dane pozyskane zostały z Lokalnej Bazy Danych prowadzonej przez GUS. Zwrócono też uwagę na rozmieszczenie przestrzenne gastronomii [Grzelak, Pająk 2010].

Do realizacji celu badawczego przyjęto procedurę wielowymiarowej analizy skupień oraz zastosowano metodę taksonomiczną Czekanowskiego do określenia grup regionów i województw podobnych. W pracy zastosowano również liniową funkcję trendu do określenia przebiegu zmian dotyczących liczby placówek oraz przychodów uzyskanych z ich działalności. Przeanalizowano również dynamikę zmian dotyczącą liczby obietków w poszczególnych województwach w latach 1999 i 2004 oraz 2005 i 2010.

WYNIKI BADAŃ

W tabeli 1 przedstawiono rozwój placówek gastronomicznych w Polsce w poszczególnych województwach w latach 1999-2010 z uwzględnieniem podziału na 2 pięcioletnie okresy: 1999-2004 i 2005-2010. Okres pierwszy to ostatnie pięciolecie przed przyjęciem Polski do Unii Europejskiej a okres drugi to pierwsze pięciolecie przynależności Polski do Unii Europejskiej. Porównanie tych okresów miało wykazać jak przystąpienie Polski do Unii Europejskiej wpłynęło na wzrost liczby placówek gastronomicznych. Uwzględniono przy tym liczbę obiektów gastronomicznych ogółem i następujące typy placówek: restauracje, bary, stołówki oraz punkty gastronomiczne. Wskaźniki dynamiki zmian wyrażone w % obliczono oddzielnie dla okresów 1999-2004 i 2005-2010.

Liczba placówek gastronomicznych ogółem w Polsce w latach 1999-2004 zmniejszyła się o 3,5%, a w okresie lat 2005-2010 wzrosła o 21,4%. W pierwszym okresie w 11 województwach zmniejszyła się liczba obiektów gastronomicznych, w tym najbardziej w województwie śląskim – o 36,4%. Natomiast w 5 województwach przybyło obiektów gastronomicznych, w tym najwięcej w województwie mazowieckim.

Tabela 1. Liczba obiektów gastronomicznych ogółem i w poszczególnych typach placówek w latach 1999-2004 i 2005-2010 oraz wskaźniki dynamiki wg województw [1]

Województwa	ogółem						restauracje						bary						stołówki						punkty gastronomiczne					
	1999	2004	dyn.	2005	2010	dyn.	1999	2004	dyn.	2005	2010	dyn.	1999	2004	dyn.	2005	2010	dyn.	1999	2004	dyn.	2005	2010	dyn.	1999	2004	dyn.	2005	2010	dyn.
	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%	ob.	ob.	%
POL	9303	8974	**96,5**	11872	14408	**121,4**	2203	2374	**107,8**	3386	4625	**136,6**	2932	2665	**90,9**	3554	3995	**112,4**	2538	2200	**86,7**	2640	2088	**79,1**	1630	1735	**106,4**	2292	3700	**161,4**
ŁD	257	318	**123,7**	432	664	**153,7**	63	99	**157,1**	140	287	**205,0**	93	106	**114,0**	166	159	**95,8**	48	55	**114,6**	58	69	**119,0**	53	58	**109,4**	68	149	**219,1**
MZ	1692	2225	**131,5**	3251	3992	**122,8**	261	415	**159,0**	635	913	**143,8**	606	730	**120,5**	1024	1025	**100,1**	333	409	**122,8**	681	495	**72,7**	492	671	**136,4**	911	1559	**171,1**
MP	930	744	**80,0**	996	1480	**148,6**	233	225	**96,6**	364	515	**141,5**	349	238	**68,2**	306	400	**130,7**	208	158	**76,0**	186	169	**90,9**	140	123	**87,9**	140	396	**282,9**
ŚL	1474	1311	**88,9**	1630	1581	**97,0**	278	316	**113,7**	419	570	**136,0**	418	346	**82,8**	448	449	**100,2**	537	394	**73,4**	438	187	**42,7**	241	255	**105,8**	325	375	**115,4**
LB	409	499	**122,0**	584	509	**87,2**	69	90	**130,4**	98	123	**125,5**	76	98	**128,9**	110	72	**65,5**	223	254	**113,9**	285	227	**79,6**	41	57	**139,0**	91	87	**95,6**
PK	425	338	**79,5**	432	511	**118,3**	92	99	**107,6**	144	183	**127,1**	137	109	**79,6**	138	178	**129,0**	119	79	**66,4**	74	68	**91,9**	77	51	**66,2**	76	82	**107,9**
PL	347	305	**87,9**	356	334	**93,8**	62	66	**106,5**	78	107	**137,2**	102	81	**79,4**	99	108	**109,1**	140	110	**78,6**	107	53	**49,5**	43	48	**111,6**	72	66	**91,7**
ŚK	214	136	**63,6**	204	283	**138,7**	45	42	**93,3**	65	93	**143,1**	84	49	**58,3**	67	94	**140,3**	59	27	**45,8**	34	37	**108,8**	26	18	**69,2**	38	59	**155,3**
LS	172	214	**124,4**	225	297	**132,0**	71	73	**102,8**	80	107	**133,8**	61	65	**106,6**	77	108	**140,3**	23	34	**147,8**	34	42	**123,5**	17	42	**247,1**	34	40	**117,6**
WP	782	609	**77,9**	829	968	**116,8**	257	220	**85,6**	326	432	**132,5**	260	192	**73,8**	275	249	**90,5**	144	138	**95,8**	140	114	**81,4**	121	59	**48,8**	88	173	**196,6**
ZP	343	326	**95,0**	453	646	**142,6**	86	74	**86,0**	124	210	**169,4**	100	97	**97,0**	124	163	**131,5**	102	110	**107,8**	135	158	**117,0**	55	45	**81,8**	70	115	**164,3**
DŚ	799	708	**88,6**	779	1219	**156,5**	257	278	**108,2**	357	414	**116,0**	203	184	**90,6**	187	399	**213,4**	217	155	**71,4**	135	145	**107,4**	122	91	**74,6**	100	261	**261,0**
OP	278	226	**81,3**	302	254	**84,1**	68	68	**100,0**	104	94	**90,4**	96	74	**77,1**	109	87	**79,8**	87	59	**67,8**	51	43	**84,3**	27	25	**92,6**	38	30	**78,9**
KP	270	309	**114,4**	412	524	**127,2**	104	105	**101,0**	145	181	**124,8**	99	93	**93,9**	117	154	**131,6**	31	57	**183,9**	72	72	**100,0**	36	54	**150,0**	78	117	**150,0**
PM	526	391	**74,3**	570	702	**123,2**	163	119	**73,0**	175	234	**133,7**	165	128	**77,6**	203	240	**118,2**	109	56	**51,4**	94	104	**110,6**	89	88	**98,9**	98	124	**126,5**
WM	385	315	**81,8**	417	444	**106,5**	94	85	**90,4**	132	162	**122,7**	83	75	**90,4**	104	110	**105,8**	158	105	**66,5**	116	105	**90,5**	50	50	**100,0**	65	67	**103,1**

Symbole województw: DŚ – dolnośląskie, KP – kujawsko-pomorskie, LB – lubelskie, LS – lubuskie, ŁD – łódzkie, MP – małopolskie, MZ – mazowieckie, OP – opolskie, PK – podkarpackie, PL – podlaskie, PM – pomorskie, ŚL – śląskie, ŚK – świętokrzyskie, WM – warmińsko-mazurskie, WP – wielkopolskie, ZP – zachodniopomorskie.

Źródło: obliczenia własne na podstawie Banku Danych Lokalnych, www.stat.gov.pl/bdl

[1] obiekty zatrudniające powyżej 9 osób

W pięcioleciu po przystąpieniu do Unii Europejskiej w 12 województwach wzrosła liczba placówek gastronomicznych w tym największy przyrost nastąpił w województwie dolnośląskim – o 56,5%. W tym samym okresie liczba obiektów gastronomicznych zmniejszyła się w 4 województwach, w tym najbardziej w województwie opolskim – o 15,9%.

Liczba restauracji w Polsce w obydwu analizowanych okresach wzrosła, przy czym ich przyrost przed akcesją był znacznie niższy i wynosił 7,8% a po akcesji 36,6%. W pierwszym okresie liczba restauracji wzrosła w 10 województwach w tym najbardziej w województwie mazowieckim – o 59%. Natomiast liczba restauracji zmniejszyła się w 6 województwach w tym najbardziej w województwie pomorskim – o 27%. W drugim okresie liczba restauracji wzrosła w 15 województwach w tym najbardziej w łódzkim – 105%. Zmniejszenie liczby restauracji w tym czasie nastąpiło tylko w jednym województwie opolskim i wyniosło 9,6%.

Liczba barów w Polsce w pięcioleciu przed akcesją zmniejszyła się o 9,1% a w następnym pięcioleciu wzrosła o 12,4%. W latach 1999-2004 liczba barów zwiększyła się tylko w 4 województwach w tym najbardziej w województwie lubuskim – o 28,9%. Natomiast w 12 województwach liczba barów uległa zmniejszeniu w tym najbardziej w województwie śląskim – o 41,7%. W pięcioleciu po akcesji Polski do Unii Europejskiej w 12 województwach nastąpił wzrost liczby barów w tym najwyższy w województwie dolnośląskim o 113,4%. W 4 pozostałych województwach liczba barów uległa zmniejszeniu którego największy rozmiar miał miejsce w województwie lubelskim – o 34,5%.

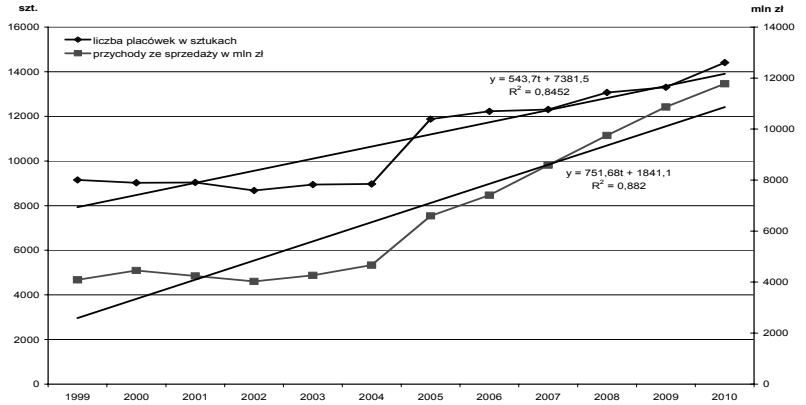
Ilość stołówek w całym analizowanym okresie wykazywała w Polsce tendencje malejące. Wiąże się to z dążeniem jednostek prowadzących stołówki do ograniczenia kosztów. W okresie przed akcesją w 6 województwach liczba stołówek wzrosła a w 10 województwach spadła.

Najwyższy wzrost ilości stołówek miał miejsca w województwie kujawsko-pomorskim – o 83,9% a największy spadek w województwie podkarpackim – o 33,6%. W okresie po akcesji wzrost ilości stołówek nastąpił w 6 województwach w tym największy w województwie lubuskim – o 23,5%. Natomiast liczba stołówek zmniejszyła się w 10 województwach w tym najbardziej w województwie podlaskim – o 50,5%.

Liczba punków gastronomicznych w Polsce w pierwszym okresie wzrosła o 6,4% a w drugim okresie a 61,4%. W pierwszym okresie przyrost liczby punktów gastronomicznych nastąpił w 8 województwach w tym największy w województwie mazowieckim – 36,4%. Także w 8 województwach ilość punków gastronomicznych zmniejszyła się w tym najbardziej w województwie podkarpackim – o 30,8%. W okresie poakcesyjnym w 13 województwach nastąpił przyrost liczby punktów gastronomicznych w tym najwyższy w województwie małopolskim – o 182,9. Natomiast w 3 województwach zmniejszyła się liczba punktów gastronomicznych w tym najbardziej w województwie opolskim – o 20,1%.

Z uzyskanych danych przedstawiono 11- elementowy szereg czasowy liczby placówek gastronomicznych oraz przychodów z ich działalności w Polsce (rys.1). Zarówno liczba placówek jak i ich przychody w latach 1999-2010 miały tendencję wzrostową. Za pomocą arkusza kalkulacyjnego wykreślono obraz liczby tych placówek jak i ich przychodów oraz funkcję liniową trendu dla tych dwóch zmiennych.

Rysunek 1. Liczba placówek gastronomicznych (zatrudniających powyżej 9 osób) i ich przychody (po uwzględnieniu inflacji) w latach 1999-2010 [szt., mln zł]



Źródło: opracowanie własne na podstawie Roczników Statystycznych RP 1999-2011, GUS oraz Bank Danych Lokalnych, www.stat.gov.pl/bdl

Podobnie obliczono współczynniki determinacji R^2 dla każdej z funkcji trendu, w celu zbadania czy poprawnie opisują one rzeczywiste prawidłowości. Założono, że dla $R^2 \geq 0{,}60$ dopasowanie funkcji do danych empirycznych jest dobre. W przypadku analizowanych danych liniowe funkcje trendu bardzo dobrze odzwierciedlały rzeczywistość ($0{,}845 < R^2 < 0{,}882$).

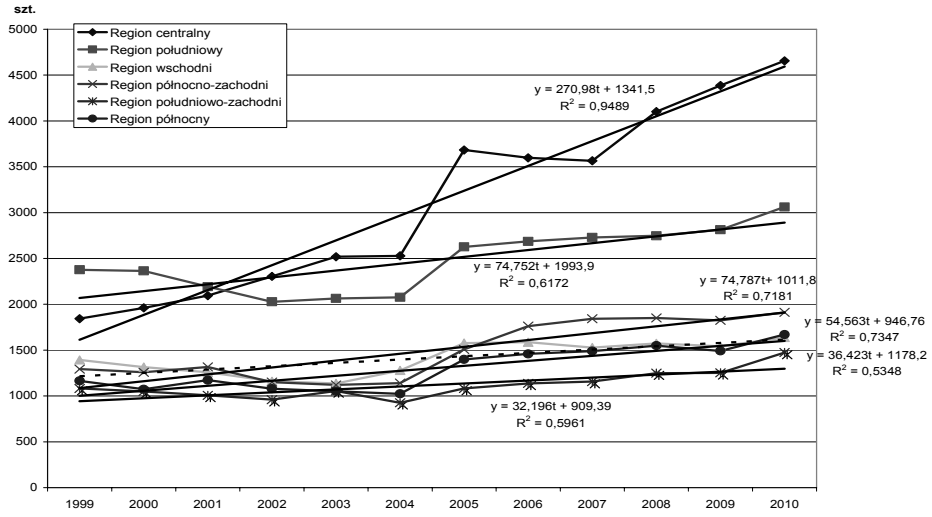
Obliczanie R^2 jest zasadne również ze względu na prognozowanie wartości cechy. Na podstawie szeregu czasowego obejmującego dane roczne można sporządzić prognozę wartości badanej cechy na rok albo dwa lata do przodu. Prognozę na dalsze okresy (3 do 5 lat naprzód) opracowuje się tylko wtedy, gdy R^2 jest bliskie 1 a dane zawierają powyżej 20 elementów [Wawrzynek 2007]. W naszym przypadku prognoza zmiany cechy dla liczby placówek oraz ich przychodów, została określona na 2 lata.

Na podstawie wykreślonych funkcji można zauważyć, że ogólne tendencje rozwojowe liczby placówek oraz przychodów z nich, w ostatniej dekadzie, były rosnące. Z równań wynika, że z każdym rokiem liczba placówek rosła średnio o 544, a przychody o 752 mln zł.

Przedstawiono również szereg czasowy liczby placówek gastronomicznych w poszczególnych regionach w latach 1999-2010, gdzie wyznaczono też liniowe funkcje trendu (rys. 2). W przypadku analizowanych danych funkcje trendu dobrze

odzwierciedlały rzeczywistość dla regionu centralnego, północnego i północno-zachodniego oraz południowego ($0{,}617<R^2<0{,}949$). Natomiast trochę słabiej wygląda to dla regionu południowo - zachodniego i wschodniego ($0{,}534<R^2<0{,}596$).

Rysunek 2. Liczba placówek gastronomicznych (zatrudniających powyżej 9 osób) w latach 1999-2010 [szt.]



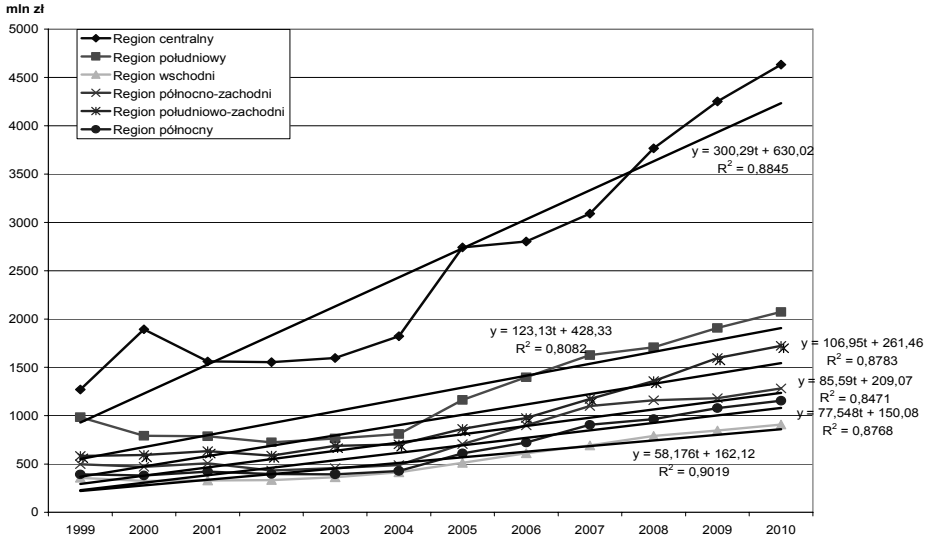
Źródło: opracowanie własne na podstawie Roczników Statystycznych RP 1999-2011, GUS oraz Banku Danych Lokalnych, www.stat.gov.pl/bdl

Również w wykreślonych funkcjach dla poszczególnych regionów można zauważyć, że trendy dla liczby obiektów zarówno w badanych obszarach jak i latach, były rosnące. Z równań wynika, że z każdym rokiem liczba placówek we wszystkich regionach rosła w tym najwięcej w regionie centralnym, gdyż średnio o 271, najmniej natomiast w południowo - wschodnim o 32.

Mimo, iż zwyczaje Polaków w kwestii jedzenia poza domem nadal znacznie odbiegają od tego co można zaobserwować w innych krajach Unii Europejskich, z danych uzyskanych w Banku Danych Lokalnych wynika, iż widoczny jest ogólny wartościowy wzrost rynku gastronomicznego w Polsce jak i we wszystkich jego regionach (rys. 3).

Z równań linii trendu wynika, że największe, gdyż o ponad 300 mln zł rocznie, wzrastały przychody gastronomii w badanych latach w regionie centralnym Polski, natomiast najmniejsze we wschodnim średnio około 58 mln zł rocznie. Duże średnioroczne przyrosty przychodów widoczne były również w regionie południowym(o ok. 123 mln zł) oraz południowo-zachodnim (o prawie 107 mln zł).

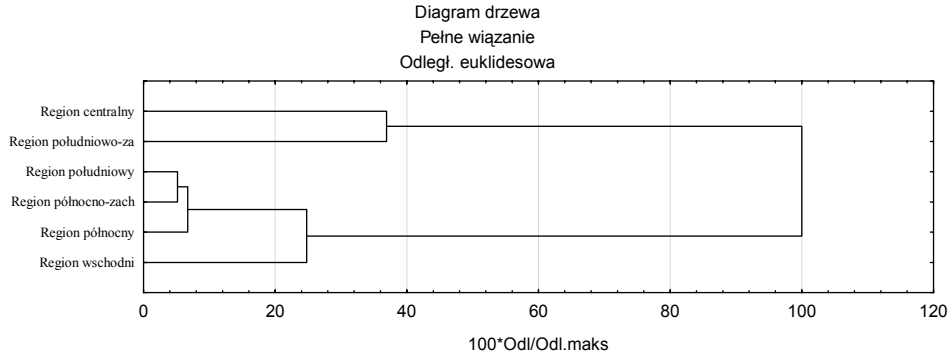
Rysunek 3. Przychody placówek gastronomicznych (zatrudniających powyżej 9 osób) w latach 1999-2010 [mln zł] (przychody stałe)



Źródło: opracowanie własne na podstawie Roczników Statystycznych RP 1999-2011, GUS oraz Bank Danych Lokalnych, www.stat.gov.pl/bdl

Stosując analizę skupień wyróżniono grupy, w których średnie przychody przypadające na jedną placówkę w danym regionie były zbliżone (rys.4).

Rysunek 4. Zróżnicowanie regionów z zależności od średniego przychodu przypadającego na jedną placówkę w danym regionie w latach 1999-2010



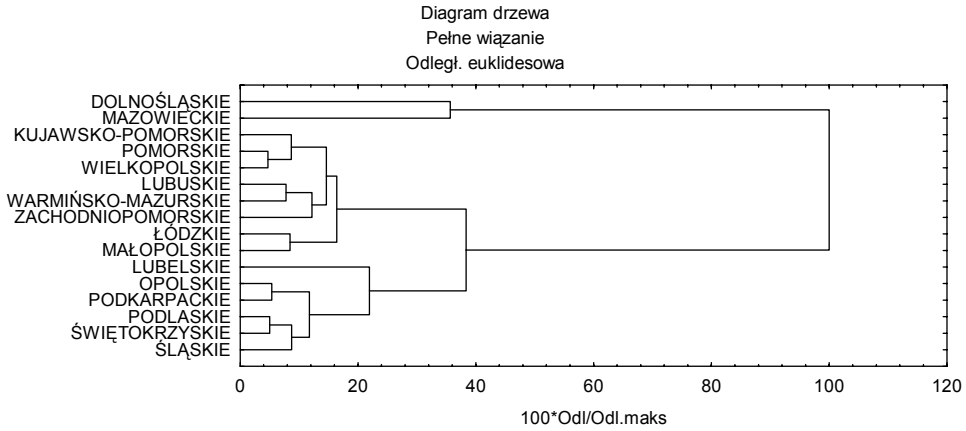
Źródło: opracowanie własne na przykładzie Roczników Statystycznych RP 2000-2011, GUS oraz Bank Danych Lokalnych, www.stat.gov.pl/bdl

Dzięki tej analizie można wyodrębnić dwie takie grupy. W pierwszej z nich znajdują się dwa regiony: centralny i południowo-zachodni, gdzie średni przychód przypadający na jedną placówkę wyniósł w badanym okresie prawie 100 tys. zł,

podczas gdy w pozostałych regionach ich wielkość była nawet o połowę niższa. Wśród pozostałych regionów najbardziej podobne były regiony południowy oraz północno-zachodni. Nieznacznie od nich odbiegał region północny i wschodni.

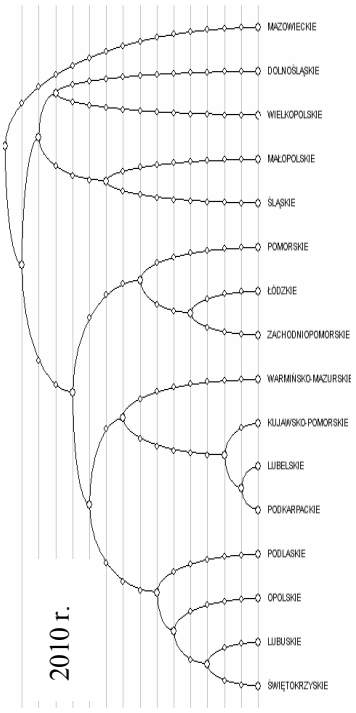
Za tak duży średni przychód z usług gastronomicznych, w latach 1999-2010, w tych dwóch regionach odpowiedzialne były przede wszystkim dwa województwa, dolnośląskie i mazowieckie, wpływ pozostałych województw był niewielki (rys.5). Wśród województw można wydzielić trzy takie grupy. Do pierwszej zaliczono województwa dolnośląskie i mazowieckie, które należą do regionów centralnego oraz południowo-zachodniego i wykazały się największym średnim przychodem przypadającym na jedną placówkę gastronomiczną. Grupę drugą stanowią województwa z regionu północno – zachodniego (wielkopolskie, lubuskie i zachodnio-pomorskie) oraz północnego (kujawsko-pomorskie i warmińsko-mazurskie), jak również województwa łódzkie i małopolskie. Do trzeciej grupy zaliczono wszystkie województwa z podregionu wschodniego (podkarpackie, podlaskie, lubelskie i świętokrzyskie) oraz województwa opolskie i śląskie.

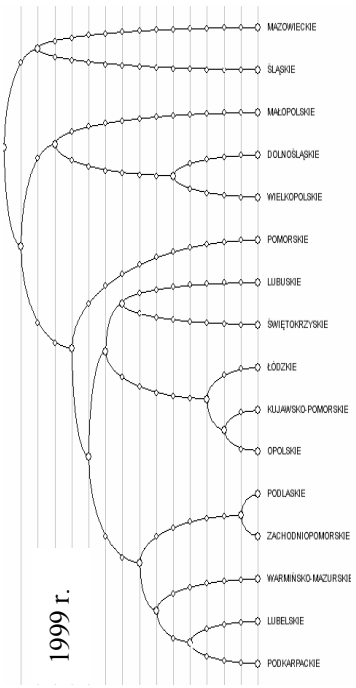
Rysunek 5. Zróżnicowanie województw z zależności od średniego przychodu przypadającego na jedną placówkę w danym regionie w latach 1999-2010



Źródło: opracowanie własne na podstawie Roczników Statystycznych RP 2000-2011, GUS oraz Bank Danych Lokalnych, www.stat.gov.pl/bdl

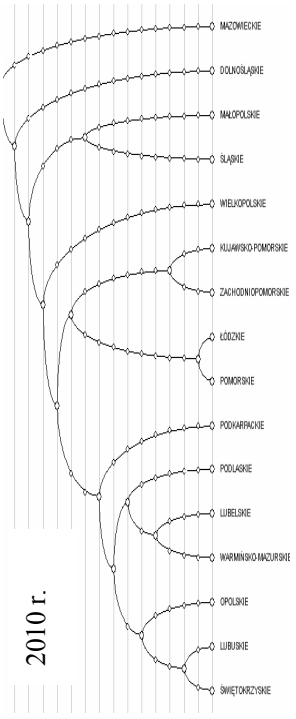
Porównując zróżnicowanie liczby placówek gastronomicznych w poszczególnych województwach oraz przychody z ich usług w latach 1999 i 2010 roku wykorzystując do tego metodę taksonomiczną Czekanowskiego można zauważyć pewne zmiany. Grupowanie przeprowadzono za pomocą programu zamieszczonego w książce M. Kolendy: Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza wielowymiarowa.

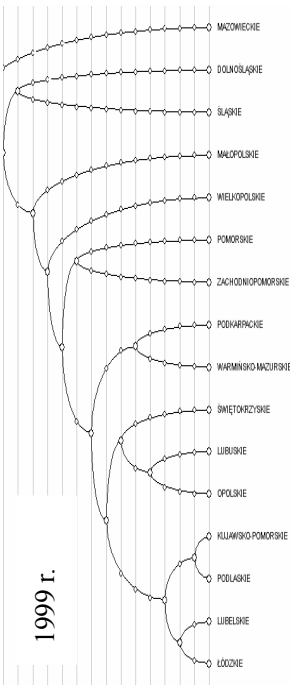




Rys. 6. Rozmieszczenie placówek gastronomicznych w poszczególnych województwach 1999 i 2010 roku

Źródło: opracowanie własne na podstawie Banku Danych Lokalnych www.stat.gov.pl/bdl





Rys. 7. Przychody z usług gastronomicznych w poszczególnych województwach 1999 i 2010 roku

Źródło: opracowanie własne na podstawie Banku Danych Lokalnych www.stat.gov.pl/bdl

Placówki gastronomiczne można podzielić na cztery grupy (rys. 6). Pierwszą z nich stanowią województwa, w których liczba placówek w 1999 roku była najmniejsza, tzn. poniżej 300 obiektów. Do grupy tej należy zaliczyć 6 województw z czego największe podobieństwa widoczne są między opolskim i kujawsko-pomorskim oraz już z większym zróżnicowaniem łódzkim, lubuskim, świętokrzyskim i pomorskim, gdzie liczba placówek była najmniejsza. Przyczyny tak małej liczby placówek gastronomicznych można upatrywać w tym, iż większość z tych województw znajduje się w części wschodniej Polski, gdzie atrakcyjność miejscowości i zamożność mieszkańców nie jest wysoka. Wpływa to ujemnie na przychody z gastronomii. Duże podobieństwo w liczbie placówek występowało również w województwach dolnośląskim i wielkopolskim oraz małopolskim. Natomiast największą liczbą placówek gastronomicznych wyróżniały się województwa śląskie i mazowieckie.

Po 11 latach nastąpiły pewne zmiany między poszczególnymi grupami województw. Bardzo dynamicznie rozwinęły się pod tym względem województwo mazowieckie, również choć już znacznie odbiegające od niego województwo dolnośląskie i wielkopolskie jak również śląskie i małopolskie. Z grupy o najmniejszej liczbie placówek gastronomicznych odłączyły się województwa kujawsko-pomorskie i łódzkie. W dalszym ciągu najsłabiej rozwija się gastronomia w województwie świętokrzyskim, lubuskim oraz opolskim. Do tej grupy dołączyło również województwo podlaskie.

Na podstawie uzyskanych przez placówki przychodów w poszczególnych województwach podzielono je na cztery grupy (rys.7). W pierwszej z nich znalazły się województwa, które w 1999 roku miały najniższe przychody z działalności gastronomicznej, tj. łódzkie, lubelskie, podlaskie, kujawsko-pomorskie, opolskie, lubuskie, świętokrzyskie warmińsko-mazurskie oraz podkarpackie. Do grupy drugiej weszły: zachodniopomorskie, pomorskie, wielkopolskie oraz małopolskie. Do grupy trzeciej: śląskie i dolnośląskie, a w grupie czwartej o najwyższych przychodach znalazło się województwo mazowieckie.

W 2010 roku widoczne są pewne roszady w tych grupach. Województwa łódzkie i kujawsko-pomorskie przeskoczyły z grupy o najniższych przychodach gastronomii do grupy drugiej. Do grupy trzeciej przeszło województwo małopolski, w którym przychody przez ten okres zwiększyły się o ponad 805 mln zł. Nie uległa zmianie grupa czwarta w której województwo mazowieckie umocniło jeszcze bardziej swoją pozycję w stosunku do pozostałych, a przychody zwiększyły się prawie czterokrotnie.

PODSUMOWANIE I WNIOSKI

Gastronomia w Polsce w okresie 1999-2010 szybko rozwijała się, przy czym dynamika rozwoju uległa przyspieszeniu po przystąpieniu Polski do Unii Europejskiej. Zmieniła się przy tym struktura polskiej gastronomii. Zwiększyła się ilość punktów gastronomicznych i restauracji a zmniejszeniu uległa liczba stołówek. Zmieniło się też przestrzenne rozmieszczenie obiektów gastronomicznych. Największy rozwój gastronomii miał miejsce w województwach mazowieckim, łódzkim i lubuskim. Największe zmniejszenie liczby lokali gastronomicznych wystąpiło w województwach opolskim i śląskim, znanych wcześniej z dobrze rozwiniętej sieci gastronomicznej.

Przeprowadzone badania potwierdziły przydatność metod wielowymiarowej analizy skupień i taksonomii Czekanowskiego do określenia grup województw podobnych pod względem wyposażenia w obiekty gastronomiczne.

BIBLIOGRAFIA

Grzelak A., Pająk K.(red.) (20101) Nowe trendy w metodologii nauk ekonomicznych. Tom I. Wydawnictwo Uniwersytet Ekonomiczny w Poznaniu.

Kolenda M. (2006) Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza obiektów wielocechowych. Wyd. AE we Wrocławiu.

Roczniki Statystyczne Rzeczpospolitej Polskiej 1999-2011, GUS, Warszawa 2000-2012

Wawrzynek J. (2007) Metody opisu i wnioskowania statystycznego Wyd. AE we Wrocławiu, 2007

www.stat.gov.pl - Bank Danych Lokalnych, GUS

STATISTICAL EVALUATION OF CATERING SERVICES MARKET IN POLAND

Abstract: The aim of this paper is to assess the development of catering services in Poland. There was presented the different levels of development in different regions and provinces using a cluster analysis, a taxonomy, a linear trend function and dynamics of change.

Keywords: cluster analysis, gastronomy, taxonomy

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 243 – 252

O PEWNYCH KRYTERIACH INWESTOWANIA W OPCJE NA AKCJE

Tomasz Warowny
Katedra Metod Ilościowych w Zarządzaniu
Politechnika Lubelska
e-mail: t.warowny@pollub.pl

Streszczenie: W artykule scharakteryzowano opcje na akcje. Wyjaśniono podstawowe pojęcia, takie jak: termin wykonania, termin wygaśnięcia, cena wykonania, cena opcji. Do opisu ewolucji cen akcji wykorzystano geometryczny ruch Browna. Sformułowano kilka problemów dotyczących inwestowania w opcje na akcje otrzymując zadania programowania stochastycznego. Korzystając z własności ruchu Browna pokazano, w jaki sposób szacować prawdopodobieństwa zdarzeń polegających na osiągnięciu przez inwestora zysków na żądanym poziomie lub przy ustalonym poziomie ryzyka. Dla każdego z zadań dokonano przykładowych obliczeń.

Słowa kluczowe: opcje na akcje, standardowy ruch Browna, geometryczny ruch Browna

WSTĘP

Opcje na akcje są jednym z najpopularniejszych instrumentów pochodnych na świecie. Umożliwiają one inwestorom osiągać zarówno ponadprzeciętne zyski, gdy cena wykonania opcji jest lepsza niż cena oferowana na wolnym rynku, jak i zabezpieczać się przed nadmiernymi stratami związanymi ze zmianami kursów akcji. Celem niniejszej pracy jest pokazanie, w jaki sposób można szacować prawdopodobieństwa osiągnięcia tych zysków z uwzględnieniem oczekiwań inwestora co do inwestycji w opcje.

CHARAKTERYSTYKA OPCJI

Opcja kupna (ang. call) jest kontraktem, który daje nabywcy prawo do kupna ustalonej ilości instrumentu podstawowego, na który opcja została wystawiona, po

określonej cenie i w ustalonym terminie. Opcja sprzedaży (ang. put) daje prawo do sprzedaży instrumentu podstawowego po określonej cenie i w ustalonym terminie. Dla nabywcy opcja jest prawem, a nie obowiązkiem. Skorzysta z tego prawa, gdy będzie mu się to opłacało. Natomiast sprzedający (wystawiający) opcje ma obowiązek, na życzenie nabywcy opcji, odsprzedać (w przypadku opcji kupna) lub odkupić (w przypadku opcji sprzedaży) instrument podstawowy, na który opcja jest wystawiona.

Z inwestycjami w opcje wiąże się duże ryzyko. W związku z tym wystawca opcji musi złożyć depozyt zabezpieczający (ang. margin), który ma zagwarantować spełnienie jego ewentualnych zobowiązań [Luenberger 2003].

Termin, w którym posiadacz opcji wykorzystał swoje prawo nazywamy terminem wykonania (ang. exercise date). Termin, po którym opcja traci swoją ważność i nie może być wykonana nazywamy terminem wygaśnięcia opcji (ang. maturity, expiration date).

Wyróżnia się dwa typy opcji:

- europejskie, które mogą być wykonane tylko w dniu wygaśnięcia opcji,
- amerykańskie, które mogą być wykonane w dowolnym dniu do terminu wygaśnięcia opcji.

Nazwy obu typów nie są w żaden sposób związane z miejscem obrotu. Terminy te w przeszłości odnosiły się do różnych zasad handlu opcjami, jakie obowiązywały w Europie i Ameryce. Większość opcji, którymi handluje się na światowych rynkach to opcje amerykańskie. Powodem tego, między innymi, jest fakt, że opcje europejskie są bardziej narażone na manipulacje w okresie bliskim terminowi wygaśnięcia. [Weron, Weron 1998].

Cena instrumentu podstawowego ustalona w kontrakcie nazywa się ceną wykonania **(**ang. strike, exercise price). Jeżeli opcję opłaca się wykonać mówimy, że opcja jest w cenie (ang. in the money), np. w przypadku opcji sprzedaży dzieje się tak, gdy cena wykonania jest wyższa niż cena instrumentu podstawowego. Gdy cena instrumentu podstawowego jest równa cenie wykonania mówimy, że opcja jest po cenie (ang. at the money). Opcja nie jest w cenie (ang. out of the money), gdy nie opłaca się jej wykonać, np. w przypadku opcji sprzedaży dzieje się tak, gdy cena wykonania jest niższa niż cena instrumentu podstawowego.

Opcjami na światowych rynkach handluje się od dawna. Na Giełdzie Papierów Wartościowych w Warszawie oferowane są wyłącznie opcje na indeks WIG 20. Opcje na akcje były notowane od października 2005 r., ale od 4 lipca 2007 r. zawieszono wprowadzanie do obrotu kolejnych serii opcji na akcje oraz zawieszono obrót wszystkimi seriami opcji na akcje.[1]

[1] www.gpw.pl/opcje_instrumenty

PROCES WIENERA I MODEL EWOLUCJI CENY AKCJI

Do opisu ewolucji ceny akcji posłuży proces geometrycznego ruchu Browna postaci:

$$p_t = p_o e^{\sigma w_t + \left(m - \frac{\sigma^2}{2}\right)t}, \; t \geq 0,$$

gdzie:

p_t - cena akcji w chwili t,

p_o - cena akcji w chwili początkowej,

w_t - standardowy ruch Browna,

m, σ^2 - parametry modelu reprezentujące odpowiednio wartość średnią i wariancję (na jednostkę czasu) stopy zmiany ceny akcji.

Proces geometrycznego ruchu Browna ma następujące własności:

- jeżeli p_0 jest wartością dodatnią, to dla każdego $t > 0$ jest $p_t > 0$,
- dla każdego ustalonego $t > 0$ zmienna losowa p_t ma rozkład lognormalny,
- wariancja $Var[p_t] \to \infty$, gdy $t \to \infty$.

Jeżeli trajektoria ruchu Browna startuje od dodatniej wartości, to proces osiąga tylko wartości dodatnie. Fakt ten, jak również to, że dla procesów opartych na rozkładzie Gaussa jest dobrze rozwinięty aparat matematyczny, zadecydował, że właśnie ten proces wielu badaczy rynku przyjmuje do opisu ewolucji cen papierów wartościowych. Wykorzystali go, między innymi, Osborne, Samuelson, których prace w latach pięćdziesiątych dokonały przełomu w matematyce finansowej. Z modelowaniem stochastycznym rynków kapitałowych związane są też takie nazwiska jak: Merton, Blaska, Scholesa.

Proces geometrycznego ruchu Browna jest rozwiązaniem stochastycznego równania różniczkowego.

$$dp_t = p_t[mdt + \sigma dw_t].$$

Więcej na temat ruchu Brona można znaleźć w [Banek 2000], [Luenberger 2003], [Sobczyk 1996], [Weron, Weron 1998].

Proces stochastyczny $\{w_t, t \geq 0\}$, w literaturze spotyka się też oznaczenie $w(t,\omega)$, nazywamy standardowym procesem Wienera lub standardowym ruchem Browna, jeżeli:

1. $w_0 = 0$.
2. Przyrost standardowego procesu Wienera na przedziale o długości Δt ma rozkład normalny z wartością oczekiwaną 0 i wariancją równą długości tego przedziału, czyli dla każdego $t \geq 0$, $\Delta t > 0$ jest $(w_{t+\Delta t} - w_t) \sim N(0, \Delta t)$.
3. Jeżeli $t_1 < t_2 \leq t_3 < t_4$, to zmienne losowe $(w_{t_2} - w_{t_1})$ i $(w_{t_4} - w_{t_3})$ są niezależne.
4. Funkcja $t \to w(t,\omega) \in \Re$ jest z prawdopodobieństwem równym 1 ciągła względem t, czyli proces Wienera ma ciągłe trajektorie.

W dalszej części pracy wykorzystane zostaną poniższe własności procesu Wienera.

Własność 1
Niech $a > 0$ i $b \in R$, wtedy

$$P\left(\max_{t \leq T} w_t \geq bt + a\right) = 1 - F\left(\frac{a + bT}{\sqrt{T}}\right) + e^{-2ab}\left(1 - F\left(\frac{a - bT}{\sqrt{T}}\right)\right),$$

gdzie F jest dystrybuantą standardowego rozkładu normalnego.
Dowód powyższej własności można znaleźć w [Szirajev, Kabanov, Kramkov, Mielnikov 1994].
W szczególności, gdy $b = 0$ otrzymujemy

Własność 2 [Billingsley 1987]

$$P\left(\max_{t \leq T} w_t \geq a\right) = 2\left(1 - F\left(\frac{a}{\sqrt{T}}\right)\right) = 2P(w_T \geq a).$$

W tym miejscu warto przytoczyć wzór Blacka-Scholesa na wycenę opcji na akcje. Rozważmy europejską opcję kupna z ceną wykonania q i terminem wygaśnięcia T. Zakładamy, że akcja, na którą została wystawiona opcja nie daje dywidendy w okresie $[0,T]$, stopa wolna od ryzyka (r) jest stała i ma miejsce kapitalizacja ciągła. Wtedy wzór na cenę opcji (cena ta nazywana jest też premią) w chwili $t \in [0,T]$ jest następujący [Luenberger 2003]:

$$c_t^E = p_t F(d_1) - qe^{-r(T-t)} F(d_2),$$

gdzie:

c_t^E - cena europejskiej opcji kupna w chwili t,

p_t - cena akcji w chwili t,

F - dystrybuanta standardowego rozkładu normalnego,

$$d_1 = \frac{\ln \frac{p_t}{q} + \left(r + \frac{\sigma^2}{2} \right)(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = \frac{\ln \frac{p_t}{q} + \left(r - \frac{\sigma^2}{2} \right)(T-t)}{\sigma\sqrt{T-t}}.$$

Znając cenę europejskiej opcji kupna (c_t^E) cenę europejskiej opcji sprzedaży (s_t^E) można wyznaczyć z parytetu kupna-sprzedaży [Weron, Weron 1998]:

$$c_t^E - s_t^E = p_t - qe^{-r(T-t)}.$$

W [Weron, Weron 1998] pokazano, że w przypadku opcji amerykańskich prawdziwe są nierówności:

$$p_t - q < c_t^A - s_t^A \leq p_t - qe^{-r(T-t)},$$

gdzie c_t^A, s_t^A oznaczają odpowiednio ceny amerykańskich opcji kupna i sprzedaży akcji.

WYBRANE KRYTERIA INWESTOWANIA W OPCJE NA AKCJE

Sformułowanych zostanie kilka problemów związanych z inwestowaniem w opcje na akcje. Oszacujemy prawdopodobieństwa zdarzeń osiągnięcia przez inwestora zysku na żądanym poziomie lub przy ustalonym poziomie ryzyka.

Zadanie 1 – ustalony poziom zysku, europejska opcja kupna

Inwestor posiada europejską opcję kupna akcji z terminem wygaśnięcia T i ceną rozliczenia opcji q. Obliczmy

$$P\left(e^{-rT}(p_T - q) \geq z\right).$$

Jest to prawdopodobieństwo tego, że rozliczając opcje w chwili T inwestor osiągnie zysk co najmniej z. Wielkość zysku jest zdyskontowana na chwilę

obecną ze stopą kapitalizacji ciągłej r. Ustalając wartość z inwestor powinien uwzględnić poniesione koszty, takie jak: koszty transakcji, cenę opcji.

W przypadku europejskiej opcji sprzedaży należałoby rozważyć prawdopodobieństwo $P\left(e^{-rT}\left(q-p_T\right)\geq z\right)$. Zajmijmy się przypadkiem opcji kupna.

Mamy

$$P\left(e^{-rT}\left(p_T-q\right)\geq z\right)=P\left(e^{-rT}\left(p_o e^{\sigma w_T+\left(m-\frac{\sigma^2}{2}\right)T}-q\right)\geq z\right)=$$

$$=P\left(w_T\geq\frac{1}{\sigma}\left[\ln\left(\frac{ze^{rT}+q}{p_o}\right)-\left(m-\frac{\sigma^2}{2}\right)T\right]\right)=1-F\left(C_T\right),$$

gdzie

$$C_T=\frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{ze^{rT}+q}{p_o}\right)-\left(m-\frac{\sigma^2}{2}\right)T\right],$$

F jest dystrybuantą standardowego rozkładu normalnego.

Przykład

Niech początkowa cena akcji będzie równa 50 zł. Inwestor nabywa europejską opcję kupna tej akcji z ceną wykonania 50 zł i terminem wygaśnięcia za rok. Stopa wolna od ryzyka dla tego okresu to 6%.

Zatem $p_o=50$, $q=50$, $r=0{,}06$, $T=1$.

Załóżmy ponad to, że stopa zmiany ceny akcji i odchylenie standardowe tej stopy są równe odpowiednio $m=0{,}08$, $\sigma=0{,}1$.

Korzystając ze wzoru Blacka-Scholesa otrzymujemy, że cena opcji wynosi $c^E=3{,}73$ zł.

Przyjmijmy $z=5$. Obliczymy, zatem prawdopodobieństwo tego, że za rok inwestor wykona posiadaną opcję z ceną wykonania 50 zł i natychmiast sprzedając akcję po aktualnej cenie rynkowej osiągnie zysk zdyskontowany na chwilę obecną na poziomie co najmniej 5 zł. Rozważane prawdopodobieństwo ma wartość

$$P\left(e^{-rT}\left(p_T-q\right)\geq z\right)=0{,}398$$

Zadanie 2 - ustalony poziom ryzyka, europejska opcja kupna

Inwestor nieskłonny do dużego ryzyka może ustalić jego maksymalną wartość i szukać takiej wielkości zysku, która może być zrealizowana z prawdopodobieństwem nie mniejszym niż ustalił. Zadanie można sformułować w następującej postaci.

Niech $\rho \in [0,1]$ będzie ustaloną przez inwestora liczbą, którą będziemy nazywać poziomem ryzyka. Należy wyznaczyć taką wartość z_ρ, że

$$z_\rho = \max\left\{z : P\left(e^{-rT}\left(p_T - q\right) \geq z\right) \geq 1 - \rho\right\}.$$

Powyższa wartość jest największym poziomem zysku, jaki można osiągnąć przy ustalonym poziomie ryzyka ρ.

Z poprzedniego zadania wiemy, że

$$P\left(e^{-rT}\left(p_T - q\right) \geq z\right) = 1 - F\left(C_T\right),$$

gdzie

$$C_T = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{ze^{rT} + q}{p_o}\right) - \left(m - \frac{\sigma^2}{2}\right)T\right].$$

Zatem

$$1 - F\left(C_T\right) \geq 1 - \rho$$

lub równoważnie $C_T \leq F^{-1}\left(\rho\right)$.

Mamy

$$\frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{ze^{rT} + q}{p_o}\right) - \left(m - \frac{\sigma^2}{2}\right)T\right] \leq F^{-1}\left(\rho\right),$$

więc

$$z \leq p_o e^{\sigma\sqrt{T}F^{-1}(\rho) + \left(m - \frac{\sigma^2}{2} - r\right)T} - qe^{-rT}.$$

Ostatecznie otrzymujemy

$$z_\rho = p_o e^{\sigma\sqrt{T}F^{-1}(\rho) + \left(m - \frac{\sigma^2}{2} - r\right)T} - qe^{-rT}.$$

Przykład

Dla wartości takich jak w poprzednim przykładzie: $p_o = 50$, $q = 50$, $r = 0{,}06$, $T = 1$, $m = 0{,}08$, $\sigma = 0{,}1$ przyjmijmy, że inwestor ustalił poziom ryzyka $\rho = 0{,}3$. z_ρ jest zatem maksymalną wartością (zdyskontowaną) zysku jaki można osiągnąć z prawdopodobieństwem równym co najmniej 0,7. Z powyższego wzoru otrzymujemy $z_\rho = 1{,}07$.

Zadanie 3 – ustalony poziom zysku, amerykańska opcja kupna

Rozważmy amerykańską opcję kupna akcji z terminem wygaśnięcia T i ceną wykonania opcji q. Oszacujemy następujące prawdopodobieństwo:

$$P\left(\max_{o\le t\le T} e^{-rt}\left(p_t - q\right)\ge z\right).$$

Jest to prawdopodobieństwo tego, że do terminu wygaśnięcia opcji inwestor będzie mógł wykonać opcję i osiągnie zysk zdyskontowany na chwilę obecną na poziomie co najmniej z.
Mamy

$$P\left(\max_{o\le t\le T} e^{-rt}\left(p_t - q\right)\ge z\right) = P\left(\max_{o\le t\le T} w_t \ge \frac{1}{\sigma}\left[\ln\left(\frac{ze^{rT}+q}{p_o}\right) - \left(m - \frac{\sigma^2}{2}\right)t\right]\right)$$

Oznaczmy

$$B(t) = \frac{1}{\sigma}\left[\ln\left(\frac{ze^{rt}+q}{p_o}\right) - \left(m - \frac{\sigma^2}{2}\right)t\right]$$

i niech $B = \min_{o\le t\le T} B(t)$

Prawdopodobieństwo, że trajektoria procesu Wienera „dojdzie" do krzywej $B(t)$ jest nie większe niż prawdopodobieństwo, że trajektoria procesu Wienera „dojdzie" do B. Stąd i z własności 2 otrzymujemy

$$P\left(\max_{o\le t\le T} w_t \ge B(t)\right) \le P\left(\max_{o\le t\le T} w_t \ge B\right) = 2P\left(w_T \ge B\right) = 2\left(1 - F\left(\frac{B}{\sqrt{T}}\right)\right),$$

gdzie F jest dystrybuantą standardowego rozkładu normalnego.

Łatwo sprawdzić, że

$$B''(t) = \frac{qzr^2e^{rt}}{\sigma\left(ze^{rt}+q\right)^2},$$

czyli, dla każdego $t \ge 0$ jest $B''(t) > 0$.
Druga pochodna funkcji $B(t)$ jest dodatnia, a więc funkcja jest wypukła. Jej wykres na przedziale $[0,T]$ znajduje się pod prostą zawierającą punkty $(0,B(0))$ i $(T,B(T))$. Przyjmijmy, że prosta ta ma równanie $l(t) = a + bt$.
Wartość współczynnika a to $B(0)$, czyli

$$a = \frac{1}{\sigma}\ln\frac{z+q}{p_o}.$$

Natomiast wartość współczynnika b wyliczymy z równania $bT + a = B(T)$. Mamy

$$bT + \frac{1}{\sigma}\ln\frac{z+q}{p_o} = \frac{1}{\sigma}\left[\ln\left(\frac{ze^{rT}+q}{p_o}\right) - \left(m - \frac{\sigma^2}{2}\right)T\right],$$

co daje

$$b = \frac{1}{\sigma T}\left(\ln\frac{ze^{rT}+q}{z+q} - \left(m - \frac{\sigma^2}{2}\right)T\right).$$

Prawdopodobieństwo, że na przedziale $[0,T]$ trajektoria ruchu Browna „dojdzie" do krzywej $B(t)$ jest, zatem większe niż prawdopodobieństwo, że trajektoria ruchu Browna „dojdzie" do prostej $l(t)$. Korzystając z własności 1 mamy

$$P\left(\max_{o\le t\le T} e^{-rt}(p_t - q) \ge z\right) = P\left(\max_{o\le t\le T} w_t \ge B(t)\right) \ge P\left(\max_{o\le t\le T} w_t \ge a + bt\right) =$$

$$= 1 - F\left(\frac{a+bT}{\sqrt{T}}\right) + e^{-2ab}\left(1 - F\left(\frac{a-bT}{\sqrt{T}}\right)\right).$$

Oczywiście, musi być $a > 0$, czyli $\frac{1}{\sigma}\ln\frac{z+q}{p_o} > 0$ więc $z + q > p_o$.

Gdyby było inaczej, $z + q \le p_o$, oznaczałoby to, że już w chwili zerowej inwestor rozliczając opcję i natychmiast sprzedając akcje osiągnąłby wymagane z.
Ostatecznie otrzymaliśmy

$$1 - F\left(\frac{a+bT}{\sqrt{T}}\right) + e^{-2ab}\left(1 - F\left(\frac{a-bT}{\sqrt{T}}\right)\right) \le P\left(\max_{o\le t\le T} e^{-rt}(p_t - q) \ge z\right) \le 2\left(1 - F\left(\frac{B}{\sqrt{T}}\right)\right)$$

ze wszystkimi oznaczeniami jak wcześniej.

Przykład
Inwestor posiada amerykańską opcję kupna z terminem wygaśnięcia jeden rok, z ceną rozliczenia 50 zł, taką samą jak obecna cena akcji. Mamy, zatem $p_o = 50$, $q = 50$, $T = 1$. Załóżmy ponad to, że stopa zmiany ceny akcji i odchylenie standardowe tej stopy są równe odpowiednio $m = 0{,}08$, $\sigma = 0{,}1$, stopa wolna od ryzyka $r = 0{,}06$. Inwestor ustala $z = 5$.

Oszacujemy, zatem prawdopodobieństwo tego, że w ciągu roku inwestor będzie mógł wykonać opcję i osiągnie zysk zdyskontowany na chwilę obecną na poziomie co najmniej 5 zł. Dla powyższych wartości otrzymujemy

$$0{,}58 \le P\left(\max_{o\le t\le T} e^{-rt}(p_t - q) \ge z\right) \le 0{,}79.$$

ZAKOŃCZENIE

Opcje dają możliwość tworzenia różnych strategii inwestycyjnych na wypadek różnych scenariuszy rozwoju wydarzeń na giełdzie. Zaprezentowane kryteria mają zastosowanie do oceny inwestycji w opcje na akcje. Mogą one być pomocne inwestorowi przy ocenie wartości prawdopodobieństw osiągnięcia oczekiwanych zysków lub szacowaniu przyszłych zysków przy ustalonym poziomie ryzyka. Pokazano, w jaki sposób szacować te prawdopodobieństwa zarówno dla opcji amerykańskich jak i europejskich.

BIBLIOGRAFIA

Banek T. (2000) Rachunek ryzyka, Centrum Badawczo-Szkoleniowe WSZiA w Zamościu, Lublin.

Billingsley P. (1987) Prawdopodobieństwo i miara, PWN, Warszawa.

Luenberger D., G. (2003) Teoria inwestycji finansowych, Wydawnictwo Naukowe PWN, Warszawa.

Sobczyk K. (1996) Stochastyczne równania różniczkowe, Wydawnictwo Naukowo-Techniczne, Warszawa.

Szirajev A.N., Kabanov J.M., Kramkov D.O., Mielnikov A.B., (1994) K teorii raszczotov opcionov Evropejskogo i Amerikanskogo tipov, Nepreryvnoje vremia - Teoria verojat. i promen., Tom 39.

Weron A., Weron R., (1998) Inżynieria finansowa, Wydawnictwo Naukowo-Techniczne, Warszawa.

SOME CRITERIA OF INVESTMENT IN STOCKS OPTIONS

Abstract: This article describes the stock options. It explains the basic concepts, such as settlement date, expiration date, strike price and premium. To describe the evolution of share prices we used the geometrical Brownian motionWe presented the several criteria for investment in options and shares and then obtained the exercises of stochastic programming. Using the properties of Brownian motion, we explained how to estimate the probability of achieving the profit of desired amount or on fixed level of risk. For each of these criteria we presented the sample calculations

Keywords: stock options, Brownian motion, geometrical Brownian motion

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 253 – 261

ESTYMACJA KRZYWEJ DOCHODOWOŚCI STÓP PROCENTOWYCH DLA POLSKI

Adam Waszkowski
Katedra Ekonomiki Rolnictwa i Międzynarodowych Stosunków Gospodarczych
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: adam_waszkowski@sggw.pl

Streszczenie: W artykule podjęty został problem estymacji krzywej dochodowości dla Polski. Przedstawiono w nim dwie metody najczęściej stosowane przez banki centralne, które publikują takiego typu dane- metodę Nelsona- Siegela oraz Svenssona. Rozważono również możliwość stosowania tych metod w polskich warunkach a w dalszej konsekwencji przedstawiono wyniki estymacji struktury terminowej stóp procentowych. Porównano również oszacowania długookresowej oraz krótkookresowej stopy procentowej obiema metodami w latach 2001- 2012 w Polsce.

Słowa kluczowe: krzywa dochodowości, model Nelsona- Siegela, model Svenssona

WSTĘP

W artykule podjęto próbę oszacowania struktury terminowej (krzywej dochodowości, ang. yield curve) stóp procentowych dla Polski. W krajach o dojrzałym rynku kapitałowym struktura terminowa szacowana jest przez banki centralne. Jak dotąd jednak Narodowy Bank Polski nie publikuje takich projekcji. Artykuł ma być próbą sprawdzenia, czy estymacja krzywej dochodowości jest również możliwa w polskich warunkach.

Artykuł składa się z pięciu części. Część druga zawiera definicję krzywej dochodowości oraz zarys wybranych pozycji literatury dotyczących estymacji struktury stóp procentowych. W części trzeciej omówiono najpopularniejsze modele estymacji krzywej: model Nelsona- Siegla oraz metodę Svenssona. Część czwarta przedstawia wyniki estymacji dla Polski dla wybranych okresów oraz oszacowania krótko- i długookresowej stopy procentowej. W części ostatniej zawarto podsumowanie oraz wnioski płynące z przeprowadzonych badań.

STRUKTURA TERMINOWA STÓP PROCENTOWYCH

Krzywa dochodowości przedstawia [Kliber, 2009] zależność między terminem wykupu określonego instrumentu wolnego od ryzyka (np. bonów skarbowych czy obligacji rządowych) a stopą procentową. Opisuje ona stopy zwrotu wolne od ryzyka dla różnych terminów inwestycji. Rozważmy bieżącą cenę obligacji zerokuponowej w terminie wykupu t z wykorzystaniem dyskonta ciągłego. Jest ona równa:

$$P = Ce^{-tr(t)}, \tag{1}$$

gdzie:

C- kwota do wypłaty posiadaczowi obligacji w chwili t,

r(t)- stopa procentowa za okres od chwili obecnej do chwili t, liczona przy kapitalizacji ciągłej ($r(t) = \ln(1 + i(t))$.

Wynika z tego, że:

$$r(t) = \frac{\ln C - \ln P}{t}. \tag{2}$$

Stopy otrzymane z powyższego wzoru to stopy rynku spot, zaś funkcja r(t) określa ich strukturę terminową. Z kolei stopę forward można wyznaczyć korzystając z kontraktów FRA (ang. Forward Rate Agreement). Niech f(s,t) oznacza stopę forward na okres 0 < s < t. Stopa forward jako stopa procentowa w kontrakcie FRA na pożyczkę od s do t wynosi zatem:

$$f(t) = \frac{tr(t) - sr(s)}{t - s}. \tag{3}$$

Chwilowa stopa forward określana jako stopa forward dla nieskończenie krótkiego okresu w przyszłości zaczynającego się w okresie t wynosi zatem:

$$f(t) = \lim_{s \to t} f(s,t) = r(t) + tr'(t). \tag{4}$$

Powyższa funkcja f(t) nazywana jest struktura terminową stóp forward. Z powyższego wzoru (4) wynika zależność między strukturą terminową cen spot i forward:

$$r(t) = \frac{1}{t}\int_0^t f(u)du. \tag{5}$$

Znajomość struktury terminowej stóp procentowych jest przydatna, ponieważ obserwowane ceny obligacji najczęściej różnią się od cen fundamentalnych („prawdziwych”) z powodu obecności szumu rynkowego, gdzie obserwowaną cenę obligacji P^{obs} można przedstawić jako:

$$P^{obs} = P + \zeta \tag{6}$$

gdzie: P - cena fundamentalna, ζ - zakłócenie losowe (szum). Szacując stopy procentowe dla różnych okresów próbie poddaje się eliminacje wpływu szumu. Ponadto w wielu modelach dla rynku kapitałowego zakładana jest znajomość funkcji struktury terminowej. Jest ona niezbędna m.in. do wyceny instrumentów pochodnych[1], prognozowania zmienności cen obligacji czy w modelach dynamiki stóp procentowych[2].

Problematyka szacowania krzywej dochodowości była podejmowana w kilku pracach. Metodologia jej estymacji została szeroko przedstawiona w opracowaniu Marciniaka [2006]. W badaniach empirycznych wykorzystane zostały funkcje sklejane oraz metoda Svenssona. Tematyka ta zawarta jest również w artykule Gurazdowskiego [2003]. Stamirowski [2003] przedstawił empiryczne wyniki szacowania krzywej dochodowości dla Polski, USA oraz strefy euro. Wykorzystał również w tym celu model Vasicka oraz Coxa-Ingersolla-Rossa.

MODELE KRZYWEJ DOCHODOWOŚCI

Krzywa dochodowości powinna posiadać pewne właściwości [Nelson i Siegel 1987]. Przede wszystkim funkcja f powinna być na tyle elastyczna, aby odzwierciedlać występujące w rzeczywistości różne kształty krzywych dochodowości. Warunkiem koniecznym aby otrzymać strukturę terminową stóp forward jest różniczkowalność struktury stóp spot. Dodatkowo funkcja f powinna mieć granice w nieskończoności a jej wartości powinny być dodatnie i ograniczone z góry.

Istnieje wiele metod estymacji struktury stóp procentowej. W niniejszym artykule uwaga jednak zostanie zwrócona tylko na dwie, stosowane powszechnie przez banki centralne.[3] Obie te metody mają swoje wady jak i zalety. Modele te bazują na podstawach teoretycznych, co powoduje naturalna interpretację oszacowań parametrów w kategoriach ekonomicznych. Dużym problemem jest jednak sama estymacja modeli. Najczęściej stosuje się nieliniową metodę najmniejszych kwadratów. W praktyce sprowadza się to do rozwiązania nieliniowego zadania minimalizacji. Używa się do tego metod numerycznych, ale ze względu na złożoność postaci funkcyjnej nie ma gwarancji znalezienia minimum globalnego.

Metoda Nelsona- Siegla

W modelu Nelsona- Siegela [1987] punktem wyjścia jest określenie równania stóp procentowych forward postaci:

[1] Przykładowo wycena instrumentów *floor* lub *cap.*
[2] Model Heatha- Jarrowa- Mortona (HJM).
[3] W praktyce banków centralnych wykorzystywana jest dodatkowo metoda funkcji sklejanych (ang. *splines*), ale nie jest to metoda estymacji lecz interpolacji.

$$f(t) = \beta_1 + \beta_2 \exp\left(\frac{-t}{\lambda}\right) + \beta_3 \frac{t}{\lambda} \exp(\frac{-t}{\lambda}), \tag{7}$$

gdzie: β_1, β_2, β_3, λ to parametry. Zgodnie zatem ze wzorem (5) struktura terminową stóp spot przedstawia relacja:

$$r(t) = \beta_1 + (\beta_2 + \beta_3)\lambda \frac{1 - \exp(-t/\lambda)}{t} - \beta_3 \exp(\frac{-t}{\lambda}). \tag{8}$$

Każdy z elementów równania znajdujący się po prawej stronie ma swoją interpretacje. Stała β_1 odzwierciedla długookresowy poziom stopy procentowej[4]: $\lim_{t\to\infty} r(t) = \beta_1$. Parametr β_2 opisuje wpływ krótkookresowych czynników na stopę terminową: $\lim_{t\to 0} r(t) = \beta_1 + \beta_2$. Ostatni, trzeci składnik, odpowiedzialny jest za średniookresową zmienność stóp terminowych i pozwala uzyskać „zgarbioną" postać krzywej dochodowości- dla wartości dodatniej β_3 - jest to maksimum, dla ujemnej- minimum. Parametry modelu można uzyskać szacując je np. nieliniową MNK[5].

Metoda Svenssona

Svensson [1994] zaproponował rozbudowę modelu Nelsona- Siegla poprzez dodanie kolejnego składnika opisującego zmienność stopy terminowej. W ten sposób zwiększył elastyczność i poprawił dopasowanie funkcji struktury terminowej stóp forward, uzyskując w ten sposób równanie:

$$f(t) = \beta_1 + \beta_2 \exp\frac{-t}{\lambda_1} + \beta_3 \frac{1}{\lambda_1} \exp\frac{-t}{\lambda_1} + \beta_4 \frac{t}{\lambda_2} \exp\frac{-t}{\lambda_2}. \tag{9}$$

Podstawiając do równania (5) uzyskujemy strukturę terminową stóp spot:

$$r(t) = \beta_1 + \beta_2\lambda_1 \frac{1 - \exp(-t/\lambda_1)}{t} + \beta_3\left(\frac{(1 - \exp(-t/\lambda_1))\lambda_1}{t} - \exp(\frac{-t}{\lambda_1})\right) +$$

$$\beta_4\left(\frac{(1 - \exp(-t/\lambda_2))\lambda_2}{t} - \exp(\frac{-t}{\beta_2})\right). \tag{10}$$

[4] Oprocentowanie obligacji konsolowej.

[5] Wadą takiego podejścia jest fakt, że funkcja celu (minimalizacja sumy kwadratów różnic między cenami teoretycznymi obligacji a cenami rynkowymi) jest funkcją silnie nieliniową i do rozwiązania zadania należy zastosować metody numeryczne poszukiwania minimum.

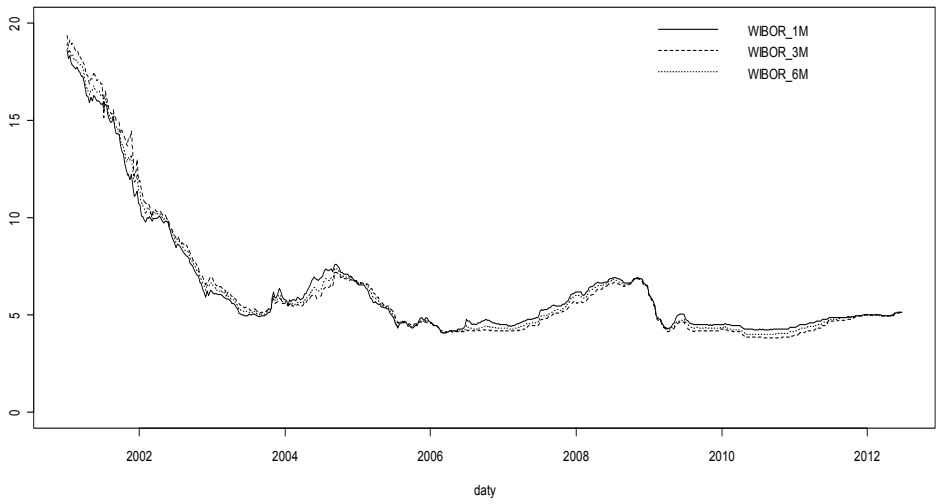
Wprowadzone przez Svenssona rozszerzenie powoduje, że w modelu tym jest możliwe do otrzymania dwóch „garbów". Parametry λ_1 i λ_2 odpowiedzialne są z kolei za ich rozmieszczenie.

WYNIKI ESTYMACJI

W tym punkcie przedstawione zostaną wyniki estymacji powyższych modeli krzywej dochodowości stóp procentowych dla polskiej gospodarki. W tym celu oszacowane zostały parametry modeli postaci (8) oraz (10) oraz szeregi dla stóp długookresowych oraz krótkookresowych. W analizie[6] wykorzystano dane tygodniowe dotyczące walorów WIBOR 1M, 3M, oraz 6M z okresu 5.01.2001 do 22.VI.2012 zaczerpnięte z portalu stooq.pl. Łącznie w obliczeniach wykorzystane zostały szeregi czasowe składające się z 599 obserwacji.

Wykres 1 pokazujący szeregi czasowe stopy WIBOR wskazuje, że w badanym okresie nastąpił istotny spadek oprocentowania w Polsce, z poziomu blisko 20% w roku 2001 do prawie 5% w roku bieżącym.

Rysunek 1. Stopy WIBOR



Źródło: obliczenia własne (wykres programu R)

Aby przedstawić różnicę między metodami estymacji struktury terminowej oszacowano parametry modeli dla pierwszego oraz ostatniego momentu analizowanego szeregu czasowego. Uzyskano następujące oszacowania:

[6] Analizę przeprowadzono korzystając z programu R.

Model Nelsona- Siegela

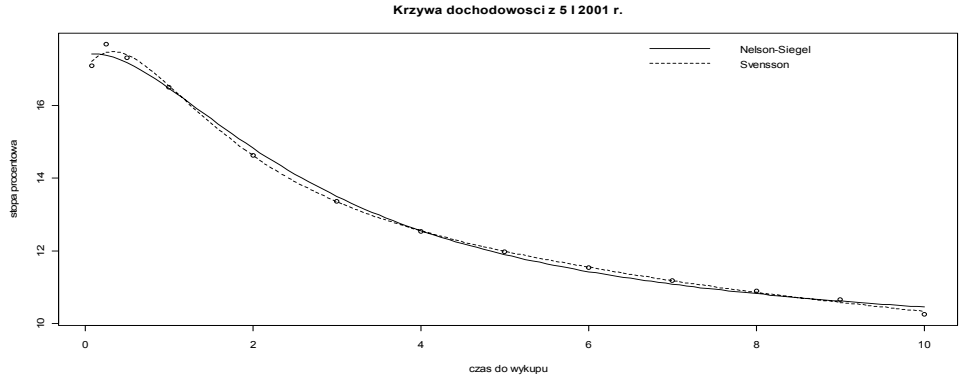
data	beta_1	beta_2	beta_3	lambda
05.01.2001	8.984	8.410	9.189	0.099
22.06.2012	9.534	7.291	9.617	0.098

Model Svenssona

Data	beta_1	beta_2	beta_3	beta_4	lambda_1	lambda_2
05.01.2001	7.237	9.715	12.422	7.654	6.691	33.458
22.06.2012	2.393	2.732	1.974	10.237	3.624	66.916

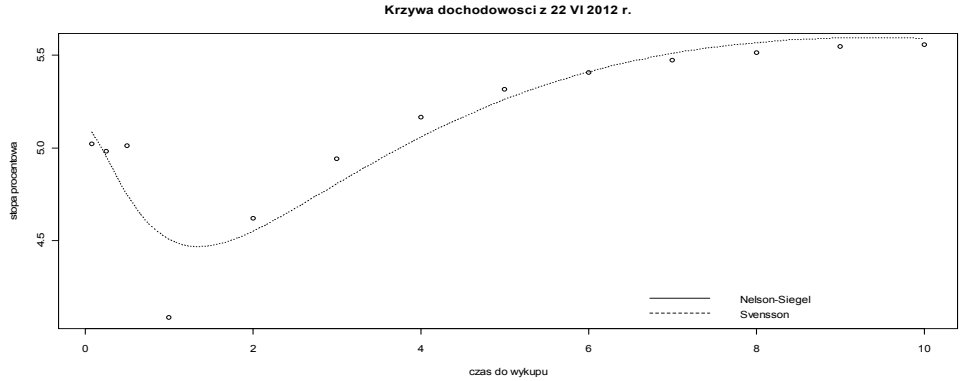
Wyniki estymacji wskazują, że oszacowania parametrów β_1 oraz β_2 dla tych samych momentów są różne mimo jednakowej ich ekonomicznej interpretacji w przypadku obu modeli. Współczynniki te jednak nie prowadzą do innych kształtów krzywej dochodowości. Rysunki 2 i 3 wskazuje, że w obu przypadkach dopasowanie do danych jest poprawne.

Rysunek 2. Krzywe dochodowości dla 5.01.2001



Źródło: obliczenia własne (wykres programu R)

Rysunek 3. Krzywe dochodowości dla 22.06.2012



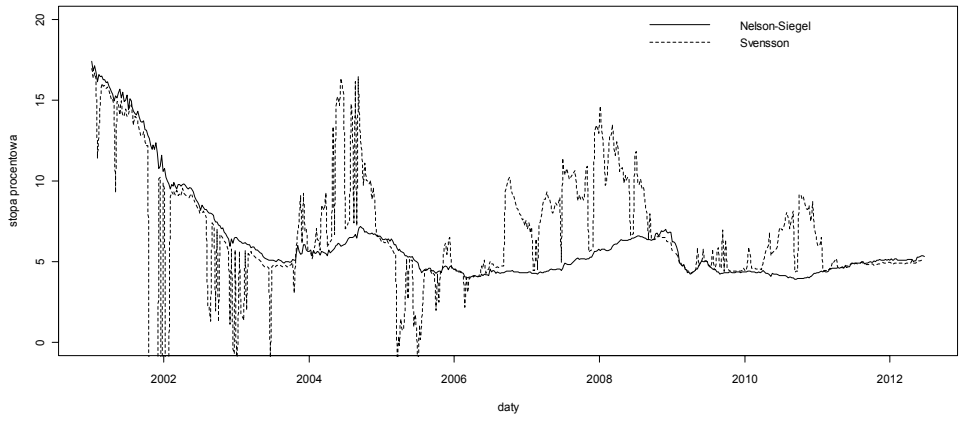
Źródło: obliczenia własne (wykres programu R)

Krzywa dochodowości ze stycznia 2001 roku (rysunek 2) wskazuje na ujemną zależność między stopą procentową a terminem zapadalności. Określa się ją w literaturze przedmiotu jako krzywą odwróconą (ang. inverted yield curve). Zgodnie z teorią oczekiwań[7] wskazuje ona na spadek w krótkim okresie stóp procentowych. Jest to zgodne z szeregiem czasowym stóp procentowych zaprezentowanym na rysunku 1. Z kolei zależność między zmiennymi dla krzywej z czerwca 2012 (rysunek 3) jest ujemna dla pierwszego okresu. Od drugiego okresu można mówić o normalnym kształcie krzywej dochodowości. W takim przypadku krzywa dochodowości z punktem przegięcia określana jest jako S - kształtna (ang. S- shaped).

W dalszej części analizie poddano kształtowanie się stóp procentowych w całym badanym okresie. Dla każdego momentu szeregu czasowego oszacowano parametry β_1 oraz β_2 dla modeli postaci (8) i (10) oraz stworzono szeregi czasowe stóp krótkookresowych (rysunek 4) oraz długookresowych (rysunek 5).

Wykres dla krótkookresowej stopy procentowej wskazuje na spadek z poziomu 18% w roku 2001 do blisko 5% w roku 2012. Dodatkowo krzywa oszacowana metodą Svenssona wykazuje duże wahania wokół tendencji spadkowej, która zarysowana jest równaniem krzywej Nelsona- Siegela. Od roku 2000 oszacowania metodą Svensona znajdują się regularnie powyżej wartości uzyskanych metodą Nelsona- Siegela.

Rysunek 4. Oszacowania wartości krótkookresowej stopy procentowej

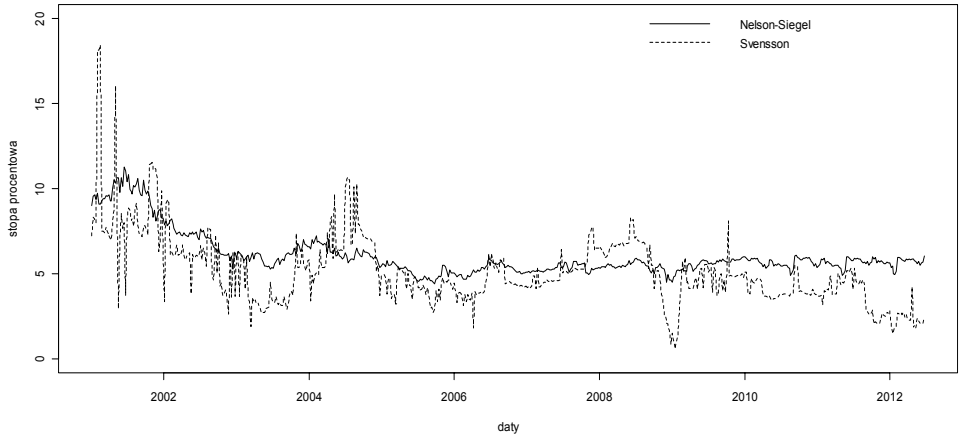


Źródło: obliczenia własne (wykres programu R)

[7] Teoria oczekiwań jest podstawową teorią opisującą kształt krzywej dochodowości. Wskazuje ona, że głównym determinantem kształtu krzywej są oczekiwania dotyczące przyszłego poziomu stóp procentowych.

W przypadku stopy długookresowej różnice w uzyskanych oszacowaniach obiema metodami nie są aż tak wyraźne jak w przypadku stopy krótkookresowej. Metoda Svenssona charakteryzuje się większymi wahaniami i mimo podobnego trendu spadkowego obie te metody estymacji dają istotnie różne oszacowania stopy długookresowej.

Rysunek 5. Oszacowania wartości długookresowej stopy procentowej



Źródło: obliczenia własne (wykres programu R)

PODSUMOWANIE I WNIOSKI

W pracy podjęto próbę estymacji krzywej dochodowości dla gospodarki Polski. W tym celu wykorzystano metody zaproponowano przez Nelsona- Siegela oraz Svenssona. Z przedstawionych badań wynika, że estymacja struktury terminowej stóp procentowych dla Polski nie jest zadaniem łatwym. Co więcej, stosowane w praktyce banków centralnych metody dają różne oszacowania parametrów, które podlegają takim samym interpretacjom ekonomicznym. Mimo to kształty uzyskane obiema metodami są podobne. Problematycznym jest fakt, że oszacowania stóp długo- oraz krótkookresowych daje różne wyniki- metoda Svenssona charakteryzuje się większymi wahnięciami względem wspólnej tendencji. Utrudnione zatem staje się wnioskowanie na temat przyszłego poziomu stopy procentowej. Zarówno wycena kontraktów na przyszłą wartość stopy procentowej jak i konstrukcja przyszłej ścieżki stóp procentowych na podstawie otrzymanych wyników może prowadzić do znacznych błędów.

BIBLIOGRAFIA

Gurazdowski E. (2003) Wykorzystanie modelu zmiennej sztywności krzywej stóp terminowych do przybliżania krzywej rynku pieniężnego, Bank i Kredyt, nr 2, str. 87 - 92.

Kliber P. (2009) Estymacja struktury terminowej stóp procentowych w Polsce, Bank i Kredyt, nr 40(1), str. 109 - 126.

Marciniak M. (2006) Yield Curve Estimation at the National Bank of Poland, Bank i Kredyt, nr 10, str. 52 - 74.

Nelson C. R., Siegel A. E. (1987) Pasimonious Modeling of Yield Curves, Jurnal of Business, Nr 60, str. 473 - 489.

Stamirowski M. (2003) Jednoczynnikowe modele Vasicka oraz CIR- analiza empiryczna na podstawie danych z polskiego rynku obligacji skarbowych, Bank I Kredyt, nr 7, str. 35 - 46.

Svensson L. E. (1994) Estimating and interpreting forward interest rates: Sweden 1992 - 1994, Working Paper, Nr 4871 NBER, Cambridge, str. 3 - 50.

ESTIMATION OF THE YIELD CURVE OF INTEREST RATES IN POLAND

Abstract: Tha aim of this article is the estimation of the yield curve of interest rates in Poland. Two methods applied at central banks publishing this type of data have been presented- Nelson-Siegel method and Svensson method. The article reports the results of applying these methods for Poland subsequently.

Key words: yield curve, Nelson-Siegel model, Svensson

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 262 – 275

BUDOWA MIERNIKÓW SYNTETYCZNYCH DO OCENY EFEKTYWNOŚCI EUROPEJSKICH FUNDUSZY INWESTYCYJNYCH[1]

Dorota Witkowska
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: dorota_witkowska@sggw.pl

Streszczenie: Celem prowadzonych badań jest konstrukcja agregatowych mierników, które umożliwią kompleksową ocenę efektywności funduszy inwestycyjnych działających na paneuropejskim rynku. Do budowy mierników wykorzystano m.in. współczynniki Sharpe'a, Treynora i Jensena. Badania empiryczne dotyczyły 18 funduszy inwestycyjnych, 4 indeksów giełdowych (Euro STOXX 50, DAX, CAC 40 i FTSE 100) oraz 6 przedziałów czasowych zawartych w okresie od 15.09.2006 r. do 20.01.2012 r. Wykazano, że zastosowanie mierników syntetycznych pozwala na jednoznaczną ocenę funduszy inwestycyjnych.

Słowa kluczowe: miernik syntetyczny, efektywność inwestycyjna, wskaźniki Sharpe'a, Treynora i Jensena

WPROWADZENIE

Fundusze inwestycyjne będące, poprzez połączenie środków finansowych inwestorów, instytucjami zbiorowego inwestowania, mają już niemal 150-letnią historię, bowiem ich rozwój został zapoczątkowany powstaniem w 1863r. w Wielkiej Brytanii Szkocko-Amerykańskiego Trustu Inwestycyjnego [Czekaj 2008, s. 54]. W końcu 2011r. aktywa funduszy inwestycyjnych na świecie wyniosły 23,8 tryliona USD, z czego 49% to aktywa funduszy amerykańskich, a 30% - funduszy europejskich. Zatem rynek funduszy inwestycyjnych jest

[1] Badania realizowane w ramach projektu MNiSW „Efektywność rynku otwartych funduszy inwestycyjnych w warunkach zmiennej koniunktury rynkowej" N N111 277638.

ogromny i będzie dalej się rozwijał, zasilany chociażby środkami zgromadzonymi w funduszach emerytalnych.

Wśród szeregu zalet wspólnego inwestowania, wymienia się [Dębski 2005, s. 560 – 562] m.in. profesjonalne zarządzanie portfelem. Klienci funduszy inwestycyjnych oczekują zatem zysku z inwestycji przy zadanym poziomie ryzyka. W budowaniu strategii marketingowych funduszy wykorzystuje się awersję znacznej grupy inwestorów do ryzyka, związanego z samodzielnym inwestowaniem, podsycając jednocześnie zaufanie inwestorów do profesjonalizmu zarządzających i czynnika skali przedsięwzięć instytucji zbiorowego inwestowania.

Autorzy wielu prac: [Ippolito 1992], [Sirri, Tufano 1993], [Patel i in. 1994], [Gruber 1996]; [Goetzmann i Peles 1997], [Edelen 1999], [Bergstresser i Poterba 2002], [Deaves 2004] oraz [Busse i Irvine 2006] wykazali, że osiągane w przeszłości wyniki inwestycyjne stanowią dla inwestorów istotne kryterium wyboru funduszu. Można zatem twierdzić, że świadomość klientów dotycząca szans i zagrożeń związanych z oszczędzaniem w funduszach jest niewielka. Co więcej, wydaje się, że jest ona „na rękę" samym funduszom, które „mamią" potencjalnych klientów jedynie historycznymi lub – jeśli te ostatnie nie były dość atrakcyjne – potencjalnymi stopami zwrotu [Hendricks i in. 1993]. Dlatego decyzje podejmowane przez klientów funduszy inwestycyjnych bywają często nietrafione.

Trwała koniunktura gospodarcza i hossa na rynku kapitałowym, panująca do połowy 2007 roku, istotnie przyczyniły się do rozwoju rynku funduszy inwestycyjnych, co było spowodowane stałym wzrostem wartości jednostek uczestnictwa i certyfikatów inwestycyjnych. Jednakże o ile w okresie dobrej koniunktury najprostsza nawet strategia inwestycyjna, np. „kup i trzymaj" jest skuteczna i generuje wysokie zwroty, o tyle w okresie dekoniunktury efektywne ekonomicznie (choćby w sensie minimalizacji strat) są jedynie te inwestycje, w których czynniki skali i profesjonalizm zostają umiejętnie wykorzystane do zarządzania aktywami. Naturalne w tym kontekście wydaje się więc pytanie o efektywność funduszy inwestycyjnych. Popularnymi narzędziami są wskaźniki, przedstawione w pracach: [Sharpe 1966], [Treynor 1965] i [Jensen 1968 i 1969], które pozwalają ocenić jakość zarządzania funduszami i tworzyć rankingi uwzględniające zarówno stopy zwrotu osiągane przez fundusze, jak i ryzyko[2] podejmowanych inwestycji.

Globalizacja rynków finansowych, w szczególności integracja w ramach Unii Europejskiej, przyczyniły się do ekspansji funduszy inwestycyjnych na rynki zagraniczne. Oznacza to, że zarówno klienci instytucji zbiorowego inwestowania, jak i instrumenty wchodzące w skład portfeli nie ograniczają się do pojedynczych rynków krajowych i dotyczą rynków międzynarodowych.

[2] Ten drugi czynnik towarzyszący inwestycjom jest często ignorowany zarówno przez prospekty informacyjne funduszy inwestycyjnych, jak i publikacje finansowe, dostępne potencjalnym klientom funduszy.

Dlatego celem prowadzonych badań jest budowa agregatowych mierników, które umożliwią kompleksową ocenę efektywności funduszy inwestycyjnych działających na paneuropejskim rynku. Do konstrukcji mierników zastosowano klasyczne wskaźniki efektywności inwestycyjnej, dzięki czemu możliwe jest porównanie osiąganych przez fundusze wyników inwestycyjnych jednocześnie w wielu aspektach.

BUDOWA MIERNIKÓW AGREGATOWYCH

Wymienione wcześniej wskaźniki efektywności Sharpe'a, Treynora i Jensena, wykorzystują w swej konstrukcji tzw. indeks rynku oraz instrument wolny od ryzyka, które stanowią podstawę do oszacowania współczynnika beta oraz porównania efektywności inwestycyjnej badanego instrumentu lub portfela. Innymi słowy przyjmuje się, że podejmowanie ryzyka jest uzasadnione w sytuacjach, kiedy towarzyszyć mu będzie zwiększony (o premię za ryzyko) zysk, a „neutralne" poziomy zysku i ryzyka wyznaczane są przez indeks rynku. Jednakże każda ze wspomnianych miar „dokonuje oceny" pod innym kątem, a co za tym idzie oceny wyznaczone za pomocą różnych miar nie są jednoznaczne, a rankingi funduszy nie pokrywają się, jak np. w pracach [Kompa, Witkowska 2010] i [Foo, Witkowska 2012].

Istotną kwestią pozostaje zatem wybór indeksu - reprezentanta rynku oraz instrumentu wolnego od ryzyka, ponieważ ich wybór ma istotny wpływ na ocenę efektywności inwestycyjnej (por. [Roll 1978], [Reilly, Akhtar 1995], [Grinblatt, Titman 1994]). Problemy z wyborem benchmarku potęgują się w sytuacji, kiedy ocenie podlegają fundusze działające na rynku globalnym. W badaniach przyjmuje się dwa podejścia tzn. albo konstruuje się specjalnie do celów badawczych indeksy rynku jak [Karpio, Żebrowska – Suchodolska 2011], albo stosuje się istniejące na rynku indeksy „globalne" tj. Morgan Stanley Capital International world index, np. [Dutta i in. 2010] lub indeksy giełd krajowych np. CAC opisujący rynek francuski, DAX – niemiecki, Mibtel – włoski, FTSE – brytyjski jak w [Otten, Bams 2002], [Otten, Scchweitzer 2002]. Podobnie jest z wyborem reprezentanta instrumentu wolnego od ryzyka, którym zazwyczaj są bony skarbowe o różnym okresie zapadalności.

W badaniach dotyczących funduszy inwestycyjnych, działających na rynku paneuropejskim, wykorzystano 4 indeksy reprezentujące rynek tj.: Euro STOXX 50 (ES 50) - opisujący rynek kapitałowy krajów strefy euro i wcześniej wspomniane indeksy giełdowe: DAX, CAC 40 oraz FTSE 100. Za instrument wolny od ryzyka przyjęto 10-letnie niemieckie bony skarbowe Bundesrepublik Deutschland-Anleihe: 15-04-2016 (BDA).

W niektórych badaniach np. [Kompa, Witkowska 2010] zauważa się, że wyniki inwestycyjne funduszy zależą od sytuacji rynkowej, czyli są inne w okresie wzrostów i spadków. Dlatego ocenę efektywności funduszy inwestycyjnych prowadzono dla kilku okresów badawczych o różnej długości.

Podjęto zatem próbę konstrukcji mierników agregatowych (por. też [Kompa, Witkowska 2010], [Prince, Bacon 2010], a także [Kosarga 2002]), które umożliwią jednoznaczną ocenę i klasyfikację funduszy inwestycyjnych przy jednoczesnym uwzględnieniu różnych mierników efektywności, benchmarków oraz okresów analizy. Zaproponowano dwie proste miary umożliwiające wielowymiarową analizę. Pierwsza z nich jest nieważoną sumą rang N_{lj} (gdzie N_{lj} - numer określający pozycję *j*-tego funduszu w *l*-tym rankingu):

$$MS_j = \sum_{l=1}^{n} N_{lj} \tag{1}$$

uzyskanych przez każdy z funduszy w wyniku przeprowadzonych rankingów funduszy inwestycyjnych, których podstawą są wyznaczone dla każdego okresu analizy jedno-okresowe stopy zwrotu oraz współczynniki efektywności Sharpe'a, Treynora i Jensena, obliczone dla różnych benchmarków[3].

Druga z miar bazuje na skonstruowanych (w wyróżnionych okresach w oparciu o wyznaczone miary efektywności i zastosowane indeksy) klasach, określających efektywność funduszy inwestycyjnych i jest postaci:

$$CS_j = \sum_{i=1}^{k} w_i p_{ij} \tag{2}$$

gdzie: p_{ij} – liczba przypadków kiedy *j*-ty fundusz został zaklasyfikowany do *i*-tej klasy efektywności, w_i – waga określającą *i*-tą klasę efektywności.

Dla wspomnianych współczynników efektywności oraz jedno-okresowych stóp zwrotu wyznaczono 5 klas efektywności (tj. k = 5). W przypadku współczynników Treynora i Sharpe'a, które pozwalają na ocenę funduszu w porównaniu z rynkiem przyjęto, że do klasy najbardziej efektywnych funduszy należą te, które generują lepsze wyniki niż „najbardziej efektywny" indeks rynku, a do najmniej efektywnych funduszy należą te, które generują wyniki gorsze od rynku „najmniej efektywnego". Podobnie wyłoniono klasy dla strategii „kup i trzymaj", bazując na kwartylach i wynikach indeksów rynku. W pierwszej klasie efektywności znalazły się zatem fundusze inwestycyjne, które generują większe zwroty niż „najlepszy" z indeksów dla (dodatnich) wartości 3. kwartyla (w okresach T2 i H) lub uzyskują dodatnie zwroty w pozostałych przypadkach. Druga klasa zawiera fundusze o stopach zwrotu większych od 3. kwartyla, przy czym dla dodatnich wartości tego ostatniego (w okresach T2 i H) są to fundusze o stopach zwrotu mniejszych od najbardziej efektywnego indeksu rynku. Trzecia klasa to fundusze o zwrotach zawartych w przedziale między medianą a 3. kwartylem, a w czwartej znalazły się fundusze o zwrotach mniejszych od mediany i większych od wartości 1. kwartyla. Ostatnia klasa zawiera najmniej

[3] W przypadku współczynnika Jensena brano pod uwagę jego wartość podzieloną przez współczynnik beta.

efektywne fundusze o zwrotach mniejszych od wartości 1. kartyla. W przypadku alfy Jensena przyjęto, że do klasy najbardziej efektywnych funduszy należy ten o największej wartości miernika, do drugiej klasy zaliczono fundusze zajmujące 2. i 3. pozycję. Podobnie zdefiniowano ostatnie dwie klasy z tym, że chodziło o najmniejsze wartości alfy. Natomiast w klasie trzeciej znalazły się pozostałe fundusze.

Zaproponowane mierniki syntetyczne umożliwią jednoznaczną ocenę i uszeregowanie funduszy, jednocześnie według kilku wykorzystanych kryteriów oceny ich efektywności, co daje inwestorom lepszą orientację przy wyborze funduszu. Istnieje bowiem możliwość wskazania na najbardziej i najmniej efektywne fundusze inwestycyjne niezależnie od okresu analizy, wybranego miernika i indeksu rynku.

WYBÓR FUNDUSZY INWESTYCYJNYCH I OKRESU BADANIA

W badaniach przeprowadzono analizę 18 europejskich funduszy inwestycyjnych, wybranych na podstawie rankingu przeprowadzonego przez Financial Times[4], w którym uwzględniono kilka kryteriów. Po pierwsze selekcję przeprowadzono spośród funduszy inwestujących głównie (tj. minimum 75% aktywów) w akcje najbardziej płynnych przedsiębiorstw z rynku paneuropejskiego. Po drugie przyjęto, że przynajmniej 75% aktywów funduszy stanowią akcje spółek europejskich, ale inwestycje na pojedynczym (krajowym) rynku muszą być mniejsze niż 50%. Po trzecie, wybrane fundusze charakteryzowały się (na początku 2012r.) 3- i 5-letnimi dodatnimi zwrotami, a okres ich działalności wynosił przynajmniej 5 lat. W ostatnim etapie spośród około 200 funduszy inwestycyjnych, które spełniały powyższe kryteria wybrano takie, które miały najwyższą 3-letnią stopę zwrotu (por tab. 1). Dodatkowo przyjęto, że konkretne towarzystwo inwestycyjne może być reprezentowane tylko przez pojedynczy fundusz, zakładając że w ramach jednego towarzystwa prowadzona jest podobna polityka inwestycyjna i zbliżony styl zarządzania.

Warto zauważyć, że wszystkie wyselekcjonowane fundusze (tab. 1) inwestowały w akcje przynajmniej 85,8% swoich aktywów. Spółki, będące przedmiotem inwestycji, najczęściej miały swoją główną siedzibę we Francji i Niemczech (14 funduszy inwestowało w tych krajach średnio po ok. 30% aktywów), w dalszej kolejności znalazły się spółki holenderskie, hiszpańskie i włoskie. Najchętniej inwestowano w spółki finansowe (- 12 funduszy średnio 16,6% aktywów), energię (- 16 funduszy średnio po 8,4% aktywów), a także w telekomunikację, opiekę zdrowotną, przemysł i dobra konsumpcyjne. Można też wymienić kilka najbardziej „popularnych" spółek, których udział wynosił od 2,6 % do 4,8 % aktywów funduszy. Są to Total SA i Sanofi, znajdujące się w portfelach 10 funduszy, Siemens, BASF SA, SAP i Telefonica – w portfelach 8 funduszy.

[4] Wykorzystano w tym celu http://funds.ft.com/uk/Screener

Tabela 1. Lista wybranych funduszy inwestycyjnych oraz uzyskane przez nie stopy zwrotu

Nr	Nazwa funduszu	Procentowe zwroty z funduszy w analizowanym okresie						
		3-letnia	T	T1	T2	BH	B	H
A	Allianz RCM Wachstum Euroland A EUR	18,61	10,69	-10,00	28,04	-8,63	-50,16	91,64
B	AXA Rosenberg Eurobloc Equity Alpha Fund B EUR Acc	4,01	-31,63	-15,06	-16,85	-38,15	-61,45	65,59
C	BlackRock Global Funds – Euro-Markets A2 EUR	12,31	-3,11	-2,47	3,04	-17,18	-51,95	77,12
D	Credit Suisse Aktien Plus A	7,18	-39,42	-22,62	-18,80	-42,56	-64,04	65,67
E	Deka-EuroStocks CF Inc	8,20	-32,41	-12,29	-19,94	-36,97	-64,12	84,78
F	DWS Invest Top Euroland FC	11,73	-8,28	-9,63	4,84	-20,63	-51,73	68,81
G	Fidelity Funds – Euro Blue Chip Fund A-EUR	7,76	-19,40	-7,31	-10,23	-30,20	-55,71	63,92
H	First Private Euro Aktien STAUFER A	7,27	-27,27	-14,86	-12,13	-35,78	-62,20	78,34
I	Henderson Horizon Euroland Fund A2	10,25	-21,99	-14,42	-5,06	-36,29	-56,74	50,68
J	HL MaxInvest FT (FT-Frankfurt Trust)	7,76	-31,25	-16,40	-14,61	-38,14	-60,62	65,31
K	HSBC Global Investment Funds Euroland Equity A EUR	6,15	-17,24	-2,53	-11,76	-25,78	-57,62	80,95
L	Pioneer Investments Euroaktien	9,21	-3,39	-4,73	3,97	-20,01	-48,53	57,52
M	Ring Aktien Fonds DWS	14,10	1,94	0,02	4,96	-18,59	-58,57	99,82
N	Schroder ISF Euro Equity A Acc	6,55	-20,88	-11,36	-7,16	-31,98	-56,94	60,26
O	SEB EuroCompanies	8,54	-29,36	-13,09	-15,29	-31,89	-60,83	80,91
P	UniEuroAktie	7,14	-23,79	-7,45	-14,09	-32,12	-57,34	64,18
R	UniExtra: EuroStoxx-50	8,18	10,69	-10,00	28,04	-8,63	-50,16	91,64
S	VERI-EUROPA	7,17	-19,89	-11,31	-4,63	-26,62	-57,12	78,23
Benchmarki	ES 50	-34,89	-12,40	-22,68	-36,86	-59,35	55,94	-34,89
	DAX	12,31	7,53	7,38	-15,38	-53,93	82,41	12,31
	CAC 40	-34,39	-14,35	-20,39	-36,43	-57,95	52,07	-34,39
	FTSE 100	-2,01	-7,42	10,17	-12,04	-46,43	63,64	-2,01
	BDA	12,00	-1,64	13,08	15,36	9,21	5,28	12,00

Źródło: opracowanie własne na podstawie www.funds.ft.com, www.finanzen.net, [Lachowski 2012]

Badania realizowano dla notowań dziennych w okresie od 15.09.2006 r. do 20.01.2012 r., dzieląc cały okres badania na dwie podpróby datą 15.09.2008 r., kiedy Lehman Brothers Holdings Inc. (LB) ogłosił bankructwo. Dodatkowo

przeprowadzono analizy dla okresu od 2.07.2007 r. do 12.11.2010 r., w którym wydzielono podokres spadków (bessy) od 2.07.2007 r. do 6.03.2009 r. i wzrostów od 9.03.2009 r. do 12.11.2010 r. w oparciu o notowania indeksu Euro STOXX 50. Poszczególne próby zostały oznaczone symbolami T1 – okres przed upadkiem LB. (- 525 obserwacji), T2 – okres po upadku LB (- 880) i B - okres bessy (- 440), H - okres wzrostów (- 440) oraz zagregowane okresy badania T=T1+T2 i BH=B+H.

W całym badanym okresie T, jak i obu podokresach T1 i T2 oraz BH i B jedno-okresowe dzienne stopy zwrotu dla funduszy europejskich są ujemne dla większości funduszy oraz indeksów (tab. 1). Przedstawione stopy zwrotu wskazują wyraźnie, że istnieje znaczące zróżnicowanie dochodów z inwestycji w okresie wzrostów i spadków, natomiast w wyróżnionych podokresach T1 i T2 nie zauważa się odmiennej sytuacji na rynku. Potwierdziły to testy statystyczne, z których wynika, że zarówno zwroty, jaki i ich różnice wyznaczone dla różnych okresów badania istotnie różnią się od zera tylko dla okresów oznaczonych jako B i H.

Przeprowadzone badanie statystyczne (por. [Foo, Witkowska 2012]) nie pozwoliło odrzucić hipotez o równości stóp zwrotu uzyskanych z rozpatrywanych funduszy oraz indeksów giełdowych i bonu skarbowego w okresach T, T1 i T2. Nie stwierdzono również różnic w wartościach oczekiwanych zwrotów dla różnych okresów badania. Z kolei przeprowadzone testy równości średnich dla wyróżnionych trzech okresów analizy B, H i BH pozwalają stwierdzić, że w stosunku do niemieckich bonów skarbowych (pełniących rolę instrumentu wolnego od ryzyka) wartości oczekiwanych zwrotów są większe w czasie dobrej koniunktury rynkowej, ale w czasie dekoniunktury straty są również większe. Oprócz tego stwierdzono, że w większości przypadków nie było podstaw do odrzucenia hipotez o równości stóp zwrotu funduszy i indeksów rynkowych, wyjątek stanowią zwroty z funduszu Veri-EUROPA, które w okresie wzrostów są istotnie większe od CAC 40 i FTSE 100.

WYNIKI

W kolejnych tabelach przedstawiono wyniki badań efektywności wybranych funduszy, które oznaczono za pomocą kolejnych liter alfabetu jak to pokazano w tab. 1. Przeprowadzone analizy dotyczą poszczególnych funduszy inwestycyjnych, dla których - po wyznaczeniu wymienionych wcześniej mierników efektywności - utworzono rankingi zbudowane w oparciu o:

- wartości wykorzystanych mierników (tj. wszystkie wyznaczone miary efektywności uzupełnione jedno-okresowymi stopami zwrotu, co odpowiada strategii „kup i trzymaj”);
- przyjęte indeksy rynku: Euro STOXX 50, DAX, CAC 40 oraz FTSE 100;
- rozpatrywane okresy badania: T, T1, T2, B, H i BH.

Innymi słowy, liczby w tabelach 2 - 4 oznaczają pozycję funduszu w rankingu od najbardziej (- pierwsza pozycja w rankingu) do najmniej efektywnych.

Tabela 2. Ranking i klasyfikacja funduszy wg współczynnika Sharpe'a

Nr	T	T1	T2	BH	B	H
A	2 ***	9 **	2 ****	2 ****	5 ***	13 **
B	17 **	16	17 *	17	18	17 *
C	5 ***	3 ***	3 ****	5 ***	6 ***	14 **
D	18	18	18 *	18	17	9 ***
E	16 **	13 **	15 **	14 **	16	3 ****
F	6 **	8 **	7 **	7 **	8 **	15 *
G	9 **	6 ***	12 **	10 **	10 *	18
H	14 **	15	10 **	12 **	11 *	4 ****
I	13 **	14 **	16 *	16	9 **	11 **
J	15 **	17	14 **	15	15	6 ****
K	8 **	4 ***	8 **	9 **	12 *	8 ***
L	4 ***	5 ***	4 ****	3 ***	2 ****	2 ****
M	3 ***	2 ***	5 ***	4 ***	7 ***	5 ****
N	10 **	12 **	11 **	13 **	13 *	7 ***
O	11 **	11 **	9 **	8 **	4 ****	10 ***
P	12 **	7 ***	13 **	11 **	14 *	12 **
R	1 ****	1 ***	1 ****	1 ****	1 ****	16 *
S	7 **	10 **	6 **	6 **	3 ****	1 ****
ES 50	*	**	*	*	*	***
DAX	****	****	***	***	***	****
CAC 40	**	*	**	**	**	*
FTSE 100	***	***	****	****	****	**

Źródło: obliczenia własne

Dla wyróżnionych sześciu okresów analizy oraz czterech indeksów rynku wyznaczono współczynniki efektywności Sharpe'a, Treynora i Jensena. W przypadku pierwszych dwóch indeksów ocena jakości zarządzania funduszem dokonywana jest poprzez porównanie wartości wskaźnika z jego wartością wyznaczoną dla reprezentanta rynku. Zatem ocena efektywności badanych funduszy ściśle zależy od wybranego indeksu rynku oraz okresu analizy. Jak stwierdzono (tab. 2) spośród wszystkich indeksów rynku współczynnik Sharpe'a największe wartości osiągnął dla indeksów DAX (dla T, T1 i H) i FTSE 100 (dla T2, HB i B), co oznaczono czterema gwiazdkami w ostatnich 4 wierszach, podczas gdy indeks ES 50 aż w czterech okresach był najmniejszy (jedna gwiazdka), a w pozostałych dwóch (tj. T1 i H) – najsłabiej „wypadł" CAC 40. Konsekwentnie, oceniając efektywność funduszy w kolejnych okresach podano ich pozycje w rankingu i oznaczono je gwiazdkami – od jednej do czterech. Przy czym **** oznacza fundusz bardziej efektywny niż „najlepszy" z indeksów rynkowych (tj. ten, który uzyskał ****), a * - fundusz bardziej efektywny niż „najsłabszy"

z indeksów. Brak gwiazdki oznacza, że wartość miernika Sharpe'a dla tego funduszu jest mniejsza od wartości tej miary wyznaczonej dla „najsłabszego" z indeksów w danym okresie.

Tabela 3. Ranking i klasyfikacja funduszy wg współczynnika Treynora

Fundusz		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	R	S
Euro STOXX 50	T	**3**	16	**4**	18	17	**6**	**8**	15	12	13	**7**	**5**	**2**	9	14	**11**	**1**	**10**
	T1	**7**	14	**3**	18	13	**8**	**5**	16	17	15	**2**	**4**	**1**	10	12	**6**	9	11
	T2	**2**	17	**3**	18	15	**6**	**11**	14	16	13	**7**	**4**	**5**	10	**9**	12	**1**	**8**
	BH	**4**	16	**2**	18	13	**5**	9	15	17	14	7	**6**	**3**	11	10	12	**1**	8
	B	6	14	2	17	16	4	8	18	15	10	5	3	12	9	13	11	1	7
	H	**14**	**17**	**15**	**6**	**4**	**13**	**18**	**2**	**7**	**10**	**5**	**3**	**8**	**9**	**11**	**12**	**16**	**1**
DAX	T	3	16	4	18	17	6	8	13	12	14	7	5	2	9	15	11	**1**	10
	T1	7	14	3	18	13	8	5	16	17	15	2	4	1	9	12	6	10	11
	T2	**2**	17	**3**	18	16	6	11	14	15	13	7	4	5	10	9	12	**1**	8
	BH	4	16	2	18	13	6	9	14	17	15	7	5	3	10	11	12	**1**	8
	B	5	13	2	18	17	8	7	16	14	12	6	3	10	9	15	11	**1**	4
	H	**16**	**17**	**15**	**5**	**4**	**13**	**18**	**1**	**7**	**10**	**6**	**3**	**9**	**8**	**11**	**12**	**14**	**2**
CAC 40	T	3	16	4	18	17	6	8	15	12	13	7	5	2	9	14	11	**1**	10
	T1	7	14	3	18	13	8	5	16	17	15	2	4	1	10	12	6	**9**	11
	T2	2	17	3	18	15	6	11	14	16	13	7	4	5	10	9	12	**1**	8
	BH	**4**	15	**2**	18	13	**5**	9	16	17	14	7	**6**	**3**	10	11	12	**1**	8
	B	5	12	2	17	16	4	7	18	15	10	6	3	13	9	14	11	**1**	8
	H	**14**	**17**	**15**	**6**	**4**	**13**	**18**	**2**	**8**	**10**	**5**	**3**	**7**	**9**	**11**	**12**	**16**	**1**
FTSE 100	T	3	16	4	18	17	6	8	13	12	14	7	5	2	9	15	11	1	10
	T1	8	14	3	18	12	7	5	16	17	15	2	4	1	10	13	6	9	11
	T2	2	17	3	18	16	6	11	14	15	13	7	4	5	10	9	12	1	8
	BH	4	15	2	18	13	5	9	16	17	14	7	6	3	10	12	11	**1**	8
	B	5	13	2	15	16	4	7	18	12	10	6	3	14	9	17	11	1	8
	H	**14**	**17**	**15**	**6**	**4**	**13**	**18**	**2**	**9**	**10**	**5**	**3**	**8**	**7**	**11**	**12**	**16**	**1**

Źródło: obliczenia własne
Pogrubiono te fundusze, które były bardziej efektywne niż rynek

W tab. 2 zauważa się, że w przypadku zastosowania indeksu DAX tylko 13% wszystkich rozpatrywanych przypadków wskazuje, że fundusz inwestycyjny jest wg miernika Sharpe'a bardziej efektywny niż rynek, w przypadku indeksu FTSE 100 takich funduszy jest ponad 32%. Z kolei jeśli indeksem rynku jest ES

50, to fundusze będące bardziej od niego efektywne stanowią już 84% rozpatrywanych przypadków. Należy przy tym zauważyć, że indeks ES 50 zawiera w sobie spółki pochodzące z krajów, które w badanym okresie miały poważne kłopoty tj. Grecja, Portugalia, Hiszpania, Włochy i Irlandia. Natomiast biorąc pod uwagę wszystkie indeksy i okresy okazuje się, że wskazania na efektywność funduszy mniejszą niż „najsłabsza" rynkowa pojawiają się w 13% rozpatrywanych przypadków.

Warto zauważyć, że sytuacja na rynku niemieckim wydaje się znacząco lepsza niż w pozostałych krajach, o czym świadczy zarówno stopa zwrotu z indeksu DAX (por. tab. 1), jak i znikoma liczba przypadków lepszej niż rynkowa (jeśli rynek opisany jest przez DAX) efektywności. Drugim rynkiem, który charakteryzował się względną efektywnością jest londyński parkiet. Mierniki efektywności funduszy inwestycyjnych wyznaczone dla FTSE 100 również rzadko są większe od współczynników wyznaczonych dla tego indeksu. Z kolei indeks rynku francuskiego CAC 40 oraz paneuropejski indeks strefy euro ES 50 charakteryzowały się ujemnymi zwrotami we wszystkich analizowanych okresach (z wyjątkiem hossy) nie dziwi zatem, że ocena efektywności funduszy w stosunku do tych indeksów wskazuje dla większości z nich lepsze wyniki niż przeciętnie spodziewane z rynku.

W przypadku współczynnika Treynora (por. tab. 3) sytuacja jest podobna tzn. w większości przypadków fundusze są mniej efektywne niż porównywane rynki (z wyjątkiem okresu wzrostów) i liczba funduszy „bardziej efektywnych niż rynek" zwiększa się jeśli porównujemy je ze „słabszymi rynkami". Warto dodać, że wszystkie rynki charakteryzowały się ujemną wartością współczynnika Treynora we wszystkich okresach z wyjątkiem hossy oraz w przypadku indeksu DAX w okresach T i T1, a także FTSE w T2.

Analizując wartości alfy Jensena zauważa się, że żaden z funduszy nie uzyskał dodatniej wartości tego współczynnika, co może wynikać z dwóch przyczyn tj. słabego zarządzania funduszami lub wybrania niemieckiego bonu skarbowego jako instrumentu wolnego od ryzyka. W tab. 4 przedstawiono ranking funduszy przeprowadzony dla ilorazu alfy przez betę.

Tabela 4. Ranking i klasyfikacja funduszy wg współczynnika Jensena

Fundusz		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	R	S
Euro STOXX 50	T	1	15	4	18	17	6	8	16	12	13	7	5	3	9	14	10	1	11
	T1	8	14	2	18	13	7	5	15	17	16	3	4	1	10	12	6	8	11
	T2	1	15	3	18	14	6	11	16	17	12	8	5	4	10	9	13	1	7
	BH	2	15	3	17	14	5	9	16	18	13	6	7	4	11	10	12	1	8
	B	2	12	3	17	16	6	8	18	15	10	5	4	14	9	13	11	1	7
	H	14	17	15	6	4	13	18	2	8	10	5	3	7	9	11	12	16	1
DAX	T	1	14	4	18	17	6	8	15	12	13	7	5	3	9	16	10	1	11
	T1	8	14	2	18	13	7	5	16	17	15	3	4	1	10	12	6	8	11
	T2	1	16	3	18	14	6	11	15	17	12	8	5	4	10	9	13	1	7
	BH	2	16	3	18	13	8	9	15	17	14	6	5	4	10	11	12	1	7
	B	2	13	3	18	16	8	7	17	15	10	6	4	11	9	14	12	1	5
	H	16	17	15	5	4	13	18	1	7	10	6	3	9	8	11	12	14	2
CAC 40	T	1	14	4	18	16	6	8	17	12	13	7	5	3	9	15	10	1	11
	T1	8	14	2	18	13	7	5	15	17	16	3	4	1	10	12	6	8	11
	T2	1	15	3	18	14	6	11	16	17	12	7	5	4	10	9	13	1	8
	BH	2	15	3	17	14	5	9	16	18	13	6	7	4	10	11	12	1	8
	B	2	12	3	17	16	5	7	18	15	10	6	4	13	9	14	11	1	8
	H	14	17	15	6	4	13	18	2	8	10	5	3	7	9	11	12	16	1
FTSE 100	T	1	14	4	18	17	6	8	15	12	13	7	5	3	9	16	11	1	10
	T1	8	14	2	18	12	7	5	15	17	16	3	4	1	10	13	6	8	11
	T2	1	15	3	18	14	6	11	17	16	12	7	4	5	10	9	13	1	8
	BH	2	15	3	18	14	4	9	16	17	13	7	6	5	10	12	11	1	8
	B	3	13	2	15	16	5	7	18	12	10	6	4	14	9	17	11	1	8
	H	14	17	15	6	4	13	18	2	9	10	5	3	8	7	11	12	16	1

Źródło: obliczenia własne

Biorąc pod uwagę mierniki Treynora i Jensena oraz częstotliwość, z jaką poszczególne fundusze osiągają dobre wyniki inwestycyjne, za najlepsze fundusze inwestujące na rynku paneuropejskim można uznać: UniExtra: EuroStoxx-50, The BlackRock Global funds-Euro-Markets A2 EUR oraz Allianz RCM Wachstum Euroland A EUR, które osiągały wysoką efektywność we wszystkich badanych okresach z wyjątkiem okresu wzrostów (H). W tych rankingach najbardziej stabilnymi o dość wysokiej pozycji okazały się: HSBC Global Investment Funds Euroland Equity A EUR i Pioneer Investments Euroaktien. Natomiast funduszem najmniej efektywnym okazał się Credit Suisse Aktien Plus A.

Na podstawie wyznaczonych mierników syntetycznych, w których wykorzystano 60 rankingów zbudowanych na podstawie różnie zdefiniowanych miar efektywności, dokonano oceny jakości zarządzania rozpatrywanych funduszy inwestycyjnych (tab. 5). Z przeprowadzonych analiz wynika, że najbardziej efektywnym okazał się fundusz UniExtra: EuroStoxx-50, drugie miejsce zajął Pioneer Investments Euroaktien, trzecie - The BlackRock Global funds-Euro-Markets A2 EUR. Ostatnie miejsce w rankingu zajął, wcześniej wymieniany jako najmniej efektywny, fundusz Credit Suisse Aktien Plus A[5].

Tabela 5. Ranking funduszy inwestycyjnych wg mierników agregatowych

Nr	Liczba przypadków przynależności do danej klasy efektywności					Ranking wg mierników	
		*	**	***	****	wg klas CS_i	wg rankingu MS_i
A	9	6	16	22	7	2	5
B	29	9	20	1	1	17	17
C	9	5	17	28	1	3	3
D	44	7	6	2	1	18	18
E	25	10	21	1	3	13	14
F	12	5	39	3	1	6	7
G	22	5	29	3	1	11	10
H	26	15	12	4	3	15	15
I	26	13	19	1	1	16	16
J	27	6	24	1	2	14	13
K	13	7	33	6	1	7	6
L	12	4	26	14	4	5	2
M	10	5	22	18	5	4	4
N	19	6	32	2	1	9	9
O	22	7	27	2	2	12	12
P	18	8	31	2	1	10	11
R	2	11	9	16	22	1	1
S	17	13	22	5	3	8	8

Źródło: obliczenia własne

WNIOSKI

W podsumowaniu należy zauważyć, że rynek europejski charakteryzuje znaczące zróżnicowanie, które sprawia istotne kłopoty z wyborem uniwersalnego dla tego rynku indeksu oraz instrumentu wolnego od ryzyka. Prowadzone badania dowodzą, że w ocenie jakości inwestowania ma to zasadnicze znaczenie. Jednakże ocena efektywności funduszy inwestycyjnych działających na paneuropejskim

[5] Wymieniany we wcześniejszych analizach jako jeden z najbardziej efektywnych funduszy: Allianz RCM Wachstum Euroland A EUR znalazł się na 5. pozycji rankingu.

rynku jest w mniejszym stopniu zależna od przyjętego benchmaku niż od okresu analizy. Stwierdzono ponadto, że zarządzający funduszami w niewielkim stopniu uwzględniają zmiany sytuacji rynkowej. Zatem portfele funduszy inwestycyjnych są mało „elastyczne" i – z wyjątkiem najsłabszych funduszy – z trudem utrzymują swoją pozycję w rankingu w warunkach zmieniającej się koniunktury rynkowej. Przeprowadzone analizy dowiodły, że zastosowanie mierników agregatowych uwzględniających różne miary i okresy analizy pozwala na ujednolicenie rankingów i wskazanie na fundusze najbardziej i najmniej efektywne.

BIBLIOGRAFIA

Bergstresser D., Poterba J. M. (2002) Do After-Tax Returns Affect Mutual Fund Inflows?, Journal of Financial Economics 63, s. 381-414.

Busse J. A., Irvine P. J. (2006) Bayesian Alphas and Mutual Fund Persistence, Journal of Finance 61, s. 2251–2288.

Czekaj J. (2008) Rynki, instrumenty i instytucje finansowe, Warszawa, PWN.

Dębski W. (2005) Rynek finansowy i jego mechanizmy, Warszawa, PWN.

Deaves R. (2004) Data-conditioning Biases, Performance, Persistence and Flows: The Case of Canadian Equity Funds, Journal of Banking and Finance 28, s. 673-694.

Dutta A. S., Su R., Xu M. (2010) Persistence in New Zealand Growth Mutual Funds Returns**:** An Examination of New Zealand Mutual Funds from 1997-2003, Indian Journal of Economics & Business, Vol. 9, No. 2, s. 303-314.

Edelen R. M., (1999) Investor Flows and the Assessed Performance of Open-end Mutual Funds, Journal of Financial Economics 53, s. 439-466.

Foo J., Witkowska D. (2012) Mutual Funds Efficiency Comparison between U.S. and Europe, paper presented at IAES Conference, Istanbul, March 2012.

Goetzmann W. N., Peles N. (1997) Cognitive Dissonance and Mutual Fund Performance, Journal of Financial Research 20, s. 145-158.

Grinblatt M., Titman S. (1994) A Study of Monthly Mutual Fund Returns and Performance Evaluation Techniques, Journal of Financial and Quantitative Analysis, vol. 29, issue 3, s. 419-444.

Gruber, M. J. (1996) Another Puzzle: The Growth in Actively Managed Mutual Funds. Journal of Finance, vol. 51, issue 3, s.783-810.

Hendricks, D., Patel, J., Zeckhauser, R. J. (1993) Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988. Journal of Finance 48, s. 93–130.

Ippolito R. A. (1992) Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry, Journal of Law and Economics 35, s. 45-70.

Jensen, M.C. (1968) The performance of Mutual Funds in the period 1945-1964, Journal of Finance, 23, No. 2, s. 389-416.

Jensen, M. C. (1969) Risk, the pricing of capital assets, and the evaluation of investment portfolios, Journal of Business, vol. 42, issue 2, s. 167-247.

Karpio A., Żebrowska – Suchodolska D. (2011) Stability of Strategies of Polish Open-End Investment Funds Investing in Global Markets during the Tinancial Crisis, Acta Scientiarum Polonorum, Oeconomia, 10(3), s. 51-59.

Kompa K., Witkowska D. (2010) Porównanie efektywności wybranych otwartych funduszy inwestycyjnych w okresie hossy i bessy, Acta Scientiarum Polonorum, Oecomnomia, 9(3), s.169 – 180.

Kosarga R. (2002) Ranking funduszy inwestycyjnych: fundusze zrównoważone na pierwszy ogień (2002-05-27) w: http://www.bankier.pl/wiadomosc/Ranking-funduszy-inwestycyjnych-fundusze-zrownowazone-na-pierwszy-ogien-343224.html.

Lachowski S. (2012) Ocena efektywności europejskich funduszy inwestycyjnych, SGGW, Warszawa, praca magisterska przygotowana pod kierunkiem D. Witkowskiej.

Otten R., Bams D. (2002) European Mutual Fund Performance. European Financial Management, Vol.8, No.1, 2002, s. 75-101.

Otten R., Schwitzer M. (2002) A Comparison between the European and the U.S. Mutual Fund Industry. Managerial Finance, Vol. 28(1), s. 14-35, 2002.

Patel J., Zeckhauser R. J., Hendricks D. (1994) Investment Flows and Performance: Evidence from Mutual Funds, Cross-Border Investments, and New Issues, w: Sato, R., Levich, R., Ramachanday, R. V. (red.), Japan, Europe and International Financial Markets: An Analytical and Empirical Perspective, Cambridge: England, s. 34-48.

Prince T., Bacon F. (2010) Analyzing Mutual Fund Performance Against Established Performance Benchmarks: A Test of Market Efficiency, Research in Business and Economic Journal, Vol. 1, February 2010, www.aabri.com/manuscripts/08112.pdf.

Reilly, F., Akhtar, R. (1995) The benchmark error problem with global capital markets, Journal of Portfolio Management, vol. 22, issue 1, s. 33-50.

Roll R. (1978) Ambiguity When Performance is Measures by the Securities Market Line, Journal of Finance, Vol. 33, s. 1051 – 1069.

Sharpe, W. F. (1966) Mutual Fund Performance, Journal of Business, 39, No. 1, s. 119-138.

Sirri, E. R., Tufano P. (1993) Buying and Selling Mutual funds: Flows, Performance, Fees, and Services, Working Paper, Harvard Business School.

Treynor, J.L., (1965) How to Rate Management of Investment Funds", Harvard Business Review, 43, No. 1, s. 63-75.

CONSTRUCTION OF SYNTHETIC MEASURES TO EVALUATE THE EFFICIENCY OF EUROPEAN MUTUAL FUNDS

Abstract: The aim of the research is construction of aggregated measures that allow to evaluate the performance of mutual funds operating at pan-European market. The synthetic measures base on well known efficiency indexes i.e. Sharpe, Treynor and Jensen ratios as well as returns. Empirical research concerns 18 selected European mutual funds, 4 different markets indexes e.g. Euro STOXX 50, DAX, CAC 40 and FTSE 100 and 6 samples distinguished from the period 15.09. 2006 - 20.01.2012. It is showed that application of aggregated measures let the synonymous evaluation and ranking of mutual funds.

Keywords: synthetic measure, investment efficiency, Sharpe, Treynor and Jensen ratios

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 276 – 287

BADANIE ATRAKCYJNOŚCI OFERTY DOSTĘPU DO INTERNETU ZA POMOCĄ ANALIZY GRADACYJNEJ

Tomasz Ząbkowski, Wiesław Szczesny
Katedra Informatyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: tomasz_zabkowski@sggw.pl, wieslaw_szczesny@sggw.pl

Streszczenie: Celem badań była analiza atrakcyjności oferty dostępu do mobilnego Internetu wśród klientów. W pracy wykorzystano podejście gradacyjne do wizualnej prezentacji danych oraz dokonania oceny podobieństwa pakietów pod względem ceny w różnych grupach wiekowych klientów. Pozwoliło to sformułować wnioski, że atrakcyjność oferty dostępu do mobilnego Internetu różni się w zależności od grupy wiekowej klientów oraz kategorii dochodowej. Jednocześnie badanie pozwoliło zidentyfikować segmenty nabywców o podobnych cechach.

Słowa kluczowe: analiza gradacyjna, Internet, mapy nadreprezentacji

CEL BADAŃ

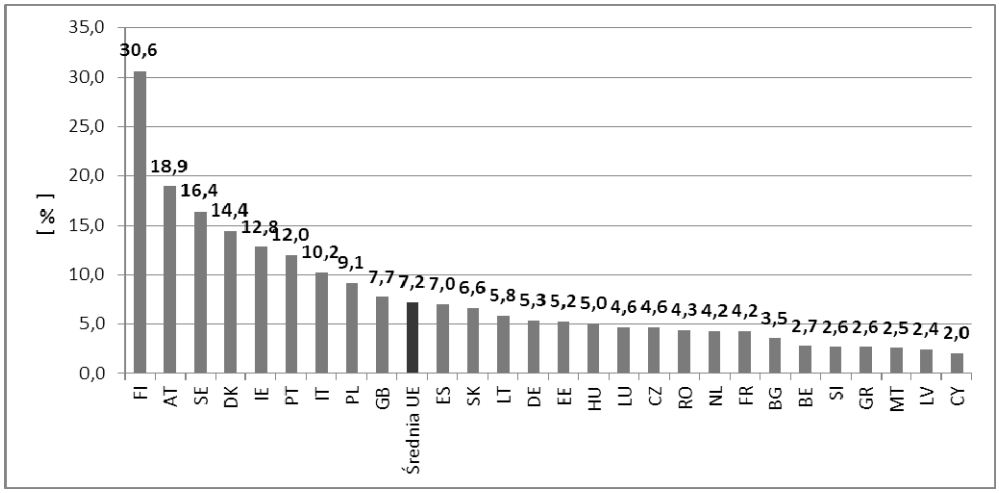
Celem prezentowanych badań była ocena atrakcyjności oferty dostępu do internetu mobilnego. W opracowaniu wykorzystano podejście gradacyjne do zebranych danych empirycznych, co umożliwiło wizualną prezentację danych oraz porównanie i ocenę podobieństwa ofert pod względem ceny w różnych grupach wiekowych oraz różnych kategoriach dochodowych. Wykorzystanie analizy gradacyjnej w pracy wynika z przesłanek o dużym znaczeniu regularności w danych. Występowanie zależności między zmiennymi i obiektami połączone z regularnością może być pomocne w opisie i segmentacji danych. Odkrycie pewnych jednorodnych podgrup klientów w danych empirycznych, mających podobne charakterystyki może zostać wykorzystane do tworzenia odpowiednio dobranych strategii marketingowych w celu budowania trwałych relacji z klientem (CRM).

OPIS PROBLEMU I DANE

Rynek dostępu do Internetu w Polsce rozwija się bardzo dynamicznie, obejmując swoim zasięgiem coraz szersze grono odbiorców. Bezpośrednim tego odzwierciedleniem jest rosnąca wartość rynku i duża liczba klientów. Według danych Urzędu Komunikacji Elektronicznej [UKE 2010] w 2010 roku blisko 6,1 mln osób posiadało stacjonarny i około 2,8 mln osób dostęp za pomocą modemów 2G/3G do Internetu. Usługi mobilnego dostępu przez modemy 2G/3G w urządzeniach przenośnych stanowiły najbardziej dynamicznie rozwijający się segment rynku, co powodowało, że była to druga (zaraz po dostępie xDSL) najczęściej wybierana forma dostępu do sieci. W dużej mierze był to efekt działalności operatora Play, który w 2010 roku, o blisko 86% zwiększył bazę klientów i tym samym osiągnął prawie 22% liczby użytkowników ogółem. Upowszechnianie się dostępu modemami 2G/3G do Internetu wynikało m.in. (a) z dostrzeżonych przez klientów zalet mobilności usługi, (b) wzrostu pokrycia siecią 3G terytorium kraju, (c) zachodzących procesów modernizacyjnych umożliwiających korzystanie z większych prędkości przesyłowych.

Porównując Polskę z krajami Unii Europejskiej pod względem wielkości wskaźnika dostępności mobilnego dostępu do Internetu, Polska, na podstawie danych z Digital Agenda Scoreboard [UKE 2010], zajęła ósme miejsce pośród państw Unii Europejskiej (por. rys. 1), przewyższając średnią unijną o 1,9 pp.

Rys. 1. Penetracja mobilnego szerokopasmowego dostępu do Internetu w UE na 100 mieszkańców (w %)



Źródło: opracowanie własne na podstawie [UKE 2010]

Dane wykorzystane w analizie stanowił zbiór 2821 klientów, którzy zdecydowali się w 2010 roku na zakup pakietu dostępu do mobilnego Internetu u operatorów świadczących usługi telefonii (por. tabela 1).

Na podstawie dostępnych danych, wydzielono sześć kategorii wiekowych: klienci w wieku do lat 20, 21-25 lat, 26-35 lat, 36-45 lat, 46-60 lat oraz powyżej 60 lat. Klienci mieli do wyboru w ofercie operatora pięć pakietów dostępu do Internetu w cenie: 39 zł, 49 zł, 69 zł, 89 zł oraz 129 zł.

Dodatkowo została wyodrębniona kategoria dochodowa, która dla A1 oznacza, że dany klient nabył pakiet o wartości wyższej niż średnia wartość jego dotychczasowych rachunków z ostatnich 6 miesięcy. A2 przedstawia sytuację odwrotną tzn., że dany klient nabył pakiet o wartości niższej niż średnia wartość jego dotychczasowych rachunków z ostatnich 6 miesięcy. Przykładowo, dla pierwszej grupy wiekowej (do lat 20), 233 klientów nabyło pakiet internetowy w cenie 39 zł, co było wartością wyższą niż średnia wartość dotychczasowych rachunków z ostatnich 6 miesięcy. Z kolei 41 klientów w tej grupie wiekowej nabyło pakiet za 39 zł, co stanowiło wartość niższą niż średnia wartość dotychczasowych rachunków z ostatnich 6 miesięcy. Innymi słowy w kategorii A1 mamy do czynienia z klientami, którzy płacą średnio niższe rachunki niż wartość nabywanego pakietu internetowego, natomiast w kategorii A2 mamy klientów, którzy płacą średnio wyższe rachunki niż wartość nabywanego pakietu.

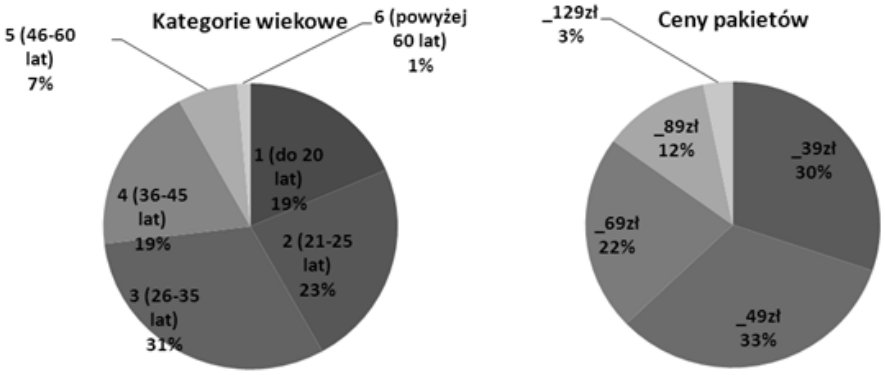
Tabela 1. Dane empiryczne wykorzystane w analizie

Grupa wiekowa	Kategoria	Wartość pakietu					Suma	Suma
		39zł	49zł	69zł	89zł	129zł		
1 (do 20 lat)	A1	233	92	38	16	2	381	523
	A2	41	37	46	17	1	142	
2 (21-25 lat)	A1	106	107	61	10	1	285	659
	A2	79	122	109	47	17	374	
3 (26-35 lat)	A1	78	160	155	138	47	578	879
	A2	34	115	95	45	12	301	
4 (36-45 lat)	A1	121	124	79	46	10	380	533
	A2	46	77	23	6	1	153	
5 (46-60 lat)	A1	68	55	11	8	0	142	186
	A2	23	17	4	0	0	44	
6 (powyżej 60 lat)	A1	21	12	1	0	0	34	41
	A2	2	3	2	0	0	7	
Suma		852	921	624	333	91	2821	2821

Źródło: opracowanie własne

W zamieszczonych w tabeli danych zwraca uwagę to, że w strukturze wiekowej najbardziej liczna jest trzecia grupa wiekowa (26-35 lat). Natomiast najmniej licznie reprezentowana jest grupa klientów powyżej 60 roku życia, gdyż stanowi ona jedynie 1%. W przypadku pakietów internetowych najbardziej popularne są pakiety najtańsze, w cenie 39 zł i 49 zł i stanowią one odpowiednio: 30% i 33% spośród wszystkich nabywanych pakietów (por. rys. 2).

Rys. 2. Udział procentowy według kategorii wiekowych oraz cen pakietów



Źródło: opracowanie własne

GRADACYJNA ANALIZA DANYCH

Gradacyjna analiza danych jest zaliczana do technik wielowymiarowej eksploracji danych, określanych także jako Data Mining. Istotną zaletą analizy gradacyjnej jest jednoczesna integracja obszarów z pogranicza teorii pomiaru, statystyki, taksonomii, wielowymiarowej analizy danych (także wizualizacji) oraz zastosowanie podejścia umożliwiającego jednolite traktowanie danych różnego typu. Szerszy jej opis można znaleźć m.in. w [Szczesny 2002], [Kowalczyk i in. 2004]. Ważną cechą analizy gradacyjnej jest jej ukierunkowanie na wizualizację, celem lepszego zrozumienia i poznania całościowej struktury danych i występujących w nich współzależności.

Wykorzystanie analizy gradacyjnej w pracy wynika z przesłanek o dużym znaczeniu regularności w danych, mogących mieć znaczenie z punktu widzenia badań rynku i kierowania odpowiedniej oferty marketingowej (pakietów dostępu do Internetu) do odpowiednich grup klientów.

Podstawowym zagadnieniem w ramach metody gradacyjnej analizy danych jest badanie zróżnicowania (koncentracji) rozkładów poprzez dążenie do takiego uporządkowania macierzy danych, aby osiągnąć maksymalny kontrast pomiędzy skrajnymi wierszami i kolumnami [Borkowski i Szczesny 2005], [Pleszczyńska i in. 2006]. Dane w postaci tabeli zostają poddane jednoczesnemu przestawieniu zarówno wierszy jak i kolumn, tak, aby w tabeli wynikowej cecha porządkująca wiersze była jak najsilniej dodatnio zależna od cechy porządkującej kolumny.

Celem przybliżenia zagadnienia zostanie przedstawiony przykład, w którym mamy dwa rozkłady: $p = (p_1,...,p_k)$, $s = (s_1,...,s_k)$ określone na tym samym zbiorze. W prezentowanym przykładzie posłużymy się danymi dotyczącymi pakietów o wartości 39 zł i 69 zł (por. tabela 2). Na ich podstawie wyznaczona została krzywa koncentracji rozkładu s względem rozkładu p. Jest to krzywa

łamana, wpisana w kwadrat jednostkowy, łącząca kolejno punkty $(0,0),(p_1,s_1),(p_1+p_2,s_1+s_2),...,(p_1+...+p_k,s_1+...+s_k),(1,1)$. Nachylenie odcinków krzywej wyznacza iloraz s_i/p_i, który może przebiegać dowolnie według naturalnego porządku (por. kol. 6 w tabeli 2) lub w sposób uporządkowany rosnąco (por. kol. 7 w tabeli 2).

Tabela 2. Tabela liczności według grup wiekowych oraz pakietów internetowych o wartości 39 zł i 69 zł

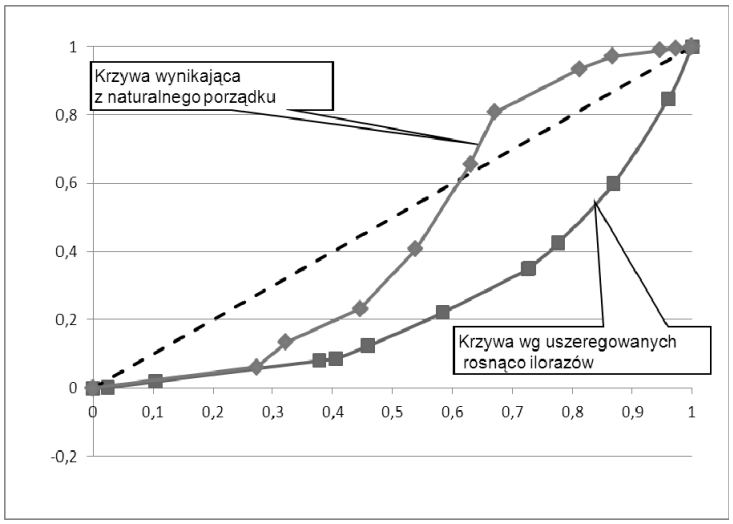
Grupa wiekowa	Pakiet		Struktura		Iloraz	Kolejność ilorazu	Punkty krzywej	
	39	69	p_i	s_i	s_i / p_i		Fp_i	Fs_i
1	2	3	4	5	6	7	8	9
1_A1	233	38	0,273	0,061	0,223	3	0	0
1_A2	41	46	0,048	0,074	1,532	9	0,025	0,002
2_A1	106	61	0,124	0,098	0,786	6	0,104	0,019
2_A2	79	109	0,093	0,175	1,884	10	0,378	0,080
3_A1	78	155	0,092	0,248	2,713	11	0,405	0,087
3_A2	34	95	0,040	0,152	3,815	12	0,459	0,123
4_A1	121	79	0,142	0,127	0,891	7	0,583	0,221
4_A2	46	23	0,054	0,037	0,683	5	0,725	0,348
5_A1	68	11	0,080	0,018	0,221	2	0,728	0,351
5_A2	23	4	0,027	0,006	0,237	4	0,776	0,425
6_A1	21	1	0,025	0,002	0,065	1	0,869	0,599
6_A2	2	2	0,002	0,003	1,365	8	0,960	0.848
Razem	852	624					1	1

Źródło: opracowanie własne

Krzywa koncentracji wynikająca z naturalnego uporządkowania może przebiegać zarówno nad, jak i pod przekątną. Jako miernik koncentracji *ar* wykorzystuje się podwójną różnicę między dwoma polami zawartymi miedzy przekątną a krzywą. Dokonując szeregowania grup wiekowych z przykładu w tabeli 2, tak aby kolejne ilorazy były rosnące (biorąc pod uwagę kolejność z kol. 7), w wyniku zostaje utworzona krzywa zróżnicowania leżąca pod krzywą zróżnicowania wyznaczoną przy dowolnym innym uporządkowaniu. Krzywą tą nazywa się maksymalną krzywą zróżnicowania (koncentracji) rozkładu *s* względem rozkładu *p*, oraz odpowiadający jej wskaźnik koncentracji nazywa się wskaźnikiem maksymalnej koncentracji z oznaczeniem $ar_{\max}$ [Borkowski i Szczesny 2005], [Pleszczyńska i in. 2006].

Ułożenie krzywej koncentracji może być dwojakie (por. rys. 3). W pierwszym przypadku, gdy dane reprezentują naturalny porządek, krzywa koncentracji przebiega zarówno nad i pod przekątną o współrzędnych (0,0) oraz (1,1). Natomiast w drugim, gdy dane są uszeregowane według niemalejących ilorazów krzywa przebiega pod przekątną.

Rys. 3. Krzywe koncentracji



Źródło: opracowanie własne

W rozważanym przykładzie przebieg krzywej koncentracji wskazuje, że klienci w ramach grup wiekowych w sposób odmienny byli zainteresowani rozważanymi pakietami dostępu do Internetu. Wskazuje na to maksymalna krzywa zróżnicowania, która jest znacznie oddalona od przekątnej.

W analizie gradacyjnej rozwiązanie określonego problemu dla danych w postaci tabeli (o nieujemnych wartościach) oznacza znalezienie takiego uporządkowania kolumn i wierszy, aby wskaźnik zależności gradacyjnej osiągnął maksimum globalne. Do tego celu wykorzystuje się współczynnik korelacji rho-Spearmana, oznaczany jako ρ* oraz τ Kendalla [Szczesny 2002]. W celu wyznaczenia maksimum lokalnego dla współczynnika ρ* należy wprowadzić pojęcie rozkładu gradacyjnego dla danej tabeli kontyngencji, oznaczone jako $P = \{p_{ij} : i = 1,..., m,\ j = 1,..., k\}$, gdzie $p_{ij} = n_{ij}/n$. Rozkład gradacyjny, oznaczony jako P* jest rozkładem ciągłym z gęstością opisaną w kwadracie jednostkowym. Zatem jeśli P jest rozkładem pary zmiennych losowych, to P* jest rozkładem łącznym tych zmiennych otrzymanym w wyniku przekształcenia gradacyjnego. Kwadrat jednostkowy <0,1> dzieli się na $m \times k$ prostokątów, o wymiarach $p_{i+} \times p_{+j}$:

$$p_{i+} = \sum_{i=1}^{k} p_{ij},\ p_{+j} = \sum_{i=1}^{m} p_{ij} \tag{1}$$

gdzie $i = 1,..., m,\ j = 1,..., k$.

Pole prostokąta o wymiarach i, j jest równe $p_{i+}p_{+j}$. Funkcja gęstości rozkładu P* jest stała w każdym prostokącie i równa jest ilorazowi $p_{ij}/p_{i+}p_{+j}$. Wartość tego ilorazu interpretuje się jako stopień nadreprezentacji w stosunku do

rozkładu jednostajnego. Im iloraz ten jest większy od 1, tym bardziej nadreprezentowana jest zawartość umieszczona w komórce p_{ij} w stosunku do tego, co wynikałoby z rozkładu proporcjonalnego. W przypadku wartości ilorazu poniżej 1, mamy do czynienia z niedoreprezentacją. Współczynniki regresji dla wierszy oraz kolumn, oznaczone jako $regr^{*}_{c|r}$ i $regr^{*}_{r|c}$, oraz współczynnik korelacji ρ* (rho) dla P*, nazywane są odpowiednio, funkcjami regresji gradacyjnej oraz współczynnik korelacji gradacyjnej dla rozkładu P. Miary zależności ρ* rho-Spearmana oraz współczynnik korelacji τ Kendalla dla kolumn wyrażają się następującymi równaniami:

$$\rho^{*}(P) = 3\sum_{t=2}^{m}\sum_{s=1}^{t-1}(S_t + S_{t-1} - S_s - S_{s-1})(p_{t+}p_{s+})ar(t:s) \tag{2}$$

oraz

$$\tau(P) = 2\sum_{t=2}^{m}\sum_{s=1}^{t-1}[(p_{t+}p_{s+})ar(t:s)] \tag{3}$$

gdzie w powyższych równaniach

$$ar(t:s) = \sum_{i=1}^{k}\sum_{j=i+1}^{k}[(p_{si}p_{tj} - p_{sj}p_{ti})/(p_{i+}p_{s+})] \tag{4}$$

oznacza wskaźnik koncentracji dla kolumny t względem kolumny s, oraz

$$S_u = \sum_{i=1}^{u} p_{i+},\ dla\ u = 1,...,k\ . \tag{5}$$

W sposób analogiczny przeprowadza się obliczenia dla wierszy.
Wartość ρ* zależy od kolejności wierszy i kolumn w tablicy o wyrazach p_{ij}. Liczba jednoczesnych przestawień wierszy i kolumn w małych tabelach jest skończona, stąd można znaleźć parę lub więcej par ustawień, przy których ρ* osiąga maksimum. Ustawienia tabeli, przy których ρ*= ρ*max są interesującym źródłem informacji o danych w niej zawartych i często daje się je interpretować [Książyk i in. 2005]. Jest to możliwe zwłaszcza wtedy, gdy tablica po maksymalizacji ρ* staje się na tyle regularna, że wszystkie wskaźniki $ar(t:s)$ dla kolumn i wierszy są równe odpowiadającym im $ar_{\max}(t:s)$.

W praktyce trudno oczekiwać znalezienia takich zależności, aby wszystkie wskaźniki $ar(t:s)$ oraz $ar_{\max}(t:s)$ były równe. W szczególności, dla tabel o dużych m i k, przeszukiwanie po wszystkich kombinacjach wierszy i kolumn jest procesem długotrwałym, dlatego do znalezienia maksimum globalnego wykorzystywane są metody symulacyjne typu Monte Carlo. W tym celu poszukuje się w pętli, rozpoczynając od losowego uporządkowania wierszy i kolumn tabeli, takiego ich ustawienia, dla którego dane ρ* osiąga maksimum lokalne. W postępowaniu tym wykorzystywany jest fakt, że jeśli któraś z regresji wierszy lub kolumn nie jest niemalejąca, to przestawiając odpowiednio wiersze lub kolumny zgodnie z wartościami regresji, otrzymujemy rozkład o większej wartości

ρ*. Rozkład ten, dla obydwu rodzajów regresji ustawionych niemalejąco, uzyskiwany jest po skończonej liczbie przestawień. Najwyższa wartość ρ* wybrana ze zbioru maksimów lokalnych jest z reguły równa maksimum globalnemu.

BADANIE ATRAKCYJNOŚCI OFERTY DOSTĘPU DO INTERNETU

Dostępny zbiór danych został poddany analizie gradacyjnej w pakiecie GradeStat opracowanym w Instytucie Podstaw Informatyki PAN. W pierwszym kroku, na podstawie danych zawartych w tabeli 1 zostały wyliczone wskaźniki nadreprezentacji dla każdej komórki leżącej na przecięciu grup wiekowych oraz cen pakietów. Przykładowo, dla kolumny z pakietem o wartości 39 zł i grupy wiekowej 1 oraz kategorii A1, wartość wskaźnika wynosi 2,025=233/(852×381/2821). Wyliczenia zawiera tabela 3.

Tabela 3. Wskaźniki nadreprezentacji

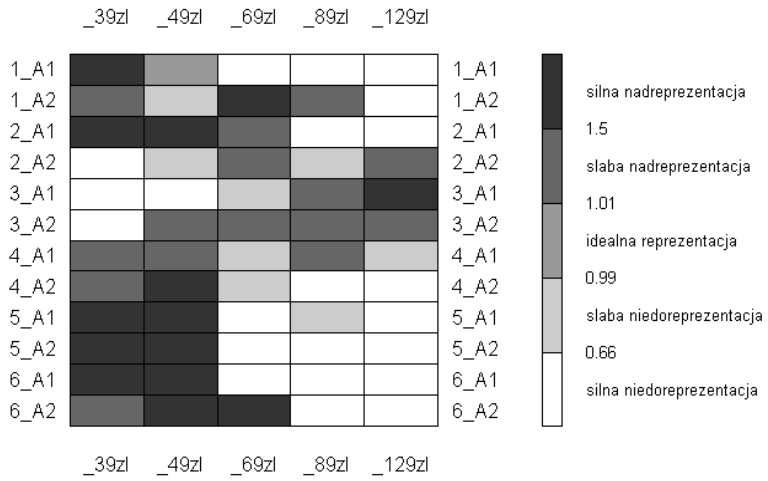
Grupa wiekowa	Kategoria	Wartość pakietu				
		39 zł	49 zł	69 zł	89 zł	129 zł
1 (do 20 lat)	A1	2,025	0,740	0,451	0,356	0,163
	A2	0,956	0,798	1,464	1,014	0,218
2 (21-25 lat)	A1	1,231	1,150	0,968	0,297	0,109
	A2	0,699	0,999	1,318	1,065	1,409
3 (26-35 lat)	A1	0,447	0,848	1,212	2,023	2,521
	A2	0,374	1,170	1,427	1,266	1,236
4 (36-45 lat)	A1	1,054	0,999	0,940	1,025	0,816
	A2	0,995	1,541	0,680	0,332	0,203
5 (46-60 lat)	A1	1,586	1,186	0,350	0,477	0,000
	A2	1,731	1,183	0,411	0,000	0,000
6 (powyżej 60 lat)	A1	2,045	1,081	0,133	0,000	0,000
	A2	0,946	1,313	1,292	0,000	0,000

Źródło: obliczenia własne

Dla wyznaczonych w tabeli 3 wskaźników nadreprezentacji można stworzyć mapę obrazującą stopień reprezentacji danych. W tym celu, za pomocą kilku odcieni szarości można określić obszary niedoreprezentacji, idealnej reprezentacji oraz nadpreprezentacji w danych. W rozpatrywanym przykładzie obszary te oznaczono w następujący sposób:
- wartość poniżej 0,66 oznacza silną niedoreprezentację;
- wartość w przedziale 0,66-0,99 oznacza słabą niedoreprezentację;
- wartość w przedziale 0,99-1,01 oznacza idealną reprezentację;
- wartość w przedziale 1,01-1,5 oznacza słabą nadreprezentację;
- wartość powyżej 1,5 oznacza silną nadreprezentację.

Reprezentację graficzną danych z tabeli 3 przedstawia rysunek 4.

Rysunek 4. Mapa nadreprezentacji dla danych z tabeli 3



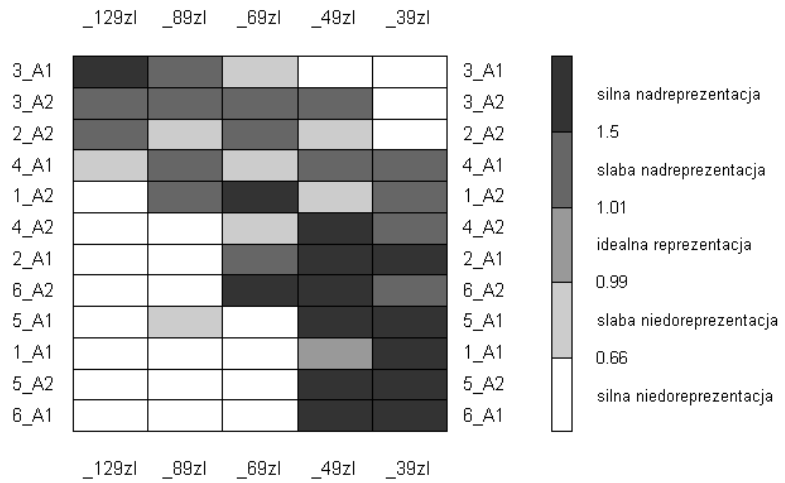
Źródło: opracowanie własne

Na rys. 4 widzimy wyraźnie, że pewne pakiety są bardzo licznie reprezentowane w określonych grupach wiekowych, a inne wprost przeciwnie. Rozważmy np. pakiety najtańsze, w cenie 39 zł i 49 zł, które są najczęściej wybierane przez klientów najmłodszych oraz tych powyżej 45 roku życia. Z kolei pakiety najdroższe są najpopularniejsze wśród klientów w trzeciej kategorii wiekowej (36-45 lat).

W następnym kroku surowe dane w postaci wskaźników nadreprezentacji zostały poddane w pakiecie GradeStat właściwej analizie polegającej na wyznaczeniu wartości współczynników regresji, zarówno dla wierszy oraz kolumn, celem reprezentacji poziomu korelacji gradacyjnej pomiędzy wymiarem kolumn, a wymiarem wierszy. W tym celu algorytm poszukuje w pętli, rozpoczynając od losowego uporządkowania wierszy i kolumn tabeli, takiego ustawienia wierszy i kolumn, dla którego dane ρ^* osiąga maksimum lokalne. W postępowaniu tym wykorzystywany jest fakt, że jeśli któraś z regresji wierszy lub kolumn nie jest niemalejąca, to przestawiając odpowiednio wiersze lub kolumny zgodnie z wartościami regresji, otrzymujemy rozkład o większej wartości współczynnika ρ^*. Najwyższa wartość ρ^* wybrana ze zbioru maksimów lokalnych jest z reguły równa maksimum globalnemu, co często uzyskiwane jest przy niewielkiej liczbie powtórzeń algorytmu.

Wynikiem ustawienia są duże wartości wskaźnika nadreprezentacji w lewym górnym rogu oraz prawym dolnym rogu tabeli. Z kolei małe wartości wskaźnika nadreprezentacji znajdują się w lewym dolnym oraz prawym górnym rogu tabeli. Wizualizację końcowego ustawienia dla rozważanych danych prezentuje mapa nadreprezentacji na rys. 5.

Rysunek 5. Mapa nadreprezentacji po przekształceniach



Źródło: opracowanie własne

Na podstawie wizualizacji z rys. 5, można zauważyć pewien związek pomiędzy grupą wiekową, a zakupem określonego pakietu internetowego. Najdroższe pakiety, za 89 zł i 129 zł, są wybierane przez klientów w trzeciej grupie wiekowej (26-35 lat), z kolei pakiety najtańsze, w cenie 39 zł i 49 zł są nabywane przez klientów najmłodszych (z pierwszej grupy wiekowej) oraz klientów najstarszych (z piątej oraz szóstej kategorii wiekowej). Można również zauważyć, że pakiety w cenie 39 zł i 49 zł są także często nabywane przez klientów z drugiej kategorii wiekowej, a dokładnie tych, którzy płacą niskie rachunki (kategoria A1). Z kolei pakiet w cenie 49 zł jest popularny (silna nadreprezentacja) wśród klientów z czwartej grupy wiekowej płacących wyższe rachunki (kat. A2).

Kolejny krok w badaniu polegał na pogrupowaniu badanej zbiorowości klientów na jednorodne grupy nabywców. W obecnych, bardzo konkurencyjnych, warunkach rynkowych segmentacja jest częstym podejściem, za którym stoi szereg przesłanek m.in. [Kotler 1999]:

a) w otoczeniu przedsiębiorstwa zachodzą ciągłe zmiany, co powoduje, że rynek nie ma charakteru homogenicznego, a postępowanie nabywców na rynku staje się coraz bardziej różnorodne;

b) z ogółu nabywców danego produktu można wyróżnić podgrupy o podobnych cechach i zachowaniach;

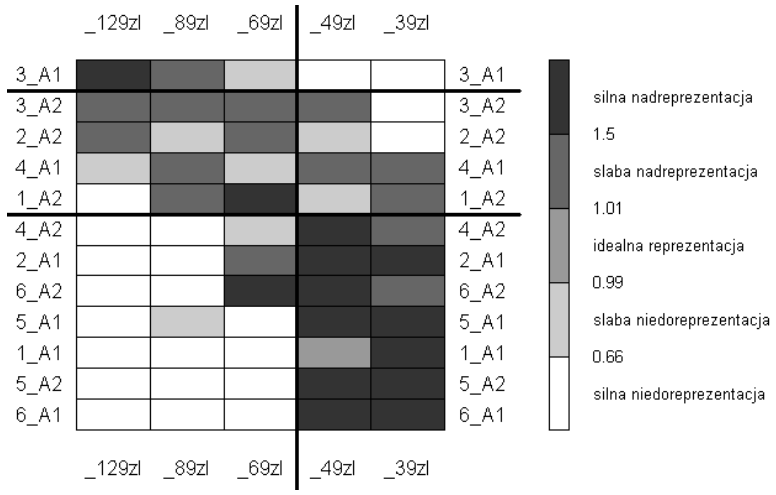
c) podgrupy te są mniej liczne i jednocześnie bardziej homogeniczne niż grupa jako całość;

d) działanie na rzecz mniejszej i podobnej podgrupy nabywców w odróżnieniu od dużych, zróżnicowanych grup nabywców jest łatwiejsze i bardziej skuteczne.

Do wyznaczenia segmentów posłużyła dostępna w GradeStat procedura analizy skupień. Proces wyszukiwania skupień dla danych był przeprowadzony

według zaimplementowanego algorytmu krzywa korelacji/rozkład brzegowy dla zmiennych [Ciok 2004], [Kowalczyk i in. 2004]. Efektem tej analizy jest rys. 6, który prezentuje sześć najbardziej jednorodnych segmentów ze względu na cenę pakietu oraz kategorię wiekową i dochodową.

Rysunek 6. Mapa nadreprezentacji po przekształceniach



Źródło: opracowanie własne

Wyróżnione na rys. 6 segmenty mogą stanowić podstawę do różnicowania oferty rynkowej dla określonych grup klientów. Dzięki czemu można zaproponować klientom te pakiety, z których będą najbardziej zadowoleni. Ponadto, w marketingu bezpośrednim, jest to niezwykle cenna informacja, dzięki której łatwiej jest przeprowadzać kampanie akwizycyjne. Przykładowo, biorąc pod uwagę wiek klienta i dotychczasową informację o wydatkach (kategoria dochodowa), łatwej jest zaproponować określony pakiet, który wpisuje się w charakterystykę danego segmentu. Trudno oczekiwać, że klient z piątej lub szóstej grupy wiekowej będzie zainteresowany pakietem dostępu do Internetu w najwyższej cenie. Działanie akwizycyjne prowadzone zgodnie z przyjętą mapą segmentów, pozwolą zaoszczędzić czas oraz środki, co w efekcie może przełożyć się na większy zwrot nakładów z kampanii.

PODSUMOWANIE

Otrzymane w badaniu rezultaty pozwalają na sformułowanie trzech zasadniczych wniosków.

1. Podejście gradacyjne, poprzez identyfikację zależności między zmiennymi i obiektami, jest pomocnym narzędziem do opisu danych wielowymiarowych. Dużą zaletą jest możliwość wizualizacji, przez co porównywanie i ocena podobieństwa w danych jest łatwiejsza.

2. Atrakcyjność oferty dostępu do mobilnego internetu różni się w zależności od grupy wiekowej klientów oraz ich kategorii dochodowej. Można zauważyć, że pakiety najtańsze wybierane są przez klientów najmłodszych i najstarszych. Z kolei pakiety najdroższe są nabywane przez klientów z grupy wiekowej 26-35 lat..
3. Wyznaczone w badaniu segmenty nabywców o podobnych cechach, mogą być podstawą do różnicowania oferty rynkowej dla określonych grup klientów. W efekcie, tak zidentyfikowane grupy mogą zostać wykorzystane do tworzenia odpowiednio dobranych strategii marketingowych w obszarze akwizycji nowych klientów.

BIBLIOGRAFIA

Borkowski B., Szczesny W. (2005) Metody wizualizacji danych wielowymiarowych jako narzędzie syntezy informacji, Roczniki Naukowe SERIA, Tom VII, zeszyt 5, s. 11-15.

Ciok A. (2004) On the number of clusters – a grade approach, IPI PAN, Warszawa.

Kotler P. (1999) Marketing. Analiza, planowanie, wdrażanie i kontrola, Felberg, Warszawa.

Kowalczyk T., Pleszczyńska E., Ruland F. eds. (2004) Grade Models and Methods of Data Analysis. With applications for the Analysis of Data Population, Studies in Fuzziness and Soft Computing, vol. 151, Springer Verlag, Berlin - Heidelberg - New York.

Książyk J., Matyja O., Pleszczyńska E., Wiech M. (2005) Analiza danych medycznych i demograficznych przy użyciu programu GradeStat. Instytut Podstaw Informatyki PAN i Instytut Pomnik – Centrum Zdrowia Dziecka, Warszawa.

Pleszczyńska E., Jarochowska E., Szczesny W. (2006) Wielowymiarowa analiza danych oparta na modelach gradacyjnych z implementacją w programie GradeStat. Materiały VI Krajowej Konferencji Inżynieria Wiedzy i Systemy Ekspertowe, Wrocław.

Szczesny W. (2002) Grade correspondence analysis applied to contingency tables and questionnaire data, Intelligent Data Analysis, vol. 6(1), IOS Press, Amsterdam, s. 17-51.

UKE (2010) Raport o stanie rynku telekomunikacyjnego w Polsce w 2010 r., dostępny na stronie www.uke.gov.pl

GRADE ANALYSIS TO ASSESS THE ATTRACTIVENESS OF INTERNET PACKAGES

Abstract: The aim of the study was to analyze the attractiveness of the mobile internet offer. In this work the grade analysis approach was used to visualize the data and to assess the similarity of internet packages in terms of price and customers age groups. This allowed formulating conclusions that the attractiveness of mobile internet packages varies depending on the customers' age and income category. At the same time research has identified segments of customers with similar characteristics.

Keywords: grade analysis, internet, overrepresentation

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 288 – 301

ZASTOSOWANIE METOD WIELOWYMIAROWEJ ANALIZY DYSKRYMINACYJNEJ DO OCENY KONDYCJI FIRM DEWELOPERSKICH

Monika Zielińska–Sitkiewicz
Katedra Ekonometrii i Statystyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: monika_zielinska_sitkiewicz@sggw.pl

Streszczenie: W pracy przedstawiono próbę ogólnej oceny kondycji wybranych firm deweloperskich notowanych na Warszawskiej Giełdzie Papierów Wartościowych. Wykorzystano modele wielowymiarowej analizy dyskryminacyjnej, opracowane w Polsce. Podjęto również próbę oceny wykorzystanych modeli. Do badania wybrano 17 dużych firm deweloperskich: GTC, ECHO, POLNORD, LCC, GANT, DOM DEVELOPMENT, RONSON, JW. CONSTRUCTION, ROBYG, TUP, TRITON, ALTERCO, CELTIC, EDINVEST, INPRO, JHMDEV oraz WIKANA. Analizie poddano dane finansowe spółek z bilansów i rachunków wyników w okresie od 1 stycznia 2005 do 31 grudnia 2011.

Słowa kluczowe: polski rynek nieruchomości, firma deweloperska, wskaźniki finansowe, analiza dyskryminacyjna

WSTĘP

W ostatnich latach wzrosło zapotrzebowanie na wiarygodną informację o rynku nieruchomości mieszkaniowych. Niewątpliwie jest to efekt światowego kryzysu finansowego oraz nagłego załamania kilkuletnich wzrostowych trendów dla wskaźników określających kondycję rynku mieszkaniowego zarówno w Polsce, jak i na świecie.

Rzetelna informacja o rynku nieruchomości mieszkaniowych ma fundamentalne znaczenie dla bezpieczeństwa relacji pomiędzy uczestnikami rynku obrotu nieruchomościami i wzajemnego zaufania wszystkich jego uczestników.

Kompleksowa analiza kondycji firm „budujących" polski rynek nieruchomości mieszkaniowych jest i będzie użyteczna dla instytucji finansowych

przy badaniu wiarygodności potencjalnych kredytobiorców, jak również może być pomocna przy tworzeniu ratingów deweloperów, które określałyby ich wiarygodność dla potencjalnych nabywców mieszkań.

Opierając się na danych z raportu BGŻ kondycja rynku nieruchomości na świecie pozostaje niekorzystna. Europejski rynek wykazuje spore zróżnicowanie, z dynamiką cen: od -15,8 % w Irlandii w 2011 r. w relacji do 2010 i do 12,8 % w stolicy Estonii w analogicznych latach. Przeważa jednak trend spadkowy, dotykając kraje silnie zadłużone, tj. Hiszpanię, Włochy czy Grecję.

W ostatnich latach na polskim rynku nieruchomości występuje spora nadpodaż mieszkań. W 2011 r. w stosunku do roku ubiegłego odnotowano ok. 65,7% wzrost liczby nowych mieszkań oddanych do użytkowania przez deweloperów. Stanowi to po części efekt odbudowywania podaży lokali z rozpoczynanych po 2009 r., czyli po okresie ograniczania, nowych inwestycji. Zwiększa się również udział mieszkań o małej powierzchni użytkowej, co świadczy o pogłębiającym się trendzie pogarszania się zdolności kredytowej Polaków.

Pomimo, iż notowani na WGPW deweloperzy mogą się pochwalić relatywnie nienajgorszymi wynikami z 2011 r., to branżowy indeks WIG-DEWELOPERZY zanotował blisko 50 % spadek swojej wartości od stycznia do grudnia 2011 r, a w okresie od stycznia do czerwca 2012 obniżył swoją wartość o blisko 8 %.

Zdaniem wielu ekspertów coraz niższe ceny mieszkań i nowe zasady realizowania inwestycji deweloperskich wg nowej „ustawy deweloperskiej" mają w ciągu kolejnych wielu miesięcy zaszkodzić całej branży.

Warto zatem przeanalizować jak faktycznie radzą sobie główne firmy deweloperskie w Polsce i jaki obecny obraz sytuacji na rynku mieszkaniowym można zobaczyć analizując ich dane finansowe.

WYKORZYSTANIE ANALIZY DYSKRYMINACYJNEJ DO OCENY KONDYCJI FIRM DEWELOPERSKICH

Modele dyskryminacyjne mogą być wykorzystywane w celu wczesnej identyfikacji symptomów pogarszającej się sytuacji finansowej przedsiębiorstwa. Ich zadaniem jest generowanie informacji o zagrożeniu efektywności gospodarowania przedsiębiorstwa i utracie zdolności płatniczej [Gołębiowski, Tłaczała 2005].

Podstawowym narzędziem prognostycznym jest funkcja dyskryminacyjna, której postać ogólna jest następująca [Prusak 2005]:

$$Z = a_0 + a_1 X_1 + a_2 X_2 + ... + a_n X_n \tag{1}$$

gdzie:
Z – wartość funkcji dyskryminacyjnej,
X_i, dla i = 1, 2, …, n – zmienne objaśniające (wskaźniki finansowe),

a_i, dla i = 1, 2, ..., n – współczynniki funkcji dyskryminacyjnej,
a_0 – stała.
Interpretacji wyniku dokonuje się na podstawie porównania obliczonej wartości Z funkcji dyskryminacyjnej z wartością graniczną, której poziom ustalany jest przez twórcę modelu. Zaklasyfikowanie podmiotów do rozłącznych grup jest prowadzone na podstawie wartości Z.

Pierwsze modele dyskryminacyjne zostały opracowane dla przedsiębiorstw działających w warunkach ustabilizowanej gospodarki rynkowej [Nowak 2005]. Najbardziej eksponowane w literaturze są modele: E.I. Altmana z 1968 r., W.H. Beavera z 1966 r. oraz G. Weinricha z 1978 r.
W Polsce zainteresowanie tymi modelami i adaptowanie do polskich realiów przypada na lata 90 XX wieku. W kolejnych latach różni autorzy mierzyli się z oszacowaniem najbardziej skutecznych modeli dyskryminacyjnych dla warunków rozwijającej się, polskiej gospodarki.

Artykuł ma na celu zaprezentowanie niektórych polskich modeli opartych na analizie dyskryminacji i wykorzystanie ich do ogólnej oceny kondycji wybranych 17 firm deweloperskich notowanych na Warszawskiej Giełdzie Papierów Wartościowych. Do analizy wybrano następujące modele: Gajdki i Stosa (1996), Hołdy (1996), Mączyńskiej (1994), Sojaka i Stawickiego (1998), Gajdki i Stosa (2003) oraz Mączyńskiej i Zawadzkiego (2006).

Przedstawiane metody zostały opracowane dla rynku polskiego i w literaturze przedmiotu oceniane są wysoko, co do swojej wartości prognostycznej.

Jednym z pierwszych modeli, który został oszacowany dla warunków polskich jest model opracowany przez J. Gajdkę i D. Stosa. Model ten został wyznaczony w wyniku przeprowadzenia badań nad 40 przedsiębiorstwami, z których 20 zaliczono do grupy bankrutów, a 20 do grupy przedsiębiorstw niezagrożonych upadłością. Próbę przedsiębiorstw niezagrożonych bankructwem stanowiły przedsiębiorstwa z branży przemysłowej, budowlanej i handlowej, które były notowane na giełdzie. Brak jest informacji z jakich branż pochodziły przedsiębiorstwa uznane za bankrutów. Analiza polegała na wyznaczeniu 20 wskaźników finansowych, obliczonych na podstawie sprawozdań finansowych z lat 1994-1995. Spośród tych 20 wskaźników do prezentowanego modelu zakwalifikowano 5: [Prusak 2004]
X_1 = przychody netto ze sprzedaży/aktywa ogółem(średnia wartość w roku)
X_2 = zobowiązania krótkoterminowe(średnia wartość w roku) /koszty wytworzenia produkcji sprzedanej *365
X_3 = zysk netto/aktywa ogółem(średnia wartość w roku)
X_4 = zysk brutto/przychody netto ze sprzedaży
X_5 = zobowiązania ogółem/aktywa ogółem

Wartość graniczna, dla funkcji dyskryminacyjnej postaci:

$$Z = 0,7732059 - 0,0856425 \cdot X_1 + 0,000774 \cdot X_2 + \\ + 0,9220985 \cdot X_3 + 0,6535995 \cdot X_4 - 0,594687 \cdot X_5 \quad (2)$$

na podstawie której dokonywana jest ocena zagrożenia upadłością wynosi **0,45**. Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 1.

Tabela 1. Wyniki modelu Gajdki i Stosa (1996) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	-0,4432	-0,0684	-0,2186	0,1105	0,3875	0,3426
POLNORD	-0,0015	0,0587	0,0788	0,0906	0,0672	0,0614
GTC	-1,7955	-0,1420	-0,9696	1,3167	2,2890	1,8031
DOMDEV	0,0344	0,0128	0,1455	0,2608	0,5467	0,6194
ECHO	0,0178	-0,0613	-0,0793	-0,0606	0,4600	0,4507
GANT	-0,1158	-0,1619	0,0225	0,0442	0,4753	0,7064
JWCONSTR	-0,1025	0,0537	0,0092	-0,0399	0,2747	*
LCCORP	0,2420	0,2569	-0,3968	0,0720	2,3167	*
RONSON	-0,1222	0,1309	-0,0151	-0,0836	0,1515	*
ROBYG	-0,2134	0,0359	-0,2967	-0,1372	*	*
TUP	-0,4961	0,2949	-0,0793	-0,1345	1,1713	0,4960
TRITON	-0,1249	-0,1279	-5,1171	0,3637	-0,0144	2,4211
ALTERCO	0,0275	0,2795	48,3182	*	*	*
CELTIC	-4,4808	-0,0022	*	*	*	*
EDINVEST	0,2793	0,3177	*	*	*	*
INPRO	0,1629	0,1503	*	*	*	*
JHMDEV	-0,0804	*	*	*	*	*
WIKANA	0,1298	0,2981	0,6057	*	*	*
ZAGROŻONE Z<=0,45				BEZPIECZNE Z > 0,45		

Źródło: obliczenia własne

Tak złe wyniki dla całej grupy analizowanych firm deweloperskich wynikają po pierwsze z grupy wskaźników wybranych do modelu, po drugie ze specyfiki branży. Aż trzy z użytych wskaźników odwołują się do wartości aktywów, które w firmach deweloperskich są wyjątkowo duże ze względu na dwie pozycje: zapasy oraz nieruchomości inwestycyjne.[1] Decydującym w ocenie jest również wskaźnik zadłużenia (X_5). Dlatego też cała grupa analizowanych firm w ostatnim okresie znalazła się w kategorii podmiotów zagrożonych bankructwem.

Kolejny model opracował A. Hołda. Okres badań obejmował lata 1993–1996. Badanych było 40 przedsiębiorstw niezagrożonych bankructwem oraz 40, które zbankrutowały. Podmioty te zawierały się w grupie 45-74 EKD[2] (m.in. budownictwo, hotele i restauracje, pośrednictwo finansowe i inne). Zostały ułożone branżowo i poddane analizie wskaźnikowej. W końcowej postaci modelu liczbę wskaźników zredukowano do 5:

X_1 = majątek obrotowy / zobowiązania krótkoterminowe

[1] Opis obu pozycji charakterystycznych dla branży znajduje się w PODSUMOWANIU.

[2] Europejska Klasyfikacja Działalności

X_2 = (kapitał obcy / pasywa ogółem) * 100
X_3 = (zysk netto / średnia wartość aktywów ogółem) * 100
X_4 = (średnia wartość zobowiązań bieżących / koszt wytworzenia sprzedanych produktów, towarów i materiałów) * 360
X_5 = przychody ogółem / średnia wartość aktywów ogółem
Jednostki gospodarcze z wartością funkcji dyskryminacyjnej postaci:

$$Z = 0,605 + 0,681 \cdot X_1 - 0,0196 \cdot X_2 + 0,00969 \cdot X_3 + \\ + 0,000672 \cdot X_4 + 0,157 \cdot X_5 \quad (3)$$

powyżej 0,1 określane są, jako te z małym prawdopodobieństwem bankructwa, a te poniżej -0,3 to te zagrożone upadłością z dużym prawdopodobieństwem. Obszar niepewności, czyli przedział wartość funkcji Z narażonych na znaczne prawdopodobieństwo błędnych klasyfikacji został określony na poziomie od -0,3 do 0,1. [Hołda 2001]
Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 2.

Tabela 2. Wyniki modelu Hołdy (1996) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	**1,3932**	**1,9276**	**1,6823**	**1,7586**	**2,1580**	**1,3034**
POLNORD	2,2181	3,1301	2,6067	3,6826	3,7372	-0,1502
GTC	1,0228	1,9372	2,0319	2,3905	3,8272	2,5486
DOMDEV	1,6323	2,3266	2,3794	3,8988	3,6897	2,2068
ECHO	1,5714	1,4685	1,1395	1,3897	1,7311	1,1873
GANT	1,0518	1,6229	1,6532	1,5646	1,8337	2,3872
JWCONSTR	0,8947	0,8487	0,5564	0,2042	0,9791	*
LCCORP	7,2214	7,5085	5,1576	2,9848	5,0334	*
RONSON	2,7667	2,4503	2,4295	2,6406	2,3837	*
ROBYG	1,9007	4,1312	2,4494	3,3890	*	*
TUP	1,0781	2,2524	6,7781	6,8355	1,9113	1,4113
TRITON	3,8224	1,1318	9,9733	1,6987	1,7997	0,6969
ALTERCO	3,5796	2,2526	1,9056	*	*	*
CELTIC	1,5269	5,1703	*	*	*	*
EDINVEST	2,5262	1,1523	*	*	*	*
INPRO	1,4484	1,2414	*	*	*	*
JHMDEV	2,4613	*	*	*	*	*
WIKANA	1,1617	1,0539	1,4021	*	*	*
ZAGROŻONE **Z < – 0,3**	NIEOKREŚLONE **– 0,3 <= Z <= 0,1**				BEZPIECZNE **Z > 0,1**	

Źródło: obliczenia własne

W ocenie deweloperów decydujące w funkcji Hołdy były wskaźniki (X_1) (podstawowy wskaźnik płynności) oraz (X_4) (wskaźnik obrotu zobowiązań), natomiast wskaźniki zyskowności majątku (X_3) oraz rotacji majątku (X_5) miały niemal zerowe wartości. Ogólnie firmy zostały zweryfikowane pozytywnie, bowiem pozycje zapasy i nieruchomości inwestycyjne nie były eksponowane w użytych wskaźnikach.

Badania nad modelem dyskryminacyjnym dostosowanym do polskich warunków prowadziła również E. Mączyńska. Dokonała ona adaptacji funkcji O. Jacobsa, służącej ocenie wiarygodności kredytowej podmiotów, dokonywanej przez zagraniczne instytucje finansowe oraz banki. Postać funkcji opracowana została relatywnie dawno, ale jej wartości prognostyczne są wysokie. Przyjęte w modelu wagi uwzględniały znaczenie poszczególnych wskaźników dla ogólnej kondycji finansowej przedsiębiorstwa. [Mączyńska, 1994]
Wskaźniki finansowe użyte do konstrukcji modelu mają postać:
X_1 = (wynik brutto + amortyzacja) /zobowiązania ogółem
X_2 = aktywa ogółem/zobowiązania ogółem
X_3 = wynik brutto/aktywa ogółem
X_4 = wynik brutto/przychody ze sprzedaży
X_5 = zapasy/przychody ze sprzedaży
X_6 = przychody ze sprzedaży/aktywa ogółem
Dokonując interpretacji funkcji dyskryminacyjnej:

$$Z = 1{,}5 \cdot X_1 + 0{,}08 \cdot X_2 + 10 \cdot X_3 + 5 \cdot X_4 - 0{,}3 \cdot X_5 + 0{,}1 \cdot X_6 \quad (4)$$

należy opierać się na następujących zasadach: Z<=0 to przedsiębiorstwo zagrożone upadłością w perspektywie 1 roku, jeśli 0<Z<1 przedsiębiorstwo jest słabe, ale niezagrożone upadłością, jeżeli 1<=Z<=2 to przedsiębiorstwo jest dość dobre, a Z>=2 oznacza przedsiębiorstwo bardzo dobre.
Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 3.

Tabela 3. Wyniki modelu Mączyńskiej (1994) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	**-2,5316**	**0,9969**	**-0,5091**	**2,7175**	**5,3170**	**5,0243**
POLNORD	0,3308	0,5871	0,9322	1,2648	-0,1268	-1,9526
GTC	-12,7231	1,2308	-6,1459	12,5433	20,7889	17,7549
DOMDEV	0,9516	0,5641	1,3785	2,3892	3,9914	3,4797
ECHO	2,6809	1,8007	1,7271	1,8845	6,5745	6,7535
GANT	0,0985	-0,6177	1,0509	1,5412	5,1738	8,0400
JWCONSTR	0,5750	1,7587	1,5085	1,0940	2,7002	*
LCCORP	2,2876	2,3262	-4,2580	0,6713	14,9333	*
RONSON	-1,3989	0,9883	0,0740	-0,5548	2,0447	*
ROBYG	-1,3276	0,8513	-1,8524	-0,5277	*	*
TUP	-3,4288	4,2816	0,0410	-0,4753	12,0238	6,4066
TRITON	-1,1679	-0,5172	-149,5060	3,9397	-0,4401	1,0920
ALTERCO	0,6334	1,5364	290,1056	*	*	*
CELTIC	-41,0098	0,5070	*	*	*	*
EDINVEST	3,2081	1,3360	*	*	*	*
INPRO	1,3833	1,5978	*	*	*	*
JHMDEV	0,0316	*	*	*	*	*
WIKANA	1,5424	2,3309	1,8342	*	*	*
BARDZO DOBRE S > 2		DOBRE 1 < S <= 2		SŁABE 0 < S <= 1		ZŁE S < 0

Źródło: obliczenia własne

W modelu Mączyńskiej wskaźniki, które najsilniej decydowały o ocenie danej firmy to rentowność obrotu (X_4) oraz odwrotność rotacji zapasów (X_5).

Widoczne jest duże zróżnicowanie w ocenie przedsiębiorstw. Silnie skrajne wyniki, które pojawiają się w tej tablicy wynikają z pewnych wartości liczbowych, które znalazły się w raportach finansowych firm pod odpowiednia kategorią, ale nie zostały w żaden sposób dodatkowo opisane.

Model Sojaka i Stawickiego w przeciwieństwie do pozostałych, składa się z trzech funkcji klasyfikacyjnych dla: przedsiębiorstw dobrych, średnich i zagrożonych upadkiem. Badacze objęli analizą 58 przedsiębiorstw, obliczając 20 wskaźników finansowych na podstawie informacji pochodzących z 1998 roku. Autorzy wytypowali następnie (za pomocą techniki zwanej analizą skupień) spośród 20 wskaźników 11, które w sposób najlepszy dyskryminują grupy. Następnie spośród tych 11 wskaźników wybrano 7 najlepszych [Prusak 2004]:

X_1 = (wynik finansowy netto/przeciętna wartość majątku obrotowego) · 100;
X_2 = (aktywa obrotowe - zapasy - rozliczenia m/o czynne) /zobowiązania krótkoterminowe;
X_3 = przeciętny kapitał pracujący/przeciętna wartość aktywów;
X_4 = (wynik finansowy netto/przeciętna wartość kapitału własnego) · 100;
X_5 = (wynik finansowy netto/przeciętna wartość majątku trwałego) · 100;
X_6 = (wynik finansowy netto + odsetki od kapitałów obcych - podatek dochodowy) /przeciętna wartość aktywów;
X7 = aktywa obrotowe/zobowiązania krótkoterminowe;
i na ich podstawie zbudowano trzy następujące funkcje klasyfikacyjne:

$$Przedsiębiorstwo_{złe} = -0{,}1144 \cdot X_1 + 0{,}5178 \cdot X_2 - 20{,}4475 \cdot X3 + \\ - 0{,}0661 \cdot X4 + 0{,}0663 \cdot X5 - 50{,}461 \cdot X6 + 1{,}8358 \cdot X7 - 11{,}6499 \quad (5)$$

$$Przedsiębiorstwo_{średnie} = -0{,}0586 \cdot X_1 - 3{,}3608 \cdot X_2 + 10{,}7088 \cdot X_3 + \\ + 0{,}01455 \cdot X_4 - 0{,}066 \cdot X_5 + 4{,}5837 \cdot X_6 + 0{,}24329 \cdot X_7 - 2{,}3393 \quad (6)$$

$$Przedsiębiorstwo_{dobre} = -0{,}0153 \cdot X_1 + 2{,}0482 \cdot X_2 + 9{,}637 \cdot X_3 + \\ + 0{,}1714 \cdot X_4 - 0{,}0091 \cdot X_5 - 15{,}78 \cdot X_6 - 0{,}0018 \cdot X_7 - 5{,}992 \quad (7)$$

O przydziale do odpowiedniej grupy przedsiębiorstw decyduje największy dodatni współczynnik. Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 4.

Grupa wskaźników wybrana do powyższych funkcji koncentruje się na rentowności, a w związku z tym bezpośrednio lub pośrednio odwołuje się do zapasów, które u dużych deweloperów stanowią w większości bank ziemi. Zatem można zauważyć, że modele te gorzej oceniły kondycję dużych firm, natomiast

zdecydowanie lepiej małych deweloperów, którzy nadwyżki finansowe lokowali w bieżącą działalność, a nie w powiększanie majątku.

Tabela 4. Wyniki modeli Sojaka i Stawickiego (1998) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	Max<0	Max<0	Max<0	Max<0	DOBRE	DOBRE
POLNORD	ŚREDNIE	ŚREDNIE	ŚREDNIE	DOBRE	DOBRE	ZŁE
GTC	ZŁE	Max<0	Max<0	Max<0	DOBRE	DOBRE
DOMDEV	ZŁE	ZŁE	ZŁE	ZŁE	ZŁE	ZŁE
ECHO	Max<0	Max<0	Max<0	Max<0	Max<0	Max<0
GANT	Max<0	ŚREDNIE	ŚREDNIE	Max<0	DOBRE	DOBRE
JWCONSTR	Max<0	Max<0	Max<0	Max<0	DOBRE	*
LCCORP	DOBRE	DOBRE	ŚREDNIE	ŚREDNIE	DOBRE	*
RONSON	ŚREDNIE	DOBRE	Max<0	ZŁE	DOBRE	*
ROBYG	ŚREDNIE	DOBRE	ŚREDNIE	ŚREDNIE	*	*
TUP	ŚREDNIE	DOBRE	DOBRE	DOBRE	DOBRE	DOBRE
TRITON	ŚREDNIE	ŚREDNIE	ŚREDNIE	DOBRE	Max<0	Max<0
ALTERCO	DOBRE	DOBRE	Max<0	*	*	*
CELTIC	ZŁE	Max<0	*	*	*	*
EDINVEST	DOBRE	Max<0	*	*	*	*
INPRO	DOBRE	DOBRE	*	*	*	*
JHMDEV	ŚREDNIE	*	*	*	*	*
WIKANA	Max<0	Max<0	DOBRE	*	*	*

Źródło: obliczenia własne

Kolejny model Gajdki i Stosa z 2003 r. powstał z myślą o ocenie kondycji finansowej przedsiębiorstw notowanych na warszawskiej Giełdzie Papierów Wartościowych. System ten opracowano na podstawie zbilansowanej próby składającej się z 34 obiektów: 17 przedsiębiorstw upadłych, którym przyporządkowano 17 jednostek „zdrowych" o podobnym profilu działalności.[Kisielińska, Waszkowski 2010]

Oszacowany liniowy model dyskryminacyjny przyjął następującą postać:

$$Z = -0{,}3342 - 0{,}000500 \cdot X_1 + 2{,}055200 \cdot X_2 + 1{,}726000 \cdot X_3 + \\ + 0{,}115500 \cdot X_4 \tag{8}$$

Wykorzystano w nim cztery wskaźniki finansowe:

X_1 = średnia wartość zobowiązań; krótkoterminowych / koszt wytworzenia produkcji sprzedanej * 360 dni;

X_2 = zysk netto / średnia wartość aktywów w roku;

X_3 = zysk brutto / przychody netto ze sprzedaży;

X_4 = aktywa ogółem / zobowiązania ogółem;

Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 5. Wartością graniczną dla modelu jest zero. Dla $Z < 0$ przedsiębiorstwo klasyfikowane jest do grupy zagrożonych upadkiem, dla $Z > 0$ do grupy jednostek o dobrym standingu. Obszar niepewności w tym modelu mieści się w przedziale <–0,49; 0,49>.

Decydującym wskaźnikiem w ocenie kondycji deweloperów był wskaźnik (X_4), czyli odwrotność stopy zadłużenia. Natomiast wskaźniki: obrotu zobowiązań (X_1) oraz zyskowności majątku (X_2) dawały niemal zerowe wyniki.

Tabela 5. Wyniki modelu Gajdki i Stosa (2003) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	**2,3646**	**3,6555**	**3,0748**	**4,0738**	**5,1645**	**4,6715**
POLNORD	4,5471	4,5961	4,6933	4,2891	5,3784	1,6868
GTC	-1,9952	3,3598	0,7549	8,1998	11,6324	10,5764
DOMDEV	3,4195	3,5350	3,6232	4,0659	4,4723	4,1629
ECHO	3,8099	3,6229	3,5196	3,4915	5,2853	5,2148
GANT	3,3203	3,4275	3,7262	3,8798	5,3720	6,2427
JWCONSTR	2,8985	2,9695	2,7167	2,3952	3,1953	*
LCCORP	6,8902	8,1411	5,0111	4,8991	12,1271	*
RONSON	4,0300	4,8198	4,0089	3,5681	3,9626	*
ROBYG	3,1437	3,9734	2,5738	2,6587		*
TUP	3,5389	6,3768	5,2169	4,9445	8,8290	6,2458
TRITON	3,2877	2,8196	-12,2798	4,9308	4,3116	2,5977
ALTERCO	4,0170	6,7213	*	*	*	*
CELTIC	-5,1358	13,1387	*	*	*	*
EDINVEST	5,3583	3,4460	*	*	*	*
INPRO	3,9562	3,2574	*	*	*	*
JHMDEV	3,4832	*	*	*	*	*
WIKANA	3,3491	3,7906	3,9198	*	*	*
ZAGROŻONE Z < 0				BEZPIECZNE Z >= 0		

Źródło: obliczenia własne

Ostatni z prezentowanych polskich modeli to model E. Mączyńskiej i M. Zawadzkiego z 2006 r. Autorzy przeprowadzili badania na zbilansowanej próbie 80 spółek notowanych na GPW w Warszawie, wykorzystując sprawozdania finansowe z lat 1997–2001 oraz obliczone na ich podstawie wskaźniki finansowe. W badaniach wykorzystano 45 wskaźników charakteryzujących rentowność, płynność, poziom zadłużenia, sprawność operacyjną oraz dynamikę wzrostu przedsiębiorstw. Do prezentowanego modelu wybrano cztery wskaźniki:
X_1 = (wynik brutto + amortyzacja) /zobowiązania ogółem
X_2 = aktywa ogółem/zobowiązania ogółem
X_3 = wynik brutto/aktywa ogółem
X_4 = wynik brutto/przychody ze sprzedaży
a funkcja dyskryminacyjna ma następująca postać:

$$Z = 9,498X_1 + 3,566X_2 + 2,903X_3 + 0,452X_4 - 1,498 \quad (9)$$

Punktem granicznym dla tego modelu jest wartość zero, a zagrożenie występuje, jeśli $Z < 0$.
Wyniki klasyfikacji dla analizowanych deweloperów prezentowane są w tablicy 6.
Podobnie jak w poprzednim modelu decydującym wskaźnikiem w ocenie kondycji deweloperów był wskaźnik odwrotności stopy zadłużenia (X_2). W niektórych szczególnych przypadkach podmiotów, u których negatywna wycena

nieruchomości inwestycyjnych miała duży wpływ na wynik finansowy firmy, odwołujące się do tegoż wyniku wskaźniki (X_1) oraz (X_4) wpływały na ogólną złą ocenę firmy.

Tabela 6. Wyniki modelu Mączyńskiej i Zawadzkiego (2006) dla spółek deweloperskich.

Dane na dzień:	2011-12-31	2010-12-31	2009-12-31	2008-12-31	2007-12-31	2006-12-31
RAZEM	3,8280	5,8827	4,8754	6,4398	9,0073	9,1300
POLNORD	7,8109	7,2934	7,7850	7,7080	9,4807	1,2915
GTC	0,4313	4,7413	2,7054	7,9992	11,6029	14,4030
DOMDEV	5,9828	6,1690	6,9384	8,5288	11,5581	10,4815
ECHO	5,2231	5,3728	5,2286	5,1084	7,7065	7,8775
GANT	5,3702	5,4780	6,6491	6,7463	8,6262	12,3544
JWCONSTR	4,7352	5,4890	4,8553	3,8992	6,5135	*
LCCORP	12,6037	15,9962	9,2024	7,9780	11,1042	*
RONSON	6,8113	9,3721	6,7377	5,5515	6,9425	*
ROBYG	5,0037	7,1325	3,6464	3,8729	*	*
TUP	6,5537	10,8718	9,3151	8,8570	13,1765	12,0708
TRITON	5,1151	4,3557	0,2769	11,1885	7,5481	5,2444
ALTERCO	6,9745	13,4071	44,6815	*	*	*
CELTIC	-7,5410	25,6817	*	*	*	*
EDINVEST	12,7355	6,5144	*	*	*	*
INPRO	7,6722	6,2745	*	*	*	*
JHMDEV	5,8136	*	*	*	*	*
WIKANA	6,1827	7,9419	8,4458	*	*	*
ZAGROŻONE Z < 0				BEZPIECZNE Z >= 0		

Źródło: obliczenia własne

PODSUMOWANIE

Zasady wyznaczane przez rynek determinujące działalność przedsiębiorstw są stale weryfikowane przez otoczenie zarówno mikro, jak i makroekonomiczne. Wczesnym wykryciem pogorszania się sytuacji firmy, w skrajnym przypadku ryzyka bankructwa, są zainteresowane podmioty bezpośrednio związane z przedsiębiorstwem, między innymi: właściciele, akcjonariusze, kooperanci, jak również instytucje finansujące działalność firmy.

W tym celu zostały opracowane liczne wskaźniki finansowe, obliczane na podstawie danych finansowych firmy, które stanowią podstawę do oceny jej sytuacji finansowej. Niestety zarówno trudność właściwego doboru, jak i złożoność interpretacji oraz znaczna liczba wskaźników analizy finansowej, utrudniają dokonanie jednoznacznej oceny sytuacji finansowej danego przedsiębiorstwa. [Lichota 2009]

Analiza zagrożenia finansowego badanych spółek wskazuje, że zastosowanie poszczególnych modeli dyskryminacji nie daje gwarancji jednoznacznej oceny ich kondycji ekonomicznej.

W funkcjach dyskryminacyjnych wykorzystuje się wskaźniki finansowe, których jest zbyt wiele, aby wybrać jedynie tych kilka najważniejszych dla danego przedsiębiorstwa. Zdarza się, że w modelu wykorzystane będą wskaźniki, które w

zestawieniu z resztą będą generować pozytywny lub negatywny wpływ na ocenę kondycji firmy. Może to całkowicie zmieniać pogląd na kondycję finansową tego samego przedsiębiorstwa.

Każda z analizowanych funkcji dyskryminacyjnych oparta jest na innym zestawie wskaźników i lepiej, bądź gorzej analizuje stan finansów danego dewelopera. Co jest bardzo charakterystyczne dla branży i powinno być brane pod uwagę w ocenie wskaźnikowej kondycji firm deweloperskich to dwie pozycje bilansowe**: zapasy** i **nieruchomości inwestycyjne**.

Firmy developerskie są w swej istocie firmami usługowymi. Rzadko kiedy same cokolwiek wytwarzają i poza nielicznymi przypadkami nie zajmują się handlem zakupionych mieszkań. W zasadzie cała ich „produkcja" polega na tym, że zlecają firmom budowlanym wybudowanie na swoim gruncie budynków. Tak wybudowane produkty są sprzedawane jako „produkty" dewelopera.

Często grunt, w zależności od przyjętych zasad rachunkowości w spółce, jest wykazywany początkowo w **zapasach**, jako 'materiały', by potem wraz z ewentualnym rozpoczęciem lub zakończeniem budowy, przejść odpowiednio do 'produkcji w toku' lub 'produktów gotowych'.

Czasem ze względu na dopuszczaną przez firmę możliwość odsprzedaży gruntów, kwalifikowane one są do 'aktywów trwałych', jako **nieruchomości inwestycyjne**.

Niektóre przedsiębiorstwa w latach prosperity tworzyły tzw. „bank ziemi" angażując w to ogromne środki finansowe. Niektóre „banki ziemi", np. w Polnord SA są tak duże, że nawet sama spółka kwalifikuje część gruntów do dziwnej z punku widzenia rachunkowości kategorii 'zapasy długoterminowe'.

Niektóre spółki wykazywały się wielką zapobiegliwością w gromadzeniu zapasów ziemi, np. Robyg (2009 rok) posiadał grunty na 85 lat „produkcji", by po dwóch latach i częściowej sprzedaży gruntów (ze stratą), zejść do skromnych 22 lat.

Z drugiej strony, są i takie spółki, które mają minimalny bank ziemi, niepozwalający na domniemanie zamiaru kontynuowania działalności deweloperskiej, np. EDINVEST (2011 rok) – zapas ziemi na 10 miesięcy.

Niezależnie od opinii i ocen rozmiarów działania poszczególnych spółek w budowaniu „banków ziemi", należy stwierdzić, że wartość zapasów gruntów stanowi wyjątkowo duży element sumy aktywów obrotowych, gdyż nawet przy bardzo ostrożnym nabywaniu gruntów, przy faktycznych procesach technologicznych i administracyjno-prawnych, spółka powinna mieć ich zapas, na co najmniej 3,5-4 lata (przy założeniu kontynuowania tej działalności).

Z kolei **Nieruchomości inwestycyjne** to takie nieruchomości, które właściciel utrzymuje licząc na wzrost ich wartości, celem sprzedaży z zyskiem innym inwestorom. Z części nieruchomości inwestycyjnych właściciele czerpią bezpośrednie korzyści z tytułu np. dzierżaw (biurowce, centra handlowe). Część nieruchomości inwestycyjnych nie jest użytkowana w ogóle i czeka na lepsze czasy.

Cechą charakterystyczną pozycji Nieruchomości Inwestycyjnych jest wyjątkowe podejście do ich wyceny, zwłaszcza wtedy, gdy wycena zwiększa ich wartość.

Co ciekawe nie ma jasnych, narzuconych przez prawo reguł dokonywania wyceny. Można jej dokonać, ale nie jest obowiązkowa. Ponadto w przeciwieństwie do innych wycen zwiększających wartość aktywów trwałych (np. środków trwałych, akcji), wynik wyceny wprost wpływa na wynik finansowy, a nie jak w innych przypadkach na Kapitały z wyceny.
Stąd olbrzymia pokusa i możliwości, by tak dobrać metodologię wyceny, tak dobrać wyceniającego, tak dobrać cenę za wycenę, by efekt wyceny poprawiał wynik finansowy, z którego Zarząd spółki jest rozliczany przez akcjonariuszy i tzw. rynek.
Nawet specjaliści zajmujący się oceną spółek zawsze analizują wynik finansowy, natomiast niemal nigdy nie zagłębiają się w pozycję bilansową Kapitały z wyceny, o ile taka pozycja jest w ogóle w sprawozdaniu finansowym wyszczególniona, a nie skumulowana z innymi kapitałami.
Analiza danych za lata 2008 - 2011, kiedy to nawet przy sporządzaniu sprawozdania finansowego za 2008 rok (około marca/kwietnia 2009 roku) było wiadomo, że jest coś takiego jak kryzys światowy, deweloperzy wykazują zaskakująco odmienne wyceny nieruchomości, zwłaszcza, że wycen dokonywano w latach o podobnej, raczej niskiej koniunkturze rynku nieruchomości. W latach 2008 - 2011 (dla ogółu spółek) wyceny roczne wahały się od optymistycznego **1.444 mld PLN**, do pesymistycznej straty **1.053 mld PLN**.
Niejednokrotnie odnosi się wrażenie, że dodatnia wycena nieruchomości inwestycyjnych stanowi swego rodzaju "koło ratunkowe" dla katastrofalnego wyniku finansowego, by w czasach, gdy wynik ten jest lepszy dokonać wyceny korygującej wcześniejsze poprawienie wyniku.
Także metodologia wycen opisana w sprawozdaniach finansowych znacząco się różni. Czasem sama wycena wygląda na po prostu rzetelną, a dopiero przestudiowanie przestrzeni czasowej i kosztowej nabytych, sąsiadujących z sobą nieruchomości wskazuje jednoznacznie na działalność dla pozoru (nieważną z mocy prawa na podstawie KC), mającą na celu zmanipulowanie danymi w celu wprowadzenia w błąd odbiorców sprawozdań finansowych. Zwłaszcza, gdy dzieje się to przeddzień emisji nowych akcji na GPW, a wynik finansowy z realnej działalności spółki trzeba ratować za wszelką cenę.

WNIOSKI

- Podsumowując wyniki 17 spółek z ich sprawozdań finansowych należy stwierdzić, iż kondycja większości deweloperów jest nienajgorsza, choć niewątpliwie panujący na rynku kryzys jest odzwierciedlony w danych finansowych analizowanych spółek.
- Poddając analizie dane deweloperów, szczególnie dotyczy to dużych firm, należy z przenikliwą dokładnością przestudiować wyżej wymienione pozycje bilansowe i uwzględnić specyfikę branży w wyciąganych wnioskach.

- We wszystkich wskaźnikach odwołujących się bezpośrednio lub pośrednio do kategorii **Zapasy**, czyli Zapasy, Aktywa obrotowe, Aktywa trwałe (w części dotyczącej 'nieruchomości inwestycyjnych') i Aktywa, wystąpią istotne różnice w ocenie sytuacji dla działalności deweloperskiej w stosunku do innych branż gospodarczych. Powinno brać się to pod uwagę poszukując gotowych metod analizujących kondycję firmy, takich jak np. modele dyskryminacji.
- Co do zaprezentowanych sześciu modeli dyskryminacyjnych to trudno w pełni i rzetelnie ocenić ich skuteczność dla branży deweloperskiej, nie posiadając spółki, w analizowanej grupie przedsiębiorstw, która ogłosiła upadłość. Porównując jedynie wyniki analizy wskaźnikowej, jaką przeprowadziła autorka, z podanymi w pracy ocenami funkcji dyskryminacyjnych dla spółek deweloperskich można stwierdzić, że najlepiej poradził sobie model Gajdki i Stosa (2003). Faktycznie wskaźniki finansowe obu wyróżnionych przez model w 2011 r. firm: GTC i CELTIC wskazują na potencjalne zagrożenia.
- Niewątpliwie największymi zaletami modeli dyskryminacyjnych są: prostota ich wykorzystania oraz duża skuteczność dla homogenicznych danych.
- Głównymi wadami tych modeli są jednak: nieprzenośność oraz niska skuteczność dla danych niehomogenicznych, co jest wyraźnie zauważalne w zróżnicowanych i nie do końca słusznych ocenach kondycji deweloperów.

BIBLIOGRAFIA

Jajuga K., Walesiak M. (1998) Klasyfikacja i analiza danych. Teoria i zastosowania, AE, Wrocław.

Kisielińska J., Waszkowski A. (2010) Polskie modele do prognozowania bankructwa przedsiębiorstw i ich weryfikacja, EiOGŻ nr 82, Wydawnictwo SGGW, Warszawa.

Lichota W. (2009) Metody wczesnego ostrzegania o zmianach sytuacji finansowej przedsiębiorstw, Wiadomości Statystyczne nr 10, Warszawa.

Gajdka J., Stos D.(1996) Wykorzystanie analizy dyskryminacyjnej w ocenie kondycji finansowej przedsiębiorstw, praca zbiorowa pod red. R. Borowieckiego: Restrukturyzacja w procesie przekształceń i rozwoju przedsiębiorstw, AE, Kraków.

Hamrol M., Chodakowski J. (2008) Prognozowanie zagrożenia finansowego przedsiębiorstwa. Wartość predykcyjna polskich modeli analizy dyskryminacyjnej, Badania Operacyjne i Decyzje nr 3, Wrocław.

Hołda A. (2001) Prognozowanie bankructwa jednostki w warunkach gospodarki polskiej z wykorzystaniem funkcji dyskryminacyjnej Z_H, Rachunkowość nr 5.

Mączyńska E. (1994) Ocena kondycji przedsiębiorstwa (uproszczone metody), Życie gospodarcze nr 38.

Mączyńska E. Zawadzki M. (2006) Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw, Ekonomista nr 2, Warszawa.

Prusak B. (2005) Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstwa, Difin, Warszawa.

Prusak B. (2004) Jak rozpoznać potencjalnego bankruta?, Wydawnictwo Politechniki Gdańskiej, Prace Naukowe Katedry Ekonomii i Zarządzania Przedsiębiorstwem, Tom 3, Gdańsk.

Raport BGŻ (2012) Kwartalna analiza rynku nieruchomości mieszkaniowych w Polsce (I kw. 2012 r.), Warszawa.

Sojak S. Stawicki J. (2001) Wykorzystanie metod taksonomicznych do oceny kondycji ekonomicznej przedsiębiorstw, Zeszyty Teoretyczne Rachunkowości, Tom 3, Warszawa.

Sukiennik M. (2007) Zastosowanie analizy dyskryminacyjnej do oceny stanu finansowego przedsiębiorstw, Referaty II KKMM, Kraków.

THE APPLICATION OF MULTIVARIATE DISCRIMINANT ANALYSIS TO EVALUATE THE CONDITION OF DEVELOPERS

Abstract: The paper contains consideration about usage discriminant analysis in financial audit of developers. 17 major developers, which are listed on the Warsaw Stock Exchange (WSE) were selected for study : GTC, ECHO, POLNORD, LCC, GANT, HOUSE DEVELOPMENT, RONSON, JW. CONSTRUCTION, ROBYG, TUP, TRITON, ALTERCO, CELTIC, EDINVEST, INPRO, JHMDEV and WIKANA. The analysis encompassed financial data from balance sheets and from profit and loss account in the period from January 1, 2005 to December 31, 2011.

Keywords: Polish real estate market, developer, financial ratios, discriminant analysis

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XIII/3, 2012, str. 302 – 311

SELEKCJA KRYTERIÓW OCENY JAKOŚCI SERWISÓW INTERNETOWYCH Z WYKORZYSTANIEM POJEMNOŚCI INFORMACYJNEJ HELLWIGA

Paweł Ziemba
Katedra Inżynierii Systemów Informacyjnych
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: pziemba@wi.zut.edu.pl
Ryszard Budziński
Katedra Zarządzania i Inżynierii Wiedzy, Uniwersytet Szczeciński
e-mail: rbudzinski@wi.zut.edu.pl

Streszczenie: Jakość strony internetowej ma istotny wpływ na postrzeganie przez użytkowników serwisu instytucji, do której dana strona należy. Wobec tego istotną kwestią jest ocena jakości serwisów internetowych oraz wybór kryteriów tej oceny. W pracy zaproponowano metodę selekcji istotnych kryteriów oceny jakości serwisów internetowych, opartą na indywidualnym wskaźniku pojemności informacyjnej Hellwiga.

Słowa kluczowe: ocena jakości serwisów internetowych, selekcja kryteriów, pojemność informacyjna Hellwiga

WSTĘP

Na świecie, wg danych z 2012 roku, z niemal 677 mln stron internetowych korzysta 2,267 mld użytkowników [Internet World Stats, 2012] [Netcraft, 2012]. W sytuacji powstawania „społeczeństwa informacyjnego" czerpiącego informacje i „załatwiającego" sprawy urzędowe za pośrednictwem Internetu, jakość strony internetowej i możliwość „załatwienia" za jej pośrednictwem określonej sprawy bądź przeprowadzenia transakcji handlowych ma istotny wpływ na postrzeganie danej instytucji przez użytkowników serwisu i na liczbę użytkowników. Teza ta w różnej formie pojawia się w wielu publikacjach dotyczących oceny serwisów internetowych różnego rodzaju. Hwang stwierdza, że „zła jakość strony może powodować utratę zarówno potencjalnych sprzedaży jak i powtórnych odwiedzin" [Hwang i in., 2011]. Lin w kontekście jakości portali internetowych zaznacza, że

„użytkownicy mają wiele opcji, które potencjalnie redukują lojalność; oni mogą przenieść się gdziekolwiek indziej (do innego portalu) i nie wrócić" [Lin i in., 2005]. Z kolei, wg Grigoroudisa "nowoczesne strony prezentują znaczne zróżnicowanie cech i złożoność struktury. Podobnie jak we wszystkich systemach informacyjnych, ocena jest aspektem ich rozwoju i działaniem przyczyniającym się do maksymalnego wykorzystania zasobów zainwestowanych w witrynę" [Grigoroudis i in., 2008]. Ocena jakości serwisów internetowych, jest więc bardzo ważna. Należy zadbać o to, aby była ona jak najbardziej wiarygodna i w maksymalnym stopniu odzwierciedlała preferencje użytkowników. Jednocześnie powinno się ją uzyskiwać jak najmniejszym nakładem kosztów i pracy.

PROBLEM BADAWCZY

W literaturze wyróżnia się wiele modeli oceny jakości witryn WWW, m.in.: eQual [Barnes i Vidgen, 2006], Ahn, e-Servqual i liczne modele relacyjne, oparte np. na procedurze AHP [Lin, 2010]. Każdy ze stosowanych modeli różni się od pozostałych: wykorzystywanymi kryteriami opisu jakości, ich liczbą, skalą oceniania oraz metodyką pozyskiwania wartości ocen.

Podstawowy problem dotyczący modeli oceny dotyczy pozyskiwania ocen. Do tego celu można się posłużyć oceną ekspercką, ankietą lub śledzeniem śladów użytkowników. Śledzenie śladów użytkowników jest metodą bardzo kosztowną finansowo, czasowo i obliczeniowo. Wymaga to specjalistycznych urządzeń lub programów, śledzących np. ruch gałek ocznych użytkownika lub wskaźnika myszy. Zebranie danych w ten sposób wymaga znacznego czasu, więc dane do analizy pochodzą z bardzo małej próby badawczej. Dodatkowo przetworzenie tych danych wymaga skomplikowanych obliczeń. Jednocześnie stosowalność tej metody jest ograniczona do badania użyteczności serwisów internetowych, a nie ich jakości. Metody ankietowa i ekspercka wykorzystują dużą ilość kryteriów, więc są kosztowne czasowo. Ponadto ankietowani i eksperci często skarżą się właśnie na dużą liczbę kryteriów oceny, która może powodować niedbałość w ocenie i obniżać poprawność ocen.

Podobny problem dotyczy metodyki tworzenia modeli oceny jakości serwisów internetowych. Podczas tworzenia nowych modeli oceny, ich autorzy bazują zazwyczaj na istniejących modelach badania jakości serwisów, wybierając z nich kryteria oceny, które wydają się odpowiednie do zastosowania w ich autorskiej metodzie. Działanie to jest wykonywane przez autorów zazwyczaj w formie niesformalizowanej, bazującej na analizie literatury i własnych przemyśleniach [Lin, 2010]. Jest to podejście eksperckie, rzadko opierane na jasno sformułowanych metodologiach i analizach, a zazwyczaj charakteryzuje się ono pewną chaotycznością. Dodatkowo istnieje ryzyko, że w tworzonych w ten sposób modelach będą funkcjonowały kryteria o niewielkiej rzeczywistej wadze, które nie są istotne dla użytkowników serwisu internetowego. Występuje tutaj pewien rodzaj

tzw. „przekleństwa wymiarowości", polegający na problemie wyboru spośród wielu wskaźników jakości tylko takich, które są przydatne w ocenie jakości serwisów określonego typu.

Niektórzy autorzy, przy wyborze kryteriów stosują bardziej sformalizowane podejście: na podstawie badania ankietowego dokonują oni grupowania kryteriów poprzez metodę statystyczną, tj. analizę czynnikową. Eliminują oni kryteria, które według wyników analizy nie przynależą do żadnej z uzyskanych grup kryteriów, czy też do żadnego z uzyskanych wymiarów jakości [Ho i Lee, 2007][Ou i Sia, 2010][Hwang i in., 2011]. Wykorzystując analizę czynnikową otrzymują oni nowe klasy kryteriów, zawierające pewne podgrupy oryginalnych kryteriów. Podejście takie jest znacznie bardziej uzasadnione od opisanego wcześniej, jednak również tutaj istnieje pewien problem, dotyczący przypisywania wag kryteriom. Ze względu na wykorzystywane w analizie czynnikowej (analiza składowych głównych) rotacje i skalowania zmiennych, trudno jest ustalić, które z kryteriów są istotne. Przekształcenia te, mogą też wprowadzać znaczące zmiany w uzyskiwanych rezultatach [Yang i in., 2011]. W tym podejściu wagi kryteriów są określane na podstawie wyników ankiet, jako uśrednione wartości wag, podanych przez ankietowanych lub ekspertów. Określanie wag kryteriów na podstawie ankiet (tj. jawnie deklarowanych preferencji użytkowników) może generować duże błędy [Zenebe i in., 2010]. Potwierdzają to badania autorskie, z których wynika, że jawnie deklarowane preferencje użytkowników mogą się różnić od kryteriów, które są rzeczywiście wykorzystywane przez użytkowników w ocenie strony internetowej [Ziemba i Budziński, 2011].

W pracy [Ziemba, 2011] sformułowano procedurę selekcji istotnych kryteriów oceny serwisów i nadawania im wag, pozwalającą sformalizować proces doboru kryteriów dla metod oceny jakości serwisów internetowych. Procedura ta wykorzystuje techniki Data Mining, a konkretnie metody filtracyjne służące do selekcji cech, umożliwiające określenie pewnych cech w zbiorze danych, jako istotnych i odrzuceniu cech nadmiarowych. Efektem zastosowania opracowanej procedury jest zmniejszenie ilości kryteriów oceny jakości serwisów internetowych poprzez wybranie najważniejszych z nich. Jednocześnie zaletą wskazanej procedury jest to, że zredukowany za jej pomocą zbiór kryteriów daje wyniki oceny zbliżone do rzeczywistych ocen użytkowników. Wobec tego sformułowana procedura pozwala na określanie niejawnych preferencji użytkowników, którymi w rzeczywistości kierują się oni oceniając serwis, a które mogą różnić się od preferencji jawnie deklarowanych.

W w/w pracy stwierdzono, że spośród badanych metod filtracyjnych, najlepsze rezultaty daje zastosowanie metody symetrycznej niepewności. Metoda ta polega na badaniu swoistej korelacji między kolejnymi cechami (kryteriami), a poszczególnymi klasami obiektów (oceną końcową). Nie uwzględnia ona jednak wzajemnych korelacji między cechami. Może więc wystąpić sytuacja, gdy dwie cechy są silnie skorelowane wzajemnie między sobą oraz zachodzi ich silna korelacja z klasami obiektów. W przypadku zaistnienia takiej sytuacji, jedna z cech

może okazać się niepotrzebna, gdyż obie cechy będą warunkowały przypisanie do tej samej klasy obiektów. Dodatkowo, podczas oceny jakości na podstawie tych dwóch cech (kryteriów), ich współistnienie w zbiorze kryteriów oceny może zawyżać wynik tejże oceny. Do rozwiązania tego problemu zdecydowano się zaadaptować metodę pojemności informacyjnej Hellwiga.

METODY POJEMNOŚCI INFORMACYJNEJ HELLWIGA

Metoda pojemności informacyjnej Hellwiga wywodzi się z ekonometrii, a polega ona na wyborze takich zmiennych objaśniających, które są silnie skorelowane z zmienną objaśnianą, a słabo skorelowane między sobą. W metodzie Hellwiga rozpatrywane są wszystkie niepuste kombinacje potencjalnych zmiennych objaśniających, tj. 2^m-1 kombinacji, gdzie m oznacza ilość zmiennych objaśniających. Pierwszym krokiem metody jest obliczenie macierzy współczynników korelacji liniowej Pearsona między zmiennymi objaśniającymi zawartymi w l-tej kombinacji, oraz wektora współczynników korelacji Pearsona między zmiennymi objaśniającymi i objaśnianymi zawartymi w l-tej kombinacji. Następnie obliczane są indywidualne wskaźniki pojemności informacyjnej zmiennych dla każdej kombinacji zgodnie z wzorem:

$$h_{lj} = \frac{r_j^2}{1 + \sum_{i=1, i \neq j}^{m} \left| r_{ij} \right|} \tag{5}$$

gdzie: h_{lj} – indywidualna pojemność informacyjna j-tej zmiennej w l-tej kombinacji, r_j – współczynnik korelacji j-tej zmiennej objaśniającej ze zmienną objaśnianą, r_{ij} – współczynnik korelacji między i-tą i j-tą zmienną objaśniającą, m –ilość zmiennych objaśniających w l-tej kombinacji. Następnie dla każdej z kombinacji obliczany jest integralny wskaźnik pojemności informacyjnej według wzoru:

$$H_l = \sum_{j=1}^{m} h_{lj} \tag{6}$$

Do modelu ekonometrycznego zalecane jest wybranie takiej kombinacji zmiennych, dla której uzyskano największą wartość integralnego wskaźnika pojemności informacyjnej [Gnat, 2006].

Jeżeli chodzi o zastosowanie metody Hellwiga, w sytuacji, gdy wykorzystywanych jest wiele cech, istnieje bardzo dużo ich możliwych kombinacji i ich obliczenie jest bardzo czasochłonne (np. dla 22 cech istnieje ponad 4 mln kombinacji). Wobec tego w pracy zdecydowano się zastosować tylko indywidualne pojemności informacyjne obliczane dla jednej kombinacji, zawierającej pełny zbiór cech. Wartości indywidualnych pojemności informacyjnych pozwoliły uzyskać ranking cech, wykorzystany w dalszych badaniach.

PROCEDURA BADAWCZA

Opracowana procedura selekcji kryteriów bazuje na założeniu, że ankietowa ocena wielokryterialna nie jest dokładna. Można tego dowieść przeprowadzając ankietę, w której ankietowani oceniają serwisy pod względem kolejnych kryteriów. Przypisują oni serwisom oceny względem kryteriów oraz oceny ogólne. Ocena jest dokładna, jeżeli oceny serwisów obliczone, jako średnie ważone ocen kryteriów, odpowiadają ocenom ogólnym tych serwisów. Drugie przyjęte założenie mówi, że: istnieje podzbiór kryteriów i ukryte wartości wag, których wykorzystanie w wielokryterialnej ocenie serwisów może dać rozwiązanie zbliżone do wartości oceny ogólnej serwisu określanej przez ankietowanych.

Autorzy sformułowali heurystykę, która mówi, że: *dla poszukiwanego „dobrego" podzbioru kryteriów istnieje model klasyfikacyjny (w sensie uczenia maszynowego), charakteryzujący się niskim stopniem sprzeczności przypadków uczących model*. Wobec tego budując modele klasyfikatorów z wykorzystaniem podzbiorów kryteriów, można wytypować podzbiory, dające rozwiązania bliskie optymalnym. W niniejszej pracy do typowania podzbiorów kryteriów wykorzystano metodę Hellwiga. Opracowaną procedurę przedstawiono wykorzystując model eQual.

Pierwszym etapem badań było zebranie ankiet użytkowników dotyczących oceny jakości trzech największych polskich portali informacyjnych, tj. wp.pl, onet.pl oraz interia.pl zgodnie z modelem eQual. W badaniach wykorzystano 85 ankiet dla każdego portalu, zebrano więc łącznie 255 ankiet. Użytkownicy poza ocenami każdego portalu względem kolejnych kryteriów eQual, podawali w ankietach także indywidualnie przez siebie przypisane wagi każdego kryterium oraz ogólną ocenę każdego serwisu. Kolejnym krokiem procedury było zastosowanie opracowanej metody selekcji cech wykorzystującej pojemność informacyjną.

Następnie walidowano zbudowany ranking poprzez zastosowanie klasyfikatora. Klasyfikator budowany był w wykorzystaniem zbioru kryteriów, z którego iteracyjnie eliminowano najmniej istotne kryterium. Wykorzystanym klasyfikatorem były drzewa decyzyjne CART stosujące do klasyfikacji miarę G-kwadrat oraz szacowanie a-priori prawdopodobieństwa przynależności do określonej klasy decyzyjnej zależnie od liczności przedstawicieli danej klasy w zbiorze treningowym [Rokach i Maimon, 2010][Webb, 2003]. Szacowanie a priori pozwoliło poprawić model klasyfikatora ze względu na to, że liczności poszczególnych klas decyzyjnych były różne. W celu stabilizacji wyników klasyfikacji, za każdym razem wykonywano 10-krotną walidację krzyżową [Rokach i Maimon, 2010]. Wpływ na wybór do klasyfikacji drzew decyzyjnych CART miał fakt, że klasyfikator ten przy pełnym zbiorze kryteriów najlepiej pozwalał uzyskać najlepsze wyniki klasyfikacji. Ponadto stosowane w drzewach decyzyjnych przycinanie pozwala uniknąć przeuczenia klasyfikatora. Minimalną

liczność w węźle drzewa ustalono na 5 przypadków. Podobnie, minimalna liczność węzła potomka, również wynosiła 5 przypadków.

Bazując na wynikach klasyfikacji, wybrano najlepiej rokujące podzbiory kryteriów. Następnie dla każdego serwisu internetowego obliczono jego średnią ocenę ogólną oceny serwisów według wzoru:

$$G_S = \left((\sum_{i=1}^{n} s_i) / 7 \right) \Big/ n * 100\% \tag{7}$$

gdzie: s_i - ogólna ocena serwisu przypisana przez *i*-tego użytkownika (zakres 1-7), n – liczba użytkowników biorących udział w ankiecie. Dla każdego uzyskanego podzbioru kryteriów obliczono też średnie znormalizowane oceny serwisów, wykorzystując oceny kryterialne i wagi zgodnie z wzorem:

$$O_S = \left((\sum_{i=1}^{n}\sum_{j=1}^{m} k_{ij} * w_{ij}) / (\sum_{i=1}^{n}\sum_{j=1}^{m} 7 * w_{ij}) \right) \Big/ n * 100\% \tag{8}$$

gdzie: k_{ij} – ocena serwisu przez *i*-tego użytkownika dla *j*-tego kryterium, w_{ij} – waga przypisana przez *i*-tego użytkownika dla *j*-tego kryterium. Ponieważ różne ilości kryteriów powodowały otrzymywanie przez serwisy różnej ilości punktów, więc punktacja dla każdego podzbioru kryteriów została znormalizowana do zakresu [0-1]. Jeżeli wybrany podzbiór kryteriów oceny dobrze odzwierciedla jakość serwisu internetowego, to między oceną ogólną serwisu a oceną kryterialną zachodzi zależność $G_S \approx O_S$. Porównanie wzorcowej wartości G_S i wartości O_S uzyskanych dla poszczególnych podzbiorów, pozwoliło wytypować optymalny podzbiór kryteriów służących ocenie serwisów informacyjnych. Porównanie tych wartości wykonano z wykorzystaniem miary bezwzględnego średniego odchylenia [Pham-Gia i Hung, 2001], według wzoru:

$$MAD = \sum_{i=1}^{n} \left\| O_i - G_i \right\| / n \tag{9}$$

gdzie: n – liczba badanych serwisów internetowych, G_i – średnia ocena ogólna i-tego serwisu internetowego, O_i – średnia ocena kryterialna i-tego serwisu internetowego. Optymalny podzbiór kryteriów wybrano bazując na wartości MAD i liczności podzbioru.

WYNIKI BADAŃ

W wyniku zastosowania metody Hellwiga otrzymano ranking kryteriów wraz z wagami. Uzyskany ranking kryteriów wykorzystano do klasyfikacji. Ranking ważności kryteriów (numery poszczególnych kryteriów i ich wagi) oraz uzyskane dla niego średnie i minimalne (dla najgorzej klasyfikowanych klas) wartości poprawnych klasyfikacji zawarto w tabeli 1. Średni poziom poprawnych klasyfikacji (C_{mean}) jest wartością procentową poprawnych przypisań do każdej

z klas decyzyjnych (1-7). Minimalny poziom poprawnych klasyfikacji (C_{min}) odzwierciedla poziom poprawnych przypisań dla klasy, dla której ten poziom jest najniższy. Wyniki przedstawiono na tle rankingu tworzonego jawnie przez użytkowników oraz rankingu utworzonego z wykorzystaniem symetrycznej niepewności. W tabeli 1 zaznaczono podzbiory kryteriów wyznaczone metodą Hellwiga (Hellwig), uwzględnione w dalszym badaniu. Zaznaczono także podzbiory wyznaczone metodą symetrycznej niepewności (Symm. Uncert.) oraz oparte na jawnych opiniach użytkowników (Users), które uwzględniono w celach porównawczych.

Tabela 1. Istotności i współczynniki poprawnych klasyfikacji dla zbiorów kryteriów

Ilość kryt.		**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Hellwig	Num.	1	18	11	3	17	14	12	4	9	13	6
	Waga	0,055	0,053	0,051	0,050	0,050	0,049	0,046	0,045	0,044	0,044	0,043
	C_{min} [%]	29,41	49,02	52,94	52,94	41,38	31,03	31,03	31,03	31,03	31,03	31,03
	C_{mean}[%]	58,04	68,24	68,63	68,63	72,55	75,29	73,73	72,94	74,90	75,29	75,29
Symm.Uncert.	Num.	5	17	1	11	9	3	6	12	10	4	19
	Waga	0,149	0,144	0,142	0,141	0,136	0,135	0,134	0,133	0,132	0,117	0,117
	C_{min} [%]	42,86	43,14	37,93	37,93	44,83	34,48	34,48	37,93	57,14	57,14	57,14
	C_{mean}[%]	66,67	69,41	67,84	69,80	74,51	74,90	74,12	74,51	74,51	74,51	74,90
Users	Num.	18	4	12	17	2	13	1	14	22	15	8
	Waga	6,13	6,12	6,04	5,75	5,74	5,71	5,68	5,67	5,46	5,34	5,07
	C_{min} [%]	34,48	41,67	41,67	41,67	41,38	41,38	41,38	55,17	55,17	55,17	41,18
	C_{mean}[%]	61,96	63,92	63,92	59,22	60,39	59,61	59,61	68,63	68,63	70,20	69,02

Źródło: opracowanie własne

Ostatni etap badania polegał na przypisaniu kryteriom, wag zawartych w tabeli 1 i obliczeniu średniego bezwzględnego odchylenia dla ocen obliczonych z użyciem wybranych podzbiorów kryteriów i ich wag. Oceny serwisów wyznaczone za pomocą kolejnych wybranych podzbiorów kryteriów i wartości średniego bezwzględnego odchylenia zawarte są w tabeli 2.

Analizując tabelę 2, należy zauważyć, że niemal każdy podzbiór kryteriów utworzony na podstawie opinii użytkowników charakteryzuje się wartością MAD większą od pełnego 22-elementowego zbioru kryteriów. Ponadto najniższe wartości MAD uzyskano dla podzbiorów, dla których istotność kryteriów wyznaczano za pomocą metody Hellwiga. Zbiory oparte na symetrycznej niepewności, charakteryzują się wartościami MAD większymi niż te, uzyskane dla zbiorów uzyskanych metodą Hellwiga.

Tabela 2. Bezwzględne średnie odchylenie uzyskane dla wybranych podzbiorów kryteriów

Podzbiór kryteriów	Onet.pl	Wp.pl	Interia.pl	MAD
Ocena ogólna	69,66 %	71,85 %	62,44 %	0,00 %
eQual 22	69,48 %	71,54 %	67,47 %	1,84 %
Users 14	70,87 %	73,00 %	68,62 %	2,84 %
Users 13	70,13 %	72,59 %	67,91 %	2,23 %
Users 12	70,47 %	73,14 %	68,28 %	2,64 %
Users 11	70,74 %	73,52 %	68,53 %	2,95 %
Users 10	69,80 %	72,48 %	67,97 %	2,10 %
Users 9	69,11 %	71,79 %	67,27 %	1,81 %
Users 8	68,43 %	71,11 %	67,06 %	2,20 %
Users 7	68,83 %	71,96 %	67,83 %	2,11 %
Users 6	70,01 %	72,55 %	68,74 %	2,45 %
SU 14	69,23 %	71,58 %	66,87 %	1,71 %
SU 13	68,76 %	71,06 %	66,41 %	1,89 %
SU 12	69,34 %	71,48 %	66,71 %	1,65 %
SU 11	69,97 %	71,75 %	67,08 %	1,68 %
SU 10	69,84 %	71,30 %	66,65 %	1,65 %
SU 9	69,33 %	70,65 %	66,12 %	1,74 %
SU 8	70,11 %	71,53 %	66,33 %	1,55 %
SU 7	69,08 %	70,07 %	64,95 %	1,62 %
SU 6	67,35 %	67,92 %	63,55 %	2,45 %
Hellwig 14	69,23%	71,41%	66,34%	1,59%
Hellwig 13	68,77%	70,92%	65,80%	1,73%
Hellwig 12	69,21%	71,49%	65,91%	1,43%
Hellwig 11	68,63%	70,85%	65,29%	1,63%
Hellwig 10	69,22%	71,07%	65,54%	1,44%
Hellwig 9	69,42%	71,31%	65,56%	1,30%
Hellwig 8	69,86%	72,03%	66,04%	1,33%
Hellwig 7	69,21%	71,30%	65,29%	1,29%
Hellwig 6	67,93%	69,62%	63,60%	1,71%

Źródło: opracowanie własne

Analizując tabelę 2, należy zauważyć, że niemal każdy podzbiór kryteriów utworzony na podstawie opinii użytkowników charakteryzuje się wartością MAD większą od pełnego 22-elementowego zbioru kryteriów. Ponadto najniższe wartości MAD uzyskano dla podzbiorów, dla których istotność kryteriów

wyznaczano za pomocą metody Hellwiga. Zbiory oparte na symetrycznej niepewności, charakteryzują się wartościami MAD większymi niż te, uzyskane dla zbiorów uzyskanych metodą Hellwiga.

PODSUMOWANIE

W artykule poruszono temat oceny jakości serwisów internetowych oraz redukcji kryteriów oceny serwisów i określania ich wag z wykorzystaniem metody pojemności informacyjnej Hellwiga. Wyniki badań porównano z wcześniejszymi rezultatami, uzyskanymi przy zastosowaniu selekcji kryteriów na podstawie opinii użytkowników oraz symetrycznej niepewności.

Opierając się na wynikach badań, w pracy wykazano, że zastosowanie pojemności informacyjnej Hellwiga pozwala uzyskać istotne kryteria oceny jakości serwisów internetowych. Wagi i kryteria wybrane za pomocą metody Hellwiga umożliwiają uzyskanie rozwiązania bliższego oczekiwanemu, niż pełny zbiór kryteriów i wagi przypisane przez użytkowników lub też symetryczna niepewność. Ponadto, jest to metoda dająca pełny ranking kryteriów, więc wykorzystując ją, ekspert zachowuje kontrolę przy wyborze podzbioru kryteriów.

Analizując uzyskane wyniki stwierdzić należy, że zbiory kryteriów uzyskane z wykorzystaniem metody Hellwiga dają lepsze wyniki klasyfikacji od zbiorów kryteriów utworzonych z pomocą użytkowników serwisów internetowych. Można z tego wnioskować, że preferencje użytkowników serwisów WWW są w pewnym stopniu niejawne. Mianowicie jawnie deklarują oni, że istotne dla nich są pewne określone cechy witryn internetowych, ale tak na prawdę przy ocenie jakości serwisów podświadomie kierują się nieco innymi kryteriami niż te, które zadeklarowali jako istotne.

BIBLIOGRAFIA

Internet World Stats (2012) World Internet Users and Population Stats, http://www.internetworldstats.com/stats.htm

Netcraft (2012) April 2012 Web Server Survey, http://news.netcraft.com/archives/2012/04/04/april-2012-web-server-survey.html

Hwang J., Yoon Y.S., Park N.H. (2011) Structural effects of cognitive and affective responses to web advertisements, website and brand attitudes, and purchase intentions: The case of casual-dining restaurants, International Journal of Hospitality Management, No.30, s. 897-907.

Lin, C.S., Wu, S., Tsai, R.J. (2005) Integrating perceived playfulness into expectation-confirmation model for web portal context. Information & Management, No.42, s. 683-693.

Grigoroudis, E., Litos, C., Moustakis, V.A., Politis, Y., Tsironis, L. (2008) The assessment of user-perceived web quality: Application of a satisfaction benchmarking approach. European Journal of Operational Research, No.187, s. 1346-1357.

Barnes S.J., Vidgen R.T. (2006) Data triangulation and web quality metrics: A case study in e-government, Information & Management, No.43, s. 767-777.

Lin H.F. (2010) An application of fuzzy AHP for evaluating course website quality, Computers & Education, No.54, s. 877-888.

Ho C., Lee Y. (2007) The development of an e-travel service quality scale, Tourism Management, No.28, s. 1434-1449.

Ou C.X., Sia C.L. (2010) Consumer trust and distrust: An issue of website design, International Journal of Human-Computer Studies, No.68, s. 913-934.

Yang Q., Shao J., Scholz M., Plant C. (2011) Feature selection methods for characterizing and classifying adaptive Sustainable Flood Retention Basins, Water Research, No.45, s. 993-1004.

Zenebe A., Zhou L., Norcio A.F. (2010) User preferences discovery using fuzzy models, Fuzzy Sets and Systems, No.161, s. 3044-3063.

Ziemba P., Budziński R. (2011) Dobór kryteriów dla oceny serwisów informacyjnych w portalach internetowych, Studia i materiały PSZW, No.37, s. 368-378.

Ziemba P. (2011) Formal procedure of selecting the significant website assessment criteria based on data mining methods**,** Young Scientists Innovations Forum **2011,** ISSN 2082-4831. www.fimb.p.lodz.pl.

Gnat S. (2006) Przegląd niektórych metod doboru zmiennych do modeli jednorównaniowych, Prace Katedry Ekonometrii i Statystyki, No. 17, s. 165-178.

Rokach L., Maimon O. (2010) Supervised Learning, Classification Trees, [In]: Maimon O., Rokach L. (eds.), Data Mining and Knowledge Discovery Handbook, Springer, s. 133-174

Webb G.I. (2003) Association Rules, [In]: Ye N. (eds.), The Handbook of Data Mining, Lawrence Erlbaum Associates, s. 25-40.

Pham-Gia T., Hung T.L. (2001) The Mean and Median Absolute Deviations, Mathematical and Computer Modeling, No. 34, s. 921-936.

CRITERIA SELECTION FOR THE WEBSITES QUALITY ASSESSMENT BASED ON THE INFORMATION CAPACITY OF HELLWIG

Abstract: Website quality is highly significant to the users' perception of an organisation. Therefore, the important issue is to assess the websites quality and selection of criteria for this assessment. The paper proposes a method for selection of important criteria assessing the websites quality, based on the individual indicator of the information capacity of Hellwig and Symmetrical Uncertainty.

Keywords: websites quality assessment, criteria selection, information capacity of Hellwig