AKADEMIA EKONOMICZNA IM. KAROLA ADAMIECKIEGO

Prace naukowe

MODELOWANIE PREFERENCJI A RYZYKO '09

Praca zbiorowa pod redakcją naukową
Tadeusza Trzaskalika

KATOWICE 2010

ISBN 978-83-7246-549-8 (wersja drukowana)
ISBN 978-83-7875-957-7 (wersja elektroniczna)

**WYDAWNICTWO AKADEMII EKONOMICZNEJ
IM. KAROLA ADAMIECKIEGO W KATOWICACH**

ul. 1 Maja 50, 40-287 Katowice, tel. 032 257-76-35, fax 032 257-76-43
www.ae.katowice.pl, e-mail: wydawnictwo@ae.katowice.pl

SPIS TREŚCI

WPROWADZENIE

RYZYKO NA RYNKU KAPITAŁOWYM

Od kilkunastu lat do opisu zjawisk ekonomicznych stosuje się koncepcje, modele i metody wypracowane w fizyce, szczególnie w fizyce statystycznej, ale również i kwantowej. Nowy dział nauk fizycznych – ekonofizyka dużą uwagę skupia na rynkach finansowych, traktowanych jako złożony układ dynamiczny. Poprzez analogię do zjawisk fizycznych uzyskano opis zjawisk finansowych, także z możliwością przewidywania nadchodzących krachów finansowych na giełdach. Praca ***Rynek finansowy z punktu widzenia fizyki statystycznej*** (**K. Bolonek-Lasoń**) zawiera modele rynku finansowego stworzone przez fizyków, gdzie uczestników giełdy traktuje się jako układ cząsteczek, które poprzez wzajemne oddziaływanie zmieniają stan układu.

Od 8 lutego 2008 roku Towarowa Giełda Energii S.A. (TGE S.A.) wprowadziła zmiany w harmonogramie sesji Rynku Dnia Następnego (RDN). Sesja RDN składała się obecnie z dwóch fixingów i notowań ciągłych. Bazując na notowaniach RDN z pierwszego roku funkcjonowania nowego harmonogramu, w pracy ***Pomiar ryzyka w systemie ceny jednolitej na Towarowej Giełdzie Energii*** (**A. Ganczarek-Gamrot**) przeprowadzono analizę porównawczą ryzyka podejmowanego przez uczestników rynku w zależności od momentu otwarcia pozycji. Do oceny ryzyka wykorzystano kwantylową miarę Value-at--Risk oszacowaną za pomocą modeli warunkowej wariancji GARCH.

Jedną z popularnych metod wielokryterialnych jest synchroniczna metoda NIMBUS, zaprojektowania głównie z myślą o poszukiwaniu rozwiązań nieróżniczkowalnych modeli wielokryterialnych. Przy jej praktycznym wykorzystaniu można zastosować oprogramowanie WWW-NIMBUS, dostępne bezpłatnie w Internecie dla celów niekomercyjnych. W pracy ***Zastosowanie synchronicznej wersji metody NIMBUS do ustalenia optymalnej struktury portfela akcji*** (**D. Kudyba, T. Trzaskalik**) przedstawiono opis metody oraz jej zastosowanie na rynku kapitałowym. Otrzymane wyniki porównano z wynikami uzyskanymi dla tych samych danych z wykorzystaniem metody satysfakcjonującego poziomu kryteriów.

Wybór funduszu inwestycyjnego jest jednym z podstawowych problemów, jaki muszą rozwiązać inwestorzy decydujący się na ulokowanie swoich kapitałów w tego typu inwestycji. Jednym z podstawowych kryteriów analizowanych przez klientów jest stopa zwrotu osiągana przez fundusz oraz jego ryzyko. Ten ostatni parametr najczęściej oceniany jest na podstawie odchylenia standardowego stopy zwrotu lub współczynnika beta. W pracy ***Ocena ryzyka funduszy inwestycyjnych z wykorzystaniem współczynnika Giniego*** (**E. Majewska**) przedstawiono wyniki zastosowania do oceny ryzyka OFI współczynnika Giniego, który proponowany jest w literaturze jako alternatywna miara ryzyka inwestycyjnego.

Kurs dolara amerykańskiego, obok kursu waluty Unii Europejskiej, ma istotne znaczenie dla polskiej gospodarki. Dolar amerykański pozostaje głównym środkiem płatniczym poza Unią Europejską, tak w sferze prywatnej, jak i instytucjonalnej. Duża zmienność kursu tej waluty w ostatnim okresie skłania do analizy przyczyn tego zjawiska. Przebieg kursu dolara amerykańskiego może być potraktowany jako realizacja procesu stochastycznego, w którym można wyróżnić tendencje rozwojowe kształtujące się w różny sposób w zależności od rozważanych okresów. Obserwacja takiego szeregu czasowego w ciągu ostatnich kilku lat pozwala dostrzec co najmniej kilka momentów, w których krótkookresowe trendy uległy gwałtownym zmianom. Może to wskazywać na występowanie punktów zwrotnych trendu. W pracy ***Próba wykrycia punktów zwrotnych w kształtowaniu się kursu walutowego dolara amerykańskiego*** (**M. Salamaga**) porównano trzy metody mające na celu wykrycie punktów zwrotnych w kształtowaniu się kursu dolara: test Perrona, test Chowa oraz analizę trendu różnic pomiędzy rzeczywistymi realizacjami kursu dolara i prognozowanymi wartościami tego kursu. Dla wykrytych punktów zwrotnych podjęto próbę ustalenia przyczyn, które mogły wywołać zmianę w kształtowaniu się kursu amerykańskiej waluty.

W przypadku asymetrycznych rozkładów stóp zwrotu papierów wartościowych, budowa portfeli inwestycyjnych jedynie na podstawie dwóch pierwszych momentów centralnych może prowadzić do nieuwzględnienia pewnych aspektów ryzyka. Ponadto, zarówno stopień dywersyfikacji portfela, jak i osiągana stopa zwrotu uzależnione są od wykorzystywanych miar ryzyka. Uwzględnienie w analizie większej liczby miar umożliwia dokonanie pełniejszej i bardziej kompletnej oceny poszczególnych papierów wartościowych. W pracy ***O pewnej propozycji zastosowania rankingów wielokryterialnych w analizie portfelowej*** (**A. Sielska**) przedstawiono metodę wprowadzenia za pomocą rankingów wielokryterialnych dodatkowych warunków ograniczających w zadaniu Markowitza. Postępowanie takie skutkuje uszeregowaniem

spółek w zależności od preferencji decydenta względem wybranych miar ryzyka, a dodatkowe warunki ograniczające nałożone na zbiór rozwiązań dopuszczalnych umożliwiają uwzględnienie tych preferencji przy konstrukcji portfela. Stopy zwrotu osiągnięte dla tak zbudowanych portfeli zostały następnie porównane z wybranymi benchmarkami.

Portfele inwestycyjne budowane na podstawie klasycznej analizy korelacji mają pewne ograniczenia. Ich ryzyko specyficzne może okazać się procesem błądzenie przypadkowego. W takim przypadku stopy zwrotu portfela mogą znacznie odchylać się od stóp zwrotu indeksu referencyjnego. Takie podejście nie uwzględnia długookresowego związku portfela z indeksem – nie wykorzystuje dostępnych informacji. Wykorzystanie kointegracji pozwala uwzględnić powyższe informacje w budowie portfela. Można w konsekwencji budować portfele zawierające stosunkowo niewielką liczbę walorów, które są równocześnie związane z indeksem rynkowym w długim okresie. Celem pracy ***Budowa portfeli efektywnych na podstawie narzędzi kointegracji*** (**S. Śmiech**) jest pokazanie technik budowania portfeli wykorzystując kointegrację na poziomie wartości walorów oraz praktyczna ilustracja na podstawie akcji notowanych na GWP.

W zagadnieniach zarządzania ryzykiem stosowane są zaawansowane metody. Przykładem jest w model Mertona (1974) służący do obliczania prawdopodobieństwa upadku firmy. W pracy ***Implementacja Modelu Mertona w wybranych środowiskach optymalizacji*** (**K. Targiel**) przedstawiono implementację tego modelu w dwu wybranych środowiskach optymalizacji, tzn. w systemie LINGO oraz GAMS. Przeprowadzono także analizę procesu zapisu modelu w tych środowiskach.

Prowadzone badania nad podejmowaniem decyzji w warunkach ryzyka i w warunkach niepewności pokazują nowe ścieżki postępowania w odróżnieniu od podejść klasycznych. W pracy ***Pesymistyczna optymalizacja portfelowa*** (**G. Trzpiot**) warunki związane z kryterium oczekiwanej użyteczności są zastąpione poprzez pesymistyczne kryteria decyzyjne.

RYZYKO W UBEZPIECZENIACH I RYZYKO BANKOWE

W pracy ***Propozycja miar do oceny dynamiki OFE*** (**J. Białek**) zaproponowano dwie miary, umożliwiające hierarchizację funduszy od najbardziej dynamicznego począwszy. Tego typu ranking wyklucza ryzyko wyboru takiego funduszu, który co prawda ma wysokie aktywa i satysfakcjonujące zwroty, ale

słabnącą dynamikę zmian wartości jednostek uczestnictwa. W perspektywie czasu utraci on pozycję na rzecz chwilowo słabszych, ale bardziej dynamicznych funduszy.

W latach 2000-2002 przez PUNU stworzony został system wczesnego ostrzegania dla polskich ubezpieczycieli, silnie bazujący na ocenie wskaźnikowej ubezpieczycieli. W związku ze zmianami na polskim rynku ubezpieczeń, takimi jak przykładowo zmiana struktury kapitałowej, zmiana prawodawstwa zmiana realiów rynkowych związanych z przestąpieniem do UE, należy zweryfikować ustalone wcześniej oceny i poddać pod dyskusję wyznaczone statystycznie nowe oceny. Temu właśnie w odniesieniu do zakładów ubezpieczeń na życie poświęcona jest praca **Nowe oceny dla wartości wskaźników polskich zakładów ubezpieczeń na życie** (**A. Jędrzychowska**). Badanie oparte zostało na danych finansowych zawartych w sprawozdaniach finansowych polskich zakładów na życie za lata 1999-2007 (badanie PUNU odnosiło się do sprawozdań za lata 1996-2000).

Celem pracy ***Pomiar korzyści z zastosowania modelu wielomianowego w ograniczaniu ryzyka kredytowego*** (**J. Marzec**) jest prezentacja pomiaru korzyści finansowych banku, wynikających z wykorzystania w procesie podejmowania decyzji kredytowych modelu wielomianowego dla kategorii porządkowanych. Udzielenie kredytu ma charakter decyzji dychotomicznej (tak lub nie), kredytobiorca zaś może dotrzymać warunków umowy kredytowej albo przestać spłacać raty kredytu lub odsetki, narażając bank na częściową lub całkowitą stratę odsetek i kapitału. Statystyczna teorii decyzji daje możliwość skonstruowania reguły, według której bank podejmuje decyzję o przyznaniu albo odmowie danego kredytu w zależności od marży kredytowej i poziomu ryzyka – wiarygodności potencjalnego kredytobiorcy. W przykładzie empirycznym wykorzystano dane o kredytach detalicznych z banku działającego w Polsce. Następnie przedstawiono wyniki, które między innymi informują, że bank stosując niestandardowe modele, np. o rozkładzie t Studenta i z aproksymacją II rzędu dla zmiennej ukrytej reprezentującej preferencje kredytobiorcy, uzyskuje – w porównaniu do modeli probitowego i logitowego – dodatkowy zysk. Ogólna konkluzja z przeprowadzonych badań sprowadza się do stwierdzenia, iż zastosowanie któregokolwiek z proponowanych modeli ryzyka kredytowego pozwala ograniczyć to ryzyko i osiągnąć wymierne korzyści finansowe.

Celem pracy ***Metody rozszerzonej analizy rentowności inwestycyjnej otwartych funduszy emerytalnych*** (**A. Mikulec**) jest prezentacja mniej znanych oraz rzadko omawianych w polskiej literaturze wskaźników rentowności i efektywności portfela inwestycji, między innymi miary M2 i M3, wskaźników

zysków i strat opartych na momentach częściowych (np. Omega) czy Indeksu Stutzera. Wybrane miary wykorzystano do analizy rynku Otwartych Funduszy Emerytalnych (OFE) w Polsce w latach 2000-2008.

Celem pracy ***Wpływ niejednorodności populacji na ryzyko demograficzne w portfelu ubezpieczeń na życie*** (**M. Papież**) jest przegląd modeli śmiertelności dla niejednorodnych populacji oraz wykorzystanie tych modeli do oceny ryzyka demograficznego dla portfela ubezpieczeń na życie. Zaprezentowano między innymi jedną z metod analizy nieobserwowalnych czynników, a mianowicie modele nazywane „frailty models". Modele te wykorzystano do badania rozkładów płatności w portfelu ubezpieczeń na życie.

Współczesny rozwój gospodarki żywnościowej nakierowany jest na produkcję roślin, które mają na celu między innymi zwiększenie wydajności, odpornych na choroby, susze, ostre zimy i nieszkodliwe dla zdrowia ludzkiego. Od wielu lat trwają prace i doświadczenia nad produkcją roślin genetycznie zmodyfikowanych (GMO). W pracy ***Ubezpieczenie odpowiedzialności cywilnej za produkty żywnościowe pochodzące z roślin genetycznie zmodyfikowanych*** (**S. Wieteska**) przedstawiono pojęcie i zakres produkcji tych roślin. Zaprezentowano argumenty zwolenników i przeciwników produkcji GMO zarówno w Polsce, jak w innych państwach. Na tle tych rozbieżnych opinii proponuje się wprowadzenie obowiązkowych ubezpieczeń odpowiedzialności cywilnej producentów za produkty pochodzenia roślinnego, które zostały genetycznie zmodyfikowane. Wymagania prawne co do produkcji tych roślin są bardzo restrykcyjne. Jednakże pomimo tych restrykcji jest pewien margines niepewności. Ubezpieczenie to jest pomostem między zwolennikami i przeciwnikami produktów GMO. Ma ono na celu pokrywać udowodnione straty wśród ludności spożywającej te produkty.

W kalkulacji składki dla niejednorodnych portfeli ubezpieczeniowych stosowana jest teoria wiarygodności (zaufania). Występuje w niej wiele modeli służących do szacowania składki wiarygodnej, bazującej na składce kolektywnej, a uwzględniającej czynniki ryzyka. W pierwszej części pracy ***Szacowanie poziomu składki wiarygodnej w wybranych modelach z zastosowaniem pakietu actuar programu statystycznego R*** (**A. Wolny-Dominiak**) przedstawiono ogólną koncepcję szacowania składki wiarygodnej w aktuarialnych modelach wiarygodności. Następnie krótko omówiono klasyczne modele: Buhlmanna oraz Buhlmanna-Strauba, ich rozszerzenie, jakim jest model hierarchiczny oraz model autoregresyjny uchylający jedno z założeń wcześniej przedstawianych modeli. W ostatniej części pracy przedstawiono możliwości pakietu

actuar programu R wyznaczania składki wiarygodnej w poszczególnych modelach oraz zaimplementowano algorytm iteracyjny dla modelu autoregresyjnego.

MODELOWANIE PREFERENCJI

W pracy ***Badanie preferencji polskiego elektoratu w wyborach parlamentarnych 2007 roku*** (**J. Hołubiec, A. Małkiewicz, G. Szkatuła, D. Wagner**) do opisu preferencji wyborców w czasie wyborów parlamentarnych w 2007 r. zaproponowano trzy grupy cech: charakteryzujące programy partii politycznych (22 cechy), opisujące partie polityczne oraz głoszone hasła i wartości (7 cech) i charakteryzujące kampanię wyborczą (4 cechy). Cechą decyzyjną było dostanie się lub nie danej partii politycznej do Sejmu. Przeanalizowano zarówno wyznaczone zbiory reduktów, jak również otrzymane reguły decyzyjne. Wyniki analizy wskazują, iż preferencje wyborców w tych wyborach skupiły się głównie na cechach opisujących specyficzną wówczas sytuację polityczną.

Praca ***Wpływ rodzaju otoczenia na decyzje konsumenckie*** (**A. Kowalska-Styczeń**) porusza problem wpływu wielkości otoczenia na podejmowanie decyzji przez konsumentów. Najbliższe otoczenie rozumiane jest tutaj jako osoby, z którymi mamy styczność na co dzień, czyli najbliższa rodzina, przyjaciele. W badaniach przeprowadzono symulacje wpływu otoczenia na podejmowanie decyzji przez konsumentów za pomocą automatów komórkowych.

Mechanizm rynkowy definiuje regułę alokacji, która dzieli oferty na przyjęte (w całości lub częściowo) i odrzucone, oraz regułę wyceny, której zadaniem jest wyznaczenie cen rozliczeniowych. Jest to złożony podproblem decyzyjny, w którym należy uwzględnić preferencje każdego z uczestników, np. osiągnięcie największych korzyści indywidualnych, oraz preferencje globalne, np. osiągnięcie sumarycznie największych korzyści ekonomicznych. Jednym z problemów w projektowaniu mechanizmów jest zapewnienie zgodności celów (preferencji) indywidualnych każdego z uczestników rynku z celami (preferencjami) globalnymi. Celem pracy ***Zgodność wybranych metod wyceny towarów z preferencjami uczestników rynku*** **(P. Pałka, E. Toczyłowski)** jest analiza porównawcza reguł wyceny na tle prostego modelu handlu giełdowego, z punktu widzenia tego, w jaki sposób poszczególne reguły wyceny (a więc i mechanizmy) spełniają preferencje indywidualne oraz preferencje globalne.

W pracy ***O pewnym dyskretnym modelu rynku*** (**Z. Świtalski**) przedstawiono model rynku z dwustronnymi preferencjami wzorowany na modelu Gale'a-Shapleya (rekrutacji kandydatów do szkół). Wprowadzono pojęcie uogólnionej równowagi na takim rynku i pokazano jej związki z pojęciem skojarzenia stabilnego zdefiniowanego przez Gale'a i Shapleya. Przedstawiono też przykład pokazujący nietypowe zachowanie się funkcji popytu na rozważanym rynku (popyt może nie ulec zmianie nawet przy wzroście wszystkich cen oferowanych dóbr).

W pracy ***Zasada ortogonalności i przykłady jej zastosowań*** (**S. Walukiewicz**) wprowadzono pojęcie ortogonalności miedzy dwoma formami (kategoriami, wymiarami itp.) w naukach społecznych, takich jak ekonomia, zarządzanie, socjologia czy nauki polityczne. Ortogonalne formy mogą współdziałać (współpracować) tak jak wartości materialne z niematerialnymi w danej firmie. Następnie w sposób indykatywny opisano cztery formy kapitału w szeroko rozumianej firmie: finansowy, materialny, ludzki i społeczny oraz dowiedziono, iż są one wzajemnie ortogonalne lub rozłączne. Stąd wynika, że wartość firmy (jej całkowity kapitał) jest równa sumie czterech wyżej zdefiniowanych składowych. Wynik ten jest podstawą modelu bilansowego do analizy kapitału społecznego – wynikające stąd zalecenia i metodologia takiej analizy (badań ankietowych) są omówione w opracowaniu. Na zakończenie uogólniono rozważania na poziom kraju/regionu, wprowadzając pojęcie nowego PKB, opisując sposób jego szacowania oraz porównując go z dotychczas używanym PKB.

RYZYKO W ORGANIZACJACH I PRZEDSIĘBIORSTWACH

W pracy ***Wpływ powiązań społecznych na preferencje konsumentów*** (**M. Anholcer**) przedstawiono wyniki badań nad zależnością preferencji konsumentów od charakteru powiązań interpersonalnych. Badania te przedstawione zostały w szerszym kontekście, jako wstęp do większego projektu, mającego na celu skonstruowanie symulacyjnego narzędzia umożliwiającego prognozowanie obrotów wybranych gałęzi gospodarki.

W pracy ***Kilka uwag o ryzyku w sektorze ochrony zdrowia*** (**S. Grzesiak**) przedstawiono problem identyfikacji i oceny ryzyka w odniesieniu do ochrony zdrowia. Wskazano na sytuacje, w których występuje konieczność uwzględnienia skali i ewentualnych konsekwencji niedoszacowania ryzyka. Przedyskutowano ogólne przyczyny powstawania i występowania takich zachowań pacjentów i innych uczestników działających w sektorze medycznym, które generują sytuacje wywołujące zwiększone ryzyko, w szczególności finansowe. Pod-

kreślono specyfikę tego rodzaju przypadków w Polsce. W tym kontekście rozważano możliwości tworzenia zabezpieczeń przed ryzykiem, minimalizacji ponoszonych kosztów oraz realność tych rozwiązań dla warunków polskich.

W przypadku analizy przestrzennych procesów ekonomicznych, kluczową rolę odgrywa identyfikacja struktury danych, której poprawne przeprowadzenie powinno zapewnić prawidłowość otrzymanych wniosków. W pracy ***Analiza danych przestrzennych a jakość informacji*** (**M. Pietrzak**) pokazano, że pominięcie ważnego składnika struktury procesów prowadzić może do błędnych interpretacji. Rozpatrzony szczegółowo został problem wpływu nieuwzględnienia autokorelacji przestrzennej na jakość testów statystycznych oraz jakość modelowania ekonometrycznego.

W pracy ***Ryzyko w działalności bankowej*** (**E.A. Studzińska**), autorka koncentruje się na zagadnieniach związanych z ryzykiem na rynku bankowym, zarówno w ujęciu ogólnym, jak i w odniesieniu do poszczególnych rodzajów ryzyka. Przedstawiono rodzaje ryzyka typowo bankowego ze względu na różne kryteria. Ryzyko jest ściśle związane z działalnością bankową i oznacza zagrożenie osiągnięcia zamierzonych zysków, co może wiązać się ze zmniejszeniem potencjalnych zysków, płynności, wiarygodności, a także trudnościami finansowymi. Istotne znaczenie mają przyczyny wystąpienia ryzyka bankowego. Nie można wyeliminować ryzyka, ale banki starają się nim zarządzać, zarówno w procesie zwiększenia przychodów, jak i redukcji kosztów. Ryzyko bankowe może stwarzać zagrożenie dla prawidłowego funkcjonowania banku i powinno być ograniczane, kontrolowane oraz mierzone, poprzez wewnętrzny sposób pomiaru ryzyka bankowego przy uwzględnieniu regulacji ostrożnościowych nadzoru bankowego.

Pierwsza części pracy ***Model ryzyka decyzji strategicznych zarządu organizacji gospodarczej*** (**J. Zemke**) jest próbą odniesienia się do definicji ryzyka w kontekście podejmowanych decyzji i działań zarządu w procesach zarządzania organizacją. Ocena skutków podejmowanych decyzji i działań dotyczy procesów kontroli realizacji planów strategicznych. Związana jest z definicją takich cech celów strategicznych, które o ile to możliwe można mierzyć, a ich zmiany pozwalają ocenić stan podjętego ryzyka, tzn. dynamikę oraz kierunek zmian zagrożeń realizowanych celów. Ta faza procesów kontroli wiąże się z identyfikacją zmiennych kontrolowanych. Zmienne te pełnią istotne znaczenie w budowie przestrzeni ryzyka, gdyż na ich podstawie definiowana jest przestrzeń zdarzeń elementarnych, stanowiąc fundament budowanego modelu ryzyka. Wyróżnienie w zdefiniowanej przestrzeni σ ciała, które stanowi zbiór o szczególnych własnościach jego elementów, są zmiennymi losowymi iden-

tyfikującymi skutki ryzyka realizowanych decyzji, pozwala na sformułowanie fundamentalnego wniosku tego opracowania, elementy wyróżnionego σ ciała są wektorami losowymi. Zdefiniowanie odwzorowania określonego na elementach σ ciała, przyjmującego wartości rzeczywiste z przedziału domkniętego [0, 1], otwiera tę część opracowania, w której zdefiniowano miary charakterystyczne wektora losowego: prawdopodobieństwo tego, że składowe wektora przyjmą wartości z określonego przedziału zmienności, wektor wartości oczekiwanych, wektor wariancji składowych, kowariancje składowych wektora ryzyka oraz miary warunkowe wektora ryzyka. Część teoretyczną opracowania uzupełnia przykład identyfikacji oraz miar modelu ryzyka obsługi zobowiązań długoterminowych.

RYZYKO – EKONOMIA, MATEMATYKA I EKONOOMETRIA

W pracy ***Zastosowanie metody autokowariancyjnej do prognozowania wskaźnika inflacji*** (**A. Przybylska-Mazur**) przedstawiono jedną z metod prognozowania wskaźnika inflacji – metodę autokowariancyjną. Przeprowadzono analizę prognozy wskaźnika inflacji, ponieważ jest ona jedną z istotnych determinant wyznaczających przyszłe trendy gospodarki. Ponadto stwierdzono, że kierunek zmian prognoz wskaźnika inflacji wyznaczonych na podstawie zaprezentowanej metody pokrywa się z kierunkiem zmian rzeczywistej wartości wskaźnika inflacji i daje coraz mniejsze błędy prognozy.

Koncepcja dominacji stochastycznej jest nieparametrycznym podejściem do porównywania wyników decyzji, stosowanym w problemach sformułowanych zarówno w kategoriach zysku, jak i strat. Do klasycznych problemów sformułowanych w kategoriach strat, w których wykorzystuje się porządkowanie wyników zgodne z relacją dominacji stochastycznej drugiego rzędu, należą np. problemy decyzyjne w sferze ubezpieczeń, problem optymalizacji wielkości zapasów. W. Ogryczak zaproponował zastosowanie metodologii wielokryterialnej do generowania decyzji efektywnych w sensie relacji dominacji stochastycznej drugiego rzędu w problemach optymalizacji portfela, sformułowanych w kategoriach zysku. Praca ***Zastosowanie koncepcji dominacji stochastycznej do wspomagania decyzji w problemach sformułowanych w kategoriach strat*** (**O. Sobotka**) jest poświęcona analizie możliwości zastosowania podobnego podejścia do problemów sformułowanych w kategoriach strat.

Tadeusz Trzaskalik

RYZYKO NA RYNKU KAPITAŁOWYM

Katarzyna Bolonek-Lasoń
Uniwersytet Łódzki

RYNEK FINANSOWY Z PUNKTU WIDZENIA FIZYKI STATYSTYCZNEJ*

Wstęp

W ostatnich latach pojawia się coraz więcej artykułów należących do jednej z najnowszych dziedzin nauki, tzw. ekonofizyki, zajmującej się badaniem zjawisk występujących w złożonych układach gospodarczych i na rynkach finansowych za pomocą narzędzi wykorzystywanych w fizyce. Metody fizyki mają swoje szczególne zastosowanie na rynkach finansowych ze względu na fakt, iż światową gospodarkę można traktować jako układ złożony z pojedynczych uczestników, którzy oddziałują ze sobą w podobny sposób, jak zbiór cząsteczek. Empirycznie poszukuje się teorii i praw, które pozwolą opisać lub pomogą zrozumieć takie kompleksowe oddziaływanie.

Podstawowa idea ekonofizyki polega na tym, że w obu dziedzinach występują zjawiska polegające na kumulacji drobnych oddziaływań elementów systemu, prowadzących do zmian kolektywnego zachowania całego układu. W fizyce zjawiska te są w mniejszym stopniu zależne od charakteru oddziaływań między cząsteczkami, a bardziej są efektem statystycznego działania prawa wielkich liczb. Powoduje to, że cechują się one dużą uniwersalnością (tzn. różne układy fizyczne wykazują podobne cechowanie przy tzw. przejściach fazowych, czyli globalnych zmianach ich własności). Ze względu na statystyczny charakter praw opisujących kolektywne zmiany układu można się spodziewać, że metody fizyki statystycznej będą użytecznym narzędziem w opisie zjawisk ekonomicznych, szczególnie związanych z dużymi zmianami jakościowymi na rynku.

Celem pracy jest przedstawienie dwóch modeli zjawisk ekonomicznych inspirowanych przez teorie fizyczne. Pierwszy model jest to tzw. gra mniejszościowa stanowiąca z matematycznego punktu widzenia układ dynamiczny modelujący zasadnicze cechy zachowań graczy giełdowych. Model choć bardzo uproszczony wydaje się nieźle oddawać zasadnicze cechy dynamiki gry giełd-

* Praca naukowa finansowana ze środków na naukę w latach 2008-2010 jako projekt badawczy Nr N N111 306335.

dowej. Drugi model stara się opisać zasadnicze cechy dynamiki indeksu giełdowego na podstawie podstawowych założeń, że charakterystyka punktu zwrotu na giełdzie i zachowania układów fizycznych w pobliżu punktu krytycznego jest bardzo podobna. Również w tym przypadku, mimo dokonania skrajnych uproszczeń, wydaje się, że model jest w stanie uchwycić zasadnicze cechy zachowania się indeksu giełdowego.

1. Gra mniejszościowa jako model rynku finansowego

Jednym z najbardziej znanych modeli rynku finansowego jest model gry mniejszościowej szeroko rozważanej w pracach [3; 10; 14; 16; 19]. Gra mniejszościowa jest modelem rynku, na którym inwestorzy kupują i sprzedają papiery wartościowe osiągając zysk wynikający jedynie z fluktuacji ceny. Podstawową ideą modelu jest to, że jeśli większość inwestorów chce kupić akcje, wówczas opłaca się sprzedawać i na odwrót, gracze z grupy mniejszościowej zawsze wygrywają. Gracze podejmując decyzję wykorzystują zarówno swoje dotychczasowe doświadczenie, jak i pewne wzory informacji.

Ogólny model gry zdefiniowany [3] jest następująco. Mamy N graczy oznaczonych przez i. W każdym kroku czasowym gracz otrzymuje jedne z P możliwych wzorów informacji $\mu(t)$, na podstawie którego musi podjąć decyzję kupna lub sprzedaży $b_i(t) \in \{-1, 1\}$ (kupno/sprzedaż). Aby tego dokonać, każdy z graczy jest wyposażony w S strategii $\boldsymbol{a}_{ig} = \{a_{ig}^{\mu}\}$ $(g = 1,...,S)$, które odwzorowują informacje $\mu \in \{1,\ldots,P\}$ w decyzje $a_{ig}^{\mu} \in \{-1, 1\}$. Każdy składnik a_{ig}^{μ} strategii w czasie $t = 0$ wybierany jest przypadkowo i niezależnie, z równym prawdopodobieństwem dla każdego i, g oraz μ i jest stały w ciągu całej gry. Gracze zachowują ślad swoich strategii za pomocą funkcji wyceny U_{ig}, która jest zapoczątkowana przez pewną wartość $U_{ig}(0)$, a jej dynamikę przedstawia wzór

$$U_{ig}(t+1) - U_{ig}(t) = \frac{-a_{ig}^{\mu(t)} A(t)}{N} \tag{1}$$

gdzie $A(t) = \sum_{i=1}^{N} b_i(t)$ jest definiowana jako nadwyżka popytu w czasie t.

W każdym kroku czasowym gracz wybiera strategię $g_i(t) = \arg\max_g U_{ig}(t)$ niosącą ze sobą najwyższą wycenę i formułującą odpowiednią decyzję – bit $b_i(t) = a_{g_i(t)}^{\mu(t)}$. W ten sposób modeluje się sytuację, w której inwestorzy w każdym kroku czasowym wybierają strategię, dzięki której oczekują, że osiągną najwyższy zysk.

Pozostaje nam określenie, co rozumiemy pod pojęciem wzoru informacji. W szczególności może to być ciąg *m* decyzji, jakie grupa mniejszościowa podejmowała w przeszłości; takie podejście opisuje zamknięty system, w którym gracze przetwarzają i reagują na informacje wytwarzane kolektywnie przez samych siebie. Taki wybór wzoru informacji definiuje się jako informacje endogeniczne. Jednak gracze mogą w sposób przypadkowy z równym prawdopodobieństwem otrzymywać informację ze zbioru $\{1,...,P\}$ w każdym kroku czasowym. Ta propozycja to model z losowymi informacjami egzogenicznymi. Zatem gra mniejszościowa jest całkowicie zdefiniowana [14] przez następujące reguły

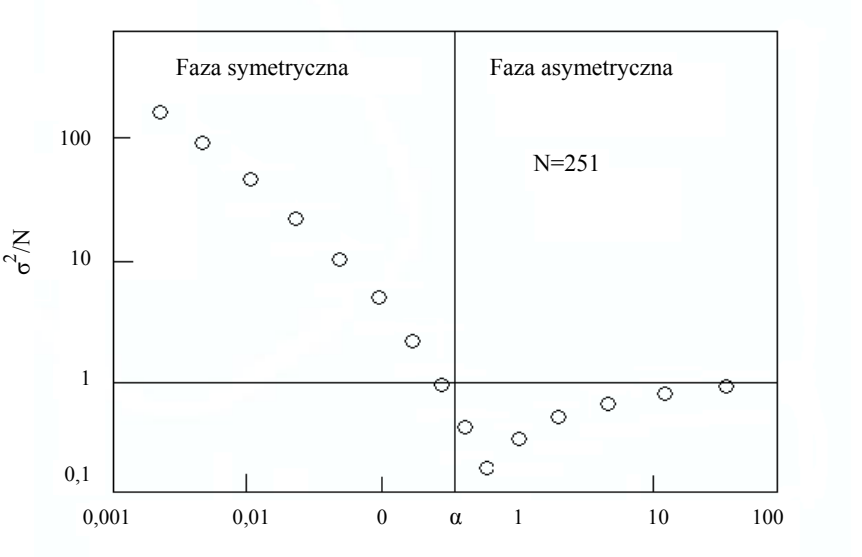
$$\begin{aligned} g_i(t) &= \arg\max_g U_{ig}(t) \\ A(t) &= \sum_{i=1}^{N} a_{ig_i(t)}^{\mu(t)} \\ U_{ig}(t+1) - U_{ig}(t) &= \frac{-a_{ig}^{\mu(t)} A(t)}{N} \end{aligned} \tag{2}$$

Równania (2) opisują układ dynamiczny, którego zachowanie jest określone jednoznacznie przez warunki początkowe i zadany wzór informacji. Jest to model bardzo uproszczony, ale oddający istotę gry rynkowej polegającą na preferowaniu strategii, które w poprzednim „rozdaniu" zapewniały przynależność do grupy mniejszościowej.

Z punktu widzenia makroskopowego, zainteresowanie naukowców koncentruje się wokół wartości $A(t)$, czyli różnicy między liczbą graczy, którzy podjęli decyzję sprzedaży, i liczbą graczy z decyzją kupna w każdym kroku czasowym. Oczywiste jest, że żadna z decyzji „kupno" czy „sprzedaż" nie może systematycznie występować w grupie mniejszościowej; oznacza to, że $A(t)$ fluktuuje wokół zera. Gdyby tak nie było, gracze z łatwością mogliby polepszać swoje wyniki poprzez wybór strategii często wygrywającej w poprzednich krokach.

Wariancja $\sigma^2 = \langle A^2 \rangle$ wielkości $A(t)$ w stanie stacjonarnym mierzy efektywność rozmieszczenia środków pieniężnych. Istotnym parametrem kontrolnym modelu jest względna liczba wzorów informacji $\alpha = P/N$. Interesująca jest zależność σ^2 od α. W przypadku gdy $\alpha >> 1$ przestrzeń informacji jest zbyt duża, by pozwolić na koordynację i gracze zasadniczo zachowują się przypadkowo. Jeśli α maleje, to oznacza, że coraz więcej graczy przyłącza się do gry lub możliwa liczba informacji maleje, zatem malejąca wartość σ^2 / N sugeruje, że gracze wykorzystując informację przechodzą do korzystniejszych stanów w porównaniu z tymi, które wynikają z przypadkowych fluktuacji.

Rysunek 1 przedstawia wyniki symulacji σ^2/N jako funkcji α w przypadku dwóch strategii. Widać, że wariancja osiąga minimum w punkcie $\alpha_c \approx 0{,}34$. Otrzymane zachowanie przypomina opis zjawisk krytycznych w fizyce statystycznej. Analogię można rozszerzyć identyfikując symetrię, której łamanie jest związane z punktem krytycznym.



Rys. 1. Wynik symulacji wielkości σ^2 / N jako funkcji parametru kontrolnego $\alpha = 2^M / N$ przy założeniu dwóch strategii (S=2) dla każdego gracza

Rozważany model można udoskonalać wprowadzając dodatkowe elementy.

Po pierwsze, można rozróżnić trzy typy graczy [8] ze względu na ich zachowanie wobec dostępnych informacji

$$b_i(t) = \begin{cases} rand & i \in N_n \\ v_i a_i^{\mu(t)} & i \in N_p \\ \omega_i a_{i,s_i(t)}^{\mu(t)} & i \in N_s \end{cases} \tag{3}$$

Pierwszy rodzaj graczy (N_n) to tzw. gracze wywołujący szum informacyjny, którzy podejmują decyzję przypadkowo bez żadnego związku z $\mu(t)$. Drugi rodzaj graczy, producenci (N_p), zachowują się w sposób deterministyczny przy danej informacji $\mu(t)$. Zmienna v_i jest wielkością ich inwestycji, a a_i^{μ} jest funkcją losową przekształcającą μ w $\{\pm 1\}$, wybieranej niezależnie dla każdego $i \in N_p$. Funkcje te nazywane są strategiami krótkoterminowymi,

ale producenci nie optymalizują swoich zachowań: mają tylko jedną strategię. Ostatnia grupa graczy to spekulanci mający S strategii $a_{i,s_i(t)}^{\mu(t)}$, z których wybierają tę, która daje najlepszy wynik; ω_i jest kwotą inwestowaną przez i-tego spekulanta.

Po drugie, można zadać pytanie jak zmieni się model gry mniejszościowej, gdy gracze będą dysponować określonym budżetem c_i. Zakładamy, że gracz *i* w czasie *t* posiada kapitał c_i, który jest zmienną dynamiczną i część ε tego kapitału każdy gracz inwestuje na rynku. Spekulanci giełdowi zyskują tylko na różnicach cenowych rynku, zatem wartość kapitału $c_i(t)$ ewoluuje w wyniku ich działań. Jednak producenci, którzy mają też inne dochody, wykorzystują rynek do ponownego rozdzielenia środków, przy czym zawsze inwestują określoną ich ilość.

Przedstawiony model gry rozważano w pracach [8; 9]. Między innymi analizowano, czy przejście fazowe będzie odporne na nową zmienną oraz czy nieregularne podziały zwrotów pieniężnych są powiązane z punktem krytycznym. Autorzy dokonali numerycznych symulacji, w wyniku których okazało się, że w miarę spadku całkowitej liczby informacji (lub wzrostu liczby graczy), rynek staje się coraz mniej przewidywalny.

Kolejna modyfikacja modelu polega na usunięciu nierzeczywistego przymusu graczy do dokonywania transakcji w każdym kroku czasowym. Wydaje się konieczne, aby pozwolić graczom nie uczestniczyć w danej transakcji, jeśli ich strategie handlowe wypadają słabo. Zatem tzw. zerowa strategia będzie oznaczać wybór $b_i = 0$. Nawet przy takiej modyfikacji okazuje się, że przejście fazowe nadal istnieje.

2. Przewidywanie krachów finansowych

Jednym z najbardziej poszukiwanych modeli rynku jest model, dzięki któremu można będzie przewidzieć moment, w którym nastąpi krach finansowy. Didier Sornette, Anders Johannes oraz J.-P. Bouchaud [17] jako jedni z pierwszych zasugerowali, że krach finansowy można uważać za pewnego rodzaju dynamiczny punkt krytyczny z charakterystycznym logarytmiczno-periodycznym zachowaniem podobnym do tego, który znaleziono dla trzęsień ziemi. Autorzy zaprezentowali model dynamicznego rynku finansowego z mikroskopowego punktu widzenia oraz pokazali, że pewnego rodzaju systemowa niestabilność może doprowadzić do krachu finansowego.

Model ten jest wyjaśniany [18] jako hierarchiczny model handlowców, oddziałujących poprzez naśladowanie procesów reakcji rynkowych. Model może być rozwiązany analitycznie i dostarczać precyzyjnych ilościowych relacji między wykładnikiem krytycznym a log-periodyczną częstotliwością z jednej strony oraz zachowaniami handlowców z drugiej.

Szeroko rozważana log-periodyczność na rynkach finansowych [1; 2; 11; 13; 17; 18] zainteresowała również polskich fizyków, a za prekursora badań tego zagadnienia uważa się Stanisława Drożdża. W pracy [11], rozważając najbardziej istotne zdarzenia finansowe z przeszłości pokazano, że obserwowany w latach 2000-2002 spadek indeksu Standard & Poor's 500 z perspektywy log-periodycznej jest tak samo znaczący jak ten, który wystąpił na początku lat 30. i z końca lat 70. XX wieku. Poza tym autorzy przewidują, że w okolicach 2025 roku nastąpi poważny krach finansowy.

Punktem wyjścia rozważań jest założenie, że właściwie zdefiniowana funkcja $F(x)$ charakteryzująca system spełnia warunek skalowania podobny do pojawiającego się przy opisie zjawisk krytycznych w fizyce statystycznej

$$F(\lambda x) = \gamma F(x) \tag{4}$$

przy czym dodatnia stała γ opisuje, jak zmieniają się własności systemu kiedy zostanie on przeskalowany przez czynnik λ. Jednym oczywistym rozwiązaniem tego równania jest

$$F_0(x) = x^{\alpha} \tag{5}$$

gdzie $\alpha = \ln(\gamma)/\ln(\lambda)$. Reprezentuje to standardowe prawo potęgowe, przy czym α jest odpowiednikiem wykładnika krytycznego. Ogólne rozwiązanie równania (4) ma postać

$$F(x) = x^{\alpha} P\left(\frac{\ln(x)}{\ln(\lambda)}\right) \tag{6}$$

P oznacza funkcję periodyczną o okresie jeden. Dominujący czynnik skalujący z równania (5) dostaje poprawkę będącą okresową funkcją $\ln(x)$. Forma funkcyjna P nie jest określona na tym poziomie. Wymaga się tylko, żeby przy

$$x = |T - T_c| \tag{7}$$

(gdzie T oznacza czas odnoszący się do szeregu czasowego ceny, reprezentuje odległość od punktu krytycznego T_c), odstęp między odpowiednimi kolejnymi powtarzalnymi strukturami w x_n spełniał warunek

$$\frac{x_{n+1} - x_n}{x_{n+2} - x_{n+1}} = \lambda \tag{8}$$

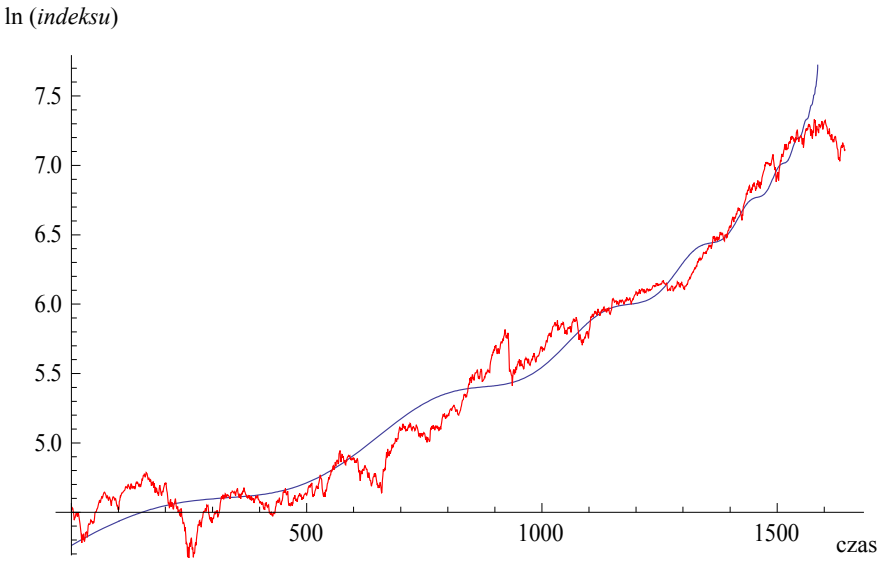
Krytyczny punkt pokrywa się z akumulacją takich oscylacji i w kontekście dynamiki finansowej fakt ten może być wykorzystywany do prognozowania, pod warunkiem że λ jest dobrze zdefiniowana i stała.

W pracy [11] S. Drożdż i inni analizowali zachowanie indeksu Standard & Poor's w latach 1970-2002. Funkcję periodyczną P z równania (6) wybrano jako pierwsze wyrażenie z rozwinięcia Fouriera

$$P\left(\frac{\ln(x)}{\ln(\lambda)}\right) = A + B\cos\left(\omega \ln(x) + \phi\right) \tag{9}$$

gdzie $\omega = 2\pi/\ln(\lambda)$. Wybierając $\lambda = 2$ próbowano otrzymać najlepszą reprezentację opisującą strukturę oscylacyjną na rzeczywistym rynku. Wynik jaki otrzymano wykazuje rosnący trend aż do 1 września 2000 roku.

Interesujący jest fakt, że słynny Czarny Poniedziałek z 19 października 1987 roku doskonale pasuje do rozważanego schematu i stanowi jeden z widocznych log-periodycznych prekursorów poważniejszego globalnego krachu, który rozpoczął się 1 września 2000 roku (rys. 2).

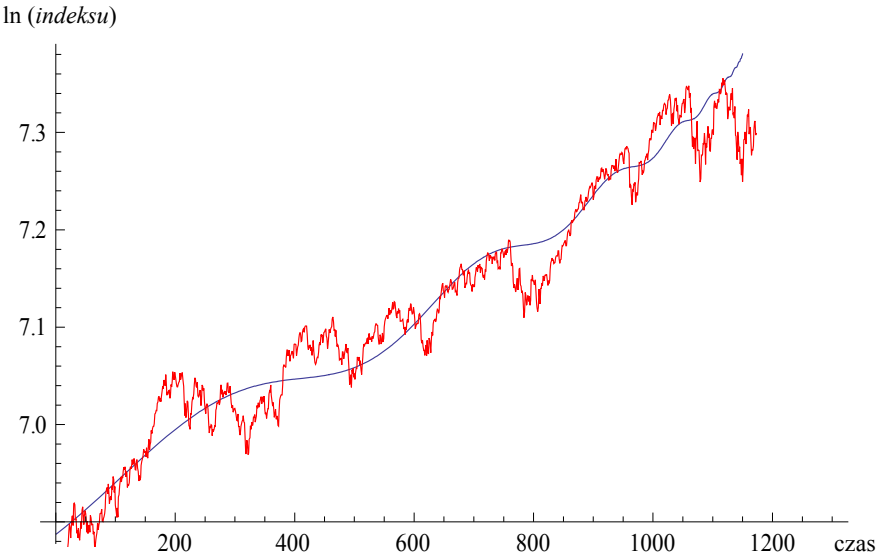


Punkt 930 na osi czasu odpowiada dacie 19.10.1987, natomiast punkt 1602 to 01.09.2000.

Rys. 2. Logarytm naturalny indeksu S&P500 w latach 1970-2001 oraz odpowiadająca mu funkcja log-periodyczna

Bliższe badania w pobliżu tego punktu (poprzez powiększenie wykresu wokół niego) pokazały dwa inne istotne elementy. Po pierwsze, moduł cosinusa w równaniu (12) zapewnia lepsze przedstawienie log-periodycznych modulacji. Po drugie, dostarczają one niezależnego dowodu, że rzeczywista data T_c, oznaczająca odwrócenie rosnącego globalnego trendu, przypada na początek września 2000 roku.

Na rys. 3 przedstawiona jest funkcja log-periodyczna odpowiadająca zmianom indeksu S&P500 w latach 2003-2008. Narastające oscylacje sugerują, iż pod koniec 2007 roku nastąpi odwrócenie rosnącego trendu.



Punkt 1150 odpowiada dacie 26.11.2007.

Rys. 3. Logarytm naturalny indeksu S&P500 w latach 2003-2008

Funkcja log-periodyczna rozważana była również dla innych indeksów giełdowych oraz dla ceny złota i ropy naftowej [1; 2; 11; 13]. Uzyskane wyniki świadczą o uniwersalności finansowej log-periodyczności, dlatego analizę tego zjawiska rozszerzono na dłuży okres rozważając indeks S&P500 od 1800 roku. W wyniku otrzymano dwa wyraźne spadki indeksu w latach 30. i 70. XX wieku oraz kolejny spadek rozpoczynający się w 2000 roku. Co więcej, ekstrapolacja w czasie wykazała, że w okolicach 2025 roku trend indeksu ulegnie spadkowi w takiej skali, jak nigdy dotąd.

Wnioski

Pomimo prostoty gry mniejszościowej okazuje się, że jest ona dobrym punktem startowym do modelowania rynku. Poprzez usuwanie kolejnych nie-realistycznych elementów modelu otrzymujemy krok po kroku rzeczywiste cechy rynku. Gra mniejszościowa jest też znakomitym narzędziem do badania oddziaływań między różnymi typami graczy oraz efektywności, która powinna zależeć od rodzaju rozważanych graczy.

Analiza dotycząca dynamiki rynku finansowego pokazuje, iż rzeczywiście na rynku tym występuje element zachowania log-periodycznego; otrzymuje się zgodność między rzeczywistością a modelem teoretycznym, co pozwala z większą wiarygodnością ekstrapolować w czasie. Należy jednak pamiętać, że log-periodyczność nie jest odporna na zewnętrzne czynniki, takie jak nieoczekiwana wojna czy wydarzenia polityczne, które zakłócają strukturę hierarchiczną organizacji rynku finansowego.

Literatura

1. Bartolozzi M., Drożdż S., Leinweber D.B., Speth J., Thomas A.W. (2000). Self-Similar Log-Periodic Structures in Western Stock Markets from. arXiv:cond-mat/0501513.
2. Cajueiro D.O., Tabak B.M., Werneck F.K. (2009). Can we Predict Crashes? The Case of the Brazilian Stock Market. Physica A 388, 1603-1609.
3. Challet D. and Zhang Y.-C. (1997). Emergence of Cooperation and Organization in an Evolutionary Game. Physica A 246, 407.
4. Challet D. and Zhang Y.-C. (1998). On the Minority Game: Analytical and Numerical Studies. Physica A 256, 514-532.
5. Challet D. and Marsili M. (1999). Symmetry Breaking and Phase Transition in the Minority Game. Phys. Rev. E 60, R6271.
6. Challet D., Marsili M. and Zhang Y.-C. (2000). Modeling Market Mechanism with Minority Game. Physica A 276, 284-315.
7. Challet D., Marsili M. and Zecchina R. (2000). Statistical Mechanics of Heterogeneous Agents: Minority Games. Phys. Rev. Lett. 84, 1824-1827.
8. Challet D., Chessa A., Marsili M., Zhang Y.-C. (2001). From Minority Games to Real Markets. Quant. Fin. 1, 168-176.
9. Challet D., Marsili M., Zhang Y.-C. (2001). Stylized Facts of Financial Markets and Market Crashes in Minority Games. Physica A 294, 514-524.
10. Challet D. (2007). Inter-pattern Speculation: Beyond Minority, Majority and $-games. J. of Economic Dynamics & Control, 32, 85-100.
11. Drożdż S., Grümmer F., Ruf F., Speth J. (2003). Log-perodic Self-similarity: an Emerging Financial Law. Physica A 324, 174-182.
12. Drożdż S., Kwapień J., Oświęcimka P. (2008). Criticality Characteristics of Current Oil Price Dynamics. Acta Physica Polonica A 114, 699-702.
13. Drożdż S., Grümmer F., Ruf F., Speth J. Prediction Oriented Variant of Financial Log-periodicity and Speculating about the Stock Market Development until 2010. arXiv:physics/0503006.
14. Martino, Marsili M. (2006). Statistical Mechanics of Socio-Economic Systems with Heterogeneous Agents. arXiv: physics/0606107v1.

15. Mosetti G., Challet D., Zhang Yi-Cheng (2005). Heterogeneous Timescales in Minority Games. Physica A 365, 529-542.
16. Savit R., Manuca R., and Riolo R. (1999). Adaptive Competition, Market Efficiency, and Phase Transitions. Physical Review Letter, 82, 2203-2206.
17. Sornette D., Johansen A., Bouchaud J.-P. (1996). Stock Market Crashes, Precursors and Replicas. Journal de Physique I France, 6, 167-175.
18. Sornette D., Johansen A. (1998). A Hierarchical Model of Financial Crashes. Physica A 261, 581-598.
19. Yeung C.H. and Hang Y.-C. (2008). Minority Games. arXiv: physics.soc-ph/0811.1479.

Alicja Ganczarek-Gamrot
Akademia Ekonomiczna w Katowicach

POMIAR RYZYKA W SYSTEMIE CENY JEDNOLITEJ NA TOWAROWEJ GIEŁDZIE ENERGII

Wstęp

Od 8 lutego 2008 roku Towarowa Giełda Energii S.A. (TGE S.A.) wprowadziła zmiany w harmonogramie sesji Rynku Dnia Następnego (RDN). Sesja RDN składała się obecnie z dwóch aukcji, na których ustalana jest cena jednolita, notowań ciągłych i transakcji pozasesyjnych.

Notowania odbywają się nadal codziennie od poniedziałku do niedzieli. Również niezmiennie uczestnikami RDN mogą być zarówno producenci, jak i odbiorcy energii elektrycznej. Mogą oni zajmować zarówno pozycję krótką, jak i długą bez względu na to, czy reprezentują grupę producentów czy odbiorców energii elektrycznej. „RDN przeznaczony jest dla tych spółek, które chcą w sposób aktywny i bezpieczny na bieżąco domykać swoje portfele zakupów/sprzedaży energii elektrycznej w poszczególnych godzinach doby" [10].

Obecnie na RDN oferowane są 24 (23 w przypadku zmiany czasu z zimowego na letni lub 25 w przypadku zmiany czasu z letniego na zimowy) kontrakty na każdą godzinę doby oraz trzy kontrakty blokowe: PASMO (cała doba), EUROSZCZYT (godziny od 8 do 22) oraz OFFPEAK (godziny od 23 do 7).

Notowania odbywają się w ciągu całego dnia według ustalonego przez TGE harmonogramu sesji. Natomiast do publicznej wiadomości podawane są jedynie ustalone niezależnie w dwóch aukcjach, tzw. ceny jednolite na kontrakty w każdej godzinie doby, kursy średnioważone wszystkich zawartych transakcji w każdej godzinie doby oraz indeksy.

Ogólnie dostępne ceny jednolite ustalane są dwukrotnie w ciągu doby:

- Aukcja 1 o godzinie 8:00,
- Aukcja 2 o godzinie 10:30.

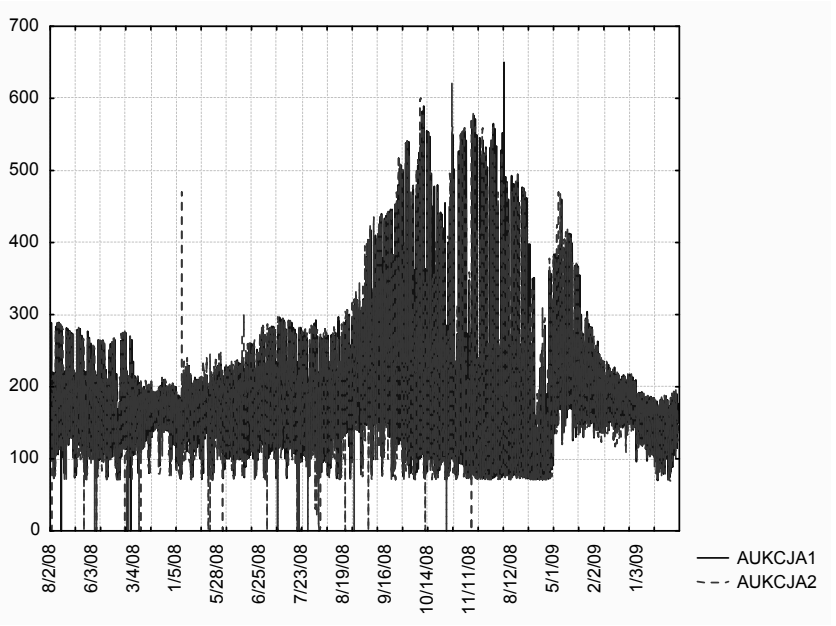
Na każdej aukcji ustalane są ceny niezależnie na każdą godzinę doby. Są to ceny równowagi złożonych ofert kupna i sprzedaży ceny energii elektrycznej. Realizacja zawartych kontraktów w systemie ceny jednolitej następuje kolejnego dnia. Pojawia się zatem pytanie, która z aukcji jest korzystniejsza dla uczestnika rynku w zależności od pozycji jaką zajmuje.

Bazując na notowaniach RDN z pierwszego roku funkcjonowania nowego harmonogramu sesji TGE, w pracy zostanie przeprowadzona analiza porównawcza ryzyka podejmowanego przez uczestników rynku w zależności od momentu otwarcia pozycji i rodzaju pozycji jaką zajmują. W tym celu przeprowadzono analizę porównawczą notowań na dwóch aukcjach. Porównano rozkłady cen oraz logarytmiczne godzinne stopy zwrotu cen na RDN z dwóch niezależnych aukcji. Wykorzystując liniowe sezonowe modele SARIMA oraz niestacjonarne modele GARCH oszacowano poziom ryzyka zmiany ceny z jednogodzinnym horyzontem czasu na każdej aukcji niezależnie. Ryzyko zostało oszacowane za pomocą kwantylowej miary zagrożenia Value-at-Risk (VaR), na podstawie której dokonano porównania poziomu ryzyka podejmowanego przez uczestników rynku w zależności od momentu otwarcia pozycji i rodzaju pozycji na RDN.

W drugi etapie badania rozważono sytuację, w której uczestnik rynku otwiera pozycję na Aukcji 1 i zamyka ją kontraktem przeciwnym na Aukcji 2 tego samego dnia. Ryzyko tej strategii oszacowano również na podstawie VaR.

1. Analiza porównawcza rozkładów notowań rozważanych aukcji

Na rys. 1 przedstawiono szeregi czasowe cen notowanych na RDN w czasie Aukcji 1 i Aukcji 2 od 8 lutego 2008 roku do 28 marca 2009 roku. Szeregi te w badanym okresie kształtowały się bardzo podobnie. Na rys. 1 możemy zaobserwować wysoki poziom cen na TGE od września do końca 2008 roku.



Rys. 1. Notowania na RDN w okresie od 8.02.2008 do 28.03.2009

Występujące przede wszystkim w pierwszym etapie funkcjonowania Aukcji 2 braki notowań usunięto parami razem z odpowiadającymi im notowaniami kontraktów zawartych na te same godziny na Aukcji 1. W tabeli 1 przedstawiono parametry rozkładów cen oraz logarytmicznych godzinnych stóp zwrotu cen dla ograniczonego brakami notowań okresu

$$z_t = \ln(\frac{y_t}{y_{t-1}}) \tag{1}$$

gdzie y_t oznacza cenę energii elektrycznej w godzinie t.

Tabela 1

Parametry rozkładu cen oraz logarytmicznych godzinnych stóp zwrotu cen notowanych na RDN w okresie od 8.02.08 do 28.03.09

Parametry	Aukcja 1		Aukcja 2	
	y_t	z_t	y_t	z_t
1	2	3	4	5
n	9898	9897	9898	9897
Średnia	193,93	0	192,5	0
Min	70	-0,8554	70	-1,0876
Max	650	0,8979	620	1,0566

cd. tabeli 1

1	2	3	4	5
$Q_{0,05}$	80	-0,2546	76	-0,2758
$Q_{0,95}$	310	0,2912	307	0,3193
Odch.std	73,61	0,1666	73,71	0,1818
Skośność	1,29	0,65	1,06	0,50
Kurtoza	4,14	4,81	3,18	5,15

Rozkłady cen, jak i rozkłady stóp zwrotu różnią się istotnie od rozkładu normalnego. Charakteryzują się wyraźną asymetrią prawostronną i leptokurtycznością. Rozkłady pomiędzy aukcjami nie różnią się znacznie między sobą, jednak na Aukcji 2 w porównaniu z Aukcją 1 obserwujemy większą zmienność, wyrażoną zarówno odchyleniem standardowym, jak i zakresem zmienności szeregu czasowego logarytmicznych stóp zwrotu. Zatem można już na pierwszym etapie analizy stwierdzić, że większe ryzyko zmiany ceny obserwujemy na Aukcji 2 ustalania ceny jednolitej.

2. Estymacja i porównanie ryzyka podejmowanego w zależności od aukcji

Szeregi czasowe na RDN charakteryzują się wyraźną cyklicznością w ciągu dnia, tygodnia oraz roczną sezonowością [5; 6]. W estymacji ryzyka na rynku energii elektrycznej należy uwzględnić cykliczność cen, a tym samy cykliczność zmienności, które bezpośrednio zależą od cykliczności zapotrzebowania na energię elektryczną. W pracy do modelowania cykliczności wykorzystano model autoregresji i średniej ruchomej SARIMA (Seasonal Auto – Regressive Integrated Moving Average) (p,d,q)× (P, D, Q) [1]

$$p(B)P_s(B^s)\nabla_s^d z_t = q(B)Q_s(B^s)\varepsilon_t \tag{2}$$

gdzie:

$$p(B) = 1 - \sum_{i=1}^{p} p_i B^i, \qquad P_s(B) = 1 - \sum_{i=1}^{P} P_{si} B^i,$$

$$q(B) = 1 - \sum_{i=1}^{q} q_i B^i, \qquad Q_s(B) = 1 - \sum_{i=1}^{Q} Q_{si} B^i,$$

s – opóźnienie sezonowe,

d – rząd zintegrowania szeregu,

z_t – empiryczne wartości szeregu (w pracy logarytmiczne stopy zwrotu cen energii elektrycznej),

B – operator przesunięcia $B^s z_t = z_{t-s}$,

∇ – operator różnicowy $\nabla^s z_t = z_t - z_{t-s} = (1 - B^s) z_t$,

ε_t – reszty modelu.

W tabeli 2 zamieszczono wyniki estymacji metodą największej wiarygodności parametrów równania (2) dla dwóch szeregów logarytmicznych godzinnych stóp zwrotu cen energii elektrycznej notowanych na Aukcji 1 oraz Aukcji 2. Szeregi nie wykazywały istotnego trendu (rząd integracji szeregu d = 0). Na podstawie obserwacji funkcji autokorelacji oraz autokorelacji cząstkowej, zróżnicowano szeregi z opóźnieniem sezonowym s = 24, a następnie dobrano rząd opóźnień części autokorelacji i średniej ruchomej[1]. Otrzymane modele różnią się nieznacznie między sobą. Dla szeregu czasowego z Aukcji 2 parametr przy średniej ruchomej z opóźnieniem rzędu dwa okazał się statystycznie nieistotny. Pozostałe parametry modeli są statystycznie istotne, ponadto bardzo do siebie zbliżone. Co oznacza, że na obydwu aukcjach procesy zmienności przebiegają podobnie, ale na poziomie ich wartości oczekiwanych. Proces notowań na Aukcji 2 ma nieco „krótszą pamięć", ponieważ sięga ona tylko jednego okresu wstecz.

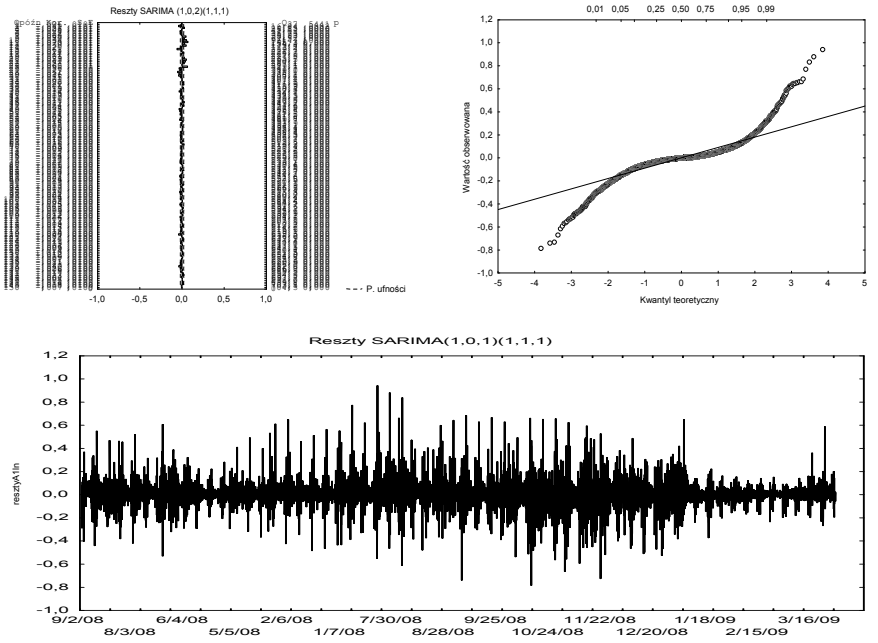
Tabela 2

Wyniki estymacji modeli SARIMA

s =24	Parametry	Oceny parametrów	p-wartość
AUKCA 1 SARIMA (1,0,2)(1,1,1) MS= 0,0098	p(1)	0,83	<0,01
	q(1)	0,86	<0,01
	q(2)	0,13	<0,01
	Ps(1)	0,21	<0,01
	Qs(1)	0,75	<0,01
AUKCJA 2 SARIMA (1,0,1)(1,1,1) MS=0 ,0154	p(1)	0,82	<0,01
	q(1)	0,99	<0,01
	Ps(1)	0,15	<0,01
	Qs(1)	0,79	<0,01

Analizując reszty modeli SARIMA brak jest podstaw do odrzucenia hipotezy o braku autokorelacji reszt. Niemniej jednak rozkłady reszt różnią się istotnie od rozkładu normalnego. Charakteryzują się asymetrią prawostronną, leptokurtozą, grubymi ogonami. Towarzyszy im również efekt skupiania się zmienności (rys. 2).

[1] Występująca w szeregach tygodniowa cykliczność danych została zaburzona przez wyeliminowanie dni brakujących notowań na Aukcji 2.



Rys. 2. Charakterystyka reszt modeli SARIMA

Ze względu na występujący w procesie reszt modeli SARIMA efekt ARCH, w kolejnym kroku szeregi te opisano za pomocą modelu [3; 4]

$$\varepsilon_t = \xi_t \sqrt{\sigma_t^2} \tag{3}$$

gdzie:

ξ_t ~ biały szum (o dowolnym rozkładzie z $E(\xi_t) = 0$, $E(\xi_t^2)=1$),

σ_t^2 – warunkowa wariancja reszt modelu (2).

Kierując się kryterium informacyjnym Schwarza [9] spośród bogatej klasy modeli GARCH [5] do opisania warunkowej wariancji reszt modeli SARIMA wykorzystano uogólniony model FIGARCH – HYGARCH(p,q) [2], którego warunkową wariancję można zapisać następująco

$$\sigma_t^2 = \omega[1-\beta(B)]^{-1} + \{1-[1-\beta(B)]^{-1}[1-\alpha(B)-\beta(B)]\{1+\alpha_H[(1-B)^d]\}\}\varepsilon_t^2 \tag{4}$$

gdzie:

ω – stały poziom wariancji,

d – rząd integracji szeregu,

α_H – parametr odpowiedzialny za długą pamięć szeregu (jeżeli $\alpha_H < 1$ szereg jest stacjonarny).

Spośród rozkładów: normalnego, GED, t-Studenta oraz skośnego rozkładu t-Studenta do opisu rozkładu białego szumu ξ_t najlepiej dopasowany był rozkład t-Studenta. W tabeli 3 przedstawiono wyniki estymacji metodą największej wiarygodności parametrów modelu (4).

Tabela 3

Wyniki estymacji modeli HYGARCH

Aukcja	Parametr	Ocena parametru	Błąd standardowy	t	p-wartość
1	d	0,4618	0,0198	23,310	<0,01
	α_1	0,7718	0,0091	84,790	<0,01
	β_1	0,6564	0,0159	41,240	<0,01
	ν_{St}	2,1285	0,0497	42,830	<0,01
	α_H	2,3453	0,3452	6,794	<0,01
2	d	0,3882	0,0212	18,330	<0,01
	α_1	0,7902	0,0109	72,230	<0,01
	β_1	0,6568	0,0188	34,920	<0,01
	ν_{St}	2,1210	0,0445	47,700	<0,01
	α_H	2,4592	0,3324	7,3980	<0,01

Wszystkie parametry obydwóch równań są statystycznie istotne, ponadto wartości parametrów są zbliżone do siebie, co świadczy o tym, że procesy na Aukcji 1 i 2 kształtują się podobnie nie tylko na poziomie wartości oczekiwanej, ale również na poziomie wariancji. Ponadto, obydwa procesy są niestacjonarne. W związku z czym nie jest wskazane prognozowanie na dłuższy okres w oparciu o szacowane modele.

Oszacowane modele wariancji warunkowej wykorzystano do estymacji wartości narażonej na ryzyko [8]

$$VaR_{(t+1)\alpha} = (z_{(t+1)\alpha} + \mu) y_t \tag{5}$$

$$z_{t\alpha} = F^{-1}(\alpha)\sqrt{\sigma_t^2} \tag{6}$$

gdzie:

$F^{-1}(\alpha)$ – kwantyl rzędu α rozkładu teoretycznego,

μ – wartość oczekiwana procesu,

y_0 – bieżąca cena waloru (cena 1MWh energii elektrycznej).

Analizując wzór (5) widzimy, że wartość narażona na ryzyko estymowana za pomocą modeli z warunkową wariancją nie jest pojedynczym parametrem, lecz szeregiem czasowym zależnym od poziomu tolerancji α i rozkładu białego szumu ξ_t w modelu (3), horyzontu, na który szacowany jest VaR oraz historii zmienności. W związku z tym, że VaR jest szeregiem czasowym w tabeli 4 zaprezentowano średnie wartości VaR oszacowane na podstawie modelu (5) z jednogodzinnym horyzontem czasu trwania inwestycji. Interpretując uzyskane wyniki możemy powiedzieć, że przeciętnie z prawdopodobieństwem 0,95 cena energii elektrycznej zakupionej na Aukcji 1 w kolejnej godzinie nie będzie niższa od ceny bieżącej o więcej niż 25,79 PLN/MWh, jak również nie będzie wyższa od ceny bieżącej o więcej niż 25,79 PLN/MWh z prawdopodobieństwem 0,95. Natomiast cena energii elektrycznej zakupionej na Aukcji 2 z prawdopodobieństwem 0,95 w kolejnej godzinie nie będzie niższa od ceny bieżącej przeciętnie o więcej niż 32,89 PLN/MWh, jak również nie będzie wyższa od ceny bieżącej o więcej niż 32,89 PLN/MWh z prawdopodobieństwem 0,95.

W tabeli 4 zamieszczono również p-wartości testu przekroczeń Kupca [7], na podstawie których brak jest podstaw do odrzucenia hipotezy, że udział przekroczeń VaR w badanym okresie jest zgodny z oczekiwanym (w pracy 5%).

Tabela 4

Wyniki estymacji wartości VaR z horyzontem godzinnym

Aukcja	α	Średnia VaR [PLN/MWh]	p-wartość testu Kupca
1	0,05	-25,79	0,0727
	0,95	25,79	0,1939
2	0,05	-32,89	0,0594
	0,95	32,89	0,0882

Analizując przeciętną wartość oszacowanego VaR (tabela 4) możemy potwierdzić wcześniejszy wniosek, że zawierając transakcję podczas Aukcji 2 uczestnicy RDN narażają się na większe ryzyko. Jednak należy pamiętać, że tak oszacowane VaR jest szeregiem czasowym i poziom ryzyka na obydwu aukcjach może również zmieniać się w czasie. Dla zobrazowania problemu na rys. 6 przedstawiono VaR oszacowane na każdą godzinę 28 marca 2009 roku.

3. Estymacja ryzyka podejmowanego między aukcjami

Aby odpowiedzieć na pytanie, na jakie ryzyko narażony jest uczestnik RDN, który zawierając transakcje na Aukcji 1 zamierza tego samego dnia zawrzeć pozycje przeciwną podczas Aukcji 2, w tej części pracy podano analizie następujący szereg

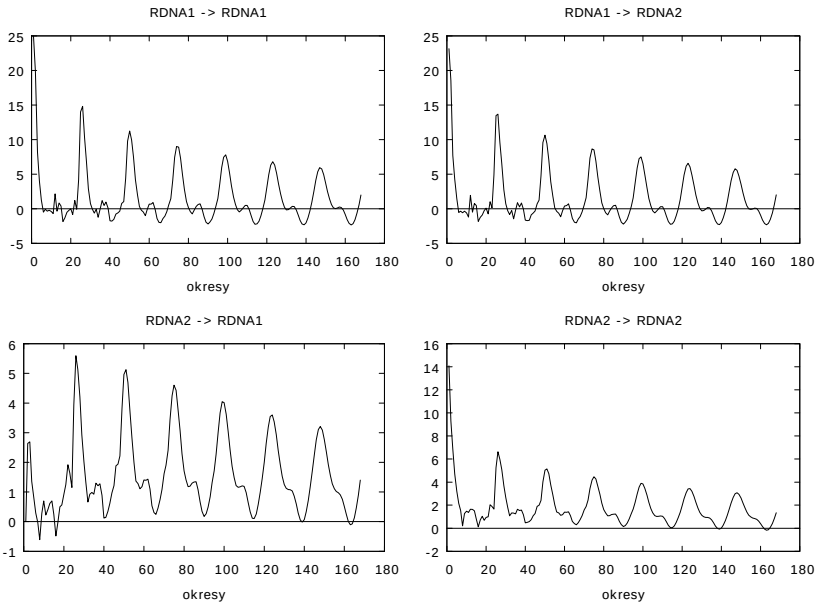
$$u_t = \frac{y_{2t} - y_{1t}}{y_{1t}} \tag{7}$$

gdzie:

y_{1t} – cena 1MWh energii elektrycznej ustalona na Aukcji 1 w godzinie t,
y_{2t} – cena 1MWh energii elektrycznej ustalona na Aukcji 2 w godzinie t.

W poprzednim rozdziale poddano analizie ryzyko zmiany ceny niezależnie na Aukcji 1 i Aukcji 2. Wykazano, że na Aukcji 2 ryzyko zmiany ceny jest przeciętnie biorąc w ciągu badanego okresu większe. Jednakże rozkłady tych szeregów, jak również procesy SARIMA-HYGARCH są bardzo podobne, ponadto są silnie ze sobą skorelowane (współczynnik korelacji liniowej Pearsona między dwoma szeregami cen wynosi 0,971). Możemy powiedzieć, że szeregi y_{1t} y_{2t} są ze sobą zintegrowane – mają wspólną ścieżkę rozwoju. Wspólna ścieżka rozwoju dla obydwu szeregów nie jest warunkowana trendem jak w przypadku danych finansowych, ale cyklicznością. Na rys. 3 przedstawiono funkcję odpowiedzi na impuls pomiędzy analizowanymi cenami (y_{1t} -RDNA1 i y_{2t} RDNA2).

Analizując funkcję odpowiedzi na impuls widzimy, że jej wartości zmieniają się w czasie odzwierciedlając cykliczność dzienną. Z dnia na dzień wpływ wartości z poszczególnych aukcji maleje, ale wciąż jest dodatni. Nieco silniejszy wpływ obserwujemy w przypadku oddziaływania Aukcji 1 na Aukcję 2 niż odwrotnie.



Rys. 3. Funkcja odpowiedzi na impuls

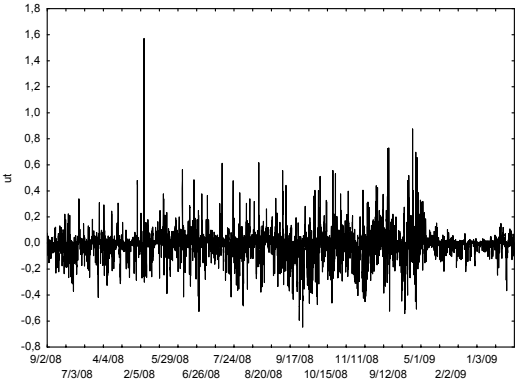
Rozkład szeregu u_t charakteryzuje się również asymetrią prawostronną, leptokurtycznścią i grubymi ogonami (tabela 5).

Tabela 5

Parametry rozkładu szeregu u_t

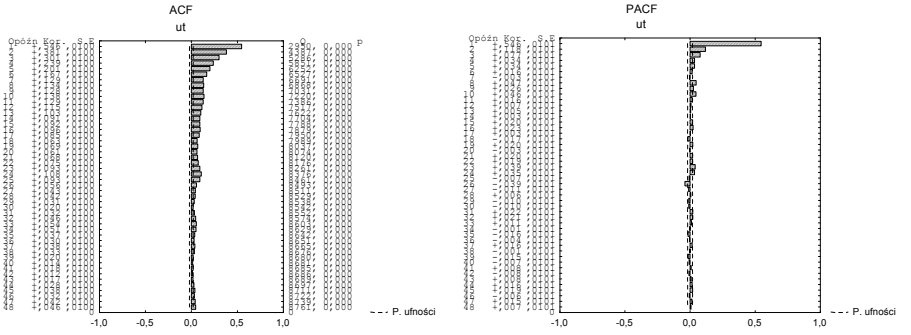
Parametry	u_t
n	9898
Średnia	-0,0071
Min	-0,6462
Max	1,5711
$Q_{0,05}$	-0,1667
$Q_{0,95}$	0,1055
Odch.std	0,0927
Skośność	0,41
Kurtoza	17,73

Wartość oczekiwana transakcji zamykanej na Aukcji 2 jest ujemna, co przemawia na korzyść otwierania pozycji krótkich na Aukcji 1, natomiast na Aukcji 2 otwierania pozycji długich. W badanym okresie można jednak również zaobserwować godziny wyższych cen na Aukcji 2 (rys. 4).



Rys. 4. Szereg czasowy u_t

Szereg czasowy dany równością (7) powinien być szeregiem z wyeliminowaną sezonowością i tak rzeczywiście jest (rys. 5). Analizując funkcję autokorelacji oraz autokorelacji cząstkowej szeregu u_t, do opisu procesu danego równością (7) możemy zaproponować model AR(3).



Rys. 5. Funkcja autokorelacji oraz autokorelacji cząstkowej procesu u_t,

W tabeli 6 zaprezentowano wyniki estymacji modelu AR(3). Wszystkie parametry modelu są statystycznie istotne. Jednak podobnie jak w przypadku reszt modeli SARIMA i tu występuje efekt skupiania się zmienności. Analogicznie jak poprzednio kierując się kryterium Schwarza [9] do estymacji warunkowej wariancji wybrano model GARCH z rozkładem GED [4]

$$\sigma_t^2 = \omega + \alpha(B)\varepsilon_t^2 + \beta(B)\sigma_t^2 \tag{8}$$

W tabeli 6 zamieszczono wyniki estymacji modelu warunkowej wariancji (8). Wszystkie parametry są statystycznie istotne. Ponadto, $\alpha_1 + \beta_1 = 1$, co oznacza, że mamy tu do czynienia z zintegrowanym modelem IGARCH, którego bezwarunkowa wariancja nie istnieje, ale sam model jest stacjonarny w szerszym sensie.

Tabela 6

Wyniki estymacji modeli dla szeregu u_t,

Model	Parametry	Oceny parametrów	p-wartość
AR(3) MS= 0,0059	p(1)	0,4735	<0,01
	p(2)	0,0817	<0,01
	p(3)	0,0785	<0,01
IGARCH(1,1)	α_1	0,0781	<0,01
	β_1	0,9219	<0,01
	v_{GED}	0,6087	<0,01

Podobnie jak w rozdziale poprzednim, estymowano wartość VaR za pomocą równania (5) z warunkową wariancją (8) i rozkładem GED. W tabeli 7 zaprezentowano średnie wartości VaR. Interpretując uzyskane wyniki możemy powiedzieć, że przeciętnie z prawdopodobieństwem 0,95 cena energii elektrycznej zakupionej w ustalonej godzinie na Aukcji 1 w tej samej godzinie na Aukcji 2 nie będzie niższa o więcej niż 20,37 PLN/MWh, jak również nie będzie wyższa od ceny bieżącej o więcej niż 20,37 PLN/MWh z prawdopodobieństwem 0,95.

Tabela 7

Wyniki estymacji wartości VaR na kolejną aukcję

α	Średnia VaR [PLN/MWh]	p-wartość testu Kupca
0,05	-20,37	<0,01
0,95	20,37	0,301

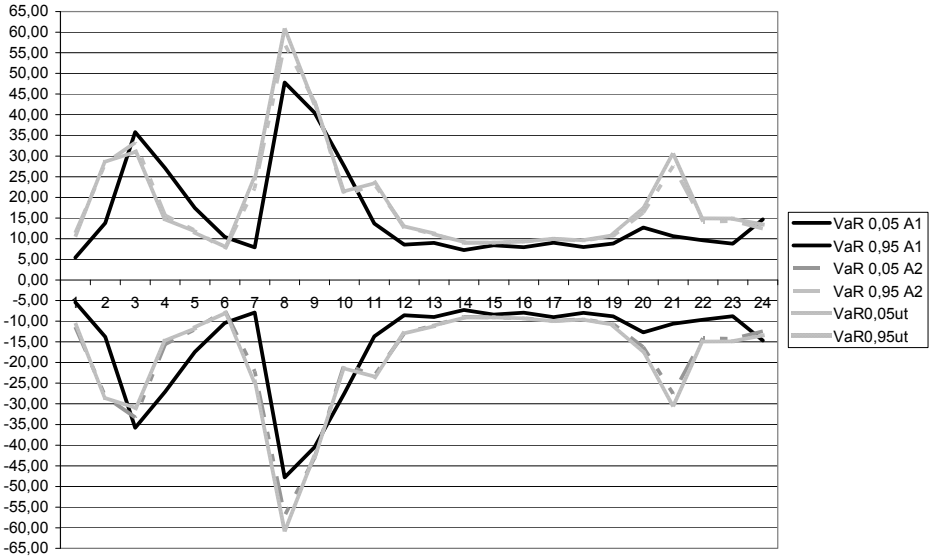
W tabeli 7 zamieszczono p-wartości testu przekroczeń Kupca [7]. W przypadku pozycji krótkiej brak jest podstaw do odrzucenia hipotezy, że udział przekroczeń VaR w badanym okresie jest zgodny z oczekiwanym (w pracy 5%). Natomiast w przypadku pozycji długiej cena energii elektrycznej w ustalonej godzinie doby na Aukcji 2 była wyższa niż cena energii elektrycznej na Aukcji 1 więcej niż w 5% przypadków.

Podsumowanie

Podsumowując wyniki uzyskane w pracy możemy powiedzieć, że przeciętnie większym ryzykiem zmiany wartości oczekiwanej transakcji obarczone są transakcje zawierane na Aukcji 2. Z drugiej części analizy wynika, że ceny ustalane na Aukcji 2 są przeciętnie niższe niż ceny ustalane na Aukcji 1, w związku z czym Aukcja 2 może być korzystniejsza dla uczestników rynku zajmujących długą pozycję. Biorąc jednak pod uwagę wyniki testów przekroczeń Kupca (tabela 7) cena energii elektrycznej w ustalonej godzinie doby w więcej niż 5% przypadków była wyższa niż oszacowana wartość VaR (tabela 7).

Analizując przeciętną wartość oszacowanego VaR (tabele 4 i 7) należy pamiętać, że tak oszacowane VaR jest szeregiem czasowym. Na rys. 6 przedstawiono VaR oszacowane na każdą godzinę 28 marca 2009 roku. VaR oszacowano niezależnie na podstawie modelu (5) dla transakcji zawieranych na Aukcji 1 (VaR0,05A1 i VaR0,95A1) na Aukcji 2 (VaR0,05A2 i VaR0,95A2) oraz w ciągu jednego dnia ($VaR0{,}05u_t$ i $VaR0{,}95\ u_t$).

Analizując fragmenty szeregów VaR możemy zauważyć, że w ostatnim dniu analizy największym ryzykiem zmiany ceny były obarczone transakcje w godzinach nocnych i przedpołudniowych. Porównując ryzyko między aukcjami w tym dniu możemy wskazać godziny, w których większe ryzyko podejmują uczestnicy zawierający transakcje na Aukcji 1 i takie godziny, w których większe ryzyko podejmują strony otwierające pozycję na Aukcji 2.



Rys. 6. VaR w każdej godzinie doby 28.03.2009

Analizując wartości VaR oszacowane na 28 marca 2009 roku można zauważyć, że ryzyko oszacowane niezależnie na podstawie notowań na Aukcji 2 jest odrobinę niższe, ale zbliżone do ryzyka na jakie mogą być narażeni uczestnicy otwierający pozycję na Aukcji 1 i zamykający pozycję tego samego dnia na Aukcji 2.

Podsumowując wyniki uzyskane w pracy możemy stwierdzić, że do oceny ryzyka na RDN w systemie ceny jednolitej można wykorzystać VaR. Wiele prac empirycznych (np. [5; 6]) pokazało, że w estymacji tej miary ryzyka należy wziąć pod uwagę cykliczność i autokorelację poziomu oraz zmienności cen energii elektrycznej. Należy również zwrócić uwagę, że otrzymane w pracy oszacowania modeli szeregów czasowych, poziomy ryzyka i zaprezentowane na ich podstawie wnioski dotyczą konkretnego okresu badawczego. W związku z czym, do pomiaru ryzyka w systemie ceny jednolitej na TGE można wykorzystać VaR estymowane z uwzględnieniem cykliczności oraz autokorelacji, jednak do oceny ryzyka należy przeprowadzić estymację na uaktualnionych szeregach z rynku energii elektrycznej.

Literatura

1. Brockwell P.J., Davis R.A. (1996). Introduction to Time Series and Forecasting. Springer-Verlag, New York.
2. Davidson J. (2004). Moment and Memory Properties of Linear Conditional Heteroscedastity Models and a new Model. Journal of Business and Economics Statistics, 22, 16-29.
3. Engle R.F. (1982). Autoregressive Conditional Heteroscedasticy with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50, 987-1007.
4. Engle R.F., Bollerslev T. (1986). Modeling the Persistence of Conditional Variance. Econometric Review, 5, 1-50.
5. Ganczarek A. (2008). Weryfikacja modeli z grupy GARCH na dobowo-godzinnych rynkach energii elektrycznej w Polsce. Rynek kapitałowy. Skuteczne inwestowanie. Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania, 9, 524-536.
6. Ganczarek A. (2008). Choosing between the Skewed Distribution and the Skewed Model in General Autoregressive Conditional Heteroscedasticity. Proceedings of 26-th International Conference Mathematical Methods in Economics 2008, 132-139.
7. Kupiec P. (1995). Techniques for Verifying the Accuracy of Risk Management Models. Journal of Derivatives, 2, 173-184.

8. Piontek K. (2001). Heteroskedastyczność rozkładu stóp zwrotu a koncepcja pomiaru ryzyka metodą VaR. W: Modelowanie preferencji a ryzyko. Red. T. Trzaskalik. AE, Katowice, 339-350.
9. Schwarz G. (1978). Estimating the Dimension of a Model. The Annals of Statistics, 6, 461-464.
10. www.polpx.pl

Dominik Kudyba

Tadeusz Trzaskalik

Akademia Ekonomiczna w Katowicach

ZASTOSOWANIE SYNCHRONICZNEJ WERSJI METODY NIMBUS DO USTALENIA OPTYMALNEJ STRUKTURY PORTFELA AKCJI

Wstęp

Problematyka zastosowania wielokryterialnego wspomagania decyzji do wyboru portfela rozpatrywana była, począwszy od lat 80. XX wieku, w wielu pracach Przegląd publikacji dotyczących wielokryterialnego wspomagania decyzji w finansach, między innymi w dziedzinie wspomagania wyboru portfela papierów wartościowych, można znaleźć w pracy [8].

Zagadnienie wielokryterialnego wspomagania wyboru portfela akcji rozpatrywane jest często jako proces składający się z dwóch zasadniczych etapów: selekcji akcji do portfela oraz alokacji aktywów w portfelu [1].

Niniejsza praca poświęcona jest wykorzystaniu metody NIMBUS do zagadnienia alokacji aktywów w portfelu. Metoda NIMBUS (**N**ondifferentiable **I**nteractive **M**ultiobjective **BU**ndle – based optimization **S**ystem) została opracowana przez zespół naukowców fińskich pod kierunkiem prof. K. Miettinen[1] w latach 1993/1994. System WWW-NIMBUS dostępny jest przez stronę internetową [11]. Jest to rozbudowane narzędzie do analizy numerycznej. Daje szerokie możliwości analizowania modeli jednokryterialnych i wielokryterialnych o różnych postaciach algebraicznych. Oferowanych jest pięć optymalizatorów do rozwiązywania zadań. System jest ogólnodostępny i bezpłatny dla celów niekomercyjnych. Opis systemu można znaleźć między innymi w pracach [5; 6; 7; 4].

Celem niniejszej pracy jest zastosowanie synchronicznej metody NIMBUS do ustalenia optymalnej struktury portfela akcji.

[1] Profesor Optymalizacji Przemysłowej Uniwersytetu w Jyväskylä, pracująca w Departamencie Matematycznych Technologii Informacyjnych.

Praca składa się z trzech rozdziałów. W rozdziale pierwszym przedstawiono synchroniczną metodę NIMBUS. W rozdziale drugim, na podstawie danych giełdowych z warszawskiej Giełdy Papierów Wartościowych z okresu 2001-2003, określono optymalną strukturę portfela, uwzględniającą preferencje decydenta (na potrzeby niniejszej pracy w rolę tę wcielił się D. Kudyba). W rozdziale trzecim otrzymane wyniki porównane zostały z wcześniejszym rozwiązaniem (wykorzystującym te same dane), opracowanym przez C. Dominiaka z wykorzystaniem interaktywnej metody satysfakcjonujących poziomów kryteriów [1].

1. Synchroniczna metoda NIMBUS

Synchroniczna metoda NIMBUS należy do grupy metod interaktywnych. W każdej iteracji decydent wyraża swoje preferencje o problemie i wyznacza kierunek poszukiwań rozwiązania tego problemu. Rozpatrywane jest zadanie wielokryterialne

$$\begin{aligned} &MIN \left\{ f_1(x),\ f_2(x), ...,\ f_k(x) \right\} \\ &x \in S \end{aligned} \tag{1}$$

gdzie:

$f_1(x),\ f_2(x), ...,\ f_k(x)$ – funkcje kryterium, takie że $f_k : R^n \to R, k \geq 2$,
$x = [x_1, x_2, ..., x_n]^T$ – wektor zmiennych decyzyjnych,
$S \subset R^n$ – zbiór rozwiązań dopuszczalnych w przestrzeni decyzyjnej.

W optymalizacji wielokryterialnej można analizować wektory wartości kryteriów z^0 w dwóch przestrzeniach: kryterialnej i decyzyjnej. Wektor wartości kryteriów uznaje się za optymalny w przestrzeni kryterialnej, jeżeli nie jest możliwa poprawa wartości żadnego z kryterium bez konieczności pogarszania wartości przynajmniej jednego innego kryterium [9]. Gdy zachodzi taka zależność, to wektor wartości kryteriów z^0 w przestrzeni kryterialnej nazywa się rozwiązaniem niezdominowanym. Rozwiązania niezdominowane spełniają warunek

$$\neg \underset{z \in Z}{\exists}\ z \overset{s}{\succ} z^0$$

gdzie $\overset{s}{\succ}$ jest relacją silnego porządku. Każde rozwiązanie niezdominowane z^0 w przestrzeni kryterialnej z jest obrazem rozwiązania sprawnego w przestrzeni decyzyjnej poprzez wektorową funkcję kryterium.

Synchroniczna wersja metody NIMBUS wykorzystuje trzy wektory do budowy zadań zastępczych oraz do poszukiwania rozwiązań niezdominowanych. Należą do nich:

- wektor idealny z^* (Ideal objective vector) składa się z najlepszych wartości kryteriów. Wartości tego wektora wyznacza się rozwiązując zadanie

$$MIN \left\{ f_i(x) \right\}, \forall\, i = 1, 2, ..., k$$
$$x \in S$$

- wektor nadir (Nadir objective vector) zawiera najgorsze oceny kryteriów w przestrzeni kryterialnej, elementy wektora nadir w zadaniach wielokryterialnej optymalizacji liniowej wyznacza się za pomocą macierzy wypłat (Payoff table)[2],
- wektor utopijny (Utopian objective vector) otrzymuje się na podstawie wzoru

$$z_i^{**} = z_i^* - \sigma_i, i = 1,...,k \tag{2}$$

gdzie:

$\sigma_i > 0$ – relatywnie mały skalar.

Lokalna klasyfikacja kryteriów używana jest przez decydenta do wyrażania preferencji podczas rozwiązywania problemu. W każdej iteracji algorytmu synchronicznej metody NIMBUS prosi się decydenta o zaklasyfikowanie indeksu każdej funkcji celu ze względu na wartość reprezentowaną w rozpatrywanym rozwiązaniu do jednego z pięciu następujących zbiorów indeksów.

a) $I^<$ – wartości kryteriów, których indeksy znajdują się w tym zbiorze, powinny być zmniejszone w stosunku do obecnej ich wartości,

b) $I^\leq$ – do tego zbioru przydziela się indeksy tych kryteriów, których wartości powinny być zmniejszone do oznaczonego przez decydenta poziomu aspiracji $\hat{z}_i$, przy czym poziom aspiracji musi być mniejszy od wartości i-tego kryterium w h-tej iteracji: $\hat{z}_i < f_i(x^h)$,

c) $I^=$ – do tego zbioru zaklasyfikowane zostają indeksy kryteriów, których wartości są obecnie na satysfakcjonującym poziomie i jeżeli nie ma takiej konieczności, nie powinny być już dalej zmieniane,

d) $I^\geq$ – indeksy dodawane do tego zbioru należą do kryteriów, których wartości mogą być zwiększone do pewnej górnej granicy ε_i, dodatkowo musi być spełniony warunek

$$\varepsilon_i > f_i(x^h)$$

[2] Przy optymalizacji nieliniowej można tylko szacować przybliżone wartości wektora nadir za pomocą macierzy wypłat.

e) $I^{\diamond}$ – indeksy w tym zbiorze dotyczą tych kryteriów, które mogą tymczasowo zmieniać się w dowolnym kierunku, nie nakłada się na nie żadnych ograniczeń.

Lokalna klasyfikacja jest dopuszczalna tylko wtedy, gdy spełnione są warunki [5]

$$I^{<} \cup I^{\leq} \cup I^{=} \cup I^{\geq} \cup I^{\diamond} = \{1,2,...,k\}$$
$$I^{<} \cap I^{\leq} \cap I^{=} \cap I^{\geq} \cap I^{\diamond} = \phi$$
$$I^{<} \cup I^{\leq} \neq \phi$$
$$I^{\geq} \cup I^{\diamond} \neq \phi$$

Jeżeli jednocześnie minimalizujemy *k* funkcji kryterium, to niemożliwe jest zaklasyfikowanie w tej samej iteracji wszystkich kryteriów do zbiorów zmniejszających ich wartości ($I^{<}, I^{\leq}$) w następnej iteracji. W zależności od zbioru, do jakiego zaklasyfikowano funkcje kryterium, wyznacza się odmienne punkty referencyjne $\bar{z}_i$ stosowane w konstrukcji zadań zastępczych

$$\bar{z}_i = z_i^*, i \in I^{<}$$
$$\bar{z}_i = \hat{z}_i, i \in I^{\leq}$$
$$\bar{z}_i = f_i(x^h), i \in I^{=}$$
$$\bar{z}_i = \varepsilon_i, i \in I^{\geq}$$
$$\bar{z}_i = z_i^{nad}, i \in I^{\diamond}$$

gdzie:

$\bar{z}_i$ – i-ty element wektora referencyjnego,

$f_i(x^h)$ – wartość i-tego kryterium w h-tej iteracji,

$\hat{z}_i$ – poziom aspiracji dla i-tego kryterium,

ε_i – górna granica dla i-tego kryterium.

Każde zadanie zastępcze w synchronicznej wersji metody NIMBUS wyjściu daje dokładnie jedno rozwiązanie niezdominowane (o ile jest ono możliwe do uzyskania). Zadania zastępcze wykorzystują wektor referencyjny[3] i funkcję osiągnięcia przybliżającą porządek [3] daną wzorem (3)

$$s_z(z) = \underset{i-1,...,k}{MAX}\left\{\left[w_i\left(f_i(x) - \bar{z}_i\right)\right] + \rho\sum_{i=1}^{k} w_i\left(f_i(x) - \bar{z}_i\right)\right\} \tag{3}$$

[3] Wektor referencyjny składa się z wartości poszczególnych kryteriów uznanych przez decydenta za pożądane w danej iteracji synchronicznej wersji metody NIMBUS.

gdzie:

w_i – waga dla i-tego elementu wektora kryterium,

ρ – relatywnie mały skalar (Augmentation term).

Twórcy synchronicznej metody NIMBUS zbudowali zadania zastępcze na bazie już istniejących metod wielokryterialnych[4]. Zadania zastępcze zbudowano tak, by pomimo tych samych preferencji dotyczących wejściowego problemu generowały odmienne rozwiązania niezdominowane. Każde zadanie zastępcze traktuje informacje o preferencjach ujawnione w lokalnej klasyfikacji kryteriów w odmienny sposób. Takie podejście daje możliwość szerszego spojrzenia na problem.

Pierwsze zadanie zastępcze ma postać

$$MIN\left\{\underset{\substack{i\in I^{<}\\ j\in I^{\leq}}}{MAX}\left[\frac{f_i(x)-z_i^{*}}{z_i^{nad}-z_i^{**}};\frac{f_j(x)-\hat{z}_j}{z_j^{nad}-z_j^{**}}\right]+\rho\sum_{i=1}^{k}\frac{f_i(x)}{z_i^{nad}-z_i^{**}}\right\}$$

$$\begin{cases} f_i(x)\leq f_i(x^h), dla \quad i\in I^{<}\cup I^{\leq}\cup I^{=} \\ f_i(x)\leq \varepsilon_i^h, dla \quad i\in I^{\geq} \\ x\in S \end{cases} \tag{4}$$

W tym zadaniu zastępczym w równaniu funkcji celu wykorzystuje się tylko kryteria zadania wejściowego zaklasyfikowane do zbiorów $I^{<}$ i $I^{\leq}$. Pozostałe kryteria są przenoszone do ograniczeń. Ograniczenia dla kryteriów nie uwzględnionych w funkcji kryterium zadania zastępczego wpływają na zwiększanie stopnia przestrzegania postanowień lokalnej klasyfikacji.

Zadanie zastępcze bazujące na metodzie STOM [5] (Satisficing Trade--Off Method) gdzie wymagane jest spełnienia warunku

$$\bar{z}_i > z_i^{**}$$

ma postać

$$MIN\left\{\underset{i=1,2,\ldots,k}{MAX}\left[\frac{f_i(x)-z_i^{**}}{\bar{z}_i-z_i^{**}}\right]+\rho\sum_{i=1}^{k}\frac{f_i(x)}{\bar{z}_i-z_i^{**}}\right\} \tag{5}$$

$$x\in S$$

[11] Są to metody STOM, RPM I GUESS.

Trzecie zadanie zastępcze wywodzi się z metody RPM (Reference Point Method) [5] i można zapisać je następująco

$$MIN\left\{\underset{i=1,2,\ldots,k}{MAX}\left[\frac{f_i(x)-\bar{z}_i}{z_i^{nad}-z_i^{**}}\right]+\rho\sum_{i=1}^{k}\frac{f_i(x)}{z_i^{nad}-z_i^{**}}\right\} \tag{6}$$

$$x \in S$$

Zadania zastępcze (5) i (6) różnią się tylko współczynnikami wagowymi. Gdyby do zadania (6) zastosować wagi $w_i = \frac{1}{\bar{z}_i - z_i^{**}}$, wtedy uzyskane wyniki będą takie same jak w przypadku zadania (5).

Ostatnie z dostępnych zadań zastępczych opisane wzorem (7) ma korzenie w metodzie GUESS [5]. Ponieważ zasada działania polega tu na minimalizacji odległości pomiędzy kryteriami i wektorem nadir, więc przypuszczalnie właśnie to zadanie daje najgorsze wyniki

$$MIN\left\{\underset{i\notin I^{<>}}{MAX}\left[\frac{f_i(x)-z_i^{nad}}{z_i^{nad}-\bar{z}_i}\right]+\rho\sum_{i=1}^{k}\frac{f_i(x)}{z_i^{nad}-\bar{z}_i}\right\} \tag{7}$$

$$x \in S$$

Jeżeli kryterium zaklasyfikuje się do $I^{<>}$, to zmienia się jego wagi na $w_i = \frac{1}{z_i^{nad} - z_i^{**}}$. Unika się wtedy dzielenia przez zero w niektórych przypadkach.

Zadania zastępcze (5) do (7) traktują informacje zawarte w lokalnej klasyfikacji kryteriów mniej rygorystycznie. Informacje z klasyfikacji nie są w tych zadaniach przenoszone na konkretne ograniczenia, ale na ich podstawie formułowany jest pewien punkt referencyjny. Względem niego znajdowane są kolejne rozwiązania, czasem lepsze od spodziewanych, pomimo tego, że nie spełniają one dokładnie założeń wcześniejszej klasyfikacji.

W synchronicznej wersji metody NIMBUS wektor startowy z^{start} (nazywany neutralnym kompromisem) oblicza się rozwiązując zadanie zastępcze(6), przy czym wektor referencyjny $\bar{z}$ wyznacza się następująco [5]

$$\bar{z} = \frac{z^{nad} + z^{**}}{2} \tag{8}$$

System komputerowy WWW-NIMBUS rozszerza możliwości obliczeniowe o rozwiązania pośrednie. Jeżeli decydent nie będzie zadowolony z otrzymanych rozwiązań w h-tej iteracji, to można wygenerować kilka rozwiązań pośrednich pomiędzy dwoma wybranymi rozwiązaniami niezdominowanymi.

Postępowanie w tym przypadku jest następujące. Odległość pomiędzy punktami wybranymi do wyznaczania rozwiązań pośrednich w przestrzeni decyzyjnej dzieli się na wymaganą ilość[5] równych odcinków. Punkty w przestrzeni kryterialnej odpowiadające odcinkom stają się jednocześnie punktami referencyjnymi dla zadania zastępczego (6). Ostatecznie należy rozwiązać ciąg zadań zastępczych (6) uwzględniając te punkty referencyjne [4].

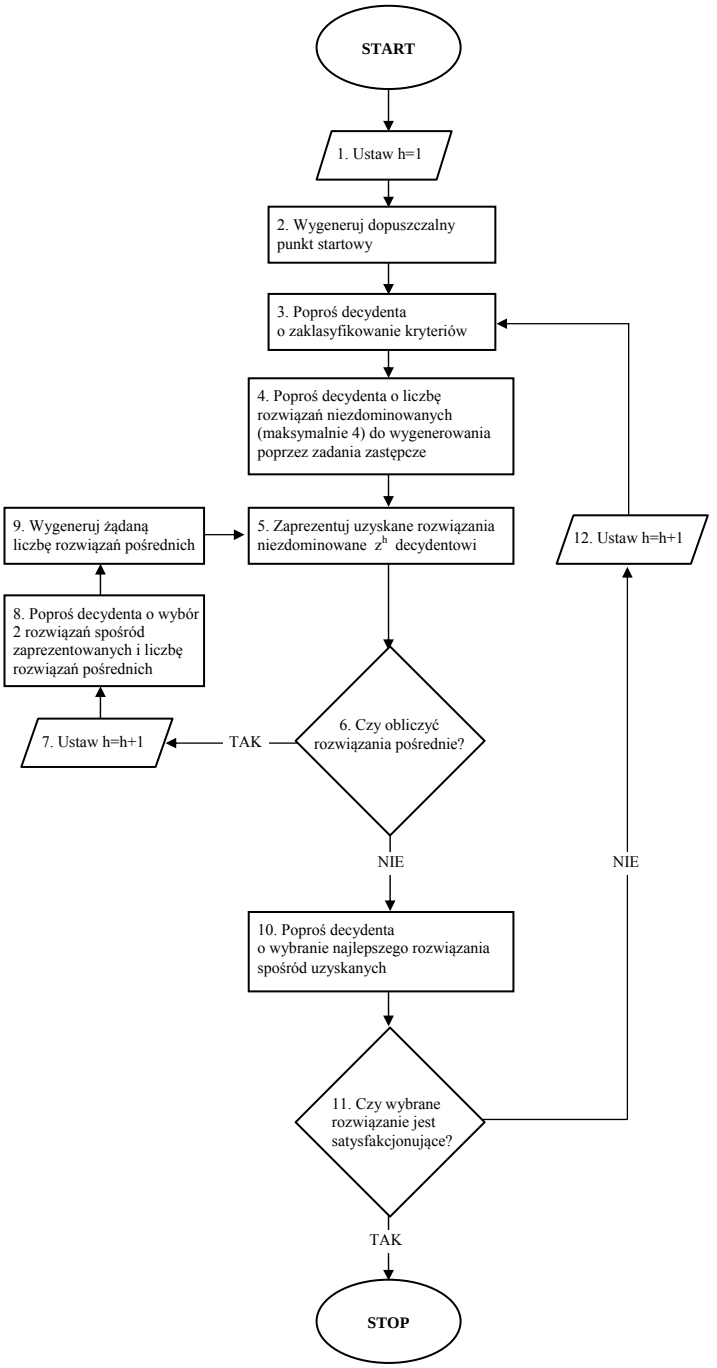
W zadaniach wielokryterialnego programowania liniowego wystarczy podzielić odległość między punktami wybranymi do wyznaczenia rozwiązań pośrednich na odcinki równej długości i obliczyć wartości kryteriów dla wyodrębnionych krańców odcinków.

Rozważania dotyczące synchronicznej wersji metody NIMBUS można przedstawić algorytmicznie (algorytm 1) oraz schematem blokowym zaprezentowanym na rys. 1.

Algorytm 1

1. Ustaw $h = 1$.
2. Wygeneruj dopuszczalny punkt startowy.
3. Poproś decydenta o zaklasyfikowanie kryteriów.
4. Poproś decydenta o liczbę rozwiązań niezdominowanych z^h (maksymalnie 4) do wygenerowania poprzez zadania zastępcze.
5. Zaprezentuj uzyskane rozwiązania niezdominowane z^h decydentowi.
6. Czy obliczyć rozwiązania pośrednie? TAK – idź do punktu 7. NIE – idź do punktu 10.
7. Ustaw $h = h + 1$.
8. Poproś decydenta o wybór 2 rozwiązań spośród zaprezentowanych i liczbę rozwiązań pośrednich.
9. Wygeneruj żądaną liczbę rozwiązań pośrednich, idź do punktu 5.
10. Poproś decydenta o wybranie najlepszego rozwiązania spośród uzyskanych.
11. Czy wybrane rozwiązanie jest satysfakcjonujące? TAK – STOP. NIE – idź do punktu 12.
12. Ustaw $h = h + 1$, idź do punktu 3.

[5] Aby wyznaczyć np. 11 rozwiązań pośrednich, trzeba dokonać podziału na 12 odcinków w przestrzeni decyzyjnej.



Rys. 1. Schemat blokowy synchronicznej metody NIMBUS

2. Zastosowanie metody NIMBUS do ustalenia optymalnej struktury portfela akcji

Korzystając z rozwiązania, zaproponowanego przez C. Dominiaka [1], w którym do selekcji spółek do portfela wykorzystano zmodyfikowana metodę BIPOLAR [2], alokacje aktywów w portfelu przeprowadzono w oparciu o metodę NIMBUS, wykorzystując dane przedstawione w tabeli 1.

Oczekiwana stopa zwrotu portfela rynkowego wynosi 0,3%, natomiast odchylenie standardowe stopy zwrotu portfela rynkowego jest na poziomie 5,7%[6].

Tabela 1

Dane akcji do ustalenia struktury portfela

Nazwa spółki na GPW	Zmienna w modelu	α	β	Var(e_i)	P/E	P/BV	ROE
Elza	x_1	–0,01	1,06	0,01	4,24	0,28	6,7
Jelfa	x_2	0	0,64	0	12,31	0,94	7,6
Grajewo	x_3	0,04	0,71	0,01	5,27	1,1	20,85
Elektrobudowa	x_4	–0,02	0,69	0,01	9,99	0,92	9,17
Elektroexp	x_5	–0,03	1,18	0,01	21,75	1,99	9,17
Pkn Orlen	x_6	0	0,84	0	15,76	0,95	6,05
Wawel	x_7	0,02	0,55	0	11,05	0,77	6,94
Forte	x_8	0,01	0,68	0,01	9,29	0,61	6,54
Stomil	x_9	0,01	0,16	0	9,82	1,55	15,83
Krosno	x_{10}	0,02	0,33	0	12,63	0,85	6,76
Irena	x_{11}	0,02	0,18	0,01	19,22	1,04	5,41
Rafako	x_{12}	0	0,74	0,01	8,8	0,54	6,12
Atlantis	x_{13}	–0,02	1,06	0,03	14,91	0,66	4,41
Mostalwar	x_{14}	–0,01	0,46	0,01	7,99	0,42	5,24
Oława	x_{15}	0,02	1	0,01	1,23	0,42	34,24

Źródło: [1].

Do obliczeń wykorzystano model decyzyjny zaproponowany przez C. Dominiaka [1]. Składa się on z pięciu kryteriów i trzech ograniczeń. Kryteria maksymalizowane, to stopa zwrotu obliczona na podstawie modelu Sharpe'a i wskaźnik ROE, natomiast kryteria minimalizowane, to ryzyko obliczone z modelu Sharpe'a i wskaźniki P/E oraz P/BV. Dwa pierwsze ograniczenia są związane z maksymalnymi udziałami poszczególnych akcji w portfelu – ograniczono łączny udział spółek ELZAB, ELEKTROBUDOWA,

[6] Dane te zostały obliczone w oparciu o miesięczne notowania indeksu giełdowego WIG.

ELEKTROEXP, MOSTALWAR do poziomu 40%, a łączny udział spółek KROSNO oraz IRENA do poziomu 30%. Trzecie ograniczenie gwarantuje, że udziały akcji w otrzymanym portfelu będą sumowały się do jedynki.

Rozwiązywane zadanie przedstawiono niżej

$$MAX\left\{\sum_{i=1}^{15}(\alpha_i x_i)+0{,}003\sum_{i=1}^{15}\left(\beta_i x_i^2\right)\right\}$$

$$MIN\left\{0{,}057^2\sum_{i=1}^{15}\left(\beta_i^2 x_i^2\right)+\sum_{i=1}^{15}\left(Var(e_i)x_i^2\right)\right\}$$

$$MIN\left\{\sum_{i=1}^{15}(P/E)_i x_i\right\}$$

$$MIN\left\{\sum_{i=1}^{15}(P/BV)_i x_i\right\}$$

$$MAX\left\{\sum_{i=1}^{15}ROE_i x_i\right\}$$

$$\begin{cases}\sum\limits_{i\in\{1,4,5,14\}} x_i \le 0{,}4\\ \sum\limits_{i\in\{10,11\}} x_i \le 0{,}3\\ \sum\limits_{i=1}^{15} x_i = 1\\ x_i \ge 0, i=1,...,15\end{cases}$$

Do obliczeń wykorzystano system komputerowy WWW-NIMBUS. Obecność ograniczenia równościowego w modelu uniemożliwiła zastosowanie optymalizatorów globalnych. System obliczył wartości wektora idealnego i nadir, które przedstawiono w tabeli 2.

Tabela 2

Zakresy zmienności dla kryteriów

Ideal Criterion Vector (estim.)	Nadir (estim.)
1.1072E-2	4.213E-2
5.941757E-5	1.3249E-2
1.23	10.88569
0.364	1.241458
0.1209635	0.342

Źródło: Obliczenia wykonane za pomocą WWW-NIMBUS.

W iteracji 1 na podstawie informacji w tabeli 2 decydent postanowił zwiększyć wartość stopy zwrotu do poziomu 3,5%, wskaźnika rentowności kapitałów do 27%. Wariancja stopy zwrotu powinna być zmniejszona do poziomu 0,05%, wskaźnik P/BV do poziomu 0,6. Decydent zgodził się na zwiększenie wskaźnika ceny do zysku do poziomu 5. Interpretacją preferencji decydenta są następujące zbiory klasyfikacyjne

$$I^{\geq} = \{1, 3, 5\}, \hat{z}_1 = 3{,}5\%, \varepsilon_3 = 5, \hat{z}_5 = 27\%$$
$$I^{\leq} = \{2, 4\}, \hat{z}_2 = 0{,}05\%, \hat{z}_4 = 0{,}6$$
$$I^{=} \cap I^{>} \cap I^{<} = 0$$

System znalazł trzy rozwiązania. Zaprezentowano je w tabeli 3.

Tabela 3

Warianty decyzyjne w iteracji 1

Wariant startowy	Stopa zwrotu	3.00198E-2	Wariant decyzyjny 3	Stopa zwrotu	1.233596E-2
	Ryzyko	4.862286E-3		Ryzyko	1.711662E-3
	P/E	4.051085		P/E	5.82458
	P/BV	0.722248		P/BV	0.6750111
	ROE	0.2557884		ROE	0.1645241
Wariant decyzyjny 2	Stopa zwrotu	1.266656E-2	Wariant decyzyjny 4	Stopa zwrotu	-7.191924E-3
	Ryzyko	1.172191E-3		Ryzyko	4.911739E-3
	P/E	6.75959		P/E	9.883164
	P/BV	0.7105865		P/BV	0.5196366
	ROE	0.1470173		ROE	5.348481E-2

Źródło: Ibid.

Uzyskane rozwiązania są gorsze od wektora z^{start} (wariantu decyzyjnego 1 w tabeli 3), nie udało się spełnić wszystkich żądań decydenta. Ryzyko portfela zmniejszyło się, ale stopa zwrotu też spadła o około połowę, decydent nie jest w stanie zaakceptować takiego rozwiązania.

W iteracji 2 decydent poprosił o wygenerowanie dziesięciu rozwiązań pośrednich pomiędzy wariantem decyzyjnym 1 (jak do tej pory najlepszym) i wariantem decyzyjnym 2 z tabeli 3. Początkowo stopa zwrotu relatywnie spadła w tempie mniejszym niż ryzyko portfela, ta tendencja utrzymywała się do wariantu decyzyjnego 3, potem spadek wartości stopy zwrotu był znaczny. Pozostałe trzy kryteria (P/E, P/BV i ROE) zmieniały się na poziomie możliwym do zaakceptowania. Decydent postanowił, że nowe rozwiązania w iteracji 3 będą tworzone na bazie wariantu decyzyjnego 3 o wartościach przedstawionych w tabeli 4.

Tabela 4

Wybrany wariant decyzyjny w iteracji 2

Wariant startowy z iteracji 1	Stopa zwrotu	3.00198E-2
	Ryzyko	4.862286E-3
	P/E	4.051085
	P/BV	0.722248
	ROE	0.2557884
Wariant decyzyjny 3	Stopa zwrotu	2.695068E-2
	Ryzyko	3.798905E-3
	P/E	4.527028
	P/BV	0.7186255
	ROE	0.2365069
Wariant decyzyjny 2 z iteracji 1	Stopa zwrotu	1.266656E-2
	Ryzyko	1.172191E-3
	P/E	6.75959
	P/BV	0.7105865
	ROE	0.1470173

Źródło: Ibid.

W iteracji 3 dalej poprawiano wartości pierwszego, drugiego i trzeciego kryterium kosztem wartości wskaźników ROE i P/BV. Interpretacją preferencji decydenta są następujące zbiory klasyfikacyjne

$$I^{\geq} = \{1, 4\},\ \hat{z}_1 = 2{,}87429 * 10^{-2},\ \varepsilon_4 = 0{,}7815263$$

$$I^{\leq} = \{2, 3, 5\},\ \hat{z}_2 = 3{,}3887 * 10^{-3},\ \hat{z}_3 = 4{,}239703,\ \varepsilon_5 = 22{,}31596$$

$$I^{=} \cap I^{>} \cap I^{<} = \phi$$

Udało się uzyskać cztery nowe warianty decyzyjne. Jedynym potencjalnym rozwiązaniem do kontynuacji obliczeń okazał się wariant decyzyjny 2. Stopa zwrotu, ryzyko, wskaźnik P/E oraz P/BV były lepsze w porównaniu z wariantem startowym (wariantem decyzyjnym 2 w tabeli 5).

Tabela 5

Wybrany wariant decyzyjny w iteracji 3

Wariant decyzyjny 2	Stopa zwrotu	2.69625E-2	Wariant decyzyjny 3 z iteracji 2	Stopa zwrotu	2.695068E-2
	Ryzyko	3.70753E-3		Ryzyko	3.798905E-3
	P/E	4.456629		P/E	4.527028
	P/BV	0.7129601		P/BV	0.7186255
	ROE	0.2245074		ROE	0.2365069

Źródło: Ibid.

W iteracji 4 obliczono 10 rozwiązań pośrednich na podstawie wariantów decyzyjnych 2 i 3 z tabeli 5. We wszystkich otrzymanych w iteracji 4 rozwiązaniach pośrednich kierunek zmian stopy zwrotu, ryzyka, P/E i P/BV był pozytywny. Odwrotnie do oczekiwań decydenta zmieniały się wartości kryterium ROE. Decydent przyjął, że może zaakceptować portfel, jeżeli wartość kryterium ROE nie będzie niższa niż 23%. Tym wymaganiom odpowiadał wariant decyzyjny 6. Na tym etapie zakończono poszukiwania rozwiązania satysfakcjonującego. Wariant decyzyjny 6 to portfel zawierający akcje sześciu spółek.

3. Porównanie rozwiązania otrzymanego metodą NIMBUS z rozwiązaniem otrzymanym interaktyną metodą satysfakcjonującego poziomu kryteriów

Udziały w portfelach akcji, a także parametry portfela otrzymanego metodą NIMBUS oraz portfela, uzyskanego przez C. Dominiaka dla tych samych danych interaktywną metodą satysfakcjonującego poziomu kryteriów (IMSPK) przedstawiono w tabelach 6 oraz 7.

Tabela 6

Udziały w portfelach akcji

Nazwa spółki	Udział	
Metoda	NIMBUS	IMSPK
ELZAB	6,82%	0%
JELFA	0%	0%
GRAJEWO	34,41%	28,32%
ELEKTROBUDOWA	0%	0%
ELEKTROEXP	0%	0%
PKN ORLEN	0%	0%
WAWEL	14,1%	20,08%
FORTE	1,751%	0%
STOMIL	1,68%	5,73%
KROSNO	0%	22,97%
IRENA	0%	8,57%
RAFAKO	0%	0%
ATLANTIS	0%	0%
MOSTALWAR	0%	0%
OŁAWA	41,24%	14,33%

Źródło: Opracowanie własne. Udziały uzyskane interaktywną metodą satysfakcjonujących poziomów kryteriów zaczerpnięto z pracy [1, s. 126].

Składy portfeli uzyskanych dwoma metodami znacznie się między sobą różnią. Dotyczy to zarówno proporcji, jak i samych akcji wchodzących do portfela. W obu przypadkach otrzymano portfel sześcioelementowy.

W skład portfela uzyskanego metodą NIMBUS wchodzą spółki ELZAB z udziałem 6,820%, GRAJEWO z drugim co do wielkości udziałem wynoszącym 34,41%, WAWEL z udziałem 14,1%, FORTE i STOMIL posiadają najmniejsze udziały wynoszące odpowiednio 1,751% i 1,68%, największy udział w portfelu w wysokości 41,24% przypada na akcje OŁAWY.

Największe udziały (odpowiednio 28,320% i 22,970%) w portfelu otrzymanym metodą SPK posiadają akcje spółki GRAJEWO i KROSNO, akcje spółki WAWEL mają 20,080% udziału w portfelu, akcje spółki OŁAWA powinny stanowić 14,330% portfela. Najmniejszy udział wynoszący odpowiednio 5,730% i 8,570% przypadł akcjom STOMIL i IRENA.

Tabela 7

Parametry portfeli akcji

Kryterium / Metoda	Stopa zwrotu	Ryzyko	P/E	P/BV	ROE
NIMBUS	2,70%	0,37%	4,50	0,72	23,11%
SPK	2,74%	5,00%	9,00	0,90	15,13%

Źródło: Opracowanie własne. Parametry portfela uzyskanego interaktywną metodą satysfakcjonujących poziomów kryteriów zaczerpnięto z pracy [1, s. 125].

Parametry uzyskanych portfeli także znacznie różnią się pomiędzy sobą. Portfel uzyskany metodą NIMBUS charakteryzuje się ogólnie rzecz biorąc lepszymi parametrami. Jego stopa zwrotu jest nieco niższa i wynosi 2,7%, ale kryteria takie jak ryzyko czy wskaźnik rentowności kapitałów własnych ROE mają o wiele bardziej pożądane wartości. Ryzyko jest bardzo małe, wynosi zaledwie 0,37% w porównaniu z 5% w portfelu z metody SPK. Wartość wskaźnika ROE dla portfela otrzymanego przy zastosowaniu metody NIMBUS jest wyższa o około 8%. Pozostałe kryteria to wskaźnik P/E niższy dla metody NIMBUS o połowę (4,5 w stosunku do 9) oraz wskaźnik P/BV niższy o prawie 0,2.

Wnioski

Synchroniczna wersja metody NIMBUS wraz z systemem komputerowym WWW-NIMBUS stanowią bardzo interesujące i nowatorskie połączenie, które ogranicza często nieskończony zbiór rozwiązań niezdominowanych do co najmniej czterech rozwiązań zadań zastępczych. Zadanie nimbusowe daje najbardziej prawidłowe wyniki z punktu widzenia postanowień klasyfikacji, zadanie GUESS poprzez minimalizację odległości kryteriów od wektora nadir odzwierciedla najgorszy scenariusz zdarzeń. Zadania zastępcze STOM i RPM są kompromisem pomiędzy dwoma wcześniejszymi skrajnościami.

Metoda NIMBUS jest dobrą alternatywą dla rozwiązywania zadań portfelowych. Składowymi portfela nie muszą być tylko akcje, można również optymalizować portfele ubezpieczeniowe lub portfele funduszy inwestycyjnych. Możliwa jest szersza analiza portfeli pod kątem ryzyka mierzonego odchyleniem przeciętnym lub inną miarą nieróżniczkowalną. Ograniczenie stanowi tutaj system komputerowy, jego optymalizatory globalne nie radzą sobie z ograniczeniami w modelach i powstaje ryzyko uzyskiwania rozwiązań optymalnych w sensie lokalnym.

Rozwiązania pośrednie oferowane przez system bardzo ułatwiają poszukiwania rozwiązania satysfakcjonującego, w toku obliczeń przedstawia się decydentowi relatywnie dużą ilość propozycji rozwiązań przy małej liczbie iteracji w porównaniu z innymi metodami wielokryterialnymi.

Literatura

1. Dominiak C. (2006). Interaktywna metoda satysfakcjonujących poziomów kryteriów. W: Metody wielokryterialne na polskim rynku finansowym. Red. T. Trzaskalik. PWE, Warszawa, 119-126.
2. Konarzewska-Gubała E. (1991). Wspomaganie decyzji wielokryterialnych: system Bipolar. AE, Wrocław.
3. Miettinen K. (1999). Nonlinear Multiobjective Optimization. Kluwer Academic Publishers, Boston.
4. Miettinen K. (2006). IND-NIMBUS for Demanding Interactive Multiobjective Optimization. W: Multiple Criteria Decision Making'05. Red. T. Trzaskalik. AE, Katowice.
5. Miettinen K., Mäkelä M.M. (2006). Synchronous Approach in Interactive Multiobjective Optimization. European Journal of Operational Research, 170(3).
6. Miettinen K., Mäkelä M.M. (1996). Comparing Two Versions of NIMBUS Optimization System. Report 23/1996. University of Jyväskylä, Department of Mathematics, Laboratory of Scientific Computing.
7. Miettinen K., Mäkelä M.M. (1998). Optimization System WWW-NIMBUS. Report 9/1998. University of Jyväskylä. Department of Mathematics, Laboratory of Scientific Computing, Jyväskylä.
8. Steuer R.E., Na P. (2003). Multiple Criteria Decision Making Combined with Finance: A Categorized Bibliographic Study. European Journal of Operational Research, 150, 496-515.

9. Trzaskalik T. (2003). Wprowadzenie do badań operacyjnych z komputerem. PWE, Warszawa.
10. http://users.jyu.fi/~miettine/engl.html, (7.03.2009).
11. http://nimbus.mit.jyu.fi, (7.03.2009).

Elżbieta Majewska
Uniwersytet w Białymstoku

OCENA RYZYKA FUNDUSZY INWESTYCYJNYCH Z WYKORZYSTANIEM WSPÓŁCZYNNIKA GINIEGO

Wprowadzenie

Wybór funduszu inwestycyjnego jest jednym z podstawowych prob-0lemów, jaki muszą rozwiązać inwestorzy decydujący się na ulokowanie swoich kapitałów w tego typu inwestycje. Kryteria brane przy tym pod uwagę mogą być bardzo różne, jednak wśród wybieranych najczęściej są podstawowe charakterystyki, czyli potencjalny zysk oraz ryzyko jego osiągnięcia. Ten ostatni parametr oceniany jest zwykle na podstawie odchylenia standardowego stopy zwrotu lub współczynnik beta. Są to wielkości powszechnie stosowane w analizie ryzyka inwestycyjnego. W pracy przedstawimy propozycję zastosowania do pomiaru ryzyka funduszy średniej różnicy Giniego oraz uogólnionego współczynnika Giniego. Wielkości te proponowane są bowiem w literaturze jako alternatywne miary ryzyka stosowane w analizie portfelowej [1; 8; 9; 10]. Zaprezentujemy rezultaty wykorzystania tych wskaźników do oceny ryzyka 15 wybranych OFI akcji. Uzyskane wyniki porównamy z ryzykiem mierzonym odchyleniem standardowym stopy zwrotu.

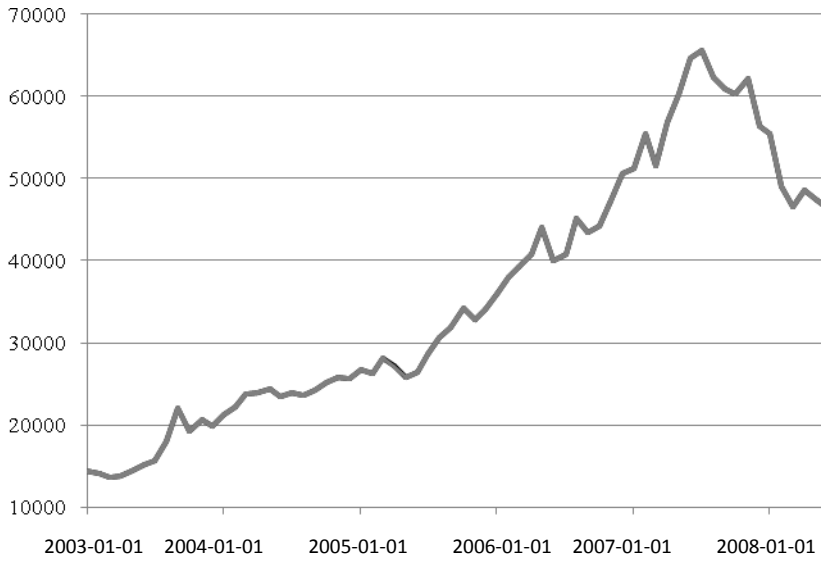
1. Ocena ryzyka funduszy inwestycyjnych z wykorzystaniem odchylenia standardowego

Jedną ze wspomnianych klasycznych miar stosowanych w ocenie ryzyka funduszy inwestycyjnych jest odchylenie standardowe stopy zwrotu (σ). Określa ono przeciętne odchylenie zrealizowanych stóp zwrotu od ich wartości oczekiwanej. W przypadku, gdy analiza prowadzona jest w oparciu o dane historyczne (szeregi czasowe stóp zwrotu) pochodzące z n okresów, wartość σ wyznacza się jako [5; s. 47]

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2} \tag{1}$$

gdzie r_t oznacza zrealizowaną stopę zwrotu z okresu *t*, natomiast $\bar{r}$ to oczekiwana stopa zwrotu (rozumiana jako średnia arytmetyczna zrealizowanych stóp zwrotu). Im większe σ, tym większego odchylenia stóp zwrotu od ich wartości oczekiwanej może spodziewać się inwestor.

Odchylenie standardowe wykorzystano do oceny ryzyka 15 otwartych funduszy inwestycyjnych akcji. Ponieważ lokowanie kapitału w OFI ma zwykle charakter długoterminowy, analizę przeprowadzono na podstawie danych z okresu od 1.01.2003 do 31.12.2008 w dwóch wariantach stóp zwrotu: dziennych oraz miesięcznych. Przyjęty w analizach 6-letni przedział czasowy obejmował zarówno okresy wzrostów, jak i znacznych spadków na rynku (rys. 1). Skutkowało to niskimi oczekiwanymi stopami zwrotu i znacznie je przekraczającymi odchyleniami standardowymi poszczególnych funduszy (tabela 1).



Rys. 1. Notowania indeksu WIG z okresu 1.01.2003-31.12.2008

Wielkości wyznaczonych charakterystyk uzyskane na podstawie miesięcznych stóp zwrotu (wariant II) są przy tym znacznie większe niż w przypadku stóp zwrotu dziennych (wariant I). Jednak łatwo zauważyć, że najlepsze wyniki (najwyższe oczekiwane stopy zwrotu) w obu wariantach osiągnął

w badanym okresie fundusz Arka, natomiast najsłabiej wypadły Pioneer oraz PKO/CREDIT. Analiza ryzyka nie pozwala już tak jednoznacznie wskazać inwestycji najbardziej i najmniej ryzykownych, ponieważ ekstremalne wartości σ (wyróżnione w tabeli 1) w każdym z wariantów stóp zwrotu odpowiadają innym funduszom. Oznacza to, że te same fundusze mogą być inaczej ocenione w horyzoncie miesięcznym niż w skali jednego dnia i odwrotnie.

Tabela 1

Oczekiwana stopa zwrotu i ryzyko wybranych funduszy inwestycyjnych w próbie 6-letniej (I – dzienne stopy zwrotu, II – miesięczne stopy zwrotu)

Fundusz inwestycyjny	$\bar{r}$ [%]		σ [%]	
	I	II	I	II
Arka BZ WBK Akcji FIO	**0,0619**	**1,4418**	1,2731	6,9212
BPH FIO Akcji	0,0398	0,8942	1,1303	5,7566
CU FIO Polskich Akcji	0,0532	1,2764	1,2020	6,4661
DWS Polska FIO Akcji	0,0300	0,6625	**1,3341**	5,8772
DWS Polska FIO Akcji Plus	0,0362	0,8139	1,2425	5,9171
DWS Polska FIO Top 25 Małych Spółek	0,0260	0,6246	1,1662	**7,0539**
ING FIO Akcji	0,0381	0,8861	1,1894	6,0791
Legg Mason Akcji FIO	0,0542	1,2253	1,1376	6,0127
Millennium FIO Akcji	0,0281	0,6621	**1,1047**	5,7858
Pioneer Akcji Polskich FIO	**0,0216**	0,5745	1,2925	6,7429
PKO/CREDIT SUISSE Akcji FIO	0,0223	**0,5633**	1,2340	6,2100
PZU FIO Akcji KRAKOWIAK	0,0302	0,7127	1,1124	5,7142
SEB 3 – Akcji FIO	0,0323	0,7545	1,2099	5,9743
Skarbiec – Akcja FIO	0,0556	1,2025	1,1303	**5,5609**
UniKorona Akcja FIO	0,0523	1,1835	1,2291	5,8681

Inwestor, który chce wybrać fundusze preferowane przez siebie powinien więc najpierw określić przewidywany czas trwania jego inwestycji po to, by niezbędne analizy przeprowadzić w odpowiednim wariancie stóp zwrotu. Potwierdzają to również różne pozycje funduszy w rankingach sporządzonych na podstawie wartości odchylenia standardowego w wariantach dziennych oraz miesięcznych stóp zwrotu (tabela 2). Zależność między pozycjami OFI mierzona współczynnikiem korelacji rang τ Kendalla[1] [6, s. 191-192] wynosi zaledwie 0,3143, co świadczy o niewielkiej zgodności uporządkowań.

[1] Współczynnik ten ocenia podobieństwo uporządkowań, przy czym wartość 1 oznacza pełną ich zgodność, 0 brak zgodności, natomiast -1 całkowitą przeciwstawność.

Tabela 2

Rankingi wybranych OFI (I – dzienne stopy zwrotu; II – miesięczne stopy zwrotu)

Fundusz inwestycyjny	Kryterium σ	
	I	II
Arka BZ WBK Akcji FIO	13	14
BPH FIO Akcji	3	3
CU FIO Polskich Akcji	8	12
DWS Polska FIO Akcji	**15**	6
DWS Polska FIO Akcji Plus	12	7
DWS Polska FIO Top 25 Małych Spółek	6	**15**
ING FIO Akcji	7	10
Legg Mason Akcji FIO	5	9
Millennium FIO Akcji	**1**	4
Pioneer Akcji Polskich FIO	14	13
PKO/CREDIT SUISSE Akcji FIO	11	11
PZU FIO Akcji KRAKOWIAK	2	2
SEB 3 – Akcji FIO	9	8
Skarbiec – Akcja FIO	4	**1**
UniKorona Akcja FIO	10	5

W dalszej kolejności sprawdzimy, czy analiza ryzyka funduszy szacowanego innymi, mniej rozpowszechnionymi miarami, dostarczy podobnych wniosków.

2. Średnia różnica Giniego jako alternatywna miara ryzyka funduszy inwestycyjnych

Współczynnik Giniego jest stosowaną od 1912 roku miarą koncentracji wykorzystywaną najczęściej do analizy nierównomierności rozkładu dochodów. Jego wartość opiera się na średniej różnicy Giniego, którą w ogólnym przypadku wyznacza się jako wartość oczekiwaną bezwzględnej różnicy między dwiema niezależnymi zmiennymi losowymi o jednakowym rozkładzie, czyli [3]

$$\Delta = E \mid X_1 - X_2 \mid \tag{2}$$

W 1982 roku Yitzhaki zaproponował wykorzystanie wielkości $\Gamma = \frac{1}{2}\Delta$ w analizie portfelowej. Pokazał [11], że konstrukcja portfela inwestycyjnego oparta na kryterium średniej arytmetycznej i Γ (metoda *MG*) łączy zalety dwóch innych znanych metod: dominacji stochastycznej (*SD*) oraz metody Markowitza opartej na średniej i wariancji (*MV*). Metoda *MG* okazuje się tak

prosta do zastosowania, jak *MV* oraz tak ogólna, jak *SD*. Zastąpienie wariancji (lub odchylenia standardowego) średnią różnicą Giniego powoduje, że spełnione są kryteria stochastycznej dominacji rzędu pierwszego oraz drugiego, co oznacza, że budowane portfele są efektywne w ogólniejszym sensie.

Stosując średnią różnicę Giniego do oceny ryzyka inwestycyjnego wykorzystuje się najczęściej formułę postaci [4; 8; 9]

$$\Gamma = 2\,\mathrm{cov}[r, F(r)] \tag{3}$$

gdzie $F(r)$ oznacza dystrybuantę rozkładu stopy zwrotu *r*.

W analizie portfelowej średnia różnica Giniego okazała się dobrą alternatywą dla odchylenia standardowego stóp zwrotu. W związku z tym, naturalne wydaje się pytanie o to, czy za pomocą Γ można również analizować ryzyko funduszy inwestycyjnych. Poszukując na nie odpowiedzi wykorzystano formułę (4) do oszacowania ryzyka badanych wcześniej funduszy inwestycyjnych[2]. Uzyskane wyniki wraz ze zbudowanym na ich podstawie rankingiem przedstawia tabela 3 (wyróżniono ekstremalne wartości Γ).

Tabela 3

Średnia różnica Giniego i rankingi wybranych OFI uzyskane dla próby 6-letniej (I – dzienne stopy zwrotu; II – miesięczne stopy zwrotu)

Fundusz inwestycyjny	Γ (%)		Pozycja w rankingu	
	I	II	I	II
Arka BZ WBK Akcji FIO	1,2711	3,7864	14	14
BPH FIO Akcji	1,1442	3,1669	4	4
CU FIO Polskich Akcji	1,2206	3,4695	11	12
DWS Polska FIO Akcji	**1,2844**	3,1508	**15**	3
DWS Polska FIO Akcji Plus	1,2198	3,2564	10	8
DWS Polska FIO Top 25 Małych Spółek	1,2032	**3,9400**	9	**15**
ING FIO Akcji	1,1462	3,4375	6	11
Legg Mason Akcji FIO	1,1453	3,2423	5	6
Millennium FIO Akcji	**1,1173**	3,2625	**1**	9
Pioneer Akcji Polskich FIO	1,2492	3,7186	12	13
PKO/CREDIT SUISSE Akcji FIO	1,2496	3,2517	13	7
PZU FIO Akcji KRAKOWIAK	1,1360	**3,0872**	2	**1**
SEB 3 – Akcji FIO	1,2006	3,1892	7	5
Skarbiec – Akcja FIO	1,1418	3,0913	3	2
UniKorona Akcja FIO	1,2016	3,2692	8	10

[2] W obliczeniach za $F(r)$ przyjęto dystrybuantę empiryczną rozkładu stóp zwrotu.

Podobnie jak w przypadku odchylenia standardowego, również Γ przyjmuje większe wartości w kontekście miesięcznych stóp zwrotu niż dziennych. Porównując uzyskane wartości Γ i σ łatwo zauważyć, że w przypadku dziennych stóp zwrotu wielkości te są bardzo zbliżone ($\sigma \in [1{,}1047\%\,;\,1{,}3341\%]$, $\Gamma \in [1{,}1173\%\,;\,1{,}2844\%]$), natomiast w wariancie stóp zwrotu miesięcznych różnią się już znacznie: σ waha się od $5{,}5609\%$ do $7{,}0539\%$, natomiast Γ od $3{,}0972\%$ do $3{,}9400\%$. Również rankingi uzyskane na podstawie tych wskaźników wykazują znaczne podobieństwo w wariancie I, co potwierdza współczynnik τ Kendalla wynoszący 0,8095. Natomiast w wariancie II podobieństwo rankingów jest znacznie mniejsze, współczynnik τ Kendalla osiąga w tym przypadku wartość 0,5048. Wyniki te są więc analogiczne do uzyskanych przy zastosowaniu odchylenia standardowego.

Istotny wydaje się jednak jeszcze jeden aspekt związany z Γ. Otóż opisana zależnością (3) średnia różnica Giniego jest szczególnym przypadkiem tzw. uogólnionego wskaźnika Giniego. Definiuje się go jako [4; 9; 10]

$$\Gamma(\nu) = -\nu \operatorname{cov}\{r, [1 - F(r)]^{\nu-1}\} \tag{4}$$

gdzie $\nu \in [1, \infty)$ jest parametrem uwzględniającym stopień awersji do ryzyka inwestora: im większa wartość ν, tym większa awersja do ryzyka. Parametr $\nu = 1$ charakteryzuje postawę neutralną wobec ryzyka, dla $\nu = 2$ otrzymujemy wzór (3), natomiast $\nu \to \infty$ oznacza inwestora, który nie dopuszcza ryzyka na żadnym poziomie i tym samym dąży do osiągnięcia stosunkowo niskiego, ale całkowicie bezpiecznego zysku.

Uwzględnienie poziomu awersji do ryzyka potencjalnego inwestora jest bardzo istotnym czynnikiem przy wyborze każdej inwestycji, również funduszu inwestycyjnego. Stąd uogólniony współczynnik Giniego może się okazać bardzo przydatnym narzędziem wspomagającym podejmowanie tego typu decyzji.

Tabela 4 prezentuje wyniki uzyskane dla badanych OFI z wykorzystaniem formuły (4) w przypadku, gdy ν równe jest kolejno 3, 4, 5, 10 oraz 20, co odpowiada inwestorom z coraz większą awersją do ryzyka. Zarówno w wariancie dziennych stóp zwrotu, jak i miesięcznych, wartości $\Gamma(\nu)$ dla poszczególnych funduszy rosną wraz ze wzrostem ν, co wydaje się naturalną tendencją, ponieważ ten sam fundusz przez inwestora z większą awersją do ryzyka będzie odbierany jako inwestycja bardziej ryzykowna.

Tabela 4

Uogólniony wskaźnik Giniego (w %) wybranych funduszy inwestycyjnych

Fundusz inwestycyjny	Γ(3)	Γ(4)	Γ(5)	Γ(10)	Γ(20)
Wariant I – dzienne stopy zwrotu					
Arka BZ WBK Akcji FIO	1,9421	2,4903	2,9750	4,8636	7,3062
BPH FIO Akcji	1,7373	2,2170	2,6382	4,2591	6,0120
CU FIO Polskich Akcji	1,8579	2,3792	2,8404	4,6439	7,0115
DWS Polska FIO Akcji	1,9498	2,4852	2,9553	4,7791	7,1641
DWS Polska FIO Akcji Plus	1,8603	2,3775	2,8315	4,5847	6,8417
DWS Polska FIO Top 25 Małych Spółek	1,8684	2,4250	2,9243	4,9135	7,5264
ING FIO Akcji	1,7489	2,2316	2,6518	4,2454	6,2330
Legg Mason Akcji FIO	1,7548	2,2525	2,6920	4,4062	6,6172
Millennium FIO Akcji	1,7151	2,2052	2,6392	4,3419	6,5958
Pioneer Akcji Polskich FIO	1,9105	2,4463	2,9172	4,7335	7,0461
PKO/CREDIT SUISSE Akcji FIO	1,9172	2,4676	2,9573	4,8892	7,4185
PZU FIO Akcji KRAKOWIAK	1,7491	2,2526	2,6991	4,4513	6,7272
SEB 3 – Akcji FIO	1,8420	2,3672	2,8319	4,6522	7,0474
Skarbiec – Akcja FIO	1,7548	2,2532	2,6924	4,3905	6,5228
UniKorona Akcja FIO	1,8355	2,3460	2,7927	4,5068	6,6962
Wariant II – miesięczne stopy zwrotu					
Arka BZ WBK Akcji FIO	5,8378	7,2981	8,4353	11,8852	14,7267
BPH FIO Akcji	4,8804	6,0454	6,9169	9,4222	11,4154
CU FIO Polskich Akcji	5,3918	6,7347	7,7587	10,7837	13,3083
DWS Polska FIO Akcji	4,9117	6,1551	7,1106	9,9573	12,3644
DWS Polska FIO Akcji Plus	5,0188	6,2296	7,1485	9,8721	12,1603
DWS Polska FIO Top 25 Małych Spółek	5,9822	7,3412	8,3490	11,1643	13,1802
ING FIO Akcji	5,1822	6,3500	7,2188	9,7169	11,7155
Legg Mason Akcji FIO	5,0135	6,2455	7,1814	9,9169	12,1770
Millennium FIO Akcji	5,0285	6,2303	7,1295	9,7074	11,7855
Pioneer Akcji Polskich FIO	5,7069	7,0854	8,1332	11,2141	13,8310
PKO/CREDIT SUISSE Akcji FIO	5,0725	6,3754	7,3928	10,5235	13,2508
PZU FIO Akcji KRAKOWIAK	4,8754	6,1260	7,0742	9,8196	12,0740
SEB 3 – Akcji FIO	4,9723	6,2186	7,1722	10,0190	12,4611
Skarbiec – Akcja FIO	4,7864	5,9630	6,8544	9,4031	11,3915
UniKorona Akcja FIO	4,9901	6,1618	7,0484	9,6503	11,6988

Interesujące wydaje się w tym kontekście pytanie, czy zmiana wartości parametru ν ma istotne znaczenie dla porządkowania funduszy, czy też nie zmienia ich kolejności w zestawieniach. Innymi słowy, czy inwestor z niewielką awersją do ryzyka będzie preferował te same fundusze akcyjne, co inwestor z awersją znacząco większą. Wyniki uzyskane we wszystkich analizowanych wariantach zawiera tabela 5.

Tabela 5

Rankingi wybranych OFI według wartości współczynnika $\Gamma(\nu)$

Fundusz inwestycyjny	$\Gamma(3)$	$\Gamma(4)$	$\Gamma(5)$	$\Gamma(10)$	$\Gamma(20)$
Wariant I – dzienne stopy zwrotu					
Arka BZ WBK Akcji FIO	14	15	15	13	13
BPH FIO Akcji	2	2	1	2	1
CU FIO Polskich Akcji	9	10	10	9	9
DWS Polska FIO Akcji	15	14	13	12	12
DWS Polska FIO Akcji Plus	10	9	8	8	8
DWS Polska FIO Top 25 Małych Spółek	11	11	12	15	15
ING FIO Akcji	3	3	3	1	2
Legg Mason Akcji FIO	5	4	4	5	5
Millennium FIO Akcji	1	1	2	3	4
Pioneer Akcji Polskich FIO	12	12	11	11	10
PKO/CREDIT SUISSE Akcji FIO	13	13	14	14	14
PZU FIO Akcji KRAKOWIAK	4	5	6	6	7
SEB 3 – Akcji FIO	8	8	9	10	11
Skarbiec – Akcja FIO	6	6	5	4	3
UniKorona Akcja FIO	7	7	7	7	6
Wariant II – miesięczne stopy zwrotu					
Arka BZ WBK Akcji FIO	1	1	1	1	1
BPH FIO Akcji	3	2	2	2	2
CU FIO Polskich Akcji	9	8	6	4	5
DWS Polska FIO Akcji	6	5	3	3	3
DWS Polska FIO Akcji Plus	13	13	13	14	14
DWS Polska FIO Top 25 Małych Spółek	10	11	11	11	12
ING FIO Akcji	12	12	12	12	13
Legg Mason Akcji FIO	8	7	7	7	7
Millennium FIO Akcji	14	14	15	15	15
Pioneer Akcji Polskich FIO	2	3	4	6	6
PKO/CREDIT SUISSE Akcji FIO	5	6	8	10	10
PZU FIO Akcji KRAKOWIAK	4	4	5	9	9
SEB 3 – Akcji FIO	7	9	9	8	8
Skarbiec – Akcja FIO	15	15	14	13	11
UniKorona Akcja FIO	11	10	10	5	4

Porównując pozycje zajmowane w rankingach (oddzielnie dla obu rozpatrywanych wariantów stóp zwrotu) widać, że w kolejnych zestawieniach zmieniają się i liderzy, i najgorzej oceniane fundusze, natomiast zajmowane przez nie pozycje w niektórych przypadkach są podobne, w innych różnią się znacznie. Daje się przy tym zauważyć pewną, skądinąd naturalną, prawidłowość: im bardziej różnią się wartości parametru ν, tym większe są

różnice w rankingach. Potwierdzają to wartości współczynnika τ Kendalla (tabela 6), które przy zbliżonych ν sięgają niemal 1, a w innych przypadkach wahają się w pobliżu 0,7 w wariancie I oraz 0,6 w wariancie II.

Tabela 6

Wartości współczynnika τ Kendalla dla rankingów uzyskanych dla różnych ν w tych samych wariantach stóp zwrotu

Wariant I – dzienne stopy zwrotu					
	$\nu = 3$	$\nu = 4$	$\nu = 5$	$\nu = 10$	$\nu = 20$
$\nu = 3$	1	0,9429	0,8476	0,6952	0,6571
$\nu = 4$		1	0,9048	0,7524	0,7143
$\nu = 5$			1	0,8476	0,8095
$\nu = 10$				1	0,9238
$\nu = 20$					1
Wariant II – miesięczne stopy zwrotu					
	$\nu = 3$	$\nu = 4$	$\nu = 5$	$\nu = 10$	$\nu = 20$
$\nu = 3$	1	0,9048	0,7905	0,5810	0,5238
$\nu = 4$		1	0,8857	0,6762	0,6191
$\nu = 5$			1	0,7905	0,7333
$\nu = 10$				1	0,9429
$\nu = 20$					1

Oznacza to, że inwestorzy o podobnym poziomie awersji do ryzyka będą preferować te same fundusze. Natomiast ci, których niechęć do ryzyka różni się znacznie, zainteresują się innymi OFI akcji.

Jeszcze mniejsze podobieństwo rankingów widoczne jest wówczas, gdy porównamy ryzyko wyznaczone w oparciu o dzienne i miesięczne stopy zwrotu (przy tych samych wartościach ν). Mają miejsce nawet sytuacje, że fundusz, który w wariancie I zajmuje najwyższe lokaty w rankingach, w II wariancie oceniany jest jako najbardziej ryzykowny i zajmuje pozycję 14 czy 15 (np. Millennium). Potwierdzają to wartości współczynnika τ Kendalla (tabela 7) wahające się od zaledwie 0,1048 do 0,3333.

Tabela 7

Wartości współczynnika τ Kendalla dla rankingów uzyskanych dla różnych wariantów stóp zwrotu

$\nu = 3$	$\nu = 4$	$\nu = 5$	$\nu = 10$	$\nu = 20$
0,2762	0,3143	0,3333	0,2000	0,1048

Związane jest to, podobnie jak w przypadku odchylenia standardowego, ze zmianą horyzontu czasowego przyjętego przy wyznaczaniu stóp zwrotu. Wydaje się jednak, że uogólniony wskaźnik Giniego wykazuje dużo większą wrażliwość na takie zmiany.

Zakończenie

Uogólniony wskaźnik Giniego oraz średnia różnica Giniego są interesującą propozycją sposobu pomiaru ryzyka inwestycyjnego. W prowadzonych analizach mogą zastąpić odchylenie standardowe stopy zwrotu. Szczególnie istotną wydaje się możliwość uwzględniania poziomu awersji do ryzyka inwestora, co pozwala na większą „indywidualizację" dokonywanych ocen. Ten aspekt zastosowań prezentowanych miar należałoby jednak poddać dalszym badaniom uwzględniającym, np. różne typy funduszy inwestycyjnych. W przeprowadzonych analizach wzięto bowiem pod uwagę tylko OFI akcji.

Literatura

1. Bey R.P., Howe K.M. (1984). Gini's Mean Difference and Portfolio Selection: An Empirical Evaluation. Journal of Financial and Quantitative Analysis, 19, 3, 329-338.
2. Dorfman R. (1978). A Formula for the Gini Coefficient. The Review of Economics and Statistics, 146-149.
3. Gluzicka A. (2008). Ocena ryzyka inwestycyjnego na podstawie miar koncentracji. W: Modelowanie preferencji a ryzyko '07. Red. T. Trzaskalik. AE, Katowice, 79-90.
4. Gregory-Allen R.B., Shalit H. (1999). The Estimation of Systematic Risk under Differentiated Risk Aversion: A Mean-Extended Gini Approach. Review of Quantitative Finance and Accounting, 12, 135-157.
5. Haugen R.A. (1996). Teoria nowoczesnego inwestowania. WIG-Press, Warszawa.
6. Kowalski J.M. (2006). Podstawy statystyki opisowej dla ekonomistów. Wyższa Szkoła Bankowa, Poznań.
7. Shalit H. (1985). Calculating the Gini Index of inequality for Individual Data. Oxford Bulletin of Economics and Statistics, 47, 2, 185-189.
8. Shalit H., Yitzhaki S. (1984). Mean-Gini, Portfolio Theory, and the Pricing of Risky Assets. The Journal of Finance, XXXIX, 5, 1449-1468.
9. Shalit H., Yitzhaki S. (1989). Evaluating the Mean-Gini Approach to Portfolio Selection. International Journal of Finance, 1, 15-31.
10. Shalit H., Yitzhaki S. (2005). The Mean-Gini Efficient Portfolio Frontier. The Journal of Financial Research, XXVIII, 1, 59-75.
11. Yitzhaki S. (1982). Stochastic Dominance, Mean Variance, and Gini's Mean Difference. The American Economic Review, 72, 1,178-185.

Marcin Salamaga
Uniwersytet Ekonomiczny w Krakowie

PRÓBA WYKRYCIA PUNKTÓW ZWROTNYCH W KSZTAŁTOWANIU SIĘ KURSU WALUTOWEGO DOLARA AMERYKAŃSKIEGO

Wprowadzenie

Kurs walutowy to stosunek, w jakim wykonywana jest wymiana określonej ilości danej waluty na jednostkę innej waluty [8]. Jest on wyznaczany na rynku walutowym przez popyt i podaż na daną walutę. Uczestnikami rynku walutowego są duże instytucje finansowe (banki, fundusze inwestycyjne), agendy rządowe, agendy samorządowe, firmy produkcyjno-usługowe i wreszcie osoby fizyczne [7].

Dla polskiej gospodarki istotne znaczenie mają kursy dwóch walut: euro i dolara amerykańskiego. Wynika to między innymi z faktu, że w tych walutach rozliczna jest większość transakcji z innymi krajami. Dolar amerykański pozostaje głównym środkiem płatniczym poza Unią Europejską tak w sferze prywatnej, jak i instytucjonalnej. Jego kurs ma bezpośrednie znaczenie przede wszystkim dla importerów i eksporterów rozliczających transakcje w dolarach. W konsekwencji znajduje to odzwierciedlenie w cenach towarów sprowadzanych z zagranicy lub wytwarzanych w Polsce na bazie komponentów zagranicznych. Te cechy waluty USA przesądziły o jej wyborze jako przedmiotu analizy w niniejszym opracowaniu.

Wśród czynników kształtujących poziom kursów walutowych można wymienić między innymi ekonomiczne, polityczne i psychologiczne [2].

W ostatnim czasie zwłaszcza grupa czynników psychologicznych związanych najczęściej z oczekiwaniami inwestorów co do przyszłej sytuacji ekonomicznej i politycznej państwa miała wpływ na dużą zmienność kursu dolara w Polsce. Należy podkreślić również znaczne oddziaływanie na kurs dolara sytuacji ekonomicznej całego regionu Europy Środkowo-Wschodniej. Niestabilność ekonomiczna (lub polityczna) któregokolwiek kraju tzw. rynków wschodzących natychmiast znajduje przełożenie w osłabianiu się waluty tego kraju, jak i innych walut regionu.

W literaturze przedmiotu istnieje wiele teorii tłumaczących kształtowanie się kursów walutowych w długim i w krótkim okresie, spośród których można wymienić teorie: parytetu siły nabywczej, parytetu stóp procentowych, oczekiwań czy efekt Fishera [8].

Duża zmienność kursu dolara w ostatnim okresie skłania do analizy przyczyn tego zjawiska. W opracowaniu podjęto próbę określenia punktów zwrotnych w kształtowaniu się kursu dolara amerykańskiego. Wzięto pod uwagę dzienne notowania średniego kursu dolara amerykańskiego ogłaszane w NBP w okresie od stycznia 2006 roku do końca 2008.

Przebieg kursu waluty obcej można potraktować jako realizację szeregu czasowego, w którym zwykle da się wyodrębnić tendencje rozwojowe kształtujące się w różny sposób w zależności od rozważanych okresów. Obserwacja szeregu czasowego kursu dolara amerykańskiego w ciągu ostatnich kilku lat pozwala dostrzec co najmniej kilka momentów, w których krótkookresowe trendy uległy gwałtownym zmianom. Może to wskazywać na występowanie punktów zwrotnych.

W opracowaniu wykorzystano dwie metody mające na celu wykrycie punktów zwrotnych w kształtowaniu się kursu dolara: test Chowa, test Perrona oraz zaproponowano analizę trendu różnic pomiędzy rzeczywistymi realizacjami kursu dolara i prognozowanymi wartościami tego kursu. Dla wykrytych punktów zwrotnych podjęto próbę ustalenia przyczyn, które mogły wywołać zmianę w kształtowaniu się kursu dolara amerykańskiego.

1. Wybrane metody badania punktów zwrotnych trendu

Jedną z bardziej popularnych procedur służących do identyfikacji punktów zwrotnych w szeregu czasowym jest test Chowa. Idea testu polega na podzieleniu wybranego odcinka czasu, w którym badany jest proces, na dwa podokresy i następnie oszacowaniu dwóch modeli trendu osobno dla każdego z tych podokresów. Potem weryfikuje się hipotezę statystyczną, czy współczynniki obu trendów różnią się istotnie od siebie. W przypadku odrzucenia hipotezy o braku istotnych różnic pomiędzy współczynnikami obu trendów, punkt rozdzielający oba podokresy jest punktem zwrotnym. Test Chowa stosuje się również do badania stabilności modelu w podpróbkach próby przekrojowej.

W teście sprawdzającym hipotezę o braku różnic pomiędzy współczynnikami trendów można wykorzystać model trendu (1) uwzględniający zmienną zero-jedynkową $R_{p,t}$ zdefiniowaną następująco: $R_{p,t}$ =1, gdy obserwacja i należy do podpróbki p oraz $R_{p,t}$ =0, gdy obserwacja i nie należy do podpróbki p.

Szacowany model trendu w przypadku dwóch podpróbek (podokresów) można zapisać następująco [5]

$$y_t = \sum_{p=1}^{2} R_{p,t} \cdot t \cdot \beta_p + \varepsilon_t \tag{1}$$

gdzie:

t =1,2,3,..,n,
β_p – parametry modelu,
ε_t – składnik losowy modelu.

Statystyka testu wykorzystywana w testowaniu punktów zwrotnych trendu ma postać [5]

$$F = \frac{(ESS_{III} - ESS_I - ESS_{II})/k}{(ESS_I + ESS_{II})/(n-2k)} \tag{2}$$

gdzie:

ESS_I – suma kwadratów reszt modelu trendu dla pierwszego podokresu,
ESS_{II} – suma kwadratów reszt modelu trendu dla drugiego podokresu,
ESS_{III} – suma kwadratów reszt modelu trendu dla całego okresu,
k – liczba szacowanych parametrów modelu.

Statystyka F ma rozkład F-Snedecora o k oraz n-$2k$ stopniach swobody.

Poszukiwanie punktów zwrotnych w kształtowaniu się kursu waluty amerykańskiej przeprowadzono także za pomocą testu Perrona. Test ten pozwala jednocześnie na ocenę stacjonarności szeregu czasowego w połączeniu z uwzględnieniem zmian w strukturze zjawiska [10]. W teście Perrona można analizować zmiany w strukturze trendu pod kątem wyłącznie współczynnika kierunkowego, wyrazu wolnego lub obu tych parametrów równocześnie. W tym opracowaniu ograniczono się do testowania zamian jednocześnie obu parametrów trendu.

W takiej sytuacji rozważany jest następujący model [3]

$$\begin{aligned} y_t &= \beta_0 + \beta_1 t + \theta D(U)_t + \gamma D(T)_t + \\ &+ \delta D(TB)_t + \alpha_1 y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-1} + \varepsilon_t \end{aligned} \tag{3}$$

przy czym
$\beta_0, \beta_1, \theta, \gamma, \delta, \alpha_1, c_i$ – parametry modelu (3),
t = 1,2,3,...,T,

$$D(U)_t = \begin{cases} 1 & dla \quad t > T_B \\ 0 & dla \quad t \le T_B \end{cases}, \quad D(TB)_t = \begin{cases} 1 & dla \quad t = T_B + 1 \\ 0 & dla \quad t \ne T_B + 1 \end{cases}$$

$$D(T)_t = \begin{cases} t & dla \quad t > T_B \\ 0 & dla \quad t \le T_B \end{cases}$$

T_B – rozważany moment wystąpienia zmian w strukturze trendu (punkt zwrotny).

Weryfikowane tu hipotezy mają postać [3]

$$\begin{aligned} &H_0: \theta=0; \beta_1=0; \gamma=0; \delta\neq 0; \alpha_1=1 \\ &H_1: \theta\neq 0; \beta_1\neq 0; \gamma\neq 0; \delta=0; \alpha_1<1 \end{aligned} \tag{4}$$

W hipotezie zerowej sprawdzane jest istnienie pierwiastka jednostkowego (zatem brak stacjonarności szeregu), któremu towarzyszy pojedyncza gwałtowna zamiana w strukturze (tzw. szok). Z kolei w hipotezie alternatywnej zakłada się brak pierwiastka jednostkowego (czyli stacjonarność szeregu) oraz fakt, że analizowana zamiana nie ma charakteru szokowego i dotyczy wyrazu wolnego i współczynnika trendu [3]. W literaturze przedmiotu znanych jest kilka testów służących do badania stacjonarności szeregu, jak test D.A. Dickeya i W.A. Fullera (test DF) czy rozszerzony test Dickeya-Fullera (test ADF) stosowany w przypadku występowania autokorelacji składnika losowego [9].

W niniejszym artykule do zweryfikowania hipotezy o pierwiastku jednostkowym ($\alpha_1=1$) wykorzystano test zaproponowany przez Perrona [3]

$$t^* = \frac{\hat{\alpha}_1 - 1}{D(\hat{\alpha}_1)} \tag{5}$$

gdzie:

$\hat{\alpha}_1$ – ocena z próby parametru α_1,
$D(\hat{\alpha}_1)$ – średni błąd oceny parametru α_1.

Wartości krytyczne testu Perrona przedstawione są np. w pracy [10].

Trzecim sposobem, jaki zaproponowano do wykrycia punktów zwrotnych w kształtowaniu się kursu waluty amerykańskiej, jest badanie istotności trendu, jaki tworzą odchylenia pomiędzy rzeczywistymi realizacjami kursu dolara w okresie następującym po wystąpieniu punktu zmiany trendu a ich prognozami uzyskanymi na podstawie trendu oszacowanego dla obserwacji w okresie poprzedzającym badany moment.

Proponowana procedura wykrycia punktów zwrotnych wymaga na początku oszacowania parametrów funkcji trendu (6) dla podpróbki bezpośrednio poprzedzającej moment wystąpienia potencjalnego punktu zwrotnego

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \tag{6}$$

przy czym
t=1,2,3,...,T_B,
T_B – moment wystąpienia punktu zwrotnego,
β_0, β_1 – parametry modelu trendu (6),
ε_t – składnik losowy równania trendu (6).

Na podstawie modelu (6) dokonujemy ekstrapolacji kursu waluty dla okresów $t = T_B + 1$, $T_B + 2$, $T_B + 3$,...,$T_B + T$ uzyskując ciąg kolejnych prognoz kursu w drugiej podpróbce: $y^p_{T_B+1}, y^p_{T_B+2}, y^p_{T_B+3}, ..., y^p_{T_B+T}$. W kolejnym korku od rzeczywistych realizacji kursów dolara w okresie następującym po wystąpieniu momentu T_B odjęto wartości otrzymanych prognoz wygasłych uzyskując ciąg różnic: $\widetilde{y}_{T_B+1}, \widetilde{y}_{T_B+2}, \widetilde{y}_{T_B+3}, ..., \widetilde{y}_{T_B+T}$.

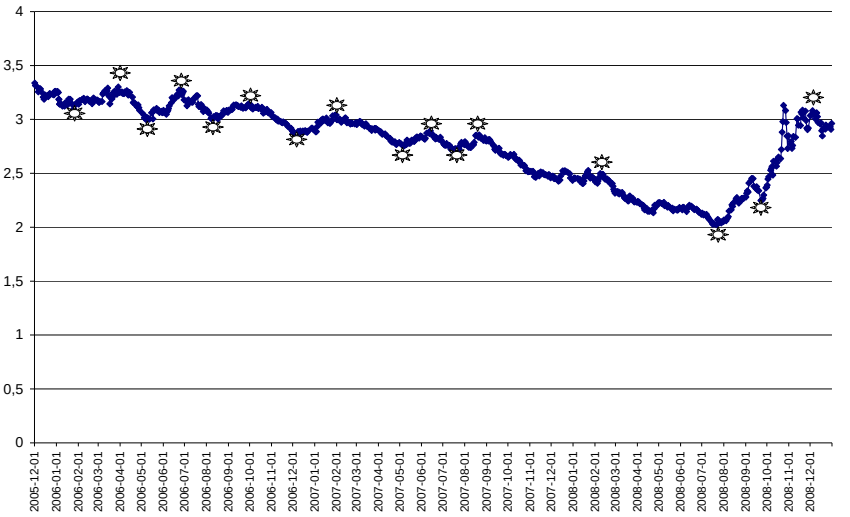
Dla obliczonych w ten sposób różnic można ponownie oszacować model trendu (7)

$$\widetilde{y}_t = \alpha_0 + \alpha_1 t + \varepsilon_t \tag{7}$$

Statystyczna istotność parametrów α_0 oraz α_1 szacowanego modelu (7) może dowodzić, że moment T_B jest punktem zwrotnym w szeregu czasowym.

2. Wyniki badań empirycznych

W badaniu uwzględniono szeregi dziennych notowań kursu dolara amerykańskiego ogłaszane przez NBP w latach 2006-2008. Analizując przebieg zmian kursu waluty amerykańskiej w tym okresie wytypowano 16 momentów zmiany współczynnika trendu, wyrazu wolnego trendu (lub obu parametrów jednocześnie), które mogły być potencjalnymi punktami zwrotnymi (rys. 1).



Rys. 1. Momenty zmiany trendu kursu dolara amerykańskiego w latach 2006-2008

Źródło: Opracowanie własne na podstawie danych NBP.

W niektórych przypadkach można wskazać daty, którym odpowiada zmiana trendu, a w innych sytuacjach są to zwykle krótkie (kilkudniowe) przedziały czasowe. Jeżeli zmianę trendu obserwowano w pewnym okresie, to jako potencjalny punkt zwrotny trendu proponowano środek właściwego przedziału czasowego. Centralny punkt przedziału czasowego, w którym szereg czasowy nie wykazuje wyraźnej tendencji rozwojowej, wydaje się być akceptowalnym „reprezentantem" takiego przedziału [3].

Dla uzyskania porównywalności wyników stosowanych testów w przypadku różnych punktów starano się dobierać równoliczne próby, dla których środkowym punktem był moment zmiany trendu. Z reguły były to próby 42-elementowe (zawierające 21 dni poprzedzających moment zmiany trendu oraz 20 dni następujących po tym momencie). O przyjęciu takiej rozpiętości czasowej odcinków szeregu czasowego do analizy zdecydowała struktura kształtowania się kursu waluty amerykańskiej i częstotliwość zmian w niej zachodzących (zastosowanie większej liczby danych mogło spowodować w niektórych przypadkach uwzględnienie w próbie np. dwóch potencjalnych punktów zwrotnych). Dla tak dobieranych prób przeprowadzono trzy procedury omówione w poprzednim paragrafie. Wyniki badania przedstawiono w tabelach 1-3.

Tabela 1 zawiera wyniki testu Chowa dla szesnastu momentów zmiany trendu w notowaniach kursu USD (w nawiasach podano wartości prawdopodobieństw testowych). Z przedstawionych danych wynika, że wszystkie momenty zmiany trendu są punktami zwrotnymi w kształtowaniu się kursu dolara

amerykańskiego za wyjątkiem ostatniego momentu, tj. 24.11.2008 (wartość statystyki *F* dla każdej z pozostałych 15 prób była statystycznie istotna). Ostrożniej należy podchodzić do interpretacji wyników testu Chowa w przypadku prób nr 13 oraz nr 15, w których współczynniki trendów oszacowane dla dwudziestu okresów poprzedzających punkty zwrotne nie były statystycznie istotne.

Tabela 1

Wyniki testu Chowa dla szesnastu momentów zmiany trendu

Nr	Moment zmiany trendu	Oceny parametrów modelu w pierwszej podpróbce		Zmiany ocen parametrów		F
		β_0	β_1	$\Delta\beta_0$	$\Delta\beta_1$	
1	25.01.2006	3,215 (0,000)	-0,004 (0,009)	-0,128 (0,000)	0,006 (0,000)	8,043 (0,001)
2	29.03.2006	3,183 (0,000)	0,003 (0,029)	0,299 (0,011)	-0,012 (0,000)	26,053 (0,000)
3	8.05.2006	3,304 (0,000)	-0,012 (0,000)	-0,360 (0,000)	0,016 (0,000)	90,376 (0,000)
4	23.06.2006	3,022 (0,000)	0,009 (0,000)	0,296 (0,000)	-0,013 (0,000)	74,395 (0,000)
5	14.08.2006	3,214 (0,000)	-0,011 (0,000)	-0,325 (0,000)	0,016 (0,000)	67,529 (0,000)
6	29.09.2006	3,094 (0,000)	0,002 (0,009)	0,096 (0,000)	-0,004 (0,000)	17,156 (0,000)
7	7.12.2006	3,018 (0,000)	-0,007 (0,000)	-0,249 (0,000)	0,011 (0,000)	82,316 (0,000)
8	30.01.2007	2,920 (0,000)	0,005 (0,000)	0,173 (0,000)	-0,008 (0,000)	26,422 (0,000)
9	8.05.2007	2,881 (0,000)	-0,006 (0,000)	-0,186 (0,000)	0,010 (0,000)	90,693 (0,000)
10	13.06.2007	2,781 (0,000)	0,004 (0,000)	0,238 (0,000)	-0,011 (0,000)	109,1 (0,000)
11	19.07.2007	2,842 (0,000)	-0,006 (0,000)	-0,194 (0,000)	0,010 (0,000)	35,860 (0,000)
12	21.08.2007	2,728 (0,000)	0,004 (0,001)	0,267 (0,000)	-0,010 (0,000)	38,725 (0,000)
13	12.02.2008	2,451 (0,000)	0,001 (0,627)	0,242 (0,000)	-0,010 (0,000)	28,277 (0,000)
14	21.07.2008	2,183 (0,000)	-0,007 (0,000)	-0,448 (0,000)	0,019 (0,000)	105,693 (0,000)
15	25.09.2008	2,338 (0,000)	0,000 (0,933)	-0,632 (0,000)	0,026 (0,000)	30,298 (0,000)
16	24.11.2008	2,880 (0,000)	0,004 (0,325)	0,241 (0,050)	-0,008 (0,122)	2,086 (0,138)

W nawiasach przedstawiono wartości prawdopodobieństw testowych.

Oceny parametrów modelu (3) oraz wyniki testu Perrona (5) dla szesnastu momentów zmiany trendu przedstawiono w tabeli 2. W sześciu przypadkach stwierdzono stacjonarność szeregów czasowych. Miało to miejsce w następujących przedziałach czasowych (tabela 2): 04.04.2006-05.06.2006 (próba nr 3), 24.05.2006-21.07.2006 (próba nr 4), 08.07.2006-12.09.2006 (próba nr 5), 08.11.2006-09.01.2007 (próba nr 7), 04.04.2007-06.06.2007 (próba nr 9), 14.05.2007-11.07.2007 (próba nr 10).

Tabela 2

Wyniki testu Perrona dla szesnastu momentów zmiany trendu

Lp.	Moment zmiany trendu	Parametry modelu (1)						Test Perrona
		β_0	β_1	θ	γ	δ	α_1	t^*
1	25.01.2006	1,598 (0,000)	-0,002 (0,112)	-0,028 (0,200)	0,002 (0,076)	-0,053 (0,050)	0,501 (0,000)	-3,366 (0,075)
2	29.03.2006	1,102 (0,010)	0,001 (0,231)	0,1 (0,066)	-0,004 (0,038)	0,028 (0,342)	0,654 (0,000)	-1,138 (0,772)
3	8.05.2006	0,680 (0,014)	-0,003 (0,001)	-0,124 (0,011)	0,004 (0,001)	0,021 (0,419)	0,798 (0,000)	-3,5575 (0,048)
4	23.06.2006	0,879 (0,011)	0,003 (0,014)	0,073 (0,117)	-0,004 (0,029)	0,058 (0,037)	0,708 (0,000)	-3,637 (0,042)
5	14.08.2006	2,281 (0,000)	-0,007 (0,000)	-0,238 (0,000)	0,012 (0,000)	0,021 (0,302)	0,289 (0,079)	-3,911 (0,022)
6	29.09.2006	1,225 (0,004)	0,001 (0,269)	0,028 (0,175)	-0,002 (0,079)	0,017 (0,216)	0,605 (0,000)	-2,770 (0,270)
7	7.12.2006	1,009 (0,012)	-0,003 (0,016)	-0,088 (0,021)	0,004 (0,009)	0,001 (0,922)	0,665 (0,000)	-4,065 (0,016)
8	30.01.2007	1,021 (0,008)	0,002 (0,038)	0,042 (0,212)	-0,003 (0,067)	0,015 (0,421)	0,651 (0,000)	-2,060 (0,550)
9	8.05.2007	1,863 (0,001)	-0,004 (0,001)	-0,106 (0,011)	0,006 (0,003)	-0,009 (0,487)	0,352 (0,047)	-4,029 (0,018)
10	13.06.2007	1,787 (0,000)	0,003 (0,000)	0,138 (0,001)	-0,007 (0,000)	0,031 (0,036)	0,357 (0,017)	-4,026 (0,018)
11	19.07.2007	0,699 (0,073)	-0,001 (0,203)	-0,050 (0,147)	0,003 (0,089)	-0,015 (0,429)	0,752 (0,000)	-0,569 (0,973)
12	21.08.2007	0,696 (0,032)	0,001 (0,141)	0,056 (0,167)	-0,003 (0,085)	0,022 (0,263)	0,746 (0,000)	-1,5396 (0,755)
13	12.02.2008	0,540 (0,044)	0,000 (0,818)	0,03 (0,419)	-0,002 (0,208)	0,013 (0,501)	0,781 (0,000)	-2,377 (0,424)
14	21.07.2008	0,340 (0,134)	-0,001 (0,170)	-0,093 (0,068)	0,004 (0,040)	-0,001 (0,942)	0,842 (0,000)	-1,446 (0,990)
15	25.09.2008	0,354 (0,222)	-0,002 (0,274)	-0,139 (0,140)	0,007 (0,044)	-0,003 (0,949)	0,857 (0,000)	-0,0017 (0,990)
16	24.11.2008	0,221 (0,081)	0,001 (0,650)	0,046 (0,462)	-0,003 (0,365)	0,061 (0,357)	0,923 (0,000)	-2,331 (0,443)

W nawiasach przedstawiono wartości prawdopodobieństw testowych.

We wszystkich wymienionych próbach odrzucono hipotezę zerową (4) na rzecz hipotezy alternatywnej zakładającej istotne zmiany obu parametrów trendu, przy czym zmiana tendencji rozwojowej kursu dolara nie miała charakteru szokowego w tych przedziałach czasowych (nie stwierdzono istnienia pierwiastka jednostkowego). W pozostałych próbach nie było podstaw do odrzucenia hipotezy zerowej (4) zakładającej brak stacjonarności szeregu i wystąpienie pojedynczej zmiany trendu (szoku) bez istotnej zmiany parametrów trendu.

W tabeli 3 podano oceny parametrów trendu (7) utworzonego przez różnice pomiędzy rzeczywistymi realizacjami kursu dolara w drugiej podpróbce (po wystąpieniu momentu zmiany trendu) a ich prognozami uzyskanymi na podstawie modelu trendu (6) oszacowanego dla pierwszej podpróbki (przed wystąpieniem momentu zmiany trendu).

Tabela 3

Wyniki testu istotności trendu różnic y_t-y_t^p dla szesnastu momentów zmiany trendu

Lp.	Punkt zwrotny	Oceny parametrów trendu y_t-y_t^p	
		α_1	α_0
1	2	3	4
1	25.01.2006	0,002 (0,060)	0,056 (0,000)
2	29.03.2006	-0,018 (0,000)	0,007 (0,617)
3	8.05.2006	0,009 (0,000)	0,048 (0,000)
4	23.06.2006	-0,009 (0,007)	-0,048 (0,043)
5	14.08.2006	0,017 (0,000)	0,067 (0,000)
6	29.09.2006	-0,006 (0,000)	-0,019 (0,044)
7	7.12.2006	0,016 (0,000)	0,003 (0,830)
8	30.01.2007	-0,009 (0,000)	-0,029 (0,000)
9	8.05.2007	0,007 (0,000)	0,074 (0,000)
10	13.06.2007	-0,011 (0,000)	-0,046 (0,000)
11	19.07.2007	0,015 (0,000)	0,013 (0,533)

cd. tabeli 3

1	2	3	4
12	21.08.2007	-0,014 (0,000)	0,011 (0,312)
13	12.02.2008	-0,006 (0,000)	-0,106 (0,000)
14	21.07.2008	0,030 (0,000)	-0,004 (0,674)
15	25.09.2008	0,037 (0,000)	0,104 (0,041)
16	24.11.2008	-0,022 (0,000)	0,099 (0,000)

W nawiasach przedstawiono wartości prawdopodobieństw testowych.

Na podstawie danych w tabeli 3 można stwierdzić, że oba parametry trendu różnic y_t-y_t^p w drugiej podpróbce były statystycznie istotne w ośmiu przypadkach, którym odpowiadają próby o numerach 3, 4, 5, 6, 8, 9, 10, 16. W pozostałych przypadkach przynajmniej jeden z parametrów trendu (7) α_0 lub α_1 nie był statystycznie istotny.

Należy tutaj podkreślić, że brak statystycznej istotności parametru α_0 w modelu (7) nie musi oznaczać braku punktu zwrotnego w kształtowaniu się kursu dolara, gdy istotny pozostaje współczynnik trendu α_1. Podobnie można rozumować w przypadku, gdy współczynnik α_1 jest statystycznie nieistotny, a istotny pozostaje wyraz wolny α_0 (np. próba nr 1), ponieważ nawet wtedy w przybliżeniu stały ciąg różnic y_t-y_t^p może również wskazywać na odmienny kierunek kształtowania się wartości y_t oraz prognoz y_t^p w okresach $t= T_B+1$, T_B+2, T_B+3,...,T_B+T. Jednoczesny brak istotności obu parametrów α_0 oraz α_1 może wskazywać na brak punktu zwrotnego w kształtowaniu się kursu dolara amerykańskiego, jednak takiego przypadku nie zaobserwowano.

Wyniki testu Chowa oraz wyniki analizy istotności parametrów trendu (7) nie podważyły jednoznacznie występowania sześciu punktów zwrotnych w przebiegu kursu dolara wyznaczonych w teście Perrona. Dlatego dalsze rozważania ograniczono do sześciu punktów zwrotnych wskazanych właśnie przy użyciu testu Perrona. Opierając się na doniesieniach prasowych, źródłach internetowych starano się wskazać głównie wydarzenia gospodarcze i polityczne, które ewentualnie mogły wpłynąć na zmianę w kształtowaniu się kursu dolara w odniesieniu do sześciu punktów zwrotnych odpowiadających datom: 08.05.2006, 23.06.2006, 14.08.2006, 07.12.2006, 08.05.2007, 13.06.2007. Badane wydarzenia zostały podzielone w układzie geograficznym na: zachodzące w Polsce, zachodzące na rynkach wschodzących i mające miejsce w Stanach Zjednoczonych. Odpowiednie dane przedstawiono w tabeli 4.

Tabela 4

Identyfikacja przyczyn wystąpienia punktów zwrotnych
w kształtowaniu się kursu dolara amerykańskiego

Punkt zwrotny	Wydarzenia w Polsce		Wydarzenia na innych rynkach wschodzących	Wydarzenia na rynku USA
	gospodarcze	polityczne		
1	2	3	4	5
08.05.2006	Realizacja zysków z fali dotychczasowego umacniania się polskiej waluty	Rekonstrukcja rządu w Polsce, dymisja wiceministra finansów	–	Oczekiwanie inwestorów na zmianę stopy procentowej przez FED
23.06.2006	Korzystne informacje makroekonomiczne o dynamice produkcji przemysłowej, dynamice sprzedaży detalicznej	Odwołanie ministra finansów	–	Informacje o możliwości spowolnienia gospodarczego w USA oraz o braku równowagi handlowej w USA
14.08.2006	–	Groźba przedterminowych wyborów w Polsce, rozpoczęcie pracy przez komisję śledczą ds. banków	Przesunięcie terminu wejścia Węgier do strefy euro, niechęć rządu Słowacji do prywatyzacji, silna presja inflacyjna w Turcji	Oczekiwania podwyżek stóp procentowych przez FED w USA
7.12.2006	–	Informacje o możliwości powołania rządu mniejszościowego, niepewność inwestorów co do budżetu na 2007 rok	–	Optymistyczne informacje makroekonomiczne o bilansie handlowym oraz wskaźnikach sprzedaży detalicznej i rynku pracy
8.05.2007	Pozytywne informacje z rynku pracy (stopa bezrobocia na poziomie 11%) oraz informacje o wzroście wartości eksportu	–	Destabilizacja sytuacji politycznej w Turcji	Brak zmian stopy procentowej przez FED, sygnały świadczące o poprawie sytuacji na rynku nieruchomości

cd. tabeli 4

1	2	3	4	5
13.06.2007	Wypłata dywidend ze spółek giełdowych, zapowiedzi niższego niż oczekiwano poziomu inflacji w Polsce	–	Publikacja danych o najwyższym od 2 lat poziomie inflacji w Chinach	Publikacja Beżowej Księgi, publikacja optymistycznych wskaźników sprzedaży detalicznej, informacja o wyższej niż oczekiwano dynamice wzrostu cen importu w USA

Źródło: Opracowanie własne na podstawie [13-25].

Z przedstawionych wyników w tabeli 4 można wnioskować, że równie często zmianie w kształtowaniu się kursu dolara towarzyszyły wydarzenia gospodarcze, polityczne w Polsce, jak i w Stanach Zjednoczonych. Duży wpływ na przebieg kursu waluty amerykańskiej miały także wydarzenia na rynkach wschodzących. W dobie silnej globalizacji rynków finansowych polityczne lub gospodarcze zawirowania tylko w jednym z krajów należących do tzw. rynków wschodzących często było wystarczającym powodem do wycofania się inwestorów z inwestycji w całym regionie geograficznym. W efekcie „umacnia" się waluta obca i „słabnie" waluta krajowa.

Najwięcej punktów zwrotnych stwierdzono w drugim i w trzecim kwartale 2006 roku. Należy zauważyć, że w tym okresie towarzyszyły im zdarzenia wskazujące na dużą niestabilność na polskiej scenie politycznej (odwołanie ministra finansów, rekonstrukcja rządu itp.). W rezultacie następowała aprecjacja amerykańskiego dolara oraz euro. Warto zwrócić uwagę na fakt, że prawie ze wszystkimi wyróżnionymi punktami zwrotnymi w kształtowaniu się kursu dolara można skojarzyć oczekiwanie inwestorów na informacje makroekonomiczne o stanie gospodarki (głównie w Polsce i w USA), a zwłaszcza na informacje o podwyżce lub obniżce stóp procentowych.

Podsumowanie

Spośród trzech metod badania punktów zwrotnych najbardziej „liberalnymi" metodami wydają się test Chowa oraz analiza istotności parametrów trendu różnic y_t-y_t^p. Te metody wskazały na największą liczbę punktów zwrotnych w kształtowaniu się kursu dolara. Wydaje się, że nie uwzględniają one w pełni zmian strukturalnych trendu badanego zjawiska ograniczając się do kontroli istotności parametrów trendu lub badania istotności zmian tych parametrów. Znacznie większe możliwości w ocenie zmian strukturalnych trendu daje test Perrona, który kontrolę istotności parametrów trendu łączy z badaniem

stacjonarności szeregu czasowego i testowaniem punktu wystąpienia nagłej zmiany w przebiegu zjawiska [3]. Na podstawie wyników tego testu stwierdzono jedynie 6 punktów zwrotnych kursu dolara amerykańskiego w latach 2006-2008. Warto zaznaczyć, że we właściwej selekcji potencjalnych punktów zwrotnych trendu może pomóc jednoznaczne zdefiniowanie czym jest punkt zwrotny i jakie kryteria powinien spełniać wyróżniony moment zmiany trendu, aby go nazwać punktem zwrotnym.

Każda z zastosowanych procedur akcentuje nieco różne aspekty związane ze zmianą trendu, stąd otrzymane wyniki przy ich użyciu nie są identyczne. Zdarzenia towarzyszące punktom zwrotnym przedstawionym w tabeli 4 w zasadzie odpowiadają determinantom kursów walutowych wymienianym w różnych teoriach ekonomicznych (np. [6]). Zwraca się w nich między innymi uwagę na to, że zmiana bieżącego kursu walutowego zależy od rozwoju sytuacji płatniczej kraju, od oczekiwanego oddziaływania przyszłych interwencji walutowych (np. poprzez bank centralny) czy od oczekiwanego kierunku zmian w polityce gospodarczej [7]. Warto też zaznaczyć, że omawiane zdarzenia zachodzące niekiedy w tym samym dniu lub w kolejnych dniach bliskich wyznaczonemu punktowi zwrotnemu mogły wzmacniać (na zasadzie interakcji) swój łączny wpływ na przebieg zmian kursu waluty, a w niektórych sytuacjach mogły neutralizować swoje oddziaływanie na kurs USD.

Literatura

1. Charemza W., Deadman D. (1997). Nowa ekonometria. PWE, Warszawa.
2. Bożyk P., Misala J., Puławski M. (1998). Międzynarodowe stosunki ekonomiczne. PWE, Warszawa.
3. Gasek A., Witkowska D. (2006). Zastosowanie testu Perrona do badania punktów zwrotnych indeksów giełdowych: WIG, WIG 20, MIDWIG i TECHWIG. Zeszyty Naukowe Ekonomika i Organizacja Gospodarki Żywnościowej. SGGW, Warszawa, nr 60.
4. Metody ekonometryczne i statystyczne w analizie rynku kapitałowego. (2000). Red. K. Jajuga. AE, Wrocław.
5. Kufel T. (2004). Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL. PWN, Warszawa.
6. Krugman P.R., Obstfeld M. (2007). Międzynarodowe stosunki gospodarcze. PWN, Warszawa.
7. Lutkowski K. (2007). Finanse międzynarodowe. Zarys problematyki. PWN, Warszawa.
8. Misztal P. (2004). Zabezpieczenie przed ryzykiem zmian kursu walutowego. Difin, Warszawa.

9. Ekonometria współczesna. (2007). Red. M. Osińska. Dom Organizatora, Toruń.
10. Perron P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. Econometrica, 57, 6, 1361-1401.
11. Piłatowska M. (2003). Modelowanie niestacjonarnych procesów ekonomicznych. Studium Metodologiczne. UMK, Toruń.
12. Welfe A., Welfe W. (2004). Ekonometria stosowana. PWE, Warszawa.

Źródła internetowe

http://www.e-rachunkowosc.pl/tms-archiwum.php?full=1
13. Gaworecki K. Dolar przeszedł do ofensywy. TMS Brokers, 13.06.2007.
14. Gaworecki K. Komentarz tygodniowy. TMS Brokers, 13.06.2007.
15. Rogalski M. Poniedziałek przyniósł nieco słabszą złotówkę i trend ten był kontynuowany także dzisiaj. TMS Brokers, 08.05.06.
16. Rogalski M. Ten tydzień był fatalny dla naszej waluty. TMS Brokers, 23.06.06.
17. Rogalski M. Czwartek przyniósł nieznaczne umocnienie się amerykańskiej waluty. TMS Brokers, 8.12.06.
http://www.rp.pl/
18. Rzeczpospolita. Rynki czekają na decyzję FED. 9.05.2006.
19. Rzeczpospolita. Dziś decyzja o stopach w USA. 10.05.2006.
20. Rzeczpospolita. Złoty traci na wartości. 13.06.2006.
21. Rzeczpospolita. Sytuacja w regionie zła dla złotego. 23.06.2006.
22. Rzeczpospolita. Złoty może dalej słabnąć. 27.06.2006.
23. Rzeczpospolita. Złoty w dół i obroty niewielkie. 16.08.2006.
24. Rzeczpospolita. Złoty bardzo spokojny. 8.12.2006.
25. Rzeczpospolita. Niewielkie osłabienie polskiej waluty. 9.05.2007.

Agata Sielska
Szkoła Główna Handlowa w Warszawie

O PEWNEJ PROPOZYCJI ZASTOSOWANIA RANKINGÓW WIELOKRYTERIALNYCH W ANALIZIE PORTFELOWEJ

Wprowadzenie

Wraz z postępującym rozwojem rynku kapitałowego zwiększająca się liczba spółek notowanych na giełdach wymaga konieczności opracowania odpowiednich narzędzi umożliwiających decydentom porównywanie ich akcji w celu zbudowania odpowiedniego portfela. Narzędzi tych dostarcza analiza portfelowa [5; 6; 14; 15; 18].

Problem konstrukcji portfela papierów wartościowych najczęściej formułowany jest jako zadanie optymalizacji wielokryterialnej, w którym decydent zainteresowany jest takim wyborem akcji, by zmaksymalizować pewien miernik zysku, minimalizując jednocześnie wartość osiąganą przez wybrany miernik ryzyka. Zazwyczaj, zgodnie z koncepcją Markowitza, jako miernikami tymi są wartość oczekiwana i odchylenie standardowe stopy zwrotu. W zależności od empirycznych rozkładów stóp zwrotu, wybierane są także inne mierniki, np. VaR [11]. Wskazuje się, że wybór odpowiednich mierników jest kluczowym elementem problemu budowy portfela, ponieważ ostateczny skład portfela jest różny dla różnych mierników [3], a ponadto, każdy z nich wnosi inne informacje o rozkładach stóp zwrotu. Dlatego zasadne wydaje się stosowanie wśród kryteriów wyboru więcej niż jednego miernika, co pozwoli decydentowi uwzględnić różne aspekty ryzyka.

Niniejsza praca jest propozycją zastosowania rankingów wielokryterialnych do wspomagania decyzji inwestycyjnych. Podejmując takie decyzje inwestorzy oceniają dostępne na rynku spółki ze względu na rozmaite kryteria, wśród których mogą się znaleźć zarówno mierniki ryzyka i zysku, jak i (w przypadku inwestycji krótkookresowych) pewne wskaźniki pochodzące z analizy technicznej generujące sygnał do zakupu. Można oczekiwać, iż w miarę zwiększania liczby kryteriów decyzyjnych stopy zwrotu z wybranych portfeli będą wyższe, ponieważ wybór dokonywany będzie na podstawie bar-

dziej kompletnego zespołu wzajemnie uzupełniających się cech. W szczególności, celem pracy jest sprawdzenie, czy zaproponowana metoda wykorzystania rankingów w konstrukcji portfela umożliwi osiągnięcie lepszego wyniku niż zastosowanie modelu Markowitza bądź wybrane benchmarki, a także czy wyniki będą zależne od użytej metody wielokryterialnej.

Praca składa się z trzech części. W pierwszej przedstawiono podstawy teoretyczne rozpatrywanego modelu. Następnie zaprezentowano wykorzystywane kryteria i omówiono metodę wyznaczenia ich wag. W ostatniej części zamieszczono i skomentowano wyniki badań eksperymentalnych.

1. Podstawy teoretyczne

1.1. Rankingi wielokryterialne

Rankingi wielokryterialne stanowią jeden z trzech podstawowych problemów rozwiązywania zadań wielokryterialnych [7]. Ich zastosowanie uznaje się za uzasadnione w przypadkach występowania kryteriów decyzyjnych niemających charakteru ilościowego, wyrażanych w różnych jednostkach, których nie można w naturalny zagregować lub też w sytuacji, w której nie jest w pełni jasny sposób, w jaki zysk na jednym z kryteriów może kompensować stratę na innych [8].

Zastosowane w niniejszej pracy metody reprezentują odmienne podejścia do rozwiązywania problemów hierarchizacji. Różnice występują także w złożoności algorytmów i możliwości dostosowania metody do potrzeb indywidualnego decydenta przez przyjęcie odpowiednich wartości parametrów. Wszystkie wykorzystane metody są jednak łatwe w interpretacji.

Metoda ELECTRE III [16] opiera się na porównywaniu wariantów decyzyjnych parami. Oprócz wag wykorzystywane są również progi weta oraz pseudokryteria, co umożliwia bogaty opis preferencji decydenta.

W metodzie TOPSIS [9] ranking wyznaczany jest na podstawie odległości względem odpowiednich punktów referencyjnych (idealnego oraz najgorszego rozwiązania rozpatrywanego problemu wielokryterialnego).

Zaletą AGREPREF [10] jest możliwość porównywania wariantów decyzyjnych parami oraz wprowadzenia progów preferencji i obojętności. Jako dodatkowy plus traktować można, podobnie jak w przypadku metody TOPSIS, intuicyjnie zrozumiały algorytm.

Metoda PROMETHEE II [1; 2] nie bazuje na przypisywaniu użyteczności dla każdego z możliwych wariantów ani kryteriów, tworząc ostateczny ranking alternatyw na podstawie porównywania ich parami. Co istotne, w algorytmie uwzględnia się odchylenia występujące pomiędzy ocenami dwóch róż-

nych wariantów ze względu na to samo kryterium zakładając, że im mniejsza różnica między wartościami danego kryterium dla poszczególnych wariantów, tym mniej preferowany jest dla decydenta najlepszy wariant z punktu widzenia tego kryterium. Im większe różnice, tym silniejsza preferencja. Należy również zwrócić uwagę na bogaty zbiór kryteriów uogólnionych, pozwalających elastycznie modelować preferencje decydenta [7]. W niniejszej pracy wykorzystano kryterium uogólnione Gaussa, gdyż nie wymaga ono przyjmowania dodatkowych założeń dotyczących preferencji decydenta.

Metoda WSA [12] może zostać uznana za jeden z najprostszych algorytmów hierarchizacji. Jej idea sprowadza się bowiem do zapewnienia porównywalności kryteriów poprzez standaryzację ich ocen, a następnie wykorzystania wag do agregacji oraz ustalenia ostatecznego rankingu.

1.2. Opis preferencji decydenta

W przypadku, w którym nie jest możliwe dokładne określenie preferencji decydenta, często przyjmuje się, że wszystkie kryteria są z jego punktu widzenia jednakowo istotne. Można jednak traktować wartości kryteriów jako źródła informacji, które powinny znajdować odzwierciedlenie w wartościach ich wag [20]. Podobnie można przyjąć, że jeśli dane kryterium w żaden sposób nie różnicuje porównywanych alternatyw, może zostać uznane za bezużyteczne dla decydenta i wyeliminowane z procesu decyzyjnego [4], gdyż nie wnosi do niego istotnych informacji.

Zgodnie z powyższymi rozważaniami, w całej niniejszej pracy utrzymywane będzie założenie, że preferencje decydenta są pochodną rzeczywistego zróżnicowania kryteriów między alternatywami. Wagi ustalono bazując na tym zróżnicowaniu, można je więc traktować jako opis preferencji decydentów opierających się na dostępnych i zrozumiałych danych dotyczących obiektywnych właściwości rozważanych alternatyw.

1.3. Model konstrukcji portfela

W pracy rozważać się będzie jednookresowy problem wyboru portfela papierów wartościowych. Zakłada się więc, że na początku okresu inwestycyjnego, w chwili t, inwestor dokonuje podziału posiadanego kapitału między wybrane instrumenty. Zbudowany w ten sposób portfel jest określony wielkościami udziałów poszczególnych walorów. Jego skład ani proporcje zaangażowanego kapitału nie ulegają zmianom w czasie okresu inwestycyjnego, a wartość całego portfela może się zmieniać jedynie za sprawą zmian kursów poszczególnych walorów. Decydent zainteresowany jest osiągnięciem maksymalnego zysku na koniec okresu inwestycyjnego przy minimalnym poziomie ryzyka.

Zgodnie z powyższym portfel inwestycyjny można zdefiniować jako

$$x = [x_1, ..., x_n]^T$$

gdzie:

n – ilość dostępnych na rynku walorów,
x_i – udział kapitału zainwestowanego w *i*-ty walor.

Wprowadzając zakaz krótkiej sprzedaży oraz warunek wykorzystania wszystkich posiadanych środków otrzymuje się wektor zmiennych decyzyjnych zdefiniowany wzorem

$$P = \left\{ x = [x_1, ..., x_n]^T ; \sum_{i=1}^{n} x_i = 1 \wedge \bigvee_{i \in \overline{1,n}} x_i \geq 0 \right\}$$

Zbiór P nazywany jest zbiorem portfeli dopuszczalnych. Z każdym wektorem należącym do tego zbioru związana jest pewna zmienna losowa R_x będąca stopą zwrotu z odpowiedniego portfela. Można pokazać, że R_x przyjmuje postać daną wzorem

$$R_x = [r_1, ..., r_n]^T [x_1, ..., x_n]$$

gdzie r_i – stopa zwrotu z waloru *i*.

W niniejszej pracy dopuszczalny zbiór papierów wartościowych ograniczono do tych, które w rankingu zostały sklasyfikowane na najwyższych pozycjach. Następnie na takim zbiorze rozwiązano zadanie Markowitza przy dodatkowych warunkach ograniczających wynikających bezpośrednio z postaci rankingu. Warunki te zapewniają, iż udział w portfelu spółki o wyższej pozycji nie może być niższy niż spółki o pozycji niższej. Jest to zgodne z intuicją, można bowiem uznać, że decydent nie powinien być skłonny zainwestować więcej w spółkę, którą uważa za gorszą niż w spółkę, która jest jego zdaniem lepsza.

Poszczególne kroki metody przedstawić można w następujących punktach:

1) wybór kryteriów,
2) ustalenie wag,
3) budowa rankingu,
4) budowa zbioru spółek dopuszczalnych (wszystkie spółki o lepszej pozycji niż pewna ustalona przez decydenta ranga graniczna),
5) ustalenie dodatkowych warunków ograniczających (spółka, która została oceniona w rankingu jako lepsza nie może mieć niższego udziału w portfelu niż spółka oceniona jako gorsza),

6) rozwiązanie zadania Markowitza na zbiorze wyznaczonym w 4) przy warunkach z 5).

Zaprezentowana metoda umożliwia zarówno uwzględnienia preferencji decydentów przy konstrukcji portfela i zmniejszenie kosztów transakcyjnych poprzez ograniczenie liczby spółek w portfelu. Wykorzystanie w tym celu rankingów wielokryterialnych wydaje się rozwiązaniem zgodnym z intuicją, ponieważ można uznać, że przed podjęciem decyzji inwestycyjnej decydent porównuje dostępne papiery wartościowe, starając się wybrać najlepsze z jego punktu widzenia.

Model Markowitza uzupełniono o dodatkowe warunki ograniczające wygenerowane na podstawie otrzymanych rankingów wielokryterialnych. Model (1) uwzględnia wyniki wcześniejszej analizy i hierarchizacji spółek według preferencji decydenta bądź zróżnicowania kryteriów.

$$\begin{aligned} & D^2(R_x) \to \min \\ & pw. \\ & \sum_{i=1}^{n} x_i = 1 \\ & \bigvee_{i \in \overline{1,n}} \quad x_i \geq 0 \\ & \bigvee_{j_1 < j_2,\ j_1, j_2 \in \overline{1,n^*},} \quad x_{j_1} \geq x_{j_2} \\ & \bigvee_{j > n^*,} \quad x_j = 0, \end{aligned} \tag{1}$$

gdzie:

$$D^2(R_x) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \operatorname{cov}(r_i, r_j),$$

j_1, j_2, j – kolejne rangi wyznaczone za pomocą metod wielokryterialnych,

n^* – ranga ostatniej spółki, w akcje której decydent skłonny jest zainwestować.

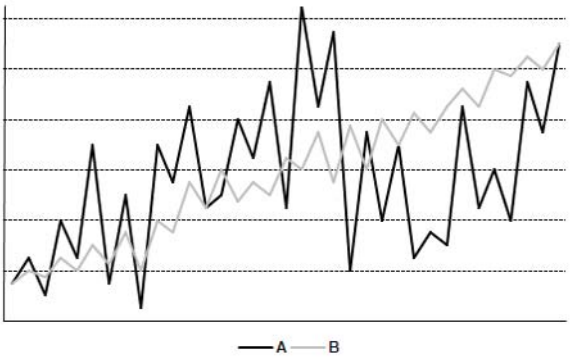
Wartość parametru n^* uzależniona jest od decydenta. Spółki o niższej randze nie są uwzględniane przy budowie portfela, można więc przyjąć, że jest to liczba niezdominowanych alternatyw. W niniejszej pracy wybrano jednak inne rozwiązanie. Opierając się na pracy M. Łuniewskiej [13], z których wynika, że na polskim rynku portfel zawierający od pięciu do dziesięciu spółek zapewnia wystarczający stopień dywersyfikacji, ustalono $n^* = 10$. Przy tak

postawionych warunkach ograniczających możliwy jest wybór do portfela mniejszej liczby spółek, jednakże nie pozwalają one na nadmierny wzrost kosztów transakcyjnych związanych z obrotem akcjami wielu różnych emitentów.

2. Kryteria decyzyjne

2.1. Wybór kryteriów

Jako pomocnicze kryteria zaproponowano zestaw mierników uwzględniający wyższe momenty rozkładów stóp zwrotu oraz mierniki oparte na kwantylach, odchylenie standardowe wolumenu, a także opisujące zmienność kursów akcji w obrębie ostatniego okresu obejmującego 5 lub 10 sesji przed podjęciem decyzji inwestycyjnej. Istotność ostatniej grupy mierników można zilustrować przedstawiając na rys. 1 przykładowe kursy akcji dwóch różnych spółek w ciągu okresu inwestycyjnego, tj. od momentu dokonania zakupu papieru wartościowego do momentu jego sprzedaży. Wydaje się oczywiste, że mimo iż stopa zwrotu z inwestycji w obu przypadkach jest taka sama, inwestor przejawiający awersję do ryzyka powinien preferować akcje spółki B, ponieważ w okresie inwestycji kształtują się one zgodnie z pewnym trendem, podczas gdy akcje spółki A ulegają gwałtownym zmianom.



Rys. 1. Przykład zmienności kursów akcji w ciągu okresu inwestycyjnego

Poszczególne kryteria to (kryteria K1-K4 dotyczą kursu akcji, K5 – wolumenu, K6-K12 stóp zwrotu):

K1 – średni kurs w ostatnim okresie (odpowiednio 5 lub 10 sesji) przed dokonaniem decyzji inwestycyjnej (wzór 2). Miernik informuje o zachowaniu się cen akcji.

$$\overline{y_{k-p}} = \frac{\sum_{i=t_p}^{t_k} y_i}{t_k - t_p} \tag{2}$$

gdzie:

y_i – kurs waloru w momencie *i*,

t_p – początek okresu,

t_k – koniec okresu.

Ponieważ uznano, że po wyborze portfela decydent nie może zmienić już swojej decyzji, jako kryterium przyjmuje się wartość tego miernika uzyskaną dla okresu inwestycyjnego bezpośrednio poprzedzającego moment zakupu papierów wartościowych.

K2 – semiodchylenie kursu w ostatnim okresie przed inwestycją

K3 – rozstęp kursu w ostatnim okresie przed inwestycją

$$R_{k-p} = \left| \max_{i \in t_p, t_k} (y_i) - \min_{i \in t_p, t_k} (y_i) \right|$$

K4 – różnica między kursem maksymalnym a kursem sprzedaży w ostatnim okresie przed inwestycją. Wzorem (3) zdefiniować można różnicę między maksymalnym kursem w okresie inwestycji, a kursem, po którym sprzedano dany papier wartościowy. Decydenta interesuje jak najmniejsza wartość poniższego wyrażenia.

$$RM_{k-p} = \max_{i \in t_p, t_k} (y_i) - (y_{t_k}) \tag{3}$$

K5 – odchylenie standardowe wolumenu,

K6 – skośność stóp zwrotu,

K7 – kurtozę stóp zwrotu,

K8 – semiodchylenie stóp zwrotu,

K9 – współczynnik zysku względnego,

K10 – CVaR przy poziomie istotności $\alpha = 0{,}05$,

K11 – VaR przy poziomie istotności $\alpha = 0{,}05$,

K12 – prawdopodobieństwo straty.

2.2. Wagi

Zgodnie z przyjętymi założeniami dotyczącymi preferencji decydenta, wagi kryteriów wyznaczono wykorzystując współczynnik zmienności wartości przyjmowanych przez kryteria (4)

$$v_i = \frac{s_i}{\bar{f}_i} \tag{4}$$

gdzie:

si – odchylenie standardowe wartości przyjmowanych przez kryterium f_i,

$\bar{f}_i$ – średnia arytmetyczna wartości przyjmowanych przez kryterium f_i,

$i = 1,\ldots,12$.

$$w_i^{\,1} = \frac{|v_i|}{\sum_{i=1}^{12} |v_i|} \tag{5}$$

Po unormowaniu zgodnie ze wzorem (5) większe wartości wag przypisano kryteriom bardziej zróżnicowanym między alternatywami.

W kolejnym kroku zbadano korelację między kryteriami, a następnie uznając, iż silnie skorelowane ze sobą zmienne przenoszą większość tych samych informacji, drugi zestaw wag wyznaczono korzystając z poniższego wzoru

$$w_i^{\,2} = \frac{\sum_{j=1}^{12} |r_{ij}|}{\sum_{i=1}^{12}\sum_{j=1}^{12} |r_{ij}|}$$

gdzie r_{ij} oznacza współczynnik korelacji kryteriów i oraz j.

Ostateczne, skorygowane wagi wyznaczono korzystając ze wzoru

$$w_i^{\,s} = \frac{\frac{w_i^{\,1}}{w_i^{\,2}}}{\sum_{i=1}^{12} \frac{w_i^{\,1}}{w_i^{\,2}}}$$

Postępując w opisany sposób uwzględniono zarówno zróżnicowanie kryteriów między cechami (im silniejsze, tym większa waga), jak i wzajemną korelację kryteriów (im silniejsza, tym niższa waga). Analizę przeprowadzono dla

dwóch różnych stóp zwrotu, by zweryfikować, czy długość okresu inwestycyjnego wpływa na strukturę wag. Uzyskane wartości wag zaprezentowano w tabelach 1 i 2.

Tabela 1

Wartości wag dla poszczególnych indeksów giełdowych dla 5-sesyjnych stóp zwrotu

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12
WIG20	0,032	0,034	0,034	0,085	0,080	0,515	0,088	0,016	0,062	0,039	0,010	0,004
mWIG40	0,053	0,043	0,041	0,062	0,090	0,081	0,083	0,038	0,404	0,055	0,045	0,005
sWIG80	0,056	0,157	0,164	0,102	0,121	0,094	0,080	0,051	0,047	0,086	0,036	0,006

Tabela 2

Wartości wag dla poszczególnych indeksów giełdowych dla 10-sesyjnych stóp zwrotu

	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12
WIG20	0,036	0,039	0,039	0,080	0,104	0,433	0,093	0,027	0,079	0,052	0,013	0,005
mWIG40	0,057	0,042	0,041	0,062	0,133	0,192	0,029	0,045	0,307	0,074	0,017	0,003
sWIG80	0,054	0,172	0,190	0,114	0,120	0,071	0,063	0,044	0,054	0,069	0,042	0,007

Na podstawie zamieszczonych wyników można wnioskować o zróżnicowaniu wybranych do analizy kryteriów. Można uznać, że kryteria posiadające największą wartość informacyjną (a zarazem mające najistotniejsze znaczenie) decydenta różnią się między indeksami. Jednak rozkład wag jest podobny w przypadku 5- i 10-sesyjnych stóp zwrotu.

3. Badanie eksperymentalne

3.1. Organizacja eksperymentu

Analizą objęto spółki tworzące indeksy WIG20, mWIG40 oraz sWIG80 w okresie od 02.01.2006 do 29.08.2008. W przypadku, w którym spółka nie była notowana przez cały ten okres, zostawała wykluczona z próby. By zniwelować wpływ przypadkowych wahań kursów na wyniki analizy, wykorzystano 5- i 10-sesyjne stopy zwrotu, otrzymując odpowiednio szeregi 133 i 66 obserwacji. Okres inwestycyjny obejmował następujące dni: od 01.09.2008 do 08.09.2008 oraz od 01.09.2008 do 15.09.2008 – odpowiednio dla inwestycji 5- i 10-sesyjnych. Poczynając od wyznaczenia wag portfele konstruowano osobno dla każdego z rozpatrywanych indeksów WIG20, mWIG40, sWIG80. Postępowanie takie zapewniało wstępną dywersyfikację spółek ze względu na sektory.

3.2. Weryfikacja normalności rozkładów stóp zwrotu

Hipotezę o normalności zweryfikowano przy użyciu testu Shapiro-Wilka [17]. Podstawowym kryterium przy wyborze testu była jego duża moc. W przypadku inwestycji w akcje nawet drobna asymetria rozkładu może wiązać się z ryzykiem istotnych strat materialnych. Z tego powodu uznano, że test o znacznej mocy może pozwolić decydentowi osiągnąć wymierne korzyści, jeżeli zastosowanie go wykaże zasadność poszerzenia szeregów kryteriów o cechy wiążące się z postacią rozkładu. W przypadku normalności rozkładu stóp zwrotu, parametry takie jak skośność bądź spłaszczenie można zaniedbać. W przeciwnym wypadku natomiast uwzględnienie takich informacji może pozytywnie wpłynąć na ograniczenie ryzyka.

Na podstawie wyników zaprezentowanych w tabeli 3 można uznać, że bez względu na przyjęty poziom istotności znaczna większość spółek nie podlega normalnemu rozkładowi. W konsekwencji w rozpatrywanej sytuacji założenia klasycznej analizy portfelowej nie są spełnione, a mierniki użyte w modelach Sharpe'a i Markowitza nie ograniczają ryzyka w wystarczającym stopniu. Uwzględnienie dodatkowych kryteriów przy konstrukcji portfeli jest więc uzasadnione.

Tabela 3

Wyniki testu Shapiro-Wilka dla rozkładów stóp zwrotu

Poziom istotności	Procent przypadków odrzucenia hipotezy zerowej	
	5-sesyjne stopy zwrotu	10-sesyjne stopy zwrotu
α=0,01	72,86%	66,43%
α=0,05	80,00%	71,43%
α=0,1	82,14%	74,29%

3.3. Benchmarki

Skuteczność skonstruowanych modeli oceniono poprzez porównanie z zestawem benchmarków. Dla potrzeb niniejszej pracy do oceny efektywności inwestycji krótkoterminowych wybrano następujące punkty odniesienia:

- Indeks – stopa zwrotu z odpowiedniego indeksu w okresie inwestycyjnym,
- Profetyczny – stopa zwrotu z portfela konstruowanego przy założeniu znajomości rzeczywistych przyszłych stóp zwrotu; portfel taki gwarantuje maksymalną możliwą do osiągnięcia w danym okresie stopę zwrotu. Inwes-

tując w ten sposób decydent mający wiedzę dotyczącą przyszłego zachowania się stóp zwrotu zainwestuje wszystkie posiadane środki tylko w jeden instrument – ten, którego przyszła stopa zwrotu rzeczywiście będzie najwyższa,
- Minimalny – minimalna stopa zwrotu możliwa do osiągnięcia w danym okresie,
- Markowitz – stopa zwrotu z portfela skonstruowanego przy takich samych założeniach, ale bez uwzględniania dodatkowych kryteriów,
- Sharpe – stopa zwrotu z portfela skonstruowanego przy wykorzystaniu wskaźnika Sharpe'a,
- SS – stopa zwrotu z portfela skonstruowanego przy optymalizacji współczynnika Sortino-Satchella [19].

4.4. Analiza inwestycji

5- i 10-sesyjne stopy zwrotu osiągnięte dla portfeli i benchmarków zamieszczono w tabelach 4 i 5. Dla inwestycji krótkoterminowych stopy zwrotu z portfeli skonstruowanych przy użyciu rankingów są wysokie. Porównanie z benchmarkami w większości przypadków wypada na ich korzyść, przy czym warte odnotowania są wyniki osiągnięte dla indeksu mWIG40 dla 5-sesyjnej stopy zwrotu. W tym przypadku portfele wykorzystujące dodatkowe warunki ograniczające są gorsze jedynie od portfela profetycznego.

Tabela 4

Realizacje 5-sesyjnych stóp zwrotu

	WIG20	mWIG40	sWIG80
Markowitz	1,27%	5,20%	5,59%
Sharpe	-0,01%	5,21%	2,27%
SS	-4,49%	3,32%	1,42%
Profetyczny	14,29%	34,11%	22,22%
Minimalny	-10,52%	-50,00%	-18,47%
Indeks	0,08%	6,36%	2,01%
PROMETHEE II	4,40%	14,04%	-16,60%
ELECTRE III	3,80%	7,03%	0,33%
WSA	4,52%	11,47%	-14,40%
TOPSIS	4,57%	10,92%	-2,30%
AGREPREF	2,22%	9,90%	2,07%

Tabela 5

Realizacje 10-sesyjnych stóp zwrotu

	WIG20	mWIG40	sWIG80
Markowitz	-5,88%	1,23%	-0,90%
Sharpe	-9,90%	6,31%	-4,19%
SS	-13,15%	1,07%	-11,06%
Profetyczny	2,44%	28,75%	15,29%
Minimalny	-16,11%	-84,19%	-87,71%
Indeks	-7,64%	-1,66%	-4,54%
PROMETHEE II	-6,15%	0,35%	-16,31%
ELECTRE III	-3,83%	-1,77%	-13,56%
WSA	-3,73%	0,69%	-15,44%
TOPSIS	-6,29%	12,56%	-16,31%
AGREPREF	-2,85%	4,41%	3,88%

W tabeli 6 zamieszczono porównanie efektywności portfeli zbudowanych przy wykorzystaniu rankingów otrzymanych różnymi metodami. Wartości procentowe informują w ilu przypadkach modele z dodatkowymi warunkami ograniczającymi przyniosły stopę zwrotu wyższą od benchmarków. W porównaniu nie uwzględniono portfeli profetycznego i minimalnego.

Tabela 6

Porównanie efektywności metod

	Markowitz	Sharpe	SS	Indeks	Średnia efektywność modelu
PROMETHEE II	33%	50%	50%	67%	50%
ELECTRE III	50%	50%	50%	50%	50%
WSA	50%	50%	50%	67%	54%
TOPSIS	50%	67%	67%	67%	63%
AGREPREF	83%	67%	100%	100%	88%
Średnia efektywność względem benchmarków	58%	58%	67%	72%	64%

Modele wykorzystujące dodatkowe warunki ograniczające okazały się lepsze od klasycznego modelu Markowitza w 58% przypadków, natomiast od indeksu już w 72%. Za średnio najbardziej efektywną uznać można metodę AGREPREF – w 88% przypadków stopa zwrotu z portfela zbudowanego na jej podstawie była wyższa od benchmarków. Jedynymi portfelami, których stopa zwrotu była wyższa były w jej przypadku portfele zbudowany przy zastosowaniu modelu Markowitza i Sharpe'a.

Wyniki uzyskane dla pozostałych metod są już zauważalnie gorsze. PROMETHEE II i ELECTRE III dają relatywnie najgorsze rezultaty. Porównując je z rezultatami otrzymanymi dla metody WSA można uznać, że złożoność obliczeniowa nie gwarantuje uzyskania dobrego wyniku. W 64% przypadków osiągnięto wynik lepszy od benchmarku.

Podsumowanie

W związku z asymetrią rozkładów stóp zwrotu, dokonując selekcji akcji decydent powinien uwzględniać szereg uzupełniających się miar ryzyka. Budowa portfela jest więc zagadnieniem wielokryterialnym, natomiast wartości poszczególnych miar przyjmowane dla różnych spółek traktować można jako wartości kryteriów. W pracy zaproponowano metodę konstruowania portfeli inwestycyjnych na podstawie rankingi wielokryterialnych. Założono, że wiedza decydenta o danych papierach wartościowych jest tym większa, im bardziej kompletny zestaw kryteriów bierze on pod uwagę. Ponadto, ze względu na dużą liczbę kryteriów przyjęto, że preferencje uzależnione są w pewien określony sposób od rzeczywistego zróżnicowania i korelacji między kryteriami. Dysponując wyznaczonymi na tej podstawie zestawami wag można rozwiązać zadanie wielokryterialne, a wynik otrzymany w ten sposób wydaje się mieć szczególne znaczenie ze względu na wymierny charakter możliwych zysków (strat) w problemach inwestycyjnych.

Dla każdego z trzech rozważanych indeksów giełdowych zbudowano rankingi spółek, a następnie na zbiorze ograniczonym dodatkowymi warunkami wynikającymi z postaci rankingów wyznaczono portfel z minimalną wariancją. Porównanie skonstruowanych w ten sposób portfeli z benchmarkami w przypadku inwestycji krótkookresowych (5- i 10-dniowych) w większości przypadków wypadło na korzyść portfeli opierających się na rankingach.

Otwartą kwestią pozostaje uniwersalność otrzymanych wyników, ich zależność od przyjętych założeń dotyczących preferencji decydenta oraz wybranych kryteriów. Wykorzystany w pracy zestaw jest jedynie propozycją. Aby zwiększyć szansę osiągnięcia zysku, zbiór kryteriów powinien spełniać pewne określone warunki, wśród których szczególną uwagę należy zwrócić na jego kompletność, różnorodność i niepowielanie tych samych informacji. Jednak nie każdy decydent ma wystarczającą wiedzę na temat miar ryzyka i rynku kapitałowego, by zbudować odpowiedni zbiór. Wykorzystywane przez niego kryteria pozostają w dalszym ciągu pochodną jego preferencji, jednak można spodziewać się, że, poprzez nieuwzględnienie pewnych kluczowych kryteriów, skuteczność metody może okazać się niższa.

W dalszej kolejności należy rozważyć zależność rezultatów od wybranych wag, które również odzwierciedlają preferencje decydenta. W pracy obliczono je wykorzystując współczynniki zmienności oraz korelacji, aby w możliwie pełny sposób uwzględnić współzależność kryteriów. Należy jednak zdawać sobie sprawę, że zaproponowany sposób ważenia kryteriów jest jedynie jedną z wielu dostępnych możliwości. Weryfikacja efektywności metod przy innych przekracza ramy tej pracy.

Przy przyjętych w pracy założeniach, zaproponowana metoda konstrukcji portfeli akcji okazała się skuteczna.

Literatura

1. Brans J.P., Vincke P. (1985). A Preference Ranking Organization Method. Management Science, 31, 647-656.
2. Brans J.P., Mareschal B., Vincke P. (1986). How to Select and How to Rank Projects : The PROMETHEE method for MCDM. European Journal of Operational Research, 24, 228-238.
3. Czupryna M. (2003). O zastosowaniu metod interaktywnych w problemie doboru optymalnego portfela inwestycyjnego. W: Modelowanie preferencji a ryzyko ’03. Red. T. Trzaskalik. AE, Katowice.
4. Diakoulaki D., Mavrotas G., Papayannakis L. (1995). Determining Objective Weights in Multiple Criteria Problems: the CRITIC Method. Computers and Operations Research, 22, 7, 763-770.
5. Dorosiewicz S. (2003). Elementy analizy portfelowej statyka – ujęcie matematyczne. SGH, Warszawa.
6. Elton E.J., Gruber M.J. (1981). Modern Portfolio Theory and Investment Analysis. John Wiley & Sons, New York.
7. Multiple Criteria Decision Analysis. State of the Art Surveys. (2005). Red. J. Figueira, S. Greco, M. Ehrgott. Springer.
8. Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory and Applications. (1999). Red. T. Gal, T.J. Stewart, T. Hanne. Kluwer Academic Publishers, Boston.
9. Hradílek Z., Juřica L., Gurecký J., Krejčí P. (2006). New Method Agrepref for the Priority Location of Remote Controlled Disconnectors in the Distribution Network. Proceedings of Power, Energy and Applications, 506(7)-506(14).
10. Hwang C.L., Yoon K. (1981). Multiple Attribute Decision Making: Methods and Applications. Springer, New York.
11. Jorion P. (2006). Value AT Risk: the New Benchmark for Managing Financial Risk. McGraw-Hill Companies.

12. Keeney R.L., Raiffa H. (1976). Decisions with Muliple Objectives. Wiley.
13. Łuniewska M. (2003). Wykorzystanie metod ilościowych do tworzenia portfela papierów wartościowych. Uniwersytet Szczeciński, Szczecin.
14. Markowitz H.M. (1952). Portfolio selection. Journal of Finance, 7.
15. Markowitz H.M. (1998). Portfolio Selection: Efficient Diversification of Investments. Blackwell, Cambridge.
16. Roy B. (1978). ELECTRE III: Un algorithme de classement fondé sur une représentation floue des préférences en présence de critčres multiples. Cahiers du Centre d'Etudes en Recherche Opérationnelle. 20(1), 3-24.
17. Shapiro S.S., Wilk M.B. (1965), An Analysis of Variance Test for Normality (Complete Samples). Biometrika, 52, 3, 4, 591-611.
18. Sharpe W.F., Alexander G.J. (1990). Investments. Prentice Hall, Englewood Cliffs New Jersey.
19. Sortino F.A. (2000). Upside-Potential Ratios Vary by Investment Style. Pensions and Investments 28, 30-35.
20. Zeleny M. (1982). Multiple Criteria Decision Making. McGraw-Hill, New York.

Sławomir Śmiech
Uniwersytet Ekonomiczny w Krakowie

BUDOWA PORTFELI EFEKTYWNYCH NA PODSTAWIE NARZĘDZI KOINTEGRACJI

Wprowadzenie

Budując portfele inwestycyjne na podstawie klasycznej analizy korelacji (dokładniej – macierzy kowariancji stóp zwrotu) ignoruje się część dostępnej informacji na temat walorów. Dzieje się tak dlatego, że wyznaczając stopy zwrotu eliminowane zostają trendy. W efekcie konstruowane portfele, które mają być co najmniej tak dobre jak rynkowe benchmarki (indeksy, które są punktami odniesienia), mogą z upływem czasu w znaczny stopniu od nich odbiegać. Konieczne staje się w takiej sytuacji nieustanne korygowanie składu portfela. Rozwiązaniem tego problemu może być wykorzystanie długookresowego związku pomiędzy cenami walorów oraz wartością indeksu. Terasvirta i Grenger [3] pokazali, że modele z długą pamięcią są adekwatne do cen akcji. Stąd techniki kointegracji mogą być wykorzystane w celu budowy portfeli w trwały sposób związanych w relacji długookresowej z wybranym benchmarkiem. Po raz pierwszy tego typu podejście zostało zaproponowane przez Lucasa [4] oraz Alexander [1].

Niniejsza praca będzie składać się z trzech części. W pierwszej zaprezentowana zostanie idea kointegracji i możliwość wykorzystania jej w kontekście budowy portfela znajdującego się w długookresowej równowadze z benchmarkiem. Zaproponowana zostanie strategia wykorzystania własności oscylowania wartości portfela inwestycyjnego wokół wartości wyznaczonych przez inwestycję referencyjną. W części empirycznej z niewielkiej liczby akcji zbudowany zostanie portfel związany w długookresowej relacji z indeksem WIG20. Oceniona zostanie także efektywność zaproponowanej wcześniej strategii. Ostatni punkt pracy to wnioski z przeprowadzonego badania.

1. Kointegracja szeregów czasowych

Pojęcie kointegracji wprowadzili do literatury Engle i Granger [2]. Zauważyli, że pomiędzy niektórymi szeregami ekonomicznymi można wyznaczyć długookresową ścieżkę równowagi, która nie zależy od czasu. Ewentualne odchylenia od tej ścieżki mają charakter procesów stacjonarnych a zatem powracających do średniej.

Formalnie kointegrację można definiować dla par procesów, a także dla wektor procesów. Ogólne sformułowanie można znaleźć np. [5]. Tutaj przedstawiona będzie definicja [7, s. 431], która jest wystarczająca w kontekście celów pracy. Niech $Y_t = (y_{1t},....,y_{nt})'$ będzie wektorem wymiaru $n\ x\ 1$ składającym się ze zintegrowanych w stopniu 1 (I(1)) procesów stochastycznych. Mówimy, że proces Y_t jest skointegrowany, jeśli istnieje wektor $\beta = (\beta_1,....,\beta_n)$ t, że

$$\beta' Y_t = \beta_1 y_{i1t} + ... + \beta_n y_{nt} \sim I(0) \tag{1}$$

Wektor niestacjonarnych procesów jest skontegrowany, jeśli istnieje stacjonarna kombinacja liniowa jego współrzędnych. Wektor β jest nazywany wektorem kointegrującym.

W wypadku, gdy niektóre ze współrzędnych wektora $\beta = (\beta_1,....,\beta_n)$ są zerami wtedy powiemy, że (odpowiedni) podzbiór wektora *Y* jest skointegrowany.

Liniowa kombinacja, występująca w równaniu (1), przedstawia długookresowy związek współrzędnych wektora *Y* (reprezentuje długookresową równowagą procesów). Ponieważ jest ona stacjonarna, realizacje procesów nie powinny w długim okresie odchylać się od jej średniej.

Wektor kointegrujący nie jest wyznaczony jednoznacznie. Dlatego wprowadza się pewne normalizacje, aby określić go w sposób jednoznaczny. Często normalizacja przebiega tak, aby wektor kointegrujący przybrał postać

$$\beta = (1,-\beta_2,....,-\beta_n)'$$

Wtedy długookresowa równowaga może być przedstawiona w postaci

$$\beta' Y_t = y_{i1t} - \beta_2 y_{2t} ... - \beta_n y_{nt} \sim I(0) \tag{2}$$

lub alternatywnie

$$y_{i1t} = \beta_2 y_{2t} + ... + \beta_n y_{nt} + u_t, \tag{3}$$

gdzie $u_t \sim I(0)$.

Ocena czy *n*-wymiarowy proces jest skointegrowany może przebiegać w oparciu o procedurę Engela-Grangera lub procedurę Johansena. Preferowana w wypadku większej liczby procesów (gdy $n > 2$) jest procedura Johansena. Procedura Engela-Grengera wymaga określenia jednego z procesów jako zmiennej zależnej. Analiza prowadzona w niniejszej pracy spełnia ten wymóg i dlatego tylko ta procedura zostanie przytoczona. Sprowadza się ona do następujących punktów:

1. Testowania stopnia integracji wszystkich składowych wektora *Y*. Stosować tutaj można takie testy, jak test Dickeya-Fullera (DF), rozszerzony test Dickeya-Fullera (ADF), test Kwiatkowskiego KPSS, test Philipsa-Perrona (PP).

2. Jeśli wszystkie składowe wektora *Y* okażą się zintegrowane (tutaj w stopniu 1), szacowane są (metodą najmniejszych kwadratów) parametry równania (3). Oceniana jest istotność ocen parametrów równania.

3. Ocenia się stacjonarność reszt równania (3). Należy przy tym pamiętać, że testy pierwiastków jednostkowych stosowane dla reszt nie mają typowych rozkładów ale asymptotycznie zmierzają do dystrybuanty Philipsa-Ouliarisa. Odpowiednie wartości krytyczne można znaleźć w [6].

2. Idea budowy portfela na podstawie kointegracji cen akcji oraz indeksu giełdowego

Tradycyjna teoria finansów (zbudowana na relacjach macierzy kowariancji i oczekiwanych stopach zwrotu) mówi, że jedynym portfelem efektywnym, zbudowanym wyłącznie z instrumentów ryzykownych (z akcji), jest portfel rynkowy. Stąd jedyną słuszną strategią pasywnego inwestowania jest próba konstrukcji portfela, który możliwie najbardziej swoim składem przypomina portfel rynkowy. Możliwe są tutaj dwa podejścia. Pierwsze polega na zbudowaniu portfela inwestycyjnego, w którym udziały odpowiadają udziałom poszczególnych akcji w indeksie giełdowym. Ze względu na dużą liczbę walorów na rynku i brak możliwości kupna ułamków akcji, odtwarzanie portfela rynkowego przez indywidualnego inwestora jest bardzo trudne. Drugie podejście, które jest pewnym uproszczeniem pierwszego, zakłada, że budowany portfel ma być maksymalnie skorelowany z indeksem rynkowym. W jego skład nie muszą już wchodzić wszystkie walory. Ograniczeniem tego podejścia okazuje się nieustanna zmiana relacji (zależności) pomiędzy stopami zwrotu (zmienności macierzy korelacji). W konsekwencji utrzymywanie portfela skorelowanego z indeksem rynkowym wymaga nieustannej korekty jego składu, która powoduje wzrost kosztów transakcyjnych. Rozwiązaniem wydającym się

wolnym od powyższych ograniczeń jest zbudowanie portfela, którego wartość będzie pozostawać w długookresowej równowadze z indeksem rynkowym. Jeśli możliwe jest skonstruowanie takiego portfela z niewielkiej liczby walorów, będzie on atrakcyjną inwestycją, której efektywność będzie dorównywać inwestycji w portfel rynkowy.

Zgodnie z procedurą Engela-Grangera, budowany portfel (zwany dalej portfelem skointegrowanym) powinien być określony przez parametry β_i, który spełniają równanie

$$\ln y_{1t} = c + \sum_{i=2}^{n} \beta_2 \ln y_{2t} + u_t \tag{4}$$

gdzie: $\ln y_{1t}$ to logarytm naturalny indeksu rynkowego, $\ln y_{it}$ dla i=2, …, n to logarytmy cen akcji, z których zbudowano portfel, c, β_i – parametry modelu, u_t to stacjonarne reszty.

W równaniu (4) zamiast logarytmów można wykorzystywać wartości (ceny) instrumentów finansowych. Użycie logarytmów powoduje, że ewentualne wzięcie przyrostów, sprowadzi analizę do poziomu stóp zwrotu i umożliwi dalszą interpretację (tutaj będzie ona pominięta).

Dodatkową użyteczną własnością modelu (4) jest stacjonarność procesu resztowego u_t. Dzięki niej wiemy, że odchylenia od długookresowej równowagi są „chwilowe". Jeśli dodatkowo wariancja procesu resztowego jest wystarczająco duża, a przy tym odpowiednio szybko proces wraca do średniej, pojawia się możliwość zastosowania strategii inwestycyjnej, która ma szansę być efektywniejsza niż pasywna strategia odtwarzania portfela rynkowego. Strategia polega na stosowaniu dwóch kolejnych kroków. Pierwszy krok jest realizowany, gdy występuje ujemne[*] (i duże) odchylenie wartości portfela (skointegrowanego) od benchmarku. Portfel skointegrowany ma wtedy mniejszą wartość niż portfel referencyjny. Można uznać, że jest on niedoszacowany w stosunku do benchmarku. Ponieważ oba instrumenty w dłuższym okresie pozostają w równowadze można się spodziewać, że nastąpi szybki „wzrost" wartości portfela skoineterowanego tak, aby powrócił on do poziomu średniego. Skoro tak, spodziewane stopy zwrotu w krótkim okresie są wyższe dla portfela skointegrowanego niż dla indeksu rynkowego. Drugi krok – sprzedaż portfela skointegrowanego następuje wtedy, gdy wartość reszt równania 4 jest dodatnia (i duża). Uzyskane ze sprzedaży środki powinny być wykorzystane do zakupu portfela rynkowego, który jest w tym momencie niedowartościowany w stosunku do portfela skointegrowanego. Jeśli (np. z wymienionych wyżej przy-

[*] Ujemna wartość reszt u_t.

czyn) nie można skonstruować portfela „odtwarzającego" indeks rynkowy można rozważyć zakup instrumentów pozbawionych ryzyka (bony). W przypadku, gdy stopy zwrotu indeksu i bonów będą w odpowiedniej relacji inwestycja może okazać się opłacalna. Kroki 1 i 2 powinny po sobie następować tak długo, jak to jest możliwe. Wybór okresów, w których powinno się realizować kroki strategii, czyli określenie, kiedy uznać, że reszty równania (4) są już odpowiednio „duże", zależy od własności procesu reszt. Można określić funkcję mierzącą ewentualne nadwyżki ze strategii, w zależności od poziomów reszt, przy których strategia jest realizowana. Sprzedając portfel w tym momencie unikamy „straty" związanej z powrotem wartości portfela do poziomu równowagi. Przedstawiona strategia powinna uwzględnić koszty transakcyjne. Dlatego sensowność jej stosowania jest uzasadniana dużą wariancją reszt równania kointegracyjnego, szybkim powrotem procesu reszt do zera oraz stosunkowo dużymi (w porównaniu do benchmarku) stopami zwrotu z bonów skarbowych.

3. Opis zbioru danych. Budowa portfela

Ocena możliwości konstruowania portfela na podstawie techniki kointegracji została przeprowadzana dla akcji wchodzących w skład indeksu WIG20. Jednocześnie benchmarkiem (naturalnym) dla budowanego portfela będzie indeks 20 największych spółek notowanych na Warszawskiej Giełdzie. Zakres czasowy analizy obejmował okres od 9 czerwca 2005 do 5 lutego 2009. Wartości indeksu oraz ceny walorów zostały zlogarytmowane. Analiza była prowadzona na kursach zamknięcia dla danych dziennych. Celem prowadzonej analizy ma być zweryfikowanie następujących hipotez:

1. Istnieje możliwość zbudowania portfela, składającego się z niewielkiej liczby akcji, który w długim okresie daje porównywalny z rynkiem dochód.

2. Skład zbudowanego portfela jest stabilny. Nie ma potrzeby częstego korygowania jego składu.

3. Zaproponowana strategia inwestowania, pozwala poprawić wyniki uzyskane za pomocą trzymania portfela kointegracyjnego.

Potwierdzenie pierwszej tezy wymaga przeszukania możliwych kombinacji doboru akcji do portfela. Teoretycznie, gdyby założyć, że portfel ma składać się z nie więcej niż 5 aktywów (z 20 wchodzących w skład indeksu WIG20), można by zaproponować ponad 21,5 tys. możliwych kombinacji akcji w portfelu. Aby ograniczyć tę liczbę zdecydowano, że o udziale danej akcji w portfelu decydować będzie zespół ograniczeń, na który składają się warunki: dostatecznie długi okres notowania spółki (skład indeksy WIG20 zmieniał się w okresie prowadzenia badań), silny związek korelacyjny pomiędzy wartoś-

ciami spółki a indeksem WIG20. W dalszej kolejności brano pod uwagę, czy istotne są oszacowane wartości parametrów w równaniu kointegracyjnym, czy znak parametru przy zmiennej jest dodatni (założono, że w warunkach w jakich działa warszawska giełda, trudno realizować krótką sprzedaż), analizowano własności reszt równania kointegracyjnego. Oceniano też, czy logarytmy cen akcji (oraz logarytm wartości WIG20) są zintegrowane w stopniu pierwszy. Ostatecznie zdecydowano, że portfel kointegracyjny, który ma być alternatywną inwestycją do inwestycji w WIG20, będzie się składał z akcji spółek PEKAO, KGHM, MOL, TVN[†].

Wartości testów ADF dla spółek, które zostały włączone do ostatecznego równania regresji zostały zestawione w tabeli 1. Przeprowadzone testy nie dają podstaw do odrzucenia hipotezy zerowej, mówiącej o występowaniu pierwiastka jednostkowego w procesach (logarytmów) cen. Postać testów ADF w każdym przypadku dopuszczała wystąpienie trendu deterministycznego, co jest typowym założeniem w przypadku analizowania zachowania cen akcji [7, s. 114].

Tabela 1

Wartości statystyki *t* dla testów *ADF* oraz wartość *p* obliczone dla (logarytmów) cen spółek wchodzących w skład portfela inwestycyjnego oraz dla wartości WIG20

	PEKAO	KGHM	MOL	TVN	WIG 20
ADF	-0.2492	-0.9985	-0.8498	-0.2374	0.1094
p-value	0.9919	0.9423	0.9594	0.9922	0.9974

Oszacowane parametry równania zostały zestawione w tabeli 2. Wszystkie współczynniki są statystycznie istotne. Największy udział[‡] w portfelu mają akcje PEKAO, najmniejszy KGHM. Pozostaje sprawdzić własności reszty równania kointeracyjnego. Wykresy szeregu reszt oraz funkcji autokorelacji i autokorelacji cząstkowej zostały zestawione na rys. 1. Szereg reszt charakteryzuje się własnością powrotu do średniej (często przecina zero). Autokorelacja maleje stosunkowo szybko, co sugeruje, że proces nie ma długiej pamięci. Wizualna analiza pozwala przyjąć wstępnie, że reszty są stacjonarne (I(0)). Formalnie ocena stacjonarności została przeprowadzona za pomocą testów ADF oraz PP.

[†] Nie twierdzi się, że jest to jedyny ani najlepszy możliwy wybór.
[‡] Udział walorów w portfelu otrzymamy po znormalizowaniu wartości parametrów tak, aby ich suma wynosiła jeden.

Tabela 2

Oszacowane wartości współczynników równania kointegracyjnego wraz ze stosownymi statystykami

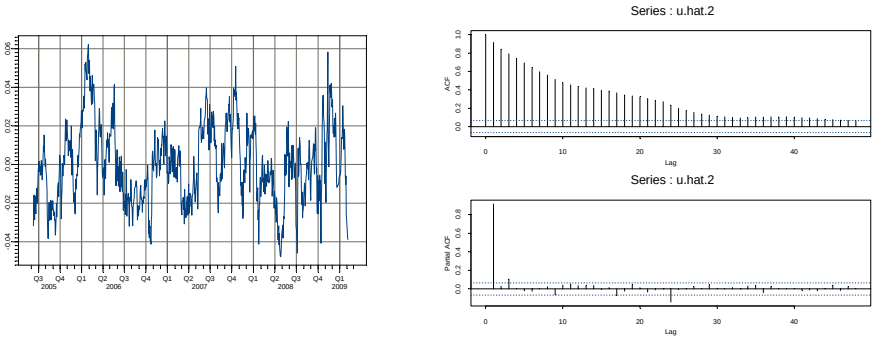
	Współczynnik	Błąd stand.	t-Student	p-value
Stała	3,6695	0,0211	173,5	0
PEKAO	0,4470	0,0098	45,36	3,28E-234
KGHM	0,0874	0,0044	19,52	5,93E-071
MOL	0,1868	0,0072	25,95	3,93E-111
TVN	0,1654	0,0081	20,33	8,26E-076

Tabela 3

Wartości statystyk *ADF, PP* oraz wartości *p* dla reszt równania kointegracyjnego

	ADF	PP
Wart. statystyki	-4.549	-5.9187
PO – p-value	0.0365	0.0003

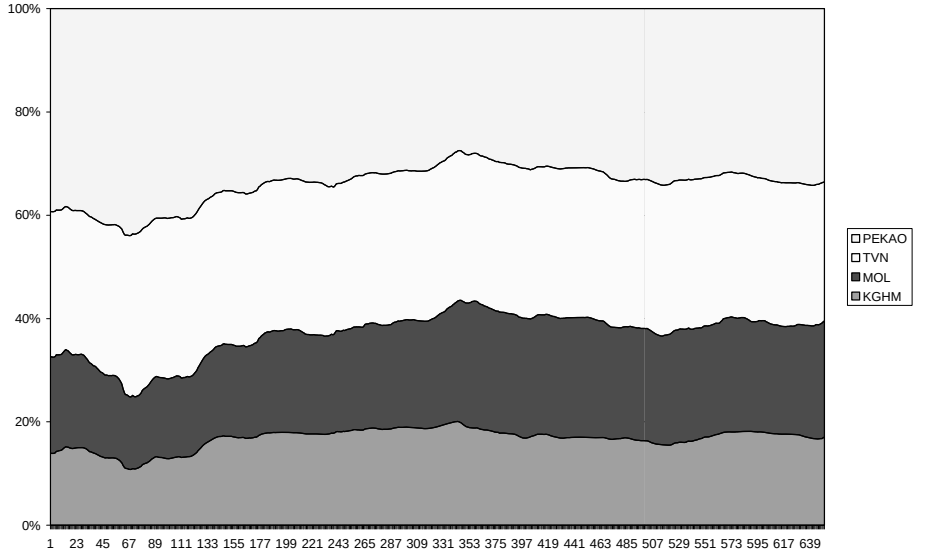
Statystyki z próby wraz z wyznaczoną na podstawie tablic rozkładu Philipsa-Ouliarisa wartością *p* zostały przedstawione w tabeli 3. Otrzymane wyniki pozwalają odrzucić hipotezę mówiącą, że proces reszt posiada pierwiastek jednostkowy. Można przyjąć, że wybrany podzbiór akcji jest skointegrowany z indeksem WIG20.



Rys. 1. Reszt równania kointegrującego oraz funkcje autokorelacji i autokorelacji cząstkowej

Weryfikacja stabilności składu portfela kointegracyjnego została przeprowadzona przez oszacowanie ciągu równań regresji. Zmienną zależną był podzbiór wartości indeksu WIG20, zaś zmienne niezależne tworzyły odpowiednie podzbiory cen akcji analizowanych spółek. Okno próby, na podstawie której szacowano parametry równań w każdym kroku było takie samo i wynosiło 250 obserwacji. W pierwszym kroku analizowano wartości zaobserwowane w okre-

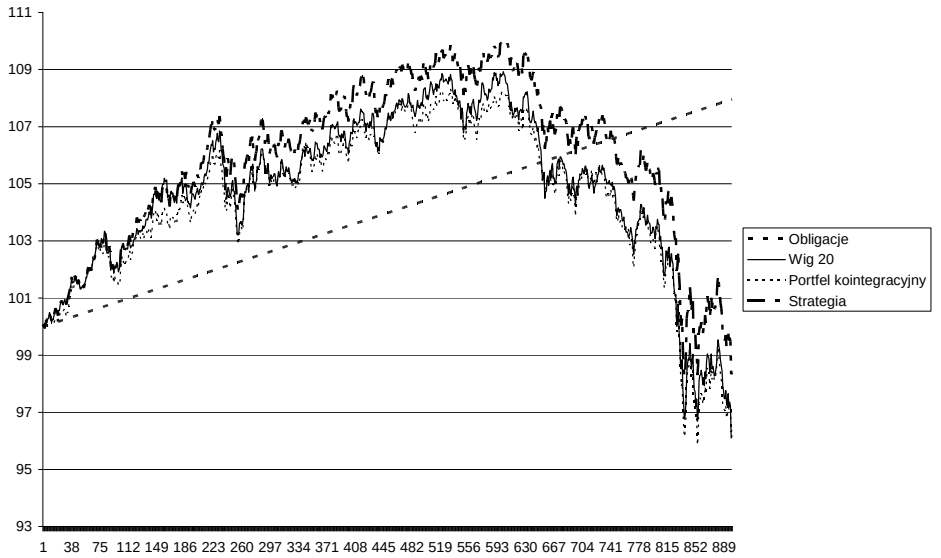
sie od 5 czerwca 2005 do 6 czerwca 2006. W kolejnych krokach okno to przesuwano o jedną obserwację w przód (o jeden dzień). W ten sposób oszacowano 651 równań. Znormalizowane wartości ich parametrów (udziały portfeli dla zadanego okresu) zostały pokazane na rys. 2. W początkowym okresie udział w portfelu akcji PEKAO był największy – przekraczał 40%. Później skład portfeli ustabilizował się. Pomimo że nie przeprowadzono formalnej oceny stabilności składu można uznać, że jego korygowanie nie było konieczne.



Rys. 2. Udziałów akcji w dynamicznie budowanych portfelach kointegracyjnych

Ostatni krok analizy polegał na porównaniu wartości portfeli, które zostały zbudowane na podstawie przedstawionych założeń. Zestawiono tutaj 4 teoretyczne portfele. Pierwszy to portfel replikujący WIG20[§], drugi to portfel kointegracyjny, trzeci portfel odpowiada przedstawionej w części 2 strategii, ostatni to inwestycja w bezpieczne bony skarbowe. Założono, że wartość wszystkich portfeli na początku jest równa 100 jednostek. W kolejnych dniach wartość portfeli zmienia się o wartość odpowiednich dla każdego portfela stóp zwrotu.

[§] Mający takie same udziały akcji jakie posiada indeks Wig20.



Rys. 3. Wartości porównywanych portfeli

Na rys. 3 pokazano jak kształtują się wartości poszczególnych portfeli. Wartości portfela WIG20 oraz portfela kointegracyjnego podążają po wspólnej trajektorii. W badanym okresie występują jedynie niewielkie odchylenia wartości na rzecz jednej czy drugiej inwestycji. Inaczej zachowuje się portfel strategia, który bardzo szybko osiąga wartość przewyższającą wartość pozostałych portfeli. Trzeba wyjaśnić, że przedstawione wartości zostały uzyskane, gdy kolejne kroki inwestycji były podejmowane dla reszt większych od 2 oraz mniejszych od -2. W całym badanym okresie nastąpiło 24-krotne zastosowanie strategii (zamiana jednego portfela na drugi). Oznacza to, że nawet uwzględniając koszty transakcyjne przedstawiona strategia dała znacząco lepsze efekty niż portfel WIG20.

Podsumowanie

Przeprowadzone badanie pokazało, że można z zaledwie kilku akcji zbudować portfel, który w długim okresie (niemal 4 lat) będzie zachowywał się podobnie jak indeks 20 największych spółek na GPW. Okazało się, że zaproponowany portfel ma stabilną strukturę. Udziały poszczególnych walorów nie zmieniały się w znaczący sposób. Zaproponowana strategia wykorzystująca własności relacji między zbudowanym portfelem oraz benchmarkiem przyniosła ponadprzeciętne zyski. Okazało się więc, że wykorzystując dość prosty mechanizm inwestowania można w budować portfele, które w długim okresie są bardziej efektywne od rynku.

W pracy przyjęto, że portfelem do którego porównuje się efektywność inwestycji jest portfel WIG20. Dalsza analiza powinna być również skierowana tak, aby odpowiedzieć na pytanie, czy można zbudować portfele skointegrowane z wartościami indeksu WIG.

Literatura

1. Alexander C.O. (1999). Optimal Hedging Using Cointegration. Philosophical Transactions of the Royal Society A 357, 2039-2058.
2. Engle R.F. and Granger C.W.J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. Econometrica, 55 (2), 251-276.
3. Granger C.W.J., Terasvirta T. (1993). Modelling Nonlinear Economic Relationships. Oxford University Press, New York.
4. Lucas A. (1997). Strategic and Tactical Asset Allocation and the Effect of Ling-Run Equilibrium Relations. Research Memorandum 1997-42. Vrije Universiteit, Amsterdam.
5. Osińska M. (2006). Ekonometria finansowa. PWE, Warszawa.
6. Philips P.C.B, Ouliaris S. (1990). Asymptotic Properties of Residual Based Test of Cointegrations. Econometrica, 58, 78-93.
7. Ziwot E., Wang J. (2002). Modeling Financial Time Series with S-Plus. Springer-Verlag, New York.

Krzysztof S. Targiel
Akademia Ekonomiczna w Katowicach

IMPLEMENTACJA MODELU MERTONA W WYBRANYCH ŚRODOWISKACH OPTYMALIZACJI

Wstęp

W wielu dziedzinach życia zarządzanie ryzykiem staje się jednym z istotnych elementów zarządzania. Ryzyko rozumiane jest obecnie, w pewnym dystansie do oryginalnej definicji F.H. Knighta, jako zagrożenie niezrealizowania założonego celu. Możemy rozróżnić wiele rodzajów ryzyka, wśród których znajdujemy ryzyko kredytowe rozumiane jako brak możliwości terminowego wywiązania się kontrahenta z całości lub części zobowiązań finansowych. W tym obszarze szczególnie istotna jest analiza ryzyka, jakościowa i ilościowa jego identyfikacja. Waga tych rozważań ma szczególne znaczenie w działalności bankowej. Środowiska finansowe dostrzegły ten problem powierzając jego standaryzację Komitetowi Bazylejskiemu do spraw Nadzoru Bankowego – międzynarodowej instytucji tworzącej standardy działalności bankowej. Ostatnie zalecenia Komitetu, odnoszące się do zarządzania ryzykiem, noszą nazwę Basel II. Dokument ten został także wprowadzone w Polsce. Uchwałą Komisji Nadzoru Bankowego z 13 marca 2007 roku, zaleceniom tym nadano nazwę Nowej Umowy Kapitałowej. Standardy te dopuszczają stosowanie przez banki własnych modeli oceny ryzyka. Rodzi to szerokie możliwości rozwoju nowych oraz modyfikacji już istniejących modeli. Jedną z pierwszych zaproponowanych metod jest model strukturalny zaproponowany przez Mertona [6]. Kluczowym elementem tej metody, jest ocena parametrów procesu stochastycznego zmienności wartości aktywów. Obecnie komercyjna wersja metody jest nazywana modelem Moody's-KMV [3].

1. Ryzyko kredytowe

Komitet Bazylejski do spraw Nadzoru Bankowego (Basel Committee on Banking Supervision) dopuścił w Nowej Umowie Kapitałowej, używanie przez banki, oprócz standardowej procedury, własnych modeli oceny ryzyka kredytowego [2]. Muszą one jednak określać pewne zdefiniowane parametry. Są nimi [3]:

- *EAD* – (Exposure At Default) zagrożona wartość umowy kredytowej,
- *LGD* – (Loss Given Default) strata w razie niedotrzymania warunków umowy (stopa straty),
- *PD* – (Probability of Default) prawdopodobieństwo niedotrzymania warunków umowy kredytowej.

Znajomość powyższych parametrów pozwala obliczyć wielkość straty banku wynikającej z niedotrzymania warunków umowy kredytowej. Wystąpienie faktu niedotrzymania warunków kredytu jest nazywane defaultem. Oznaczając przez *LI* zmienną losową zerojedynkową oznaczającą wystąpienie zjawiska defaultu, możemy obliczyć stratę *L* (Loss) jako

$$L = EAD \cdot LGD \cdot LI \tag{1}$$

Zmienna losowa *LI* ma rozkład geometryczny o wartości oczekiwanej *PD*. Stąd wartość oczekiwana wielkości straty *EL* (Expected Loss) może zostać obliczona ze wzoru

$$EL = EAD \cdot LGD \cdot PD \tag{2}$$

W zależności od wybranego przez bank podejścia, wielkości zagrożonej wartości umowy kredytowej (parametr *EAD*) i stopa straty w razie niedotrzymania warunków umowy (parametr *LGD*) mogą być szacowane przez bank lub wyliczane zgodnie z wytycznymi Komitetu Bazylejskiego. Obowiązkiem banku pozostaje oszacowanie prawdopodobieństwa wystąpienia defaultu.

2. Model Mertona

Jednym z możliwych do wykorzystania sposobów oceny prawdopodobieństwa defaultu jest model Mertona [6]. Merton zaproponował „strukturalne" podejście do oceny ryzyka, poprzez uwzględnienie struktury finansowej przedsiębiorstwa. Aktywa (A_t) firmy są finansowane z kapitałów własnych (K) oraz zobowiązań (Z).

$$A_t = K_t + Z \tag{3}$$

Należy podkreślić, iż w modelu tym jest rozpatrywana rzeczywista wartość aktywów, a nie wartość księgowa zapisana w księgach rachunkowych. Jeśli firma nie jest w stanie swoimi aktywami pokryć zobowiązań, zarząd jest zobligowany ogłosić upadłość. W takiej sytuacji uregulowania prawne sprawiają, że z pozostałej części aktywów zaspokajani są w pierwszej kolejności wierzyciele, a dopiero w ostatniej akcjonariusze. Jeśli wartość aktywów spadnie poniżej wartości zobowiązań, akcjonariusze nie otrzymają nic ze swoich kapitałów. Sprawia to, że na kapitały z punktu widzenia akcjonariuszy można spojrzeć jak na opcję kupna o następującej funkcji wypłaty (w momencie realizacji)

$$K_T = (A_T - Z)^+ \tag{4}$$

Merton zaproponował wykorzystanie modelu Blacka-Scholesa do wyceny wartości kapitałów. Zgodnie z jego ideą, jeśli wartość rzeczywistą aktywów będziemy modelować jako geometryczny proces Wienera

$$dA_t = \mu \cdot A_t\, dt + \sigma_A \cdot A_t\, dW \tag{5}$$

uzyskamy następujące wzory na wartość kapitałów

$$K_t = A_t \cdot \Phi(d_1) - Z \cdot e^{-r(T-t)} \cdot \Phi(d_2) \tag{6}$$

$$d_1 = [\, ln(A_t/Z) + (r + \sigma_A{}^2/2) \cdot (T-t)\,] / [\sigma_A \cdot (T-t)^{0,5}] \tag{7}$$

$$d_2 = d_1 - \sigma_A \cdot (T-t)^{0,5} \tag{8}$$

gdzie w powyższych wzorach poprzez σ_A została oznaczona zmienność aktywów. Jest to wielkość nieobserwowalna, podobnie jak realna wartość aktywów A_t. Natomiast wartość kapitału akcyjnego, może być bezpośrednio obliczona, jeśli tylko firma jest notowana na giełdzie. Z przebiegu notowań giełdowych możemy także oszacować zmienność historyczną kapitałów σ_K .

Konsekwencją przyjętego modelu dynamiki zmian wartości aktywów jest fakt, iż w momencie wykonania, logarytm wartości aktywów (lnA_T) ma rozkład normalny

$$lnA_T \sim N(lnA_t + (\mu + \sigma_A{}^2/2) \cdot (T-t);\ \ \sigma_A \cdot (T-t)^{0,5}) \tag{9}$$

Wykorzystując powyższy fakt można obliczyć prawdopodobieństwo defaultu, czyli prawdopodobieństwo tego, iż aktywa osiągną poziom niższy niż poziom zobowiązań

$$\mathrm{P}(default) = \mathrm{P}(lnA_T \le lnZ) = \Phi([lnZ - E(lnA_T)] / \sigma(lnA_T)) \tag{10}$$

Korzystając z dystrybuanty standardowego rozkładu normalnego Φ można obliczyć pożądany parametr

$$\mathrm{P}(default) = \Phi([lnZ - (lnA_t + (\mu + \sigma_A^2/2) \cdot (T-t)] / (\sigma_A \cdot (T-t)^{0,5}) \tag{11}$$

Aby obliczyć prawdopodobieństwo defaultu, konieczna jest znajomość parametrów geometrycznego procesu Wienera opisującego dynamikę zmian wartości aktywów. Parametr dryfu μ może być oszacowany z modelu CAPM [5]. Pozostaje problem oszacowania parametru zmienności aktywów σ_A. Wykorzystać można zależność podawaną w literaturze przedmiotu [2; 3]

$$\sigma_K = \sigma_A \cdot \Phi(d_1) \cdot A_t / K_t \tag{12}$$

W rezultacie problem sprowadza się do rozwiązania układu dwu równań

$$\sigma_K = f(A_t, \sigma_A, Z, r, t) \tag{13}$$

$$K_t = f(A_t, \sigma_A, Z, r, t) \tag{14}$$

w których niewiadomymi są wartość aktywów A_t i zmienność aktywów σ_A. Analityczne rozwiązanie tych równań nie jest możliwe, lecz można znaleźć numerycznie wartości A_t i σ_A, dla których równania są spełnione. Nie jest to jednak zadanie trywialne.

3. Metoda numeryczna rozwiązania problemu

Metoda Löfflera i Poscha [5] polega na minimalizacji sumy kwadratów procentowych różnic pomiędzy uzyskiwaną ze wzoru (14) a rzeczywistą wartością kapitałów (K_t^*) oraz pomiędzy wyliczoną ze wzoru (12) a obliczoną wartością zannualizowanej zmienności kapitałów (σ_K^*).

Postawione zadanie ma postać

$$\underset{A_t, \sigma_A}{\text{Min}} \; [\, (K_t/K_t^* - 1)^2 + (\sigma_K/\sigma_K^* - 1)^2 \,] \tag{15}$$

$$K_t = A_t \cdot \Phi(d_1) - Z \cdot e^{-r(T-t)} \cdot \Phi(d_2) \tag{16}$$

$$d_1 = [\, \ln(A_t/Z) + (r + \sigma_A{}^2/2) \cdot (T - t) \,] / [\sigma_A \cdot (T - t)^{0,5}] \tag{17}$$

$$d_2 = d_1 - \sigma_A \cdot (T - t)^{0,5} \tag{18}$$

$$\sigma_K = \sigma_A \cdot \Phi(d_1) \cdot A_t / K_t \tag{19}$$

Powyższy problem jest zadaniem typu NLP. Wiadomo jednak, kiedy uzyskujemy rozwiązanie optymalne. Ponieważ funkcja celu jest sumą kwadratów względnych różnic pomiędzy wartościami kapitału i zmienności, minimalna wartość jaką może ona przyjąć to zero.

Uzyskanie zerowej wartości funkcji celu oznacza, iż znaleziono wartości A_t, σ_A spełniające wzory (13) i (14).

Zmienność kapitałów liczy się jako zmienność historyczną z logarytmicznych stóp zwrotu zgodnie z następującymi wzorami

$$\sigma_K = [\sum_i (u_i - k)^2/(n-1)]^{0,5} \tag{20}$$

gdzie:

$$k = (\sum_i u_i)/n \tag{21}$$

$$u_i = \ln (S_i/S_{i-1}) \tag{22}$$

Sumy liczone są po wszystkich n dniach w kwartale, dla których można policzyć logarytmiczną stopę zwrotu (u_i) z cen zamknięcia (S_i) akcji spółki.

Zmienności liczone w okresach kwartalnych, są przeliczane na okresy roczne (annualizowane) zgodnie z wzorem

$$\sigma_{KkwA} = m^{0,5} \sigma_{Kkw} \tag{23}$$

gdzie m to liczba okresów kwartalnych w roku.

4. Implementacja w wybranych środowiskach

Znana jest w literaturze implementacja metody w pakiecie MS EXCEL opisana w pracy Löfflera i Poscha [5]. Do rozwiązania wykorzystuje się standardowo dodawany do pakietu Solver. Korzystanie z tego rozwiązania rodzi jednak pewne problemy. Nie zawsze uzyskuje się oczekiwane rozwiązanie, w którym wartość funkcji celu jest równa zero. Uzyskanie rozwiązania optymalnego jest zależne od wartości początkowych zmiennych, od których rozpoczyna się rozwiązywanie problemu. Ponieważ problem jest zadaniem optymalizacyjnym, zrodził się pomysł jego rozwiązywania w zintegrowanych środowiskach optymalizacyjnych. Takie podejście daje możliwość zautomatyzowania prac wyboru rozwiązania początkowego oraz ewentualnie jego ponownego rozwiązywania, gdy nie uzyskano rozwiązania optymalnego.

Do obliczeń wykorzystano dane dotyczące firmy ELEKTRIM S.A. dla pierwszego kwartału 2002 roku pochodzące z pracy [9] a przedstawione w tabeli 1.

Tabela 1

Dane do obliczeń (ELEKTRIM SA)

Okres	Z Zobowiązania (w tys. PLN)	K_t Kapitał (w tys. PLN)	σ_K Zmienność kapitału	A Aktywa zaksięgowane (w tys. PLN)	T
I kwartał 2002 roku	3 246 259,00	736 061,645	0,089800	5 989 336,00	1,00

Źródło: Opracowanie własne na podstawie [9].

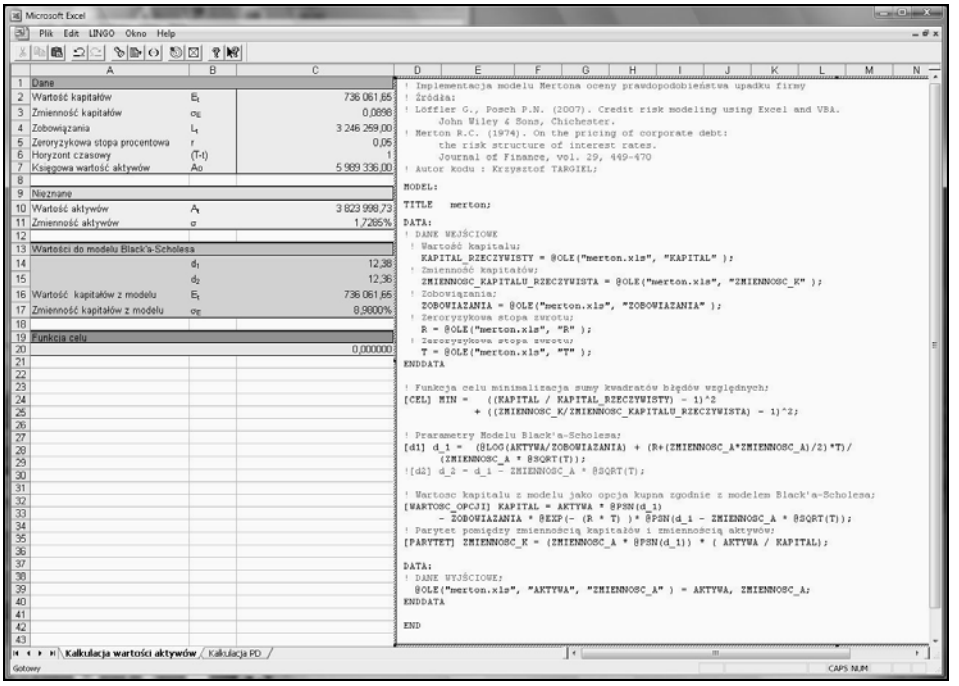
Implementację modelu Mertona w języku LINGO przedstawiono w tabeli 2. Język LINGO zdefiniowany w pracach [4; 8] został wybrany ze względu na swoją prostotę, brak konieczności definiowania zmiennych, automatyczny dobór solvera do typu zadania [10].

Tabela 2

Listing implementacji w języku LINGO

```
! Implementacja modelu Mertona oceny prawdopodobieństwa upadku firmy
! Źródła:
! Löffler G., Posch P.N.(2007). Credit risk modeling using Excel and
VBA.
       John Wiley & Sons, Chichester.
! Merton R.C. (1974). On the pricing of corporate debt:
       the risk structure of interest rates.
       Journal of Finance, vol. 29, 449-470
! Autor kodu : Krzysztof TARGIEL;
MODEL:
TITLE   merton;
DATA:
! DANE WEJŚCIOWE
 ! Wartość kapitalu;
   KAPITAL_RZECZYWISTY = @OLE("merton.xls", "KAPITAL" );
 ! Zmienność kapitałów;
   ZMIENNOSC_KAPITALU_RZECZYWISTA = @OLE("merton.xls", "ZMIENNOSC_K" );
 ! Zobowiązania;
   ZOBOWIAZANIA = @OLE("merton.xls", "ZOBOWIAZANIA" );
 ! Zeroryzykowa stopa zwrotu;
   R = @OLE("merton.xls", "R" );
 ! Zeroryzykowa stopa zwrotu;
   T = @OLE("merton.xls", "T" );
ENDDATA

! Funkcja celu minimalizacja sumy kwadratów błędów względnych;
[CEL] MIN =   ((KAPITAL / KAPITAL_RZECZYWISTY) - 1)^2
              + ((ZMIENNOSC_K/ZMIENNOSC_KAPITALU_RZECZYWISTA) - 1)^2;

! Prarametry Modelu Black'a-Scholesa;
[d1] d_1 =
(@LOG(AKTYWA/ZOBOWIAZANIA)+(R+(ZMIENNOSC_A*ZMIENNOSC_A)/2)*T)/
       (ZMIENNOSC_A * @SQRT(T));
![d2] d_2 = d_1 - ZMIENNOSC_A * @SQRT(T);

! Wartosc kapitalu z modelu jako opcja kupna zgodnie z modelem Black'a-
Scholesa;
[WARTOSC_OPCJI] KAPITAL = AKTYWA * @PSN(d_1)
       - ZOBOWIAZANIA * @EXP(-(R*T))* @PSN(d_1 - ZMIENNOSC_A*@SQRT(T));
! Parytet pomiędzy zmiennością kapitałów i zmiennością aktywów;
[PARYTET] ZMIENNOSC_K = (ZMIENNOSC_A * @PSN(d_1)) * ( AKTYWA / KAPITAL);

DATA:
! DANE WYJŚCIOWE;
  @OLE("merton.xls", "AKTYWA", "ZMIENNOSC_A" ) = AKTYWA, ZMIENNOSC_A;
ENDDATA

END
```

W przedstawionej implementacji dane pobierane są z arkusza kalkulacyjnego MS Excel zapisanego w pliku ‘merton.xls’. Sam kod programu został osadzony jako obiekt w tym arkuszu. Uzyskane wartości aktywów oraz zmienności aktywów są także eksportowane do arkusza. Widok arkusza przedstawiono na rys. 1.



Rys. 1. Arkusz ‘merton.xls’

Rozwiązanie uzyskano stosując solver globalny będący elementem środowiska LINGO. Pozwala on wykorzystać mechanizm multistartu. Dzięki tej możliwości nie było konieczne inicjalizowanie zmiennych, co było mankamentem implementacji z wykorzystaniem standardowego solvera pakietu MS Excel. Problemem był jednak nieoczywisty, zwłaszcza zapis parametrów d1 i d2 modelu Blacka-Scholesa jako ograniczeń modelu, a nie jako obliczanych ad hoc wartości.

Model zaimplementowano także w środowisku GAMS. Środowisko to wraz ze zdefiniowanym w nim językiem opisu problemów optymalizacyjnych zostało przedstawione w pracach [1; 7]. Listing implementacji przedstawia tabela 3.

Tabela 3

Listing implementacji w języku GAMS

```
$Title Merton

$Ontext
 Implementacja modelu Mertona oceny prawdopodobieństwa upadku firmy
 Źródła:
 Löffler G., Posch P.N.(2007).Credit risk modeling using Excel and VBA.
         John Wiley & Sons, Chichester.
 Merton R.C. (1974). On the pricing of corporate debt:
         the risk structure of interest rates.
         Journal of Finance, vol. 29, 449-470.
 Autor kodu : Krzysztof TARGIEL;
$Offtext
Scalars
   d1                          parametr modelu Black'a-Scholesa
   d2                          parametr modelu Black'a-Scholesa;

Parameters
    KAPITAL_RZECZYWISTY           Wartość kapitalu
                                / 736061.645 /
    ZMIENNOSC_KAPITALU_RZECZYWISTA Zmienność kapitałów
                                / 0.0898 /
    ZOBOWIAZANIA                  Zobowiązania
                                / 3246259.00 /
    R                             Zeroryzykowa stopa zwrotu
                                / 0.05 /
    T                             Czas do wykonania
                                / 1.0 /

Variables
    E                             wartość funkcji celu
    AKTYWA                        wartość aktywów z modelu
    ZMIENNOSC_A                   zmienność aktywów z modelu
    KAPITAL                       wartość kapitalu  z modelu
    ZMIENNOSC_K                   zmienność kapitałów z modelu;

Positive variables KAPITAL, ZMIENNOSC_K,ZMIENNOSC_A, AKTYWA;

* Inicjalizacja zmiennych bieżących modelu;

* Początkowa wartość aktywów jest bliska wartości księgowej;
  AKTYWA.l = KAPITAL_RZECZYWISTY + ZOBOWIAZANIA;

* Początkowa zmienność aktywów jest obliczana z parytetu ;
  ZMIENNOSC_A.l = ZMIENNOSC_KAPITALU_RZECZYWISTA *
                 (KAPITAL_RZECZYWISTY / AKTYWA.l );
* Ograniczenie dolne zmiennej Kapitał by nie dzielić przez zero
  KAPITAL.lo = 0.01;

Equations
  CEL              Funkcja celu
  WARTOSC_OPCJI    Model Black'a-Scholesa
  PARYTET          Parytet zmiennością kapitałów i aktywów;

* Funkcja celu
CEL.. E
=e= 1000*SQR( (KAPITAL/KAPITAL_RZECZYWISTY) - 1 )
```

```
   + SQR( (ZMIENNOSC_K/ZMIENNOSC_KAPITALU_RZECZYWISTA) - 1 );

d1 = (LOG(AKTYWA.l/ZOBOWIAZANIA) +
        (R + SQR(ZMIENNOSC_A.l)/2)*T)/(ZMIENNOSC_A.l * SQRT(T));
d2 = d1 - ZMIENNOSC_A.l * SQRT(T);

* Wartość kapitalu z modelu Black'a-Scholesa;
WARTOSC_OPCJI.. KAPITAL
=e= AKTYWA * ERRORF( d1 )
  - ZOBOWIAZANIA * EXP(- (R * T) )* ERRORF( d2 );

* Parytet pomiędzy zmiennością kapitałów i zmiennością aktywów;
PARYTET.. ZMIENNOSC_K
=e= ZMIENNOSC_A *(AKTYWA/KAPITAL)* ERRORF( d1 );

Model merton /all/;

Option decimals = 6, nlp=minos;

Solve merton using nlp minimizing E;

Display AKTYWA.l, ZMIENNOSC_A.l, E.l, KAPITAL.l, ZMIENNOSC_K.l;
```

W tej implementacji dane są zapisane w kodzie programu. Do rozwiązania problemu wykorzystano solver Minos. Nie każdy dołączany do pakietu GAMS solver przeznaczony do rozwiązywania problemów typu NLP dawał poprawne wyniki. Konieczne okazało się także w tym środowisku skalowanie procentowych udziałów kapitałów i zmienności kapitałów. Jest to widoczne w funkcji celu. Łatwiejszy był natomiast zapis modelu. Formuły określający wartości d1 i d2 są traktowane jako wyrażenia. Sam model, oprócz funkcji celu, ma dwa ograniczenia. Pierwsze z nich określające z modelu Blacka-Scholesa wartość kapitału firmy, drugie określające wartość zmienności kapitałów z parytetu ze zmiennością aktywów.

Podsumowanie

Jak przedstawiono wyżej, pewne zadania identyfikacji ryzyka można sprowadzić do zadań optymalizacji. Ma to miejsce w przedstawionym modelu Metrona służącym do oceny prawdopodobieństwa upadku firmy (Probability of Default). Praca przedstawiła implementacje modelu w wybranych środowiskach optymalizacji. Jednak własności tych środowisk nie zawsze pozwalają na łatwy zapis modelu. Przyjazne środowisko LINGO, w którym nie ma konieczności definiowania zmiennych, a system sam rozpoznaje typ zadania, sprawiło pewne trudności w zapisie bardziej złożonego modelu jakim jest model Mertona. Łatwiej przyszło zapisać model w środowisku GAMS, w którym nie ma takich udogodnień. Tutaj trudności wystąpiły jednak przy rozwiązywaniu, doborze właściwego solvera i jego parametrów.

Podsumowując, w pracy przedstawiono implementację modelu Mertona, który został przedstawiony jako zadanie optymalizacyjne w dwóch środowiskach optymalizacji GAMS i LINGO. Pokazano także pewne trudności w implementacji wynikające ze specyfiki danego języka opisu problemu optymalizacyjnego.

Literatura

1. Brooke A., Kendrick D., Meeraus A. (1988). GAMS: A User's Guide. Scientific Press, Redwood City, CA.
2. Gątarek D., Maksymiuk R. i inni. (2001). Nowoczesne metody zarządzania ryzykiem finansowym. WIG-Press, Warszawa.
3. Jajuga K. (2007). Zarządzanie ryzykiem. Wydawnictwo Naukowe PWN, Warszawa.
4. LINGO 10.0 User's Manual (2006). LINDO Systems Inc. Chicago, http://www.lindo.com
5. Löffler G., Posch P.N. (2007). Credit Risk Modeling using Excel and VBA. John Wiley & Sons, Chichester.
6. Merton R.C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. Journal of Finance, 29, 449-470.
7. Rosenthal R.E. (2007). GAMS – A User's Guide. GAMS Development Corporation, Washington.
8. Schrage L., Cunningham K. (1988). Demo LINGO/PC: Language for Interactive General Optimization, version 1.04a. LINDO Systems Inc., Chicago.
9. Targiel K. (2008). Wyznaczanie parametru zmienność w modelu Mertona. W: Badania operacyjne i systemowe: decyzje, gospodarka, kapitał ludzki i jakość. Red. Jan W. Owsiński, Z. Nahorski, T. Szapiro. Polska Akademia Nauk, Warszawa.
10. Targiel K. (2009). Wybrane języki opisu problemów optymalizacyjnych. AE, Katowice.

Grażyna Trzpiot
Akademia Ekonomiczna w Katowicach

PESYMISTYCZNA OPTYMALIZACJA PORTFELOWA

Wprowadzenie

Prowadzone badania nad podejmowaniem decyzji w warunkach ryzyka i w warunkach niepewności pokazują nowe ścieżki postępowania w odróżnieniu od podejść klasycznych Warunki związane z kryterium oczekiwanej użyteczności są zastąpione poprzez pesymistyczne kryteria decyzyjne. Wybrane uwagi o tym podejściu w alokacji portfelowej stanowią treść opracowania.

1. Nieklasyczna użyteczność

Rozważając problem wyboru pomiędzy dwoma zmiennymi losowymi X oraz Y o dystrybuantach odpowiednio F i G, wyznaczamy oczekiwane użyteczności i preferujemy X, jeżeli

$$E_F u(X) = \int_{-\infty}^{\infty} u(x)dF(x) > \int_{-\infty}^{\infty} u(x)dG(x) = E_G u(Y) \tag{1}$$

Funkcja użyteczności u jest związana z nastawieniem decydenta do ryzyka [10]. Powyższy warunek można zapisać równoważnie następująco

$$E_F u(X) = \int_{0}^{1} u(F^{-1}(t))dt > \int_{0}^{1} u(G^{-1}(t))dt = E_G u(Y) \tag{2}$$

gdzie wykorzystano funkcje kwantyli odpowiednio $F^{-1}(t)$ dla zmiennej losowej X oraz $G^{-1}(t)$ dla zmiennej losowej Y [6].

Teoria Choqueta zakłada, że mamy inną miarę prawdopodobieństwa i w miejsce całkowania względem dt dopuszczamy całkowanie względem $dv(t)$, gdzie v jest funkcją transformującą na [0, 1]. Preferencje są zatem opisane zarówno poprzez funkcję użyteczności u, jak i funkcję transformującą v, czyli poprzez parę (u, v), gdzie u opisuje nastawienie do przepływów pieniężnych, natomiast v przekształca rozkład prawdopodobieństwa.

Zatem możemy zapisać, że preferujemy X, jeżeli

$$E_{v,F}u(X) = \int_0^1 u(F^{-1}(t))dv(t) > \int_0^1 u(G^{-1}(t))dv(t) = E_{v,G}u(Y) \tag{3}$$

Ponieważ zamieniliśmy zmienne opisujące zachowania decydenta, porządkujemy zdarzenia, zatem wartości u($F^{-1}(t)$) oraz u($G^{-1}(t)$) są również porządkowane w odniesieniu do rosnących oczekiwań. Funkcja transformująca v odzwierciedla rosnące lub malejące prawdopodobieństwa w odniesieniu do uporządkowanych zdarzeń.

Zbiór monotonicznych funkcji użyteczności jest wykorzystywany do rozróżnienia pomiędzy zadaniami decyzyjnymi w warunkach ryzyka oraz w warunkach niepewności. Teoria Choqueta wykorzystuje często funkcją transformującą v: $[0, 1] \to [0, 1]$ z wartościami granicznymi $v(0) = 0$ oraz $v(1) = 1$ oraz miarę prawdopodobieństwa, która jest formalnym zapisem rozkładu zdarzeń. Jako zwrot otrzymujemy czytelną postać opisu preferencji w postaci łatwo interpretowalnej pary funkcji (u, v).

2. Pesymistyczne podejście inwestycyjne

Wartość oczekiwana w sensie Choqueta daje łatwą interpretację funkcji transformującej rozkładu prawdopodobieństwa jako obrazu optymizmu lub pesymizmu decydenta (inwestora). Jeżeli funkcja transformująca jest wklęsła (concave) wówczas najmniej preferowane zdarzenia otrzymują rosnące wagi a najbardziej preferowane zdarzenia otrzymują malejące wagi (są zdyskontowane), to odpowiada pesymizmowi. Aby to zilustrować, rozpatrujemy przypadek absolutnie ciągłej oraz wklęsłej (concave) funkcji v oraz rozkład prawdopodobieństwa o malejącej funkcji gęstości. Przykładowo, w miejsce rozkładu równomiernego wartość oczekiwana w sensie Choqueta podwyższa wagi niepożądanych zdarzeń, a redukuje wagi zdarzeń preferowanych. Jeżeli funkcja transformująca jest wypukła (convex) sytuacja jest odwrotna, przeważa optymizm, podnosimy wiarygodność korzystnych dla decydenta zdarzeń i umniejszamy rolę najgorszych wydarzeń. Znaną funkcją transformującą jest parametryczna rodzina

$v_\alpha(t) = \min[t/\alpha, 1]$ dla $\alpha \in [0, 1]$

W tym przypadku mamy

$$E_{v_\alpha}u(X) = \alpha^{-1}\int_0^\alpha u(F^{-1}(t))dt \tag{4}$$

Zauważamy, że prawdopodobieństwo dla α najmniej preferowanych zdarzeń jest podlega *inflacji* oraz $1 - \alpha$ część najbardziej preferowanych zdarzeń jest zdyskontowana.

3. Miary ryzyka

Odpowiadając na pytanie jak mierzyć ryzyko portfela zostały określone aksjomaty koherentnych miar ryzyka.

Definicja 1. Koherentna miara ryzyka to funkcjonał ρ: $\mathcal{L}^2 \to (-\infty,\infty)$ o następujących własnościach:

a) monotoniczność: $X, Y \in \mathcal{G}$, jeżeli $X \le Y$, to $\rho(Y) \le \rho(X)$,

b) subaddytywność: $\rho(X + Y) \le \rho(X) + \rho(Y)$, dla dowolnych $X, Y \in \mathcal{G}$,

c) dodatnia homogeniczność: $X \in \mathcal{G}$ i $\lambda \ge 0$, $\rho(\lambda X) = \lambda\rho(X)$,

d) translacja inwarintana: $c \in R$, zachodzi $\rho(X + c \cdot r_f^*) = \rho(X) - c$.

Powyższe zasady wyeliminowały zastosowania wielu konwencjonalnych miar ryzyka wykorzystywanych w finansach. W szczególności miary wykorzystujące drugi moment rozkładu stopy zwrotu ze względu na brak monotoniczności oraz VaR ze względu na brak subaddytywności.

Miara ryzyka, która jest koherentna oraz zyskuje coraz szersze zastosowanie ma postać

$$\rho_{v_\alpha}(X) = -\int_0^1 F^{-1}(t)dv(t) = -\alpha^{-1}\int_0^\alpha F^{-1}(t)dt \tag{5}$$

gdzie $v_\alpha(t) = \min[t/\alpha, 1]$ dla $\alpha \in [0, 1]$.

Różne wersje tej miary $\rho_{v_\alpha}(X)$ były przedmiotem badań na polskim rynku kapitałowym [7; 8; 9; 10]. Wielu autorów podejmowało analizy własności tej miary: shortfall [1], CVaR [5], tail conditional expectation [2]. Będziemy posługiwać się zapisem miara-α. Miara ta jest ujemną wartością oczekiwana w sensie Choqueta v_α. Zapisując miarę ryzyka mierzoną w odniesieniu do kwantyla α mamy naturalne kryterium: $\rho_{v_\alpha}(X) - \lambda\mu(X)$ lub $\lambda\mu(X) - \rho_{v_\alpha}(X)$

ponieważ

* r_f stopa zwrotu instrumentu wolnego od ryzyka.

$$\mu(X) = \int F_X^{-1}(t)dt = -\rho_{v_1}(X) \tag{6}$$

zatem to kryterium daję klasę miar ryzyka zdefiniowaną poniżej.

Definicja 2. Miara ryzyka jest pesymistyczna jeżeli dla pewnej miary probabilistycznej φ na [0, 1] zachodzi

$$\rho(X) = \int_0^1 \rho_{v_\alpha}(X)d\varphi(\alpha).$$

Aby zobaczyć dlaczego, można powiedzieć, że ta miara jest pesymistyczna. Zapiszemy

$$\rho(X) = -\int_0^1 \alpha^{-1}\int_0^\alpha F^{-1}(t)dtd\varphi(\alpha) = -\int_0^1 F^{-1}(t)\int_t^1 \alpha^{-1}d\varphi(\alpha)dt \,^{\dagger} \tag{7}$$

W najprostszym przypadku, gdy φ jest skończona sumą można zapisać wykorzystując funkcję delta Diraca $d\varphi = \sum_{i=1}^m \varphi_i \delta_{\tau_i}$, $\varphi_i \geq 0$, $\sum_{i=1}^m \varphi_i = 1$ oraz $0 = \tau_0 < \tau_1 < \cdots < \tau_m = 1$.

Zatem

$$\int_t^1 \alpha^{-1}\delta_\tau(\alpha)d\alpha = \tau^{-1}I(t<\tau) \,^{\ddagger}$$

Możemy zapisać

$$\rho(X) = \int_0^1 \rho_{v_\alpha}(X)d\varphi(\alpha) = -\varphi_0 F^{-1}(0) - \int_0^1 F^{-1}(t)\gamma(t)dt \tag{8}$$

gdzie $\gamma(t) = \sum_{i=1}^m \varphi_i \tau_i^{-1} I(t<\tau_i)$.

Dla dodatnich wartości φ_i rozkład wag względem funkcji gęstości jest malejący, zatem rezultat transformacji rozkładu implikuje wiarygodność najmniej pożądanych zdarzeń oraz deprecjonuje wiarygodność najbardziej oczekiwanych, jest to mocno „pesymistyczne".

Możemy interpretować miarę-α jako wartości ekstremalne wypukłego zbioru koherentnych miar ryzyka [4].

---

† Zmiana kolejności całkowania na podstawie twierdzenia Fubbiniego.
‡ *I(A)* = 1 jeżeli A jest prawdziwe, *I(A)* = 0 w przeciwnym przypadku.

4. Regresja kwantylowa a miary ryzyka

Rozpatrzmy zadanie decyzyjne minimalizacji koherentnej miary ryzyka, czyli równoważnie zadanie maksymalizacji wartości oczekiwanej w sensie Choqueta z wykorzystaniem liniowej postaci funkcji użyteczności oraz wypukłej funkcji transformującej v_α.

Empiryczne strategie minimalizujące ryzyko mierzone w odniesieniu kwantyla rzędu α dostarczają narzędzia w postaci regresji kwantylowej. Niech

$$\rho_\alpha(u) = u(\alpha - I(u < 0)) \tag{9}$$

będzie kawałkami liniową funkcją straty (dystrybuanta empiryczna) oraz rozważmy problem

$$\min_{\xi \in R} E_{\rho_\alpha}(X - \xi) \tag{10}$$

Rozwiązaniem tego zadania jest kwantyl rzędu α zmiennej losowej X.

Zapisując zadanie reresji kwantylowej w ogólnym przypadku rozpatrujemy problem estymacji wektora nieznanych parametrów (regresji) **b** dla próby niezależnych obserwacji na ciągu zmiennych losowych $Y_1, Y_2, ..., Y_T$, zgodnych z rozkładem

$$P(Y_t < y) = F(y - x_t\mathbf{b}), \quad t=1,\ldots T \tag{11}$$

gdzie $\{x_t, t=1,\ldots T\}$ jest wierszem w znanej macierzy obserwacji (o wymiarach $T \times K$) oraz rozkład F nie jest znany [11].

Ciąg wartości $\{y_t, t=1,\ldots T\}$ to obserwacje w próbie losowej pobranej z populacji o rozkładzie **Y** mającym dystrybuantę F. Wówczas kwantyl rzędu α, dla $0 < \alpha < 1$, może być wyznaczony jako rozwiązanie zadania

$$\min_{\beta \in \mathbf{R}} \left\{ \sum_{\mathbf{t} \in \{\mathbf{t}: \mathbf{y_t} \geq \beta\}} \alpha \left| \mathbf{y_t} - \beta \right| + \sum_{\mathbf{t} \in \{\mathbf{t}: \mathbf{y_t} < \beta\}} (1-\alpha) \left| \mathbf{y_t} - \beta \right| \right\} \tag{12}$$

Zapiszemy jako $\{x_t, t=1,\ldots T\}$ ciąg K wektorów (wierszy) macierzy obserwacji, zakładamy, że $\{y_t, t=1,\ldots T\}$ jest losową próbą procesu regresji $u_t = y_t - x_t b$ mającym dystrybuantę F. Wówczas kwantyl regresji rzędu α, dla $0 < \alpha < 1$ jest zdefiniowany jako rozwiązanie zadania

$$\min_{\beta \in \mathbf{R}} \left\{ \sum_{\mathbf{t} \in \{\mathbf{t}: \mathbf{y_t} \geq \mathbf{x_t}\beta\}} \alpha \left| \mathbf{y_t} - \mathbf{x_t}\beta \right| + \sum_{\mathbf{t} \in \{\mathbf{t}: \mathbf{y_t} < \mathbf{x_t}\beta\}} (1-\alpha) \left| \mathbf{y_t} - \mathbf{x_t}\beta \right| \right\} \tag{13}$$

W przypadku K=1 oraz $x_t = 1$dla wszystkich t, model (13) redukuje się do zadania (12). Najmniejszy błąd bezwzględny jest wówczas równy medianie. Zadanie (13) ma zawsze rozwiązanie, w przypadku rozkładów ciągłych rozwiązanie jest jednoznaczne.

Zadanie minimalizacji związane z regresją kwantylową jest równoważne następującemu zadaniu programowania liniowego

$$\min\{\alpha \mathbf{1}'\mathbf{r}^{+} + (1-\alpha)\mathbf{1}'\mathbf{r}^{-}\} \tag{14}$$

przy ograniczeniach

$$\mathrm{y} = \mathbf{X}\mathrm{b} + \mathbf{r}^{+} + \mathbf{r}^{-}$$
$$(\mathrm{b}, \mathbf{r}^{+}, \mathbf{r}^{-}) \in \mathrm{R}^{\mathrm{K}} \times \mathrm{R}_{+}^{2\mathrm{T}}$$

gdzie **1** jest wektorem jednostkowym T wymiarowym.

5. Pesymistyczna optymalizacja portfelowa

Mając próbę losową x_i, dla $i = 1,\ldots, n$ empiryczne oszacowanie miary-α możemy zanotować następująco

$$\hat{\rho}_{\nu_\alpha}(x) = (n\alpha)^{-1} \min_{\xi \in R} \sum_{i=1}^{n} \rho_\alpha (x_i - \xi) - \hat{\mu}_n \tag{15}$$

gdzie $\hat{\mu}_n$ jest estymatorem $EX = \mu$, przykładowo $\overline{x}_n$.

Przechodząc do portfela zapiszemy: $Y = X^T \pi$ to portfel składający się z aktywów X_i o wagach π_i. Obserwujemy losową próbę: $x_{i1, \ldots,} x_{ip}$ dla $i = 1,\ldots, n$ z rozkładu stopy zwrotu portfela zatem minimalizujemy funkcję Lagrange'a

$$\min_{\pi} \rho_{\nu_\alpha}(Y) - \lambda\mu(Y) \tag{16}$$

To zadanie jest równoważne minimalizacji $\rho_{\nu_\alpha}(Y)$ przy ograniczeniach przyjętych co do poziomu wartości oczekiwanej stopy zwrotu. Alternatywnie możemy maksymalizować wartość oczekiwaną stopy zwrotu oczekiwany zwrot przy ograniczeniach na miarę-α. Ponieważ $\mu(Y) = -\rho_{\nu_\alpha}(Y)$, zatem średnia jest dodatkową miarą-α i może być traktowane jako dyskretna wersja pesymistycznej miary ryzyka.

Zapisując dodatkowo ograniczenia na wagi w portfelu mamy zadanie minimalizacji

$$\min_{\pi} \rho_{\nu_\alpha}(X^T \pi) \tag{17}$$

gdzie

$$\mu(X^T \pi) = \mu_0 \quad \text{oraz} \quad \mathbf{1}^{\mathrm{T}} \pi = 1$$

Wykorzystując zadanie regresji kwantylowej możemy zapisać ostatnie zadanie równoważnie w postaci

$$\min_{(\beta,\xi)\in R^p} \sum_{i=1}^{n} \rho_\alpha (x_{i1} - \sum_{j=2}^{p} (x_{i1} - x_{ij})\beta_j - \xi) \tag{18}$$

gdzie

$$\bar{x}\pi(\beta) = \mu_0 \quad \text{oraz} \quad \pi(\beta) = (1 - \sum_{j=2}^{p} \beta_j, \beta^T)^T$$

Ponieważ miara-α określa jedno parametryczną rodzinę koherentnych miar ryzyka, możemy szukać w sposób naturalny średniej ważonej

$$\rho_\nu (X) = \sum_{k=1}^{m} \nu_k \rho_{\nu_{\alpha_k}} (X) \tag{19}$$

gdzie wagi ν_k sumują się do jedynki.

Uogólnione zadanie z kryterium ograniczającym ryzyko można zatem zapisać następująco

$$\min_{(\beta,\xi)\in R^{p+m}} \sum_{k=1}^{m} \sum_{i=1}^{n} \nu_k \rho_\alpha (x_{i1} - \sum_{j=2}^{p} (x_{i1} - x_{ij})\beta_j - \xi_k) \tag{20}$$

oraz

$$\bar{x}\pi(\beta) = \mu_0$$

Mamy zatem zadanie wyznaczenia *m* estymatorów kwantyli rozkładów stóp zwrotów aktywów z portfela. Rozwiązujemy *m* zadań regresji kwantylowej. Asymptotyczne własności wyznaczanych kwantyli znajdziemy w pracach Koenkera [3]. Ponieważ pesymistyczna funkcja transformująca może być reprezentowana przez liniową wypukłą funkcję generowaną poprzez ważoną średnią miar-α, zatem możemy powiedzieć że rozwiązanie uogólnionego zadania (20) wyznacza estymatory wag pesymistycznego portfela.

Podsumowanie

Zapisaliśmy uogólnioną postać pesymistycznych preferencji w sensie Choqueta w warunkach ryzyka z liniową funkcją użyteczności jako optymalizacyjny problem alokacji portfelowej, który może być rozwiązany z wykorzystaniem regresji kwantylowej. Taki portfel może być rozumiany jako maksymalizacja wartości oczekiwanej stopy zwrotu przy przyjętym ograniczeniu na koherentną miarę ryzyka. Pesymistyczne nastawienie do ryzyka wprowadza

uzupełnienie do konwencjonalnego spojrzenia na awersję do ryzyka bazującą na maksymalizacji wartości oczekiwanej funkcji użyteczności. Alokacja portfela jest dobrą bazą do rozwijania tego podejścia w innych zastosowaniach.

Literatura

1. Acerbi C., Tasche D. (2002). Expected Shortfall: A Natural Coherent Alternative to Value at Risk. Economic Notes, 31, 379-388.
2. Artzner P., Delbaen F., Eber J.-M., Heath D. (1999). Coherent Measure of Risk. Mathematical Finance, 9, 203-228.
3. Koenker R., Bassett G. (1978). Regression Quantiles. Econometrica 46, 33-50.
4. Kusuoka S. (2001). On Law Invariant Coherent Risk Measures. Advances In Mathematical Economics, 3, 83-96.
5. Rockafellar R.T., Uryasev S. (2000). Optimization of Conditional Value-at-Risk. Journal of Risk, 2, 21-41.
6. Trzpiot G. (2003). Analiza portfelowa z wykorzystaniem metody momentów i dominacji stochastycznych – podejście kwantylowe. Prace Naukowe. AE, Wrocław, 990, 216-224.
7. Trzpiot G. (2004a). Kwantylowe miary ryzyka. Prace Naukowe. AE, Wrocław, 1022, 420-430.
8. Trzpiot G. (2004b). O wybranych własnościach miar ryzyka. Badania operacyjne i decyzje. 3-4, 91-98.
9. Trzpiot G. (2006a). Pomiar ryzyka finansowego w warunkach niepewności. Badania operacyjne i decyzje, 2, 81-88.
10. Trzpiot G. (2006b). Dominacje w modelowaniu i analizie ryzyka na rynku finansowym. AE, Katowice.
11. Trzpiot G. (2007). Regresja kwantylowa a estymacja VaR. Prace Naukowe. AE, Wrocław, 1176, 465- 471.
12. Trzpiot G. (2008). Implementacja metodologii regresji kwantylowej w estymacji VaR. Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, Szczecin, 9, 316-323.
13. Trzpiot G. (2008). Wybrane modele oceny ryzyka. AE, Katowice.

RYZYKO W UBEZPIECZENIACH I RYZYKO BANKOWE

Jacek Białek
Uniwersytet Łódzki

PROPOZYCJA MIAR DO OCENY DYNAMIKI OFE*

Wprowadzenie

Jak wiadomo, funkcjonujące miary efektywności Otwartych Funduszy Emerytalnych (OFE) nie są doskonałe. W polskim prawie obowiązuje definicja przeciętnej stopy zwrotu grupy OFE, która wyznacza tzw. minimalny zwrot dla funduszy. Ryzyko uzyskania stopy zwrotu za ostatnie 36 miesięcy mniejszej od wymaganego ustawowo minimum pociąga za sobą poważne konsekwencje finansowe. Zgodnie z polskim prawem, w sytuacji takiej fundusz jest zobligowany do pokrycia powstałego deficytu. Jednak jak pokazali Gajek i Kałuszka [5] miara przeciętnej stopy zwrotu nie spełnia pewnych ekonomicznie zasadnych postulatów. Co więcej, nie uwzględnia ona ryzyka inwestycyjnego. Dlatego część prac z zakresu problematyki OFE koncentruje się nad zastosowaniem miar uwzględniających ryzyko (np. miary Treynora, Jensena, czy Sharpe'a), inni autorzy próbują – godząc się z kryterium minimalnej stopy zwrotu – modyfikować istniejące rozwiązanie (patrz Gajek i Kałuszka [6; 1; 3]. Ale jest jeszcze inne wyjście – można pokusić się o konstrukcję zupełnie nowej miary, która brałaby pod uwagę zwroty osiągane przez OFE, lecz dodatkowo oceniała dynamikę ich przyrostów. W niniejszej pracy proponuje się dwie tego typu miary, które umożliwiają hierarchizację funduszy od najbardziej dynamicznego począwszy. Tego typu ranking wyklucza ryzyko wyboru takiego funduszu, który co prawda ma wysokie aktywa i satysfakcjonujące zwroty, ale słabnącą dynamikę zmian wartości jednostek uczestnictwa. W perspektywie czasu utraci on pozycję na rzecz chwilowo słabszych, ale bardziej dynamicznych funduszy.

* Praca naukowa finansowana ze środków na naukę w latach 2008-2010 jako projekt badawczy Nr N N111 306335.

1. Ocena dynamiki grupy funduszy emerytalnych

Poniżej zaprezentowano autorską miarę Przeciętnej Dynamiki Funduszy (PDF), opracowaną na bazie cenowego indeksu agregatowego przedstawionego w pracy [1]. Ustalmy przedział czasowy obserwacji jako $[T_1, T_2]$ i przyjmijmy następujące oznaczenia:

- $w_i(t)$ – wartość jednostki udziałowej i-tego funduszu w chwili t,
- $k_i(t)$ – liczba jednostek udziałowych i-tego funduszu w chwili t.

Wartość $w_i(t)$ ustalana jest przez podzielenie całkowitych aktywów i-tego funduszu przez liczbę jednostek tego funduszu. Mamy zatem

$$A_i(t) = k_i(t) w_i(t) \tag{1}$$

gdzie $A_i(t)$ oznacza wartość całkowitych aktywów netto i-tego funduszu.

Przy wprowadzonych oznaczeniach proponowana miara Przeciętnej Dynamiki Funduszy na przedziale czasowym $[T_1, T_2]$ ma postać[1]

$$PDF(T_1,T_2) = \sum_{i=1}^{N}\left[\frac{\sum_{u=T_1}^{T_2} w_i(u)k_i(u)}{\sum_{k=1}^{N}\sum_{u=T_1}^{T_2} w_k(u)k_k(u)} \cdot \sum_{u=T_1+1}^{T_2} \frac{1}{2}\cdot\left(\frac{w_i(u-1)k_i(u-1)}{\sum_{y=T_1+1}^{T_2} w_i(y-1)k_i(y-1)} + \frac{w_i(u)k_i(u)}{\sum_{y=T_1+1}^{T_2} w_i(y)k_i(y)}\right)\cdot\frac{w_i(u)}{w_i(u-1)}\right] \tag{2}$$

gdzie $N-$ liczba funkcjonujących funduszy (obecnie $N = 15$).

Ze względu na ograniczenia w zakresie publikacji dziennych danych o funduszach emerytalnych nie rozważano miary *PDF* dla danych o dziennej częstotliwości, choć pozwoliłyby one na dokonanie najdokładniejszego pomiaru Przeciętnej Dynamiki Funduszy (PDF). Za jednostkowy okres czasowy przyjęto w dalszej części miesiąc.

Zauważmy, że wykorzystując zależność (1) i wprowadzając następujące oznaczenia

$$\alpha_i^u = \frac{1}{2}\left(\frac{A_i(u-1)}{\sum_{y=T_1+1}^{T_2} A_i(y)} + \frac{A_i(u)}{\sum_{y=T_1+1}^{T_2} A_i(y)}\right) \tag{3}$$

dla $i = 1,2,...,N, \quad u = T_1+1,...,T_2$

[1] Rozważamy tu jedynie przypadek czasu dyskretnego. Przyjmiemy $[T_1, T_2] = 1,36.$

$$\beta_i = \frac{\sum_{u=T_1}^{T_2} A_i(u)}{\sum_{k=1}^{N}\sum_{u=T_1}^{T_2} A_k(u)}, \quad i = 1,2,...,N \tag{4}$$

wobec formuły (2) otrzymujemy

$$PDF(T_1,T_2) = \sum_{i=1}^{N} \beta_i \sum_{u=T_1+1}^{T_2} \alpha_i^u \frac{w_i(u)}{w_i(u-1)} \tag{5}$$

Dodajmy, iż formuły (3) i (4) określają pewne wagi, ponieważ zachodzi [2]

$\alpha_i^u \in [0,1]$, $\beta_i \in [0,1]$ dla $i = 1,2,...,N$, $u = T_1 + 1,...,T_2$

oraz

$$\sum_{i=1}^{N} \alpha_i^u = 1 \text{ dla } u = T_1 + 1,...,T_2, \quad \sum_{i=1}^{N} \beta_i = 1 \tag{6}$$

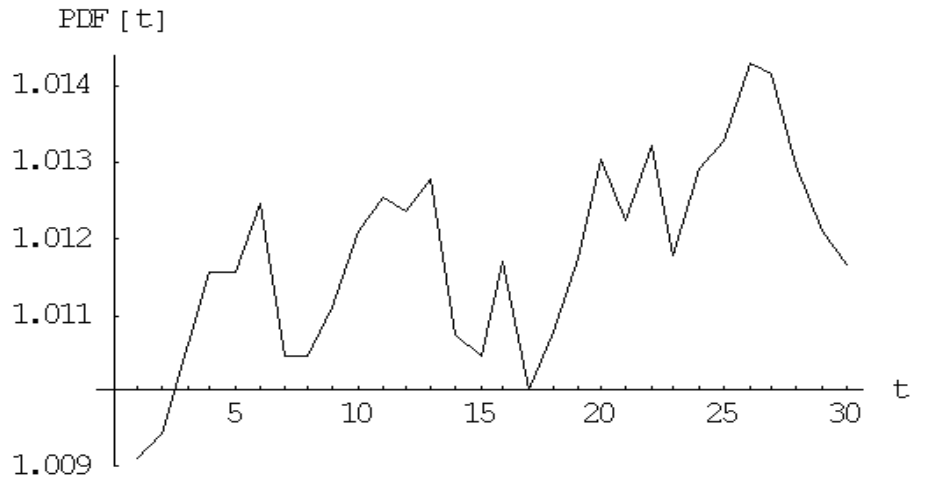
Współczynniki β_i określają udział aktywów *i*-tego funduszu względem całej grupy OFE w rozważanym przedziale czasu, natomiast α_i^u informują o tym, jak istotny był *u*-ty okres inwestycyjny dla *i*-tego funduszu. Interpretacja miary *PDF* jest następująca – wartości powyżej jedności wskazują na dodatnie tempo wzrostu (z okresu na okres) wartości jednostek uczestnictwa w grupie funduszy (po uśrednieniu). Im większa jest wówczas ta nadwyżka, tym dynamiczniej rosła średnia wartość jednostek uczestnictwa grupy funduszy w rozważanym przedziale czasowym. Wartości poniżej jedności informują o „niekorzystnej" sytuacji na rynku OFE. Należy dodać, iż dynamika rozumiana jest tu jako zmiana wartości z okresu na okres, a zatem istotną kwestią dla interpretacji i dokładności miary *PDF* jest ustalenie długości okresu obserwacyjnego. W pracy [2] zaprezentowano kilka podstawowych, statystycznych własności tej miary.

W pracy [4] dokonano analizy wpływu rozpiętości jednostkowego okresu obserwacyjnego na własności procesu $PDF(t) \equiv PDF(t, t+36)$.

Przykład 1

Rozważmy następujący okres funkcjonowania OFE w Polsce: 04.2002--09.2007. Nadmieńmy, iż był to dobry okres dla Otwartych Funduszy Emerytalnych w Polsce. W 2007 roku każdy z funduszy posiadał trzyletnią stopę zwrotu o wiele wyższą niż wymagane ustawowo minimum. Nie zdając sobie

sprawę z nadchodzącego kryzysu finansowego, OFE inwestowały w ryzykowne papiery wartościowe i póki co, pomnażały majątki swoich klientów. Obrazuje to rys. 1.



Rys. 1. Zależność miary PDF (liczonej dla 36 miesięcy na danych miesięcznych) dla okresu 04.2002-09.2007

Źródło: Obliczenia wykonane w programie Mathematica na podstawie danych z www.money.pl

Widać, iż w omawianym okresie miara Przeciętnej Dynamiki Funduszy przyjmuje wartości powyżej jedności, co wskazuje, iż był to korzystny okres dla rozwoju OFE. Co więcej, rys 1. sugeruje, iż dynamika wzrostu wartości jednostek uczestnictwa OFE posiadała wówczas trend rosnący.

2. Ocena dynamiki pojedynczego funduszu

Prezentowana miara PDF ocenia wypadkową kondycję całej grupy funduszy emerytalnych. Klienci poszczególnych funduszy są jednak zainteresowani oceną efektywności ich własnego funduszu na tle grupy. Funkcjonujące rankingi OFE nie obrazują niestety w pełni hierarchii funduszy. Najczęściej tworzy się je na podstawie wielkości aktywów, stopy zwrotu, wielkości pobieranych opłat, prowizji itd. Tymczasem nawet potężny fundusz, o znacznych aktywach, może charakteryzować się wyraźnie słabnącą dynamiką przyrostów wartości jednostki uczestnictwa. Co za tym idzie, wypracowuje on de facto coraz mniejsze profity dla swoich klientów i w perspektywie czasu może okazać się gorszy niż jego znacznie mniej zasobni konkurenci.

Zatem, poza miarodajną stopą zwrotu, istotna jest również znajomość jej dynamiki i trendu rozwoju. Wychodząc z tego punktu widzenia w niniejszej pracy zaproponowano miarę Atrakcyjności Dynamiki Funduszu (ADF). Zanim ją jednak zdefiniujemy zauważmy, iż w przypadku, gdy $N = 1$ (grupa OFE redukuje się do jednego funduszu o umownym numerze n_0) uzyskujemy wobec (2)

$$PDF_{n_0}(T_1, T_2) = \sum_{u=T_1+1}^{T_2} \alpha_{n_0}^u \frac{w_{n_0}(u)}{w_{n_0}(u-1)} \tag{7}$$

gdzie

$$\alpha_{n_0}^u = \frac{1}{2}\left(\frac{w_{n_0}(u-1)k_{n_0}(u-1)}{\sum_{y=T_1+1}^{T_2} w_{n_0}(y-1)k_{n_0}(y-1)} + \frac{w_{n_0}(u)k_{n_0}(u)}{\sum_{y=T_1+1}^{T_2} w_{n_0}(y)k_{n_0}(y)}\right). \tag{8}$$

Przyjmijmy zgodnie z przyjętym ustawodawstwem, iż interesuje nas ocena trzyletniej działalności OFE. Podzielmy ten przedział czasowy na trzy roczne interwały. Na podstawie (2) i (7) dla miesięcznych okresów obserwacji definiujemy

$$D_{1,i} \equiv \frac{PDF_i[25,36]}{PDF[25,36]},\ D_{2,i} \equiv \frac{PDF_i[13,24]}{PDF[13,24]},\ D_{3,i} \equiv \frac{PDF_i[1,12]}{PDF[1,12]} \tag{9}$$

Miary określone w (9) informują, jaka była relacja przeciętnej, jednomiesięcznej zmiany wartości jednostki danego *i*-tego funduszu w stosunku do analogicznej zmiany dla grupy odpowiednio dla minionego roku, roku wcześniejszego i jeszcze wcześniejszego. Traktując priorytetowo dane „najmłodsze" określmy, np. metodą wykładniczą, wagi związane z poszczególnymi latami.

Tabela 1

Przypisanie wag związanych z trzema minionymi latami działalności OFE

Okres	Waga
[25,36]	$\exp(-\beta)$
[13,24]	$\exp(-2\beta)$
[1,12]	$\exp(-3\beta)$

Dla $\beta > 0$ zachodzi

$$\exp(-\beta) + \exp(-2\beta) + \exp(-3\beta) = 1 \tag{10}$$

Numeryczne obliczona wartość występującego tu parametru to $\beta = 0.609358$. Ostatecznie, proponowana miara Atrakcyjności Dynamiki Funduszu (ADF) przyjmuje tutaj postać

$$ADF_i = (\sum_{k=1}^{3} \exp(-k\beta) D_{k,i} - 1) \cdot 100\% \tag{11}$$

Miara ADF_i informuje, o ile, procentowo ujmując, miesięczna zmiana wartości jednostki funduszu jest większa (mniejsza) od przeciętnej, analogicznej zmiany wśród całej grupy funduszy.

Gdy $ADF_i > 0$, to fakt ten oznacza, iż *i*-ty fundusz wykazał lepszą miesięczną dynamikę wartości jednostki uczestnictwa w porównaniu z całą grupą w rozważanym okresie.

Znak „<" oznacza, iż dynamika ta była słabsza, natomiast znak „=" oznacza, iż dynamika rozwoju wartości jednostki tego funduszu odpowiada dynamice całej grupy – w tym wypadku mamy do czynienia z zupełnie przeciętnym funduszem. Ranking OFE polegałby na traktowaniu jako najlepsze te fundusze, które mają największe wartości ADF_i.

3. Badanie empiryczne

W badaniu uwzględniono następujący okres funkcjonowania OFE: 12.2004-12.2008 (48 miesięcy). Dzieląc rozważany, czteroletni interwał czasowy na dwa mniejsze (trzyletnie): $\Delta T_1 = [1, 36]$ i $\Delta T_2 = [13, 48]$, uzyskujemy rezultaty dla stóp zwrotu funduszy i miary ADF zawarte w tabeli 2.

Tabela 2

Stopy zwrotu i wartości ADF_i dla OFE w okresie 12.2004-12.2008

Efektywność i dynamika OFE				
ΔT_1			ΔT_2	
Fundusz	Stopa zwrotu (%)	ADF_i	Stopa zwrotu (%)	ADF_i
1	2	3	4	5
AEGON	39.7773	0.234962	3.36497	0.132838
AIG	44.6831	0.292891	4.40252	-0.0829845
Allianz	38.2742	-0.259379	6.38298	0.0437167
AXA	40.6918	0.0889454	7.79857	-0.00869896
Bankowy	34.9699	-0.195804	0.692641	-0.0963167
Commercial Union	43.3602	0.324115	2.59133	-0.140731
Generali	44.4444	0.0897989	5.69276	-0.00972209
ING	43.3523	-0.00887259	1.85551	-0.105065

cd. tabeli 2

1	2	3	4	5
Nordea	39.1431	0.155459	3.49563	0.0918951
Pekao	45.7327	0.247469	6.12431	0.0606366
Pocztylion	41.6667	-0.0555121	4.12088	-0.0562537
Polsat	43.8967	0.0731422	-0.710059	-0.377207
PZU Złota Jesień	43.2933	0.01553	4.90824	-0.0364062
Warta	40.0877	-0.44016	0.817327	0.371707

Źródło: Opracowanie własne na podstawie www.money.pl

Wnioski

Widać, iż w okresie pierwszych 36 miesięcy analizowanego przedziału czasowego (a więc zanim nastąpił kryzys finansowy) dynamika przyrostów wartości jednostek uczestnictwa u większości funduszy była większa niż analogiczna średnia dynamika całej grupy. Fundusze, których dynamika była gorsza od średniej w grupie to Allianz, Bankowy, ING, Pocztylion i Warta. Co ciekawe, potężny ING, mimo wysokiej stopy zwrotu za ten okres (43,35%), cechuje się słabnącą dynamiką przyrostów wartości jednostki. Dynamika wartości jego jednostki uczestnictwa jest niemal identyczna jak średnia dynamika w całej grupie OFE (ADF_i = -0.0088). Przesuwając krańce rozważanego przedziału czasowego o 12 miesięcy do przodu obserwujemy zgoła odmienną sytuację. OFE wyraźnie odczuwają już skutki kryzysu finansowego – w okresie ΔT_2 znacznie spadają stopy zwrotu funduszy (Polsat osiąga nawet ujemną stopę zwrotu w wysokości -0,71%). Miara *ADF* wyraźnie wskazuje na załamanie rynku i złą kondycję funduszy emerytalnych. Większość z nich charakteryzuje się słabnącą dynamiką zmian wartości jednostki uczestnictwa. Co więcej, średnia ważona stopa zwrotu [1] za okres 12.2005-12.2008 (czyli 2.84%) i wyznaczona za jej pomocą minimalna stopa zwrotu, przewyższają możliwości finansowe niektórych funduszy (np. Bankowy, Polsat, Warta). Zatem istnieje realne ryzyko, iż niektóre OFE nie zdołają spełnić kryterium minimalnej stopy zwrotu i – zgodnie z ustawodawstwem polskim – zobowiązane będę do pokrycia zaistniałego deficytu ze środków własnych.

Literatura

1. Białek J. (2005). Jak mierzyć rentowność grupy funduszy emerytalnych? Model stochastyczny. W: Modelowanie preferencji a ryzyko'05. Red. T. Trzaskalik. AE, Katowice.
2. Białek J. (2006). The Average Price Dynamics and Indexes of Price Dynamics – Discrete Time Stochastic Model. Acta Universitatis Lodziensis. Folia Oeconomica WUŁ, Łódź, 196, 155-172.
3. Białek J. (2007). Wpływ fuzji funduszy emerytalnych na przecięną stopę zwrotu grupy OFE. W: Modelowanie preferencji a ryzyko'07. Red. T. Trzaskalik. AE, Katowice.
4. Białek J., Mikulec A. (2008). Statystyczna analiza wartości jednostek uczestnictwa i stóp zwrotu OFE. Wiadomości Statystyczne (w druku).
5. Gajek L., Kałuszka M. (2000). On the Average Return Rate for a Group of Investment Funds. Acta Universitas Lodziensis. Folia Oeconomica, 152, 161-171.
6. Gajek L., Kałuszka M. (2001). On Some Properties of the Average Rate of Return – A Discrete Time Stochastic Model (praca nieopublikowana).

Anna Jędrzychowska
Uniwersytet Ekonomiczny we Wrocławiu

NOWE OCENY DLA WARTOŚCI WSKAŹNIKÓW POLSKICH ZAKŁADÓW UBEZPIECZEŃ NA ŻYCIE

Wprowadzenie

U podstaw wielu analiz i ocen zakładów ubezpieczeń znajduje się często analiza wskaźnikowa. Jest ona szybką i efektywną metodą uzyskiwania wglądu w sytuację zakładu ubezpieczeń. Umożliwia ocenę przeszłej i teraźniejszej sytuacji zakładu ubezpieczeń oraz może być podstawą prognoz jego sytuacji w przyszłości.

W latach 2000-2002 został stworzony przez organ nadzoru system wczesnego ostrzegania dla polskich ubezpieczycieli, opierający się na ocenie wskaźnikowej. Elementem przeprowadzonej analizy było wyznaczenie ocen dla poszczególnych wartości badanych wskaźników. Oceny te wskazano po wyznaczeniu statystyk pozycyjnych (percentyli) dla populacji polskich zakładów ubezpieczeń (na podstawie danych finansowych z lat 1996-2000). Skale ocen zostały dopasowane do trzech podstawowych charakterów wskaźników: nominanty, stymulanty i destymulanty.

W związku z dynamicznym rozwojem rynku ubezpieczeń w Polsce oraz w związku z towarzyszącymi temu zmianami struktury kapitałowej, prawodawstwa regulującego rynek finansowy i ubezpieczeniowy, realiów rynkowych związanych z przystąpieniem do UE, należałoby zweryfikować ustalone wcześniej oceny i poddać pod dyskusję nowe, wyznaczone statystycznie ich oceny. Temu zadaniu (w odniesieniu do zakładów ubezpieczeń na życie) poświęcone zostało niniejsze opracowanie. Badanie oparto na danych finansowych[1] zawartych w sprawozdaniach finansowych polskich zakładów na życie za lata 1999-2007 oraz metodologii zbliżonej do zaproponowanej przez PUNU w 2001 roku.

[1] Dane finansowe pochodziły z Monitorów Polskich B za lata 2001-2007.

1. Opis badania

W opracowaniu przygotowanym przez PUNU dokonano oceny trzech podstawowych wskaźników – stymulanty, destymulanty, nominanty. Na użytek niniejszego opracowania wybrano po jednym wskaźniku z każdej z tych grup, tj.:

1) stymulantę – wskaźnik rentowności działalności lokacyjnej,
2) destymulantę – wskaźnik poziomu kosztów akwizycji,
3) nominantę – stopę rezerwy składki brutto.

Dla każdego typu wskaźników przygotowane są bowiem odpowiednie granice oparte na statykach pozycyjnych, według których nadawane są oceny poszczególnym ich wartościom (tabela 1).

Tabela 1

Przedziały odpowiadające ocenom wartości wskaźników

STYMULANTA				
Bardzo zła	**zła**	**średnia**	**dobra**	**bardzo dobra**
	percentyl (10)	kwartyl dolny	mediana	od kwartyl górny
Do percentyl (10)	kwartyl dolny	mediana	kwartyl górny	
DESTYMULANTA				
Bardzo dobra	**dobra**	**średnia**	**zła**	**bardzo zła**
	kwartyl dolny	mediana	kwartyl górny	percentyl (90)
Kwartyl dolny	mediana	kwartyl górny	percentyl (90)	
NOMINANTA				
Zła	**średnia**	**dobra**	**średnia**	**zła**
	percentyl (10)	percentyl (40)	percentyl (60)	percentyl (90)
Percentyl (10)	percentyl (40)	percentyl (60)	percentyl (90)	

Źródło: [5].

Przed dokonaniem wyliczeń podjęto rozważania na temat wyboru i przygotowania szeregów danych – wartości wskaźników. Wstępnie wskazano 22 metody przygotowania danych, które następnie poddano eliminacji. Wszystkie wersje przedstawione są niżej.

Wersja 1 – poszczególne wartości wskaźnika z lat 1999-2007 dla wszystkich zakładów ubezpieczeń (bez żadnej eliminacji).

Wersja 2 – poszczególne wartości wskaźnika z lat 1999-2007 dla wszystkich zakładów ubezpieczeń, po odrzuceniu wartości nietypowych.

Wersja 3 – poszczególne wartości wskaźnika z lat 2002-2007 dla wszystkich zakładów ubezpieczeń (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 4 – poszczególne wartości wskaźnika z lat 2002-2007 dla wszystkich zakładów ubezpieczeń, po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Wersja 5 – średnie wartości wskaźnika z lat 1999-2007 dla poszczególnych zakładów ubezpieczeń (bez żadnej eliminacji).

Wersja 6 – średnie wartości wskaźnika z lat 1999-2007 dla poszczególnych zakładów ubezpieczeń, po odrzuceniu wartości nietypowych.

Wersja 7 – średnie wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 8 – średnie wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń, po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Wersja 9 – średnie wartości wskaźnika z lat 1999-2007 dla poszczególnych lat (bez żadnej eliminacji).

Wersja 10 – średnie wartości wskaźnika z lat 1999-2007 dla poszczególnych lat, po odrzuceniu wartości nietypowych.

Wersja 11 – średnie wartości wskaźnika z lat 2002-2007 dla poszczególnych lat (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 12 – średnie wartości wskaźnika z lat 2002-2007 dla poszczególnych lat, po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Wersja 13 – mediana wartości wskaźnika z lat 1999-2007 dla poszczególnych zakładów ubezpieczeń (bez żadnej eliminacji).

Wersja 14 – mediana wartości wskaźnika z lat 1999-2007 dla poszczególnych zakładów ubezpieczeń, po odrzuceniu wartości nietypowych.

Wersja 15 – mediana wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 16 – mediana wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń, po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Wersja 17 – mediana wartości wskaźnika z lat 1999-2007 dla poszczególnych lat (bez żadnej eliminacji).

Wersja 18 – mediana wartości wskaźnika z lat 1999-2007 dla poszczególnych lat, po odrzuceniu wartości nietypowych.

Wersja 19 – mediana wartości wskaźnika z lat 2002-2007 dla poszczególnych lat (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 20 – mediana wartości wskaźnika z lat 2002-2007 dla poszczególnych lat, po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Wersja 21 – łączne wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń, z wagami 1/36; 2/36; 3/36; …; 6/36 (bez żadnej eliminacji); lata po zmianie prawa ubezpieczeniowego.

Wersja 22 – łączne wartości wskaźnika z lat 2002-2007 dla poszczególnych zakładów ubezpieczeń, z wagami 1/36; 2/36; 3/36; …; 6/36) po odrzuceniu wartości nietypowych; lata po zmianie prawa ubezpieczeniowego.

Uznano, że wersje, w których wykorzystane byłyby dane sprzed 2001 roku nie przyniosą odpowiednich wyników, ponieważ w tym roku dokonano znaczącej zmiany w prawie ubezpieczeniowym [2][2]. W związku z powyższym, pozostawiono wyłącznie wersje wyliczenia wskaźników na podstawie sprawozdań z lat 2002-2007. Ponadto uznano, że lepsze wyniki dla zobrazowania populacji zakładów ubezpieczeń otrzyma się przy zastosowaniu mediany niż średniej z próby. Wiąże się to z dużą dominacją dwóch zakładów na rynku polskim – PZU Życie oraz Warta Vita, które w znaczący sposób swoimi wynikami rzutują na wartości średnie z rynku. Jednocześnie uznano, że pojedyncze wartości wskaźników, które wynikać mogą z nietypowych działań zakładów ubezpieczeń w poszczególnych latach[3] pozostawiono, celem uwzględnienie drobnych wahań wiążących się z rozwojem i dynamiką rynku. Ostatecznie wyliczenia statystyk pozycyjnych oparto na danych według wersji: 15, 16, 19, 20, 21, 22.

Po przygotowania danych, wyznaczono statystyki pozycyjne i na ich podstawie, według zasady przyjętej w opracowaniu PUNU (tabela 1), przygotowano po 6 wersji kryteriów oceny dla poszczególnych wskaźników. Następnie, celem utworzenie nowych przedziałów oceny, wyznaczono mediany poszczególnych przedziałów.

[2] Do równie ważnych zmian zaliczyć należy zmianę w Rozporządzeniu Ministra Finansów w sprawie wyliczanie wysokości marginesu wypłacalności oraz minimalnej wysokości kapitału gwarancyjnego dla działów i grup ubezpieczeniowych. Dz.U. z 12.12.2003 r. Rozporządzenie to reguluje kwestie wymogów wypłacalności zgodnie z aktualnym prawem ubezpieczeniowym Unii Europejskiej.

[3] Przykładowo może wiązać się to z przeprowadzaniem zmian w profilu działalności czy systemie likwidacji szkód. Do takich działań można zaliczyć również duże kampanie reklamowe, proces łączenia czy przejęcia przez inny zakład.

2. Wyliczenia dla poszczególnych wskaźników

2.1. Stymulanta – rentowność działalności lokacyjnej

W odniesieniu do ubezpieczeń na życie należy, ze względu na ich na długoterminowy charakter, monitorować wielkości wydatków lokacyjnych w odniesieniu do wydatków ogółem. Odgrywa to bowiem istotną rolę w zachowaniu płynności zakładu ubezpieczeń. Ważną kwestią jest też ocena sposobu finansowania wydatków z działalności lokacyjnej wpływami netto z działalności operacyjnej i finansowej [4]. Służy tej analizie wskaźnik rentowności działalności lokacyjnej. Informuje on o tym jaki jest udział przychodów z lokat w lokatach ogółem zakładu ubezpieczeń. Wyraża się on wzorem

$$\text{Wskaźnik rentowności działalności lokacyjnej} = \frac{\text{dochody z lokat}}{\text{lokaty}}$$

Wśród zakładów ubezpieczeń wyższą rentowność prowadzonej działalności lokacyjnej osiągają zakłady ubezpieczeń działu I w strukturze lokat, w których przeważają lokaty długoterminowe dające wyższą wartość przychodów z lokat niż lokaty krótkoterminowe, które dominują w przypadków zakładów ubezpieczeń działu II. Nierzadko zakłady ubezpieczeń dążąc do zwiększenia rentowności działalności, lokując aktywa w instrumenty finansowe o wyższym ryzyku[4]. Dlatego też zagadnienie struktury portfela oraz poziomu jego bezpieczeństwa powinno być uzupełnieniem analizy wskaźnika rentowności działalności lokacyjnej.

Tabela 2 zawiera zestawienie statystyk porządkowych dla wartości wskaźnika w poszczególnych latach (2002-2007) tworzących przedziały oceny dla poszczególnych jego wartości. W pierwszym wierszu zaprezentowano wartości pochodzące z badania dokonanego przez PUNU. W ostatnim wierszu tabeli znajdują się natomiast przedziały wyznaczone jako mediana wartości wskaźnika obliczonych w sześciu wersjach badania.

Tabela 2

Przedziały ocen wskaźnika rentowności działalności lokacyjnej

Wersja	B. zła	Zła		Średnia		Dobra		B. dobra
1	2	3	4	5	6	7	8	9
Stara	do 0,0566	0,0566	0,0903	0,0903	0,1185	0,1185	0,1518	od 0,1518
15	do 0,054102	0,054102	0,061829	0,061829	0,118179	0,118179	0,188034	od 0,188034

[4] Decyzje te muszą być jednak zgodne z ramami dopuszczalnymi przez ustawę o działalności ubezpieczeniowej.

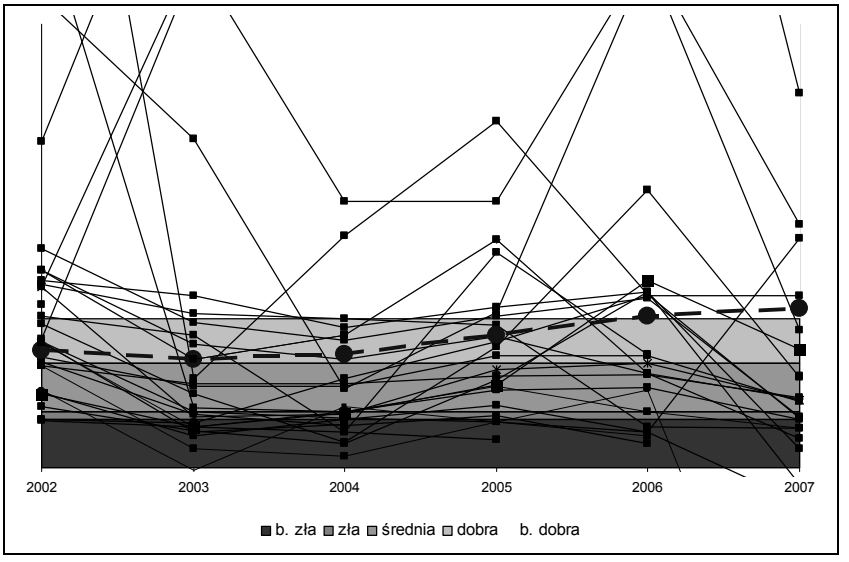
cd. tabeli 2

1	2	3	4	5	6	7	8	9
16	do 0,054102	0,054102	0,061829	0,061829	0,119179	0,119179	0,167549	od 0,167549
19	do 0,075818	0,075818	0,091664	0,091664	0,131003	0,131003	0,169275	od 0,169275
20	do 0,066337	0,066337	0,087604	0,087604	0,118008	0,118008	0,16975	od 0,16975
21	do 0,028914	0,028914	0,065285	0,065285	0,104975	0,104975	0,150052	od 0,150052
22	do 0,028514	0,028514	0,037844	0,037844	0,104887	0,104887	0,149467	od 0,149467
Nowa	do 0,0541	0,0541	0,063557	0,063557	0,118094	0,118094	0,168312	od 0,168312

Nowe przedziały ocen wskazują na to, że aby zakład ubezpieczeń otrzymał najwyższą ocenę, musi w lepszy niż dotychczas sposób zarządzać swoją polityką kolacyjną (wzrost progu przedziału o około 0,017). Na wykresie 1 znajduje się ilustracja tego, jak na tle nowo wyznaczonych ocen plasuje się polski sektor ubezpieczeń na życie. Przerywana linia na wykresie wskazuje na wartości wskaźnika dla całego sektora w omawianych latach.

Wykres 1

Przebieg wskaźnika rentowności działalności lokacyjnej w latach 2002-2007 na tle nowo ustalonych granic jego oceny



2.2. Destymulanta – wskaźnik poziomu kosztów akwizycji

Bezpośrednie przełożenie na wynik finansowy i rentowność działalności ubezpieczyciela mają koszty związane z prowadzeniem działalności. Jednym z podstawowych wskaźników badających poziom ponoszonych przez zakład ubezpieczeń kosztów jest

$$\text{Wskaźnik poziomu kosztów akwizycji} = \frac{\text{koszty akwizycji}}{\text{składka przypisana brutto}}$$

Wskaźnik ten (zwany również stopą kosztów akwizycji) informuje, jaki stopień składki przypisanej brutto stanowią koszty akwizycji[5], czyli określa wysokość kosztów akwizycji, która przypada na jednostkę składki przypisanej brutto. Wartość wskaźnika uzależniona jest też od kanałów dystrybucji produktów. W przypadku zakładów ubezpieczeń na życie, w konstrukcji wskaźnika należy uwzględnić jedną z dwóch zmian:

- koszty akwizycji skorygować zmianą wysokości odroczonych kosztów akwizycji,
- koszty akwizycji odnieść do składki zarobionej brutto (czyli po uwzględnieniu zmiany stanu rezerwy składek i rezerwy na niewygasłe ryzyko).

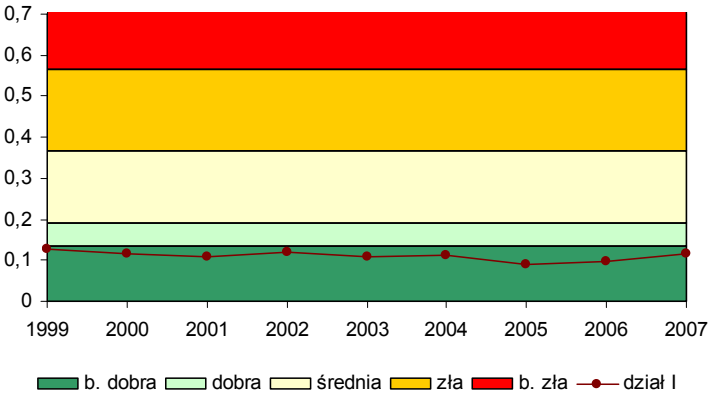
Zakłady ubezpieczeń mając trudności z pozyskaniem nowych klientów, a zatem ze zwiększaniem składki przypisanej brutto, dążą do obniżenia kosztów [1]. Takie działania mogą jednak dotyczyć tylko tych pozycji kosztów, które w bezpośredni sposób nie warunkują zwiększania składki przypisanej brutto. W przeciwnym wypadku można poprzez nieracjonalne obniżanie kosztów akwizycji doprowadzić do znacznego zmniejszenia udziałów w rynku, które będzie miało bezpośrednie przełożenie na wielkość składki przypisanej brutto w przyszłości. Powodami zmiany poziomów kosztów akwizycji mogą być: zmiana struktury wykorzystywanych kanałów dystrybucji, zastosowanie wyższych stawek prowizji, ponoszenie wyższych kosztów badania ryzyka lub innych kosztów bezpośrednich bądź rozszerzenia działalności promocyjnej (w tym przypadku przydatna byłaby analiza skuteczności przeprowadzonych działań). Sam fakt wzrostu wskaźnika jedynie dla jednego roku nie powinien być oceniany jednoznacznie negatywnie – np. koszty poniesione na promocję mogą dać efekty w postaci zwiększonego przypisu składek w kolejnym okresie. Wskaźnik ten wymaga zatem głębszych analiz, w tym również marketingowych.

[5] Koszty akwizycji są to wszelkie koszty bezpośrednio związane z pozyskiwaniem klienta, zawarciem ubezpieczenia oraz indeksem składki (prowizje agencyjne i brokerskie, koszty badań lekarskich, koszty ekspertyz i atestów przy ocenie ryzyka itp.). Koszty akwizycji obejmują również prowizje reasekuracyjne i udziały w zyskach płacone cedentom.

Dotychczasowy przebieg wartości wskaźnika dla działu ubezpieczeń na życie został na tle dotychczasowych jego ocen zaprezentowany na wykresie 2.

Wykres 2

Przebieg wskaźnika poziomu kosztów akwizycji dla zakładów ubezpieczeń działu I na tle granic jego oceny ustalonych przez PUNU



Wartości przedziałów oceny wskaźnika, wyznaczone na podstawie statystyk pozycyjnych prezentuje tabela 3, zaś wykres 3 stanowi ilustrację tego, jak wartości wskaźnika dla poszczególnych ubezpieczycieli przebiegają na tle nowych przedziałów jego oceny.

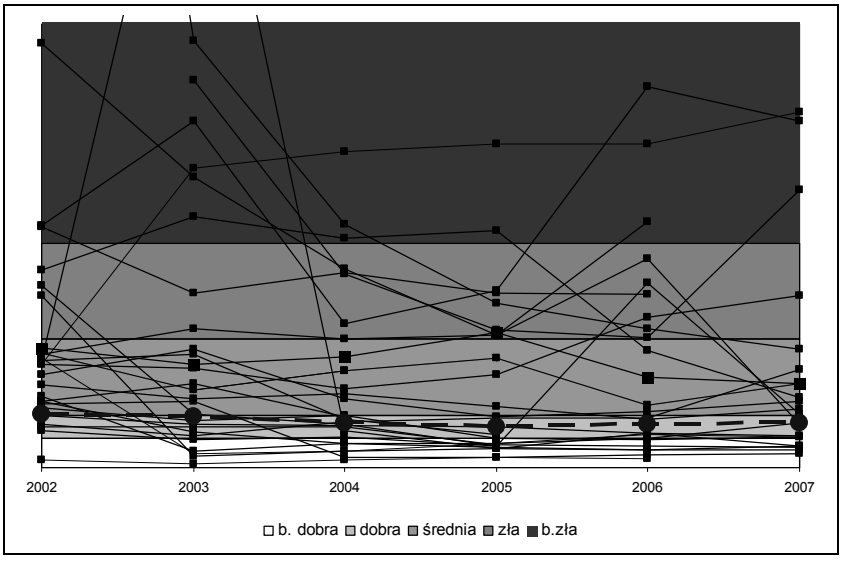
Tabela 3

Przedziały ocen wskaźnika poziomu kosztów akwizycji

Wersja	B. dobra	Dobra		Średnia		Zła		B. zła
Stara	do 0,1362	0,1362	0,1904	0,1904	0,3664	0,3664	0,564	od 0,564
15	do 0,066076	0,06676	0,118525	0,118525	0,293699	0,293699	0,543484	od 0,543484
16	do 0,066076	0,066076	0,118525	0,118525	0,293699	0,293699	0,517252	od 0,517252
19	do 0,109917	0,109917	0,114976	0,114976	0,175482	0,175482	0,240857	od 0,240857
20	do 0,109917	0,109917	0,114976	0,114976	0,175482	0,175482	0,240857	od 0,240857
21	do 0,063336	0,063336	0,137408	0,137408	0,289366	0,289366	0,551627	od 0,551627
22	do 0,063336	0,063336	0,137408	0,137408	0,289366	0,289366	0,495606	od 0,495606
Nowa	do 0,06642	0,06642	0,118525	0,118525	0,289366	0,289366	0,506429	od 0,506429

Wykres 3

Przebieg wskaźnika poziomu kosztów akwizycji w latach 2002-2007
na tle nowo ustalonych granic jego oceny



Analizując tabelę 3 oraz wykresy 2 i 3 należy zauważyć, że nastąpiło na rynku obniżenie poziomu kosztów akwizycji, przez co również za dobrą sytuację dla zakładu ubezpieczeń uznaje się koszty, których wskaźnik znajduje się pośnież wartości 0,06642. W stosunku do poprzednich ocen, jest to wartość dwukrotnie niższa. Wciąż bowiem w przypadku ubezpieczeń na życie sprzedaż polis odbywa się najczęściej poprzez bezpośredni kontakt ze sprzedawcą, a nie jak dzieje się to w przypadku ubezpieczeń non-life – przez kanały dystrybucji on-line czy direct[6]. Jednak w obliczu silnej konkurencji wśród osób, które sprzedają polisy oraz instytucji pośredniczących w sprzedaży ubezpieczeń spadek kosztów akwizycji jest znaczący.

3.3. Nominanta – stopa rezerwy składki brutto

Dla oceny adekwatności rezerw techniczno-ubezpieczeniowych zakładu ubezpieczeń wykorzystuje się wskaźnik informujący o udziale rezerwy składki brutto w składce przypisanej brutto, ma on postać

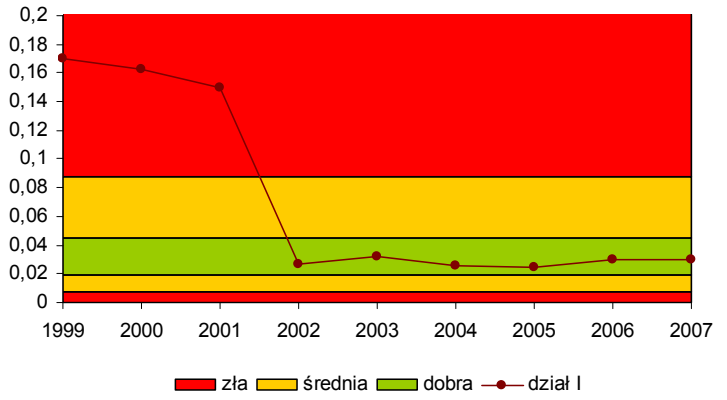
[6] Szerzej w raportach Głównego Urzędu Statystycznego (Ubezpieczenia dla gospodarstw domowych. GUS, Warszawa 2004 www.stat.gov.pl) oraz w opracowaniach Komisji Nadzoru Finansowego.

$$\text{Wskaźnik rezerwy składki brutto do składki przypisanej brutto} = \frac{\text{rezerwa składki brutto}}{\text{składka przypisanej brutto}}$$

Wartość tego wskaźnika zależna jest od daty zawierania umów ubezpieczeń oraz okresu trwania (jaką część okresów sprawozdawczych w sobie mieszczą). Konieczność tworzenia rezerw wynika bowiem z podstawowej zasady gospodarki finansowej zakładu ubezpieczeń – współmierności przychodów i rozchodów z danego okresu ubezpieczenia do okresu sprawozdawczego. Na rynkach rozwiniętych, jeśli zakład ubezpieczeń prowadzi rokrocznie stabilną sprzedaż rocznych polis, które co roku są odnawiane, wartość tego wskaźnika powinna wynosić 0,5. Dla ubezpieczycieli prowadzących długookresowe ubezpieczenia na życie, wartość ta winna być zdecydowanie niższa. Dotychczasowe oceny i poziom wskaźnika prezentuje wykres 4.

Wykres 4

Przebieg wskaźnika stopy rezerwy składki brutto dla zakładów ubezpieczeń działu I na tle granic jego oceny ustalonych przez PUNU



Zauważalna na wykresie zmiana wartości wskaźnika w latach 2001 i 2002 wiąże się ze zmianami w ustawie o działalności ubezpieczeniowej, między innymi zmianami dotyczącymi zasad tworzenia rezerw techniczno-ubezpieczeniowych i aktywów stanowiących ich zabezpieczenie.

Podobnie jak dla poprzednich wskaźników, także dla omawianego wykonano obliczenia przedziałów ocen wartości wskaźnika (tabela 4) oraz zaprezentowano wyniki zakładów na ich tle (wykres 5).

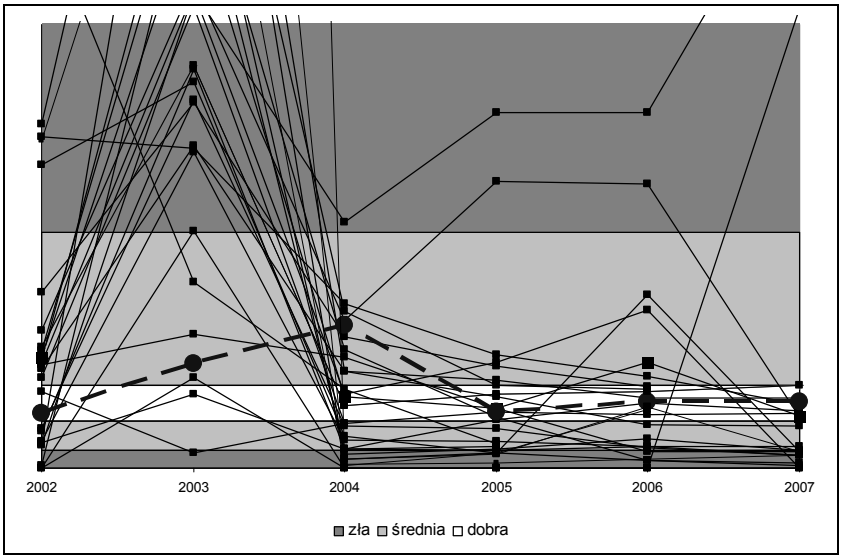
Tabela 4

Przedziały ocen wskaźnika poziomu kosztów akwizycji

Wersja	Zła	Średnia		Dobra		Średnia		Zła
Stara	do 0,0079	0,0079	0,0192	0,0192	0,0446	0,0446	0,0877	od 0,0877
15	do 0,003247	0,003247	0,019042	0,019042	0,037251	0,037251	0,066675	od 0,066675
16	do 0,003247	0,003247	0,017725	0,017725	0,037251	0,037251	0,066675	od 0,066675
19	do 0,009892	0,009892	0,021086	0,021086	0,025418	0,025418	0,207987	od 0,207987
20	do 0,009892	0,009892	0,021086	0,021086	0,025418	0,025418	0,194702	od 0,194702
21	do 0,008069	0,008069	0,044002	0,044002	0,049061	0,049061	0,118217	od 0,118217
22	do 0,007891	0,007891	0,042964	0,042964	0,047231	0,047231	0,094962	od 0,094962
Nowa	do 0,00798	0,00798	0,021086	0,021086	0,037251	0,037251	0,10659	od 0,10659

Wykres 5

Przebieg wskaźnika stopy rezerwy składki brutto w latach 2002-2007
na tle nowo ustalonych granic jego oceny



Podsumowanie

Dynamika rynków finansowych, w tym również sektora ubezpieczeń, wymaga ciągłej aktualizacji metod nadzoru nad nim. Wiąże się to nie tylko z wprowadzaniem coraz bardziej rozbudowanych modeli matematycznych, ale również z koniecznością korygowania wyliczonych parametrów opartych na danych historycznych. Ocena zakładów ubezpieczeń nie może być jednak automatycznym wyliczeniem i ocenieniem wartości poszczególnych wskaźników. Musi być osadzona na znajomości branży oraz zawierać analizę wzajemnych wpływów poszczególnych obszarów działalności zakładu ubezpieczeń.

Dokonane w opracowaniu obliczenia – wyznaczenie wartości oceny poszczególnych wskaźników, stanowią pierwszą próbę przygotowania nowego, aktualnego zestawu kryteriów oceny wartości poszczególnych wskaźników związanych z bezpieczeństwem finansowym zakładów ubezpieczeń. Natomiast docelowo wyliczenia te mogą znaleźć zastosowanie w klasyfikacji i ocenie (również ratingowej) ubezpieczycieli działających na polskim rynku.

Literatura

1. Borda M. (2004). Analiza wskaźnikowe jako instrument oceny sytuacji finansowej zakładu ubezpieczeń. W: Zarządzanie finansami w zakładach ubezpieczeń. Red. W. Ronka-Chmielowiec. Oficyna Wydawnicza Branta, Bydgoszcz-Wrocław, 206.
2. Ustawa o działalności ubezpieczeniowej. (2003). Dz.U., 6.07.
3. Jaworski W. (2002). Rating ubezpieczeniowy. AE, Poznań.
4. Karmańska A. (2000). Przepływy pieniężne w analizie finansowej ubezpieczyciela. Wiadomości Ubezpieczeniowe, 9-10, 27.
5. Metodologia analizy finansowej zakładów ubezpieczeń. (2001). PUNU, Warszawa.

Jerzy Marzec
Uniwersytet Ekonomiczny w Krakowie

POMIAR KORZYŚCI Z ZASTOSOWANIA MODELU WIELOMIANOWEGO W OGRANICZANIU RYZYKA KREDYTOWEGO

Wprowadzenie

Jednym z głównych założeń, leżących u podstaw modeli ekonomicznych opisujących działalność banków, jest przyjęcie, że działają one na rynku niedoskonałym, który charakteryzuje się niepełną informacją i niepewnością. Na podstawową (tradycyjną) działalność banków składa się przyjmowanie depozytów i transformowanie je w kredyty z wykorzystaniem pracy ludzkiej i kapitału fizycznego [30]. Zatem z ekonomicznego punktu widzenia zarządzanie bankiem sprowadza się w uproszczeniu do zarządzania aktywami i pasywami. Ryzyko kredytowe, podobnie jak ryzyko płynności, jest związane z bilansem banku i „[…] oznacza potencjalną niemożność odzyskania pełnej wartości księgowej aktywów" [22, s. 69]. Sytuacja ta może być wynikiem niespłacenia przez pożyczkobiorców rat kredytowych lub należnych odsetek. Bank jest przedsiębiorstwem maksymalizującym zysk, jednakże problem decyzyjny różni się od tego, który charakteryzuje typowe przedsiębiorstwo. Ryzyko, które towarzyszy udzielaniu kredytów, czyli sprzedaży podstawowego produktu bankowego, powoduje występowanie zjawiska ich racjonowania. To z kolei jest w sprzeczności z założeniami modeli popytu i podaży, w których zakłada się istnienie ceny równowagi, czyli takiego oprocentowania kredytów, po którym wszyscy potencjalni kredytobiorcy mogą zakupić kredyty. Teorie endogenicznego racjonowania kredytów należą do tych rozwiązań, które są spójne z zasadą maksymalizacji zysku i opisują racjonalne zachowanie się banków w sytuacji niedoskonałej informacji [11; 13] i pozycje bibliograficzne zawarte [9] oraz [22]. Modele te zakładają występowanie we wzajemnych relacjach między bankiem a kredytobiorcą takich zjawisk, jak asymetria informacji, negatywna selekcja, negatywne bodźce i pokusa nadużycia. Bank jest zainteresowany osłabieniem skutków powyższych czynników, szczególnie pierwszych trzech. Modele punktowej oceny zdolności kredytowej (credit scoring) są właśnie tymi narzędziami,

które umożliwiają ex ante określenie stopnia ryzyka związanego ze spłatą kredytu, czyli rozróżnienie na podstawie wcześniej zebranych informacji o kredytobiorcach bezpiecznych i ryzykownych, co przyczynia się do ograniczenia ryzyka kredytowego.

W niniejszym opracowaniu prezentuje się pomiar korzyści finansowych z zastosowania modeli wielomianowych dla kategorii uporządkowanych w procesie podejmowania decyzji kredytowych. W tym celu wykorzystano elementy statystycznej teorii decyzji, która w warstwie normatywnej pozwala określić decyzje, które są optymalne ze względu na cele decydenta przy dostępnej, często ograniczonej informacji. Jeżeli bank pragnie maksymalizować swój oczekiwany zysk z działalności kredytowej, to powinien postępować według określonych zasad. W tym opracowaniu proponuje się, aby w ramach systemu scoringowego stosować reguły udzielania kredytu, które zależą od marż kredytowych i prawdopodobieństw niespłacenia kredytów przez potencjalnych kredytobiorców. Ekonometryczny model wielomianowy dla kategorii uporządkowanych jest opisem zachowania się kredytobiorcy wobec spłaty długu. Jego estymacja na podstawie danych historycznych umożliwia wyznaczenie ocen a posteriori parametrów rozkładu prawdopodobieństwa zmiennej charakteryzującej niespłacalność kredytobiorcy. Niniejsze badania stanowią kontynuację tematyki dotyczącej wykorzystania modeli danych jakościowych w podejmowaniu decyzji kredytowych i oceny korzyści finansowych z tego tytułu dla banku [20; 21].

Proponowane podejście różni się istotnie od prezentowanego w literaturze polskiej. Wyniki badań empirycznych dotyczących oceny wiarygodności polskich kredytobiorców przedstawiają m.in. Chrzanowska i Witkowska [4] Witkowska i Chrzanowska [35; 36], Misztal [24]. Wykorzystują oni przede wszystkim sieci neuronowe, drzewa klasyfikacyjne bądź analizę dyskryminacyjną. Prace te mają charakter czysto empiryczny, główny nacisk jest położony na porównywanie wyników klasyfikacji otrzymanych za pomocą wspomnianych metod. W literaturze światowej, takie podejście można znaleźć między innymi [3; 10; 34]. Natomiast niniejsze badania nawiązują do dyskusji prowadzonej między innymi przez Thomasa [32] oraz Thomasa, Olivera i Handa [33], którzy uważają, że w przyszłości istotną rolę dla banku będzie pełnić scoring zysku, gdyż umożliwia on na poziomie pojedynczego klienta szacowanie zysku ze sprzedaży różnych produktów. Wobec powyższego, interdyscyplinarne ujęcie oceny przydatności modeli statystycznych w ograniczaniu ryzyka kredytowego – prezentowane tutaj – wydaje się ciekawym od strony praktycznej i poprawnym od strony metodycznej spojrzeniem na jakże ważny element działalności w banku – podejmowanie decyzji kredytowych.

1. Modele wielomianowe – definicja

W literaturze ekonometrycznej modele dla jakościowych zmiennych endogenicznych lub inaczej modele dyskretnego wyboru (quantal response lub discrete choice models) przedstawiają zależność między wynikiem dokonywanych wyborów a egzogenicznymi zmiennymi objaśniającymi, które opisują cechy możliwych alternatyw lub indywidualne charakterystyki podmiotów podejmujących decyzje.

Podstawowa definicja tych modeli opiera się na zaproponowanej przez McFaddena koncepcji stochastycznej funkcji użyteczności (random utility function) [31][1]. Funkcja użyteczności związana z decyzją o numerze j, opisująca preferencje konsumenta t, reprezentowana jest przez nieobserwowalną (ukrytą) zmienną losową z_{tj}. Podmiot dokonuje takiego wyboru, który przynosi mu najwięcej korzyści, czyli maksymalizuje funkcję użyteczności. Zmienna z_{tj} zależy nie tylko od charakterystyk możliwych wyborów i cech decydenta, ale także od składnika losowego, który reprezentuje wpływ na podjęte decyzje zakłóceń losowych, błędów pomiaru, nieodpowiedniej specyfikacji postaci funkcji użyteczności oraz błędnego postrzegania wielkości kosztów jednostkowych charakteryzujących poszczególne wybory. Podjęta decyzja jest reprezentowana przez endogeniczną zmienną losową, przyjmująca skończoną liczbę wartości.

Najprostszym przypadkiem modelu dyskretnego wyboru jest model dychotomiczny, gdy $J = 2$. Jeżeli zmienna objaśniana mierzona jest na skali porządkowej, to otrzymujemy wielomianowy model dla kategorii uporządkowanych.

W niniejszym opracowaniu przedmiotem analizy jest wielomianowy model dla kategorii uporządkowanych przy założeniu jednakowej liczby alternatyw oraz posiadania danych charakteryzujących jedynie podmiot dokonujący wybór. Wprowadzając ciągłe zmienne z_t, których wartości określają obserwowaną kategorię zmiennej y_t – podjętą decyzję, otrzymujemy model o następującej postaci [23]

$$\begin{cases} z_t = \mathbf{x}_t \cdot \boldsymbol{\beta} + \varepsilon_t \\ y_{tj} = 1 \quad \text{gdy} \quad \alpha_{j-1} < z_t < \alpha_j \ \text{ dla } \quad t = 1,\ldots,T \quad j = 1,\ldots,J \\ y_{tj} = 0 \quad \text{w przeciwnym przypadku,} \end{cases} \tag{1}$$

gdzie α_j spełniające warunek $\alpha_{j-1} < \alpha_j < \alpha_{j+1}$ są tzw. punktami granicznymi (ucięcia) zmiennej z_t, zaś $\boldsymbol{x}_t$ jest wektorem zmiennych egzogenicznych (lub ich znanych funkcji) charakteryzujących jednostkę podejmującą wybór. Pomocnicza zmienna y_{tj} przyjmuje wartość jeden, gdy $y_t = j$ albo zero w pozostałych

[1] Retrospektywne ujęcie podstaw teoretycznych modeli danych dyskretnych przedstawiono np. w [14].

przypadkach. Zmienna y_t jest zmienną skalarną, której wartości odpowiadają umownym numerom kategorii, czyli 1,2,…,J. O składnikach ε_t najczęściej zakłada się, jak w modelach dwumianowych, że są niezależnymi zmiennymi losowymi i posiadają identyczne rozkłady o wartości oczekiwanej równej zero i ustalonej wariancji, co jest jednym ze sposobów narzucenia identyfikowalności parametrów. Przyjmuje się zatem, że $\alpha_0 = -\infty$ i $\alpha_J = +\infty$ oraz $\alpha_1 = 0$, jeżeli w równaniu dla zmiennej z_t występuje wyraz wolny β_1.

Prawdopodobieństwa zaobserwowania kategorii o numerach j = 1,…,J, czyli parametry rozkładu zmiennej dyskretnej y_t, są równe różnicy dystrybuant

$$p_{tj} \equiv \Pr(y_{tj} = 1) = \Pr(\alpha_{j-1} < z_t < \alpha_j) = F(\alpha_j - \mathbf{x}_t\boldsymbol{\beta}) - F(\alpha_{j-1} - \mathbf{x}_t\boldsymbol{\beta}) \qquad (2)$$

gdzie $F(a)$ jest wartością dystrybuanty zmiennej losowej ε_t w punkcie a. Najbardziej znanymi przypadkami modelu (1) są modele probitowy i logitowy, które otrzymujemy, gdy $F(\cdot)$ jest dystrybuantą standaryzowanej zmiennej losowej o rozkładzie normalnym albo zmiennej o standardowym rozkładzie logistycznym.

W przypadku tych modeli, podstawową i najczęściej stosowaną metodą estymacji, gdy wykorzystuje się dane indywidualne, jest metoda największej wiarygodności, która ma charakter asymptotyczny [2; 15; 29]. W przypadku niestandardowych modeli – np. z rozkładem t Studenta – w pełni probabilistycznym podejściem do zagadnienia konstrukcji modelu, estymacji jego parametrów i testowania przyjętych założeń na podstawie zarówno dla małej jak i dużej próby, jest wnioskowanie bayesowskie [1; 17; 18].

2. Definicja modeli zastosowanych w badaniach

W przykładzie empirycznym zostaną przedstawione wyniki porównania zyskowności modeli wielomianowych, zastosowanych w scoringu kredytowym. W tym celu wykorzystano, obok modeli probitowego i logitowego, także specyfikację z rozkładem t Studenta.[2]

Rozszerzenie standardowych modeli, czyli logitowego i probitowego, może postępować w dwóch kierunkach. Pierwsza proponowana modyfikacja polega na przyjęciu innej postaci zależności między zmienną ukrytą z_t a zmiennymi objaśniającymi. W niniejszych badaniach zostanie wykorzystana formuła wielomianu drugiego stopnia względem zmiennych egzogenicznych w_{th}

$$\mathbf{x}_t \cdot \boldsymbol{\beta} = \beta_1 + \sum_h \beta_h \cdot w_{th} + \sum_h \sum_{i \geq h} \beta_{hi} \cdot w_{th} \cdot w_{ti}. \qquad (3)$$

[2] Rezultaty uzyskane na podstawie modeli dwumianowych przedstawiono w [20; 21].

Osiewalski i Marzec [28] zaproponowali nazywać specyfikację (3) modelem II rzędu (z aproksymacją liniową), w odróżnieniu do przypadku, w którym $\mathbf{x}_t \cdot \boldsymbol{\beta} = \beta_1 + \sum_h \beta_h \cdot w_{th}$, czyli modelu I rzędu. Takie rozszerzenie spotyka się zwłaszcza w badaniach empirycznych [6; 8; 12; 25]. O korzyściach z konstrukcji opisanej wzorem (3) pisał między innymi Marzec [17].

Druga modyfikacja – z punktu widzenia konstrukcji modeli danych jakościowych jest oryginalna i rzadko stosowana – polega na przyjęciu dla składnika ε_t rozkładu z szerszej klasy. W przypadku tej rodziny modeli, Albert i Chib [1] zaproponowali rozkład *t* Studenta o nieznanej liczbie stopni swobody $\nu \in (0; +\infty)$ [16; 18; 28]. Jednym z motywów zastosowania rozkładu *t* Studenta jest spostrzeżenie, że rozkład logistyczny może być aproksymowany powyższym rozkładem o stopniach swobody między 7 a 9 [26]. Zatem model z rozkładem *t* Studenta stanowi proste uogólnienie modeli probitowego ($\nu \to +\infty$) i logitowego ($\nu \in (7; 9)$).

W niniejszym opracowaniu rozważamy dwa modele: model z rozkładem *t* Studenta z aproksymacją II rzędu (M_1) oraz model z aproksymacją liniową: probitowy (M_2). Wyniki zaprezentowane w [17; 19] pokazały, że w modelu M_1 ocena parametru ν wynosi około 5,7 a odchylnie standardowe 0,44, otrzymano więc model zbliżony do logitowego. W konsekwencji dla uproszczenia prezentowanej analizy pominięto model logitowy, gdyż uzyskane na jego podstawie wyniki są bardzo zbliżone do tych z modelu M_1. Warto wspomnieć, że w ujęciu statystycznym model M_1 zdecydowanie lepiej opisuje zjawisko niespłacalności kredytów niż pozostałe dwa prostsze, które jednakże w zagadnieniach scoringu kredytowego są stosowane powszechnie.

3. Model wielomianowy w podejmowaniu decyzji kredytowych – pomiar korzyści

3.1. Opis problemu decyzyjnego

Model wielomianowy może stanowić rdzeń systemu scoringowego, gdy rozważa się kilka typów zachowań potencjalnego kredytobiorcy w odniesieniu do spłaty rat kapitałowo-odsetkowych. Przyjęto, że bank jest zainteresowany podjęciem decyzji o przyznaniu bądź odmowie kredytu w zależności od stanu natury, czyli podejścia kredytobiorcy do spłaty kredytu. W warunkach niepełnej informacji podjęcie decyzji o udzieleniu albo odmowie kredytu jest uzależnione od prawdopodobieństw zaobserwowania poszczególnych zachowań pożyczko-

biorców oraz cząstkowych wypłat pieniężnych będących konsekwencjami decyzji banku i postaw klientów wobec spłaty zaciągniętych długów. Warto przypomnieć, że w modelu dychotomicznym rozróżnia się wyłącznie kredyty niespłacane i spłacane, co najczęściej oznacza, że w pierwszym przypadku bank traci w całości zarówno kapitał, jak i marżę kredytową, a tylko te drugie są źródłem przychodów odsetkowych. W omawianym przypadku dopuszcza się – w odróżnieniu do powyższego modelu – częściową utratę kapitału i odsetek. Istotną kwestią staje się więc zdefiniowanie zmiennej wielomianowej y_t, która reprezentuje niespłacalność kredytów, czyli stany natury.

Z uwagi na dostępność danych proponuje się, aby w modelu wielomianowym definicje zmiennej y_t oprzeć na klasyfikacji należności, która jest podstawą do tworzenia rezerw celowych. Posiadając dane o kredytach detalicznych z lat 2000-2001 rozróżniono cztery kategorie należności: normalne, poniżej standardu, wątpliwe i stracone. Przyjmuje się, iż bank traci cały kapitał i odsetki tylko w przypadku ostatniej kategorii, więc stopa odzysku wynosi zero. Należności o kategorii poniżej standardu i wątpliwe przynoszą tylko część oczekiwanych zysków. Dla uproszczenia zakłada się[3], że stopa odzysku dla należności o kategorii poniżej standardu wynosi 0,8, w przypadku zaś należności wątpliwych kształtuje się na poziomie 0,5. Gdy kredyty są spłacane w terminie, to bank otrzymuje marżę kredytową równą m (w punktach procentowych), w krańcowym przypadku zaś traci marżę i kapitał.[4] Decyzję banku odzwierciedla dwupunktowa zmienna d_t. Bank udziela kredytu ($d_t = 1$), gdy oczekiwana wypłata z tytułu jego udzielenia jest większa od zera albo większa od wypłaty otrzymanej, gdy odmawia kredytu ($d_t = 0$). Wybór wariantu zależy od szczegółowej konstrukcji funkcji wypłat, od tego, czy rozważamy wyłącznie wypłaty pieniężne czy także koszty utraconych korzyści. Funkcję nagród, wynikających z potencjalnych decyzji banku w zależności od stanów natury (kategorii należności), czyli opis rzeczywistego problemu decyzyjnego, prezentuje tabela 1. Podstawy teoretyczne statystycznej teorii decyzji, wykorzystywanej w tej części badań, prezentowane są między innymi w [5; 7]. W kontekście statystycznej teorii decyzji problem udzielania kredytu w przypadku dwóch stanów natury omawia między innymi Osiewalski [27].

[3] Stopy odzysku zostały ustalone arbitralnie, ale mogą być szacowane na podstawie informacji o przebiegu dotychczasowych spłat rat i odsetek od kredytów zagrożonych.

[4] Przez marżę kredytową ($m > 0$) rozumie się różnicę miedzy realnym oprocentowaniem kredytu, a kosztem pozyskania finansujących go środków, tj. oprocentowaniem depozytów. Dla uproszczenia przyjęto, że marża jest identyczna dla wszystkich kredytów.

Tabela 1

Funkcja wypłat w przypadku czterech kategorii należności (m – marża kredytowa)*

Decyzje	Kategorie (stany natury)			
	j= 1 normalne	j= 2 poniżej standardu	j= 3 wątpliwe	j= 4 stracone
Udzielić $d_t = 1$	m	0,8 m-0,2 (1+m)	0,5 m-0,5 (1+m)	-(1+m)
Odmówić $d_t = 0$	-m	-0,8 m	-0,5 m	0
Prawdopodobieństwo	p_{t1}	p_{t2}	p_{t3}	p_{t4}

* Uwzględnienie kosztów alternatywnych jest umotywowane specyfiką działalności banku jako pośrednika między klientami mającymi nadwyżkę środków finansowych oraz tymi, którzy w danym momencie odczuwają ich brak. Deponenci lokują w banku swoje oszczędności, za które bank płaci im odsetki. Depozyty mają charakter czynnika produkcji, który generuje koszt, kredyty zaś są aktywami generującymi przychód. Bank nie może odmówić przyjęcia depozytu, co najwyżej może zniechęcić potencjalnego deponenta proponując mu niskie oprocentowanie. Zatem podejmując decyzję o odmowie kredytu bank traci możliwość uzyskania środków pieniężnych, które pokryłyby koszt pozyskania depozytów.

Niech prawdopodobieństwa zdarzeń, że potencjalny kredyt o numerze t należy do pierwszej, drugiej, trzeciej albo czwartej kategorii należności, wynoszą odpowiednio p_{t1}, p_{t2}, p_{t3} i p_{t4}, przy czym $p_{t1}+p_{t2}+p_{t3}+p_{t4}=1$. Przy podejmowaniu decyzji kredytowych w warunkach niepewności zakłada się, że bank jest w stanie ocenić szanse wystąpienia różnych stanów natury. Informacja ta może być wyrażona w postaci rozkładu prawdopodobieństwa dla y_t – rozkładu a priori lub a posteriori. Ten pierwszy wyrażają subiektywną wiedzę posiadaną przez bank na ten temat. Jednakże parametry p_{tj} (j = 1, …, 4) nie są znane, więc w niniejszym opracowaniu proponuje się je szacować według wzoru (2) na podstawie danych o udzielonych kredytach i modelu (1). W tym celu wykorzystano podejście bayesowskie, które umożliwia na podstawie informacji z próby aktualizację wstępnej wiedzy reprezentowanej przez rozkład a priori, która ostatecznie znajduje odzwierciedlenie w formie rozkładu posteriori.

Przy ustalonej funkcji wypłat i znanych wielkościach p_{tj} oczekiwana wypłata z tytułu udzielenia kredytu wynosi

$$EW(d_t = 1) = m \cdot p_{t1} + (0{,}6m - 0{,}2)p_{t2} + 0{,}5 - (1+m)p_{t4} \tag{4}$$

w przypadku zaś odmowy

$$EW(d_t = 0) = -m \cdot p_{t1} - 0{,}8m \cdot p_{t2} - 0{,}5m \cdot p_{t3} \tag{5}$$

Optymalną decyzją kredytową jest ta, która przynosi największą wypłatę (najmniejszą stratę). Zatem bank udzieli kredytu klientowi ($d_t = 1$), gdy oczekiwany zysk z tego tytułu jest wyższy od kosztów utraconych korzyści, co zachodzi, gdy

$$EW(d_t = 1) > EW(d_t = 0) \tag{6}$$

W przeciwnym przypadku klient spotka się z odmową przyznania kredytu. Przyjmijmy, że zyskiem z tytułu przyznaniu kredytu o jednostkowej wartości jest oczekiwana wypłata $EW(d_t = 1)$, w przypadku zaś odmowy $EW(d_t = 0)$.

Wypłata $EW(d_t = 1)$ jest realną pieniężną korzyścią z działalności kredytowej, więc z punktu widzenia adekwatności kapitałowej może ona być podstawą do obliczenia minimalnego poziomu kapitału regulacyjnego, gdy bank stosuje metodę wewnętrznych ratingów. Wówczas $EW(d_t = 1)$ wzięta z przeciwnym znakiem i pomnożona przez wartość kredytu stanowi oczekiwaną stratę z tytułu niespłacenia pojedynczego kredytu.

Z ekonomicznego punktu widzenia, przewaga proponowanego podejścia wynika z możliwości dokonania oceny przydatności rozważanych modeli w kontekście korzyści finansowych, które można osiągnąć w ramach scoringu kredytowego. Wielkość zysku można określić dla pojedynczego wniosku kredytowego lub portfela. Proponowana reguła pozwala na podejmowanie optymalnych decyzji kredytowych na podstawie czynników ekonomicznych i probabilistycznych, tj. marży kredytowej, prawdopodobieństwa częściowego i całkowitego niespłacenia długu przez kredytobiorcę. Zastosowanie tego narzędzia wspomagającego podejmowanie decyzji w warunkach niepewności prowadzi do zmniejszenia asymetrii informacji między pożyczkobiorcą a pożyczkodawcą, dodatkowo zaś może osłabić napływ nieuczciwych klientów, czyli ograniczyć niepożądane skutki negatywnej selekcji. W ramach tego podejścia możliwe jest indywidualne ustalenie – na podstawie wzorów (4) i (5) – ceny kredytu dla danego kredytobiorcy, która jest adekwatna do poziomu ryzyka tegoż klienta czy produktu. Przeciwdziała to zjawisku rezygnacji wiarygodnych kredytobiorców z ubiegania się o kredyty z powodu negatywnych bodźców, gdy jednakowa, wysoka jego cena zawiera koszty ryzyka wszystkich potencjalnych pożyczkobiorców.

3.2. Prezentacja wyników

Przedmiotem analizy jest portfel kredytów detalicznych o łącznej wartości 422,6 mln zł, udzielonych w latach 2000-2001. Wartość kredytów należących do kategorii należności normalnych wynosi 379,7 mln zł, do kategorii poniżej standardu 13,6 mln zł, kredyty wątpliwe i stracone zaś kształtują się na poziomie 13,9 i 15,4 mln zł. Marża kredytowa (m) wynosi 10,7 punktu procentowego[5], a macierz wypłat prezentuje tabela 2. Zauważmy, że wyłącznie

[5] Średnią wartość marży kredytowej obliczono na podstawie danych dla gospodarstw domowych z 2001 roku. W tym okresie średnie ważone oprocentowanie kredytów detalicznych wynosiło 22,2 punktu procentowego, cena depozytów gospodarstw domowych zaś wynosiła 11,5 punktu procentowego. Obliczeń dokonano na podstawie danych pochodzących z NBP.

w przypadku kredytów normalnych wypłata z tytułu przyznania kredytu jest większa od kosztów utraconych korzyści w przypadku odmowy. Dla pozostałych kategorii należności decyzją korzystniejszą jest rezygnacja z udzielenia kredytu.

Tabela 2

Macierz wypłat dla marży kredytowej równej 10,7 punktu procentowego

Decyzje	Kategorie (stany natury)			
	j= 1 normalne	j= 2 poniżej standardu	j= 3 wątpliwe	j= 4 stracone
Udzielić $d_t = 1$	0,107	-0,136	-0,5	-1,107
Odmówić $d_t = 0$	-0,107	-0,086	-0,054	0

W wariancie optymistycznym, gdyby wszystkie kredyty zostały spłacone w całości, to przy marży równej 10,7 punktu procentowego bank osiągnąłby zysk maksymalny w kwocie 45,2 mln zł. Faktyczna wypłata (zysk, Z_{Fakt}) z tytułu udzielenia tych kredytów liczona według stanów natury wynosi 14,8 mln zł. Składa się na nią: 40,6 mln zł zysku z udzielenia kredytów z kategorii należności normalnych i straty w kwocie 1,8 mln zł, 7 mln zł i 17 mln zł w przypadku kredytów poniżej standardu, wątpliwych i straconych. Faktyczna wartość utraconych korzyści z tytułu odmowy udzielenia kredytów wynosi zero, gdyż badano wyłącznie wnioski kredytowe, które zostały zaakceptowane przez bank.

Oczekiwane wypłaty

Tabela 3 przedstawia oczekiwane wypłaty z podjętych decyzji kredytowych. Oczekiwana wypłata z tytułu udzielenia kredytów kształtuje się na poziomie 13,5 mln zł w modelu M_1 i 13,9 mln zł w M_2, gdy wypłata faktyczna wynosi 14,8 mln zł. Wypłata uzyskana na podstawie modelu M_2 jest nieznacznie wyższa, o 0,5 mln zł. W trzech modelach oczekiwana wartość utraconych korzyści z tytułu odmowy kształtuje się na poziomie 41,2 mln zł. Przyjęcie za p_{t1}, p_{t2}, p_{t3} i p_{t4} częstości występowania poszczególnych kategorii w próbie, tj. 80,3%, 6%, 6,3% i 7,4%, odpowiada modelowi (1), w którym występuje jedynie wyraz wolny. W tym szczególnym modelu oczekiwana wypłata z tytułu udzielenia kredytów byłaby stratą równą 15 mln zł. W odniesieniu do rezultatów otrzymanych na podstawie proponowanych modeli (M_1 i M_2) oraz faktycznej wypłaty, równej 14,8 mln zł, model ten jest całkowicie nieskutecznym narzędziem prognozowania.

Tabela 3

Oczekiwane wypłaty z podjętych decyzji kredytowych
na podstawie modeli wielomianowych (w mln zł)

Decyzje	Model	
	M_1	M_2
Udzielić kredytu	13,5	13,9
Odmówić	-41,2	-41,2

Pomiar korzyści ex post

W celu obliczenia korzyści wynikających z zastosowania modelu wielomianowego do podejmowania decyzji kredytowych można wykorzystać rzeczywiste wartości zmiennych y_t. Uwzględnienie informacji o kategoriach należności poszczególnych kredytów powoduje, że analiza ta ma charakter ex post. W pierwszej kolejności wyznaczono optymalny podział portfela kredytów, co ilustruje tabela 4. Jak wcześniej zauważono, udzielone powinny być wyłącznie kredyty, które w przyszłości należałyby do kategorii należności normalnych. Bank zna postawy klientów wobec spłaty kredytów (stany natury), więc może określić najlepsze decyzje ex post. Na podstawie macierzy wypłat (z tabeli 2) i informacji o optymalnym podziale portfela otrzymuje się maksymalny zysk (Z_{Max}), który wynosi 38,71 mln zł.

Różnica między zyskiem maksymalnym a faktycznym ($Z_{Max} - Z_{Fakt}$) stanowi wielkość dodatkowego zysku, który bank mógłby osiągnąć, gdyby w odniesieniu do badanego portfela mógł powtórnie zastosować model scoringowy do podjęcia decyzji o udzieleniu albo odmowie kredytu. Różnica ta wynosi 23,91 mln zł.

Tabela 4

Optymalny (ex post) podział portfela kredytów i maksymalne wypłaty (w mln zł)

Decyzje	Stany natury				Suma
	j= 1 normalne	j= 2, poniżej standardu	j= 3 wątpliwe	j= 4 stracone	
	Podział portfela				
Udzielić	379,7	0	0	0	379,7
Odmówić	0	13,6	13,9	15,4	42,9
	Maksymalne wypłaty				
Udzielić	40,63	0	0	0	40,63
Odmówić	0	-1,17	-0,75	0	-1,92

Następnie, na podstawie wyników modeli M_1 i M_2, dokonano podziału rozważanych rachunków kredytowych na dwie grupy w zależności od tego, czy powinny zostać otwarte czy nie. Wykorzystano w tym celu formułę (6). W kolejnym kroku, korzystając z macierzy nagród, zaprezentowanej w tabeli 2, dla każdej z czterech grup rachunków kredytowych obliczono korzyści finansowe wynikające z podjęcia decyzji o przyznaniu albo odmowie kredytu.

Tabela 5 przedstawia w ujęciu wartościowym podział badanych rachunków kredytowych w zależności od tego, czy powinny być otwarte czy nie. Gdyby decyzje kredytowe były podejmowane na podstawie modelu M_1, to całkowita wartość udzielonych kredytów byłaby niższa o 19%, czyli o 79,2 mln zł. Wartość udzielonych kredytów w poszczególnych kategoriach zmniejszyłaby się o 13% w pierwszej grupie ryzyka (j=1), 63% w drugiej oraz 79% w trzeciej i czwartej. Wyniki otrzymane dla modelu probitowego (M_2) są lepsze w odniesieniu do kategorii należności normalnych, ale gorsze w przypadku kredytów zagrożonych.

Tabela 5

Podział i portfela kredytowego w zależności od decyzji kredytowej (w mln zł)

Decyzje	Stany natury				Suma
	j= 1 normalne	j= 2, poniżej standardu	j= 3 wątpliwe	j= 4 stracone	
	Model M_1				
Udzielić	332,2 (87%)	5,0 (37%)	2,9 (21%)	3,2 (21%)	343,3 (81%)
Odmówić	47,5 (13%)	8,6 (63%)	11,0 (79%)	12,1 (79%)	79,2 (19%)
Suma	379,7	13,6	13,9	15,4	422,6
	Model M_2				
Udzielić	334,3 (88%)	5,1 (38%)	3,5 (25%)	4,0 (26%)	346,9 (82%)
Odmówić	45,4 (12%)	8,5 (62%)	10,4 (75%)	11,4 (74%)	75,7 (18%)
Suma	379,7	13,6	13,9	15,4	422,6

Tabela 6 prezentuje korzyści finansowe wynikające z podjętych decyzji kredytowych w zależności od kategorii należności (stanów natury). Stanowią one iloczyn elementów tabel 2 i 5. Stosując model t Studenta (M_1), bank osiągnie zysk w kwocie 29,8 mln zł z tytułu udzielenia kredytów i jednocześnie poniesie stratę 6,4 mln zł z tytułu odmowy. Łączny zysk wyniesie 23,4 mln zł i będzie o 0,6 mln zł wyższy niż w przypadku modelu probitowego. Stanowi to 599 zł w przeliczeniu na jeden kredyt o średniej wartości równej około 11 tys. zł. Faktyczny zysk z tytułu udzielenia badanego portfela kredytów wynosi 14,8 mln zł. Zastosowanie modelu M_1 pozwoliłoby zatem na zwiększenie zysku

o dodatkowe 8,6 mln zł, co stanowi 58% przyrost. Jeżeli odniesie się kwotę 8,6 mln zł do różnicy $Z_{Max} - Z_{Fakt} = 23{,}91$ mln zł, to otrzyma się względny miernik efektywności modelu w scoringu. W modelu tym wynosi on 36%. W przypadku modelu probitowego (M_2) łączne korzyści wynikające z klasyfikacji są o 0,6 mln zł mniejsze, czyli o 2,5% w stosunku do specyfikacji M_1. Skuteczność modelu M_2 jest więc niższa i wynosi 33,5%.

Tabela 6

Korzyści finansowe wynikające z zastosowania modeli wielomianowych w scoringu (w mln zł)

Decyzje	Stany natury				Suma
	j= 1 normalne	j= 2, poniżej standardu	j= 3 wątpliwe	j= 4 stracone	
	Model M_1				
Udzielić	35,5	-0,7	-1,5	-3,6	29,8
Odmówić	-5,1	-0,7	-0,6	0,0	-6,4
Suma	30,5	-1,4	-2,0	-3,6	23,4
	Model M_2				
Udzielić	35,8	-0,7	-1,8	-4,4	28,9
Odmówić	-4,9	-0,7	-0,6	0,0	-6,1
Suma	30,9	-1,4	-2,3	-4,4	22,8

Wnioski

W niniejszych badaniach przedstawiono wykorzystanie modelu wielomianowego dla kategorii uporządkowanych w procesie udzielania kredytów w warunkach niepewności. Zastosowanie to przyniosło wiele korzyści z punktu widzenia banku. Umożliwiło ocenę ex ante i ex post skuteczności systemu scoringowego w kategoriach zysków i utraconych korzyści wynikających z podjętych decyzji kredytowych. Użycie modelu o rozkładzie t Studenta przyniosło o 2,5% wyższe korzyści finansowe w odniesieniu do modelu probitowego. Efektywność tego pierwszego modelu wyniosła 36%, drugiego zaś 33,5%.

Spojrzenie na problem konstrukcji systemu scoringowego w kategoriach statystycznej teorii decyzji umożliwia obiektywne porównanie badanych modeli ograniczających ryzyka pojedynczego kredytu. Wówczas pomiar korzyści z ich zastosowania opiera się na przesłankach ekonomicznych, co zaproponowano w niniejszym opracowaniu.

Literatura

1. Albert J., Chib S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. Journal of the American Statistical Association, 88, 669-679.
2. Amemiya T. (1985). Advanced Econometrics. Harvard University Press, Cambridge (Massachusetts).
3. Bensic M., Sarlija N., Zekic-Susac M. (2005). Modelling Small-business Credit Scoring by Using Logistic Regression, Neural Networks and Decision Trees. International Journal of Intelligent Systems in Accounting and Finance Management, 13 (3), 133-150.
4. Chrzanowska M., Witkowska D. (2005). Zastosowanie sztucznych sieci neuronowych do klasyfikacji ryzyka kredytowego podmiotów gospodarczych. W: Metody ilościowe w badaniach ekonomicznych V. SGGW, Warszawa, 80-89.
5. Coombs C.H., Dawes R.M., Tversky A. (1977). Wprowadzenie do psychologii matematycznej. PWN, Warszawa.
6. Cramer J.S. (2004). Scoring Bank Loans That May go Wrong: a Case Study. Statistica Neerlandica, 58, 3, 365-380.
7. DeGroot M. (1981). Optymalne decyzje statystyczne. PWN, Warszawa.
8. Discrimination, Competition, and Loan Performance in FHA Mortgage Lending. (1998). Red. J.A. Berkovec, G.B. Canner, S.A. Gabriel, T.H. Hannan. Review of Economics and Statistics, 80 (2), 241-250.
9. Freixas X., Rochet J.C. (2007). Mikroekonomia bankowa. CeDeWU, Warszawa.
10. Fritz S., Hosemann D. (2000). Restructuring the Credit Process: Behaviour Scoring for German Corporates. International Journal of Intelligent Systems in Accounting, Finance and Management, 9 (1), 9-21.
11. Hodgman D. (1960). Credit Risk and Credit Rationing. Quarterly Journal of Economics, 74, 258-278.
12. Hosmer D., Lemeshow S. (2000). Applied Logistic Regression, Wiley, New York.
13. Jaffee D.M., Russell T. (1976). Imperfect Information, Uncertainty and Credit Rationing. Quarterly Journal of Economics, 4, 651-666.
14. Kapłon R. (2006). Analiza danych dyskretnych w ujęciu retrospektywnym. Badania Operacyjne i Decyzje, 1, 55-72.
15. Maddala G.S (1983). Limited-dependent and Qualitative Variables in Econometrics. Cambridge University Press, Cambridge.
16. Marzec J. (2003). Bayesowska analiza modeli dyskretnego wyboru (dwumianowych). Przegląd Statystyczny, 50, 129-146.

17. Marzec J. (2006). Bayesowski model wielomianowy z rozkładem t Studenta dla kategorii uporządkowanych. W: Metody ilościowe w naukach ekonomicznych. Szóste Warsztaty Doktorskie z Zakresu Ekonometrii i Statystyki. Red. A. Welfe. SGH, Warszawa, 123-144.
18. Marzec J. (2008a). Bayesowska analiza i testowanie modeli dwumianowych z rozkładem t Studenta. Przegląd Statystyczny, 55, 2, 129-146.
19. Marzec J. (2008b). Bayesowskie modele zmiennych jakościowych i ograniczonych w badaniach niespłacalności kredytów. Uniwersytet Ekonomiczny, Kraków.
20. Marzec J. (2009). Zdolności dyskryminacyjne modelu dwumianowego ze skończoną mieszanką rozkładów normalnych w ocenie niespłacalności kredytów. W: Współczesne problemy statystyki, ekonometrii i matematyki stosowanej. Studia i Prace Uniwersytetu Ekonomicznego. Kraków, 3, 129-144.
21. Marzec J. Modele dwumianowe w scoringu kredytowym – pomiar korzyści z ich zastosowania. (maszynopis).
22. Matthews K., Thomson J. (2007). Ekonomika bankowości. PWE, Warszawa.
23. McKelvey R.D., Zavoina W. (1975). A Statistical Model for the Analysis of Ordinary Level Dependent Variables. Journal of Mathematical Sociology, 4, 103-120.
24. Misztal M. (2006). O zastosowaniu metody rekurencyjnego podziału w analizie ryzyka kredytowego. W: Modelowanie preferencji a ryzyko '05. Red. T. Trzaskalik. AE, Katowice, 453-468.
25. Moffatt P.G. (2005). Hurdle Models of Loan Default. Journal of the Operational Research Society, 56, 1063–1071.
26. Mudholkar G., George E. (1978). A Remark on the Shape of the Logistic Distribution. Biometrika, 65, 667-668.
27. Osiewalski J. (2007). Bayesowska statystyka i teoria decyzji w analizie ryzyka kredytu detalicznego. W: Finansowe warunkowania decyzji ekonomicznych. Red. D. Fatuła. Krakowska Szkoła Wyższa, Kraków, 15-28.
28. Osiewalski J., Marzec J. (2004). Model dwumianowy II rzędu i skośny rozkład Studenta w analizie ryzyka kredytowego. Folia Oeconomica Cracoviensia, 45, 63-84.
29. Pratt J.W. (1981). Concavity of the Log Likelihood. Journal of the American Statistical Association, 76, 373, 103-106.
30. Sealey C.W., Lindley J.T. (1977). Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions. The Journal of Finance, 3, 1251-1277.
31. Structural Analysis of Discrete Data with Econometric Applications (1981). Ed. C. Manski, D. McFadden. MIT Press, Cambridge.

32. Thomas L.C. (2000). A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers. International Journal of Forecasting, 16, 149-172.
33. Thomas L.C., Oliver R.W., Hand D.J. (2005). A Survey of the Issues in Consumer Credit Modelling Research. Journal of the Operational Research Society, 56 (9), 1006-1015.
34. Trinkle B., Baldwin A. (2007). Interpretable Credit Model Development Via Artificial Neural Networks. International Journal of Intelligent Systems in Accounting and Finance Management, 15 (3-4). 123-147.
35. Witkowska D., Chrzanowska M. (2005). Wybrane metody klasyfikacji kredytobiorców: modele logitowe i sieci neuronowe. W: Modelowanie preferencji a ryzyko '04. Red. T. Trzaskalik. AE, Katowice, 531-540.
36. Witkowska D., Chrzanowska M. (2006). Drzewa klasyfikacyjne jako metoda grupowania klientów banku. W: Modelowanie preferencji a ryzyko '05. Red. T. Trzaskalik. AE, Katowice, 485-496.

Artur Mikulec
Uniwersytet Łódzki

METODY ROZSZERZONEJ ANALIZY RENTOWNOŚCI INWESTYCYJNEJ OTWARTYCH FUNDUSZY EMERYTALNYCH*

Wstęp

W bogatej literaturze przedmiotu z zakresu analizy inwestycji i zarządzania portfelem, oprócz najbardziej znanych klasycznych wskaźników rentowności inwestycji uwzględniających ryzyko, tj. Sharpe'a, Treynora, α-Jensena [8; 9], odnaleźć można wiele innych wskaźników służących do oceny rentowności i efektywności wyników inwestycyjnych[1]. Studiując literaturę przedmiotu należy zwrócić uwagę na zróżnicowany stopień popularności metod wykorzystywanych w praktyce. Powszechnie prezentowane są wyniki analiz na podstawie miar z wartością narażoną na ryzyko (VaR), przy czym stosunkowo słabo, także w Polsce, rozpowszechniona jest metodologia RiskMetrics, będąca uniwersalnym standardem oceny wyników inwestycyjnych (także umożliwiająca analizę z wykorzystaniem VaR)[2]. Z kolei wskaźnik IR okazuje się niewłaściwy w zakresie oceny efektywności inwestycji, a jego maksymalizacja nie musi świadczyć o wyższych umiejętnościach osób zarządzających portfelem inwestycyjnym [13].

Głównym celem opracowania jest prezentacja mniej znanych oraz rzadko omawianych w polskiej literaturze, a mających już swoje miejsce i znaczenie w historii, wskaźników rentowności i efektywności portfela inwestycji. Wśród około 30 wymienionych wskaźników znalazły się propozycje stosunkowo

* Praca naukowa finansowana ze środków na naukę w latach 2008-2010 jako projekt badawczy Nr N N111 306335.

[1] Ze względu na różne określenia i podziały wskaźników występujące w literaturze warto zwrócić uwagę, że pojęcie rentowności powinno raczej odnosić się do miar oceny działalności inwestycyjnej – skuteczności, które nie uwzględniają ryzyka inwestycyjnego (np. analiza stóp zwrotu). Natomiast termin efektywność, powinien być zarezerwowany dla wskaźników opartych jednocześnie na rentowności i poniesionym ryzyku. Zatem o ile dopuszczalne jest ogólne traktowanie i nazywanie wskaźnika efektywności, wskaźnikiem rentowności (z uwzględnieniem ryzyka), to niewłaściwe jest zastępowanie rentowności pojęciem efektywności.

[2] Miarę ryzyka RiskGrade oraz ryzyko-zwrot ReturnGrade przedstawiono w [10, s. 195-210].

nowych miar oceny działalności inwestycyjnej występujących w literaturze, natomiast część z nich to różnego rodzaju modyfikacje znanych wskaźników, o których jednak warto mówić, gdyż są wśród nich podejścia użyteczne, nieefektywne lub zdezaktualizowane.

W zakresie zewnętrznej oceny działalności inwestycyjnej podmiotów rynku finansowego i kapitałowego (external performance analysis), wyróżnia się między innymi: 1) wskaźniki analizujące ryzyko i zwrot (risk-adjusted performance measures), 2) wskaźniki zysków i strat (gains-to-losses measures), 3) inne miary oceny zewnętrznej (other external measures) oraz 4) szerokie podejście modelowe (market timing models, factor models, style analysis).

W ramach poszczególnych grup wskaźników od 1) do 3), które są tematem niniejszego opracowania można wymienić:

1. Wskaźnik Sharpe'a, αSharpe'a (Sharpe, αSharpe Ratio), Treynora, αTreynora (Treynor, αTreynor Ratio), α-Jensena (α-Jensen Ratio), wskaźnik informacyjny IR (Information Ratio, Appraisal Ratio), M2 oraz M3 (M2, M3 Measure), Sortino (Sortino Ratio), wskaźnik Sharpe'a opaty na VaR (VaR Ratio), warunkowy wskaźnik Sharpe'a (CVaR Ratio), zmodyfikowany wskaźnik Sharpe'a (MVaR Ratio), czy wskaźnik min-max (MinMax Ratio), będący przypadkiem CVaR Ratio.

2. Zysku-straty (Gain-to-Loss Ratio), zysków i strat (Gains-to-Losses), oczekiwanych zysków i strat (Expected Gains-to-Losses), wskaźnik Farinelli-Tibiletti (Farinelli-Tibiletti Ratio), Omega (Omega Ratio), potencjalnych zysków U-P (Upside Potential Ratio), Kappa (Kappa Ratio), wskaźnik Calmara (Calmar Ratio), Sterlinga (Sterling Ratio), Burke'a (Burke Ratio), pewnego ekwiwalentu (Certainty Equivalent), maksymalnego spadku (Maximum Drawdown), wskaźnik Racheva (Rachev Ratio, R-Ratio).

3. Indeks Stutzera (Stutzer Index) – agencji ratingowej Morningstar.

Ze względu na liczne powiązania i zależności pomiędzy wymienionymi wskaźnikami powyższy podział należy traktować umownie, gdyż część wskaźników zysków i strat – przy pewnych założeniach preferencji odnośnie do inwestycji – może być rozważanych w kategorii zwrotu i ryzyka.

1. Wstępna analiza danych

Za pomocą wybranych wskaźników dokonano oceny efektywności inwestowania Otwartych Funduszy Emerytalnych w latach 2000-2008. Na potrzeby analizy skonstruowano portfel rynkowy *M* zawierający dopuszczone dla inwestycji OFE aktywa, których udział w portfelu wzorcowym – stanowiącym punkt odniesienia dla funduszy – określono na podstawie średniego zaangażo-

wania miesięcznych, skumulowanych inwestycji OFE w poszczególne aktywa finansowe w całym w analizowanym okresie. W skład portfela rynkowego wchodziły: obligacje hurtowe (64,0%), indeks WIG (30,6%), stopę oprocentowania WIBID 1M (4,3%) oraz indeksy zagraniczne (razem 1,1%) DAX, FTSE-100 i DJIA. Wstępna analiza danych wykazała, że szeregi czasowe stóp zwrotu OFE na poziomie istotności $\alpha = 0{,}01$ (z wyjątkiem Bankowy OFE) mają rozkład normalny, a szeregi czasowe stóp zwrotu z aktywów wolnych od ryzyka i portfela inwestycyjnego OFE oraz portfela rynku nie są skorelowane.

2. Wskaźniki ryzyko-zwrot

Największą liczbę wariantów i modyfikacji posiada klasyczny wskaźnik Sharpe'a $\left(S_i = \left(\overline{R_i} - \overline{RFR}\right)/\hat{\sigma}_i\right)$. Przyjmując następujące oznaczenia dla wartości miesięcznych stóp zwrotu wyrażonych w skali roku[3]: $\overline{R_i}$, $\overline{R}_M$ – średnia stopa zwrotu portfela inwestycyjnego, odpowiednio *i*-tego OFE oraz portfela rynkowego *M*; $\hat{\sigma}_i$, $\hat{\sigma}_M$ – odchylenie standardowe stóp zwrotu portfela inwestycyjnego, odpowiednio i – tego OFE i portfela rynkowego *M*; $\overline{RFR}$ – średnia stopa zwrotu aktywów wolnych od ryzyka[4], pierwszy z wymienionych wskaźników αSharpe'a, będący miarą różnicową można zapisać w postaci

$$\alpha S_i = \overline{R_i} - \overline{RFR} - \frac{\hat{\sigma}_i}{\hat{\sigma}_M}\left(\overline{R}_M - \overline{RFR}\right) \tag{1}$$

Dla wyników inwestycyjnych OFE lepszych (gorszych) od portfela rynkowego αS_i przyjmuje wartości dodatnie (ujemne), a dzięki swojej konstrukcji umożliwia „natychmiastową" ocenę OFE względem rynku.

Wskaźnik Treynora $\left(T_i = \left(\overline{R_i} - \overline{RFR}\right)/\hat{\beta}_i\right)$ uwzględniający ryzyko systematyczne również można przedstawić w postaci wskaźnika αTreynora za pomocą wzoru (2)

$$\alpha T_i = \overline{R}_i - \overline{RFR} - \frac{\hat{\beta}_i}{\hat{\beta}_M}\left(\overline{R}_M - \overline{RFR}\right) \tag{2}$$

Wyniki porównań OFE w latach 2000-2008 według omówionych dotychczas wskaźników Sharpe'a i Treynora przedstawiono w tabeli 1.

[3] W analizie – o ile nie zaznaczono inaczej – wykorzystano miesięczne logarytmiczne lub średnie logarytmiczne stopy zwrotu poszczególnych OFE wyznaczone na podstawie wartości ich jednostki rozrachunkowej i wyrażone w skali roku – wówczas odchylenie standardowe stóp zwrotu także jest wyrażone w skali roku [4, s. 174-179; 11, s. 1-4].

[4] Ważona stopa zwrotu 52-tygodniowych bonów skarbowych dostępnych dla OFE (rynek pierwotny i wtórny). Ze względu na jej dużą zmienność do obliczeń wymagających średniej wartości stopy zwrotu z aktywów wolnych od ryzyka przyjęto ich wartość środkową – medianę.

Tabela 1

Wyniki efektywności inwestycyjnej OFE według wskaźników αS, S, αT oraz T

OFE	Stopa zwrotu $\overline{R_i}$ (w skali roku w %)	S DEV $\hat{\sigma}_i$ (w skali roku) [a]	αS_i	αS_i POZYCJA	S_i	S_i POZYCJA	β_i	αT_i	αT_i POZYCJA	T_i	T_i POZYCJA
$\overline{RFR}$	6,16	0,00	x	x	x	x	x	x	x	x	x
$\overline{R}_M$	6,38	7,56	0,000	x	0,029	x	1,000	0,000	x	0,220	x
ING	8,98	9,46	2,545	1	0,298	3	1,152	2,567	1	2,448	3
Generali	8,89	8,25	2,490	2	0,331	1	1,008	2,508	2	2,708	1
Polsat	8,78	8,52	2,372	3	0,308	2	1,020	2,396	3	2,570	2
PZU	8,56	8,25	2,160	4	0,291	4	1,007	2,179	4	2,384	4
AXA	8,44	8,16	2,043	5	0,279	6	1,010	2,058	5	2,258	6
Allianz	8,32	7,73	1,935	6	0,279	5	0,934	1,955	6	2,313	5
Nordea	8,24	8,06	1,845	7	0,258	7	0,955	1,870	7	2,178	7
CU	8,20	8,44	1,794	8	0,242	8	1,050	1,809	8	1,943	8
Pekao	8,02	8,19	1,622	9	0,227	9	0,973	1,646	9	1,912	9
WARTA	8,02	8,56	1,611	10	0,217	10	1,020	1,636	10	1,823	10
AIG	7,88	8,17	1,482	11	0,211	11	1,004	1,499	11	1,714	11
Pocztylion	7,83	8,30	1,428	12	0,201	12	1,021	1,445	12	1,635	12
AEGON	7,64	7,72	1,255	13	0,192	13	0,950	1,271	13	1,558	13
Bankowy	7,46	10,51	0,994	14	0,124	14	1,187	1,039	14	1,096	14

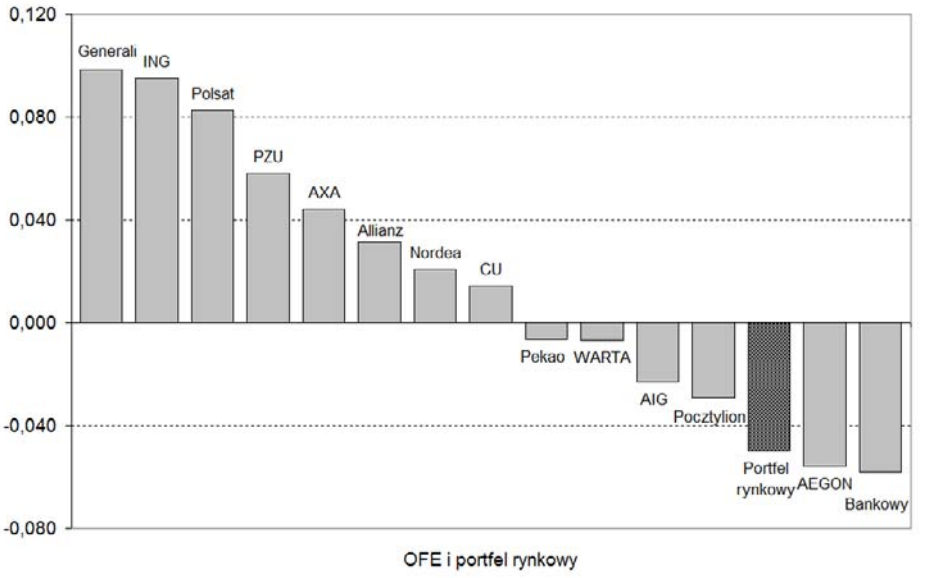
a S DEV – odchylenie standardowe stóp zwrotu.

Klasyczny wskaźnik Sharpe'a opracowany w 1966 roku oparty na wartości oczekiwanej stóp zwrotu i aktywów wolnych od ryzyka został w 1994 roku zmodyfikowany przez autora, a jego nowa wersja dopuszcza zmienne w czasie oprocentowanie aktywów wolnych od ryzyka $\left(RFR_t\right)$ [16, s. 23]. Zmiana powoduje, że średnia nadwyżkowa stopa zwrotu jest liczona na podstawie wartości oczekiwanej różnic stóp zwrotu $\left(R_{i,t} - RFR_t\right)$, a tak uzyskana różnica jest odnoszona nie do ryzyka całkowitego portfela *i*-tego funduszu, lecz ryzyka związanego z osiągnięciem przez niego wspomnianej nadwyżkowej stopy zwrotu $\left(R_{i,t} - RFR_t\right)$, czyli do odchylenia standardowego stóp zwrotu ponad stopę zwrotu aktywów wolnych od ryzyka (wykres 1)

$$S_i^{(94)} = \frac{\overline{R_{i,t} - RFR_t}}{\sigma\left\{R_{i,t} - RFR_t\right\}} \tag{3}$$

Wykres 1

Wskaźnik Sharpe'a (excess-return Sharpe Ratio) dla OFE i rynku
w latach 2000-2008



Kolejny wskaźnik, $M2_i$ definiuje „ekwiwalentną" stopę zwrotu jaką osiągnąłby *i*-ty fundusz, gdyby ryzyko jego portfela inwestycyjnego było równe z ryzykiem portfela rynkowego $\hat{\sigma}_M$ (RAP, Risk-Adjusted Portfolio). Fundusz z najwyższą wartością wskaźnika M2, podobnie jak dla wartości wskaźnika Sharpe'a, powinien mieć najwyższą stopę zwrotu niezależnie od poziomu ryzyka. Omawiany wskaźnik dany jest wzorem

$$M2_i = \hat{\sigma}_M \frac{\overline{R_i} - \overline{RFR}}{\hat{\sigma}_i} + \overline{RFR} \quad \Rightarrow \quad M2_i = \hat{\sigma}_M S_i + \overline{RFR} \tag{4}$$

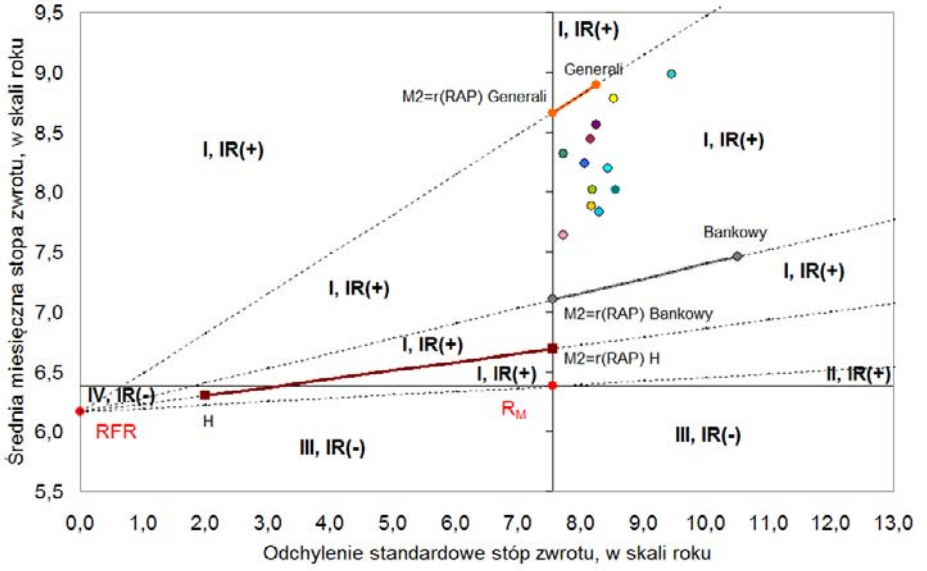
lub przyjmując $d_{i,M} = \hat{\sigma}_M / \hat{\sigma}_i$, jako $M2_i = d_{i,M}\overline{R}_i + (1 - d_{i,M})\overline{RFR}$.

Informacja jaką zawiera wskaźnik M2 daje takie samo uporządkowanie funduszy jak wskaźnik Sharpe'a, gdyż jest on Jest on jego liniową transformacją $M2_i = \hat{\sigma}_M S_i + \overline{RFR}$. M2 nie zawiera żadnej dodatkowej informacji, natomiast stanowi bardziej przyjazną użytkownikowi formę oceny rentowności (z uwzględnieniem ryzyka) ich portfeli względem rynku [12, s. 63-65].

Analogicznie do wzoru (4) można zapisać wskaźnik Treynora uwzględniający ryzyko systematyczne, jako $T2_i$

$$T2_i = \hat{\beta}_M \frac{\overline{R_i} - \overline{RFR}}{\hat{\beta}_i} + \overline{RFR} \quad \Rightarrow \quad T2_i = \hat{\beta}_M T_i + \overline{RFR} \tag{5}$$

a przyjmując $d_{i,M} = \hat{\beta}_M / \hat{\beta}_i$, jako $T2_i = d_{i,M}\,\overline{R}_i + (1 - d_{i,M})\overline{RFR}$



Rys. 1. Profil ryzyko-zwrot oraz wskaźnik M2 dla OFE w latach 2000-2008

Rozszerzonym wskaźnikiem rentowności służącym do wyznaczania ekwiwalentnej stopy zwrotu – skorygowanej nie tylko pod względem oceny ryzyka, lecz również pod względem stopnia korelacji porównywanych portfeli inwestycyjnych – jaką osiągnął *i*-ty fundusz jest wskaźnik $M3_i$ (CAP, Correlation-Adjusted Portfolio) [12, s. 66-69]

$$M3_i = a\overline{R_i} + b\overline{R}_M + (1 - a - b)\overline{RFR} \tag{6}$$

przy czym

$$c = \rho_{i,M}, \qquad TE_t = ?, \qquad \rho_{t,M} = 1 - \frac{TE_t^2}{2\hat{\sigma}_M^2}, \qquad a = \frac{\hat{\sigma}_M}{\hat{\sigma}_i}\sqrt{\frac{\left(1 - \rho_{t,M}^2\right)}{\left(1 - \rho_{i,M}^2\right)}}$$

$$b = \rho_{t,M} - (a)\frac{\hat{\sigma}_i}{\hat{\sigma}_M}(c) \quad \Rightarrow \quad b = \rho_{t,M} - \sqrt{\frac{\left(1 - \rho_{t,M}^2\right)}{\left(1 - \rho_{i,M}^2\right)}}\rho_{i,M}$$

Dysponując danymi na temat $\overline{R}_i$, $\overline{R}_M$, $\hat{\sigma}_i$, $\hat{\sigma}_M$ oraz $c = \rho_{i,M}$ należy przyjąć akceptowalny, docelowy poziom błędu odwzorowania benchmarku przez portfel funduszu (TE_t) i wyznaczyć $\rho_{t,M}$, a następnie oszacować parametry a i b. Wartości wskaźnika $M3_i$ pozwalają na porównywanie funduszy pod względem efektywności inwestycji, gdyż uwzględniają nie tylko skorygowane względem ryzyka (absolutnego i relatywnego) stopy zwrotu z portfeli inwestycyjnych, lecz także korelacje porównywanych portfeli względem portfela rynkowego. Ich zmienność (volatility) jest wówczas taka sama jak portfela rynkowego, gdyż błąd odwzorowania benchmarku $TE_i = TE_t$ dla wszystkich funduszy jest równy docelowemu (tabela 2)[5].

Tabela 2

Wyniki efektywności inwestycyjnej OFE według wskaźników M2 i M3

OFE	Stopa zwrotu $\overline{R}_i$ (w skali roku w %)	S DEV $\hat{\sigma}_i$ (w skali roku) [a]	Współczynnik korelacji $\rho_{i,M}$ (c)	$d_{i,M}$ (w %)	$M2_i$ r(RAP)	$M2_i$ POZYCJA	TE_i (w skali roku)	(a)	(b)	(1-a-b)	$M3_i$ r(CAP)	$M3_i$ POZYCJA
								TE_t =	4,0%	$\rho_{t,M}$ =	0,855	
$\overline{RFR}$	6,16	0,00	(-0,085)	x	x	x	x	x	x	x	x	x
$\overline{R}_M$	6,38	7,56	x	100,0	6,38	x	0	x	x	x	x	x
Generali	8,89	8,25	0,924	91,6	8,67	1	3,16	1,223	-0,373	0,150	9,42	1
Polsat	8,78	8,52	0,905	88,7	8,48	2	3,64	1,065	-0,225	0,160	8,90	5
ING	8,98	9,46	0,920	79,9	8,41	3	3,88	1,041	-0,338	0,297	9,02	3
PZU	8,56	8,25	0,923	91,7	8,36	4	3,18	1,216	-0,364	0,148	9,00	4
Allianz	8,32	7,73	0,914	97,8	8,27	5	3,18	1,230	-0,289	0,059	8,75	7
AXA	8,44	8,16	0,935	92,6	8,27	6	2,89	1,333	-0,485	0,152	9,09	2
Nordea	8,24	8,06	0,896	93,8	8,11	7	3,60	1,078	-0,170	0,092	8,37	8
CU	8,20	8,44	0,940	89,5	7,98	8	2,90	1,340	-0,546	0,206	8,77	6
Pekao	8,02	8,19	0,898	92,3	7,88	9	3,60	1,071	-0,181	0,110	8,11	11
WARTA	8,02	8,56	0,900	88,3	7,80	10	3,73	1,034	-0,194	0,160	8,04	13
AIG	7,88	8,17	0,929	92,6	7,76	11	3,02	1,276	-0,421	0,145	8,26	9
Pocztylion	7,83	8,30	0,930	91,0	7,68	12	3,06	1,265	-0,431	0,166	8,18	10
AEGON	7,64	7,72	0,930	97,9	7,61	13	2,87	1,360	-0,431	0,071	8,08	12
Bankowy	7,46	10,51	0,853	71,9	7,09	14	5,67	0,703	0,026	0,271	7,08	14

a S DEV – odchylenie standardowe stóp zwrotu.

[5] Idea tej miary polega na sprowadzeniu do porównywalności wyników ocenianych portfeli inwestycyjnych pod względem stóp zwrotu, ryzyka i stopnia korelacji z portfelem rynkowym. Uzyskuje się to przez przyjęcie wspólnego dla wszystkich funduszy docelowego poziomu TE. Średnia wartość TE dla OFE wyznaczona na podstawie miesięcznych stóp zwrotu z lat 2000-2008 i wyrażona w skali roku wyniosła 3,5%, a do obliczeń przyjęto umownie poziom 4,0%.

Kolejny wskaźnik, z zakresu analizy efektywności inwestycyjnej OFE to Sortino Ratio (SOR_i). Jest podobny do klasycznego wskaźnika Sharpe'a z tym, że dzieli zmienność stóp zwrotu na „dodatnią" i „ujemną", tzn. odpowiednio powyżej i poniżej pewnego ustalonego progu m. W przypadku wskaźnika Sharpe'a ryzyko jest określane przez odchylenie standardowe (standard deviation) stóp zwrotu wyrażone w skali roku (volatility), a w przypadku wskaźnika Sortino przez część downside-deviation. Wyższa wartość wskaźnika Sortino świadczy o lepszych wynikach inwestycyjnych danego funduszu.

W przypadku downside-deviation minimalna akceptowalna stopa zwrotu (próg m) może być dla celów porównawczych definiowana w sposób dowolny. Miara ryzyka ocenia odchylenie standardowe stóp zwrotu funduszu o wartościach niższych od przyjętego progu m względem tego progu, przy czym $R_{i,t}$ to stopa zwrotu *i*-tego funduszu w momencie t (np. miesięczna), T to całkowita liczba obserwacji szeregu, natomiast d_t to funkcja o wartościach $d_t = 0$, jeżeli $R_{i,t} \geq m$ oraz $d_t = 1$, jeżeli $R_{i,t} < m$.

$$DD(R_i)_m^2 = \frac{1}{T-1}\sum_{t=1}^{T} d_t \left(m - R_{i,t}\right)^2 \Rightarrow DD(R_i)_m = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T} d_t \left(m - R_{i,t}\right)^2} \qquad (7)$$

Oryginalna wersja wskaźnika zaproponowana przez F. Sortino (8) porównuje średnią stopę zwrotu osiągniętą przez fundusz ponad średnią stopę zwrotu aktywów wolnych od ryzyka do jej „ujemnej" zmienności, tj. poniżej stopy $\overline{RFR}$ (wówczas dla wzoru (7) $m = \overline{RFR}$)

$$SOR\left(\overline{R_i}, \overline{RFR}\right) = \frac{\overline{R_i} - \overline{RFR}}{DD(R_i)_{\overline{RFR}}} \qquad (8)$$

Wskaźnik SOR_i wykorzystując DD mierzy bezpieczeństwo zarządzania funduszem, bez wpływu na ten wynik wzrostu cen aktywów. To pozwala w lepszy sposób oceniać jego wynik, gdyż nie jest on „zniekształcony" przez zmienność wynikającą ze wzrostu cen. Można powiedzieć, że inwestor z reguły kojarzy ryzyko ze „złymi" wynikami, tj. ujemnymi stopami zwrotu lub ogólnie ze stopami zwrotu poniżej jego oczekiwań i właśnie w takiej kategorii ocenia wyniki wskaźnik Sortino.

Ogólnie zdefiniowane jednostronne ryzyko oraz szczegóły obliczeniowe i definicja progu m sprawiają, że w literaturze spotkać można różne wersje wskaźnika Sortino, mające różną interpretację[6]. Ogólną wersję wskaźnika Sortino, który do obliczeń wykorzystuje ustalony próg m przedstawia wzór (9), a wyniki uzyskanych analiz tablica 3

[6] Estrada [2, s. 123-124] definiuje wskaźnik Sortino w stosunku do stopy zwrotu z portfela rynkowego. Z kolei A. Chaudhry, H. Johnson, [1, s. 488-489] wykorzystują w liczniku wzoru (9) współczynnik α-Jensena, będący różnicą między stopami zwrotu portfeli inwestycyjnych zarządzanych aktywnie i pasywnie.

$$SOR\left(\overline{R_i}, m\right) = \frac{\overline{R_i} - m}{DD\left(R_i\right)_m} \tag{9}$$

Tabela 3

Ranking OFE w latach 2000-2008 według wskaźnika Sortino

OFE	Sortino $m = \overline{RFR}$ (8)	POZYCJA	Sortino $m = \overline{RFR}$ (w skali roku) (8)	POZYCJA	Sortino $m = 0$ (9)	POZYCJA	Sortino $m = 0$ (w skali roku) (9)	POZYCJA	Sortino $m = \overline{R}_M$ (9)	POZYCJA	Sortino $m = \overline{R}_M$ (w skali roku) (9)	POZYCJA
Generali	0,144	1	0,500	1	0,556	1	1,926	1	0,132	1	0,456	1
ING	0,129	2	0,445	2	0,472	7	1,636	7	0,118	2	0,408	2
Polsat	0,126	3	0,437	3	0,486	6	1,685	6	0,115	3	0,398	3
PZU	0,120	4	0,415	4	0,498	5	1,726	5	0,108	4	0,374	4
Allianz	0,118	5	0,407	5	0,536	2	1,857	2	0,105	5	0,363	5
AXA	0,116	6	0,403	6	0,504	4	1,747	4	0,104	6	0,362	6
Nordea	0,110	7	0,380	7	0,512	3	1,774	3	0,097	7	0,338	7
CU	0,099	8	0,343	8	0,463	10	1,605	10	0,088	8	0,303	8
Pekao	0,094	9	0,324	9	0,47	8	1,629	8	0,082	9	0,284	9
WARTA	0,090	10	0,313	10	0,458	12	1,585	12	0,079	10	0,274	10
AIG	0,084	11	0,291	11	0,445	13	1,543	13	0,073	11	0,252	11
Pocztylion	0,083	12	0,288	12	0,461	11	1,597	11	0,072	12	0,248	12
AEGON	0,077	13	0,265	13	0,465	9	1,611	9	0,065	13	0,224	13
Bankowy	0,054	14	0,185	14	0,351	14	1,215	14	0,044	14	0,153	14

Kolejny wskaźnik służący do oceny rentowności portfela funduszu w kategorii ryzyka i zwrotu to Information Ratio, który wykorzystuje miarę względnego ryzyka Tracking Error, będącą odchyleniem standardowym stóp zwrotu portfela funduszu względem portfela rynkowego. Wskaźnik informacyjny IR jest ilorazem „dodatkowej" stopy zwrotu uzyskanej z portfela inwestycyjnego OFE ponad portfel rynkowy oraz błędu odwzorowania benchmarku TE, wyrażonych w skali roku

$$TE_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(R_{i,t} - R_{M,t}\right)^2} \quad , \quad TE_i = \sqrt{\hat{\sigma}_i^2 - 2\rho_{i,M}\hat{\sigma}_i\hat{\sigma}_M + \hat{\sigma}_M^2} \tag{10}$$

$$IR_i = \hat{\alpha}_i / TE_i \tag{11}$$

Wskaźnik IR jako miara „zmienności" stóp zwrotu osiąganych z inwestycji może służyć do oceny umiejętności osób zarządzających portfelem inwestycyjnym – ma to znaczenie bardziej z punktu widzenia oceny dokony-

wanej przez zarząd czy inwestorów instytucjonalnych. Wskaźnik ten nie posiada jednak informacji przydatnej dla oceny efektywności inwestycji, a posługiwanie się nim w tym kontekście może wprowadzać w błąd [14, s. 28-43]. Niepoprawność wskaźnika IR, bazującego na TE polega na tym, że:

- występuje bezpośredni, dodatni związek pomiędzy TE_i, a $\hat{\sigma}_M$ i $\rho_{i,M}$, dlatego też portfel z najwyższym wskaźnikiem IR może nie być odzwierciedleniem najwyższych umiejętności zarządzania,
- miara IR nie mówi nic na temat wyników inwestycyjnych, tzn. nie daje wskazówek na temat konstrukcji portfela, a wskaźniki typu *RAP* powinny być miarami bezpośredniego „sukcesu" inwestycyjnego,
- inwestor powinien przede wszystkim maksymalizować stopę zwrotu na jednostkę ryzyka całkowitego portfela, a nie wskaźnik IR dla portfela zarządzanego aktywnie,
- wskaźnik informacyjny IR może wprowadzać w błąd, podczas gdy maksymalizacja M2 zawsze prowadzi do poprawnego wyboru portfela, także z punktu widzenia umiejętności osób nim zarządzających. Błędne wskazanie IR ilustruje rys. 1, na którym hipotetyczny portfel H znajdujący się w obszarze „IV, IR(-)" ma ujemną wartość IR i będzie ignorowany przy maksymalizacji IR, a z punktu widzenia ryzyka i zwrotu portfel H jest efektywniejszy, niż portfel rynkowy, czy niejeden inny hipotetyczny portfel z dodatnią wartością IR. Sytuacja odwrotna ma miejsce w obszarze „II, IR(+)" [13, s. 233-244].

3. Wskaźniki zysków i strat

Istnieje duża liczba wskaźników, które porównują „zyski" do „strat" według przyjętych kryteriów. Ich zaletą jest to, że „bezpośrednio dotykają" analizowanych szeregów czasowych stóp zwrotu i nie są abstrakcyjnymi charakterystykami, jak np. volatility. Jednak bardziej skomplikowane wskaźniki – ze względu na wysublimowany sposób pomiaru – mogą stanowić instrument oceny dla wąskiego kręgu osób zarządzających inwestycjami funduszy. Proste wskaźniki rozpatrujące osiągnięcia inwestycyjne w kategoriach zysków i strat to Gain-to-Loss Ratio (liczba wystąpień zysku i straty), Gains-to-Losses (suma zysków i strat) oraz Expected Gains-to-Losses (wartość oczekiwana zysków i strat)[7]:

$$GLR_i = count(G_i)/count(L_i) \tag{12}$$

[7] Wyznaczone przy przyjęciu pewnego progu *m* według którego rozstrzyga się, co stanowi zysk, a co stratę dla funduszu http://www.andreassteiner.net/performanceanalysis.

$$G : L(i) = \sum_{t=1}^{T} G_{i,t} / \sum_{t=1}^{T} L_{i,t} \tag{13}$$

$$E(G_i, L_i) = \frac{\sum_{t=1}^{T} G_{i,t}}{count(G_i)} / \frac{\sum_{t=1}^{T} L_{i,t}}{count(L_i)} \tag{14}$$

gdzie G_i – oznacza zysk (zyski), L_i – oznacza stratę (straty).

Definiując górny i dolny częściowy moment (Upper, Lower Partial Moment) rzędu p i q przy progu m dla analizowanych stóp zwrotu – oznaczone odpowiednio przez $UPM(R_i, p)_m$ i $LPM(R_i, q)_m$ i nazywane jednostronnymi momentami (one-sided moments) – można przejść do wyznaczania bardziej skomplikowanych wskaźników oceny działalności inwestycyjnej funduszy. Jeśli zatem $R_{i,t}$ jest bazową stopą zwrotu w momencie t (np. miesięczna), m ustalonym progiem granicy zysku i straty, T jest całkowitą liczbą obserwacji w szeregu $t = 1,...,T$, a p i q są współczynnikami determinującymi rząd momentu częściowego dla i-tego funduszu

$$UPM(R_i, p)_m = \frac{1}{T-1} \sum_{t=1}^{T} \left(d_t \left(R_{i,t} - m \right)^p \right) \tag{15}$$

przy czym $d_t = 1$, jeżeli $R_{i,t} \geq m$ oraz $d_t = 0$, jeżeli $R_{i,t} < m$ oraz

$$LPM(R_i, q)_m = \frac{1}{T-1} \sum_{t=1}^{T} \left(d_t \left(m - R_{i,t} \right)^q \right) \tag{16}$$

przy czym $d_t = 0$, jeżeli $R_{i,t} \geq m$ oraz $d_t = 1$, jeżeli $R_{i,t} < m$.

Ogólny wzór wskaźników oceniających wyniki inwestycyjne przy zastosowaniu jednostronnej miary ryzyka dla dowolnego momentu częściowego rzędu p i q oraz progu m, definiującym poziom zysku i straty, ma postać wskaźnika Farinelli i Tibiletti

$$\Phi(R_i)_m^{p,q} = \frac{\sqrt[p]{UPM(R_i, p)_m}}{\sqrt[q]{LPM(R_i, q)_m}} \tag{17}$$

Wskaźniki $\Phi(R_i)_m^{p,q}$ dla odpowiedniego rzędu p i q to stosunek „pozytywnych zdarzeń" związanych z osiągnięciem zakładanego zysku do „zdarzeń negatywnych" przynoszących stratę dla i-tego funduszu. Wskaźnik $\Phi(R_i)_m^{p,q}$ nadal może mieć interpretację ekonomiczną w postaci ceny przypadającej na jednostkę ryzyka związanego z osiągnięciem nadwyżkowej stopy zwro-

tu [3, s. 1-15]. Dla ustalonego „punktu odniesienia" m preferowana jest wyższa wartość wskaźnika $\Phi(R_i)_m^{p,q}$, przy czym punktem odniesienia może być $\overline{RFR}$, $\overline{R}_M$, 0 lub inna zaakceptowana wartość.

Szczególnym przypadkiem symetrycznej preferencji odchyleń stóp zwrotu od ustalonego progu m jest wskaźnik Omega (Omega Ratio) [6, s. 59-84]. Przy ustalonym progu m można go zapisać, jako $\Phi(R_i)_m^{1,1}$, a ponieważ $p = q = 1$ ze wzoru (17) otrzymujemy

$$\Phi(R_i)_m^{1,1} = \frac{UPM(R_i,1)_m}{LPM(R_i,1)_m} \tag{18}$$

Wskaźnik Omega ma naturalną interpretację, gdyż jest to stosunek wartości oczekiwanych zysków do wartości oczekiwanych strat, przy czym w odróżnieniu od wskaźnika (14) opiera się na danych szczegółowych szeregu czasowego stóp zwrotu. Jest wiele przekształceń, na podstawie innych wskaźników, które jako wynik końcowy dają w przybliżeniu wartość $\Phi(R_i)_m^{1,1}$. Przede wszystkim istnieje możliwość dokładnego, analitycznego wyznaczenia wartości Omega przy ustalonym progu m za pomocą metody nieparametrycznej, która nie wymaga znajomości rozkładu stóp zwrotu, a opiera się na dystrybuancie empirycznej wartości analizowanych stóp zwrotu. Wartość Omega może być interpretowana w kategorii „ważonego prawdopodobieństwem" wyznacznika zysków i strat przy danym poziomie oczekiwanej stopy zwrotu, tj. ustalonym progu m.

Kolejną wersją wskaźnika, służącego do oceny działalności funduszy w kontekście zysków i strat przy minimalnym, akceptowalnym progu m jest wskaźnik potencjalnych zysków *Upside Potential Ratio* zaproponowany przez F. Sortino [15, s. 15] – wskaźnik $\Phi(R_i)_m^{p,q}$, gdzie $p = 1$, $q = 2$

$$U - P\ Ratio = \frac{UPM(R_i,1)_m}{\sqrt[2]{LPM(R_i,2)_m}} \tag{19}$$

Natomiast uogólnieniem podejścia mierzenia wyników inwestycyjnych w stosunku do „ujemnej zmienności" stóp zwrotu (Downside Risk-Adjusted Performance) jest wskaźnik Kappa, $K(R_i)_m^q$ rzędu q $(q > 0)$ opracowany przez agencję ratingową Morningstar [5, s. 1-17]

$$K(R_i)_m^q = \frac{\overline{R_i} - m}{\sqrt[q]{LPM(R_i,q)_m}} \tag{20}$$

przy czym porównując wzory (9) i (20) można zauważyć, że $SOR\left(\overline{R_i}, m\right) = K\left(R_i\right)_m^2$, gdyż ze wzorów (7) i (16) wynika następująca własność $DD\left(R_i\right)_m = \sqrt[2]{LPM\left(R_i, 2\right)_m}$, natomiast mając na uwadze wzór (20) wzór (18) można zapisać, jako $\Phi\left(R_i\right)_m^{1,1} = K\left(R_i\right)_m^1 + 1$.

Pewną podgrupę wskaźników zysków i strat tworzą trzy wskaźniki: Calmara (Calmar Ratio), Sterlinga (Sterling Ratio) i Burke'a (Burke Ratio) wykorzystujące średnie stopy zwrotu $\overline{R_i}$ i $\overline{RFR}$ oraz ryzyko mierzone przez maximum drawdown (*MDD*), czyli maksymalny spadek stopy zwrotu, tj. największą różnicę pomiędzy najwyższą i najniższą wyrażoną w skali roku stopą zwrotu w analizowanym okresie [7, s. 99-102]

$$Calmar\left(\overline{R_i}, \overline{RFR}\right) = \frac{\overline{R_i} - \overline{RFR}}{MDD} \tag{21}$$

Początek tym wskaźnikom dał pierwszy z wymienionych, Calmar Ratio [18] postaci $\left(Calmar\left(\overline{R_i}\right) = \overline{R_i} / MDD\right)$, przy czym w analizie portfelowej rozpowszechniona jest wersja odnosząca stopę zwrotu $\overline{R_i}$ ponad aktywa wolne od ryzyka $\overline{RFR}$ podana wzorem (21), przez co wszystkie trzy wskaźniki stają się podobne do wskaźników zaproponowanych przez Sharpe'a, czy Sortino. Generalnie pożądane są wyższe wartości wskaźnika Calmara, a do jego obliczeń proponuje się wykorzystywać okres 36 miesięcy.

$$Sterling\left(\overline{R_i}, \overline{RFR}\right) = \frac{\overline{R_i} - \overline{RFR}}{\overline{MDD}_{n=5}} \tag{22}$$

Wskaźnik Sterlinga porównuje średnie wyrażone w skali roku stopy zwrotu z inwestycji i aktywa wolne od ryzyka do średniego poziomu ryzyka wyrażonego przez drawdown. W ogólnej wersji podanej wzorem (22) $\overline{MDD}$ jest średnią z $n = 5$ wartości maksymalnych spadków w całym badanym okresie. Przyjęcie jako miary ryzyka średniej wartości spadków w analizowanym okresie czyni ten wskaźnik mniej wrażliwy na wartości odstające, czy nietypowe[8].

$$Burke\left(\overline{R_i}, \overline{RFR}\right) = \frac{\overline{R_i} - \overline{RFR}}{\sqrt{\sum_{n=1}^{5} \left(MDD\right)^2}} \tag{23}$$

[8] W wersji wskaźnika np. dla trzech ostatnich lat średnia wartość MDD jest określana, jako średnia z „największych" spadków nadwyżkowej stopy zwrotu odnotowanych w każdym roku, pomniejszonych finalnie o próg 10,0%.

Ostatni z tej grupy *Burke Ratio* to stosunek „nadwyżki" średnich stóp stopy zwrotu z inwestycji $\overline{R_i}$ i aktywów wolnych od ryzyka $\overline{RFR}$ (w skali roku) do pierwiastka sumy kwadratów $n = 5$ największych spadków odnotowanych w analizowanym okresie[9] (drawdowns), będących swego rodzaju miarą zaobserwowanego ryzyka. W tabeli 4 zaprezentowano wyniki wybranych wskaźników zysków i strat dla OFE w latach 2000-2008.

Tabela 4

Wyniki wskaźników zysków i strat i indeksu Stutzera dla OFE w latach 2000-2008

OFE	Omega $m = \overline{RFR}$ (18)	POZYCJA	U-P $m = \overline{RFR}$ (19)	POZYCJA	Calmar (w skali roku) (21)	POZYCJA	Sterling (w skali roku) (22)	POZYCJA	Burke (w skali roku) (23)	POZYCJA	Indeks Stutzera $\left(\overline{RFR}\right)$ (24)	POZYCJA
Generali	1,284	1	0,659	1	0,0048	8	0,0050	13	0,0022	12	0,0046	1
Polsat	1,266	2	0,606	9	0,0056	6	0,0063	6	0,0028	6	0,0039	2
ING	1,261	3	0,626	2	0,0072	1	0,0085	1	0,0037	1	0,0037	3
PZU	1,246	4	0,612	6	0,0062	3	0,0073	3	0,0032	3	0,0035	4
Allianz	1,243	5	0,607	8	0,0062	4	0,0071	4	0,0032	4	0,0033	5
AXA	1,237	6	0,612	4	0,0064	2	0,0075	2	0,0033	2	0,0033	6
Nordea	1,221	7	0,612	5	0,0057	5	0,0066	5	0,0029	5	0,0028	7
CU	1,200	8	0,598	10	0,0053	7	0,0062	7	0,0027	7	0,0024	8
Pekao	1,190	9	0,591	11	0,0045	12	0,0052	11	0,0023	11	0,0022	9
WARTA	1,174	10	0,614	3	0,0048	9	0,0055	9	0,0024	8	0,0020	10
AIG	1,173	11	0,575	13	0,0047	10	0,0055	8	0,0024	9	0,0019	11
Pocztylion	1,160	12	0,608	7	0,0046	11	0,0054	10	0,0024	10	0,0017	12
AEGON	1,154	13	0,579	12	0,0044	13	0,0050	12	0,0022	13	0,0015	13
Bankowy	1,107	14	0,559	14	0,0034	14	0,0037	14	0,0017	14	0,0006	14

4. Indeks Stutzera

Spośród innych miar oceny efektywności warto zwrócić uwagę na indeks Stutzera (Stutzer Index), który stanowi alternatywę dla powszechnie stosowanego wskaźnika Sharpe'a. Porządkuje portfele inwestycyjne zgodnie ze wskaźnikiem Sharpe'a, gdy stopy zwrotu mają rozkład normalny oraz uwzględnia przy ocenie asymetrię rozkładu – wartości odstające, skośność, spłaszczenie, gdy stopy zwrotu nie mają rozkładu normalnego. Jest odpowiedni do oceny portfeli inwestycyjnych różnego typu funduszy przez osoby nimi zarządzające, jak i budowy ich rankingów [17, s. 52-61]. Indeks Stutzera I_i jest stopą zwrotu,

[9] http://www.andreassteiner.net/performanceanalysis

przy której prawdopodobieństwo nieosiągnięcia założonego poziomu „nadwyżki", tj. niedoszacowania w analizowanym okresie czasu docelowego poziomu stopy zwrotu będzie zmierzało do zera, czyli prawdopodobieństwo osiągnięcia tej „nadwyżki" będzie największe. Inwestycje funduszy należy uszeregować według malejącej wartości indeksu, który wyznacza się numerycznie (np. Solver) maksymalizując wyrażenie

$$I_i = \max_{\theta}\left(-\ln E\left(e^{\theta \overline{R_i^*}(T)}\right)\right) \Rightarrow \hat{I}_i = \max_{\theta}\left(-\ln\left(\frac{1}{T}\sum_{t=1}^{T} e^{\theta R_{i,t}^*}\right)\right) \qquad (24)$$

gdzie: $\theta < 0$, a $R_{i,t}^*$ to różnica stóp zwrotu *i*-tego portfela oraz benchmarku dla każdego momentu t (tabela 4, poziom docelowy $\overline{RFR}$). Jeśli stopy zwroty mają rozkład normalny to $I_i = 0{,}5\left(S_i^2\right)$, gdzie S_i to wskaźnik Sharpe'a.

Podsumowanie i wnioski

Wymienione na wstępie wskaźniki omówiono według przyjętych grup tematycznych, przechodząc od prostych do coraz bardziej skomplikowanych oraz biorąc pod uwagę kryterium ich popularności. Świadomie zrezygnowano z prezentowania niektórych z nich, np. opartych na *VaR*, których szczegółowy przegląd będzie tematem oddzielnej pracy oraz wskaźnika Certainty Equivalent, bazującego na teorii perspektywy.

Zaprezentowane miary mogą być wykorzystywane do analizy rentowności inwestycyjnej OFE, przy czym w przypadku analiz wielowymiarowych z pewnością zaistnieje konieczność ich merytorycznej selekcji i redukcji z uwagi na występowanie pomiędzy nimi zależności, czy ich zbliżonej interpretacji. Lista wskaźników, które wymieniono w pracy nie jest z pewnością zamknięta, a spośród omówionych na szczególną uwagę zasługują wskaźniki: M2, M3, Omega, U-P oraz indeks Stutzera.

Literatura

1. Chaudhry A., Johnson H. (2008). The Efficacy of the Sortino Ratio and Other Benchmarked Performance Measures Under Skewed Return Distributions. Australian Journal of Management, 32, 3, 485-502.
2. Estrada J. (2006). Downside Risk in Practice. Journal of Applied Corporate Finance, 18, 1, 117-125.
3. Farinelli S., Tibiletti L. (2002). Sharpe Thinking with Asymmetrical Preferences. Social Science Research Network. http://papers.ssrn.com
4. Jajuga K. (2007). Inwestycje. Instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa. PWN, Warszawa.

5. Kaplan P., Knowels J. (2004). Kappa: A Generalized Downside Risk--Adjusted Performance Measure. http://corporate.morningstar.com
6. Keating C., Shadwick W. (2002). A Universal Performance Measure. Journal of Performance Measurement, 63, 59-84.
7. Magdon-Ismail M., Atiya A. (2004). Maximum Drawdown. Risk Magazine, 17, 10, 99-102.
8. Mikulec A. (2004). Zastosowanie wskaźników rentowności portfela inwestycji do oceny działalności funduszy inwestycyjnych akcji (cz. I, cz. II). Nasz Rynek Kapitałowy, 82-86, 6, 84-87, 7.
9. Mikulec A. (2004). Ocena efektywności inwestowania Otwartych Funduszy Emerytalnych. Wiadomości Statystyczne, 9, 26-39.
10. Mikulec A. (2008). Metodologia RiskMetrics jako alternatywne podejście do mierzenia ryzyka aktywów finansowych. W: Modelowanie preferencji a ryzyko '07. Red. T. Trzaskalik. AE, Katowice.
11. Morningstar Methodology: Standard Deviation and Sharpe Ratio. (2005). Methodology Paper. http://corporate.morningstar.com
12. Muralidhar A. (2000). Risk-Adjusted Performance: The Correlation Correction. Financial Analysts Journal, 56, 5, 63-71.
13. Muralidhar A. (2005). Why Maximising Information Ratios is Incorrect. Derivatives Use, Trading & Regulation, 11, 3, 233-244.
14. Muralidhar A. (2002). Skill, History and Risk-Adjusted Performance. Journal of Performance Measurement, 48, 28-43.
15. Satchell S., Sortino F. (2001). Managing Downside Risk. Butterworth--Heinemann, Oxford.
16. Sharpe W. (1998). Morningstar's Risk-Adjusted Ratings. Financial Analysts Journal, 54, 4, 21-33.
17. Stutzer M. (1998). A Portfolio Performance Index. Financial Analysts Journal, 56, 3, 52-61.
18. Young T. (1991). Calmar Ratio: A Smoother Tool. Futures Magazine. http://www.allbusiness.com

Monika Papież
Uniwersytet Ekonomiczny w Krakowie

WPŁYW NIEJEDNORODNOŚCI POPULACJI NA RYZYKO DEMOGRAFICZNE W PORTFELU UBEZPIECZEŃ NA ŻYCIE

Wprowadzenie

Portfele ubezpieczeniowe są tak konstruowane przez aktuariuszy, aby czynniki oddziałujące na polisy w danym portfelu były takie same. Jednak nie zawsze jest to możliwe, gdyż nie zawsze można wyodrębnić wszystkie czynniki ryzyka. Dla portfeli budowanych w ubezpieczeniach majątkowych zostały szeroko rozwinięte modele opisujące np. całkowitą szkodowość, gdy nie są znane wszystkie czynniki ryzyka. Natomiast dla portfeli budowanych w ubezpieczeniach życiowych wykorzystuje się w dużym stopniu modele uwzględniające niejednorodność ryzyka spowodowaną tylko obserwowalnymi czynnikami (np. płeć, wiek, indywidualny styl życia, nawyk palenia), a w małym stopniu modele uwzględniające niejednorodność wynikającą z nieobserwowalnych czynników (np. cechy dziedziczne, genetyczne). Obserwowalne czynniki ryzyka są powszechnie uwzględniane w wycenie produktów ubezpieczeniowych – jest to tzw. underwriting. Natomiast nieobserwowalne czynniki ryzyka są z reguły pomijane.

W ostatnich dwóch dekadach nastąpił duży spadek współczynnika śmiertelności dla osób dorosłych i starszych, stąd do wyceny produktów w ubezpieczeniach na życie powinno się wykorzystywać dynamiczne tablice trwania życia. Do budowy takich tablic życia proponuje się stosować stochastyczne modele opisujące umieralność (np. model Lee-Cartera). Tablice te pozwalają eliminować systematyczną różnicę wynikającą z obserwowanej śmiertelności, a oszacowanej przez aktuariusza. Jednym z powodów różnicy pomiędzy obserwowaną a oszacowaną długością trwania życia może być również nieuwzględnienie nieobserwowalnych czynników.

Opracowanie ma na celu przegląd modeli śmiertelności dla niejednorodnych populacji oraz wykorzystanie tych modeli do oceny ryzyka demograficznego dla portfela ubezpieczeń na życie. Zostanie w nim zaprezentowana między innymi jedna z metod analizy nieobserwowalnych czynników,

a mianowicie modele nazywane „frailty models", czyli modele opisujące zróżnicowanie umieralności. W modelach tego typu zakłada się, że Z jest nieobserwowalną, nieujemną zmienną losową (frailty random variable), która charakteryzuje ryzyko danej grupy spowodowane zróżnicowaniem pod względem umieralności. Modele te zostaną wykorzystane do wyceny jednorazowej składki netto w bezterminowym ubezpieczeniu na życie oraz analizy ryzyka demograficznego związanego z tą wyceną, gdy na badany portfel działają czynniki nieobserwowalne.

1. Czynniki wpływające na niejednorodność populacji

Niejednorodność populacji pod względem umieralności spowodowana jest indywidualnymi różnicami pomiędzy osobami wynikającymi zarówno z obserwowalnych, jak i nieobserwowalnych czynników.

Indywidualne zróżnicowanie trwania życia osoby może być spowodowane czynnikami obserwowalnymi. Czynniki te, wpływające na niejednorodoność populacji, to między innymi czynniki biologiczne i fizjologiczne, takie jak płeć, wiek, otoczenie i środowisko naturalne: strefa klimatyczna, zanieczyszczenie środowiska, gęstość populacji, warunki higieniczne i sanitarne, warunki życia, wykonywany zawód, wykształcenie, indywidualny styl życia, sposób odżywiania, fizyczna aktywność, stan zdrowia.

Natomiast nieobserwowalne czynniki, które wpływają na niejednorodność populacji, to między innymi cechy dziedziczne i genetyczne. Od dawna można zaobserwować, że długowieczność jest cechą dziedziczną. Również część chorób jest dziedziczna lub genetyczna, co oznacza, że w momencie urodzenia wiadomo, że taka osoba w większym stopniu jest narażona na zachorowania [4, s. 153]. Jednak badania genetyczne są zabronione w celu selekcji ubezpieczonych.

2. Metody analizy niejednorodności populacji

Praktyka ubezpieczeniowa, a także literatura ubezpieczeniowa dotycząca ubezpieczeń na życie, w niewielkim stopniu porusza zagadnienia niejednorodności populacji wynikające z nieoberwowalnych czynników.

Od lat 50. XX wieku aktuariusze zdawali sobie sprawę z wpływu niejednorodności populacji spowodowanej nieobserwowalnymi czynnikami na zmiany śmiertelności w ubezpieczonej grupie osób. Baadania nad niejednorodnością populacji w analizie przeżycia zostały przedstawione między innymi w pracy [10]. W pracy tej zostały zaprezentowane dwa podejścia:

- dyskretne, w którym niejednorodność jest modelowana przez mieszankę odpowiednich funkcji,
- ciągłe, w którym do modelowania niejednorodności wykorzystuje się nieobserwowalną nieujemną zmienną losową (frailty random variable) opisującą zróżnicowanie umieralności; modele „fraility", czyli indywidualnego zróżnicowania populacji pod względem umieralności mają za zadanie opisać wszystkie czynniki nieobserwowalne wpływające na przebieg indywidualnej umieralności.

Modele „fraility" – zróżnicowania populacji

W artykule [10] zdefiniowano „fraility", czyli indywidualne zróżnicowanie populacji pod względem umieralności, jako nieujemną zmienną losową Z, która wyraża poziom nieobserwowalnego czynnika ryzyka oddziaływującego na indywidualny przebieg umieralności. Ponadto założono, że osoby z wyższą wartością zmiennej losowej Z umierają średnio wcześniej niż pozostałe, czyli im wyższa wartość zmiennej losowej Z, tym większe prawdopodobieństwo wcześniejszego zgonu, a zatem krótsza długość trwania życia spowodowana nieobserwowanymi czynnikami [5].

Niech x oznacza wiek osób, które tworzą niejednorodną grupę z powodu nieobserwowalnych czynników. Załóżmy ponadto, że dla pojedynczej osoby czynniki te opisane są przez nieujemną zmienną losową opisująca indywidualne zróżnicowanie umieralności. Niech Z_x będzie nieobserwowalna zmienna losowa dla wieku x o rozkładzie ciągłym opisanym funkcją gęstości $g_x(z)$. Zdefiniujmy warunkową funkcję intensywności umieralności dla osoby w wieku x przy danym poziomie indywidualnego zróżnicowania umieralności następującym wzorem

$$\mu_x(z) = \lim_{t\to\infty} \frac{P(T_x \le t \mid Z_x = z)}{t} \tag{1}$$

Oznacza to, że badanie zależności pomiędzy $\mu_x(z)$ a standardową funkcją intensywności umieralności μ_x wymaga analizy łącznego rozkładu zmiennych losowych $(T_x Z_x)$. W pracy [10] został zaproponowany multiplikatywny model dla intensywności umieralności

$$\mu_x(z) = z\mu_x \tag{2}$$

gdzie μ_x można interpretować, jako intensywność umieralności osoby w wieku x, gdy $z = 1$, czyli na osobę nie działają czynniki nieobserwowalne.

Rozkład i charakterystyki zmiennej losowej Z_x

Podstawową charakterystyką nieobserwowalnej zmiennej losowej opisującej zróżnicowanie ze względu na umieralność w populacji będącej w wieku *x* jest wartość oczekiwana zdefiniowana wzorem

$$\bar{z}_x = E(Z_x) = \int_0^\infty z g_x(z) dz \tag{3}$$

Wartość ta opisuje średnie zróżnicowanie badanej populacji pod względem umieralności. Do opisu populacji można też zdefiniować $\bar{\mu}_x(z)$, czyli średnią funkcję intensywności umieralności dla populacji będącej w wieku *x*, wzorem 4, s. 155]

$$\bar{\mu}_x(z) = \int_0^\infty \mu_x(z) g_x(z) dz \tag{4}$$

Korzystając ze wzoru (2) i (3) średnią funkcję intensywności można wyznaczyć w następujący sposób

$$\bar{\mu}_x(z) = \mu_x \bar{z}_x \tag{5}$$

Oznacza to, że średnie zróżnicowanie ze względu na umieralność maleje wraz ze wzrostem wieku oraz intensywność umieralności dla poszczególnych osób rośnie szybciej z wiekiem niż średnia intensywność dla populacji.

Natomiast średnią liczbę osób dożywających wieku *x* można zdefiniować wzorem [3]

$$\bar{S}(x) = \int_0^\infty S(x \mid z) g_0(z) dz \tag{6}$$

gdzie: $S(x \mid z)$ – to warunkowa funkcja przeżycia dana wzorem

$$S(x \mid z) = e^{-\int_0^x \mu_t(z) dt} = e^{-zH(x)} \tag{7}$$

$H(x)$ – to skumulowana funkcja intensywności umieralności postaci

$$H(x) = \int_0^x \mu_t dt \tag{8}$$

Zależność pomiędzy indywidualną umieralnością a umieralnością populacji zależy od rozkładu zmiennej losowej Z_x. Przyjmuje się następujące założenia dotyczącego indywidualnego zróżnicowania. Zakłada się, że parametr indywidualnego zróżnicowania jest wynikiem działania wielu czynników nie-

obeserwowalnych, nie jest stały w ciągu całego życia oraz że kumuluje on czynniki mające wpływ na umieralność [4, s. 154]. Stąd można przyjąć, tak jak sugerują autorzy w pracy [10], żeby w badaniach dotyczących niejednorodności populacji zmienna losowa Z_x miała rozkład gamma. Jeżeli przyjmiemy, że $(x=0)$ wówczas zmienna losowa opisująca zróżnicowanie umieralności w momencie urodzenia Z_0 ma rozkład gamma o parametrach δ i θ, gdzie δ jest parametrem kształtu tego rozkładu a θ jest parametrem skali, zatem $Z_0 \sim Gamma(\delta,\theta)$.

Funkcja gęstości tej zmiennej ma następującą postać

$$g_0(z)=\frac{\theta^{\delta} z^{\delta-1}}{\Gamma(\delta)} e^{-\theta z} \tag{9}$$

a jej wartość oczekiwana i wariancja dane są odpowiednio wzorami

$$E(Z_0)=\bar{z}_0=\frac{\delta}{\theta} \tag{10}$$

$$Var(Z_0)=\frac{\delta}{\theta^2} \tag{11}$$

Na podstawie worów (10) i (11) w współczynnik zmienności wynosi

$$V(Z_0)=\frac{\sqrt{Var(Z_0)}}{E(Z_0)}=\frac{1}{\sqrt{\delta}} \tag{12}$$

Stąd parametr kształtu δ można interpretować również, jako poziom niejednorodności populacji. Jeżeli $\delta \to \infty$ wówczas $V(Z_0)\to 0$, czyli można przyjąć, że populacja jest jednorodna. Jeżeli $\delta \to 0$ to $V(Z_0)\to \infty$, oznacza to, że populacja staje się wysoce niejednorodna.

Przyjęty do analizy rozkład gamma ma następującą własność [3; 4; s. 157]. Jeżeli założymy, że zmienna losowa Z_0 opisująca zróżnicowanie populacji na zachorowanie w momencie urodzenia, czyli w chwili $(x=0)$ ma rozkład gamma, to rozkład wartości zmiennej losowej Z_x opisującej zróżnicowanie populacji na zachorowanie w dowolnym wieku ma również rozkład gamma z takim samym parametrem kształtu jak w momencie $x=0$. Natomiast wartość parametru skali jest zwiększona o skumulowaną intensywność umieralności.

Zatem jeżeli $x>0,$ wówczas $Z_x \sim Gamma(\delta,\theta+H(x))$.

Wówczas funkcja gęstości ma następującą postać

$$g_x(z)=\frac{(\theta+H(x))^{\delta} z^{\delta-1}}{\Gamma(\delta)} e^{-(\theta+H(x))z} \tag{13}$$

a wartość oczekiwana oraz wariancja są wyrażone wzorami

$$E(Z_x) = \bar{z}_x = \frac{\delta}{\theta + H(x)} \tag{14}$$

$$Var(Z_x) = \frac{\delta}{(\theta + H(x))^2} \tag{15}$$

Stąd wartość współczynnika zmienności jest taka sama, jak dla zmiennej losowej Z_0

$$V(Z_x) = \frac{\sqrt{Var(Z_x)}}{E(Z_x)} = \frac{1}{\sqrt{\delta}} \tag{16}$$

Podstawiając do wzoru (6) wartości ze wzoru (7) i (9) wynika, że średnia liczba osób dożywających wieku *x* jest postaci

$$\bar{S}(x) = \left(\frac{\theta}{\theta + H(x)} \right)^{\delta} \tag{17}$$

lub

$$\bar{S}(x) = \left(\frac{\bar{z}_x}{\bar{z}_0} \right)^{\delta} \tag{18}$$

Oznacza to, że średnia liczba osób dożywających wieku *x,* zależy tylko porównania średniej wartości zróżnicowania populacji ze względu na umieralność w wieku *x* i średniej wartości zróżnicowania populacji ze względu na umieralności w wieku 0. Natomiast na jej wartość nie ma wpływu założony rozkład „standardowej" funkcji przeżycia (np. rozkład Gompertza).

Natomiast ze wzoru (5) wynika, że średnia funkcja intensywności umieralności dla populacji w wieku *x* jest wyrażona wzorem

$$\bar{\mu}_x(z) = \mu_x \frac{\delta}{\theta + H(x)} \tag{19}$$

3. Składka i ryzyko dla niejednorodnego portfela ubezpieczeń na życie

Zdefiniujmy zmienną losową $Y^{(j)}$ opisującą obecną wartość świadczenia z bezterminowej polisy na życie na sumę 1, z której wypłata następuje w chwili śmierci *j*-tej osoby w wieku *x* następująco

$$Y^{(j)} = e^{-rT_x^{(j)}} \tag{20}$$

gdzie:

$T_x^{(j)}$ – oznacza zmienną losową opisującą dalsze trwanie życia *j*-tej osoby w wieku *x*,

r – intensywność oprocentowania.

Wielkość portfela firmy ubezpieczeniowej, który jest jednorodny ze względu na obserwowalne ryzyka (np. wiek, płeć) oznaczymy symbolem

$$\Pi = \{j, j = 1,\ldots, n_0\} \tag{21}$$

Jeżeli założymy, że zmienne losowe $\{T_x^{(j)}, j = 1,\ldots n_0\}$ mają taki sam rozkład i są niezależne, to wówczas zmienna losowa

$$Y^{(\Pi)} = \sum_{j\in\Pi} Y^{(j)} \tag{22}$$

opisuje wartość obecną świadczenia z całego portfela dla bezterminowej polisy na życie. Jeżeli dodatkowo założymy, że portfel jest niejednorodny ze względu na nieobserwowalne ryzyka, to wówczas zmienne losowe $\{T_x^{(j)}, j = 1,\ldots n_0\}$ są zależne od nieobserwowalnej zmiennej losowej Z_x, ale zmienne losowe $\{T_x^{(j)}, j = 1,\ldots n_0 \mid Z_x = z\}$ są warunkowo niezależne i mają taki sam rozkład. W tym przypadku postać portfela nie ulegnie zmianie. Dla takiego portfela można wyznaczyć jednorazową składkę oraz ryzyko demograficzne związane z nią. Warunkowa jednorazowa składka netto w bezterminowym ubezpieczeniu na życie dla *j*-tej osoby w wieku *x* przy założeniu wartości indywidualnego zróżnicowania ze względu na umieralność wynosi [3]

$$\overline{A}_{x|z} = E\left(Y^{(j)} \mid Z_x = z\right) = \int_0^\infty e^{-\delta t}\, \frac{S(x+t \mid z)}{S(x \mid z)}\, \mu(x+t \mid z)dt \tag{23}$$

Natomiast jednorazowa składka netto w bezterminowym ubezpieczeniu na życie *j*-tej osoby w wieku *x*, którą można interpretować, jako średnią jednorazową składkę netto w bezterminowym ubezpieczeniu na życie *j*-tej osoby w wieku *x* z niejednorodnej populacji ze względu na nieobserwowalne ryzyka, wynosi

$$\overline{A}_x = E\left(Y^{(j)}\right) = E\left(E\left(Y^{(j)} \mid Z_x = z\right)\right) = \int_0^\infty e^{-\delta t}\, \frac{\overline{S}(x+t)}{\overline{S}(x)}\, \overline{\mu}(x+t)dt \tag{24}$$

Odpowiednio momenty zwykłe drugiego rzędu wynoszą

$$^2\overline{A}_{x|z} = E\left(\left(Y^{(j)} \mid Z_x = z\right)^2\right) = \int_0^\infty e^{-2\delta t}\, \frac{S(x+t \mid z)}{S(x \mid z)}\, \mu(x+t \mid z)dt \tag{25}$$

$$^2\overline{A}_x = E\left(\left(Y^{(j)}\right)^2\right) = \int_0^\infty e^{-2\delta t}\,\frac{\overline{S}(x+t)}{\overline{S}(x)}\,\overline{\mu}(x+t)dt \tag{26}$$

Wobec tego wariancje odpowiednich jednorazowych składek netto dla *j*-tej osoby w wieku *x* można wyznaczyć ze wzorów

$$Var\left(Y^{(j)} \mid Z_x = z\right) = {}^2\overline{A}_{x|z} - \left(\overline{A}_{x|z}\right)^2 \tag{27}$$

$$Var\left(Y^{(j)}\right) = {}^2\overline{A}_x - \left(\overline{A}_x\right)^2 \tag{28}$$

Wartość jednorazowej składki netto dla całego portfela Π wynosi

$$\overline{A}_x^{(\Pi)} = E\left(Y^{(\Pi)}\right) = \sum_{j\in\Pi} E\left(Y^{(j)}\right) = nE\left(Y^{(j)}\right) = n\overline{A}_x \tag{29}$$

a jej wariancja

$$Var\left(Y^{(\Pi)}\right) = E\left(\left(Y^{(\Pi)}\right)^2\right) - \left(E\left(Y^{(\Pi)}\right)\right)^2 \tag{30}$$

lub, wykorzystując metodę łącznego podziału opartą na analizie wariancji, wariację dla całego portfela można przedstawić w następujący sposób [6; 7]

$$\begin{aligned} Var\left(Y^{(\Pi)}\right) &= E\left(Var\left(Y^{(\Pi)} \mid Z_x = z\right)\right) + Var\left(E\left(Y^{(\Pi)} \mid Z_x = z\right)\right) = \\ &= nE\left(Var\left(Y^{(j)} \mid Z_x = z\right)\right) + n^2 Var\left(E\left(Y^{(j)} \mid Z_x = z\right)\right) \end{aligned} \tag{31}$$

Jeżeli założymy, że populacja jest jednorodna, wówczas pierwszy składnik sumy (31) – $E\left(Var\left(Y^{(\Pi)} \mid Z_x = z\right)\right)$ redukuje się do wartości $Var\left(Y^{(\Pi)}\right)$, a drugi składnik wynosi zero. Tak więc pierwszy składnik można traktować, jako składnik ryzyka demograficznego, który mierzy ryzyko spowodowane przypadkowymi odchyleniami od wartości oczekiwanej umieralności, czyli mierzy ryzyko ubezpieczeniowe. Natomiast, gdy założymy niejednorodność populacji w portfelu wynikającą z czynników nieobserwowalnych, to drugi składnik sumy $Var\left(E\left(Y^{(\Pi)} \mid Z_x = z\right)\right)$ mierzy systematyczne odchylenia pomiędzy rzeczywistą liczbą zgonów a oczekiwaną wynikającą z działania czynników nieobserwowalnych.

Do wyznaczenia udziału ryzyka całego portfela mierzonego odchyleniem standardowym w stosunku do jednorazowej składki netto można obliczyć indeks ryzyka w zależności od liczby polis w portfelu

$$r(n) = \frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)} = \sqrt{\frac{E\left(Var\left(Y^{(j)} \mid Z_x = z\right)\right)}{nE^2\left(Y^{(j)}\right)} + \frac{Var\left(E\left(Y^{(j)} \mid Z_x = z\right)\right)}{E^2\left(Y^{(j)}\right)}} \tag{32}$$

oraz wielkość indeksu ryzyka, gdy liczba polis w portfelu rośnie do nieskończoności

$$\lim_{n\to\infty} r(n\ ) = \sqrt{\frac{Var\left(E\left(Y^{(j)} \mid Z_x = z\right)\right)}{E^2\left(Y^{(j)}\right)}} \tag{33}$$

Ze wzorów (32) i (33) wynika, że udział pierwszego składnika sumy (31) w składce może być zmniejszony, gdy zostanie zwiększona wielkość portfela. Natomiast udział drugiego składnika w składce nie ulega zmianie wraz ze wzrostem liczby polis w portfelu. Ze wzoru (33) wynika, że wraz ze wzrostem liczby polis w portfelu udział całkowitego ryzyka demograficznego w składce redukuje się tylko do udziału ryzyka związanego z niejednorodnością populacji. Udział ryzyka ubezpieczeniowego redukuje się wówczas do zera.

4. Analiza składki i ryzyko dla niejednorodnego portfela ubezpieczeń na życie – przykład empiryczny

W celu analizy wpływu niejednorodności populacji na wysokość jednorazowej składki w bezterminowym ubezpieczeniu na życie oraz wpływu na ryzyko demograficzne w portfelu ubezpieczeń na życie przyjmijmy, że w rozpatrywanym portfelu indywidualna intensywność umieralności opisana jest za pomocą funkcji Gompertza. Z badań empirycznych wynika, że funkcja ta dobrze opisuje intensywność zgonów w populacji osób w wieku od 30 do 90 lat. Intensywność umieralności rozkładu Gompertza opisana jest wzorem

$$\mu_x = \alpha e^{\beta x} \tag{34}$$

a skumulowana intensywność umieralności jest postaci

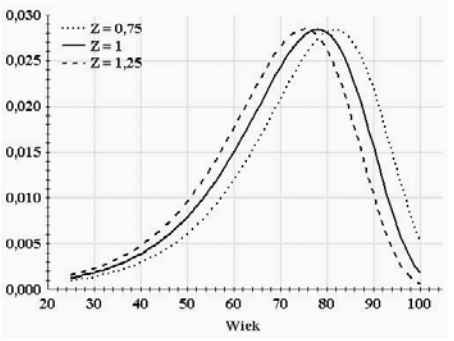
$$H(x) = \int_0^x \alpha e^{\beta t}\, dt = \frac{\alpha}{\beta}\left(e^{\beta x} - 1\right) \tag{35}$$

Wykorzystując wzór (5) średnia intensywność umieralności dla populacji przy założeniu, że Z_x ma rozkład gamma wynosi

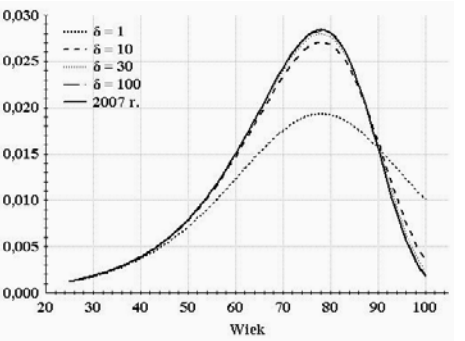
$$\overline{\mu}_x(z) = \frac{\alpha\delta e^{\beta x}}{\left(\theta - \frac{\alpha}{\beta}\right) + \frac{\alpha}{\beta} e^{\beta x}} \tag{36}$$

Do dalszej analizy parametry rozkładu Gompertza oszacowano na podstawie tablic trwania życia dla mężczyzn w Polsce w 2007 roku. Oszacowania parametrów dokonano za pomocą programu Statistica 8.0. Oszacowane wartości wynoszą odpowiednio: $\alpha = 0{,}0001878$ i $\beta = 0{,}07713$. Wyniki dotyczące wysokości jednorazowej składki w bezterminowym ubezpieczeniu na życie

oraz wartości ryzyka związanego z tą składką znajdują się w tabelach 1-5. Rysunek 1 przedstawia warunkowe funkcje gęstości $f(x|z)$ w zależności od wartości zmiennej losowej Z_x. Do dalszej analizy przyjęto następujące wartości zmiennej Z_x: $z = 0{,}75$ (obniżona umieralność), $z = 1$ („standardowa" umieralność) i $z = 1{,}25$ (zwiększona podatność na choroby). Z wykresów wynika, że im wyższa wartość zmiennej losowej, tym modalna jest dla wyższego wieku.



Rys. 1. Wykres warunkowych funkcji gęstości $f(x|z)$ w zależności od wartości zmiennej losowej Z_x: $z = 0{,}75$ (obniżona umieralność), $z = 1$ („standardowa" umieralność) i $z = 1{,}25$ (zwiększona podatność na choroby)



Rys. 2. Wykres funkcji gęstości $\bar{f}_x(t)$ zmiennej losowej T_x w zależności od parametru δ oraz wykres funkcji gęstości na podstawie oszacowanych wartości dla 2007 roku

Jeżeli założymy, tak jak proponuje literatura [3], że średnie zróżnicowanie indywidualne umieralności w chwili urodzenia jest równe 1, czyli $E(Z_0) = \frac{\delta}{\theta} = 1$, wówczas parametr skali jest równy parametrowi kształtu w rozkładzie gamma ($\delta = \theta$). Wobec tego rys. 2 przedstawia wykresy funkcji gęstości postaci $\bar{f}_x(t)$ zmiennej losowej T_x w zależności od parametru δ. Im niższe wartości δ, tym większa niejednorodność populacji, a im wyższe wartości δ, tym mniejsza. W pracach [1; 2] autorzy sugerują, że do wyceny ubezpieczeń na życie powinno zakładać się wartości δ z przedziału $(20{,}40)$.

Do analizy wysokości jednorazowej składki w bezterminowym ubezpieczeniu na życie i wielkości ryzyka demograficznego w portfelu ubezpieczeń na życie w zależności od indywidualnego zróżnicowania osób ze względu na umieralność oraz stopnia niejednorodności populacji ze względu na nieobserwowalne czynniki przyjmijmy wysokość intensywności oprocentowania na poziomie $r = 0{,}0198$. Tabela 1 przedstawia wartości jednorazowej składki

netto w bezterminowym ubezpieczeniu na życie w zależności od wieku *x* przy założeniu wartości indywidualnego zróżnicowania ze względu na umieralność oraz wartość ryzyka mierzonego odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto dla pojedynczej polisy. Do analizy założono różne wartości zmiennej losowej opisującej zróżnicowanie indywidualne pod względem umieralności. Przyjęto następujące wartości zmiennej: $z = 0{,}75$, co oznacza zwiększoną indywidualną odporność na zachorowanie, $z = 1{,}25$ – oznacza zwiększoną indywidualną podatność na choroby oraz $z = 1$, które można traktować jako „standardową" umieralność. Analiza wartości jednorazowej składki netto wskazuje, że wzrasta ona zarówno z wiekiem, jak i ze wzrostem wartości nieobserwowalnej zmiennej losowej Z_x opisującej indywidualne zróżnicowanie umieralności. Natomiast ryzyko mierzone odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto dla pojedynczej polisy maleje wraz ze wzrostem wieku i wartości zmiennej losowej Z_x.

Tabela 1

Wartości jednorazowej składki netto w bezterminowym ubezpieczeniu na życie w zależności od wieku *x* przy założeniu wartości indywidualnego zróżnicowania ze względu na umieralność $z = 0{,}75$, $z = 1$ i $z = 1{,}25$ oraz wartość współczynnika zmienności dla pojedynczej polisy

	$\overline{A}_{x\|z}$			$\frac{\sqrt{Var\left(Y^{(j)} \mid Z_x = z\right)}}{E\left(Y^{(j)} \mid Z_x = z\right)}$ w [%]		
Wiek	$z = 0{,}75$	$z = 1$	$z = 1{,}25$	$z = 0{,}75$	$z = 1$	$z = 1{,}25$
25	0,3893	**0,4166**	0,4387	33,69	**32,58**	31,66
30	0,4262	**0,4553**	0,4788	32,18	**30,96**	29,95
35	0,4655	**0,4963**	0,5210	30,52	**29,19**	28,10
40	0,5071	**0,5394**	0,5651	28,71	**27,27**	26,10
45	0,5506	**0,5840**	0,6104	26,76	**25,22**	23,97
50	0,5956	**0,6297**	0,6562	24,68	**23,05**	21,75
55	0,6413	**0,6755**	0,7018	22,48	**20,79**	19,45

Tabela 2 przedstawia wartości jednorazowej składki netto przy założeniu niejednorodności populacji w portfelu ze względu na nieobserwowalne ryzyka przy założeniu różnych parametrów kształtu rozkładu gamma będącego rozkładem zmiennej losowej opisującej zróżnicowanie populacji ze względu na umieralność. Do analizy założono następujące wartości parametru kształtu: $\delta = 1$ (oznacza to, że populacja jest wysoce niejednorodna), $\delta = 10$; $\delta = 30$; $\delta = 100$. Tabela 2 zawiera również wartości ryzyka mierzonego odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto dla pojedynczej polisy.

Tabela 2

Wartości jednorazowej składki netto przy założeniu niejednorodności populacji w portfelu ze względu na nieobserwowalne ryzyka oraz wartość współczynnika zmienności dla pojedynczej polisy

	$\overline{A}_x$				$\frac{\sqrt{Var\left(Y^{(j)}\right)}}{E\left(Y^{(j)}\right)}$ [%]			
Wiek	$\delta = 1$	$\delta = 10$	$\delta = 30$	$\delta = 100$	$\delta = 1$	$\delta = 10$	$\delta = 30$	$\delta = 100$
25	0,3752	0,4124	0,4152	0,4162	42,86	33,51	32,89	32,67
30	0,4093	0,4507	0,4538	0,4548	41,32	31,90	31,27	31,05
35	0,4452	0,4912	0,4946	0,4958	39,67	30,14	29,50	29,28
40	0,4823	0,5336	0,5375	0,5388	37,91	28,24	27,59	27,36
45	0,5202	0,5776	0,5819	0,5834	36,07	26,20	25,54	25,31
50	0,5581	0,6224	0,6273	0,6289	34,18	24,06	23,38	23,15
55	0,5951	0,6674	0,6728	0,6747	32,29	21,83	21,13	20,89

Analiza wartości w tabeli 2 wskazuje, że wraz ze wzrostem wartości parametru kształtu oraz wysokości wieku wzrasta wysokość jednorazowej składki i maleje ryzyko mierzone odchyleniem standardowym w stosunku do składki.

Tabela 3 przedstawia stosunek wielkości jednorazowej składki netto dla portfela o jednorodnej populacji ($z = 1$) do wielkości jednorazowej składki netto dla portfela niejednorodnego w zależności od wartości parametru δ oraz wieku wyznaczony wzorem: $\frac{\overline{A}_{x|z=1}}{\overline{A}_{x(\delta)}} - 1$. Należy zauważyć, że różnica pomiędzy wielkością jednorazowej składki netto dla portfela jednorodnego i niejednorodnego wzrasta wraz ze wzrostem wieku oraz im wyższa wartość parametru δ tym portfel staje się bardziej jednorodny oraz wartości składek się wyrównują.

Tabela 3

Stosunek wielkości jednorazowej składki netto dla portfela o jednorodnej populacji ($z = 1$) do wielkości jednorazowej składki netto dla portfela niejednorodnego w zależności od wartości parametru δ oraz wieku w (%)

Wiek	$\delta = 1$	$\delta = 10$	$\delta = 30$	$\delta = 100$
25	11,03	1,01	0,33	0,10
30	11,23	1,03	0,34	0,10
35	11,49	1,05	0,35	0,10
40	11,83	1,08	0,36	0,11
45	12,27	1,11	0,37	0,11
50	12,82	1,16	0,38	0,11
55	13,51	1,22	0,40	0,12

W tabeli 4 zostało przedstawione porównanie ryzyka mierzonego odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto dla portfela o jednorodnej populacji ($z=1$) oraz niejednorodnej (δ=30) w zależności od liczby polis w portfelu i wieku. Analiza tych wartości wskazuje, że udział jednorazowej składki netto w odchyleniu standardowego jest wyższy, gdy portfel jest niejednorodny. Oznacza to, że gdy na populację działają nieobserwowalne czynniki, to ryzyko związane ze składką wzrasta. Ponadto, wraz ze wzrostem liczby polis w portfelu ryzyko to redukuje się do zera w przypadku populacji jednorodnej, a w wypadku populacji niejednorodnej nie można go całkowicie zredukować.

Tabela 4

Porównanie ryzyka mierzonego odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto dla portfela o jednorodnej populacji ($Z_x=1$) oraz niejednorodnej ($\delta=30$) w zależności od liczby polis w portfelu i wieku

Wiek	$x=25$		$x=40$		$x=55$	
N	$\frac{\sqrt{Var\left(Y^{(\Pi)} \mid Z_x\right)}}{E\left(Y^{(\Pi)} \mid Z_x\right)}$ $z=1$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$ $\delta=30$	$\frac{\sqrt{Var\left(Y^{(\Pi)} \mid Z_x\right)}}{E\left(Y^{(\Pi)} \mid Z_x\right)}$ $z=1$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$ $\delta=30$	$\frac{\sqrt{Var\left(Y^{(\Pi)} \mid Z_x\right)}}{E\left(Y^{(\Pi)} \mid Z_x\right)}$ $z=1$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$ $\delta=30$
1	32,58	32,89	27,27	27,59	20,79	21,13
10	10,30	11,17	8,62	9,47	6,57	7,36
100	3,26	5,38	2,73	4,75	2,08	3,85
1000	1,03	4,41	0,86	3,98	0,66	3,30
100000	0,10	4,29	0,09	3,89	0,07	3,24

Tabele 5a-5c przedstawiają porównanie wielkości ryzyka ubezpieczeniowego i ryzyka związanego z niejednorodnością populacji w stosunku do całkowitego ryzyka oraz indeksu ryzyka w zależności od liczby polis w portfelu dla osób w wieku 25, 40 i 55 lat. Analiza wartości w tabeli 5a-5c (kolumny 1 i 2) wskazuje, że wraz ze wzrostem liczby polis w portfelu niezależnie od wieku ulega zmianie rodzaj ryzyka oddziaływujący na całkowite ryzyko demograficzne. W wypadku jednej polisy całkowite ryzyko demograficzne w ok. 98% zależy od przypadkowych odchyleń pomiędzy oszacowaną, a rzeczywistą długością trwania życia i tylko w ok. 2% zależy od niejednorodności populacji. Jeżeli liczba polis w portfelu wzrośnie, to wówczas udział ryzyka ubezpieczeniowego w ryzyku demograficznym redukuje się prawie do zera. Natomiast wzrasta udział ryzyka związanego z niejednorodnością populacji.

Analiza wartości indeksu ryzyka (kolumny 3-5), czyli wielkości ryzyka mierzonego odchyleniem standardowym w stosunku do wielkości jednorazowej składki netto w ubezpieczeniu na życie wskazuje, że wraz ze wzrostem wieku maleje wartość tego indeksu niezależnie od liczby polis oraz rodzaju ryzyka. Również niezależnie od liczby polis indeks ryzyka związanego z niejednorodnością populacji nie ulega zmianie, a jedynie zmniejsza się wraz ze wzrostem wieku, w którym jest wystawiana jednorazowa składka netto. Tak więc zmiana liczby polis nie powoduje, że wielkość ryzyka związanego z niejednorodnością populacji w stosunku do składki ulegnie zmianie (tabele 5a-5c). Natomiast wraz ze wzrostem liczby polis maleje wielkość ryzyka ubezpieczeniowego w stosunku do składki. Redukuje się ono prawie do zera. Podobnie maleje wielkość ryzyka demograficznego w stosunku do składki, ale tylko do poziomu udziału ryzyka związanego z niejednorodnością populacji w składce.

Tabela 5a

Porównanie wielkości ryzyka ubezpieczeniowego i ryzyka związanego z niejednorodnością populacji w stosunku do całkowitego ryzyka oraz indeksu ryzyka w zależności od liczby polis w portfelu dla osób w wieku 25 lat w (%)

N	$\frac{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$
1	98,30	1,70	32,61	4,28	32,89
10	85,28	14,72	10,31	4,28	11,17
100	36,67	63,33	3,26	4,28	5,38
1000	5,47	94,53	1,03	4,28	4,41
100000	0,06	99,94	0,10	4,28	4,29

Tabela 5b

Porównanie wielkości ryzyka ubezpieczeniowego i ryzyka związanego z niejednorodnością populacji w stosunku do całkowitego ryzyka oraz indeksu ryzyka w zależności od liczby polis w portfelu dla osób w wieku 40 lat w (%)

N	$\frac{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$
1	98,02	1,98	27,31	3,89	27,59
10	83,17	16,83	8,64	3,89	9,47
100	33,07	66,93	2,73	3,89	4,75
1000	4,71	95,29	0,86	3,89	3,98
100000	0,05	99,95	0,09	3,89	3,89

Tabela 5c

Porównanie wielkości ryzyka ubezpieczeniowego i ryzyka związanego z niejednorodnością populacji w stosunku do całkowitego ryzyka oraz indeksu ryzyka w zależności od liczby polis w portfelu dla osób w wieku 55 lat w (%)

N	$\frac{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}{Var\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{E\left(Var\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(E\left(Y^{(\Pi)} \mid Z_x\right)\right)}}{E\left(Y^{(\Pi)}\right)}$	$\frac{\sqrt{Var\left(Y^{(\Pi)}\right)}}{E\left(Y^{(\Pi)}\right)}$
1	97,65	2,35	20,88	3,24	21,13
10	80,62	19,38	6,60	3,24	7,36
100	29,38	70,62	2,09	3,24	3,85
1000	3,99	96,01	0,66	3,24	3,30
100000	0,04	99,96	0,07	3,24	3,24

Uwagi końcowe

Analiza wysokości jednorazowej składki w bezterminowym ubezpieczeniu na życie i wielkości ryzyka demograficznego w portfelu ubezpieczeń na życie w zależności od indywidualnego zróżnicowania osób ze względu na umieralność oraz stopień niejednorodności populacji ze względu na nieobserwowalne czynniki wskazuje, że nieuwzględnienie tych czynników ma wpływ na wysokość składki, jak i ryzyka z nią związanego. Wyniki empiryczne potwierdzają fakt, iż częściową redukcję ryzyka demograficznego można dokonać zwiększając liczbę polis w portfelu. Jednak całkowita redukcja jest niemożliwa, gdyż niezależnie od liczby polis udział ryzyka związanego z niejednorodnością populacji jest stały.

Literatura

1. Butt Z., Haberman S. (2002). Application of Frailty-based Mortality Models to Insurance Data. Actuarial Research Paper No. 142, Department of Actuarial Science & Statistics, City University, London.
2. Butt Z., Haberman S. (2004). Application of Frailty-based Mortality Models Using Generalized Linear Models. ASTIN Bulletin, 34(1), 175-197.
3. Olivieri A. (2006). Heterogeneity in Survival Models. Applications to Pensions and Life Annuities. Belgian Actuarial Bulletin, 6, 1, 23-39.
4. Ostasiewicz S., Dębicka J., Jasiulewicz H., Mazurek E., Ostasiewicz W. (2000). Metody oceny i porządkowania ryzyka w ubezpieczeniach życiowych. AE, Wrocław.

5. Papież M., Wanat S. (2006). Wybrane metody analizy i modelowania zależnych zmiennych losowych wykorzystywanych w ubezpieczeniach. AE, Kraków.
6. Papież M. (2008). Ryzyko długowieczności dla portfela ubezpieczeń na życie i zakładów emerytalnych. W: Ubezpieczenia w XXI wieku. Red. W. Ronka-Chmielowiec. AE, Wrocław.
7. Papież M. (2008). Metody oceny ryzyka demograficznego i inwestycyjnego dla portfela ubezpieczeń na życie. W: Modelowanie preferencji a ryzyko `07. Red. T. Trzaskalik. AE, Katowice, 361-376.
8. Pitacco E. (2004a). Survival Models in a Dynamic Context: a Survey. Insurance: Mathematics and Economics, 35, 279-298.
9. E. Pitacco (2004b). From Halley to „Frailty": A Review of Survival Models for Actuarial Calculations. Giornale dell'Istituto Italiano degli Attuari, LXVII (1-2), 17-47.
10. Vaupel J.W., Manton K.G., Stallard E. (1979). The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. Demography, 16(3), 439-454.

Stanisław Wieteska
Uniwersytet Łódzki

UBEZPIECZENIE ODPOWIEDZIALNOŚCI CYWILNEJ ZA PRODUKTY ŻYWNOŚCIOWE POCHODZĄCE Z ROŚLIN GENETYCZNIE ZMODYFIKOWANYCH

Wprowadzenie

Już od początku lat 90. obserwujemy procesy w rolnictwie mające na celu zwiększenie wydajności produkcji rolniczej. Jednym ze sposobów zwiększenia wydajności produkcji rolniczej jest produkcja roślin genetycznie zmodyfikowanych (GMO). Produkcja tej roślinności ma na celu nie tylko podwyższenie wartości odżywczych żywności, ale także komercyjny jej charakter, tzn. jej wzrost sprzedaży na inne rynki światowe.

Wydawać by się mogło z jednej strony, że cele stawiane są bardzo ambitne, jednak z drugiej spotykamy głosy sprzeciwu, jako że żywność pochodząca z tych roślin jest szkodliwa dla zdrowia. Argumenty po jednej i drugiej stronie są przekonujące na tyle, że proces ten jest rozszerzany pod warunkiem stosowania ostrych przepisów o ochronie konsumenta.

Ponieważ całkowite uniknięcie ryzyka nie jest możliwe, koniecznością jest więc zaproponowanie instrumentu ekonomicznego, który by chronił procesy produkcji żywności od przypadków losowych, a także ludność przed negatywnymi konsekwencjami ich spożycia. Takim instrumentem może być ubezpieczenie odpowiedzialności cywilnej za produkty żywnościowe.

Celem tego opracowania jest przedstawienie podstawowych informacji na temat GMO, a także zaproponowanie ubezpieczenia jej producentów i przetwarzających produkty roślinne genetycznie zmodyfikowane.

Opracowanie jest jednym z pierwszych głosów włączających się w dyskusję nad problematyką GMO. Z punktu widzenia ubezpieczeń majątkowych, myślą przewodnią jest propozycja nowego produktu ubezpieczeniowego.

1. Definicja GMO

Skrót GMO pochodzi z języka angielskiego – Genetically Modified Organisms. GMO definiuje się jako organizmy zmodyfikowane genetycznie: rośliny, zwierzęta i drobnoustroje, których geny zostały celowo zmienione przez człowieka. Manipulowanie kodami DNA rośliny i inne podobne techniki pozwalają tworzyć organizmy o odmiennych właściwościach niż macierzysty gatunek. Ma to na celu maksymalizację plonów przy jak najmniejszy kosztach.

Ustawa z 22 czerwca 2001 roku definiuje GMO jako: „organizm zmodyfikowany genetycznie to organizm inny niż organizm człowieka, w którym materiał genetyczny został zmieniony w sposób niezachodzący w naturalnych warunkach krzyżowania lub naturalnej rekombinacji".

Według Rozporządzenia (WE) nr 1829/2005 Parlamentu Europejskiego i Rady, jeśli środek spożywczy zawiera GMO ponad przewidzianą przepisami normę, powinien być oznaczony jako genetycznie zmodyfikowana żywność.

Według normy KT287 ds. biotechnologii, przyjmuje się następującą definicję GMO: „Organizm, którego materiał genetyczny został zmieniony w sposób, w jaki nie zdarza się w warunkach naturalnych wskutek krzyżowania i/lub naturalnej rekombinacji".

Warto także podkreślić, że konwencja o różnorodności biologicznej posługuje się pojęciem biotechnologii „żywych organizmów zmodyfikowanych" (living modified organisms – LMO) [34]. Pojęcie to nie zostało zdefiniowane, co utrudnia interpretację treści konwencji przez poszczególne państwa ratyfikujące ją. W naszych rozważaniach będziemy używali pojęcia genetycznie zmodyfikowanych organizmów zakładając, że jest ono szersze niż LMO.

2. Zarys historii rozwoju GMO

Już od bardzo dawna badania naukowe zmierzają do produkcji coraz to nowych odmian produktów żywnościowych. Celem tych badań było wyprodukowanie wydajniejszych roślin po to, by zaspokoić potrzeby żywnościowe świata. Nowe zastosowania między innymi nawozów sztucznych, pestycydów, znacznie podwyższyły poziomy plonów wielu roślin. Większość produktów roślinnych obecnie spożywanych różni się od swoich naturalnych przodków.

W ciągu ostatnich 20 lat ubiegłego stulecia rozwinęły się skuteczne metody transformowania roślin, czyli wprowadzenie genów odpowiedzialnych za pożądane do genomu biorcy. Genetyczna transformacja roślin miała na celu wzbogacenie ich nowych cech użytkowych, między innymi odpornych na choroby, szkodniki oraz niekorzystne warunki środowiska.

W literaturze przedmiotu wyróżnia się kilka typów żywności genetycznie zmodyfikowanej:

1) żywność w całości jest genetycznie zmodyfikowana (np. pomidory),

2) żywność zawiera w swoim składzie GMO, np. koncentraty pomidorowe,

3) żywność, którą wyprodukowano z zastosowaniem GMO (np. produkty z zastosowaniem drożdży transgenicznych),

4) żywność, która jest pochodną GMO, tzn. nie zawiera żadnych komponentów transgenicznych, np. olej rzepakowy otrzymywany z transgenicznego rzepaku [12, s. 44].

Jak widzimy, określenie żywności GMO dotyczy jego składu.

Łatwo zauważyć, że aby wyprodukować żywność GMO, należy najpierw wyprodukować (surowiec) roślinny, w których przeprowadzono modyfikacje genetyczne. Wyhodowanie roślin GMO to proces wielu lat produkcji doświadczalnej.

W połowie lat 90. ubiegłego stulecia produkcja roślin zmodyfikowanych genetycznie (GMO) osiągnęła 1,7 mln ha. W ciągu 11 lat (ok. 2006) powierzchnia wzrosła 60-krotnie, co stanowiło ok. 6,8% całkowitej powierzchni upraw na świecie. W 2006 roku produkcję GMO prowadziły 22 kraje, w tym 6 w UE. Największe ilości GMO to: ryż, soja, kukurydza, bawełna, rzepak [23].

3. Stan prawny i normalizacyjny w zakresie GMO

Wśród wielu aktów prawnych dotyczących GMO stanowionych w różnych państwach na uwagę zasługują ustalenia prawne zapoczątkowane w latach 90. Do nich zaliczyć należy między innymi:

- Agendę 21 (dokumenty końcowe konferencji Narodów Zjednoczonych „Środowisko i rozwój", Rio de Janerio 3-14 czerwca 1992),
- Konwencję o różnorodności biologicznej; ważny jest tutaj Protokół Kartageński sporządzony w Montrealu 29 stycznia 2000, którego celem jest między innymi przyczynienie się do zapewnienia odpowiedniego poziomu ochrony w dziedzinie bezpiecznego przemieszczenia, przechowywania i wykorzystania innych zmodyfikowanych organizmów stanowiących wynik prac nowoczesnej biotechnologii [18],
- Dyrektywa Parlamentu Europejskiego i Rady 2001/18/WE z 12 marca 2001 r. o celowym wprowadzaniu do środowiska genetycznie zmodyfikowanych organizmów.

Polska ratyfikowała wyżej wymieniony Protokół Kartagoński 10 grudnia 2003 roku.

Oprócz tego, od 2001 obowiązują akty prawne regulujące: zamknięte użycie organizmów GMO, obrót krajowy i zagraniczny, także właściwości organów administracji państwowej [31].

W Polsce mamy następujące akty prawne normujące zagadnienia organizmów genetycznie zmodyfikowanych:

- Ustawę z 22 czerwca 2001 r. o organizmach genetycznie zmodyfikowanych (Dz.U., nr 76, poz. 811 z późn. zm.,
- Ustawę z 26 czerwca 2003 r. o nasiennictwie (nowelizowana 27.4.2006); przepisy art. 5 ust. 4 i art. 57 ust. 3 ustanawiają zakaz rejestracji odmian genetycznie zmodyfikowanych oraz wprowadzania do obrotu materiału siewnego odmian genetycznie zmodyfikowanych wraz z późniejszymi zmianami,
- Ustawa z 22 lipca 2006 r. o paszach; art. 15 ust. 1 pkt 4 wraz z późniejszymi zmianami zakazuje wytwarzania, wprowadzania do obrotu i stosowania w żywieniu zwierząt pasz genetycznie zmodyfikowanych wraz z późniejszymi zmianami.

Dodatkowy akt prawny zajmuje się aspektami zdrowotnymi żywności. Akt ten głosi, że żywność nie może stanowić zagrożenia dla zdrowia lub życia człowieka oraz środowiska [29; 30].

Kolejny ważny akt prawny w Polsce nawiązuje do przepisów rozporządzenia (UE) nr 178/2002 Parlamentu Europejskiego i Rady Europejskiej z 28 stycznia 2002 ustanawiającego ogólne zasady i wymagania prawa żywnościowego powołującego Europejski Urząd do spraw Bezpieczeństwa Żywności oraz ustanawiającego procedury w sprawie bezpieczeństwa żywności [32]. Nowa ustawa o bezpieczeństwie żywności stawia wysokie wymagania między innymi zdrowotne i higieniczne. Nakłada na organy właściwości kontrolne, a także odpowiedzialność za wyrządzone szkody przez środki spożywcze. Ponadto, Unia Europejska wydała dwa rozporządzenia dotyczące organizmów genetycznie zmodyfikowanych, w których określa zasady przemieszczania tych produktów, w tym takie obrotu pasz i żywności [27; 28].

Również w Polsce w 2006 roku prezydent podpisał ustawę zakazującą wytwarzanie obrotu i wysiewu pasz genetycznie zmodyfikowanych. W 2007 roku rząd Polski złamał obietnice w zakresie żywności GMO akceptując Ustawa z 25 sierpnia 2006 r.

Do wykrycia GMO w żywności Unia Europejska wprowadziła w 2003 roku następujące rozporządzenia:

- Rozporządzenie (WE) nr 1829/2003 Parlamentu Europejskiego i Rady Unii Europejskiej z 22.9.2003 r. w sprawie genetycznie zmodyfikowanej żywności i paszy,

– Rozporządzenie (WE) nr 1830/2003 Parlamentu Europejskiego i Rady Unii Europejskiej z 22.9.2003 r. dotyczące możliwości śledzenia i etykietowania organizmów zmodyfikowanych genetycznie oraz możliwości śledzenia żywności i produktów paszowych wyprodukowanych z organizmów zmodyfikowanych genetycznie (zmiana dyrektywy 2001/18/WE).

W 2001 roku wprowadzono 3 normy służące do wykrywania GMO w żywności o nazwie „Biotechnologia – organizmy do zastosowania w środowisku – Wytyczne do określenia organizmów zmodyfikowanych genetycznie poprzez analizę: funkcjonalnej ekspansji modyfikacji genowej (PN-EN 12682 : 2001), stabilność molekularnej modyfikacji genowej (PN-EN 12683 : 2001), modyfikacji genowej (PN-EN 12687 : 2001).

Dalsze prace dotyczące analitycznych metod organizmów zmodyfikowanych genetycznie i produktów pochodnych w artykułach żywnościowych w Polsce wprowadzono za pomocą 5 norm w latach 2005-2007 [12, s. 44-46]. Normy te szczegółowo opisują procedury i metody wykrywania GMO w produktach żywnościowych.

Jak widzimy, polityka normalizacyjna i legislacyjna nad wykryciem GMO w żywności, dopuszczenie jej do obrotu, oznakowania tej żywności są dalece zaawansowane. Normy te skutecznie ograniczają ryzyko spożycia żywności genetycznie zmodyfikowanej.

W sumie polskie regulacje prawne należą do najbardziej surowych [3] pod względem wymogów stawianych żywności i produktom GMO. W Polsce od 2008 roku przygotowywany jest nowy projekt ustawy dotyczący roślin GMO. Będzie on zgodny z ustaleniami UE.

4. Reakcje społeczeństwa na żywność GMO

Ważne są także opinie Europejczyków na temat żywności GMO. Badania opinii publicznej [10, s. 25] (w okresie listopad-grudzień 2007) w 27 krajach UE wykazują, że 2/3 krajów członkowskich nie chce GMO ani w uprawach, ani w artykułach spożywczych. Około 96% Europejczyków uznaje ochronę środowiska za bardzo ważny obszar i nie może on być zanieczyszczany dodatkowo przez GMO (ok. 20% respondentów uznało GMO za zagrożenie dla środowiska). W poszczególnych krajach UE jest więcej przeciwników niż zwolenników GMO. Najwięcej przeciwników GMO jest w Słowenii (81%), na Cyprze (80%), Polsce (67%), Włoszech (74%). W marcu 2007 roku „Gazeta Wyborcza" przeprowadziła sondaż na temat GMO. Z sondażu wynika, że ok. 60% uważa GMO za szkodliwe dla zdrowia, 45% jest za zakazem upraw transgenicznych, 60% deklaruje, że nie będzie kupowało produktów GMO, pomimo

że cena byłaby konkurencyjna. Z sondażu przeprowadzonego w serwisie www.biotechnolog.pl wynika, że Polacy nie zgadzają się na istnienie stref geograficznych produkujących żywność GMO. W USA i Kanadzie, liderach w produkcji GMO, odpowiednio 61% i 55% jest gotowe zaakceptować GMO. W marcu 2008 roku firma Martin i Jakobs opublikowała wyniki badań z których wynika, że przedstawiciele nauki i eksperci oraz producenci roku są za wprowadzeniem GMO.

W literaturze przedmiotu spotykamy liczne opinie sprzeciwu dla produkcji GMO. Oto kilka opinii:

- produkcja GMO zakłóci naturalne relacje w środowisku [24, s. 23],
- społeczeństwo europejskie popiera biotechnologię czerwoną (związaną z medycyną), biotechnologię białą (związaną z przemysłem) i sprzeciwia się biotechnologii zielonej, czyli związanej z rolnictwem [7, s. 19],
- odbywają się manifestacje protestujące przeciwko stosowaniu GMO w Niemczech, a nawet w Polsce [13],
- w Polsce nie istnieje żaden system ochrony naszych unikalnych w skali światowej zasobów [24, s. 16-17] przed skażeniem biologicznym jaki niesie GMO,
- im więcej pestycydów użytych w hodowli roślin GMO, tym gwałtowniejsze jest jałowienie gleb, tym samym konieczność bardziej agresywnego ich nawożenia [9].

Dyrektywa UE 2001/18/WE Parlamentu Europejskiego i Rady przewiduje, że „państwa członkowskie nie mogą zakazywać, ograniczać ani utrudniać wprowadzenia do obrotu GMO" [26, s. 24].

Zgodnie z traktatową zasadą swobodnego przepływu towarów, Polska nie może zabronić na swoim terytorium obrotu żywnością GMO, która jest w sprzedaży na rynkach UE zgodnie z decyzją Komisji Europejskiej. W 2007 roku było zarejestrowanych 26 rodzajów modyfikacji znajdujących się w Rejestrze Żywności GMO i Pasz GMO. Istnienie tej żywności w rejestrze powoduje, że żywność GMO może być przedmiotem obrotu handlowego [18]. W rejestrze są takie rośliny, jak kukurydza, soja, rzepak i bawełna. Jeśli nie dopuścimy w Polsce GMO do obrotu rynkowego, to za każdy dzień zwłoki grożą surowe kary.

Według stanu na koniec 2008 roku liczba produktów żywnościowych GMO na półkach sklepowych jest niewielka. Klienci w zasadzie nie zwracają uwagi na sposoby etykietowania, w których oznaczone są produkty żywnościowe GMO.

5. Wpływ spożycia GMO na zdrowie człowieka

Opublikowany raport Greenpeace pt. *Rok 2006 rokiem globalnego sprzeciwu konsumentów, rolników i rządów wobec GMO* prezentuje dane o zaprzestaniu jej produkcji. Przekonuje się, że istnieją niezaprzeczalne dowody na to, że produkcja GMO jest szkodliwa dla zdrowia ludności. Jako przykłady podaje się produkcję soi oraz zmutowanego ryżu.

Unia Europejska zaostrzyła system zabezpieczeń żywności przed nielegalnym obrotem GMO w 2006 roku Ministrowie UE większością głosów podtrzymali zakaz upraw kukurydzy [2, s. 27].

W czerwcu 2007 roku Rada Ministrów d/s Rolnictwa w UE przegłosowała propozycję Komisji Europejskiej w sprawie przepisów prawnych dotyczących zanieczyszczenia żywności ekologicznej przez GMO ustalając próg na poziomie 0,9%. Domagano się jednak progu 0,2%, tj. dopuszczalnego zanieczyszczenia produktów ekologicznych na poziomie technicznych możliwości wykrywalności [17, s. 11-12]. Pojawią się także zdecydowane wypowiedzi pracowników nauki (S.K. Wiechowski, T. Żarski) o szkodliwości żywności genetycznie zmodyfikowanej [20, s. 23-24].

Bardzo często zadaje się pytanie o ocenę ryzyka stosowania żywności genetycznie zmodyfikowanej. Ryzyko utożsamiane jest z zagrożeniem dla zdrowia i środowiska. Zgodnie z Dyrektywą Unii Europejskiej, ocena ryzyka w środowisku jest zdefiniowana jako „identyfikacja zarówno bezpośrednich, jak i pośrednich natychmiastowych bądź występujących z opóźnieniem skutków które mogą być wynikiem zamierzonego uwolnienia do środowiska lub wprowadzenia do obrotu. Ocena ryzyka jest to proces szacowania prawdopodobieństwa wystąpienia negatywnych skutków obecnie i w przyszłości dla człowieka i środowiska, biorąc pod uwagę wiele czynników. Trzeba jednak mocno podkreślić, że pomimo dołożenia wielu starań na etapie badania GMO całkowite wyeliminowanie ryzyka nie jest możliwe [19, s. 44; 15]. W myśl tej Dyrektywy zobowiązuje się instytucje wprowadzające GMO do monitorowania skutków dla ludzi i środowiska. Jednocześnie w tej Dyrektywie zwraca się uwagę, że gdy produkt GMO dostał pozwolenie do obrotu, to państwa członkowskie nie powinny ograniczać wprowadzania go na własny rynek.

W praktyce mogą się pojawić niezamierzone zanieczyszczenia GMO trudne do zidentyfikowania. Podkreśla się, że skutki mogą być negatywne i nieodwracalne dla człowieka i środowiska.

Wśród wielu negatywnych zjawisk oraz procesów wymienia się między innymi:

- procesy modyfikacji mogą być niedostatecznie rozpoznawane i przebadane,

- mogą być pomijane negatywne skutki ujawnione w trakcie długoletnich badań,
- mogą powstawać nieznane w naturze biologiczne jednostki hybrydowe o nieprzewidywalnych formach zachowania się w naturalnym środowisku [5].

Spożywanie żywności GMO może oddziaływać w sposób ciągły. Możemy wówczas mieć do czynienia z tzw. zdarzeniami ciągłymi, w zasadzie o długotrwałym nieprzerwanym oddziaływaniu czynnika szkodliwego na organizm człowieka. Niektórzy autorzy wskazują możliwość powstawania nowych rodzajów szkodników, wyrządzania szkód gatunkom już istniejącym, zakłócenie naturalnych procesów, także cykli pokarmowych i pogodowych, zniszczenie bioróżnorodności przez zajęcie miejsca gatunkom rodzimym [25, s. 87].

Pomimo wielu starań w zakresie produkcji GMO podmiotów gospodarczych mogą powstawać zjawiska przypadkowe i technicznie nieuniknione. Za takie przypadki producenci powinni ponosić ciężar odpowiedzialności. Jednostkowe obciążenie producenta skutkami negatywnymi produkcji GMO stawiałoby go w bardzo trudnej sytuacji ekonomicznej. Stąd ciężar skutków powinien być zdywersyfikowany, tj. rozłożony na wszystkich producentów.

Praktyka gospodarcza wskazuje także na to, że oprócz żywności niemodyfikowanej będzie występowała także żywność genetycznie zmodyfikowana [5]. Współistnienie obu rodzajów żywności stawia potencjalnego klienta w sytuacji wyboru spożycia żywności. To z kolei powoduje przerzucenie ryzyka (w sensie zagrożenia) na klienta, który powinien być świadomy jaką żywność spożywa. To z kolei decyduje o rozwoju rynku konsumpcji w Polsce.

6. Korzyści ze stosowania GMO

W Instytucie Badawczym Leśnictwa w Sękocinie pod Warszawą przeprowadzono cykl szkoleń na temat GMO, udowadniając liczne korzyści z jej produkcji [20, s. 21-23]. Udowadnia się ogromny postęp jaki daje biotechnologia w zakresie żywności genetycznie zmodyfikowanej. Zdaniem profesora Andrzeja Anioła z Instytutu Hodowli i Aklimatyzacji Roślin (IHAR), nie powinno się rezygnować z prac nad produkcją GMO, gdyż Polska stanie się zacofanym krajem. Zdaniem profesora T. Twardowskiego, blokowanie postępów biotechnologii jest sprzeczne z nauką i gospodarką [22, s. 20].

Dzięki żywności genetycznie zmodyfikowanej mogą lub powinny [19, s. 33-36; 1, s. 59-64]:

- zmniejszyć się straty spowodowane przez ich przedwczesne się psucie,
- być wzbogacone wartości odżywcze i smakowe,
- ulec rozkładowi (biodegenerowane plastiki),

- wzrosnąć produkcja biopaliw,
- obniżyć koszty produkcji żywności, lekarstw,
- zaistnienie możliwości oczyszczania środowiska,
- być zwalczane choroby człowieka, np. alergie, niedobory witamin,
- być odporne na szkodniki i choroby [18, s. 134],
- być odporne na mróz, suszę, stres, wyleganie [33].

Niektórzy autorzy wskazują na dużą odporność roślin zmodyfikowanych genetycznie na chwasty, herbicydy. Obecnie można mówić o zastosowaniu GMO w żywieniu ludności, produkcji konkretnych wyrobów spożywczych, produkcji różnego rodzaju dodatków, np. do pasz. Wskazuje się także na zastosowanie GMO w przemyśle farmaceutycznym, drzewnym, spożywczym, a także w produkcji kwiatów [25, s. 90].

7. Koncepcja obowiązkowego ubezpieczenia odpowiedzialności cywilnej za produkty genetycznie zmodyfikowane

Wcześniej wskazywano zwolenników i przeciwników produkcji żywności GMO. Presja społecznych potrzeb żywnościowych z jednej strony i argumentacja o szkodliwości dla zdrowia ludności z drugiej powoduje, że podejmowane decyzje są obarczone dużą dawką niepewności. Argumentacja, że skutki spożycia GMO mogą być widoczne po wielu latach skłania do poszukiwania innych instrumentów zabezpieczających. Takim instrumentem mogłyby być ubezpieczenia odpowiedzialności cywilnej, których celem jest rozłożenie ciężaru skutków na wiele podmiotów gospodarujących za skalkulowaną składkę. Jest to instrument w pewnym sensie godzący zwolenników i przeciwników GMO w praktyce. Jak każde ubezpieczenie, tak i w tym przypadku odpowiedzialność cywilna za produkt powinna dotyczyć przypadków zdarzeń losowych zdefiniowanych w ustawie o działalności ubezpieczeniowej z 22 maja 2003 roku.

Ochrona konsumenta stanowi część prawa obowiązującego na jednolitym rynku UE. Firmy działające w państwach członkowskich muszą przestrzegać 14 dyrektyw, których celem jest ochrona praw konsumenta [8]. Podstawową rolę odgrywają tu dyrektywy dotyczące odpowiedzialności za produkt. Dyrektywy te zostały przekazane do stosowania we wszystkich krajach, które przystąpiły do Unii Europejskiej w 2004 roku. W myśl tych dyrektyw firmy muszą wdrażać odpowiednie procedury w łańcuchu dostaw, jeżeli chcą mieć prawo do ubezpieczenia odpowiedzialności cywilnej za produkt.

Zakłada się, że ubezpieczenie odpowiedzialności cywilnej za produkt będzie w przyszłości obowiązkowe. Początkowo w warunkach tego ubezpieczenia nie zakłada się kosztów wycofania produktu z rynku, ponieważ byłoby ono bardzo drogie. Przewiduje się natomiast system odpowiedzialności gwarancyjnej co oznacza, że dostawcy i producenci sami muszą dowieść, że zrobili wszystko co mogli, aby zapobiec lub zminimalizować ryzyko straty lub szkody zdrowia człowieka.

Niezwykle trudno określić minimalną sumę gwarancyjną za ubezpieczenie odpowiedzialności cywilnej za produkt. Brakuje w tej kwestii doświadczenia i badań naukowych. Ponieważ spodziewać się należy dużych sum gwarancyjnych, koniecznością będzie zastosowanie reasekuracji ubezpieczeniowej. Reasekuracja ubezpieczeniowa powinna obejmować duże ryzyko po stronie akwizycji ubezpieczeniowej, jak i duże szkody ubezpieczeniowe. Z pewnością powinny być stosowane metody reasekuracji nieproporcjonalnej.

W proponowanym ubezpieczeniu ważnym elementem jest odległość czasowa skutków spożywania żywności GMO. Powstaje wówczas pytanie o datę powstania szkody. Jest to problem bardzo trudny do rozwiązania, gdyż spożywanie żywności GMO przez różne osoby może dać bardzo różne skutki dla zdrowia: od żadnych, poprzez różnego rodzaju schorzenia, do zgonów włącznie. Uzależnione jest od bardzo wielu czynników zdrowotnych.

Z punktu widzenia ubezpieczeniowego, zagadnienie to wiąże się z tzw. triggerami zdarzeń [14, s. 14-15]. Dotyczy problemów powstania, ujawniania, zgłaszania i formułowania roszczeń szkód na osobie. Zagadnienie dyskutowane jest przez środowisko prawników i w dalszym ciągu pojawia się tu wiele wątpliwości [4, s. 7-9].

Z punktu widzenia zakładu ubezpieczeń ma to znaczenie nie tylko dla konstrukcji ogólnych warunków ubezpieczeń, ale i dla procesu kalkulacji stopy składki. W proponowanym ubezpieczeniu możemy spotkać się z dostatecznie długim okresem odpowiedzialności za powstałe szkody wynikające ze spożywania produktów GMO. Obecnie obowiązujące przepisy prawne (Ustawa z 16.II.2007 r. Dz.U., 2007, nr 80, poz. 538) dają możliwość dochodzenia roszczeń nawet po 20 latach.

Ochroną ubezpieczenia odpowiedzialności cywilnej powinny być objęte wszystkie ogniwa żywnościowego łańcucha dostaw i produkcji żywności genetycznie zmodyfikowanej. Innymi słowy ubezpieczeniem należałoby objąć te podmioty gospodarcze, które produkują, przetwarzają i dystrybuują żywność genetycznie zmodyfikowaną. W sumie byłaby to odpowiedzialność cywilna za łańcuch logistyczny dla produktów żywnościowych GMO.

Ubezpieczenie proponowane może być rozpatrywane na zasadzie winy lub też na zasadzie ryzyka. Na zasadzie winy nieumyślnej określona może być miarą niedbalstwa, niedołożonej wystarczającej staranności w stosunku do

GMO. Na zasadzie ryzyka polega na uniezależnieniu odpowiedzialności odszkodowawczej określonych podmiotów od tego, czy ponoszą winę za spowodowane szkody czy też nie. Ta zasada ma również zastosowanie do GMO traktowanych jako produkty niebezpieczne dla środowiska i człowieka. W tym przypadku poszkodowany powinien udowodnić poniesioną stratę. W ubezpieczeniu tym także możemy się dopatrywać zarówno odpowiedzialności kontraktowej (za nienależyte wywiązanie się z produkcji roślinnej GMO), jak i odpowiedzialności deliktowej z tytułu czynu niedozwolonego.

Uwagi końcowe

Z przeprowadzonych rozważań wynika, że produkcja żywności genetycznie zmodyfikowanej będzie kontynuowana w krajach Unii Europejskiej. Obok produktów GMO będą także produkty niezmodyfikowane genetycznie. Równocześnie europejskie normy ochrony konsumenta zwiększą ryzyko odpowiedzialności za produkt. Dlatego też coraz więcej firm będzie poszukiwać odpowiednich rozwiązań w zakresie zarządzania ryzykiem. Coraz częstsze będzie przenoszenie tego ryzyka na rynek ubezpieczeń.

Na tle ustaleń w Dyrektywach Unii Europejskiej Polska w pełni podporządkowuje się ich realizacji. Można by powiedzieć, że polskie regulacje prawne są o wiele bardziej rygorystyczne niż Unii Europejskiej.

Proponowane ubezpieczenie odpowiedzialności cywilnej ma za zadanie chronić sprawcę, jak i poszkodowanego za szkody spowodowane przez GMO. Wprowadzenie i konstrukcja tego ubezpieczenia powinna być przedmiotem legislacji Parlamentu Europejskiego i Rady Unii Europejskiej.

Przedstawione w opracowaniu problemy nie wyczerpują całości kwestii ubezpieczeniowych w nowym obszarze badawczym. Warto tutaj podkreślić, że produkcja żywności genetycznie zmodyfikowanej to także problemy ubezpieczeń zdrowotnych i ubezpieczeń życiowych.

Literatura

1. Bąkowska-Żywiecka K., Tyczewska A., Twardowski T. (2008). Organizmy genetycznie zmodyfikowane – ekonomia i bezpieczeństwo. Nauka, 3.
2. GMO zagraża zdrowej żywności. (2007). Środowisko, 12.
3. Jóźwiak Z. (2007). Zbyt duża odpowiedzialność badaczy zmutowanej kukurydzy. Rzeczpospolita, 13.X.
4. Kosicka E. (2008). Triger w praktyce. Miesięcznik Ubezpieczeniowy, 2.

5. Maciejczak M. (2005). Ekonomiczne i rynkowe aspekty współistnienia producentów modyfikowanych genetycznie i niezmienionych w łańcuchu dystrybucji żywności i pasz. Zagadnienia Ekonomiki Rolnej, 3.
6. Makowiecki M. (2008). Kto się boi GMO. Nowe Życie Gospodarcze, 2 (dodatek specjalny).
7. Nowicki M. (2008). GMO stanowi zagrożenie. Środowisko, 11 (871), 19.
8. Odpowiedzialność za produkt. (2006). Dziennik Ubezpieczeniowy, Gazeta Ubezpieczeniowa, 3.X.
9. Polacy się boją GMO i bardzo słusznie. (2008). Ekopartner, 5.
10. Priwieziencew E. (2008). Opinie Europejczyków o GMO. Środowisko, 11, 25.
11. Serwach M. (2007). Triggery wolność kontraktowa. Miesięcznik Ubezpieczeniowy, 11.
12. Sosnowska T. (2008). Polskie Normy – narzędzia do wykrywania GMO w żywności. Wiadomości PKN, Normalizacja, 5.
13. Społeczeństwo przeciw GMO. Symboliczny protest. (2008). Ekoświat, lipiec-sierpień.
14. Sukiennik P. (2008). Brokerzy i triger „zdarzenia". Miesięcznik Ubezpieczeniowy, 7/8.
15. Szopa J., Łukaszewicz M. (2003). Stan i perspektywy rozwoju roślin transgenicznych w Polsce. Wieś Jutra, 7.
16. Tolik M. (2008). Konwencja o różnorodności biologicznej wobec żywych zmodyfikowanych organizmów i wyzwań rozwoju biotechnologii. Prawo i Środowisko, 2, 63-79.
17. Uprawy GMO na świecie. (2007). Środowisko, 2.
18. Zębek E. (2007). Rośliny genetycznie zmodyfikowane w środowisku na tle regulacji prawnych. Prawo i Środowisko, 1.
19. Zimny J. (2007). Żywność zmodyfikowana genetycznie i bezpieczeństwo jej stosowania. Postępy Nauk Rolniczych, 1.
20. Zysk A. (2007). Czy GMO to żywność Frankensztajna? Środowisko, 7.
21. Zysk A. (2008). Czy ustawa chroni przyrodę przed GMO. Środowisko, 17, 16-17.
22. Zysk A. (2007). Zdrowa żywność bez GMO. Środowisko, 4.
23. Zysk A. (2007). Nasiona kłamstwa. Środowisko, 22/358.
24. Żarski T. (2008). Zagrożenia bioróżnorodności przez GMO. Środowisko, 11 (371).
25. Żuk D. (2003). Inżynieria genetyczna. Inżynieria procesów bioagrotechnologicznych. Warszawa.
26. Żywność genetycznie modyfikowana (ekoportal.pl). (2007). Biznes i Ekologia, 64.

27. Rozporządzenie (WE) nr 11946/2003 Parlamentu Europejskiego i Rady z 15.07.2003 r. w sprawie transgranicznego przemieszczania organizmów genetycznie zmodyfikowanych. Dz.Urz. UEL 287 z 05.11.2003 r., s. 1. Dziennik Urzędowy Unii Europejskiej Polskie wydanie specjalne, rozdz. 17, t. 7, s. 650.
28. Rozporządzenie (WE) nr 1829/2003 Parlamentu Europejskiego i Rady z 22.09.2003 r. w sprawie genetycznie zmodyfikowanej żywności i paszy. Dz.Urz. UEL 268 z 18.10.2003 r., s. 1, Dz. Urz. UE Polskie wydanie specjalne, rozdz. 13, t. 32, s. 432.
29. Ustawa z 11 maja 2001 r. o warunkach zdrowotnych żywności i żywienia. Dz.U. 2001, nr 63.
30. Ustawa z 21 maja 2003 r. o zmianie ustawy o organizmach genetycznych oraz ustawy o warunkach zdrowotnych żywności i żywienia. Dz.U. 2003, nr 130 z późn. zm.
31. Ustawa z 22 czerwca 2001 r. o organizmach genetycznie zmodyfikowanych. Dz.U. 2001, nr 76, poz. 811 z późn. zm.
32. Ustawa z 25 sierpnia 2006 r. o bezpieczeństwie żywności i żywienia. Dz.U. 2006, nr 171 z późn. zm.
33. http://www.ihar.edu.pl/gfl2000/zimny.html
34. www.cbd.int

Alicja Wolny-Dominiak
Akademia Ekonomiczna w Katowicach

SZACOWANIE POZIOMU SKŁADKI WIARYGODNEJ W WYBRANYCH MODELACH Z ZASTOSOWANIEM PAKIETU ACTUAR PROGRAMU STATYSTYCZNEGO R

Wstęp

Proces szacowania składki netto w ubezpieczeniach majątkowych może wykorzystywać różne podejścia stosowane w tej problematyce. Jedno z podejść polega na budowie taryf dla portfela niejednorodnego stosując dwuetapowy proces taryfikacji, dzielący portfel na subportfele jednorodne. Subportfel jednorodny rozumiany jest jako zbiór polis generujących szkody niezależnie od siebie oraz zmienne losowe opisujące szkodowość o tych samych rozkładach prawdopodobieństwa. W subportfelu jednorodnym każdej polisie przypisywana jest taka sama składka netto. Proces budowy taryf uwzględnia: taryfikację a'priori oraz taryfikację a'posteriori. Taryfikacja a'priori polega na kalkulacji składki bazowej biorąc pod uwagę czynniki opisujące osobę ubezpieczającą się (np. wiek, płeć) oraz charakterystyczne cechy przedmiotu ubezpieczenia (np. dla ubezpieczeń komunikacyjnych pojemność silnika pojazdu, rok produkcji). Traktując te czynniki jako jakościowe zmienne losowe, zwane zmiennymi taryfikacyjnymi, możliwe jest przeprowadzanie statystycznej analizy danych badając wpływ poszczególnych zmiennych na poziom szkodowości. Z kolei taryfikacja a'posteriori to określenie składki dodatkowej (liczonej najczęściej jako procent składki bazowej), gdzie uwzględniana jest liczba szkód spowodowanych w przeszłości przez osobę ubezpieczającą się. W tym celu wykorzystuje się najczęściej system bonus-malus, system zniżek-zwyżek.

Inne podejście wykorzystywane jest w tzw. teorii wiarygodności (teorii zaufania), gdzie szacowana jest składka wiarygodna uwzględniająca jednocześnie historię szkodowości w całym portfelu oraz różnorodne czynniki ryzyka reprezentowane przez parametr ryzyka Θ. W literaturze przedmiotu zaproponowano wiele różnorodnych metod szacowania składki wiarygodnej oraz występującego w niej tzw. współczynnika wiarygodności. W pracy przedstawiono

wybrane aktuarialne modele wiarygodności (głównie modele hierarchiczne) i ich implementacja w programie R. Zaprezentowano działanie funkcji `cm()` dostępnej w pakiecie programu R o nazwie `actuar` oraz zaimplementowano algorytm iteracyjny szacowania składki wiarygodnej w modelu autoregresyjnym.

1. Ogólna charakterystyka aktuarialnych modeli wiarygodności

Aktuarialne modele wiarygodności stanowią część teorii wiarygodności, której istotą jest wyznaczanie poziomu tzw. składki wiarygodnej na okres $n+1$, przy danych obserwacjach z okresów wcześniejszych $1,...,n$. Składka wiarygodna ustalana jest dla subportfela, który rozumiany jest jako zbiór polis traktowanych jako identyczne jednostki ryzyka, tj. generujące straty niezależnie od siebie i charakteryzujące się jednakowymi rozkładami szkód oraz rozkładami struktury ryzyka (w szczególności subportfel może się składać z jednej polisy). W efekcie, w *k*-tym subportfelu wszystkie jednostki ryzyka mają przypisany taki sam poziom składki. W ramach subportfela dopuszcza się niewielkie zróżnicowanie w wartościach odszkodowań.

Zbiór polis należących do różnych subportfeli, ale przyjmujących zbliżone wartości zmiennych taryfikacyjnych opisujących cechy ryzyka tworzy portfel. Generalnie, jednostki ryzyka w portfelach najczęściej posiadają dodatkowe cechy nieuwzględniane w zmiennych taryfikacyjnych, co powoduje niejednorodność. Dlatego w modelach wprowadza się zmienną charakteryzującą ryzyko, tzw. parametr ryzyka Θ.

W dalszej części pracy przyjmowane są następujące oznaczenia: X_{ki}-całkowita wartość odszkodowań w *k*-tym subportfelu i w i-tym okresie, w_{ki}-waga dla zmiennej X_{ki} (np. liczba szkód), θ_{ki}-parametr ryzyka reprezentujący wszystkie nie uwzględnione w zmiennych taryfikacyjnych cechy ryzyka dla *k*-tego subportfela w i-tym okresie, gdzie $k=1,..,K$, $i=1,...,n$. Wskaźnik k pokazuje liczbę realizacji parametru ryzyka (np. kobieta, mężczyzna) jednocześnie wskazując liczbę subporfeli dla całego portfela. W prezentowanych w pracy modelach zakłada się, iż parametr ryzyka dla subportfeli nie zmienia się w czasie, dlatego występuje $\theta_1,...,\theta_K$ realizacji parametru ryzyka.

Zgodnie z zasadą równowagi wypłaconych odszkodowań i zebranych składek, składka netto uwzględniająca nieznane cechy ryzyka dla *k*-tego subportfela wyraża się wzorem

$$\mu_{n+1}(\theta_k) = E(X_{k,n+1} \mid \theta_k)$$

Tak zdefiniowana składka nazywana jest składką wiarygodną, w odróżnieniu od składki kolektywnej dla portfela: $\mu = E(X_{ki})$. Jako że wartość składki wiarygodnej jest nieznana, z uwagi na nieznany poziom parametru ryzyka niezbędne jest szacowanie jej poziomu.

W aktuarialnych modelach wiarygodności poszukuje się najlepszych liniowych nieobciążonych estymatorów BLUE (best linear unbiased estimator) składki wiarygodnej stosując znaną technikę minimalizacji średniego błędu kwadratowego

$$\min_f E(\mu(\theta_k) - f(X_i))^2$$

gdzie f jest dowolną funkcją rzeczywistą a $X_i = (X_{k1},...,X_{kn})$ wektorem historycznych obserwacji wartości odszkodowań w okresach $1,...,n$.

W poniższych modelach uzyskiwany jest estymator składki wiarygodnej dla *k*-tego subportfela, który można zapisać w postaci ogólnej

$$\mu_{n+1}^k = Z_k \overline{X}_k + (1 - Z_k)\mu$$

gdzie Z oznacza współczynnik wiarygodności, $\overline{X}$ estymator wiarygodności, a μ składkę kolektywną. W zależności od przyjmowanych założeń (poza założeniem ogólnym o istnieniu drugich momentów dla zmiennej X), uwzględnianych wielkości oraz innych czynników, estymatory nieznanych wielkości Z, $\overline{X}$ oraz μ przyjmują różną postać.

2. Model Buhlmanna oraz Buhlmanna-Strauba

Model Buhlmanna jest najstarszym modelem, rozwijanym dalej na przestrzeni lat, gdzie przyjmowane są następujące założenia: dla ustalonego parametru ryzyka θ_i zmienne $X_{k1},...,X_{kn_{pj}}$ są warunkowo niezależne o jednakowych rozkładach, wektory $(\theta_1, X_{1i}),\ldots,\ (\theta_K, X_{Ki})$ są wzajemnie niezależne o tych samych rozkładach. Zatem w modelu odszkodowania występujące dla subportfeli w danym roku nie zależą od tego, co działo się w latach przeszłych oraz wartość odszkodowań ma ten sam rozkład w różnych latach. Ponadto, subportfele generują odszkodowania w danym roku niezależnie od siebie oraz wszystkie subportfele mają taką samą historię n lat. Estymator składki wiarygodnej ma postać [1]

$$\hat{\mu}_{n+1}^k = \hat{Z}\ \overline{X}_k + (1 - \hat{Z}\ )\hat{\mu}$$

gdzie

$\hat{Z} = \dfrac{n\psi}{n\psi + \varphi}$, $\overline{X}_k = \sum_{i=1}^{n} X_{ki}$ oraz $\hat{\mu} = \sum_{k=1}^{K} \overline{X}_k$. We wzorze współczynnika wiarygodności występują tzw. parametry struktury interpretowane następująco

(i) $\varphi = E(D^2(X_{ki} \mid \theta_k))$ – miara niejednorodności wewnątrz subportfela,

(ii) $\psi = E(D^2(X_{ki} \mid \theta_k))$ – miara niejednorodności pomiędzy subportfelami.

W modelu Buhlmanna-Strauba wprowadza się dodatkowo wagi w_{ki} dla zmiennej X_{ki}. Najczęściej wagą jest liczba odszkodowań dla danej zmiennej. Wtedy do modelu szacowania składki wiarygodnej wprowadza się dodatkowe założenie o istnieniu funkcji s^2, zależnej od parametru ryzyka i takiej, że $D^2(X_{ki} \mid \theta_k) = \frac{s^2(\theta_k)}{w_{ki}}$. Oznacza to, iż zmienne losowe $X_{k1},...,X_{kn_{pj}}$ mogą przyjmować różne wariancje rozkładów. Wtedy odpowiednie estymatory BLUE dane są wzorami [4]

$$\hat{\mu}_{n+1}^k = \hat{Z}_k \overline{X}_k + (1 - \hat{Z}_k)\hat{\mu}$$

gdzie

$$\hat{Z}_k = \frac{\sum_{i=1}^{n} w_{ki}\psi}{\sum_{i=1}^{n} w_{ki}\psi + \varphi},\ \overline{X}_k = \sum_{i=1}^{n} \frac{w_{ki}}{\sum_{i=1}^{n} w_{ki}} X_{ki} \text{ oraz } \hat{\mu} = \sum_{k=1}^{K} \frac{Z_k}{\sum_{k=1}^{K} Z_k} \frac{w_{ki}}{\sum_{i=1}^{n} w_{ki}} X_{ki}$$

Przyjmując wszystkie wagi równe 1 uzyskuje się model Buhlmanna.

3. Hierarchiczny model wiarygodności i jego szczególne przypadki

W modelach Buhlmanna oraz Buhlmanna-Strauba przyjmuje się, że ryzyko w portfelu charakteryzuje jedna cecha opisywana przez parametr ryzyka Θ. Wprowadzając dwa i więcej parametrów ryzyka przechodzi się do modeli hierarchicznych, które klasyfikują jednostki w portfelu ze względu na jedna lub więcej cechy. Do prezentacji modelu załóżmy dwupoziomową strukturę hierarchii. Niech $p = 1,...,P$ oznacza liczbę realizacji pierwszej cechy ryzyka tworząc pierwszy poziom hierarchii dzielący portfel na sektory, $k = 1,...,K_p$ oznacza liczbę realizacji drugiej cechy ryzyka tworząc poziom drugi, w którym występują subportfele, $i = 1,...,n_{pk}$ oznacza liczbę okresów dla kontraktu na pierwszym poziomie p oraz drugim poziomie k. Ponadto, zmienna X_{pki} reprezentuje wartość odszkodowań dla kontraktu na poziomach (p,k) w i-tym okresie czasu natomiast θ_p oraz θ_{pk} oznaczają realizację parametrów ryzyka dla poziomu p oraz poziomów (p,k).

W celu uzyskania estymatora BLUE składki wiarygodnej w modelu hierarchicznym, minimalizując przeciętny błąd kwadratowy niezbędne jest przyjęcie następujących założeń: wektory $(\theta_1,\theta_{1j},X_{1ki}),\ldots,\ (\theta_P,\theta_{Pj},X_{Pki})$ są niezależne, dla ustalonego parametru θ_p wektory $(\theta_{1k},X_{1ki}),\ldots,\ (\theta_{Pk},X_{Pki})$ są niezależne warunkowo, wszystkie pary parametrów ryzyka (θ_p,θ_{pk}) mają jednakowe rozkłady, dla ustalonych parametrów ryzyka zmienne $X_{pk1},...,X_{pkn_{pk}}$ są warunkowo niezależne. Uzyskuje się następujące estymatory składki wiarygodnej [4]

$$\hat{\mu}_{pk} = Z_{pk}\overline{X}_{pk} + (1-Z_{pk})\hat{\mu}_p$$
$$\hat{\mu}_p = Z_p\overline{X}_p + (1-Z_p)\hat{\mu}$$

gdzie

współczynniki wiarygodności $\hat{Z}_{pk} = \dfrac{\sum_{i=1}^{n_{pk}} w_{pki}}{\sum_{i=1}^{n_{pk}} w_{pki} + \dfrac{\varphi}{\psi}}$, $\hat{Z}_p = \dfrac{\sum_{k=1}^{K_p} Z_{pk}}{\sum_{k=1}^{K_p} Z_{pk} + \dfrac{\psi}{\eta}}$

W modelu występuje dodatkowy parametr strukturalny $\eta = D^2(E(X_{pki} \mid \theta_p))$ – miara niejednorodności pomiędzy sektorami.

4. Model autoregresyjny jako rozszerzenie modelu Buhlmanna

Wyznaczając estymator składki wiarygodnej $\hat{\mu}_{n+1}^k$ w modelu hierarchicznym, w szczególności w modelu Buhlmanna (jeden poziom hierarchii oraz wszystkie wagi $w=1$), przyjmuje się założenie o warunkowej niezależności w czasie wartości wypłacanych odszkodowań dla k-tego subportfela. Pomijając to założenie można przedstawić problem szacowania składki wiarygodnej jako model autoregresyjny przyjmując, że na wartość odszkodowań w okresie $n+1$ wpływają wartości wypłacane w okresach wcześniejszych $1,...,n$.

Model autoregersyjny składki wiarygodnej dla k-tego kontraktu można przedstawić jako autoregresję pierwszego rzędu AR(1) [6]

$$e_{ki} = \rho_k e_{k,i-1} + \xi_{ki},\ 1,...,n$$

gdzie $e_i = X_{ki} - \mu(\theta_k)$ oraz $\mu(\theta_k)$ nie zależy od okresu i, dla każdego $k = 1,..,K$. W powyższym równaniu współczynnik ρ_k jest współczynnikiem autokorelacji natomiast składnik losowy ξ_{ki} spełnia klasyczne założenia:

(i) $E(\xi_{ki} \mid \theta_k) = 0,$

(ii) $D^2(\xi_{ki} \mid \theta_k) = \sigma_k^2(\theta_k),$

(iii) $\text{cov}(\xi_{ki}, \xi_{kj} \mid \theta) = 0$, $i, j = 1,...,n,$

(iv) dla danego θ_k składnik losowy ξ_{kj} jest niezależny od e_{ki}, gdzie $j > i.$

Ponadto, jako że proces generujący obserwacje dotyczące wypłaconych odszkodowań w portfelu jest stacjonarnym procesem stochastycznym, niezbędne jest założenie $|\rho_k| < 1$.

W powyższym modelu estymator składki wiarygodnej przyjmuje postać

$$\hat{\mu}_{k,n+1} = Z_k \overline{X}_k^{\rho} + (1 - Z_k)\mu$$

gdzie Z_k jest współczynnikiem wiarygodności, $\overline{X}_k^{\rho}$ estymatorem zaufania a μ składką kolektywną dla portfela. Zgodnie z twierdzeniem nieznane wielkości szacuje się następująco [5]

$$\hat{Z}_k = \frac{\psi[n(1-\rho) + 2\rho]}{\frac{\varphi}{1-\rho} + \psi[n(1-\rho) + 2\rho]}$$

$$\hat{\overline{X}}_k^{\rho} = \frac{X_{k1} + (1-\rho_k)\sum_{i=2}^{n-1} X_{ki} + X_{kn}}{n(1-\rho_k) + 2\rho_k}$$

gdzie parametry strukturalne wynoszą: $\varphi = E(D^2(X_{ki} \mid \theta_k))$ oraz $\psi = D^2(E(X_{ki} \mid \theta_k)).$

Widać, iż dla wyznaczenia składki wiarygodnej w okresie $n+1$ niezbędne jest oszacowanie parametrów struktury występujących w modelu: μ oraz $\rho_k, \varphi_k, \psi_k$. W celu wykorzystania wszystkich informacji dla całego portfela przyjąć należy założenie, iż dla każdego $k = 1,..,K$ parametry $\rho_k, \varphi_k, \psi_k$ są niezmienne. Zatem $\rho_k = \rho$, $\varphi_k = \varphi$, $\psi_k = \psi$. Jako iż poszczególne parametry zależą od siebie wzajemnie niezbędne jest zastosowanie algorytmu iteracyjnego [5]. W kroku startowym algorytmu zakłada się

(a) $\rho^0 = 0$ – w tym przypadku model autoregresyjny sprowadza się do modelu Buhlmanna,

(b) $\mu^0 = \frac{1}{Kn}\sum_{k=1}^{K}\sum_{i=1}^{n} X_{ki}$ – średnia wartość wypłacanych odszkodowań z całego portfela,

(c) $\varphi^0 = \frac{1}{K(n-1)}\sum_{k=1}^{K}\sum_{i=1}^{n}(X_{ki} - \frac{1}{n}\sum_{l=1}^{n} X_{kl})^2$,

(d) $\psi^0 = \frac{1}{K-1}\sum_{k=1}^{K}(\frac{1}{n}\sum_{i=1}^{n} X_{ki} - \mu)^2 - \frac{1}{n}(\frac{\varphi^0}{1-\rho^0})$,

(e) $\hat{Z}^0 = \frac{\psi^0 n}{\varphi^0 + \psi^0 n}$,

(f) $\hat{\bar{X}}_k^0 = \frac{X_{k1} + \sum_{i=2}^{n-1} X_{ki} + X_{kn}}{n}$,

(g) $\hat{\mu}^0{}_{k,n+1} = \hat{Z}^0_{kk}\,\hat{\bar{X}}_k^0 + (1-\hat{Z}_k^0)\mu^0$,

(h) $\hat{e}_{ki}^0 = X_{ki} - \hat{\mu}_{k,n+1}^0$.

W kroku $j+1$ uzyskiwane są następujące estymatory

(a) $\hat{\rho}^{j+1} = \frac{1}{K}\sum_{k=1}^{K}(\frac{\sum_{i=2}^{n}\hat{e}_{ki}^j\hat{e}_{k,i-1}^j}{\sum_{i=2}^{n}\hat{e}_{k,i-1}^2})$

(b) $\hat{\mu}^{j+1} = \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{k,n+1}^j$

(c) $\hat{\varphi}^{j+1} = \frac{1}{K(n-2)}\sum_{k=1}^{K}\sum_{i=2}^{n}(\hat{e}_{ki}^j - \hat{\rho}^{j+1}\hat{e}_{k,i-1}^j)^2$

(d) $\hat{\psi}^{j+1} = \frac{1}{K-1}\sum_{k=1}^{K}(\hat{\mu}_{k,n+1}^j - \hat{\mu}^{j+1})^2 - \frac{1}{n}(\frac{\hat{\varphi}^{j+1}}{1-(\hat{\rho}^{j+1})^2})$

(e) $\hat{Z}^{j+1} = \frac{\hat{\psi}^{j+1}[n(1-\hat{\rho}^{j+1}) + 2\hat{\rho}^{j+1}]}{\frac{\hat{\varphi}^{j+1}}{1-\hat{\rho}^{j+1}} + \hat{\psi}^{j+1}[n(1-\hat{\rho}^{j+1}) + 2\hat{\rho}^{j+1}]}$

(f) $\hat{\bar{X}}_k^{\rho} = \dfrac{X_{k1} + (1-\hat{\rho}^{j+1})\sum_{i=2}^{n-1} X_{ki} + X_{kn}}{n(1-\hat{\rho}^{j+1}) + 2\hat{\rho}^{j+1}}$

(g) $\hat{\mu}^{j+1}{}_{k,n+1} = \hat{Z}_{kk}^{j+1}\hat{\bar{X}}_k^{j+1} + (1-\hat{Z}_k^{j+1})\mu^{j+1}$

(h) $\hat{e}_{ki}^{j+1} = X_{ki} - \hat{\mu}_{k,n+1}^{j+1}$

Kryterium stopu algorytmu może być przyjmowane przykładowo jako różnica w kolejnych wartościach współczynnika wiarygodności.

5. Zastosowanie pakietu `actuar` programu `R` w analizie porównawczej aktuarialnych modeli wiarygodności

Pakiet `actuar` programu `R` posiada funkcję służącą do szacowania składki wiarygodnej `cm()`. Zastosowanie tej funkcji wymaga danych ujętych w następującym formacie (model jednopoziomowy lub model dwupoziomowy):

Tabela 1

Macierz danych wejściowych

Sektory p	Subportfele k	Wartości odszkodowań w latach			Wagi odpowiadające wartościom odszkodowań		
1	1	X_{111}	…	$X_{11n_{11}}$	w_{111}	…	$w_{11n_{11}}$
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
P	K	X_{PK1}	…	$X_{PKn_{PK}}$	w_{PK1}	…	$w_{PKn_{PK}}$

Ogólna składnia funkcji `cm()` dla modelu hierarchicznego przedstawia się następująco [3]

```
cm(formula, data, ratios, weights,
   method = c("Buhlmann-Gisler", "Ohlsson", "iterative"),
   tol = sqrt(.Machine$double.eps), maxit = 100,
   echo = FALSE)
```

W celu zobrazowania działania funkcji `cm()` dla modelu hierarchicznego, a dalej dla porównania modeli szacowania składki wiarygodnej, wykorzystano w części (nie wszystkie zmienne są uwzględniane) bazę danych zaczerpniętej z pracy [2], gdzie wprowadzono do modelu jednopoziomowego następujące zmienne:

1) wiek ubezpieczonego (8 grup wiekowych): 17-20, 21-24, 25-29, 30-34, 35-39, 40-49, 50-59, 60+,

2) wiek pojazdu (4 grupy): 0-3, 4-7,8-9, 10+ .

Dla modelu dwupoziomowego zagregowano wiek pojazdu do dwóch grup 0-7 oraz 8-10+. Dane szczegółowe przedstawia tabela 2.

Tabela 2

Format danych dla modelu jednopoziomowego

Sektor	Subportfel	Wiek_ubezpieczonego								Wagi							
	Wiek_poj	x1	x2	x3	x4	x5	x6	x7	x8	w1	w2	w3	w4	w5	w6	w7	w8
1	1	1613	1397	1260	1235	899	1098	1106	1142	30	145	416	523	563	1070	850	537
2	2	1669	1100	1107	1105	900	971	934	973	51	184	406	444	466	873	667	458
3	3	600	829	684	657	717	820	1757	764	6	32	77	92	103	190	180	142
4	4	171	405	1160	386	391	474	549	457	2	9	31	42	45	105	94	109

Źródło: [2].

Tabela 3

Format danych dla modelu dwupoziomowego

Sektor	Subportfel	Wiek_ubezpieczonego								Wagi							
	Wiek_poj	x1	x2	x3	x4	x5	x6	x7	x8	w1	w2	w3	w4	w5	w6	w7	w8
1	1	1613	1397	1260	1235	899	1098	1106	1142	30	145	416	523	563	1070	850	537
1	2	1669	1100	1107	1105	900	971	934	973	51	184	406	444	466	873	667	458
2	3	600	829	684	657	717	820	1757	764	6	32	77	92	103	190	180	142
2	4	171	405	1160	386	391	474	549	457	2	9	31	42	45	105	94	109

Po zastosowaniu funkcji uzyskano następujące wyniki:

a. Model Buhlmanna

```
cm(formula = ~wiek_poj, data = dane, ratios = x1:x8)
```

Tabela 4

Wyniki dla modelu Buhlmanna

Składka wiarygodna	Współczynnik wiarygodności
1187,759552	0,897446294
1076,588392	0,897446294
859,9672931	0,897446294
541,9347626	0,897446294

Parametry strukturalne: składka kolektywna $\mu = 916.56$; wariancja wewnątrzgrupowa $\varphi = 82410.4$, $\psi = 90146.56$.

b. Model Buhlmanna-Strauba

```
cm(formula = ~wiek_poj, data = dane, ratios = x1:x8,
weights = w1:w8)
```

Tabela 5

Wyniki dla modelu Buhlmanna-Strauba

Składka wiarygodna	Współczynnik wiarygodności
1110,477634	0,911962073
998,4066479	0,898917215
963,2066964	0,673172774
721,7697083	0,522675223

Parametry strukturalne: składka kolektywna $\mu = 948{,}46$; wariancja wewnątrzgrupowa $\varphi = 29464067$, $\psi = 23714.52$.

c. Model hierarchiczny dwupoziomowy

```
X<-cbind(sektor=c(1,1,2,2),dane)
fit <- cm(~sektor+sektor:wiek_poj, data=X, ratios = x1:x8,
weights = w1:w8, method = "iterative")
```

Tabela 6

Wyniki dla modelu hierarchicznego dla sektorów

Składka wiarygodna	Współczynnik wiarygodności
994,6764676	0,4950648
856,112783	0,4263002

Tabela 7

Wyniki dla modelu hierarchicznego dla subportfeli

Składka wiarygodna	Współczynnik wiarygodności
1119,04722	0,9462073
1003,44207	0,9378911
944,96083	0,7776568
634,12839	0,6502768

Parametry strukturalne: składka kolektywna $\mu = 925.39$; wariancja wewnątrzgrupowa $\varphi = 9464067$, $\psi = 20955.26$, $\eta = 40268.94$.

Implementacja modelu autoregresyjnego w programie `R` przedstawia poniższy algorytm iteracyjny, gdzie zmienna `dane` oznacza dane wejściowe.

```
# krok 0
ro=0
mi=0
fi=0
ksi=0
K=nrow(dane)
n=ncol(dane)
mi=sum(dane)/(K*n)
sr.wiersz=rowSums(dane)/n
dane1=dane
for (i in 1:K) {
      dane1[i,]=(dane1[i,]-sr.wiersz[i])^2
}
fi=sum(dane1)/(K*(n-1))
ksi=sum((sr.wiersz-mi)^2)/(K-1)-fi/n
Z=(ksi*n)/(fi+ksi*n)
X_k=rowSums(dane)/n
mi_k=Z*X_k+(1-Z)*mi
e=dane-mi_k
# krok j+1
Z_wektor=c(100, Z) ### liczba 100 jest umowna, wymusza
zrealizowanie przynajmniej jednej pętli
p=2
epsylon=0.0000001
while (abs(Z_wektor[p]-Z_wektor[p-1])>=epsylon) {
e_pom1=e
e_pom2=e
iloraz=c()
for (k in 1:K) {
      for (i in 2:n) {
            e_pom1[k,i]=e[k,i]*e[k,i-1]
            e_pom2[k,i]=e[k,i-1]^2
                        }}
suma1=0
suma2=0
suma1=rowSums(e_pom1[,2:n])
suma2=rowSums(e_pom2[,2:n])
iloraz=suma1/suma2
ro=sum(iloraz)/K
mi=sum(mi_k)/K
e_pom=e
for (k in 1:K) {
      for (i in 2:n) {
            e_pom[k,i]=(e[k,i]-ro*(e[k,i-1]))^2
                        }}
e_pom=e_pom[,-1]
fi=sum(e_pom)/(K*(n-2))
ksi=(sum((mi_k-mi)^2)/(K-1))-(fi/(n*(1-ro^2)))
Z=(ksi*(n*(1-ro)+2*ro))/(fi/(1-ro)+(ksi*(n*(1-ro)+2*ro)))
```

```
Z_wektor=c(Z_wektor, Z)
suma=0
X_k=c()
for (k in 1:K) {
      for (i in 2:(n-1)) {
            suma=suma+dane[k,i]
                        }
X_k=rbind(X_k,(dane[k,1]+(1-ro)*suma+dane[k,n])/(n*(1-
ro)+2*ro))
suma=0
}
mi_k=Z*X_k+(1-Z)*mi
mi_k_wektor=cbind(1,mi_k)
e=dane-mi_k
p=p+1
}
```

Dla kryterium stopu < 0,000001 algorytm wykonał 9 iteracji w efekcie otrzymując wyniki zawarte w tabeli 8.

Tabela 8

Wyniki dla modelu autoregresyjnego

Składka wiarygodna	Współczynnik wiarygodności
1185,66345	0,880071155
1077,97534	0,880071155
857,69476	0,880071155
545,55044	0,880071155

Podsumowanie

Model autoregresyjny do wyznaczania składki wiarygodnej jest uzupełnieniem dużej grupy aktuarialnych modeli wiarygodności. Założenie o zależności wartości szkód dla ryzyka w portfelu w kolejnych okresach wydaje się bardziej realistyczne niż klasyczne założenia modelu Buhlmanna. Dzięki implementacji w programie R modelu autoregreszjnego możliwe jest przeprowadzanie testów na rzeczywistych bazach danych w celu określenia sytuacji, w których model jest najbardziej użyteczny dla ubezpieczyciela.

Literatura

1. Buhlmann H. (1967). Experience Rating and Credibility. ASTIN Bulletin, 4.
2. Coutts S.M. (1984). Motor Insurance Rating, an Actuarial Approach. Journal of the Institute of Actuaries, 111.

3. Dutang Ch., Goulet V., Pigeon M. (2008). Actuar: An R Package for Actuarial Science. Journal of Statistical Software, 25, 7
4. Jasiulewicz H. (2005). Teoria zaufania – modele aktuarialne. AE, Wrocław.
5. Kremer E. (1999). An Extension of the BuhlmannCredibility Model. ASTIN Colloquium International Actuarial Association, Brussels, Belgium. Toyko.
6. Welfe A. (2009). Ekonometria: metody i ich zastosowanie. PWE, Warszawa.

MODELOWANIE PREFERENCJI

Jerzy Hołubiec
Grażyna Szkatuła
Dariusz Wagner
Polska Akademia Nauk w Warszawie
Andrzej Małkiewicz
Uniwersytet Zielonogórski

BADANIE PREFERENCJI POLSKIEGO ELEKTORATU W WYBORACH PARLAMENTARNYCH 2007 ROKU

Wstęp

W wyniku współpracy z politologami do analizy wyników wyborów parlamentarnych w 2007 roku przyjęto 37 cech opisujących partie polityczne. Cechy zostały podzielone na następujące grupy:

I) cechy charakteryzujące formułowane programy partii politycznych (22 cechy),
II) cechy opisujące partie polityczne oraz głoszone w czasie kampanii wyborczej hasła i wartości (7 cech),
III) cechy charakteryzujące kampanię wyborczą (4 cechy),
IV) cecha opisująca wynik wyborów (1 cecha),
V) cechy charakteryzujące elektorat negatywny poszczególnych partii politycznych (3 cechy).

Szczegółowy wykaz wszystkich rozpatrywanych cech zawiera załącznik. W tabeli 1 zamieszczono wartości cech w podziale na wyżej wymienione grupy. Cecha a34 opisująca wynik wyborów stanowi cechę decyzyjną.

Tabela 1

Wartości cech wybranych do opisu partii politycznych

Wartości, które przyjmują cechy należące do grupy I

Partie\cechy	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12
PO	2	2	1	2	1	1	1	1	1	1	1	1
PiS	1	1	1	1	1	2	2	3	3	2	1	2
LiD	3	2	2	2	2	3	2	2	2	2	2	3
PSL	2	2	1	2	2	1	2	3	3	3	3	1
Sam.	3	3	2	3	1	2	3	3	3	3	2	1
LPR	1	1	1	1	3	2	3	3	3	1	1	2

Partie\cechy	a13	a14	a15	a16	a17	a18	a19	a20	a21	a22
PO	2	2	1	1	1	1	1	1	1	1
PiS	1	1	1	1	1	2	1	2	2	2
LiD	1	2	2	2	3	1	2	3	1	3
PSL	2	2	3	1	1	3	1	3	3	2
Sam.	2	3	3	1	2	1	3	3	3	2
LPR	3	3	3	3	1	1	3	3	2	2

Wartości, które przyjmują cechy należące do grupy II

Partie\cechy	a23	a24	a25	a26	a27	a28	a29
PO	1	2	1	1	2	1	1
PiS	1	2	1	2	1	2	2
LiD	2	2	2	1	3	2	3
PSL	1	1	2	2	3	2	2
Sam.	2	2	2	3	1	2	4
LPR	2	2	2	3	1	3	5

Wartości, które przyjmują cechy należące do grupy III, IV i V

Partie\cechy	a30	a31	a32	a33	a34	a35	a36	a37
PO	1	1	1	2	1	2	3	3
PiS	1	1	1	1	1	2	3	1
LiD	2	3	3	3	2	2	1	3
PSL	2	2	2	2	2	1	1	3
Sam.	3	2	2	1	3	3	2	2
LPR	3	2	1	1	3	3	2	1

Cecha decyzyjna a34 dzieli zbiór partii politycznych na trzy klasy: 1: partia wchodzi do Sejmu z dużym poparciem elektoratu, 2: partia wchodzi do Sejmu z małym poparciem elektoratu, 3: partia nie wchodzi do Sejmu. Nie wszystkie z pozostałych 36 cech warunkowych opisujących partie polityczne są równie istotne dla rozpatrywanej klasyfikacji. Podjęto próbę rozważenia

znaczenia poszczególnych cech i rozpatrzenia możliwości zrezygnowania z mało znaczących. Zbyt mała liczba przykładów nie pozwoliła na zastosowanie metod czysto statystycznych. Zastosowano do tej analizy teorię zbiorów przybliżonych, zaproponowanych przez Zdzisława Pawlaka na początku lat 80. XX wieku. Wyznaczono zbiory reduktów, czyli takie podzbiory zbioru cech, które zapewniają taką samą jakość klasyfikacji jak pełny zbiór cech. Następnie zbiory te poddano analizie.

1. Omówienie uzyskanych wyników

Na podstawie przeprowadzonych wstępnych obliczeń, po wyznaczeniu zbioru reduktów dla rozpatrywanej klasyfikacji oraz wygenerowaniu zbioru reguł elementarnych do rozpatrywanych klas, do dalszej analizy przyjęto cechy z grupy I, III, IV i V, pomijając cechy z grupy II, jako mało istotne.

Dla nowego, zredukowanego zestawu cech wyznaczono 283 redukty, w tym 2 redukty jednoelementowe, 125 dwuelementowych, 141 trójelementowych i 15 czteroelementowych.

W tabeli 2 podano częstości występowania poszczególnych cech w zbiorze wszystkich wyznaczonych reduktów.

Tabela 2

Częstości występowania cech w zbiorze reduktów

Cecha	Częstość
a1	28
a2	21
a3	36
a4	21
a5	28
a6	26
a7	27
a8	32
a9	32
a10	18

Cecha	Częstość
a11	12
a12	25
a13	62
a14	21
a15	19
a16	52
a17	12
a18	23
a19	36
a20	25

Cecha	Częstość
a21	31
a22	32
a30	1
a31	19
a32	24
a33	26
a35	24
a36	1
a37	21

Aby wyciągnąć poprawne wnioski z analizy danych zawartych w powyższej tabeli, należy je skojarzyć z danymi przedstawionymi w tabeli 3, która zawiera częstości występowania poszczególnych cech w reduktach, w rozbiciu na redukty jedno-, dwu-, trzy- oraz czteroelementowe.

Tabela 3

Częstości występowania cech w zbiorze reduktów

Cechy: / W reduktach:	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10
1-elementowych	-	-	-	-	-	-	-	-	-	-
2-elementowych	6	7	3	7	18	7	17	3	3	13
3-elementowych	22	14	24	14	10	13	10	24	24	5
4-elementowych	-	-	9	-	-	6	-	5	5	-
Razem	28	21	36	21	28	26	27	32	32	18

Cechy: / W reduktach:	a11	a12	a13	a14	a15	a16	a17	a18	a19	a20
1-elementowych	-	-	-	-	-	-	-	-	-	-
2-elementowych	12	3	3	13	14	4	12	23	11	12
3-elementowych	-	13	59	8	5	33	-	-	25	13
4-elementowych	-	9	-	-	-	15	-	-	-	-
Razem	12	25	62	21	19	52	12	23	36	25

Cechy: / W reduktach:	a21	a22	a30	a31	a32	a33	a35	a36	a37
1-elementowych	-	-	1	-	-	-	-	1	-
2-elementowych	4	3	-	14	6	7	18	-	7
3-elementowych	27	24	-	5	18	13	6	-	14
4-elementowych	-	5	-	-	-	6	-	-	-
Razem	31	32	1	19	24	26	24	1	21

Z porównania danych z obu tych tabel wynika, że dwie cechy, cecha a36: zmiana wielkości elektoratu negatywnego w latach 2005-2007 oraz cecha a30: organizacja komitetów wyborczych, stanowią jedyne redukty jednoelementowe, co wyraźnie przesądza o ich dużej istotności dla rozpatrywanej klasyfikacji.

Następnie rozważono cechy, które występują tylko w reduktach dwuelementowych. Są to: cecha a11: polityka społeczna, a17: stosunek do Stanów Zjednoczonych oraz a18: stosunek do wojny w Iraku, przy czym cecha a18 występuje aż 23 razy. W reduktach tego typu żadna inna cecha nie występuje z taką częstością. Fakt ten wydaje się zrozumiały biorąc pod uwagę, że społeczeństwo polskie nie miało jednoznacznego stanowiska, co do wysyłania polskich żołnierzy do Iraku; istotną rolę odgrywały tu informacje o ponoszonych stratach.

Przeanalizowano, które cechy występują w reduktach dwuelementowych razem z cechami: a11, a17 lub a18. Wyniki zamieszczono w tabeli 4. Znak x oznacza, że istnieje redukt dwuelementowy zawierający daną cechę i odpowiednio jedną z wyżej wymienionych cech.

Tabela 4

Cechy występujące w reduktach dwuelementowych

Cechy:	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a12	a13	a14	a15
a11					x		x			x		x	x	x
a17					x		x			x		x	x	x
a18		x		x	x	x	x	x	x	x	x	x	x	x

Cechy:	a16	a19	a20	a21	a22	a30	a31	a32	a33	a35	a36	a37
a11	x	x	x				x			x		
a17	x	x	x				x			x		
a18	x	x		x	x		x	x	x	x		x

W reduktach dwuelementowych zawierających cechę a11: polityka społeczna występowały głównie cechy opisujące program partii politycznych, co jest zrozumiałe.

Cecha a17: stosunek do Stanów Zjednoczonych, występowała w reduktach dwuelementowych głównie z cechami dotyczącymi polityki zagranicznej Polski (a14, a15, a16, a17, a18 lub a19) oraz prowadzonej polityki gospodarczej (a2, a5, a6, a7, a8, a9 lub a10).

W reduktach dwuelementowych zawierających cechę a18: stosunek do wojny w Iraku, występowały cechy odnoszące się bezpośrednio do spraw gospodarczych (a2, a4, a5, a6, a7, a8, a9, a10, a11, a12, a13 lub a14) oraz cechy dotyczące ogólnie rozumianej polityki zagranicznej (a15, a16 lub a17). Nie występowały natomiast cechy: a1: sposób rządzenia, a3: stosunek do lustracji oraz a20: walka z korupcją. Z racji tego, że same stanowią redukty jednoelementowe nie występowały cechy: a30: organizacja komitetów wyborczych oraz a36: zmiana wielkość elektoratu negatywnego w latach 2005-2007.

Z przeprowadzonej analizy wynika, że polityka zagraniczna państwa w połączeniu z polityką gospodarczą miały duży wpływ na decyzje wyborców.

W reduktach trójelementowych największą częstością występowania charakteryzuje się cecha a13: stosunek do administracji i samorządów, pojawia się bowiem aż w 59 reduktach. Biorąc pod uwagę, że w reduktach dwuelementowych występuje tylko 3 razy, a w reduktach czteroelementowych nie pojawia się wcale, można uznać, że jest to również cecha istotna.

Następnie przeanalizowano, które cechy występują w reduktach trójelementowych zawierających cechę a13. Wyznaczono częstości ich występowania w tych reduktach, wyniki zamieszczono w tabeli 5.

Tabela 5

Częstości cech w reduktach trójelementowych z cechą a13

Cechy:	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15
a13	6	6	9	6	5	6	6	10	10	-	-	-	-	2	1

Cechy:	a16	a17	a18	a19	a20	a21	a22	a30	a31	a32	a33	a35	a36	a37
a13	1	-	-	4	2	8	10	-	1	7	6	5	-	6

Jeśli chodzi o cechy opisujące program partii politycznych, to cecha a13 występowała najczęściej wraz z cechą a8: podatki, a9: ochrona zdrowia oraz cechą a22: charakter państwa. To zestawienie wydaje się uzasadnione, bowiem te trzy cechy dotyczą polityki państwa mającej bezpośredni wpływ na działalność samorządów.

Z cechą a13 w rozpatrywanych reduktach często występowały również cechy dotyczące prowadzenia kampanii wyborczej: a32: ukierunkowanie kampanii wyborczej i a33: charakterystyka kampanii wyborczej. Było to uzasadnione faktem, iż społeczności lokalne były zainteresowane, czy partie polityczne w kampanii wyborczej starały się do nich dotrzeć ze swoimi programami.

Analizując zestaw cech wybrany do opisu partii politycznych można zauważyć, że cecha a2: ocena dwulecia 2005-2007 oraz cecha a37: ocena poprzedniego rządu mają w zasadzie taki sam rozkład wartości, można więc jedną z nich pominąć.

Podsumowanie

W stosunku do poprzednich prac autorów, istotnym elementem było wprowadzenie nowej grupy cech charakteryzujących elektorat negatywny poszczególnych partii politycznych startujących w wyborach parlamentarnych w 2007 roku (grupa V).

Przeprowadzona analiza istotności cech wybranych do opisu partii politycznych pozwoliła lepiej poznać zależności przyczynowo-skutkowe pomiędzy preferencjami wyborców a wynikiem wyborów.

W wyniku przeprowadzonej analizy utworzonego zbioru reduktów stwierdzono między innymi, że do cech istotnie decydujących o wynikach wyborów należy cecha a36, ilustrująca zmiany wielkości elektoratu negatywnego poszczególnych partii w latach 2005-2007 (z grupy V) oraz cecha a30, dotycząca organizacji komitetów wyborczych (z grupy III).

Odczucia społeczne, a także analiza wyników wyborów wykonana przez ekspertów zagranicznych (S. Brams), wskazywały, iż ostatnie wybory miały charakter „protest voting". Te odczucia potwierdza zresztą stwierdzony mało istotny wpływ cech z grupy II na wynik wyborów.

Zmniejszenie wymiarowości rozpatrywanego zadania poprzez pominięcie cech mniej istotnych dla rozpatrywanej klasyfikacji pozwala tworzyć mniejszą liczbę reguł decyzyjnych, opartych na cechach bardziej istotnych.

ZAŁĄCZNIK 1

Cechy charakteryzujące kampanię wyborczą w wyborach do Sejmu 2007r.

I. Cechy charakteryzujące programy partii politycznych

a1- Sposób rządzenia
1- umacnianie państwa i radykalna walka z patologiami — PiS, LPR
2- budowanie konsensusu społecznego — PO, PSL
3- brak konkretnych propozycji — LiD, Sam.

a2- Ocena dwulecia 2005-2007
1) okres skutecznego uzdrawiania państwa — PiS, LPR
2) okres narastającej arogancji władzy i podważania demokracji — PO, PSL, LiD
3) stanowisko ambiwalentne — Sam.

a3 – Stosunek do lustracji
1) potrzeba konsekwentnej lustracji — PiS, PO, PSL, LPR
2) zakończenie lustracji — LiD, Sam.

a4 – Stosunek do kościoła i religii
1) potrzeba ścisłego współdziałania państwa i Kościoła — PiS, LPR
2) rozdzielenie kościoła i państwa, szacunek dla wartości religijnych — PO, PSL, LiD
3) brak stanowiska — Sam.

a5 – Walka z bezrobociem
1) obniżenie pozapłacowych kosztów pracy, w tym obniżenie składki rentowej ZUS — PO, PiS, Sam
2) ulgi w opłatach na ubezpieczenie społeczne i podatku dochodowym — LiD, PSL
3) brak konkretnych propozycji — LPR

a6 – Edukacja i nauka
1) spójna podstawa programowa dla wszystkich przedmiotów przy jednoczesnym zwiększaniu autonomii szkół, aby dostosować się do potrzeb rynku — PO, PSL
2) zwiększanie nakładów finansowych na edukację i naukę — PiS, LPR, Sam
3) obniżenie wieku szkolnego oraz rozdział kościoła od szkolnictwa — LiD

a7 – Gospodarka
1- wolność gospodarcza oparta na własności prywatnej — PO
2- rozwój małych i średnich przedsiębiorstw, w tym ulgi i dostępność kredytów oraz wprowadzenie systemu poręczeń i gwarancji dla przedsiębiorców — PiS, Lid, PSL
3- wstrzymanie prywatyzacji i lustracja sprywatyzowanych przedsiębiorstw — LPR, Sam.

a8 – Podatki
4- uproszczenie podatków poprzez wprowadzenie podatku linowego — PO
5- abolicja podatkowa dla Polaków wracających z zagranicy — LiD
6- obniżenie stawki podatku od osób fizycznych ze szczególnym uwzględnieniem osób najbiedniejszych, polityka pro rodzinna, w tym wprowadzenie ulg mieszkaniowych — Sam., LPR, PiS, PSL

a9 – Ochrona zdrowia
1- podział NFZ na kilka konkurencyjnych funduszy i określenie koszyka świadczeń gwarantowanych — PO
2- wprowadzenie odpłatności za usługi medyczne — LiD
3- w miarę powszechny dostęp do podstawowej opieki zdrowotnej, w tym utworzenie funduszu charytatywnego i przekazanie części środków z Funduszu Pracy na ochronę zdrowia — LPR, Sam., PSL, PiS

a10 – Bezpieczeństwo wewnętrzna państwa
1- stworzenie scentralizowanego i skoordynowanego systemu zwalczania najpoważniejszych zagrożeń państwa — PO, LPR
2- polepszenie uzbrojenia i wyposażenia służb mundurowych oraz stworzenie ogólnopolskiego systemu łączności służb ratowniczych — PiS, LiD
3- brak propozycji — PSL, Sam.

a11 – Polityka społeczna
1- wprowadzenie ulg rodzinnych, nowych świadczeń dla dzieci i wydłużenie urlopu macierzyńskiego dla obojga rodziców — PO, PiS, LPR
2- wprowadzenie dodatków mieszkaniowych i reforma emerytalna, wzrost zasiłków i pomocy najuboższym — LiD, Sam.
3- wprowadzenie podatku rodzinnego i wspólnego rozliczania się rodzin — PSL

a12 – Rolnictwo
1- korzystanie z funduszy unijnych przy rozbudowie polskiego rolnictwa — PO, PSL, Sam.
2- wspieranie niskotowarowych i nisko rentowych gospodarstw i umacnianie gospodarstw rodzinnych — Pis, LPR
3- przywrócenie rent strukturalnych i urealnienie składek KRUS — LiD

a13 – Stosunek do administracji i samorządów
1- decyzje administracyjne powinny być zgodne z literą prawa i ich transparentność — LiD, PiS
2- decentralizacja i wzmocnienie podstaw majątkowych samorządów — PO, PSL, Sam
3- ograniczenie samodzielności samorządów — LPR

a14 – Wprowadzenie euro
1- wprowadzenie Polski do strefy Euro później niż w 2015 r. — PiS
2- wprowadzenie Polski do strefy Euro jak najszybciej — PO, PSL, LiD
3- zachowanie odrębnej waluty — LPR, Sam

a15 – Polityka zagraniczna
1- popieranie Ukrainy, Gruzji, Mołdawii w członkostwie w UE, wzmacnianie roli Polski w kontaktach z sąsiadami, współpraca w ramach trójkąta Weimarskiego — PiS PO
2- dobre stosunki z wszystkimi sąsiadami, zwłaszcza Ukrainą i Niemcami — LiD
3- poprawa stosunków z Rosją — Sam., LPR, PSL

a16 –Stosunek do UE
1- pogłębianie związku z UE i współodpowiedzialność za jej rozwój — PO, PiS, PSL, Sam.
2- wspólna polityka zagraniczna i stanowisko wobec Rosji — LiD
3- ograniczenie współpracy z UE — LPR

a17 – Stosunek do USA
1- pielęgnowanie partnerstwa w ramach bezpieczeństwa i strefy gospodarczej — PO, PiS, LPR, PSL
2- rozluźnienie związków — Sam.
3- brak propozycji — LiD

a18 – Stosunek do wojny w Iraku
1- wywiązanie się z misji bez przedłużania obecności polskich wojsk, powiązane z zaangażowaniem gospodarczym i politycznym — PO, LiD, Sam., LPR
2- dalszy udział polskich wojsk w misji oraz wzmacnianie pozycji i bezpieczeństwa Polski — PiS
3- brak propozycji — PSL

a19 – Bezpieczeństwo energetyczne kraju
1- dywersyfikacja źródeł energii i odwołanie się do własnych zasobów, w tym rozwój czystej energii — PO, PiS, PSL
2- promowanie wspólnej polityki energetycznej z UE — LiD
3- znormalizowanie współpracy z Rosją — Sam., LPR

a20 – Walka z korupcją
1- przejrzyste procedury administracyjne — PO
2- kontynuowanie działalności CBA — PiS
3- brak propozycji — PSL, LiD, Sam., LPR

a21 – Stosunek do aborcji
1- aborcja tylko wówczas gdy ciąża zagraża życiu i zdrowi matki — PO, LiD
2- całkowity zakaz aborcji — PiS, LPR
3- brak propozycji — PSL, Sam.

a22-Charakter państwa
1- tanie, zdecentralizowane i prospołeczne — PO
2- solidarne i socjalne — PiS, Sam., LPR, PSL
3- praworządne i ponadpartyjne — LiD

II. Cechy opisujące partie polityczne oraz głoszone hasła i wartości.

a23 - Organizacja partii politycznej
1- partia z mocno rozbudowanymi strukturami — PSL, PO, PiS
2- partia ze słabo rozbudowanymi strukturami — LiD, Sam., LPR

a24 - Doświadczenie w polityce
1- partia z dużym stażem politycznym — PSL
2- partia z małym stażem politycznym — PO, PiS, LiD, LPR, Sam.

a25 - Udział w poprzednim parlamencie
1- tak, na silnych pozycjach — PO, PiS
2- tak, na słabych pozycjach — LiD, PSL, LPR, Sam.

a26 - Wizerunek Partii
1- partia nowoczesna — PO, LiD
2- partia konserwatywna — PSL, PiS
3- partia koniunkturalna — LPR, Sam.

a27 - Pozycja lidera
1. wyrazisty, dominujący nad partią — PiS, LPR, Sam.
2. wyrazisty, współpracujący z partyjnymi kolegami — PO
3. mało wyrazisty — LiD, PSL

a28 - Deklarowane poglądy
1- liberalne — PO
2- centrowe — PiS, PSL, LiD, Sam.
3- prawicowe — LPR

a29 - Głoszone hasła i wartości
1- liberalno-socjalne — PO
2- narodowo-socjalne — PiS, PSL
3- lewicowo-socjalne — LiD
4- socjalliberalne — Sam.
5- narodowo-chrześcijańskie — LPR

III. Cechy charakteryzujące kampanię wyborczą

a30 - Organizacja komitetów wyborczych
1- profesjonalna — PO, PiS
2- niedopracowana — PSL, LiD
3- amatorska — LPR, Sam.

a31 - Formy prowadzenia kampanii wyborczej
1- prowadzona przy użyciu najnowszych rozwiązań medialnych — PO, PiS
2- prowadzona z naciskiem na bezpośrednie dotarcie do wyborcy — PSL, LPR, Sam.
3- prowadzona z naciskiem na internet — LiD

a32 - Ukierunkowanie kampanii wyborczej
1- skierowana do ogółu wyborców — PiS, PO, LPR
2- skierowana do mieszkańców wsi i małych miast — PSL, Sam.
3- skierowana do ludzi młodych i twardego elektoratu lewicowego — LiD

a33 - Charakterystyka kampanii wyborczej
1- kampania agresywna, oparta na hasłach i wartościach — PiS, LPR, Sam.
2- kampania spokojna, oparta na hasłach i wartościach — PO, PSL
3- kampania spokojna, oparta na znanych osobistościach — LiD

IV. Cecha decyzyjna

a34 - Wynik wyborów
1) wejście do Sejmu z dużym poparciem elektoratu — PO, PiS
2) wejście do Sejmu z małym poparciem elektoratu — LiD, PSL
3) nie wejście do Sejmu — LPR, Sam.

V. Cechy charakteryzujące elektorat negatywny poszczególnych partii

a35 – Wielkość elektoratu negatywnego
1) mała — PSL
2) średnia — PO, PiS, LiD
3) duża — Sam., LPR

a36 – Zmiana wielkość elektoratu negatywnego w latach 2005-2007
1) bez większych zmian — PSL, LiD
2) mały wzrost — Sam., LPR
3) duży wzrost — PO, PiS

a37 – Ocena poprzedniego rządu

1) pozytywna — PiS, LPR
2) słabo negatywna — Sam.
3) silnie negatywna — PO, LiD, PSL

Literatura

1. Garbień A., Małkiewicz A. (2008). Uwarunkowania wyników wyborów sejmowych w Polsce w 2007 (maszynopis).
2. Greco S., Matarazzo B., Slowinski R. (2001). Rough Sets Theory for Multicriteria Decision Analysis. European Journal of Operational Research, 129, 1-47.
3. Hołubiec J., Szkatuła G., Wagner D. (1998). Próba zastosowania uczenia maszynowego do prognozowania wyników głosowań sejmowych. W: Metody i zastosowania badań operacyjnych. Red. T. Trzaskalik. AE, Katowice, 117-127.
4. Hołubiec J., Szkatuła G., Wagner D. (2002). Modelowanie preferencji wyborców w postaci reguł decyzyjnych. W: Modelowanie preferencji a ryzyko ’01. Red. T. Trzaskalik. AE, Katowice, 133-144.
5. Hołubiec J., Szkatuła G., Wagner D. (2002). Analiza obietnic wyborczych ugrupowań politycznych. W: Metody i techniki analizy informacji i wspomagania decyzji. Badania operacyjne i systemowe wobec wyzwań XXI wieku. Red. Z. Bubnicki, O. Hryniewicz, R. Kulikowski. AOW EXIT, Warszawa, III, 63-74.
6. Hołubiec J., Szkatuła G., Wagner D. (2002). Modelling Electorate Preferences by Machine Learning. In: Proceedings of 8th IEEE International Conference on Methods and Models in Automation and Robotics 2002. 2-5.09.2002. Technical University, Szczecin, 1383-1388.
7. Hołubiec J., Szkatuła G., Wagner D. (2003). New Aspects in Electorate Preferences Modelling Using Machine Learning. In: Proceedings of the Conference: 9th IEEE International Conference on Methods and Models in Automation and Robotics Międzyzdroje, 25-28 August 2003, 1271-1276.
8. Hołubiec J., Małkiewicz A., Szkatuła G., Wagner D. (2003). Próba uwzględnienia dodatkowych atrybutów w analizie kampanii wyborów do Sejmu w 2001 r. W: Modelowanie preferencji a ryzyko '03. Red. T. Trzaskalik. AE, Katowice, 133-150.
9. Pawlak Z. (1982). Rough Sets. International Journal of Computer and Information Sciences, 11, 5, 341-356.
10. Pawlak Z. (1991). Rough Sets. Theoretical Aspect of Reasoning about Data. Kluwer Academic Publishers.

11. Szkatuła G., Hołubiec J., Wagner D. (1997). Machine Learning from Examples for Forecasting Voting Behaviour. In: Methods and Models in Automation and Robotics. Międzyzdroje 26-29. 08.1997. Technical University, Szczecin, 385-389.
12. Szkatuła G., Hołubiec J., Wagner D. (2000). Forecasting Voting Behaviour Using Machine Learning – Poland in Transition. Annals of Operations Research, 97, 31-41.
13. Szkatuła G., Wagner D. (2003). Programmes of Parties Versus Their Location on the Political Scene. Application of Decision Rules to Describe the Differences. W: Group Decisions and Voting. Red. J. Kacprzyk, D. Wagner. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 227-238.

Agnieszka Kowalska-Styczeń
Politechnika Śląska w Gliwicach

WPŁYW RODZAJU OTOCZENIA NA DECYZJE KONSUMENCKIE

Wprowadzenie

Zachowania konsumentów są procesem bardzo złożonym. Wynika to w dużej mierze z tego, że są one kształtowane pod wyraźnym wpływem otoczenia. Konsument żyje w konkretnym środowisku, do którego się przystosowuje i które również w pewnym stopniu zmienia swoim postępowaniem. Na jego zachowanie mają wpływ zarówno osoby należące do bliższego jak i dalszego otoczenia.

Jedną z najważniejszych grup czynników kształtujących zachowanie konsumentów są wpływy interpersonalne, czyli oddziaływania pochodzące od innych konsumentów [1]. Podatność konsumenta na wpływy innych osób to jego potrzeba identyfikowania się ze znaczącymi osobami lub poprawa własnego wizerunku w ich oczach, która przejawia się przez:

- nabywanie określonych produktów i określonych marek,
- dążenie w zakresie decyzji nabywczych do dostosowania się do oczekiwań osób znaczących,
- gromadzenie wiedzy o produktach poprzez obserwację innych konsumentów i pozyskiwanie od nich informacji [2].

Decyzja o zakupie produktów lub usług nie jest zatem tylko decyzją poszczególnych osób. Najczęściej mamy do czynienia z grupowym podejmowaniem decyzji, czego przykładem są decyzje rodzinne lub decyzje w gronie przyjaciół, znajomych, którzy z czasem zastępują tradycyjną rodzinę z rodzicami. Podejmowania decyzji w takiej grupie jest procesem wzajemnego oddziaływania osób ją tworzących. Natomiast siła takiego oddziaływania zależy od trzech grup czynników związanych z:

- grupą lub jednostką wywierającą wpływ,
- konsumentem, który jest pod wpływem oddziaływania,
- nabywanym produktem [7].

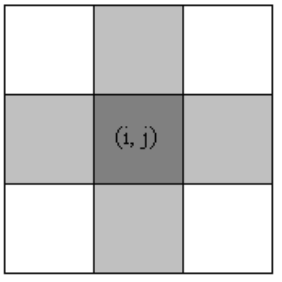
Wśród najważniejszych czynników związanych z grupą wywierającą wpływ należy wymienić:
- dużą zwartość grupy,
- wielkość grupy,
- wiarygodność grupy, doświadczenie i wiedzę.

Grupą wywierającą najsilniejsze oddziaływanie jest grupa zwarta (czyli członkowie grupy są sobie bliscy) oraz posiadająca odpowiednią wiedzę na interesujący konsumenta temat. Taką grupę stanowi rodzina lub grupa bliskich przyjaciół. To właśnie w takiej grupie spędzamy większość swojego życia i to w takiej grupie podejmowane jest większość decyzji o zakupowych [5].

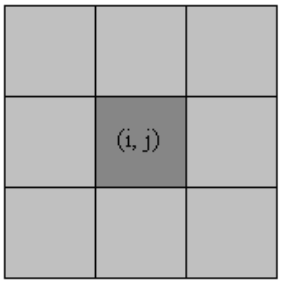
W opracowaniu podjęto zatem próbę symulowania wpływu wielkości najbliższego otoczenia konsumenta na podejmowanie decyzji. Do symulacji wykorzystano automat komórkowy dwuwymiarowy [3; 4; 9] z otoczeniem von Neumanna i Moorea.

Automat komórkowy można zdefiniować jako obiekt matematyczny składający się:
- z sieci komórek $\{i\}$ przestrzeni D-wymiarowej,
- ze zbioru stanów pojedynczej komórki $\{s_i\}$, zawierającego k elementów,
- z reguły F określającej stan komórki w chwili t+1 w zależności od stanu w chwili t-tej komórki i komórek ją otaczających $s_i(t+1) = F(\{s_j(t)\})$, $j \in O(i)$, gdzie $O(i)$ jest otoczeniem i-tej komórki [4].

Zatem jest to dynamiczny model matematyczny procesów zachodzących w czasie. D-wymiarowa przestrzeń, w której zachodzi ewolucja automatu komórkowego podzielona jest na jednakowe komórki, z których każda może przyjąć jeden ze stanów (przy czym liczba stanów jest skończona). Otoczenie dla każdej komórki jest takie samo. Stan komórki zmienia się w czasie zgodnie z regułą F i zależy od jej poprzedniego stanu i stanu jej sąsiadów, czyli otoczenia. Ważne zatem parametry dla automatu komórkowego to: wymiar sieci D, ilość stanów pojedynczej komórki k. Często używanym parametrem jest również promień r, którego wielkość zależy od postaci otoczenia danego automatu komórkowego. Jeśli otoczeniem $O(i)$ są najbliżsi sąsiedzi i-tej komórki to $r = 1$.



Rys. 1. Otoczenie von Neumanna komórki (i, j) automatu dwuwymiarowego dla r = 1 (otoczenie 4-elementowe)



Rys. 2. Otoczenie Moore'a komórki (i, j) automatu dwuwymiarowego dla r = 1 (otoczenie 8-elementowe)

Środowisko konsumentów w poniższych symulacjach przedstawione jest jako dyskretna przestrzeń dwuwymiarowa, którą tworzy skończona kwadratowa sieć komórek. Ustala się zatem warunki brzegowe, które opisują zachowanie się komórek na brzegach sieci. Komórki reprezentujące konsumentów, podczas symulacji po dojściu do brzegu sieci pojawiają się z drugiej strony sieci. Komórki, które są na skraju sieci mają za sąsiadów komórki po drugiej stronie sieci. Taki sposób modelowania oddaje nasze rzeczywiste otoczenie – żyjemy, przemieszczamy się na kuli ziemskiej, która jest swego rodzaju przestrzenią periodyczną.

1. Model zachowań konsumentów

Przestrzeń konsumentów przedstawiono za pomocą kwadratowej sieci *n* na *n* (automat komórkowy dwuwymiarowy). Dana jest gęstość zaludnienia *p* (w %) przez konsumentów zajmujących komórki sieci, pozostałe komórki są puste. Każdego konsumenta reprezentowanego przez komórkę (i, j) charakteryzują dwie zmienne. Pierwsza zmienna określa kierunek, w którym jest zwrócony (północ, południe, wschód, zachód), a druga zmienna to lista preferencji o *m* elementach. System rozwija się w czasie poprzez daną liczbę etapów lub do momentu, gdy żaden konsument nie zmienia już swoich preferencji.

Pojedyncza komórka przyjmuje *k* stanów, w zależności od ilości preferencji konsumentów, natomiast reguła wykorzystana w tym automacie składa się z dwóch etapów przejściowych.

Pierwszy etap przejściowy obejmuje sprawdzenie preferencji otoczenia i ewentualną zmianę preferencji przez wybranego konsumenta. Drugi etap to przemieszczanie się jednostek. Zmiana jest jednokierunkowa i wybrany konsument przejmuje preferencję, która dominuje w jego otoczeniu. Jeśli brak do-

minującej preferencji, to nie dochodzi do zmiany i jest to drugi etap symulacji, w którym konsument przemieszcza się do najbliższej pustej komórki, w której kierunku jest zwrócony.

Sieć obrazującą konsumentów zapełnia się przez:

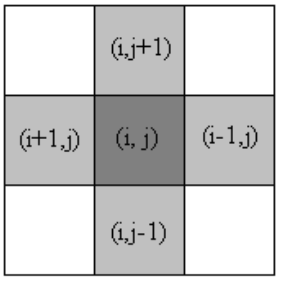
- puste komórki,
- komórki zajęte, które charakteryzują dwa elementy: kierunek w jaki zwrócona jest jednostka oraz preferencje konsumenta.

W celu przeprowadzenia symulacji podaje się następujące wartości:

- n – rozmiar kraty,
- p – gęstość zaludnienia,
- m – liczba możliwości wyborów jakie ma konsument,
- t – liczba etapów.

Symulacje przeprowadzono dla różnego rodzaju otoczenia, ponieważ grupa, która wpływa na nasze decyzje może mieć różną wielkość. Wybrano otoczenie von Neumanna dla promienia r = 1 (czyli sąsiedztwo 4-elementowe) i r = 2 (sąsiedztwo 12-elementowe) oraz otoczenie Moore'a dla r = 1 (8 – sąsiadów) i r = 2 (24 – sąsiadów). Do napisania programu symulującego wykorzystano środowisko Delphi.

Pierwsze symulacje przeprowadzono dla otoczenia von Neumanna, gdy r = 1. Otoczenie O_{ij} konsumenta k_{ij}, którego reprezentuje komórka *(i, j)* tworzą komórki *(i, j+1), (i+1, j), (i, j-1) oraz (i-1, j)* (rys. 3).



Rys. 3. Otoczenie von Neumanna komórki (i, j) dla r = 1

Zachowanie konsumenta k_{ij} w czasie $t+1$ zależy od jego zachowania i zachowania jego otoczenia w czasie t, czyli

$$k_{ij}(t+1) = \left(k_{ij}(t); O_{ij}(t)\right)$$

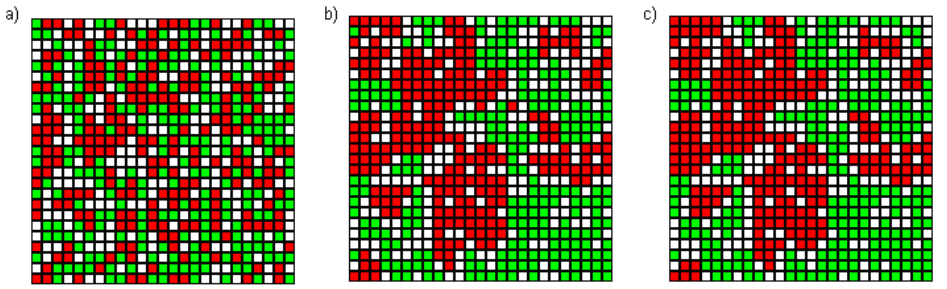
Przyjęto następujące wartości:

- $n = 25$,
- $p = 0{,}7$,
- $m = 2$,
- $t = 4$.

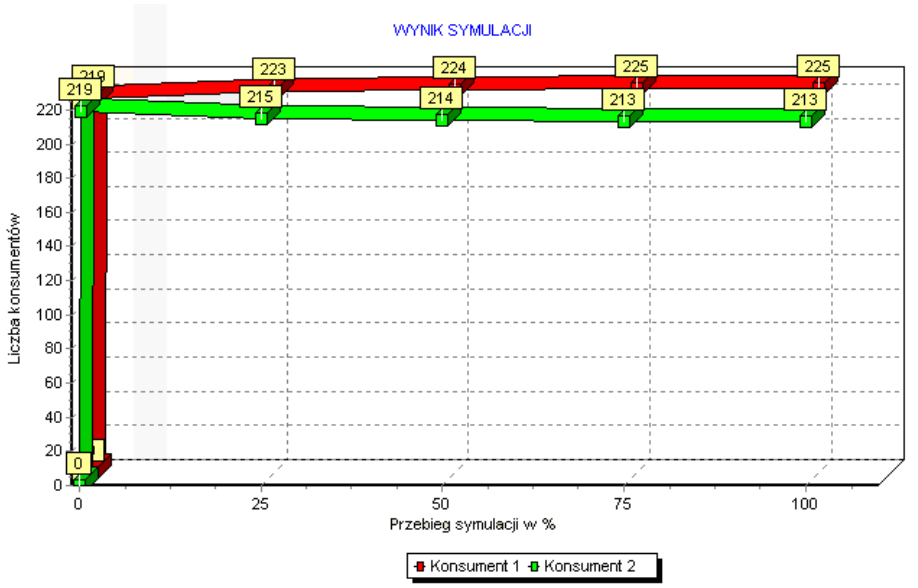
Symulacje przeprowadzono od stanu początkowego t = 0, który stanowi losowe ustawienia komórek. Stan początkowy dla wszystkich symulacji jest taki sam. Wyniki symulacji w kolejnych etapach przedstawione zostały na rys. 4 i 5. Rysunek 5 pokazuje, jak zmieniała się liczba konsumentów każdej z preferencji podczas symulacji. Dwa kolory: czerwony i zielony oznaczają odmienne preferencje konsumentów, a kolor biały oznacza puste komórki.

Wnioski dla symulacji przeprowadzonej dla $1 \leq t \leq 4$ etapów:

- tworzą się skupiska konsumentów o tych samych preferencjach, nie ma tendencji do przyjmowania wspólnego systemu zachowań,
- liczebność konsumentów o danej preferencji nie ulega znacznej zmianie (rys. 5).



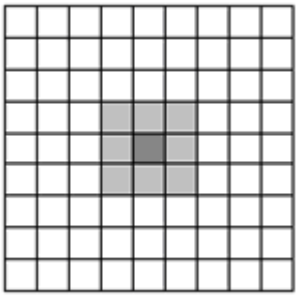
Rys. 4. Wynik symulacji dla stanu początkowego: losowe ustawienie konsumentów – a, t = 2 etapy – b, t = 4 etapy – c



Rys. 5. Wynik symulacji od stanu początkowego do t = 4 etapy

W symulacjach przyjęto t = 4, ponieważ dla t = 4 system jest stabilny i żaden z konsumentów nie zmienia już swojej preferencji. Przeprowadzono również symulacje dla większej ilości preferencji. Wyniki były analogiczne jak w przypadku symulacji dla $m = 2$.

Kolejną symulację przeprowadzono dla otoczenia Moore'a, gdy r = 1. Sąsiedztwo jest wówczas ośmioelementowe (rys. 6).



Rys. 6. Otoczenie Moore'a automatu komórkowego dwuwymiarowego dla r = 1

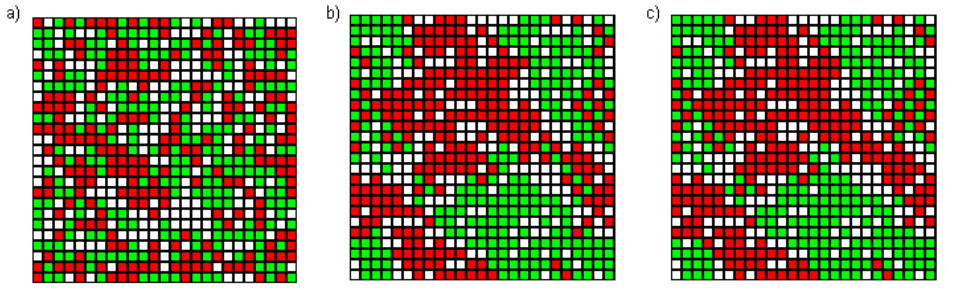
Przyjęto następujące wartości:

- $n = 25$,
- $p = 0{,}7$,
- $m = 2$,
- $t = 4$.

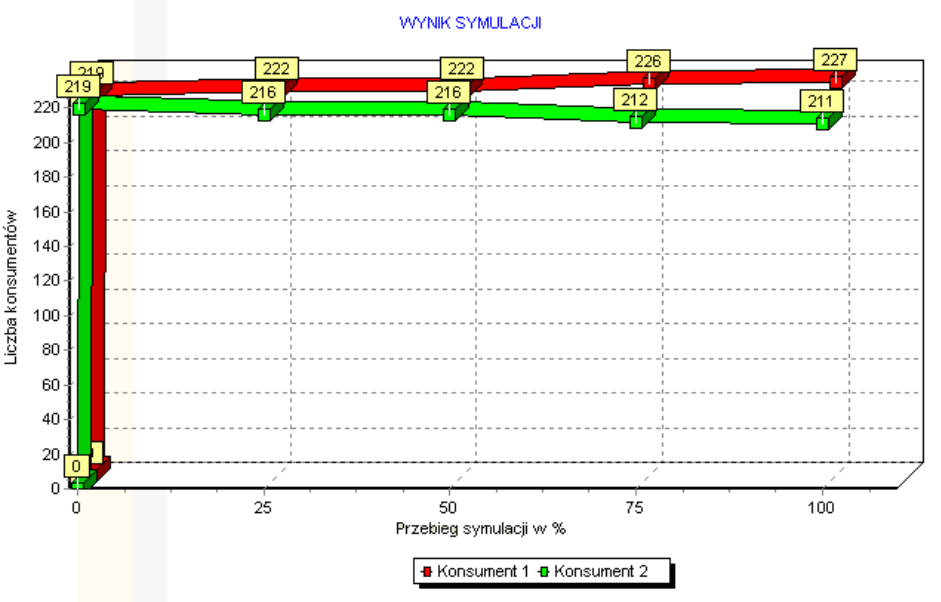
W symulacjach przyjęto t = 4, ponieważ dla t = 4 system jest stabilny i żaden z konsumentów nie zmienia już swojej preferencji.

Wnioski dla symulacji przeprowadzonej dla $1 \leq t \leq 4$ etapów:

- tworzą się skupiska konsumentów o tych samych preferencjach, nie ma tendencji do przyjmowania wspólnego systemu zachowań,
- skupiska konsumentów o tych samych preferencjach nie są tak zwarte jak w przypadku poprzedniej symulacji,
- liczebność konsumentów o danej preferencji nie ulega znacznej zmianie (rys. 8).



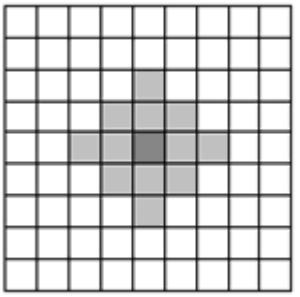
Rys. 7. Wynik symulacji dla stanu początkowego: losowe ustawienie konsumentów – a, t = 2 etapy – b, t = 4 etapy – c



Rys. 8. Wynik symulacji od stanu początkowego do t = 4 etapy

Z przeprowadzonych symulacji wynika, że konsumenci tworzą skupiska osób, dokonujących tych samych wyborów – brak tendencji do przyjmowania tego samego systemu zachowań. Dla otoczenia 4-elementowego i 8-elementowego dość szybko konsumenci wybierają ostatecznie swoją preferencję. Dyskusja odbywa się maksymalnie przez cztery etapy symulacji. Oznacza to, że aby ostatecznie się zdecydować, konsument zmienia swoją pozycję w systemie maksymalnie cztery razy, czyli wymienia poglądy w domu z rodziną, jeśli to nie wystarcza do podjęcia decyzji, podejmuje dalsze dyskusje z kolegami w pracy, ze znajomymi i przyjaciółmi.

Następne symulacje przeprowadzono dla otoczenia von Neumanna, gdy r = 2 (rys. 9).

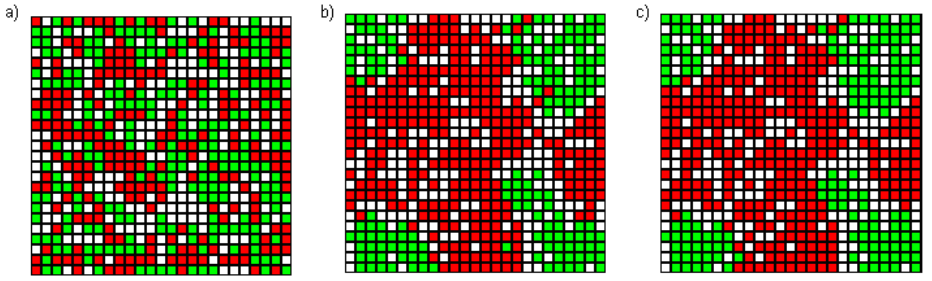


Rys. 9. Otoczenie von Neumanna automatu komórkowego dwuwymiarowego dla r = 2

W celu przeprowadzenia symulacji przyjęto następujące wartości:

– $n = 25$,
– $p = 0{,}7$,
– $m = 2$,
– $t = 8$.

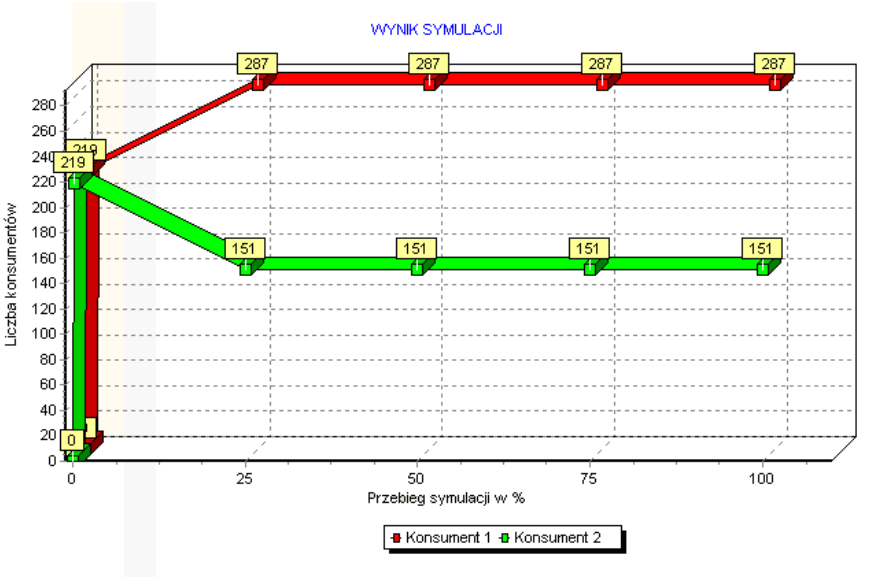
Wyniki takiej symulacji przedstawiają rys. 10 i 11.



Rys. 10. Wynik symulacji dla stanu początkowego: losowe ustawienie konsumentów – a, t = 4 etapy – b, t = 8 etapy – c

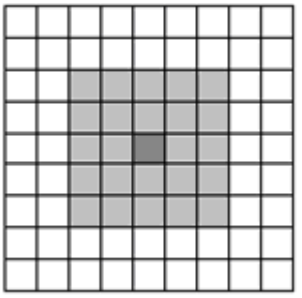
Wnioski dla symulacji przeprowadzonej dla $1 \leq t \leq 8$ etapów:

- tworzą się bardzo wyraźne skupiska konsumentów o tych samych preferencjach,
- liczebność konsumentów o danej preferencji ulega znacznej zmianie jedna z preferencji wyraźnie dominuje (rys. 11).



Rys. 11. Wynik symulacji od stanu początkowego do t = 8 etapy

Przeprowadzono również symulacje dla otoczenia Moore'a, gdy r = 2 (rys. 12).

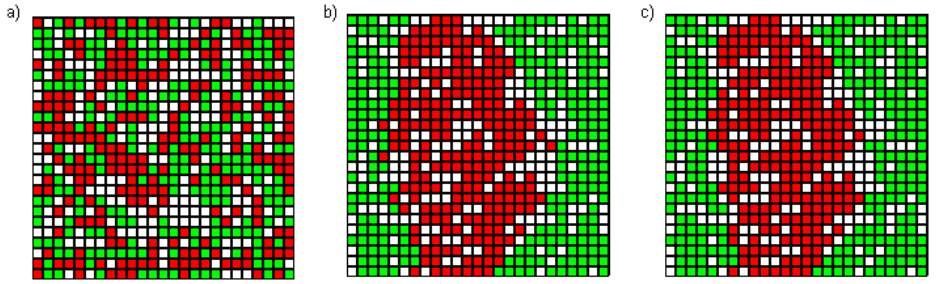


Rys. 12. Otoczenie Moore'a automatu komórkowego dwuwymiarowego dla r = 2

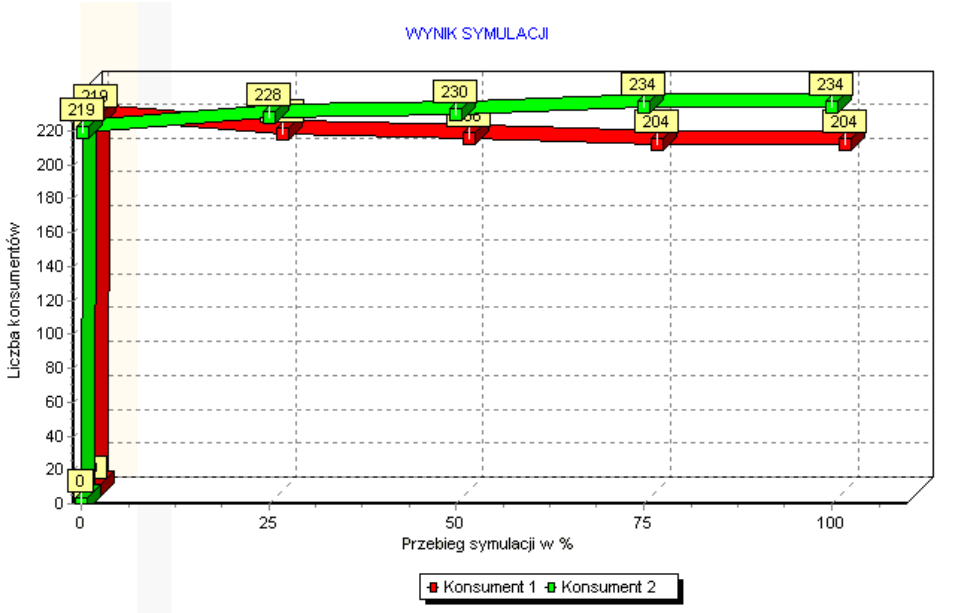
W celu przeprowadzenia symulacji przyjęto następujące wartości:

- $n = 25$,
- $p = 0{,}7$,
- $m = 2$,
- $t = 8$.

Wyniki symulacji przedstawione zostały na rys. 13 i 14.



Rys. 13. Wynik symulacji dla stanu początkowego: losowe ustawienie konsumentów –a, t = 4 etapy –b, t = 8 etapy – c



Rys. 14. Wynik symulacji od stanu początkowego do t = 8 etapy

Wnioski dla symulacji przeprowadzonej dla $1 \leq t \leq 8$ etapów:

- tworzą się bardzo wyraźne skupiska konsumentów o tych samych preferencjach,
- przeprowadzając wiele symulacji okazuje się, że liczebność konsumentów o danej preferencji ulega zmianie oraz w każdej symulacji dominuje jedna z preferencji.

Po przeprowadzeniu symulacji dla otoczeń bardziej licznych okazuje się, że konsumenci podobnie jak w modelach dla r = 1 tworzą skupiska osób, dokonujących tych samych wyborów. Grupy te są jednak bardziej liczne. Odmiennie niż w poprzednich symulacjach jedna z preferencji dominuje. Ponadto, potrzeba więcej czasu na podjęcie ostatecznej decyzji o wyborze danej preferencji. Dyskusja odbywa się za każdym razem w szerszym gronie niż w przypadku symulacji dla r = 1.

Podsumowanie

Decyzje dotyczące nabywania towarów konsumpcyjnych opierają się w znacznym stopniu na wywieraniu wpływów i głosowaniu w najbliższym otoczeniu konsumenta [8]. Powoduje to powstanie nowego trendu projektowania strategii marketingowych na podstawie analizy grup, a nie tylko pomiarów dotyczących jednostek.

W tym opracowaniu zasymulowano zatem, jak otoczenie wpływa na decyzje konsumentów. Wykorzystano automaty komórkowe, które dzięki swojej budowie pozwalają stworzyć taki model.

W symulacjach wykorzystano automat komórkowy dwuwymiarowy. Stworzono cztery modele symulacyjne uwzględniające dwa najczęściej stosowane otoczenia w modelach dwuwymiarowych: otoczenie von Neumanna i Moore'a oraz promień otoczenia r = 1 i r = 2. Taki dobór otoczeń powoduje, że w modelu mamy do czynienia, tak jak w rzeczywistości, z różną wielkością grup, jakie mają na nas wpływ (w rzeczywistości rodziny i grupy przyjaciół mogą liczyć kilka lub nawet kilkanaście osób). Zachowanie konsumentów opisano regułami, które definiują postępowanie konsumenta na etapie t + 1, w zależności od zachowania jego otoczenia na etapie t, oraz wprowadzają mobilność konsumentów w systemie. Okazuje się, że we wszystkich modelach konsumenci mają tendencję do tworzenia większych grup, skupisk osób o tych samych preferencjach. Tworzenie takich skupisk jest bardziej wyraźne w przypadku większych otoczeń dla r = 1 i r = 2. Może stanowić to przesłankę do tworzenia strategii marketingowych skierowanych do większych grup, ale ze względu na przynależność do takiej grupy. Ponadto, duży wpływ rodziny i najbliższych znajomych oraz przyjaciół powoduje, że decyzje zakupowe i preferencje z czasem są podobne w całej takiej grupie. Coraz silniej należy odchodzić od stereotypów i podziałów na przykład na produkty typowo męskie i typowo kobiece, a strategie kierować do całej grupy. W znacznej części decyzji dotyczących przede wszystkim towarów trwałego użytkowania podejmo-

wane one są wspólnie [5], nawet jeśli każdy uczestnik procesu decyzyjnego dokonuje własnej oceny. Czas zatem na reklamę samochodów w kobiecych czasopismach i pralek w męskich.

Literatura

1. Burgiel A. (2005). Znaczenie naśladownictwa i wpływów społecznych w zachowaniach konsumentów. AE, Katowice.
2. Bearden W.O., Netemeyer R.G, Teel J.E. (1989). Measurement of Consumer Susceptibility to Interpersonal Infuence. Journal of Consumer Research, 15.
3. Kowalska-Styczeń A. (2007). Symulowanie złożonych procesów ekonomicznych za pomocą automatów komórkowych. Politechnika Śląska, Gliwice.
4. Kułakowski K. (2000) Automaty komórkowe. AGH, Kraków.
5. Lambkin M, Foxall G., van Raaij F., Heilbrunn (2001). Zachowanie konsumenta. Koncepcje i badania europejskie. Wydawnictwo Naukowe PWN, Warszawa.
6. Ławnicki J.S. (2005). Marketing sukcesu – partnering. Difin, Warszawa.
7. Schiffman L., Kanuk L. Consumer Behavior. Prentice Hall, Englewood Cliffs, 377-382.
8. Smyczek S., Sowa I. (2005). Konsument na rynku. Zachowania, modele, aplikacje. Difin, Warszawa.
9. Wolfram S. (2002). A New Kind of Science. Wolfram Media, Inc.

Piotr Pałka

Eugeniusz Toczyłowski

Instytut Automatyki i Informatyki Stosowanej w Warszawie

ZGODNOŚĆ WYBRANYCH METOD WYCENY TOWARÓW Z PREFERENCJAMI UCZESTNIKÓW RYNKU

Wstęp

Mechanizmy giełdowe i aukcyjne stosują różne reguły wypłat. Od najprostszej, polegającej na wyznaczeniu ceny rynkowej przecięcia krzywych jako punkt popytu i podaży, poprzez bardziej złożone, stosowane w aukcjach podwójnych, oparte na wyznaczeniu jednolitej ceny rozliczeniowej jako kombinacji wypukłej ostatnich przyjętych cen kupna i sprzedaży. Istnieje cała rodzina reguł wyceny opartych na drugich cenach lub podatku Grovesa – do takich należy aukcja Vickreya [11], ale także jej modyfikacje z dziedziny aukcji podwójnych [12]. W końcu bardziej złożone reguły wyceny, jak stosowana w aukcji Yoona [13] reguła opłat określająca rozchylone ceny rynkowe czy też oparte na analizie parametrycznej metody wyceny towarów [7].

Różne rodzaje wyceny służą do zapewnienia między innymi zgodności celów (preferencji) indywidualnych każdego z uczestników rynku z celami (preferencjami) globalnymi, jakimi są osiągnięcie sumarycznie największej nadwyżki rynkowej. Preferencje indywidualne uczestników rynku to dla każdego uczestnika chęć uzyskania największych zysków indywidualnych. Wyznaczanie cen rozliczeniowych stanowi więc złożony problem decyzyjny, w którym należy uwzględnić zarówno preferencje każdego z uczestników (którzy często działają strategicznie, ukrywając swoje prawdziwe preferencje), jak i preferencje globalne, definiowane przez wiele właściwości, których spełnienie jest równie ważne. Powyższe właściwości definiuje teoria mechanizmów. Niektóre z nich to zbilansowanie budżetu, indywidualna racjonalność dla każdego uczestnika rynku, odporność na spekulacje czy sprawiedliwość wyników rozliczeń. Preferencje globalne są narzucane przez ogół uczestników lub też przez pewnego odgórnego regulatora (przez przepisy, ustawy itd.). Często preferencje globalne nie są zgodne z preferencjami indywidualnymi, np. preferencją globalną może być maksymalizacja sumarycznych korzyści ekonomicznych, zaś indywidualną maksymalizacja korzyści własnych.

Przedmiotem pracy jest analiza reguł wyceny na tle prostego modelu handlu giełdowego, z punktu widzenia tego, w jaki sposób spełniają one preferencje uczestników oraz preferencje globalne. Niektóre właściwości są analizowane dla przypadku oligopolu (niekorzystnego ze względu na osiąganie pewnych własności, np. zgodności motywacji).

1. Reguły alokacji

Aby analizować niektóre właściwości mechanizmów, konieczne jest zdefiniowanie reguł alokacji towarów. Zakładamy, że w handlu będą uczestniczyć dwie strony: strona popytu i podaży. Niech w handlu uczestniczy pewna liczba sprzedawców (przez indeks $l \in S$ oznaczmy ich numerację) oraz kupujących (przez indeks $m \in B$ oznaczmy ich numerację). Składają oni swoje użyteczności w postaci pewnych sparametryzowanych ofert.

W handlu aukcyjnym podmioty zgłaszają jedynie ceny ofertowe jakie są oni w stanie zapłacić lub za jakie są oni w stanie zakupić dany towar. Innymi słowy, każdy kupujący składa ofertę kupna, o określonej przez oferenta maksymalnej cenie, jaką jest on w stanie zapłacić za dany towar (e_m). Podobnie, każdy sprzedający składa ofertę sprzedaży, o określonej przez niego minimalnej cenie, za jaką gotów jest on sprzedać dany towar (s_l). Zakłada się handel niepodzielnym towarem (dla ustalenia uwagi wolumen równy 1 jednostce).

W przypadku handlu giełdowego, każdy podmiot może oferować różny wolumen towaru do kupna lub sprzedaży. Także przyjęcie oferty może być częściowe. Oferent oferuje oprócz cen ofertowych (notacja jak wyżej), maksymalny wolumen oferowany na sprzedaż ($p_l^{\max}$) lub maksymalny wolumen towaru jaki podmiot chce zakupić ($d_m^{\max}$).

Zakładamy także, że towar sprzedawany przez każdego sprzedającego jest jednorodny, czyli kupujący nie dba o to, od którego sprzedawcy kupi towar. Reguła alokacji, czyli reguła określająca, które z ofert zostaną przyjęte, a które odrzucone (określają to zmienne p_l dla sprzedawców oraz d_m dla kupujących), polega na przyjęciu tych ofert sprzedaży, których cena ofertowa nie jest większa od cen ofertowych tych ofert kupna, które także zostają przyjęte.

Reguła alokacji przedstawia się następująco. Uporządkujmy oferty sprzedaży w porządku niemalejącym $s_1 \le s_2 \le \cdots \le s_{|S|}$, a oferty kupna w porządku nierosnącym $e_1 \ge e_2 \ge \cdots \ge e_{|B|}$. Jeśli zachodzi warunek $\left\langle s_1, s_{|S|} \right\rangle \cup \left\langle e_{|B|}, e_1 \right\rangle \notin \varnothing$, wówczas zachodzi także warunek $s_k \le e_j$, $s_{k+1} > e_{j+1}$. Oznacza on, iż pewna liczba ofert zostanie przyjęta, zaś pewna ilość ofert

zostanie odrzucona. W szczególności przyjęte zostaną oferty sprzedaży o numerach mniejszych bądź równych od *k* oraz oferty kupna o numerach mniejszych bądź równych od *j*. Warunek ten można zapisać także w inny sposób $\{e, j\} = \arg\max_{l \in S, m \in B} s_l \le e_m$.

2. Analiza metod wyceny

W niniejszym rozdziale analizujemy kilka metod wyceny, począwszy od najprostszych, wyznaczających jednolitą cenę rynkową, do złożonych, wyznaczających cenę na podstawie złożonych zależności.

2.1. Reguła ceny dualnej

Jedna z najprostszych metod wyceny polega na wyznaczaniu ceny rozliczeniowej (w dalszej części pracy będziemy ją oznaczać jako π) jako rzut punktu przecięcia krzywych popytu i podaży na oś cen. Z punktu widzenia modelu matematycznego, cena ta odpowiada wartości dualnej ograniczenia bilansowego. Wyznaczona cena jest ceną jednolitą dla wszystkich podmiotów uczestniczących w handlu.

2.2. Reguły κ pierwszych cen

Na rynkach aukcyjnych punkt przecięcia krzywych popytu i podaży może być niejednoznaczny – jest to odcinek łączący cenę ofertową ostatniej przyjętej (najdroższej przyjętej) oferty sprzedaży z ceną ofertową ostatniej przyjętej (najtańszej przyjętej) oferty kupna. W teorii aukcji podwójnych (które definiują jedynie regułę alokacji) istnieje pojęcie aukcji κ podwójnej, gdzie zdefiniowana jest także ogólna reguła wyceny [1; 12]. Zdefiniujmy parametr $s_k = \max_{l \in S}(s_l : p_l > 0)$ będący ceną ostatniej przyjętej (najdroższej wśród przyjętych) ofert sprzedaży, a także $e_j = \min_{m \in B}(e_m : d_m > 0)$ będący ceną ostatniej przyjętej (najtańszej wśród przyjętych) ofert kupna. Cena rozliczeniowa w teorii aukcji κ podwójnych stanowi kombinację wypukłą cen s_k i e_j (1), której konkretna wartość zależy od parametru κ: $0 \le \kappa \le 1$. W dalszej części pracy tę ogólną regułę wyceny będziemy nazywali regułą κ pierwszych cen.

$$\pi = \kappa \cdot e_j + (1 - \kappa) \cdot s_k, \quad 0 \le \kappa \le 1 \tag{1}$$

Widzimy więc, że cena rozliczeniowa może kształtować się w granicach wyznaczanych przez s_k oraz e_j, a jej wartość jest ustalana przez wartość parametru κ – sformułowana reguła wyceny jest ogólna. W literaturze istnieje wiele prac dotyczących tego tematu, w których rozważa się właściwości różnych mechanizmów aukcyjnych w zależności od parametru κ. Specyficzne wartości tego parametru określają specyficzne właściwości aukcji podwójnej. Dla $\kappa = 0$ mamy do czynienia z aukcją SBDA (Seller's Bid Double Auction) [2; 3]. Cena rozliczeniowa jest w tym przypadku równa cenie ofertowej ostatniej przyjętej oferty sprzedaży $\pi = s_k$. W literaturze istnieje także aukcja BBDA (Buyer's Bid Double Auction) [10]. Parametr κ w tym przypadku jest równy $\kappa = 1$, tak więc cena rozliczeniowa jest równa cenie ofertowej ostatniej przyjętej oferty kupna $\pi = e_j$. Nazwy tych aukcji wynikają z faktu, która strona (popytu czy podaży) wyznacza cenę rozliczeniową.

2.3. Reguły κ drugich cen

Specyficzną rodzinę metod wyceny są aukcje, w których cena rozliczeniowa wyznaczana jest nie za pomocą cen ofertowych ostatnich przyjętych ofert, lecz tzw. drugich cen. Ceny te zdefiniowane są jako ceny ofertowe pierwszych nieprzyjętych ofert. Uszczegóławiając, „druga cena" sprzedaży jest zdefiniowana jako cena ofertowa odpowiadająca nieprzyjętej ofercie sprzedaży o najniższej cenie ofertowej. Oznaczmy ją przez $s_{k+1} = \min_{l \in S}(s_l : p_l = 0)$. Podobnie, „druga cena" kupna będzie zdefiniowana jako cena ofertowa odpowiadająca nieprzyjętej ofercie kupna o najwyższej cenie ofertowej. Oznaczmy ją przez $e_{j+1} = \min_{m \in B}(e_m : d_m = 0)$. Uogólnijmy metodę wyznaczania ceny podobnie jak zostało to zdefiniowane poprzednio. W dalszej części pracy tę ogólną regułę wyceny będziemy nazywali regułą κ drugich cen.

$$\pi = \kappa \cdot e_{j+1} + (1-\kappa) \cdot s_{k+1}, \quad 0 \le \kappa \le 1 \tag{2}$$

Wśród metod wyceny charakteryzujących się powyższymi właściwościami można znaleźć między innymi modyfikację aukcji BBDA, nazwaną przez autorów [4] MBBDA (Modified Buyer's Bid Double Auction) lub MDA (Modified Double Auction) [9]. Jest to aukcja w której cena rozliczeniowa jest obliczana jako $\pi = e_{j+1}$, czyli jest równa cenie ofertowej odpowiadającej nieprzyjętej ofercie kupna o najwyższej cenie ofertowej.

Jednak metody te nie są tak interesujące jak zaproponowana przez McAfeego w pracy [6] metoda, polegająca na wycenie towarów zgodnie z zależnością (2) z parametrem $\kappa = \frac{1}{2}$. Aukcja ta została nazwana DSA (Dominant Strategy Auction). Aukcja ta ma wiele interesujących właściwości[1].

2.4. Reguła rozchylonych cen MVDA

Ostatnią z omawianych metod wyceny wywodzących się z teorii aukcji jest opisana przez Yoona [13] modyfikacja aukcji Vickreya [11] MVDA (Modified Vickrey Double Auction). Proponuje on wyznaczenie rozchylonych cen kupna i sprzedaży. Cena sprzedaży jest wyznaczana przez mniejszą z wartości pierwszej odrzuconej oferty sprzedaży i ostatniej przyjętej oferty kupna (3), zaś cena zakupu jest wyznaczana przez większą z wartości pierwszej odrzuconej oferty zakupu i ostatniej przyjętej oferty sprzedaży (4).

$$\pi^S = \min\{s_{k+1}, e_j\} \tag{3}$$

$$\pi^K = \max\{e_{j+1}, s_k\} \tag{4}$$

Metoda wyceny zaproponowana przez Yoona ma również bardzo interesujące właściwości. W szczególności, ceny wyznaczone za pomocą tej reguły mają tę własność, że poszczególni uczestnicy rynku nie są w stanie osiągnąć wyższej ceny rozliczeniowej za pomocą (indywidualnych) działań strategicznych. Mimo że przedstawione powyżej metody wyceny wywodzą się z aukcji podwójnych, mogą z powodzeniem być stosowane w handlu giełdowym.

2.5. Parametryczna reguła wyceny

Zaproponowana przez autorów tej pracy metoda wyceny opierająca się na analizie parametrycznej [8] (w skrócie parametryczna metoda wyceny) stanowi rozszerzenie metody zaproponowanej przez Yoona. Rozpatrywana jest w kontekście handlu giełdowego, w którym metoda Yoona traci własność zgodności motywacyjnej [7]. Parametryczna metoda wyceny polega na wyznaczeniu cen, których wartość zależy od wyników analizy parametrycznej rozwiązania zadania bilansowania rynku. Ceny wyznaczane mogą być w sposób indywidualny

[1] McAfee w swojej pracy [6] zmodyfikował regułę alokacji w taki sposób, aby jego mechanizm był indywidualnie racjonalny, za to nie posiadał on własności neutralności finansowej. W niniejszej pracy analizujemy zaproponowaną przez niego regułę wyceny, przyjmując klasyczny aukcyjny model alokacji. W tym przypadku neutralność finansowa operatora jest zachowana, natomiast nie spełniona jest własność indywidualnej racjonalności.

dla każdego uczestnika lub też, jak w metodzie Yoona, mogą stanowić rozchylone ceny jednolite dla popytu i podaży. Analiza parametryczna to dokonywana globalnie analiza wrażliwości rozwiązania dla każdego podmiotu rynkowego.

Analiza wrażliwości przeprowadzona dla modelu giełdowego, mówi nam o tym, o ile można zwiększyć (bądź zmniejszyć) ceny ofertowe poszczególnych podmiotów, aby nie zmieniła się optymalna alokacja towarów. Można to interpretować jako maksymalne zmiany na plus (dla ofert sprzedaży s_l^+) bądź na minus (dla ofert zakupu e_m^-) poszczególnych cen ofertowych, jakich może dokonać oferent, aby rozwiązanie się nie zmieniło. Podmioty, które zmienią swoje ceny o więcej niż wynoszą wyniki analizy mogą liczyć jedynie na zmniejszenie swojego udziału w rynku.

Dla każdej oferty sprzedaży suma ceny ofertowej s_l i współczynnika s_l^+ oznaczającego maksymalną wartość o jaką możemy podnieść cenę daje nam dla każdego podmiotu wartość ceny ofertowej, po przekroczeniu której zmieni się alokacja, a więc sprzedawca straci. Podobnie dla ofert zakupu, różnica ceny ofertowej e_m i współczynnika e_m^- daje nam cenę ofertową, po przekroczeniu której zmniejsza się udział nabywcy. Istota analizy parametrycznej polega na iteracyjnym przeprowadzaniu analizy wrażliwości (dla każdego podmiotu z osobna). Po każdym kroku *i* przeprowadzamy ponownie analizę wrażliwości zwiększając (dla sprzedawcy) cenę ofertową o wartość $s_l^{+,(i)} + \varepsilon,\ \varepsilon > 0$, zaś dla nabywcy zmniejszając cenę ofertową o wartość $e_m^{-,(i)} - \varepsilon,\ \varepsilon > 0$. Otrzymujemy nowe wyniki rozliczenia oraz nowe wartości współczynników $s_l^{+,(i+1)}$ oraz $e_m^{-,(i+1)}$. Przeprowadzamy analizę tak długo, jak długo podmiot jest przyjmowany. Ilość przeprowadzonych kroków dla każdego podmiotu oznaczamy przez parametr i^*.

$$\pi_l^S = s_l + \sum_{i=1}^{i^*} s_l^{+,(i)} \tag{5}$$

$$\pi_m^K = e_m - \sum_{i=1}^{i^*} e_m^{-,(i)} \tag{6}$$

Na podstawie analizy parametrycznej otrzymujemy indywidualne ceny rozliczeniowe dla każdego podmiotu. Jest to maksymalna cena (dla sprzedawców, patrz (5)) jaką może zgłosić dany podmiot aby nie zostać odrzuconym, lub jako minimalna cena (dla nabywców, patrz (6)) jaką może zgłosić dany nabywca aby nie został odrzucony.

W ten sposób otrzymujemy ceny indywidualne dla każdego podmiotu. Aby uzyskać ceny rozchylone dla strony popytu i podaży, należy wyznaczyć maksymalną cenę sprzedaży (patrz (7)) oraz minimalną cenę zakupu (patrz (8)) z cen indywidualnych.

$$\pi^S = \max_{l \in S} \pi_l^S \tag{7}$$

$$\pi^K = \min_{m \in B} \pi_m^K \tag{8}$$

3. Spełnienie preferencji indywidualnych

3.1. Ceny jednolite a indywidualne

Do metod wycen, które wyznaczają jednolitą cenę rozliczeniową należą: metoda opierająca się na wartości dualnej ograniczenia bilansowego, metody oparte na parametrze κ – zarówno z pierwszymi cenami, jak i z drugimi. Również aukcja McAfeego charakteryzuje się wyznaczeniem jednolitej ceny rozliczeniowej. Rozchylone ceny rozliczeniowe wyznaczają metody wyceny Yoona oraz parametryczna metoda wyceny. W końcu parametryczna metoda wyceny wyznacza także ceny indywidualne dla każdego z uczestników rynku.

3.2. Własność indywidualnej racjonalności

Preferencją indywidualną każdego podmiotu jest także to, czy mechanizm rynkowy wyznaczy takie ceny i przydziały dóbr, aby każdy podmiot osiągał nieujemną wartość swojej użyteczności (należy zauważyć, że podmioty muszą działać racjonalnie, czyli składać odpowiednie oferty). Własność ta jest nazywana w teorii mechanizmów własnością indywidualnej racjonalności. Rozważmy, czy własność ta zachodzi dla wymienionych reguł wycen.

Reguła wyceny oparta na cenie dualnej ograniczenia bilansowego zakłada wyznaczanie ceny rozliczeniowej w punkcie s_k lub e_j. Jako że $s_k \le e_j$, tak więc $s_k \le \pi \le e_j$. Załóżmy, że użyteczność każdego sprzedawcy będzie wynosiła (9), zaś dla każdego nabywcy będzie wynosiła (10).

$$Q_l(\boldsymbol{\theta}) = Q_l\left(s_l^0, p_l, \pi\right) = p_l \cdot \left(\pi - s_l^0\right) \tag{9}$$

$$Q_m(\boldsymbol{\theta}) = Q_m\left(e_m^0, d_m, \pi\right) = d_m \cdot \left(e_m^0 - \pi\right) \tag{10}$$

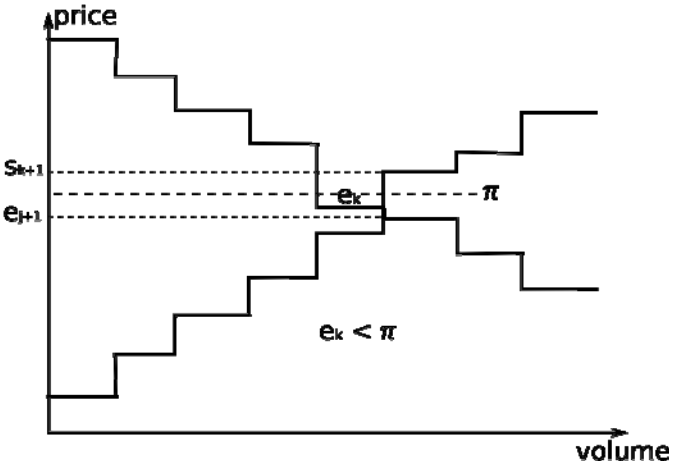
Tak więc, jeśli oferta zostanie przyjęta (czyli $l \le k$ oraz $p_l > 0$), użyteczność danego podmiotu będzie zależała od różnicy $\pi - s_l^0$. Dodatkowo wiemy, że cena s_k jest najwyższą przyjętą ceną ofertową, tak więc dla dowolnego sprzedawcy l zachodzi $s_l \le s_k$, czyli $s_l \le \pi$. Z założenia, że podmioty działają racjonalnie możemy stwierdzić, że $s_l^0 \le s_l$ – czyli sprzedawcy nie będą oferowali ceny niższej niż wynosi cena wynikająca z ich użyteczności. Ostatecznie widzimy, że $s_l^0 \le \pi$, co stanowi o tym, że użyteczność każdego sprzedawcy przy tej regule wyceny jest zawsze nieujemna. Podobne rozumowanie można przeprowadzić dla nabywców z podobnym wynikiem.

Podobne rozumowanie możemy przeprowadzić dla każdej reguły wyceny, dla której zachodzi własność $s_k \le \pi \le e_j$. Do takich reguł wyceny należy ogólna reguła wyceny κ pierwszych cen. Tak więc widzimy, iż te reguły wyceny spełniają własność indywidualnej racjonalności.

Rozważmy teraz, czy własność indywidualnej racjonalności zachodzi także dla innych reguł wyceny. Rozważmy regułę κ drugich cen. Dla ustalenia uwagi weźmy pod uwagę regułę McAfeego. Dla klasycznej reguły alokacji reguła wyceny McAfeego (ale także inne reguły κ drugich cen) nie spełniają własności indywidualnej racjonalności. Sytuacja taka może zajść w dwóch przypadkach: kiedy cena rozliczeniowa jest wyższa od cen ofertowych przyjętych ofert kupna $\exists_{m \in B, d_m > 0} e_m < \pi$ (rys. 1), kiedy cena rozliczeniowa jest niższa od cen ofertowych przyjętych ofert sprzedaży $\exists_{l \in S, p_l > 0} s_l > \pi$ (rys. 2).

Dla różnych parametrów κ jedna z sytuacji przedstawionych na rys. 1 i 2 nie będzie zachodziła, lecz będzie częściej zachodziła druga sytuacja.

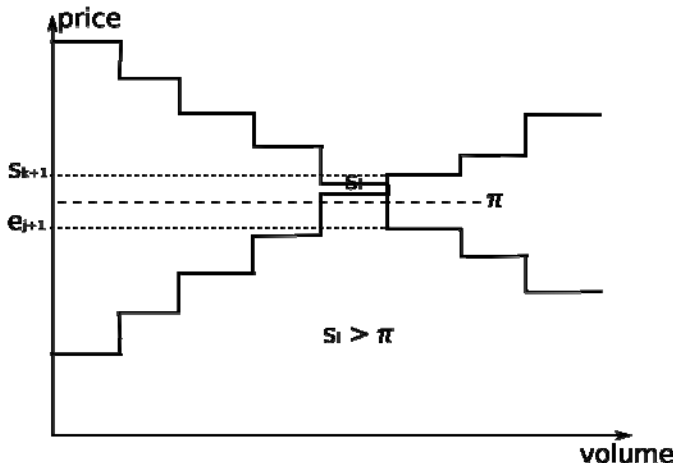
W pracy [13] Yoon zaproponował nieco inną regułę alokacji, która spełniała własność indywidualnej racjonalności. Zaproponował on metodę wyznaczania rozchylonych cen sprzedaży i kupna. Cena sprzedaży jest równa $\pi^S = \min\{s_{k+1}, e_j\}$, tak więc spełnia ona warunek $s_k \le \pi^S$ gdyż: $s_k \le s_{k+1}$ oraz $s_k \le e_j$ – cena rozliczeniowa sprzedaży wyznaczana jest na podstawie mniejszej z tych wartości. Podobnie dla ceny rozliczeniowej zakupu, która jest równa $\pi^K = \max\{e_{j+1}, s_k\}$, warunek $\pi^K \le e_j$ jest spełniony, gdyż $e_{j+1} \le e_j$ oraz $s_k \le e_j$ – cena rozliczeniowa zakupu jest wyznaczana jako większa z tych wartości. Tak więc reguła wyceny Yoona jest indywidualnie racjonalna.



Rys. 1. Rozliczenie metodą McAfeego, nie jest spełniona właściwość indywidualnej racjonalności dla nabywcy *j*

Parametryczną regułę wyceny można traktować jako uogólnienie metody Yoona na rynki giełdowe. Wyznacza ona ceny indywidualne dla każdego podmiotu, na podstawie analizy parametrycznej matematycznego zadania bilansowania rynku. Ceny indywidualne mogą zostać sprowadzone do cen rozchylonych (jak w metodzie Yoona) bądź też do ceny jednolitej. W niniejszej pracy zostaną rozważone dwie pierwsze możliwości.

Indywidualna cena rozliczeniowa dla sprzedawcy *l*-tego wynosi $\pi_l^S = s_l + \sum_{i=1}^{i^*} s_l^{+,(i)}$, gdzie i^* oznacza ostatni krok analizy parametrycznej. Tak więc spełnienie własności indywidualnej racjonalności jest oczywiste, gdyż $\forall_i s_l^{+,(i)} \geq 0$. Tak więc zachodzi również $\pi_l^S \geq s_l$. Podobnie dla każdego kupującego m indywidualna cena rozliczeniowa wynosi $\pi_m^K = e_m - \sum_{i=1}^{i^*} e_m^{-,(i)}$ oraz $\forall_i e_m^{-,(i)} \geq 0,$ tak więc spełniony jest również wymóg $\pi_m^K \leq e_m$.



Rys. 2. Rozliczenie metodą McAfee'go, własność indywidualnej racjonalności nie jest spełniona dla sprzedawcy *l*

Należy pamiętać, że rozważamy tutaj tylko sytuację podmiotów przyjętych do rozliczenia, gdyż dla podmiotów odrzuconych wartość użyteczności (zależności (9) oraz (10)) jest równa zeru.

4. Preferencje globalne

Dla poszczególnych metod wyceny w różny sposób spełnione są preferencje globalne, takie jak zbilansowanie budżetu, maksymalizacja nadwyżki ekonomicznej, sprawiedliwość, niewrażliwość na spekulacje.

4.1. Niewrażliwość na spekulacje

Własność niewrażliwości na spekulacje można utożsamiać z własnością zgodności motywacji. W niniejszej pracy nie zostanie przeprowadzona pełna analiza zgodności motywacji każdej z reguły wyceny. Zostaną podane proste przykłady, które stwierdzą, czy reguła jest wrażliwa na spekulacje bądź też podane zostaną odwołania do badań przeprowadzonych w innych pracach. Analizowana będzie niewrażliwość na spekulacje w sytuacji oligopolu.

Metody wyceny κ-pierwszych cen, jak i metoda ceny dualnej są wrażliwe na spekulacje, w szczególności w sytuacji oligopolu. Sprzedawca (nabywca) marginalny może podnosić (zaniżać) swoją cenę ofertową zwiększając (zmniejszając) cenę rozliczeniową, a tym samym zwiększając swój przychód. Tak więc

proste metody wyceny są wrażliwe na spekulacje. Bardziej zaawansowane reguły wyceny – reguły κ drugich cen, są bardziej odporne na spekulacje, gdyż cena rozliczeniowa jest wyznaczana na podstawie cen zaoferowanych przez odrzucone podmioty. Tak więc podmioty przyjęte nie mają sposobności wpływania na ceny rozliczeniowe (tylko w przypadku aukcji; w modelu giełdowym podmioty o dużej sile rynkowej mogą wpływać na cenę rozliczeniową, choć w mniejszym stopniu). McAfee w swojej pracy [6] przedstawił mechanizm jako zgodny motywacyjnie w strategiach dominujących. Również reguła wyceny Yoona jest, według pracy [13], zgodna motywacyjnie w strategiach dominujących.

Parametryczna reguła wyceny spełnia własność zgodności motywacji w strategiach dominujących (dowód w pracy [7]). Jej cechy wspólne z regułą Yoona wpływają na jej dobre właściwości w tej materii. Ponadto, reguła ta projektowana specjalnie dla rynku giełdowego posiada lepsze własności osłabiające siłę rynkową niż reguła Yoona.

4.2. Neutralność finansowa operatora

Własność neutralności finansowej operatora bądź zbilansowania budżetu zachodzi, jeśli suma wydatków pochodząca z wpłat za poszczególne towary jest równa sumie wypłat dla poszczególnych dostawców towarów. Innymi słowy, budżet jest zbilansowany, jeśli operator rynku nie czerpie nieuzasadnionych zysków z wymiany towarowej ani też nie musi do niej dopłacać. Jest to ważna własność, gdyż operator czerpiący nieuzasadnione zyski z handlu, zmniejsza zyski poszczególnych uczestników, tym samym zniechęcając ich do handlu. Jednak operator, który musi dopłacać do rozliczenia nie jest w stanie prowadzić działalności (chyba że otrzymuje subwencje z budżetu państwa).

Istnienie jednolitej ceny rozliczeniowej powoduje, że własność ta zachodzi. Tak więc wszystkie reguły wyceny wyznaczające jednolitą cenę rozliczeniową posiadają również własność zbilansowania budżetu.

Metody opisywane w naszych rozważaniach, wyznaczające rozchylone ceny kupna i sprzedaży, charakteryzują się niezbilansowaniem budżetu i to takim, w których operator musi dopłacać do rozliczenia. Reguły te to reguła Yoona i metoda analizy parametrycznej. Jednak można złagodzić niezbilansowanie budżetu przy pewnych założeniach. W pracach [13; 14] Yoon proponuje mechanizm opłat za uczestnictwo w aukcji, który to mechanizm spowoduje zrównoważenie budżetu w sensie wartości oczekiwanej. Mechanizm taki wymaga jednak znajomości rozkładów użyteczności podmiotów oraz wprowadza bariery wejścia na rynek.

W przypadku giełd, w których wielokrotnie uczestniczą te same podmioty (np. giełda energii) możliwe jest łagodzenie niezbilansowania budżetu poprzez agregowanie pewnych opłat za uczestnictwo w danej giełdzie, płacone np. co miesiąc przez każdego uczestnika. Metoda ta różni się od podejścia Yoona tym, że w metodzie Yoona każdy uczestnik płaci tyle samo za możliwość uczestniczenia w handlu (mimo że mógł nic nie kupić/sprzedać), zaś w drugiej metodzie wielkość opłaty za uczestnictwo w handlu może zależeć od wolumenu obrotu danego podmiotu. Tak więc reguła Yoona oraz reguła parametryczna z cenami rozchylonymi nie spełniają własności neutralności finansowej operatora, lecz istnieją metody łagodzące tę wadę.

Metoda parametryczna wyznaczająca indywidualne ceny rozliczeniowe także jest niezbilansowana budżetowo (choć w mniejszym zakresie niż metoda parametryczna wyznaczająca ceny rozchylone). Możemy jednak zastosować te same metody łagodzenia niezbilansowania jak opisane powyżej.

4.3. Maksymalizacja nadwyżki ekonomicznej

Własność ta zależy od reguły alokacji. Zarówno klasyczna reguła alokacji dla aukcji podwójnych, jak i reguła alokacji dla rynku giełdowego maksymalizują nadwyżkę ekonomiczną – są regułami efektywnymi. Wynika to ze sformułowania zadania. Maksymalizacja nadwyżki ekonomicznej wraz ze sprawiedliwym jej podziałem są istotną cechą mechanizmów rynkowych.

4.4. Sprawiedliwość

Przez sprawiedliwość będziemy rozumieli, czy zachodzą następujące właściwości. Po pierwsze, jeśli dwa podmioty złożą takie same oferty, to zostaną potraktowane jednakowo. Po drugie, czy za jednostkę tego samego towaru sprzedawcy będą otrzymywać tę samą kwotę oraz czy za jednostkę towaru nabywcy będą płacić tę samą kwotę.

Jeśli chodzi o pierwsze kryterium, to jego spełnienie zależy od reguły alokacji towarów. Dla przedstawionych tutaj modeli alokacji kryterium to nie jest spełnione. Jednakże można wyobrazić sobie poprawioną regułę alokacji, która spełnia to kryterium. Jeśli chodzi o drugie kryterium, jest ono spełnione dla reguł, które wyznaczają cenę jednolitą dla wszystkich podmiotów (czyli reguła oparta na cenie dualnej, reguły κ ceny pojedynczej i podwójnej). Dla reguł wyznaczających ceny rozchylone (czyli reguła Yoona i jedna z odmian reguły parametrycznej) to kryterium także jest spełnione – każdy sprzedawca otrzymuje za jednostkę towaru tę samą cenę, także każdy nabywca otrzymuje za jednostkę towaru tę samą cenę. W końcu reguły, które wyznaczają indywidualne ceny rozliczeniowe dla każdego podmiotu tego kryterium nie spełniają.

Podsumowanie

Przeprowadzona została analiza różnych metod wyceny towarów na rynku aukcyjnym i giełdowym. Analiza ta była przeprowadzana pod kątem spełnienia różnych preferencji, zarówno indywidualnych, jak i globalnych, pożądanych przez ogół uczestników rynku bądź też przez pewne organy nadzorcze (np. regulatora rynku). Proste reguły wyceny (wyznaczające jednolitą cenę ofertową reguły κ pierwszych cen czy reguła ceny dualnej) spełniają preferencje indywidualnej racjonalności, sprawiedliwości, a także posiadają zbilansowany budżet. Jednakowoż, reguły te są wrażliwe na spekulacje podmiotów, które posiadają siłę rynkową. Badacze opracowali reguły wyceny poprawiające tą własność. Reguła wyceny McAfeego jest zgodna motywacyjnie dla rynków aukcyjnych, jednakże traci właściwość indywidualnej racjonalności. Reguła Yoona, także zgodna motywacyjnie dla rynków aukcyjnych, mająca własność indywidualnej racjonalności, nie posiada własności neutralności finansowej operatora, jednakże przedstawione zostały mechanizmy łagodzące tę wadę. Pomimo dobrych własności metod wyceny McAfeego oraz Yoona dla rynków aukcyjnych, po przeniesieniu tych reguł wyceny na rynki giełdowe tracą one swoje dobre właściwości. Opracowana przez autorów reguła parametryczna dla handlu giełdowego powoduje zachowanie dobrych własności uzyskiwanych dla metody Yoona na rynku aukcyjnym. Opracowana przez autorów reguła wyceny będzie rozwijana, badane będą dokładniej jej własności.

Literatura

1. Chatterjee K. and Samuelson W. (1983). Bargaining under Incomplete Information. Operations Research, 31, 835-851.
2. Jain R. (2004). Efficient Market Mechanisms and Simulation-based Learing for Multi-Agent Systems. PhD thesis, University of California, Berkeley.
3. Jain R. and Varaiya P. (2004). An efficient Incentive-Compatible Combinatorial Market Mechanism. Allerton Conf.
4. Kagel J.H. and Vogt W. (1993). Buyer's Bid Double Auctions: Preliminary Experimental Results. 10, 285-305. Perseus Publishing, Cambridge.
5. Krishna V. (2002). Auction Theory. Academic Press.
6. McAfee R.P. (1992). A Dominant Strategy Double Auction. Journal of Economic Theory, 56, 434-450.
7. Pałka P. (2009). Analiza zgodności motywacji mechanizmów wieloagentowej platformy wymiany towarowej. Politechnika Warszawska, Warszawa.

8. Pałka P., Toczyłowski E. (2009). Mechanizmy wyceny dóbr za pomocą uogólnionej metody Yoona i metody analizy parametrycznej. Automatyka, 13(2), 539-550.
9. Satterthwaite M.A. and Williams S.R. (1989). The Rate of Convergence to Efficiency in the Buyer's bid Double Auction as the Market Becomes Large. The Review of Economic Studies, 56(4), 477-498.
10. Satterthwaite M.A. and Williams S.R. (1993). The Bayesian Theory of the k-Double Auctions. 4, 99-123. Perseus Publishing, Cambridge.
11. Vickrey W. (1961). Counterspeculation, Auctions, and Competitive Sealed Tenders. Journal of Finance, 16(1), 8-37.
12. Wilson R. (1985). Incentive Efficiency of Double Auctions. Econometrica, (53), 1101-1115.
13. Yoon K. (2001). The Modified Vickrey Double Auction. Journal of Economic Theory, 101, 572-584.
14. Yoon K. (2008). The Participatory Vickrey-Clarke-Groves Mechanism. Journal of Mathematical Economics, 44, 324-336.

Zbigniew Świtalski
Uniwersytet Zielonogórski

O PEWNYM DYSKRETNYM MODELU RYNKU

Wstęp

Praca Gale'a i Shapleya z 1962 roku opisująca model rekrutacji kandydatów do szkół [2] zwróciła uwagę wielu ekonomistów na tzw. rynki z dwustronnymi preferencjami [3]. Na rynku takim mamy do czynienia z dwoma rodzajami uczestników: kupującymi i sprzedającymi, przy czym zakłada się, że każdy kupujący ma preferencje w zbiorze sprzedających, a każdy sprzedający ma preferencje w zbiorze kupujących. Przykładem takiego rynku może być „rynek szkolny", na którym kandydaci konkurują o miejsca w szkołach, i który może być opisany za pomocą modelu Gale'a-Shapleya. Na rynku tym rolę „kupujących" odgrywają kandydaci, a rolę „sprzedających" – szkoły.

Za pomocą analogicznego modelu można też opisywać bardziej „konwencjonalne" rynki, na których sprzedaje się tradycyjne towary, np. mieszkania, samochody lub inne dobra niepodzielne. Jeśli założymy, że każdy sprzedający dysponuje tylko jednym rodzajem towarów, to preferencje kupujących w zbiorze sprzedających mogą być określone w oczywisty sposób (sprzedający X jest lepszy od sprzedającego Y, jeśli towar oferowany przez X jest, dla danego kupującego, lepszy od towaru oferowanego przez Y). Z kolei preferencje sprzedających w zbiorze kupujących mogą być określone za pomocą ofert cenowych składanych przez kupujących (kupujący A jest dla danego sprzedającego lepszy od kupującego B, jeśli A oferuje za posiadany przez tego sprzedającego towar wyższą cenę niż B).

Podstawowym pojęciem używanym w teorii Gale'a-Shapleya jest pojęcie skojarzenia stabilnego. Skojarzenie stabilne (definicja 3) może być traktowane jako pewien zbiór transakcji gwarantujący, że żadna para (kupujący, sprzedający) nie ma motywacji do zmiany istniejącego układu. Jest to więc pewien stan równowagi układu złożonego z kupujących i sprzedających.

Jednak, jeśli na takim rynku zostaną ustalone ceny towarów (na rynku szkolnym rolę cen mogłyby odgrywać minimalne liczby punktów gwarantujące dostanie się do danej szkoły), to moglibyśmy zdefiniować klasyczne pojęcie równowagi rynkowej (konkurencyjnej).

W opracowaniu definiujemy właśnie tego rodzaju równowagę rynkową i badamy jej związki z pojęciem skojarzenia stabilnego.

Warto zauważyć, że w literaturze naukowej wiele uwagi poświęca się ostatnio badaniom związków między skojarzeniami stabilnymi a równowagami w modelach wzorowanych na modelu Shapleya-Shubika (jest to pewien rodzaj modelu kojarzenia różny od modelu Gale'a-Shapleya, [1; 4]. Nie ma jednak opracowań, w których tego rodzaju związki byłyby analizowane dla modelu Gale'a-Shapleya.

1. Model rynku z dwustronnymi preferencjami

Załóżmy, że $K = \{K_1,K_2,...,K_n\}$ jest skończonym zbiorem kupujących, a $S = \{S_1,S_2,...,S_m\}$ – skończonym zbiorem sprzedających. Sprzedający S_j posiada q_j jednostek pewnego niepodzielnego towaru. Każdy kupujący chce kupić dokładnie jedną jednostkę któregoś z towarów.

Przyjmujemy, że każdy kupujący K_i ma ustalone preferencje w zbiorze S wyznaczone za pomocą ostrego liniowego porządku $>_i$ (tzn. K_i jest w stanie ustawić sprzedających ze zbioru S w kolejności od najlepszego do najgorszego i żadni dwaj sprzedający nie są równoważni). Każdy sprzedający S_j ma z kolei preferencje w zbiorze K wyznaczone za pomocą ostrego liniowego porządku $>_j$.

Weźmy jako przykład „rynek" szkolny, który odpowiada modelowi Gale'a-Shapleya [2]. Na rynku tym kupującymi ze zbioru K są kandydaci do szkół ze zbioru S (czyli szkoły są sprzedającymi). Liczba q_j oznacza limit przyjęć dla szkoły S_j (tzn. maksymalną liczbę kandydatów, którą szkoła S_j gotowa jest przyjąć). Określone są też preferencje kandydatów w zbiorze S i preferencje szkół w zbiorze K (wyznaczone za pomocą odpowiedniej punktacji; zakładamy dla uproszczenia, że dowolni dwaj kandydaci są rozróżnialni przez daną szkołę, tzn. że nie mogą mieć tej samej liczby punktów).

Jako drugi przykład rozważmy następujący model rynku nieruchomości. Zbiór S jest zbiorem deweloperów. Deweloper S_j oferuje q_j identycznych mieszkań. Każdy kupujący chciałby kupić dokładnie jedno mieszkanie. Kupujący K_i ma preferencje w zbiorze S (czyli preferencje w zbiorze oferowanych mieszkań, bo wszystkie mieszkania oferowane przez S_j są traktowane jako identyczne). Zakładamy, że każdy kupujący K_i gotowy jest zapłacić co najwyżej cenę a_{ij} za mieszkanie oferowane przez S_j oraz, że $a_{ij} \neq a_{lj}$, jeśli $K_i \neq K_l$. Relację $>_j$ określamy następująco

$$K_i >_j K_l \Leftrightarrow a_{ij} > a_{lj}$$

Zdefiniujemy teraz skojarzenia stabilne dla podanego modelu rynku. Niech

$$E(K,S) = \{\{K_i,S_j\}:\ K_i \in K,\ S_j \in S\ \}$$

oznacza zbiór par nieuporządkowanych złożonych z elementów zbiorów K i S (inaczej jest to zbiór krawędzi w grafie dwudzielnym pełnym łączącym zbiory K i S). Dla dowolnego zbioru krawędzi $\mu \subset E(K,S)$, niech $\mu(K_i)$ oraz $\mu(S_j)$ oznaczają zbiory wierzchołków sąsiadujących z wierzchołkiem K_i oraz S_j odpowiednio, tzn.

$$\mu(K_i) = \{S_j: \{K_i, S_j\} \in \mu\}$$

$$\mu(S_j) = \{K_i: \{K_i, S_j\} \in \mu\}$$

Podzbiór $\mu \subset E(K,S)$ możemy interpretować jako zbiór transakcji dokonanych na rynku. Wówczas $\mu(K_i)$ jest zbiorem sprzedających, z którymi K_i zawarł transakcję, a $\mu(S_j)$ – zbiorem kupujących, którzy zawarli transakcję z S_j. Skojarzeniem będziemy nazywali zbiór transakcji $\mu \subset E(K,S)$, który spełnia określone w modelu ograniczenia (każdy kupujący kupuje tylko jedno dobro, a każdy sprzedający sprzedaje nie więcej niż q_j dóbr).

Definicja 1. Skojarzenie jest to zbiór $\mu \subset E(K,S)$ taki, że

$$\forall i,j \qquad |\mu(K_i)| \le 1, \qquad |\mu(S_j)| \le q_j$$

(symbol $|X|$ oznacza liczbę elementów zbioru X).

Definicja 2. Para $\{K_i,S_j\} \in E(K,S)$ nazywa się parą blokującą dla skojarzenia μ, jeśli

1) $S_j >_i \mu(K_i)$

2) $|\mu(S_j)| < q_j$ lub $(\exists K_l \in \mu(S_j))$: $K_i >_j K_l$)

(zauważmy, że zgodnie z definicją skojarzenia zbiór $\mu(K_i)$ jest co najwyżej jednoelementowy, w związku z tym nierówność 1) interpretujemy następująco: jeśli $|\mu(K_i)| = 1$, to S_j jest lepszy (dla K_i) od jedynego elementu zbioru $\mu(K_i)$, a jeśli $\mu(K_i) = \varnothing$, to zakładamy, że nierówność 1) jest zawsze prawdziwa).

Definicja 3. Skojarzenie μ nazywa się skojarzeniem stabilnym, jeśli nie istnieją pary blokujące dla μ.

2. Równowaga na rynku z dwustronnymi preferencjami

W typowych modelach równowagi rynkowej równowaga wyznaczana jest przez układ cen. Zmiany cen powodują dostosowanie popytu do podaży, a rynek zbliża się w ten sposób do stanu równowagi. W modelu kojarzenia kandydatów ze szkołami nie ma explicite określonych cen. Jeśli szkoły oceniają kandydatów za pomocą punktacji z testów, to rolę „ceny" jaką „płaci" kandydat mogłaby odgrywać minimalna liczba punktów potrzebna do dostania się do danej szkoły (próg punktowy albo inaczej liczba punktów, którą otrzymał najgorszy z przyjętych kandydatów). Kandydaci przyjęci to właśnie ci, którzy przekroczyli ten próg. Dana szkoła mogłaby, zamiast ustalania rankingu kandydatów, wyznaczyć (przed rozpoczęciem rekrutacji) próg punktowy i przyjmować wszystkich, którzy przekroczyli ten próg. Równowaga byłaby osiągana, gdyby dla każdej szkoły liczba przyjętych w ten sposób kandydatów była równa ustalonemu na początku limitowi przyjęć, a zmiana progów (w trakcie kolejnych rekrutacji) byłaby sposobem zbliżania się do równowagi.

Szkoła, stosując taką metodę przyjmowania kandydatów, podejmowałaby ryzyko zgłoszenia się zbyt wielu chętnych i przekroczenia limitu (przyjęcie zbyt wielu chętnych mogłoby okazać się niemożliwe z powodu ograniczeń fizycznych). Warto zauważyć, że ryzyko takie podejmuje każdy sprzedający, który określa cenę swojego towaru i z góry nie wie, czy popyt nie przekroczy tej ilości towaru, którą posiada. Szkoły (na ogół) nie stosują podanego sposobu przyjmowania kandydatów, a możliwość taką rozważamy tutaj czysto hipotetycznie po to, aby móc objaśnić pojęcie równowagi i udowodnić, ze każde skojarzenie stabilne jest pewnym rodzajem równowagi rynkowej.

Możemy przyjąć jeszcze ogólniejsze podejście. Określenie progu punktowego jest wyznaczeniem pewnego warunku, który muszą spełnić kandydaci, aby dostać się do danej szkoły. Na rynku, na którym funkcjonują ceny, wyznaczenie ceny przez sprzedającego jest określeniem pewnego warunku, który musi spełnić kupujący, aby móc otrzymać dane dobro (tzn. musi on być gotowy zapłacić tą cenę za to dobro). Istnieje wiele rynków, na których zapłacenie danej ceny nie jest jedynym warunkiem, który gwarantuje otrzymanie danego dobra. W szczególności może się zdarzyć, że szkoły będą wyznaczały inne niż progi punktowe warunki potrzebne, aby dostać się do danej szkoły.

Sformalizujemy teraz pojęcie równowagi wyznaczanej za pomocą układu warunków, które muszą spełnić kupujący, aby otrzymać dobra, którymi dysponują sprzedający. Załóżmy, że sprzedający S_j określa pewne warunki W_j gwarantujące kupującemu otrzymanie dobra posiadanego przez S_j. Formalnie, W_j utożsamiamy ze zbiorem kupujących, którzy spełniają warunki W_j, tzn.

$$W_j = \{K_i \in K: K_i \text{ spełnia } W_j \}$$

Dopuszczamy możliwość $W_j = \emptyset$, tzn. warunek może być tak silny, że żaden kupujący go nie spełnia. Zdefiniujmy dla danego kupującego K_i zbiór sprzedających dopuszczalnych przy warunkach $W = (W_1, W_2, ..., W_m)$ jako

$$D_W(K_i) = \{S_j \in S : K_i \in W_j\}$$

Jest to więc zbiór sprzedających, których warunki dany kupujący spełnił. Jeśli $D_W(K_i) \neq \emptyset$, to definiujemy sprzedającego najlepszego dla K_i jako

$$M_W(K_i) = \max \{S_j : S_j \in D_W(K_i)\}$$

$M_W(K_i)$ jest maksymalnym elementem w zbiorze $D_W(K_i)$ ze względu na liniowy porządek $>_i$. Zbiór $D_W(K_i)$ odpowiada zbiorowi wiązek dopuszczalnych, a element $M_W(K_i)$ odpowiada wiązce optymalnej w teorii konsumenta (przy założeniu, że dane są ceny dóbr i dochody konsumenta).

Zdefiniujemy teraz funkcję, którą można nazwać „funkcją popytu". Niech dany będzie sprzedający S_j. Definiujemy zbiór $\varphi_W(S_j)$ jako

$$\varphi_W(S_j) = \{K_i \in K : D_W(K_i) \neq \emptyset \wedge M_W(K_i) = S_j\}$$

Zbiór $\varphi_W(S_j)$ jest więc zbiorem kupujących, dla których sprzedający S_j jest najlepszy wśród wszystkich sprzedających, których warunki dany kupujący spełnił. O kupujących ze zbioru $\varphi_W(S_j)$ będziemy mówili, że tworzą popyt na dobro oferowane przez S_j. Wartością funkcji popytu będzie liczba $|\varphi_W(S_j)|$, czyli liczba kupujących K_i, którzy tworzą popyt na dobro oferowane przez S_j. Jeśli $D_W(K_i) = \emptyset$, to $|\varphi_W(S_j)| = 0$. Zauważmy, że liczba $|\varphi_W(S_j)|$ zależy od wszystkich warunków $W_1, W_2, ..., W_m$ (wynika to z definicji zbioru $D_W(K_i)$ i elementu $M_W(K_i)$).

Warunki $W_1, W_2, ..., W_m$ wyznaczają więc popyt na dobro posiadane przez sprzedającego S_j, czyli liczbę $|\varphi_W(S_j)|$. Podażą sprzedającego S_j jest q_j. Możemy więc zdefiniować równowagę następująco:

Definicja 4. Ciąg $(W_1, W_2, ..., W_m)$ wyznacza równowagę w modelu rynku z dwustronnymi preferencjami, jeśli

$$\forall j \quad |\varphi_W(S_j)| = q_j$$

Żeby istniała równowaga, musi być odpowiednio dużo kupujących, którzy będą tworzyli popyt na dobra oferowane przez sprzedających. Łatwo zauważyć, że warunkiem koniecznym i wystarczającym istnienia równowagi jest spełnienie warunku

$$n \geq \sum_j q_j$$

3. Równowaga a stabilność

Czy jest jakiś związek między pojęciem równowagi rozumianym jako ciąg warunków (W_1, W_2, ..., W_m), dla którego popyt zrównuje się z podażą, a pojęciem skojarzenia stabilnego rozumianym jako zbiór par $\{K_i, S_j\}$, który nie zawiera pary blokującej (definicja 3)? Okazuje się, że jest pewna odpowiedniość między tymi pojęciami, jeśli będziemy brać pod uwagę tylko tzw. skojarzenia zupełne (są to skojarzenia, przy których wszystkie towary zostały sprzedane) oraz równowagi spełniające pewne dodatkowe warunki.

Wprowadzimy teraz pewną liczbę niezbędnych definicji.

Definicja 5. Równowaga (W_1,W_2, ..., W_m) nazywa się równowagą podstawową, jeśli

$$\forall j \quad W_j = \varphi_W(S_j)$$

W równowadze podstawowej każdy kupujący spełnia co najwyżej jeden warunek W_j (bo zbiory $\varphi_W(S_j)$ są rozłączne). W ogólnym przypadku zbiory W_j nie muszą być rozłączne. Zauważmy, że bezpośrednio z definicji $\varphi_W(S_j)$ wynika, że

$$\forall j \quad \varphi_W(S_j) \subset W_j.$$

Definicja 6. Skojarzenie $\mu \subset E(K,S)$ nazywa się skojarzeniem zupełnym, jeśli

$$\forall j \quad \left| \mu(S_j) \right| = q_j .$$

Skojarzenie zupełne jest to więc skojarzenie takie, że wszyscy sprzedający sprzedali wszystkie swoje towary. Okazuje się (wynika to w prosty sposób z podanych definicji), że skojarzeniom zupełnym odpowiadają właśnie równowagi podstawowe, a mianowicie mamy:

Twierdzenie 1 (skojarzenia a równowagi)

Istnieje wzajemnie jednoznaczna odpowiedniość między skojarzeniami zupełnymi a równowagami podstawowymi.

Przedstawimy teraz związek między skojarzeniami stabilnymi a pewnymi rodzajami równowagi, które można nazwać równowagami stabilnymi.

Definicja 7. Jeśli $W = (W_1, W_2, ..., W_m)$ jest równowagą, to rozszerzeniem poprawiającym W nazywamy ciąg $V = (V_1, V_2, ..., V_m)$ taki, że

1) $\forall j \quad W_j \subset V_j$,

2) $(\forall K_i \in V_j \setminus W_j) \quad (\exists K_l \in W_j)$: $K_i >_j K_l$

Definicja 8. Równowaga podstawowa nazywa się równowagą stabilną jeśli dowolne jej rozszerzenie poprawiające jest równowagą.

Twierdzenie 2 [6]

Istnieje wzajemnie jednoznaczna odpowiedniość między skojarzeniami zupełnymi i stabilnymi a równowagami stabilnymi.

4. Równowagi porządkowe

Można określić jeszcze jedną ciekawą zależność między równowagami, a skojarzeniami stabilnymi dla podanego modelu rynku. Będziemy teraz rozpatrywali tzw. równowagi porządkowe. Równowaga porządkowa to rodzina zbiorów W_j, które są przedziałami początkowymi w zbiorze K ze względu na porządek $>_j$. Inaczej mówiąc, każdy zbiór W_j zawiera wszystkich kupujących, którzy są lepsi (dla sprzedającego S_j) od ustalonego kupującego K_i (łącznie z tym kupującym). Na przykład w modelu rynku nieruchomości będzie to zbiór kupujących, którzy gotowi są zapłacić za towar oferowany przez S_j ustaloną cenę p (warunek W_j jest więc tutaj po prostu klasycznym warunkiem cenowym w tradycyjnym modelu rynku).

Sformułujemy teraz formalną definicję równowagi porządkowej.

Definicja 9. Niech $W_j \subset K$ będzie warunkiem wyznaczonym przez sprzedającego S_j. Warunek W_j nazywamy warunkiem porządkowym, jeśli dla dowolnych $K_i, K_l \in K$ zachodzi

$$K_i \in W_j \ \wedge K_l >_j K_i \ \Rightarrow \ K_l \in W_j \, .$$

Definicja 10. Równowagę $(W_1, W_2, ..., W_m)$ nazywamy równowagą porządkową, jeśli wszystkie zbiory W_j ($j = 1, 2..., m$) są warunkami porządkowymi.

Wśród równowag porządkowych wyróżnimy tzw. równowagi brzegowe. Nazwijmy kupującego $K_i \in W_j$ kupującym minimalnym dla S_j, jeśli K_i jest minimalnym elementem zbioru W_j ze względu na relację $>_j$ (tzn. z punktu widzenia S_j kupujący K_i jest gorszy od wszystkich innych kupujących, którzy spełniają warunek W_j). Oznaczmy minimalnego kupującego dla S_j symbolem $minW_j$.

Definicja 11. Równowaga porządkowa $(W_1, W_2, ..., W_m)$ nazywa się równowagą brzegową, jeśli dla dowolnego j = 1, 2, ..., m, minimalny kupujący dla S_j tworzy popyt na dobro oferowane przez S_j.

Warunek brzegowości równowagi można też zapisać w postaci $minW_j \in \varphi_W(S_j)$. Oznacza to, że kupujący najniżej oceniany przez sprzedającego S_j (spośród tych, którzy spełniają warunki postawione przez S_j) uważa towar oferowany przez S_j za najlepszy wśród wszystkich towarów oferowanych na rynku (spośród tych, które są dla niego dostępne przy warunkach (W_1, W_2, ..., W_m)).

Twierdzenie 3 [5]

Istnieje wzajemnie jednoznaczna odpowiedniość między skojarzeniami stabilnymi i zupełnymi, a równowagami brzegowymi.

5. Przykłady

Przedstawimy teraz dwa przykłady ilustrujące zdefiniowane powyżej pojęcia.

Przykład 1

Załóżmy, ze mamy kupujących *A, B, C, D, E* oraz sprzedających *X, Y, Z*. Preferencje kupujących i sprzedających przedstawione są poniżej.

A: XYZ	X: BCADE
B: ZXY	Y: BACED
C: YXZ	Z: DACBE
D: ZXY	
E: ZYX	

Zapis *A: XYZ* oznacza, że kupujący *A* najchętniej zawarłby transakcję ze sprzedającym *X*, w drugiej kolejności z *Y*, a sprzedającego *Z* ustawia na końcu swojej listy preferencji. Niech podaże sprzedających będą równe $q_X = 2$, $q_Y = 2$, $q_Z = 1$. Określamy skojarzenie $\{AX, BX, CY, DZ, EY\}$ (symbol AX jest skróconym zapisem pary $\{A, X\}$ itd.). Łatwo można sprawdzić, że jest to skojarzenie stabilne, któremu odpowiada równowaga podstawowa $W_X = \{B, A\}$, $W_Y = \{C, E\}$, $W_Z = \{D\}$. Można tutaj określić wiele innych równowag wyznaczających to samo skojarzenie, np. zbiory $V_X = \{B, C, A\}$, $V_Y = \{A, C, E\}$, $V_Z = \{D\}$ tworzą rozszerzenie poprawiające dla W, zbiory $U_X = \{B, C, A\}$, $U_Y = \{B, A, C, E\}$, $U_Z = \{D\}$ tworzą równowagę porządkową, która jest równowagą brzegową, a zbiory $T_X = \{B, C, A, D\}$, $T_Y = \{B, A, C, E, D\}$, $T_Z = \{D, A\}$ tworzą równowagę porządkową, która nie jest równowagą brzegową.

Przykład 2

Załóżmy, że preferencje kupujących A, B, C i sprzedających X, Y, Z są następujące:

A: YZX X: ABC

B: YXZ Y: CAB

C: XZY Z: BAC

Zakładamy, że każdy sprzedający dysponuje jedną jednostką pewnego towaru oraz, że każdy kupujący gotowy jest zapłacić pewną maksymalną sumę za towar oferowany przez danego sprzedającego. Te maksymalne kwoty podane są w tabeli 1.

Tabela 1

Maksymalne kwoty oferowane przez kupujących

	X	Y	Z
A	130	130	115
B	120	120	120
C	110	140	110

Na przykład A gotowy jest zapłacić za towar oferowany przez X maksymalnie 130 zł. Zauważmy, ze preferencje sprzedających są wyznaczone przez kolejność ofert cenowych przedstawionych przez kupujących. Jeśli sprzedający określą ceny swoich towarów następująco: p_X (cena wyznaczona przez X) = 120, $p_Y = 130$, $p_Z = 110$, to otrzymamy warunki (zbiory kupujących, którzy są gotowi zapłacić cenę danego sprzedającego) $W_X = \{A, B\}$, $W_Y = \{C, A\}$, $W_Z = \{B, A, C\}$. Warunki te tworzą równowagę, gdyż odpowiadają skojarzeniu stabilnemu $\{AY, BX, CZ\}$. Zauważmy, że podwyższenie wszystkich cen do wartości $p_X = 130$, $p_Y = 140$, $p_Z = 120$ znowu daje równowagę odpowiadającą innemu skojarzeniu stabilnemu, a mianowicie $\{AX, BZ, CY\}$. Przykład ten pokazuje więc ciekawą własność funkcji popytu dla rozważanego modelu rynku: nawet przy wzroście wszystkich cen oferowanych na rynku dóbr popyt na te dobra może nie ulec zmianie (co jest niemożliwe w klasycznych, ciągłych modelach rynku).

Podsumowanie

W pracy wprowadzono pojęcie równowagi rynkowej dla modelu rynku wzorowanego na modelu kojarzenia Gale'a-Shapleya. Okazuje się, ze istnieje ścisły związek między skojarzeniami stabilnymi w sensie Gale'a-Shapleya

szczególnymi rodzajami równowag rynkowych (są to zdefiniowane w pracy równowagi stabilne i równowagi porządkowe). Niektóre z tych równowag mogą być traktowane jako równowagi cenowe (warunki cenowe są warunkami porządkowymi, a więc w oczywisty sposób wyznaczają równowagi porządkowe). Przykład 2 pokazuje, ze popyt w przedstawionym modelu rynku może zachowywać się nieco paradoksalnie, tzn. może nie reagować na wzrost cen wszystkich dóbr oferowanych na rynku. Model ten może więc opisywać zjawiska, których nie da się opisać za pomocą tradycyjnych, ciągłych modeli rynku.

Literatura

1. Camina E. (2006). A Generalized Assignment Game. Mathematical Social Sciences, 52, 152-161.
2. Gale D., Shapley L.S. (1962). College Admissions and the Stability of Marriage. American Mathematical Monthly, 69, 9-15.
3. Roth A.E., Sotomayor M.A. (1992). Two-sided Matching. A Study in Game-Theoretic Modeling and Analysis. Cambridge University Press.
4. Sotomayor M. (2007). Connecting the Cooperative and Competitive Structures of the Multiple-partners Assignment Game. Journal of Economic Theory, 134, 155-174.
5. Świtalski Z. Równowagi na rynkach z dwustronnymi preferencjami. (w druku).
6. Świtalski Z. (2008). Stability and Equilibria in the Matching Models. Scientific Research of the Institute of Mathematics and Computer Science, Czestochowa University of Technology, 2(7), 77-85.

Stanisław Walukiewicz
Polska Akademia Nauk w Warszawie

ZASADA ORTOGONALNOŚCI I PRZYKŁADY JEJ ZASTOSOWAŃ

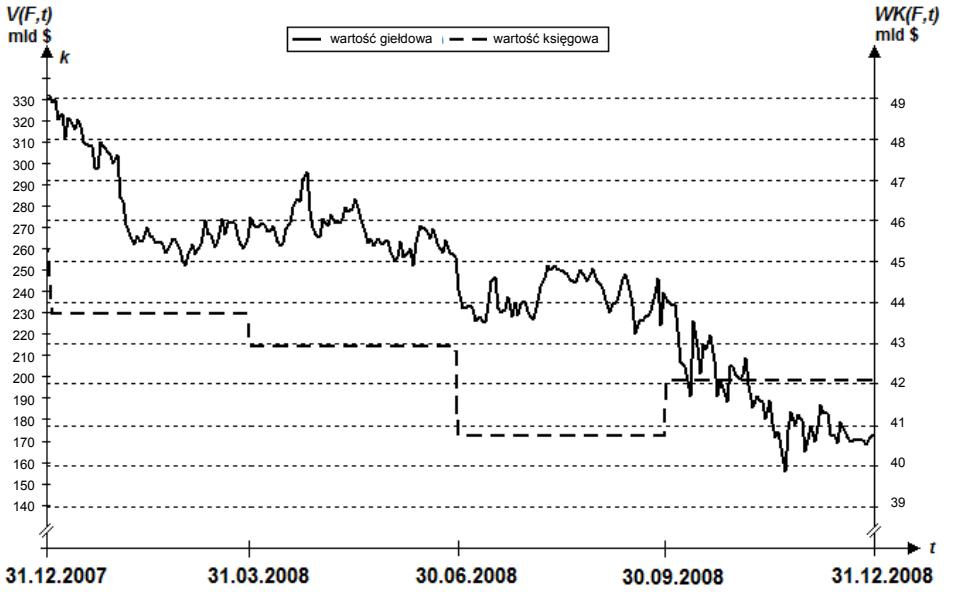
Wstęp

Pytanie „Ile to jest warte?" („Ile to kosztuje?") jest zadawane od zarania ludzkości, z tym, że w miarę rozwoju coraz bardziej skomplikowany i złożony jest podmiot w tym pytaniu. W warunkach równowagi rynkowej, gdy popyt równoważy podaż, a tylko taki przypadek będziemy rozpatrywać w tym opracowaniu, odpowiedź na to pytanie jest oczywista w przypadku prostej rzeczy, np. jabłka, lub prostej usługi, np. przejazdu pociągiem z Warszawy do Katowic. Odpowiedź ta bardzo komplikuje się, gdy spytamy, ile warta jest dana firma czy instytucja, np. Microsoft czy IBS PAN lub gdy będziemy chcieli wiedzieć, ile są warte wszystkie zasoby danego kraju czy regionu. Nikt poważnie myślący nie zamierza kupować czy sprzedawać danego kraju; chodzi tu o wycenę lub oszacowanie wartości wszystkich, ale absolutnie wszystkich zasobów zarówno materialnych, jak i niematerialnych danego kraju/firmy. Jest to problem stawiany w ekonomii i naukach o zarządzaniu od wielu dziesięcioleci, a ostatnio wzięła go na warsztat Komisja ds. Mierzenia Wyników Gospodarczych i Postępu Społecznego, kierowana przez noblistę Josepha Stiglitza. Powołał ją w lutym 2009 prezydent Francji Nicolas Sarkozy, a we wrześniu 2009 ukazał się jej raport.

W przypadku Microsoftu, firmy notowanej na giełdzie, odpowiedź na to pytanie wydaje się na pierwszy rzut oka prosta: wartość (rynkowa, giełdowa) Microsoftu (np. 31.03.2008) jest równa iloczynowi liczby akcji będących w obrocie (9.566.808.000 sztuk) i ceny akcji na zamknięciu notowań w tym dniu (28,38 dol. USA), co daje ponad 271 mld dol. (dokładnie 271.506.011.040 dol.), niemal tyle samo, ile wynosił PKB Polski w 2008 roku[1]. Na rys. 1 linią ciągłą, lewa skala, pokazano, jak zmieniała się wartość rynkowa Microsoftu od 31.12.2007 do 31.12.2008, a linią przerywaną, prawa skala, pokazano jak zmieniała się w tym czasie jej wartość księgowa, czyli to, co zarejestrowano

[1] Wszystkie dane o firmie Microsoft zostały zaczerpnięte z Wikipedii w lutym 2009.

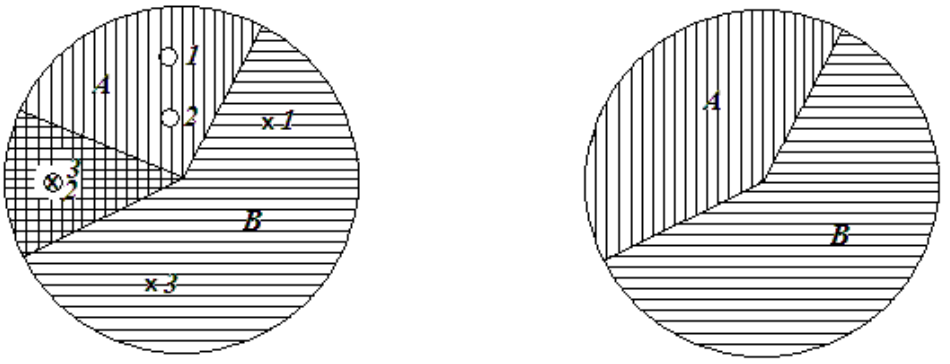
w bilansach Microsoftu, sporządzanych na koniec każdego kwartału zgodnie ze wszystkimi regułami (klasycznej) rachunkowości. Z rys. 1 wynika, iż zawsze, nawet w czasie obecnego kryzysu ekonomicznego, wartość rynkowa jest większa od wartości księgowej, zwykle 5-6 razy. Wtedy w naturalny sposób pojawia się pytanie, dlaczego inwestorzy (giełdowi) przepłacają wielokrotnie wartość księgową lub bardziej dobitnie, czego, jakich wartości (kapitałów, zasobów) nie uwzględnia (klasyczna) rachunkowość w swoich bilansach? Odpowiedź na to pytanie jest zasadniczym celem tej pracy.



Rys. 1. Wartość rynkowa i wartość księgowa Microsoftu w 2008 roku

By przybliżyć cel tej pracy rozważmy wszystkie, ale to absolutnie wszystkie, wartości (zasoby, kapitały) danej firmy *F* w chwili *t*. Inaczej mówiąc, analizujemy wycinek rzeczywistości *X* (wszystkie wartości firmy *F*) w chwili *t*. Fakt, że analizujemy wszystkie wartości będziemy na rysunkach oznaczać za pomocą koła. Załóżmy, ze analiza wszystkich wartości tego wycinka rzeczywistości jest dla nas zbyt skomplikowana i decydujemy się podzielić *X* na dwie formy (kategorie, pojęcia, wymiary itp.) *A* oraz *B*. Wtedy możliwe są dwa przypadki, przedstawione na rys. 2: albo forma *A*, zakreskowana pionowo, zachodzi na formę *B*, zakreskowaną poziomo, co przedstawiono na rys. 2a jako część koła zakreskowaną w kratkę, albo tak, jak na rys. 2b formy te są rozłączne, tworząc rozbicie *X* na *A* oraz *B*. Różnicę między podziałem a rozbiciem można zobrazować następującym przykładem: jeżeli talerz upuszczony na pod-

łogę rozbije się, to mamy jego rozbicie na co najmniej dwie części, a nie podział. Inaczej mówiąc, każde rozbicie jest podziałem, ale nie każdy podział jest rozbiciem.



a) podział *X* na *A* oraz *B* b) rozbicie *X* na *A* oraz *B*

Rys. 2. Różnica między podziałem a rozbiciem

Załóżmy dalej, że wyznaczamy wartość najpierw formy *A* dodając wartości jej poszczególnych elementów (pozycji bilansowych) oznaczonych kółkami, a później *B* dodając wartości elementów oznaczonych krzyżykami. Na rys. 2a drugi element formy *B* pokrywa się, jest tożsamy, z trzecim elementem *A*, a więc mamy tutaj przykład naruszenia podstawowej zasady rachunkowości, która mówi, że każdy element z analizowanego wycinka rzeczywistości jest uwzględniany w bilansie jeden i tylko jeden raz. Taki przypadek nie może zaistnieć na rys. 2b, gdy formy są rozłączne. W naukach społecznych, takich jak ekonomia, zarządzanie, rachunkowość, socjologia czy nauki polityczne nie potrafimy dziś przedstawić analizowanego wycinka rzeczywistości *X* jako rys. 2a albo 2b. To opracowanie należy traktować jako jeden z pierwszych kroków [7; 8] w tym kierunku.

Dalej formułujemy Zasadę Ortogonalności, która pozwala stwierdzić, kiedy dwie formy są rozłączne, tak jak na rys. 2b oraz omawiamy jej matematyczną i graficzną interpretację. Jest to zasadniczy rezultat tego opracowania. Pozostałe rozdziały tej pracy zawierają przykłady zastosowań Zasady Ortogonalności w analizie wszystkich, ale absolutnie wszystkich zasobów (kapitałów, wartości) danej szeroko rozumianej firmy lub też kraju/regionu. Analiza ta jest prowadzona zgodnie z zasadą od ogółu do szczegółu.

Odpowiedź na pytanie sformułowane na początku opracowania pada zwykle w momencie zakupu/inwestycji. By analizować takie inwestycje/zakupy, potrzebny nam będzie model inwestycyjny [7], tj. pewien ogólny schemat

postępowania inwestorów. W takim modelu wyróżniamy trzy okresy (fazy): przeszłość lub „wczoraj" (zawsze ujęte w cudzysłów), teraźniejszość lub „dziś" oraz przyszłość, czyli „jutro" i inwestor zawsze podejmuje decyzję „dziś" na podstawie danych/informacji z „wczoraj", aby „jutro" osiągać zyski lub jak najmniej stracić. Na trzech przykładach pokażemy, jak ten model działa w praktyce.

Gdy kupuję jabłka na rynku, to decyzję o ich zakupie podejmuję, powiedzmy, w 10 sekund („dziś") na podstawie mojego doświadczenia w przeszłości („wczoraj") po to, by zjeść je w najbliższych dniach („jutro"). Podobnie inwestor kupujący akcje Microsoftu na giełdzie „dziś" na podstawie oficjalnych i nieoficjalnych informacji, np. plotek o tej firmie z "wczoraj", po to, aby odnieść sukces (ekonomiczny) „jutro". Długości tych trzech faz mogą być bardzo różne i zależą od inwestora. Ważne jest w modelu inwestycyjnym to, że „dziś" to granica, być może nieskończenie krótka (mgnienie oka), między praktycznie nieskończonym „wczoraj" i takim samym „jutro", i że zawsze te trzy fazy występują.

Jako trzeci przykład rozważmy finansowanie działalności statutowej IBS PAN przez Ministerstwo Nauki i Szkolnictwa Wyższego, które w tym roku („dziś") ocenia nasze publikacje i nasz dorobek z ostatnich dwóch lat („wczoraj"), by określić wysokość dofinansowania w 2010 roku („jutro"). Tak więc, w tym przykładzie „dziś" oraz „jutro" trwają rok, natomiast „wczoraj" – 2 lata. Inaczej mówiąc, Ministerstwo w tym roku przygotowuje na podstawie informacji z lat 2007-2008 odpowiedź na pytanie, ile wart jest IBS PAN w 2010 roku.

1. Zasada Ortogonalności

Osie współrzędnych (wymiary, zmienne – oznaczmy je jako t oraz k) na rys.1 przecinają się pod kątem prostym, czyli są wzajemnie ortogonalne, prostopadłe. Oznacza to, że zmienne te są (liniowo) niezależne, co w konsekwencji praktycznie oznacza, iż niezależnie od tego, jak dokładnie będziemy mierzyć zarówno czas t (w sekundach, godzinach czy w dniach), jak też kapitał/wartość k (w tysiącach czy też w milionach dolarów), to zależność wartości rynkowej, czy też wartości księgowej od czasu będzie zawsze taka sama, tj. jej kształt będzie taki sam, np. wartość księgowa będzie zawsze funkcją schodkową. Przypadek, gdy te osi nie są prostopadłe analizujemy nieco dalej na rys. 3. Dziś w naukach społecznych mamy niewiele par zmiennych (form, pojęć, kategorii, wymiarów itp.), o których możemy twierdzić, że są wzajemnie ortogonalne lub rozłączne. Sformułujemy teraz Zasadę Ortogonalności, która pozwala rozszerzyć katalog takich pojęć.

Zasada Ortogonalności. Dwa pojęcia (formy, kategorie, wymiary, koncepcje, procesy itp.) *A* oraz *B* są ortogonalne lub rozłączne wtedy i tylko wtedy, kiedy istnieje obiektywna, prosta, jednowymiarowa reguła decyzyjna typu tak-nie, pozwalająca zdecydować o przynależności dowolnego elementu/obiektu z rozważanego wycinka rzeczywistości *X* albo do *A*, albo do *B*.

Omówimy teraz cztery warunki, jakie nałożyliśmy na regułę decyzyjną w Zasadzie Ortogonalności.

Wymaganie, iż reguła decyzyjna ma być obiektywna oznacza, że wynik jej stosowania nie zależy ani od osoby stosującą tę regułę, ani też od czasu i miejsca, gdzie to zachodzi. W naszym rozumieniu reguła decyzyjna jest prosta, jeżeli można ją zdefiniować w 1-2 zdaniach powszechnie, potocznie używanego języka, bez odwoływania się do specjalistycznej terminologii. Na przykład, niech *X* będzie bilansem danej firmy sporządzanym przez księgowych. Wtedy nie jest prostą regułą decyzyjna opisana na, powiedzmy, 10 stronach, której zrozumienie wymaga studiowania podręczników akademickich.

Podobne założenia legły w naszym wymaganiu, aby reguła decyzyjna była jednowymiarowa i typu tak-nie. Kontynuując przykład z księgowymi, dzisiaj i w dającej się przewidzieć przyszłości trudno jest wymagać od nich stosowania analizy wielokryterialnej (wielowymiarowej) czy też rachunku zbiorów rozmytych.

Ponieważ Zasada Ortogonalności jest fundamentem, podstawą podstaw naszych badań, to podamy jej matematyczną interpretację, tj. zapiszemy ją w języku teorii zbiorów (matematyki). Niech *X* będzie rozważanym wycinkiem rzeczywistości, tj. zbiorem wszystkich rozważanych elementów/obiektów, wtedy reguła decyzyjna rozbija ten zbiór na dwa rozłączne, niepuste podzbiory *A* oraz *B* takie, że

$$X = A \cup B \text{ oraz } A \cap B = \varnothing,\ A \neq \varnothing,\ B \neq \varnothing. \tag{1}$$

Wzór (1) można zilustrować następująco: jeżeli talerz (nasz zbiór *X*) upuszczony na podłogę rozbije się, to będzie to rozbicie zbioru *X* na co najmniej dwa niepuste podzbiory, takie, że dowolne dwa z nich nie mają elementów wspólnych.

Z Zasady Ortogonalności wynika, iż formy/zmienne, które są mierzone w różnych, liniowo niezależnych jednostkach są ortogonalne. Regułą decyzyjną jest tu właśnie fakt, że zmienne te są mierzone w tych różnych (liniowo niezależnych) jednostkach. Z Zasady Ortogonalności wynika, że zmienne *t* oraz *k* z rys. 1 są ortogonalne. Najciekawszy jest przypadek, kiedy dwie zmienne/formy mierzone w tych samych jednostkach są ortogonalne

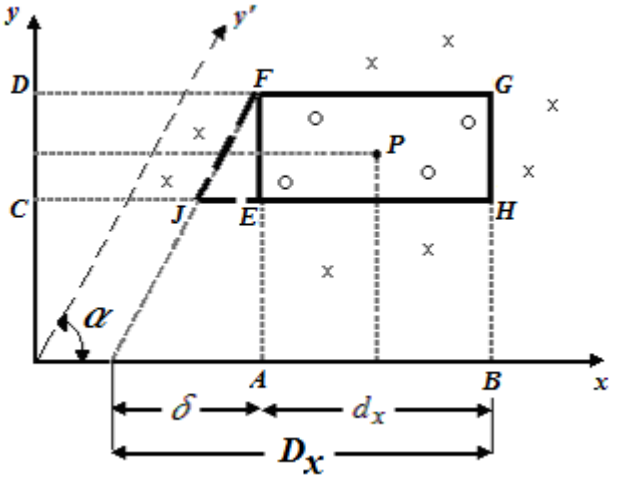
Z klasycznej mechaniki wiemy, że zgodnie z zasadą zachowania energii, energia E rozważonego ciała jest równa sumie energii potencjalnej E_p oraz energii kinetycznej E_k

$$E = E_p + E_k$$

Energia potencjalna ma ciało znajdujące się w spoczynku (bezruchu), a energia kinetyczna, to energia ruchu. Łatwo można sprawdzić, że reguła decyzyjna „spoczynek (bezruch)-ruch" spełnia wszystkie cztery wymagania sformułowane w Zasadzie Ortogonalności. Wynika stąd, iż energia potencjalna jest ortogonalna do energii kinetycznej, a więc ich wartości możemy dodawać, bez obawy popełnienia jakiegokolwiek błędu (rys. 2b). Udowodniliśmy w ten sposób jeszcze raz powszechnie znane prawo zachowania energii w przyrodzie. To opracowanie ma na celu wykazanie, że istnieje analog tego prawa w szeroko rozumianej gospodarce.

Na zakończenie omówimy tu graficzną interpretację Zasady Ortogonalności. W naukach społecznych i nie tylko, wiele pomiarów jest obarczonych pewnym błędem. Przykładowo, przewidywania wyników wyborów, np. prezydenckich, są obarczone statystycznym błędem ± 3%.

By uczynić nasze rozważania bardziej konkretnymi rozważmy problem, w którym poszukujemy odpowiedzi na pytanie: które z badanych polskich małych i średnich przedsiębiorstw (MŚP) są innowacyjne, a które nie? Załóżmy, że innowacyjność MŚP można opisać za pomocą dwóch kryteriów (wymiarów, osi) x oraz y. To ograniczenie do dwóch wymiarów wynika tylko i wyłącznie z chęci graficznej ilustracji rozważanego zjawiska. Załóżmy na początek, że osie x oraz y są prostopadłe, a więc kryteria im odpowiadające są ortogonalne zgodnie z Zasadą Ortogonalności, jak to przedstawiono na rys. 3.



Rys. 3. Graficzna interpretacja Zasady Ortogonalności

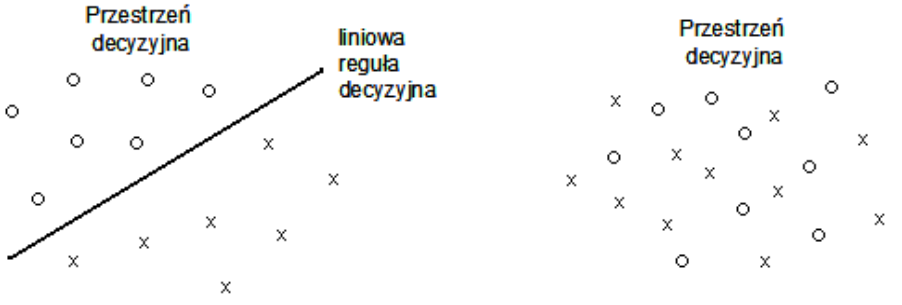
Niech pomiary dla innowacyjnych MŚP dla kryterium *x* mieszczą się między punktami *A* oraz *B,* a dla kryterium *y* – między *C* oraz *D*. Wtedy innowacyjnym MŚP odpowiadają kółka w prostokącie *EFGH,* a firmom nieinnowacyjnym – krzyżyki poza tym prostokątem. Zauważmy jeszcze, że w miarę wzrostu dokładności pomiarów, prostokąt ten będzie malał, a w granicznym przypadku będzie on punktem *P* na rys. 3.

Co będzie, gdy osie *x* oraz *y* nie są wzajemnie prostopadłe, gdy przecinają się one pod pewnym, ostrym lub rozwartym kątem α, różnym od *90°*, tak jak to pokazano na rys. 3 linią przerywaną (oś *y'*)? Wtedy nasze innowacyjne MŚP znajdują się w trapezie *FGHJ* i co więcej ten brak ortogonalności kryteriów indukuje dodatkowy błąd δ tak, że teraz łączny błąd wzdłuż osi *x* wynosi

$$D_x = \delta + d_x.$$

Zauważmy jeszcze, że dla α równego *0°* lub *180°* mamy nieskończenie duży błąd dodatkowy. Taki błąd popełniamy, gdy chcemy zjawisko dwuwymiarowe opisać za pomocą jednego wymiaru. Z nieskończenie wielu wartości α dla jednej i tylko jednej wartości α*=90°,* gdy kryteria są ortogonalne, błąd dodatkowy jest równy zeru. Z trygonometrii wiemy, jak przeliczać współrzędne dla dowolnych wartości α i jest to zadanie dla ucznia szkoły średniej. Niestety, w naukach społecznych nie mamy odpowiednika trygonometrii, dlatego potrzebna jest Zasada Ortogonalności.

Ortogonalność kryteriów *x* oraz *y* oznacza, że istnieje reguła decyzyjna, opisana w Zasadzie Ortogonalności, która pozwala oddzielić przedsiębiorstwa innowacyjne (kółka) od nieinnowacyjnych (krzyżyki). Jeżeli przyjmiemy, że graficzną reprezentacją tej reguły w przestrzeni decyzyjnej (to w ogólnym przypadku nie jest układ *x0y)* jest linia prosta, to w przypadku ortogonalności kryteriów mamy sytuację jak na rys. 4a, gdzie po jednej stronie tej prostej mamy innowacyjne MŚP, a po drugiej nieinnowacyjne firmy.



a) istnieje liniowa reguła decyzyjna b) nie istnieje liniowa reguła decyzyjna

Rys. 4. Graficzna ilustracja ortogonalności (4a) i nieortogonalności (4b) kryteriów

Na rys. 4b przedstawiono sytuację, gdy kryteria nie są ortogonalne. Wtedy zgodnie z Zasadą Ortogonalności nie istnieje reguła decyzyjna (prosta) rozdzielająca te dwie grupy firm (kółka od krzyżyków). Dalej podamy zarys metodologii budowy takich w maksymalnym stopniu ortogonalnych kryteriów w projektowanych kwestionariuszach badawczych.

2. Wartości materialne i niematerialne – przykład 1

By lepiej zrozumieć, jak złożonym pojęciem jest „wartość firmy" rozważmy, co kryje się za nim na przykładzie Microsoftu, naszej szeroko rozumianej firmy *F*.

Microsoft, światowy potentat w oprogramowaniu, zatrudnia 76 tys. pracowników w swoich filiach zlokalizowanych w 102 krajach. Kapitałem tej firmy są jej zasoby finansowe lub jej kapitał finansowy (oszczędności, akcje, obligacje itp., ale również długi, kredyty i inne zobowiązania), jak też jej zasoby materialne lub jej kapitał materialny (budynki, samochody, meble, komputery itp., ale też oprogramowanie, gdyż firma ta korzysta z oryginalnego oprogramowania innych firm, a więc posiada/kupiła licencje na takie oprogramowanie, jak też być może ma inne prawa własności intelektualnej). Microsoft zatrudnia wybitnych programistów/analityków z ich wiedzą, doświadczeniem, kompetencjami, talentami itp., co stanowi bardzo istotny zasób tej firmy i co krótko będziemy nazywać jej kapitałem ludzkim. Dzisiaj oprogramowanie, a szczególnie to, które dostarcza Microsoft, to efekt wspólnej pracy twórczej tysięcy i dziesiątków tysięcy programistów, ich wzajemnego zaufania, profesjonalnej etyki itp., co będziemy nazywać jej kapitałem społecznym. Powszechnie uważa się, że o sile i potędze Microsoftu decyduje około 50 czołowych programistów/analityków i menedżerów, ich kompetencje, talenty, wzajemne zaufanie i współpraca, gdyż to oni decydują, co i jak Microsoft produkuje i to oni tworzą atmosferę pracy twórczej w tej firmie.

Z tego przykładu wynika, iż analiza kapitału wszystkich, ale to absolutnie wszystkich zasobów firmy *F* jest w ogólnym przypadku bardzo trudna. Kapitał firmy *F*, jest mierzony w jednostkach monetarnych. Stąd wnika, iż wszystkie części tego kapitału są również mierzone w jednostkach monetarnych. Dlatego rozbijamy kapitał firmy *F*, zgodnie z Zasadą Ortogonalności na dwie formy/kategorie, a mianowicie: wartości (kapitały, zasoby) materialne (WM) oraz wartości niematerialne (WNM). Regułą decyzyjną jest tu stwierdzenie, że wartości materialne, takie jak pieniądze, budynki, samochody, komputery, oprogramowanie, patenty itp., można dotknąć i może to zrobić każdy, zawsze i wszędzie, natomiast wartości niematerialnych, takich jak wiedza, doświadczenie, talent, zaufanie, współpraca itp., nie można dotknąć.

Może pojawić się pytanie, dlaczego oprogramowanie zaliczamy do wartości materialnych, czy naprawdę oprogramowanie można dotknąć? Mówimy o leganie zakupionym oprogramowaniu, którego właściciel ma (kupił) licencję (prawo) na jego użytkowanie. Zauważmy, że zawsze, nawet w dobie elektronicznych podpisów i elektronicznych licencji, można taką licencję wydrukować na papierze, następnie podpisać i papier ten można dotknąć. Zupełnie analogiczne rozumowanie pokazuje, że patenty, jak również wszelkie prawa własności intelektualnej (prawa autorskie) zgodnie z Zasadą Ortogonalności powinny być zaliczone do wartości materialnych.

W naszych badaniach zakładamy, że współczesną (klasyczną) rachunkowość należy zmieniać stopniowo, ewolucyjnie, a nie rewolucyjnie, jak sugerują np. Edvinsson i Malone [1] w swej książce na temat rosnącego znaczenia kapitału społecznego w gospodarce opartej na wiedzy. Odpowiedź na to pytanie jest więc zgodna z ewolucyjnością naszego podejścia, a mianowicie z faktem, iż dzisiaj wydatki na oprogramowanie i patenty księguje się jako wartości materialne. Powyższe rozważania są dobrym przykładem na to, że księgowi już dziś, choć w sposób niedoskonały, stosują Zasadę Ortogonalności, nie używając tej nazwy.

Ktoś może postawić pytanie, dlaczego wiedzę i doświadczenie zaliczamy do wartości niematerialnych, przecież różnorodne świadectwa, dyplomy i zaświadczenia są wydawane na papierze. Naszą odpowiedź sformułujemy w trzech punktach: po pierwsze, człowiek, jako pojedyncza istota ludzka czy też zespół ludzi to zagadnienie znacznie bardziej złożone i skomplikowane niż nawet najbardziej rozbudowane oprogramowanie albo też najbardziej oryginalna myśl (patent); a więc mamy tu do czynienia z nową jakością. Po drugie, współczesna rachunkowość, jak i wszystkie nauki społeczne, nie potrafi wartościować wymienionych wyżej wartości niematerialnych, poprzestając na stwierdzeniu, że są one trudno mierzalne bądź też stwierdzając wprost, że są one niemierzalne [12] w odróżnieniu od oprogramowania i patentów. Po trzecie, założenie, że każda wiedza, doświadczenie, talent, każdy akt zaufania czy współpracy ma swój „papier" jest naiwne i nierealistyczne i opisywałoby zawsze sytuację z „wczoraj", a nas przede wszystkim interesuje „dziś" oraz „jutro", zdefiniowano we wstępie.

Nasze zainteresowanie „dziś" oraz „jutro" zmusza nas do częstego używania zwrotów typu itp. oraz itd. lub też zwrotu wszystkie, ale to absolutnie wszystkie, gdyż dziś nie można przewidzieć, jakie nowe produkty, usługi czy formy współpracy pojawią się za 5 czy 25 lat. Chyba najlepszym przykładem jest tu telefon komórkowy: 15-20 lat temu to urządzenie było całkowicie nieznane, dziś ma je prawie każdy, niektórzy nawet po kilka sztuk, a jego wpływ na sposób porozumiewania się i współpracy ludzi trudno przecenić.

3. Kapitał ludzki i społeczny – przykład 2

Zastosujemy tu Zasadę Ortogonalności do analizy wartości niematerialnych firmy *F*, wartości, które w powszechnej opinii są, w ogólnym przypadku, bardzo złożone i skomplikowane, gdyż są one nierozerwalnie związane z człowiekiem, jako pojedynczą istotą ludzką, jak również ze wszystkimi możliwymi pozytywnymi i/lub negatywnymi oddziaływaniami między ludźmi w danej społeczności/grupie lub społeczeństwie, np. wśród pracowników firmy *F*.

Spośród wszystkich wartości niematerialnych wyróżniamy wartości, takie jak kompetencje, doświadczenie, wiedza, talent, zdolności np. przywódcze, zdrowie, energia życiowa itp., które są związane z pracownikami firmy *F* traktowanymi jako oddzielne istoty ludzkie i nazywamy je kapitałem ludzkim (KL) firmy *F*, a całą resztę wartości niematerialnych związanych z relacjami co najmniej dwóch osób (współpraca, zaufanie itp.) nazywamy kapitałem społecznym (KS) tej firmy. Obrazowo rzecz ujmując: kapitał ludzki idzie po pracy z pracownikiem do domu, a kapitał społeczny zostaje w miejscu pracy, które obecnie w dobie Internetu należy traktować bardzo szeroko, włączając do niego np. sieć telekomunikacyjną łączącą pracowników firmy *F*. Jeszcze inaczej: kapitał ludzki jest zawsze związany z pojedynczą istotą ludzką, a kapitał społeczny powstaje tam, gdzie jest oddziaływanie, w pozytywnym (zaufanie, solidarność, wyrozumiałość itp.) bądź negatywnym (podejrzliwość, zakłamanie itp.) sensie, co najmniej dwóch osób. Odkryliśmy tym samym regułę decyzyjną do Zasady Ortogonalności: jedna osoba – kapitał ludzki, co najmniej dwie osoby – kapitał społeczny. Tak więc kapitał ludzki i kapitał społeczny są rozłączne lub ortogonalne. Łatwo można sprawdzić, ta reguła decyzyjna spełnia wszystkie warunki podane w Zasadzie Ortogonalności.

Z całą mocą należy podkreślić, że słowo „kapitał" w terminach „kapitał ludzki" oraz „kapitał społeczny" używamy z pełną świadomością. Terminy te powstały przez rozbicie formy kapitały (wartości) niematerialne na właśnie te dwie grupy. Tak więc podstawową, uniwersalną miara tych kapitałów są jednostki monetarne (pieniądze). Pokażemy, że zarówno kapitał ludzki, jak i kapitał społeczny mają wiele cech wspólnych z kapitałem finansowym (pieniędzmi) lub kapitałem materialnym (rzeczami). Jednak nie jest tak zawsze, np. często ocenia się kandydata do pracy (jego kapitał ludzki) w punktach, które potem jakoś przelicza się na pieniądze (jego wynagrodzenie). Podobnie, wyniki IBS PAN w ostatnich dwóch latach wycenia się najpierw w punktach, które są potem przeliczane na dotację budżetową na 2010 rok. Pokażemy, że jest to w zasadzie kapitał ludzki i kapitał społeczny Instytutu.

Ponieważ kapitał społeczny powstaje w grupie co najmniej dwóch osób, to relacje miedzy tymi osobami stanowią podstawowy element (składową) tego kapitału. Wszystkie relacje możemy zawsze podzielić na formalne i nieformalne i wtedy w sposób naturalny pojawia się pytanie, czy ten podział jest rozbiciem zbioru wszystkich relacji (zbioru *X*) w sensie Zasady Ortogonalności, tj., czy relacje formalne są ortogonalne (rozłączne, niezależne) do relacji nieformalnych?

Całkowicie niezależnie od naszych badań Li [2, s. 227-246] pokazał, że potrzeba pięciowymiarowej reguły decyzyjnej, aby odróżnić relacje formalne od nieformalnych. Te pięć wymiarów to:

1. Kodyfikacja. Relacje formalne są skodyfikowane/zapisane w regulaminach, spisach procedur, konstytucjach, kodeksach itp., podczas gdy relacje nieformalne nie są skodyfikowane.

2. Źródło. Relacje formalne mają swoje zasadnicze źródło wewnątrz firmy lub kraju (np. w historii), a na relacje nieformalne ogromny wpływ ma otoczenie, szczególnie teraz w dobie Internetu i globalnej wioski telewizyjnej.

3. Egzekutywa. Relacje formalne są ściśle egzekwowane lub powinny być tak egzekwowane (przepisy ruchu drogowego, kodeks karny itp.), natomiast podejście do relacji nieformalnych jest elastyczne, na luzie itp.

4. Władza. Relacje formalne wprowadzają hierarchiczną (pionową) strukturę władzy, a w relacjach nieformalnych wszyscy są sobie równi, a więc mamy horyzontalną (poziomą) strukturę władzy.

5. Personalizacja. Istotą relacji nieformalnych jest człowiek, traktowany jako samodzielna istota ludzka z jego wszystkimi zaletami, wadami itp., podczas gdy relacje formalne są z definicji odpersonalizowane, dotyczą stanowisk itp.

Wynika stąd, że wartości relacji formalnych, niezależnie od tego, jak będą one mierzone/szacowane (np. w punktach, pieniądzach itp.), nie można dodawać do wartości relacji nieformalnych, gdyż mamy, w ogólnym przypadku, sytuację przedstawioną na rys. 2a (podział), a nie na rys. 2b (rozbicie).

Sport, zwłaszcza profesjonalny, dostarcza doskonałych przykładów na to, że warto odróżniać kapitał ludzki od kapitału społecznego. W sportach zespołowych mecz to punktowe (w danym czasie i miejscu) porównanie kapitałów społecznych dwóch grających drużyn (ich współdziałania, zaufania itp.), które jest wyceniane zgodnie z zasadami tego meczu/dyscypliny sportowej i zwykle nie są to wartości kapitału społecznego mierzone w pieniądzach. Historia zna tysiące przykładów, gdy w danym meczu najlepszy był, tj. posiadał największy kapitał ludzki, zawodnik drużyny przegranej, bo w meczu porównujemy kapitały społeczne, a nie ludzkie. Podobnie, w punkcie, np. kilka dni później i/lub w innym miejscu wynik meczu między tymi dwoma drużynami może być odwrotny. Dlatego wprowadzono różnorodne tabele i rankingi (sezonowe, roczne,

wszech czasów, lokalne, krajowe, kontynentalne, światowe itp.), które eliminują punktowość takich porównań. Wtedy liderzy tych rankingów mają największy kapitał ludzki/społeczny liczony w pieniądzach, a gdy przez dłuższy czas przestają być liderami, to wartości ich kapitału ludzkiego/społecznego spadają. Podsumowując: w sportach zespołowych nie grają nazwiska (kapitały ludzkie), lecz kapitały społeczne, a rankingi odzwierciedlają kapitały ludzkie zawodników i kapitały społeczne ocenianych drużyn.

4. Kapitał finansowy i materialny – przykład 3

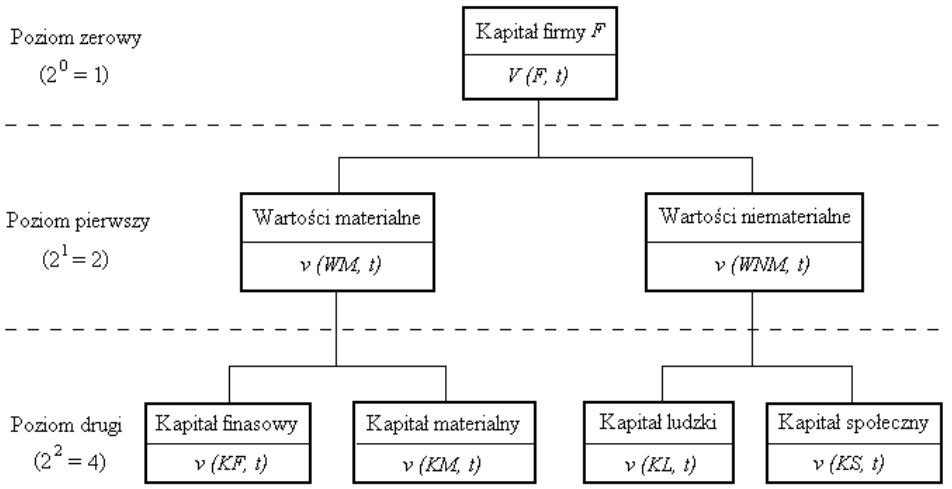
Zupełnie analogicznie rozważymy teraz wartości/kapitały materialne naszej firmy *F*. Spośród wszystkich wartości materialnych firmy *F* wyróżniamy te, które mają bezpośredni charakter finansowy (oszczędności, gotówka, akcje, obligacje, oczekiwane wpłaty itp., ale również długi, kredyty i wszelkie inne zobowiązania) i nazywamy je kapitałem finansowym (KF) firmy *F*, a wszystkie pozostałe wartości materialne (budynki, maszyny, meble, komputery, oprogramowanie, patenty itp.) nazywamy kapitałem materialnym (KM) tej firmy. Łatwo można sprawdzić, że reguła „bezpośredni charakter finansowy" spełnia wszystkie warunki Zasady Ortogonalności. Dla tych, którzy uważają, że „bezpośredni charakter finansowy" jest niezbyt precyzyjną regułą decyzyjną można odpowiedzieć, że obecnie znamy kilkanaście rodzajów kapitału finansowego (definiując go wymieniliśmy 8 takich rodzajów), a więc w ostateczności można zawsze sporządzić ich listę, którą będzie się rozszerzać w miarę pojawiania się nowych produktów finansowych. Tak więc kapitał finansowy i kapitał materialny są rozłączne lub ortogonalne. Kapitał finansowy mierzymy w jednostkach monetarnych, a kapitał materialny jako część wartości materialnych mierzymy również w pieniądzach wykorzystując do tego techniki amortyzacji tego kapitału.

Nasze rozważania w przykładach 1-3 mają wielopoziomowy charakter (rys. 5). Na poziomie zerowym (2^0=1) mamy tylko jedną formę (rodzaj) kapitału, który tworzą wszystkie, ale to absolutnie wszystkie zasoby i aktywa naszej szeroko rozumianej firmy *F*. Jest to kapitał tej firmy, a $V(F,t)$ jest jego wartością w jednostkach monetarnych w chwili *t*. Ponieważ analiza tak rozumianego kapitału jest trudna, to zgodnie z Zasadą Ortogonalności rozbijamy go, tak jak opisaliśmy to wyżej, na dwie (mniejsze) formy: wartości materialne i wartości niematerialne. Dlatego na poziomie pierwszym (2^1=2) mamy dwie formy kapitału firmy *F*, które są rozłączne (nie mają wspólnych elementów), czyli są ortogonalne.

Ponieważ zarówno wartości materialne, jak i niematerialne są nadal zbyt trudne, aby oszacować/określić ich wartości, to zgodnie z Zasadą Ortogonalności rozbijamy każdą z nich na dwie mniejsze formy i dlatego na poziomie drugim (2^2=4) mamy cztery formy kapitału firmy *F*: kapitał finansowy, kapitał materialny, kapitał ludzki i kapitał społeczny. Rysunek 5 wprowadza oczywisty system oznaczeń, którego dalej będziemy używać.

Udowodniliśmy w ten sposób następujący lemat.

Lemat 1. Zdefiniowane powyżej w indykatywny sposób kapitał finansowy, kapitał materialny, kapitał ludzki oraz kapitał społeczny są wzajemnie rozłączne lub ortogonalne i tworzą rozbicie wszystkich zasobów na te cztery i tylko cztery formy kapitału firmy *F*.



Rys. 5. Ilustracja powstawania czterech form kapitału

Wartość każdej z form kapitału możemy rozważać jako pojęcie statyczne bądź dynamicznie, dlatego poczynając od tego miejsca będziemy wymiennie używać terminów forma oraz wymiar (zmienna) kapitału. Nasze rozważania podsumowuje następujący

Lemat 2. Kapitał traktowany jako pojęcie statyczne bądź dynamiczne ma cztery wymiary.

Lemat 2 jest odpowiedzią na pytanie: ile wymiarów ma kapitał? Te cztery formy (wymiary) ściśle ze sobą współpracują, tak jak energia potencjalna i kinetyczna danego ciała, chociaż są one wzajemnie rozłączne/ortogonalne.

5. Równanie Fundamentalne

Z lematu 1 oraz punktu 4 wynika, iż cztery wyżej zdefiniowane formy nie mają elementów wspólnych, a więc możemy dodawać wartości odpowiadających im kapitałów. Udowodniliśmy w ten sposób zależność

$$V(F,t) = v(KF,t) + v(KM,t) + v(KL,t) + v(KS,t)$$
$$\text{dla każdego } t \text{ z „wczoraj“, „dziś” oraz „jutro”} \qquad (2)$$

którą będziemy dalej nazywać Równaniem Fundamentalnym. Mówi ono, że w gospodarce rynkowej, w stanie równowagi, gdy popyt równoważy podaż, wartość firmy w dowolnej chwili *t* z „wczoraj”, „dzisiaj” oraz „jutro” jest równa sumie wartości czterech i tylko czterech form kapitału wyznaczonych dla tej chwili, a mianowicie kapitału finansowego, materialnego, ludzkiego oraz społecznego. Nazwaliśmy zależność (2) Równaniem Fundamentalnym, gdyż jest ono podstawą teoretyczną, fundamentem modelu bilansowego (dla analizy kapitału firmy *F*), naszej metodologii opracowywania/pisania bilansów (księgowych) tej firmy.

Równanie Fundamentalne mówi, że każdy bilans składa się z czterech zasadniczych rozdziałów lub wspiera się na czterech jego filarach: kapitale finansowym, kapitał materialnym, kapitale ludzkim i kapitale społecznym. Z lematu 1 oraz punktu 4 wynika, iż na drugim poziomie analizy kapitału (firmy *F*) są tylko cztery takie rozdziały/filary, a nie pięć, czy sześć. Oznacza to, że ta analiza powinna iść w głąb, a nie w szerz: na poziomie trzecim ($2^3 = 8$) mamy 8 podform kapitału, na poziomie czwartym ($2^4 = 16$) mamy 16 podpodform kapitału itd. Chociaż celem tego opracowania nie jest analiza kapitału na poziomie trzecim, to o ile zdefiniowanie tych 8 podform w skali całej gospodarki może być dzisiaj (wiosna 2009) bardzo trudne, to już zrobienie tego dla konkretnej firmy lub wąsko rozumianego sektora – stosunkowo łatwe. Inaczej mówiąc, zalecamy w tych przyszłych pracach postępowanie zgodnie z zasadą od ogółu do szczegółu.

W tym miejscu warto zauważyć, że dość powszechnie używany termin „kapitał intelektualny”, będący w naszej terminologii połączeniem kapitału ludzkiego i kapitału społecznego, nie posuwa tej analizy do przodu. Z naszych rozważań wynika, że szacownie wartości kapitału intelektualnego jest tak samo trudne jak szacowanie wartości zasobów niematerialnych [12].

Równanie Fundamentalne daje ogólną odpowiedź na pytanie sformułowane we wstępie, za co płacą inwestorzy na giełdzie kupując akcje Microsoftu lub za co zapłaciło Ministerstwo przyznając dotację budżetową na działalność statutową IBS PAN? Otóż ogólna odpowiedź na takie pytania brzmi: za kapitał finansowy, materialny, ludzki i społeczny rozważanej firmy.

Równanie Fundamentalne należy czytać tak, jak każde równanie matematyczne: niewiadome po lewej, a widome po prawej stronie znaku równości. W równaniu

$$x + y = a + b$$

x oraz *y* są niewiadomymi, natomiast *a* oraz *b* są danymi lub parametrami, których wartości, tak jak w przypadku *v(KL, t)* oraz *v(KS, t),* kiedyś będą określone/oszacowane. Służy ono do opisu (wycinka) rzeczywistości, a nie do tworzenia abstrakcyjnych bytów, np. niech *v(KF,t) = -10mln PLN* (długi), a wartości wszystkich pozostałych form kapitału niech będą równe 1 mln PLN. Co to będzie za firma?

W tym opracowaniu pojęcie „firma" traktujemy maksymalnie szeroko jako byt ekonomiczny, w którym ludzie łączą swoje wysiłki, aby zrealizować jego mniej lub bardziej precyzyjnie określony cel i w którym te wysiłki są rejestrowane przez mniej lub bardziej sformalizowany system księgowości. W naszym rozumieniu typowa rodzina jest firmą z jej mniej lub bardziej rozbudowaną listą wydatków i dochodów, ale nie jest nią Linux, społeczny ruch programistów tworzących oprogramowanie otwarte i wolne. Oznacza to, w największym skrócie, że użytkownicy nie płacą za to oprogramowanie, autorzy nie otrzymują wynagrodzenia za ich dzieła i nikt nie ponosi odpowiedzialności za ewentualne błędy w takim oprogramowaniu. Wieloletnie doświadczenia pokazują, że Linux, jako taki, stawia wysokie wymagania takim firmom jak Microsoft produkującym oprogramowanie. Inaczej mówiąc, różnica między Microsoftem a Linuksem jest taka jak między oficjalnie zarejestrowaną partią polityczną a ruchem społecznym.

Na koniec, Równanie Fundamentalne mówi, że wartości poszczególnych form kapitału wzajemnie się uzupełniają, np. wartość jednej z form może rosnąć, podczas, gdy wartość innej formy może spadać.

Jak wiemy, kapitał społeczny tworzą formalne i nieformalne relacje między pracownikami firmy *F*. Relacja jest funkcją dwóch argumentów/zmiennych, a więc aby zaistniała relacja firma *F* musi zatrudniać co najmniej dwie osoby. Możemy więc sformułować następujący

Lemat 3. W jednoosobowej firmie nie ma kapitału społecznego, tj. *v(KS,t)=0* dla każdego *t.*

Wynik ten został opublikowany w marcu 2006 [4] i wzbudził spore kontrowersje. Krytycy tego rezultatu uważali, że skoro sukces jednoosobowej firmy istotnie zależy od relacji jej właściciela z klientami, włączając w to jego/jej relacje towarzyskie, po angielsku „social", z klientami, to taka firma ma niezerowy kapitał społeczny. Błąd tego rozumowania polega na braku precyzji

w definicji kapitału społecznego. Jeżeli mówimy o kapitale społecznym firmy *F*, to formalne i nieformalne relacje, z których on się składa, muszą być wewnątrz tej firmy. Fakt, na ile właściciel tej firmy jest, mówiąc po angielsku, „social" ze swoimi klientami uwzględniamy w wycenie/szacowaniu jego/jej kapitału ludzkiego, a więc nic tu nie ginie. Należy z całą mocą podkreślić, że nie jest to zwykły zabieg techniczny przerzucenia danej pozycji bilansowej z jednego rozdziału bilansu (kapitał społeczny) do drugiego (kapitał ludzki). W analizie firmy *F* ta zaleta właściciela jest cechą pojedynczej istoty ludzkiej, a więc jest to element kapitału ludzkiego. W naszym głębokim przekonaniu firma jednoosobowa z jej zerowym kapitałem społecznym ($v(KS, t) = 0$) będzie odgrywać w ekonomii taką samą rolę jak zero stopni w skali Kelvina w fizyce. Lemat 3 jest jeszcze jednym dowodem na to, że brak precyzji myśli może być źródłem błędnych wniosków.

Zauważmy, że w powyższej analizie kapitału firmy *F* traktujemy ją, zgodnie z zasadami badań systemowych, jako system względnie odosobniony/zamknięty znajdujący się w otoczeniu klientów. Podobnie postępujemy w przypadku obliczania PKB (swego rodzaju bilansu), np. Polski, kiedy analizujemy eksport/import do lub z jego otoczenia, dodając wartość eksportu ze znakiem plus, a wartość importu ze znakiem minus do PKB, a nie powiększając Polskę o kraje, do których eksportujemy/importujemy.

Fakt, iż w firmie jednoosobowej nie ma kapitału społecznego, a więc relacji współpracy, można wykorzystać w analizie instytutów naukowo-badawczych, uczelni itp. instytucji, gdzie taka współpraca jest stosunkowo słabo rozwinięta. W pierwszym przybliżeniu takiej analizy traktujemy rozważany instytut z jego *n* pracownikami naukowymi, jako zbiór *n* niezależnych firm jednoosobowych, z których każda dąży do maksymalizacji wartości swojego kapitału ludzkiego. Zauważmy, że tak w zasadzie postępuje Ministerstwo oceniając w punktach *2n* najlepszych prac danego instytutu i potem przeliczając te punkty na dofinansowanie działalności statutowej w następnym roku. W przypadku, gdy np. jeden z dwójki autorów danej pracy nie pracuje w rozważanym instytucie, to do jego dorobku liczy się tylko 50% punktów tej pracy.

Możemy teraz odpowiedzieć bardziej dokładnie na pytanie ze wstępu: Ministerstwo w przypadku instytutów PAN ocenia i płaci/kupuje w zasadzie kapitał ludzki danego instytutu, gdyż wszystkie pozostałe oceniane elementy (uzyskane stopnie i tytuły naukowe, projekty UE itp.) odgrywają w nich zwykle marginalną rolę. Taki sposób finansowania nauki doprowadził do atomizacji środowiska naukowego i tą atomizację konsekwentnie pogłębia. Mamy sytuację, w której wielu pisze, niewielu czyta (prace kolegów), a bardzo nieliczni chcą o nich dyskutować.

6. Nowy PKB

Uogólnimy tu nasze rozważania na poziom kraju/regionu i będziemy traktować gospodarkę danego kraju jako jedno ogromne przedsiębiorstwo, megafirmę *MF,* w której ludzie pracują (wytwarzają jej kapitał materialny i finansowy), uczą się i doskonalą swoje umiejętności, rozwijając kapitał ludzki tego kraju oraz współpracują, ale także spierają się, np. w parlamencie, co składa się na jego kapitał społeczny. Najbardziej ogólną miarą ich wysiłku jest szeroko rozumiane bogactwo lub dobrobyt danego kraju, co w naszej terminologii odpowiada wartości megafirmy *MF* w chwili *t,* co będziemy oznaczać symbolem *V(MF,t).* Na podstawie Równania Fundamentalnego wiemy, iż to bogactwo jest równe sumie wartości czterech form kapitału, a mianowicie: kapitału finansowego (szeroko rozumiane zasoby finansowe, dzieła sztuki, złoto, kamienie szlachetne itp.), kapitału materialnego (szeroko rozumiane zasoby surowcowe i infrastruktura, kapitał materialny (własność) poszczególnych firm, organizacji lub osób fizycznych, patenty, oprogramowanie itp.), kapitału ludzkiego (wykształcenie, kompetencje, talent, zdrowie, energia życiowa itp. obywateli danego kraju traktowanych jako oddzielne istoty ludzkie) oraz kapitału społecznego (system polityczny i system wartości, zaufanie, sposoby rozwiązywania konfliktów, współpraca itp.), co zapiszemy jako

$$V(MF,t) = f(KF,KM,KL,KS,t) = v(KF,t) + v(KM,t) + v(KL,t) + v(KS,t) \quad \text{dla każdego } t \text{ z „wczoraj”, „dziś” oraz „jutro”} \tag{3}$$

Równanie Fundamentalne (9) mówi, że tak rozumiane bogactwo kraju jest funkcją pięciu zmiennych: czterech form kapitału i czasu oraz, że w warunkach stabilnej gospodarki, gdy popyt równoważy podaż, jest ono równe sumie wartości tych kapitałów wyznaczonych dla dowolnej chwili *t* z odpowiednio zdefiniowanych „wczoraj”, „dziś” oraz „jutro” tego kraju.

W (3) bogactwo traktujemy statycznie, jako zasób w danej chwili *t,* np. na koniec danego roku, mierzonyw jednostkach monetarnych. Będą to w przypadku Polski tysiące miliardów złotych, co jest nieporęczne w codziennym stosowaniu. Dlatego dalej będziemy analizowali tak zdefiniowane bogactwo kraju jako pojęcie dynamiczne, a mianowicie zdefiniujemy *nowyPKB(t)* (zawsze pisane we wzorach jako jeden wyraz – zmienna, bez spacji, natomiast w zwykłym tekście ze spacją, jako nowy PKB), jako względną zmianę bogactwa danego kraju w danym roku *t* w stosunku do roku poprzedniego *t-1,* co zapisujemy jako

$$nowyPKB(t) = \frac{V(MF,t) - V(MF,t-1)}{V(MF,t-1)} 100\% \tag{4}$$

Inaczej mówiąc, *nowyPKB(t)* podaje procentowy wzrost lub spadek, tj. ujemny wzrost, bogactwa w danym roku w stosunku do roku poprzedniego. Z pomocą (3) możemy go rozpisać tak

$$nowyPKB(t) = \frac{v(KF,t)-v(KF,t-1)}{V(MF,t-1)}100\% + \frac{v(KM,t)-v(KM,t-1)}{V(MF,t-1)}100\% + \\ + \frac{v(KL,t)-v(KL,t-1)}{V(MF,t-1)}100\% + \frac{v(KS,t)-v(KS,t-1)}{V(MF,t-1)}100\% \quad (5)$$

Pierwszy składnik w (5) nazwiemy wkładem kapitału finansowego (w *nowyPKB(t)*) w roku *t* i oznaczymy symbolem *R(KF,t).* Podobnie definiujemy wkład kapitału materialnego – *R(KM,t),* wkład kapitału ludzkiego –*R(KL,t)* oraz wkład kapitału społecznego **–** *R(KS,t).* Stąd

$$nowyPKB(t) = R(KF,t)+R(KM,t)+R(KL,t)+R(KS,t). \quad (6)$$

Poszczególne składniki mogą mieć w (6) różne znaki. Mówimy, że w danym kraju w roku *t* był/jest rozwój zrównoważony, jeśli wszystkie składniki w (6) są ściśle dodatnie (**>**0).

Encyklopedyczna definicja (klasycznego) PKB to „wartość dóbr i usług nowo wytworzonych na terenie kraju w ciągu roku". Bardzo często zamiast tak zdefiniowanego pojęcia używa się stopy wzrostu PKB obliczanej analogicznie jak (4). Z powyższej definicji, jak też z metodologii obliczania wynika, iż zarówno PKB, jak i stopa jego wzrostu obejmują kapitał materialny i finansowy, a nie uwzględniają kapitału ludzkiego i społecznego, co ma miejsce w definicji nowego PKB (4).

Na zakończenie omówimy trzy zasadnicze różnice między PKB a nowym PKB.

1. Metodologiczne. Nowy PKB w sposób bardziej dokładny i pełny opisuje zmiany w danym roku nie tylko w produkcji materialnej i usługach, ale również w wykształceniu, kompetencjach, talentach itp. (kapitał ludzki), jak też zmiany we wzajemnym zaufaniu, nastrojach społecznych, w rozwiązywaniu konfliktów np. politycznych itp. (kapitał społeczny). Z lematu 1 wynika, iż wszystkie te cztery formy kapitału są rozłączne, a więc ich wartości możemy sumować tak jak w (6).

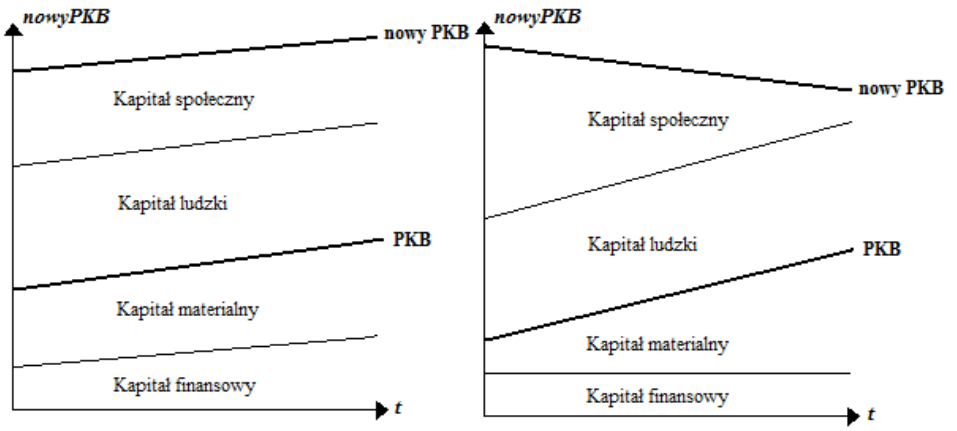
2. Techniczne. Jeden z podstawowych zarzutów w stosunku do nowego PKB brzmi: tego (*V(MF,t))* się nie da zmierzyć ani oszacować. Odpowiedź na ten zarzut wymaga oddzielnego artykułu, dlatego tu ją tylko naszkicujemy. Co 10-15 lat przeprowadza się spis powszechny, w którym między innymi bada się poziom wykształcenia, zatrudnienie itp. dane o kapitale ludzkim i społecznym, które mogą być pomocne przy szacowaniu *V(MF,T),* gdzie *T* jest rokiem spisu powszechnego. W przyszłych spisach powszechnych można będzie uwzględniać więcej takich składowych. Podobnie jak w przypadku stopy

wzrostu PKB, eksperci mogą szacować przyrost bądź spadek wkładu poszczególnych form kapitału w (6), szacując ich wartość w % w *nowyPKB(t)* w roku *t*. Wtedy bogactwo danego kraju w dowolnym roku *t>T* można obliczyć wykonując kolejno obliczenia

$$V(MF,t) = V(MF,T)\left[1 + \frac{nowyPKB(t)}{100\%}\right] \tag{7}$$

dla $t=T+1,T+2,\ldots$

3. Poznawcze. Nowy PKB pozwala głębiej wejrzeć w istotę procesów zachodzących w gospodarce danego kraju. Na rys. 6 przedstawiono, jak naszym zdaniem, zmieniał się nowy PKB w latach 2006-2007 w takich krajach jak Niemcy i Francja (rys. 6a) oraz Polska (rys. 6b). Otóż, dwa pierwsze kraje są przykładem zrównoważonego rozwoju: ich PKB rósł około 1% rocznie i w zgodnej opinii ekspertów zwiększał się w tym czasie ich kapitał społeczny i ludzki. Natomiast lata 2006-2007 w Polsce to okres tzw. IV RP, czas gdy jej kapitał społeczny był niszczony (sztucznie generowane konflikty polityczne, lustracja itp.) w tempie znacznie przekraczającym 5-6% wzrostu (klasycznego) PKB. Stawiamy tezę, iż Polska w ostatnich latach jest przykładem niezrównoważonego rozwoju. Często kapitał społeczny traktuje się jako smar w trybach gospodarki. W naszej opinii te lata są przykładem, gdy zamiast smaru sypano w te tryby piasek i efekty tego widać do dziś na każdym niemal kroku.



a) rozwój zrównoważony b) rozwój niezrównoważony

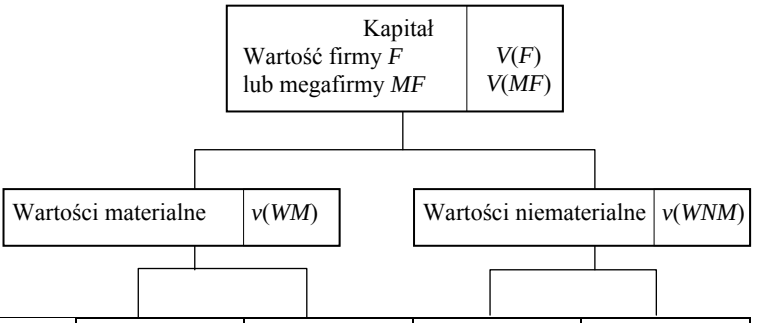
Rys. 6. Przykłady zrównoważonego i niezrównoważonego rozwoju

7. Synteza

Wykażemy, że kapitał społeczny ma wiele cech wspólnych z innymi formami kapitału (patrz tabela 1).

Tabela 1

Synteza



Lp.	Cecha	Kapitał finansowy $v(KS)$	Kapitał materialny $v(KM)$	Kapitał ludzki $v(KL)$	Kapitał społeczny $v(KS)$
1	Znak wartości	$v(KF) \gtrless 0$	$v(KM) \gtrless 0$	$v(KL) \geq 0$	$v(KS) \geq 0$
2	Efekt skali	Tak	Tak	Tak	Tak
3	Wpływ czasu	Ujemny	Ujemny	Ujemny	Ujemny
4	Liczba rodzajów	Około 20	Praktycznie ∞	Około 20	Praktycznie ∞
5	Prywatny/publiczny	W zasadzie prywatny	W zasadzie publiczny	Tylko prywatny	Tylko publiczny
6	Szybkość zmian	Ogromna, co 1 msek	Mała, zwykle raz na rok	Mała, zwykle raz na rok lub miesiąc	Mała, zwykle raz na rok lub miesiąc
7	Efekt rocznika	Brak	Brak	?	?
8	Złożoność	Mała	Średnia	Duża	Ogromna

1. Znak wartości. Kapitał finansowy może być dodatni (np. oszczędności, gotówka) lub ujemny (np. długi, zobowiązania), podobnie jak kapitał materialny (dodatni – własność budynków, ziemi, maszyn itp., ujemny – koszty utylizacji, koszty przywrócenia stanu poprzedniego wynajmowanych pomieszczeń itp.) Z wcześniejszych rozważań oraz z [7] wiemy, iż wartości zarówno kapitału społecznego, jak i kapitału ludzkiego są zawsze nieujemne.

2. Efekt skali. Większe oszczędności są zwykle wyżej oprocentowane (kapitał finansowy), większe inwestycje są efektywniejsze (kapitał materialny), podobnie jak studiowanie/nauka pokrewnych kierunków/przedmiotów (kapitał ludzki) czy też współpraca specjalistów/osób o zbliżonych zainteresowaniach (kapitał społeczny) Tak więc, we wszystkich tych formach występuje efekt skali (synergia).

3. Wpływ czasu na wszystkie formy kapitału jest, generalnie rzecz biorąc, ujemny: pieniądze (kapitał finansowy) tracą z czasem wartość, budynki, maszyny się zużywają (kapitał materialny), nieużywana wiedza, zdolności itp., zanikają (kapitał ludzki), a niepodtrzymywana współpraca w naturalny sposób obumiera (kapitał społeczny). Znane są środki zaradcze: wyższe oprocentowanie, remonty i modernizacje, kursy aktualizujące wiedzę, okresowe spotkania uaktualniające współpracę itp.

4. Liczba rodzajów. W [7] wykazano, że obecnie znanych jest około 20 produktów finansowych (kapitał finansowy), podobnie jak w przypadku rodzajów kapitału ludzkiego. Natomiast, jeżeli chodzi o kapitał materialny, jak też kapitał społeczny, to tu liczba rodzajów jest praktycznie nieskończenie wielka.

5. Prywatny/publiczny. Podpis prywatnej osoby (ministra, prezesa itp.) jest potrzebny, aby uruchomić publiczne środki (ministerstwa, firmy itp.). Dlatego kapitał finansowy jest w istocie prywatny, odwrotnie niż kapitał materialny, gdzie w zasadzie każda, nawet prywatna inwestycja musi spełniać wiele publicznych regulacji, np. w zakresie ekologii. Wiemy, że kapitał ludzki jest tylko prywatny – idzie z pracownikiem po pracy do jego domu, a kapitał społeczny tylko publiczny – trzeba co najmniej dwóch osób by zaistniała relacja, podstawowy składnik tego kapitału.

6. Szybkość zmian. Zmiany wartości kapitału finansowego mogą być rejestrowane co milisekundę dzięki współczesnym systemom komputerowym, podczas gdy zmiany w kapitale materialnym są zwykle rejestrowane raz do roku przy okazji sporządzania tzw. spisów z natury, ksiąg inwentarzowych itp. Zmiany w kapitale ludzkim i społecznym wynikają głównie z umów o pracę i porozumień o współpracy, a te zwykle są pisane na lata lub miesiące. Pojawia się tu problem odpowiedniego wygładzania/wyrównywania różnych częstotliwości rejestracji zmian poszczególnych form kapitału. Rozwiązanie tego problemu zależy od celu analizy wszystkich zasobów – inne będzie dla Microsoftu, gdzie bilanse sporządza się co kwartał, a inne dla kraju, gdzie spisy powszechne przeprowadza się co 10-15 lat. Rozwiązanie to powinno uwzględniać deprecjację/aprecjację danej waluty, tj. zmianę wartości pieniądza w czasie.

7. Efekt rocznika. Termin ten zaczerpnięto od znawców wina, którzy twierdzą, iż winogrona z pewnych roczników dają lepsze wino. Oczywiste, że efekt ten nie występuje w kapitale finansowym i materialnym. Westlund [10]

sugeruje, że niektóre roczniki, np. pokolenie 1968 roku może być zdolniejsze (kapitał ludzki) bądź też bardziej prospołecznie zorientowane (kapitał społeczny). My uważamy tę tezę za wysoce wątpliwą i dlatego w odpowiednich rubrykach tabeli 1 postawiliśmy znaki zapytania.

8. Złożoność. Z naszych rozważań wynika, iż kapitał finansowy jest najprostszą formą kapitału, jeżeli chodzi o skomplikowanie procedur wyznaczania jego wartości *v(KF,t),* a kapitał społeczny najbardziej złożoną formą. Wartość *v(KF,t)* można wyznaczyć dla każdego *t* z „wczoraj", „dziś" oraz „jutro" korzystając ze znanych technik deprecjacji/aprecjacji wartości danej waluty w czasie, przeliczając inne waluty na pieniądze przyjęte do obliczeń i sumując ze znakiem plus poszczególne elementy kapitału finansowego, gdy jest to gotówka, oszczędności, akcje itp., a ze znakiem minus, gdy są to długi, zobowiązania itp. Jeżeli chodzi o kapitał materialny, to znane techniki amortyzacji i kosztorysowania pozwalają wyznaczyć/oszacować *v(KM,t)* dla każdego *t* z „wczoraj", „dziś" oraz „jutro". Szacując wartość kapitału ludzkiego danej firmy *F* możemy przyjąć w pierwszym przybliżeniu, że *v(KL,t)* ma co najmniej tyle składników, ile jest pracowników *F* i że miarą, być może bardzo zgrubną, kapitału ludzkiego danego pracownika jest jego wynagrodzenie. Szacowanie wartości kapitału społecznego rozważanej firmy jest najbardziej złożone i skomplikowane, gdyż chcemy tu oszacować wartość grup/zespołów/projektów. Jest oczywiste, że może być pracownik, który równocześnie uczestniczy w kilku projektach. Analiza projektów z „wczoraj" (ich budżetów, zysków/strat jakie one przyniosły itp.) może być pomocna przy wyznaczaniu *v(KS,t)* dla każdego *t* z „wczoraj", „dziś" oraz „jutro" .

Jak z tego widać, nasza wielopoziomowa analiza kapitału firmy *F* lub bogactwa danego kraju, tj. kapitału megafirmy *MF,* przebiega zgodnie z zasadą od ogółu do szczegółu. Inaczej mówiąc, zgłębiamy naszą wiedzę od góry do dołu. Na każdym poziomie stosujemy Zasadę Ortogonalności rozbijając rozważaną formę kapitału na dwie i tylko dwie mniej ogólne podformy, które są rozłączne/ortogonalne. Ponieważ Zasada Ortogonalności w zakresie jej stosowania jest prawem obiektywnym, tak jak prawo powszechnego ciążenia w fizyce, to wspomniane rozbicie (1) ma formę, mówiąc obrazowo, następującą: stosunkowo „mały", tj. zawierający mało elementów, zbiór *A* oraz „olbrzymi" zbiór *B* zwierający resztę, tj dopełnienie zbioru *A* w zbiorze *X* , który zawiera, praktycznie rzecz biorąc, nieskończenie wiele elementów.

Podobne postępowanie możemy zastosować w projektowaniu badań ankietowych: zawsze pytania porządkujemy zgodnie z regułą od ogółu do szczegółu każde pytanie staramy się sformułować tak, aby odpowiedzi na nie były w maksymalnym stopniu rozłączne.

Zakończenie

W naszych badaniach wychodzimy z założenia, iż wszystko na tym świecie ma swoją wartość (cenę), określaną za pomocą modelu inwestycyjnego „dziś" na podstawie informacji z „wczoraj", po to, aby odnieść sukces „jutro". Chociaż nikt rozsądnie myślący w dającej się przewidzieć rzeczywistości nie będzie proponował kupna/sprzedaży uczelni/szkoły, partii politycznej czy kraju/regionu, to znajomość *V(F,t)* czy *V(MF,t)* jako rezultatu tych umownych lub rzeczywistych transakcji jest niezwykle cenna i potrafi już dziś w sposób ścisły objaśnić fakty, które wielu z nas czuje „przez skórę" (atomizacja badań naukowych, niezrównoważony rozwój tzw. IV RP, gwałtowny wzrost wartości/znaczenia kapitału społecznego w dobie Internetu itp.).

We wstępie sformułowaliśmy pytanie, za co się płaci w tych umownych bądź rzeczywistych transakcjach? Nasza odpowiedź brzmi: płaci się za cztery i tylko cztery kapitały: finansowy, materialny, ludzki oraz społeczny, które dzięki Zasadzie Ortogonalności tworzą piękną, wielopoziomową i zrównoważoną strukturę, przedstawioną w tabeli 1.

Nasze badania pozwalają nie tylko zdefiniować zrównoważony rozwój, ale również precyzyjnie określić, co należy rozumieć pod terminem „gospodarka oparta na wiedzy" lub „nowa ekonomia". Otóż, w gospodarce opartej na wiedzy przyrost wartości niematerialnych przynajmniej przez kilka lat wyprzedza przyrost wartości materialnych i, co więcej, wśród tych wartości niematerialnych przyrost wartości kapitału społecznego jest większy niż przyrost wartości kapitału ludzkiego[2]. Zauważmy, że o ile zasoby materialne naszego świata są ograniczone, gdyż z oczywistych względów ograniczona jest ilość pieniądza na rynku (kapitał finansowy), jak też ograniczone są wszystkie rodzaje kapitału materialnego (surowce, infrastruktura, łącznie z liczbą stosowanych w praktyce programów i patentów), o tyle rozwój zarówno kapitału ludzkiego jak i społecznego jest nieograniczony. Internet, w ogólności rozwój telekomunikacji, te możliwości rozwojowe znakomicie powiększa, i to zarówno gdy chodzi o pojedynczą istotę ludzką, jak też o zespół zgranych, rozumiejących się w pół słowa specjalistów.

W tej pracy pokazaliśmy, jak wykorzystać Zasadę Ortogonalności w analizie wartości niematerialnych szeroko rozumianej firmy lub kraju traktowanych w sposób statyczny jako zasób, którego wartość *v(WNM,t)* lub *v(KL,t)* czy też *v(KS,t)* chcemy określić, zdefiniować i oszacować. W 2006 roku zaproponowaliśmy Wirtualną Taśmę Produkcyjną (WTP), jako naturalne rozwinięcie powszechnie znanej idei Henryego Forda, która może być wykorzystana do

[2] Tezę tę opublikowano po raz pierwszy w kwietniu 2008 w Raporcie Projektu Zintegrowanego 6. Programu Ramowego UE o kryptonimie EURODITE [9].

analizy kapitału ludzkiego i społecznego traktowanych w sposób dynamiczny, jako pewne procesy przebiegające w czasie [4; 5]. Szczególnie interesuje nas rola jaką odgrywa kapitał społeczny i kapitał ludzki w szeroko rozumianych procesach twórczych. Oryginalne wykorzystanie WTP do analizy kapitału ludzkiego w edukacji opisano w [11]. W naszym głębokim przekonaniu Zasada Ortogonalności i Wirtualna Taśma Produkcyjna stanowią łącznie trwały fundament, podstawę podstaw, naszych przyszłych badań nad kapitałem społecznym i ludzkim zarówno w szeroko rozumianej firmie, jak też w kraju czy regionie.

Literatura

1. Edvinsson L. and Malone M.S. (2001). Kapitał intelektualny. Poznaj prawdziwą wartość swojego przedsiębiorstwa odnajdując jego ukryte korzenie. Wydawnictwo Naukowe PWN, Warszawa.
2. Li P.P. (2007). Social Tie, Social Capital, and Social Behavior: Toward an Integrative Model of Informal Exchange. Asia Pacific J. Management, 24, 227-246.
3. Pawar M. (2006). „Social Capital"? The Social Science Journal, 43, 211-226.
4. Walukiewicz S. (2006a). Systems Analysis of Social Capital at the firm Level. Working Paper WP-1-2006, Systems Research Institute, Warsaw.
5. Walukiewicz S. (2006b). Trzy modele do analizy kapitału społecznego. W: BOS 2006 – wiedza systemowa. Red. J. Stachowicz, A. Straszak, S. Walukiewicz. IBS PAN, Warszawa, 25-40.
6. Walukiewicz S. (2007). Four Forms of Capita and Proximity. Working Paper WP-3-2007, Systems Research Institute, Warsaw.
7. Walukiewicz S. (2008a). Piękno liczby cztery (w naukach społecznych). Working Paper WP-2-2008. Instytut Badań Systemowych PAN, Warszawa.
8. Walukiewicz S. (2008b). The Dimensionality of Capital and Proximity. Proceedings of ERSA 2008, Liverpool.
9. Walukiewicz S. and Dewalska-Opitek A. (2008). WP5: Regional Case Studies – SILESIA. Working Paper WP-1-2008, Systems Research Institute, Warsaw.
10. Westlund H. (2006). Social Capital in the Knowledge Economy. Sprinter-Verlag, Berlin, Heidelberg.
11. Wiktorzak A.A. (2009). Analiza systemowa i obliczenia inteligentne w modelowaniu kapitału ludzkiego i społecznego na przykładzie szkoły ponadgimnazjalnej. IBS PAN, Warszawa.
12. Young G.K. (2007). A Survey on Intangible Capital. CEI Working Paper Series, No. 2007-10. Hitotsubashi Univwersity, Tokyo, Japan.

RYZYKO W ORGANIZACJACH I PRZEDSIĘBIORSTWACH

Marcin Anholcer

Uniwersytet Ekonomiczny w Poznaniu

WPŁYW POWIĄZAŃ SPOŁECZNYCH NA PREFERENCJE KONSUMENTÓW

1. Sieci społeczne

Sieć społeczna to formalny model zbioru osób („agentów", „aktorów") i powiązań między nimi w języku teorii grafów. Osoby stanowią przy tym wierzchołki grafu, powiązania zaś modelowane są za pomocą krawędzi.

Można wyróżnić kilka typów sieci społecznych. Innym rodzajem jest zbiór użytkowników komunikatora internetowego, innym zaś zbiór absolwentów szkół średnich. W tym pierwszym przypadku relacja (czyli występowanie krawędzi między dwiema osobami) oznacza obecność każdej z osób na liście kontaktów drugiego użytkownika. W drugim przypadku relacja może oznaczać na przykład uczęszczanie w przeszłości do tej samej szkoły lub klasy. W dalszej części pracy interesować nas będzie sieć, w której relacja odpowiada znajomości dwóch osób.

Badania nad sieciami społecznymi zapoczątkowane zostały w dwudziestym wieku (np. [4; 5]). Pojęcie to stało się jednak bardziej popularne w ostatnich latach, kiedy zwiększająca się nieustannie moc obliczeniowa komputerów, a przede wszystkim powstanie i rozbudowa olbrzymich baz danych pozwoliły na rozpoczęcie mniej lub bardziej dogłębnej ich analizy.

Każdą sieć można scharakteryzować za pomocą dwóch parametrów (np. [5; 6]). Pierwszy z nich to średnia długość ścieżki. Załóżmy, że osoba X zna osobę Y. Mówimy wtedy, że długość ścieżki między nimi wynosi 1. Jeżeli osoba Z zna osobę Y, ale X i Z nie znają się osobiście, odległość między X i Z wynosi 2. Postępując analogicznie, mierzymy odległość między każdą parą osób w danej sieci społecznej i obliczamy ich średnią, otrzymując w ten sposób pożądaną wartość. Drugi parametr w literaturze anglosaskiej określany jest jako „clustering". Jest on równy prawdopodobieństwu, że jeśli X zna Y i Y zna Z, to również X zna Z. Aby go wyznaczyć, należy zbadać wszystkie trójki X, Y i Z takie, że X zna Y i Y zna Z i podzielić liczbę wszystkich relacji X – Z przez liczbę trójek.

Parametry te są niezależne, a ich wartości w rzeczywistym świecie wynoszą odpowiednio [6] około 6 i około 0.75. W szczególności, wartość pierwszego z parametrów dla społeczności użytkowników komunikatora Gadu-Gadu w sierpniu 2008 roku wynosiła 5,78 [3]. Co istotne, zachowanie się złożonych systemów związanych z siecią zależy przeważnie tylko od wartości tych dwóch parametrów, nie zaś od ilości wierzchołków [1; 6].

2. Systemy agentowe

System agentowy to model symulacyjny, w którym główną rolę odgrywają agenty. Każdy agent reprezentuje osobę (obiekt) lub grupę osób (zbiór obiektów) i scharakteryzowany jest poprzez zestaw prostych reguł zachowań. Łączne działanie zbioru agentów, pomimo ich prostoty, pozwala dość dokładnie modelować zachowanie się złożonych systemów.

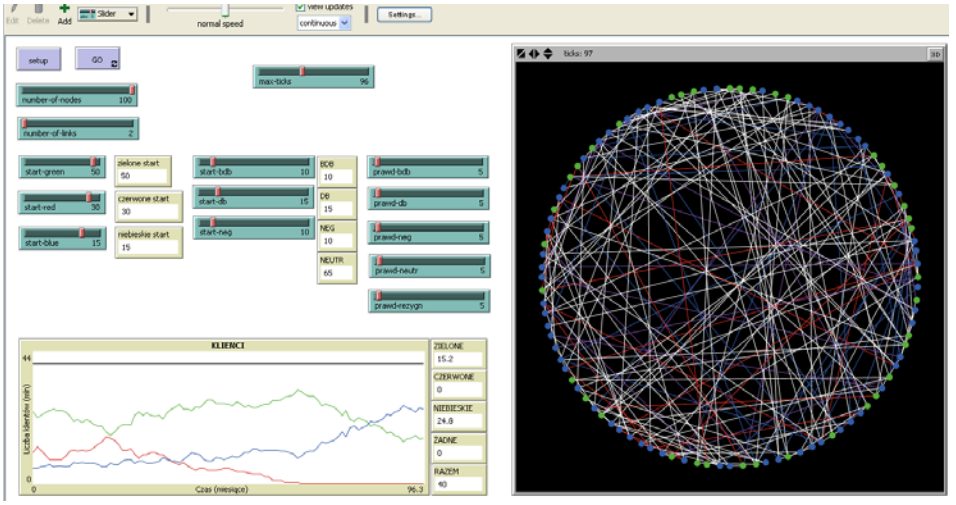
Modele agentowe używane są między innymi do badania własności złożonych systemów (w tym sieci społecznych), optymalizacji (jako przykład mogą posłużyć algorytmy mrówkowe [2] i prognozowania. Ze względu na charakter prowadzonych badań, interesuje nas przede wszystkim ta trzecia grupa zastosowań.

Systemy agentowe działają najczęściej w turach (tzw. model synchroniczny). W każdej turze każdy agent sprawdza stan swojego otoczenia (na przykład określone cechy agentów, które są z nim połączone) i na tej podstawie podejmuje działania zmieniające jego własne parametry. Przykładem mógłby być system symulujący zakupy, oparty na efekcie naśladownictwa. W każdej turze agent „sprawdza" swoich „znajomych" i z określonym prawdopodobieństwem zmienia markę kupowanych przez siebie produktów na markę produktów kupowanych przez któregoś ze „znajomych". Okazuje się, taki prosty system dobrze się nadaje do symulowania zakupów niektórych dóbr (jak na przykład zakup abonamentu u określonego operatora sieci telefonii komórkowej).

3. System prognozujący preferencje konsumentów

Biorąc pod uwagę powyższe, postanowiono skonstruować system agentowy modelujący i prognozujący zachowanie się konsumentów. Celem było skonstruowanie modelu jak najogólniejszego, który nadawałby się do prognozowania preferencji (i w konsekwencji wielkości zakupów) dowolnego produktu lub usługi.

Pierwszym krokiem było opracowanie prototypu systemu. Założono, że na decyzje konsumentów wpływ mają zarówno cechy samych produktów lub usług, jak i preferencje innych członków sieci społecznej. Po wstępnej analizie uznano, że z punktu widzenia prowadzonych badań konieczna jest rozszerzona charakterystyka sieci, zawierająca nie tylko informacje o zachodzących relacjach, ale również o ich sile. Prototyp systemu został zaprojektowany i zaimplementowany w środowisku NetLogo [7]. Przykładowy wygląd jego okna przedstawia rys. 1.



Rys. 1. Przykładowy wygląd okna prototypu systemu agentowego

Źródło: Opracowanie własne przy użyciu edytora NetLogo.

Biorąc pod uwagę przytoczone założenia, w drugim etapie badań wstępnie rozpoznano dziedziny, w przypadku których istotne jest zróżnicowanie ze względu na stopień znajomości. W tym celu na początku 2008 roku przeprowadzono ankietę wśród studentów jednej ze specjalności Wydziału Zarządzania Akademii Ekonomicznej w Poznaniu. Bardziej szczegółowe informacje na ten temat znajdują się w kolejnym podrozdziale.

Kolejny etap projektu (aktualnie w fazie przygotowań) obejmuje przeprowadzenie znacznie szerszego badania, jednak tylko dla określonych dziedzin (tych, które dały „pozytywny" wynik po pierwszym, pilotażowym, badaniu ankietowym). Badania te pozwolą dokładniej oszacować podstawowe parametry modelu, przede wszystkim prawdopodobieństwa przejścia.

Ostatnim krokiem będzie wprowadzenie ewentualnych poprawek i rozszerzeń do prototypu systemu i postawienie prognoz rozwoju wybranych gałęzi gospodarki na podstawie jego działania.

4. Badania empiryczne

W styczniu 2008 roku studenci jednej ze specjalności Wydziału Zarządzania Akademii Ekonomicznej w Poznaniu wypełnili ankietę na temat swoich preferencji. Składała się ona z dwóch części.

W pierwszej każdy ze studentów widział pełną listę nazwisk osób ze swojej specjalności. Przy każdym nazwisku musiał określić stopień znajomości w kilkustopniowej skali (od „nie znam", poprzez „spotykamy się wyłącznie na zajęciach" aż do „znamy się bardzo dobrze") oraz swój stosunek do danej osoby (negatywny, neutralny lub pozytywny, również w kilkustopniowej skali).

Druga część składała się z 383 pytań podzielonych na pięć działów: „finanse i ubezpieczenia", „komunikacja", „rozrywka", „telekomunikacja i informatyka" i „życie codzienne". Pytania dotyczyły zarówno aktualnych decyzji, jak i ich ewentualnych zmian w najbliższej przyszłości (np. pytanie o posiadanie pralki towarzyszyło pytanie o zamiar jej zakupu). Dodatkowo każda osoba wypełniła krótką „metrykę".

W przypadku części studentów ankieta została również wypełniona przez jedną lub dwie dodatkowe osoby spoza specjalności: wybranego członka rodziny lub wybranego przyjaciela. Te osoby wypełniały jednak tylko drugą część ankiety (każda z nich została automatycznie połączona z odpowiednim studentem relacją typu „pozytywna").

Ostatecznie udało się zebrać informacje na temat preferencji i ponad 3000 relacji od prawie 200 osób.

Wyniki zostały następnie przeanalizowane pod kątem występowania zależności między charakterem znajomości między dwiema osobami, a zbieżnością ich preferencji zakupowych.

Z oczywistych względów nie można zaprezentować tutaj pełnych wyników analizy. Niżej przedstawiono jednak bardziej szczegółowo kilka przykładów oraz bardziej ogólne wnioski wynikające z badań. Dla każdego z omówionych pytań pary odpowiedzi pogrupowano ze względu na dwie cechy: stopień zażyłości (-1 oznacza stosunek negatywny, 0 neutralny, zaś 1 i 2 pozytywny, przy czym 2 oznacza częste kontakty, 1 zaś rzadsze) i zbieżność odpowiedzi (taka sam albo inna).

Należy podkreślić, że celem badań było znalezienie pytań, na które pary osób cechujące się różnym stopniem zażyłości odpowiadały tak samo z różnymi częstościami. Nie były natomiast obiektem moich zainteresowań bezwzględne wartości omawianych częstości. Te wartości będzie można dokładnie oszacować dopiero po przeprowadzeniu bardziej szczegółowych badań dla większej i bardziej zróżnicowanej grupy osób (wspomniany wcześniej trzeci etap projektu).

Zależności zostały przetestowane statystycznie za pomocą testu niezależności χ^2.

Wyniki badań świadczą o tym, że stopień znajomości raczej nie wpływa na decyzje związane z finansami. Jedynym wyjątkiem były odpowiedzi na pytania odnośnie rachunków oszczędnościowo – rozliczeniowych w niektórych bankach.

W grupie pytań związanych z komunikacją, w większości przypadków stopień zbieżności odpowiedzi zależy istotnie od stopnia znajomości. Jako przykład posłużyć może odpowiedź na pytanie o sposób poruszania się po swojej miejscowości (tabela 1). Wyraźnie widać zależność między stopniem znajomości a odsetkiem zbieżnych odpowiedzi: w przypadku negatywnego stosunku odsetek zbieżnych odpowiedzi wynosi 4%, w przypadku neutralnego stosunku lub słabej znajomości przy pozytywnym stosunku około 25%, zaś w przypadku dobrej znajomości – 45%.

Tabela 1

Zbieżność odpowiedzi na pytanie: „Jakim środkiem komunikacji najczęściej poruszasz się po swojej miejscowości"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,96	0,77	0,75	0,55	2687
TA SAMA ODPOWIEDŹ	0,04	0,23	0,25	0,45	892
SUMA	46	2078	1258	197	3579

$\chi^2 = 57{,}04$; $p = 0{,}00000$

Z podobną sytuacją mamy do czynienia w przypadku pytania o sposób poruszania się poza rodzinną miejscowością. Tym razem odsetek zbieżnych odpowiedzi waha się mniej, między 28% a 40%, jednak zależność pozostaje statystycznie istotna. Wyniki przedstawione zostały w tabeli 2.

Tabela 2

Zbieżność odpowiedzi na pytanie: „Jakim środkiem komunikacji najczęściej poruszasz się poza swoją miejscowością"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,72	0,65	0,69	0,60	2377
TA SAMA ODPOWIEDŹ	0,28	0,35	0,31	0,40	1202
SUMA	46	2078	1258	197	3579

$\chi^2 = 10{,}93$; $p = 0{,}01210$

W grupie pytań dotyczących rozrywki jedynie odpowiedzi na pytania dotyczące ulubionego sportu okazały się zależeć od charakteru znajomości (wyniki zawiera tabela 3). Zbieżność odpowiedzi na pytania związane z preferencjami odnośnie filmów, czy książek nie zależały od częstości kontaktów i stosunku do drugiej osoby.

Tabela 3

Zbieżność odpowiedzi na pytanie: „Jaki sport lubisz uprawiać najbardziej"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	1,00	0,92	0,89	0,80	2626
TA SAMA ODPOWIEDŹ	0,00	0,08	0,11	0,20	275
SUMA	34	1636	1062	169	2901

$\chi^2 = 32{,}81$; $p = 0{,}00000$

Odpowiedzi na pytania z działu „telekomunikacja i informatyka" w większości istotnie zależały od charakteru znajomości. Przykładowo, zbieżność odpowiedzi na pytanie o rodzaj dostępu do komputera dało się zaobserwować u 46% par znających się dobrze i tylko u 11% par których stosunek do siebie był negatywny (wyniki zawiera tabela 4).

Tabela 4

Zbieżność odpowiedzi na pytanie: „Do jakiego komputera masz dostęp"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,89	0,73	0,78	0,54	2452
TA SAMA ODPOWIEDŹ	0,11	0,27	0,22	0,46	876
SUMA	45	1889	1205	189	3328

$\chi^2 = 53{,}52$; $p = 0{,}00000$

Z podobnym zróżnicowaniem mamy do czynienia w przypadku odpowiedzi na pytanie o dostęp do Internetu (tabela 5).

Tabela 5

Zbieżność odpowiedzi na pytanie: „Jaki masz dostęp do Internetu"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,87	0,64	0,67	0,53	2148
TA SAMA ODPOWIEDŹ	0,13	0,36	0,33	0,47	1180
SUMA	45	1889	1205	189	3328

$\chi^2 = 24{,}30$; $p = 0{,}00002$

Nieco słabsza, ale również statystycznie istotna zależność cechuje odpowiedzi na pytania o markę telefonu komórkowego i operatora, z którego usług korzystają respondenci. Dane na temat tego ostatniego pytania zawiera tabela 6.

Tabela 6

Zbieżność odpowiedzi na pytanie: „Z usług którego operatora telefonii komórkowej korzystasz na codzień"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,98	0,88	0,86	0,74	2880
TA SAMA ODPOWIEDŹ	0,02	0,12	0,14	0,26	448
SUMA	45	1889	1205	189	3328

$\chi^2 = 53{,}52$; $p = 0{,}00000$

Stopień znajomości nie ma wpływu na zbieżność odpowiedzi odnośnie do najważniejszej cechy telefonu komórkowego, jest natomiast istotny w przypadku szczegółowej oceny ważności poszczególnych cech aparatu. Istotne jest zróżnicowanie odpowiedzi na pytania o najważniejsze cechy operatora.

Zróżnicowana jest również zgodność odpowiedzi na pytania dotyczące najważniejszych cech operatora telefonii stacjonarnej, mimo że sam wybór operatora nie zależy od preferencji znajomych.

W grupie pytań dotyczących życia codziennego, ocena cech pojazdu nie zależy od preferencji znajomych. Są jednak od nich zależne preferencje odnośnie do sposobu dokonywania zakupów spożywczych oraz odzieży i obuwia, a także wielkość wydawanych kwot. Za przykład mogą posłużyć dane na temat zbieżności odpowiedzi na pytanie odnośnie do sposobu dokonywania zakupów odzieży i obuwia zawarte w tabeli 7.

Tabela 7

Zbieżność odpowiedzi na pytanie: „W jaki sposób najczęściej kupujesz odzież i obuwie"?

	-1	0	1	2	SUMA
INNA ODPOWIEDŹ	0,93	0,71	0,76	0,61	2654
TA SAMA ODPOWIEDŹ	0,07	0,29	0,24	0,39	1015
SUMA	45	2119	1290	215	3669

$\chi^2 = 33{,}82$; $p = 0{,}00000$

Istotne okazały się również różnice w przypadku pytań na temat posiadania i chęci zakupu niektórych sprzętów gospodarstwa domowego (np. telewizora).

Podsumowanie

W opracowaniu przedstawiono wyniki badań nad zależnością preferencji konsumentów od charakteru powiązań interpersonalnych. Badania te przedstawione zostały w szerszym kontekście, jako wstęp do większego projektu, mającego na celu skonstruowanie symulacyjnego narzędzia umożliwiającego prognozowanie obrotów wybranych gałęzi gospodarki.

Wyniki badań wskazują na fakt, iż rozbudowa obecnie stosowanych modeli o parametry opisujące stopień znajomości ma sens głównie w przypadku preferencji odnośnie do usług telekomunikacyjnych i technologii informatycznych. Charakter znajomości ma również wpływ na preferencje odnośnie do zakupu artykułów spożywczych, odzieży i obuwia oraz wyboru środków komunikacji. Dalsze badania będą się więc koncentrować wokół tych obszarów. Ich sprecyzowanie umożliwi przeprowadzenie bardziej skonkretyzowanych badań na większej grupie osób i w konsekwencji dokładną kalibrację parametrów systemu agentowego dla różnych działów gospodarki. Ostatecznie pozwoli to na postawienie prognoz, które powinny być lepsze od uzyskiwanych na podstawie obecnie stosowanych modeli.

Literatura

1. Colbaugh R., Glass K. (2006). Identifying Facility Function using Limited Observations. Journal of Intelligence Community Research and Development.
2. Dorigo M., Stützle T. (2004). Ant Colony Optimization. MIT Press.
3. http://www.internetstandard.pl/news/162232/Gadu.Gadu.dowolna.osoba.w.Polsce.nie.7.a.6.krokow.od.ciebie.html (29.03.2009).
4. Mitchell J.C. (1969). Social Networks in Urban Situations: Analysis of Personal Relationships in Central African Towns. Manchester University Press, Manchester.
5. Scott J. (2000). Social Networks Analysis. A Handbook. Sage Publications, London-Thousand Oaks-New Delhi.
6. Strogatz S. (2003). SYNC. The Emerging Science of Spontaneous Order. Hyperion, New York.
7. Wilensky U. (1999). NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

Stefan Grzesiak
Uniwersytet Szczeciński

KILKA UWAG O RYZYKU W SEKTORZE OCHRONY ZDROWIA

Wprowadzenie

Podejmowanie i akceptowanie ryzyka to nieodłączny element działalności gospodarczej. Dyskusja nad jego identyfikacją i szacowaniem prowadzona jest od dawna w ekonomicznej literaturze światowej i polskiej [1; 2; 5; 6; 8; 10; 12]. Dyskusje i spory wywołuje kwestia oceny stopnia niepewności przy podejmowaniu decyzji w konkretnych uwarunkowaniach gospodarczych. Dlatego o ryzyku będziemy mówić, gdy po podjęciu wielu decyzji:

- rezultat, jaki będzie osiągnięty w przyszłości, nie jest znany, ale możliwe jest zidentyfikowanie przyszłych sytuacji,
- znane jest prawdopodobieństwo zrealizowania się poszczególnych możliwości w przyszłości [10, s. 12].

Wydaje się, że w odniesieniu do dyskutowanego w opracowaniu tematu częściej będzie miała miejsce pierwsza sytuacja. Mówiąc o przejawach ryzyka mamy na myśli przede wszystkim oczekiwane w przyszłości skutki tak dla sytuacji ekonomicznej podmiotów prowadzących działalność gospodarczą, jak i innych w jakikolwiek sposób od nich uzależnionych. Jeżeli akceptujemy przekonanie F.H. Knighta [10], że ryzyko to niepewność wymierna, a więc mogąca podlegać kwantyfikacji, to oczekujemy od ekonomistów co najmniej umiejętności jego oszacowania.

Występowanie ryzyka wiąże się ze wszystkimi procesami ekonomicznymi, generującymi niepewność przyszłych skutków. Jednak dociekania ekonomistów odwołujące się do występowania, przyczyn i skutków uwzględniania czy pomijania ryzyka w działalności gospodarczej związane są w pierwszym rzędzie z rynkiem kapitałowym, ubezpieczeniami oraz bankowością. W tych segmentach gospodarki ryzyko jest obecne na codzień, a procedury jego rozpoznawania, szacowania i oceny są najbardziej rozwinięte. Powstało wiele prac naukowych, dogłębnie analizujących przejawy i skutki oceny ryzyka na rynku kapitałowym, ubezpieczeń i w bankowości [4; 7; 9; 10; 12]. Rynek ubezpieczeń kształtował się od początku pod wpływem obiektywnej konieczności dywersy-

fikacji ryzyka wystąpienia wielu niespodziewanych zjawisk, gwałtownych i dotkliwych w skutkach dla branży ubezpieczeniowej. Konieczność oceny wystąpienia i skali ryzyka ujawniła się w też działalności bankowej, gdzie istnieje naturalna możliwość pojawienia się obiektywnych i subiektywnych trudności w kontaktach banków jako wierzycieli z dłużnikami. Wydawałoby się, że tym samym suma wieloletnich zebranych doświadczeń powinna być pomocna w opanowaniu i przezwyciężeniu kolejnych recesji i kryzysów. Doświadczenie z przeszłości oraz obecne wskazuje, że jednak nie zapobiegło to występowaniu turbulencji w gospodarce światowej. Wiele wskazuje, że być może właśnie niewłaściwe wnioski z przeszłości połączone z długoletnim forsowaniem rozwiązań poprawnych politycznie, ale ułomnych ekonomicznie pomogły je wywoływać.

1. Organizacja rynku usług medycznych a ryzyko

W niektórych sferach działalności gospodarczej i społecznej kwestia występowania i postrzegania przejawów ryzyka jest wyjątkowo ważna i jednocześnie trudna. Chodzi o takie segmenty rzeczywistości społeczno-ekonomicznej, w których określenie i wycena ryzyka stwarza trudności nie tylko rachunkowe, ale i merytoryczne. Tam, gdzie ryzyko kojarzy się z zagrożeniem zdrowia, a nawet życia, precyzyjny, ale i trudny do akceptacji wywód na temat kosztów, możliwości i trudności nie jest łatwo poddać chłodnej analizie.

Niewątpliwie do takich należy zestaw problemów obejmujących uwarunkowania, decyzje i działania podejmowane w sektorze ochrony zdrowia. Sektor ten ma swoją specyfikę, gdyż występują tam obok siebie i niemal jednocześnie sytuacje typowo rynkowe, jak również zjawiska i procesy o charakterze społeczno-moralnym, niedające się podporządkować zwykłym regułom równowagi popytu i podaży na rynku towarów i usług. Ta swoista mieszanka prowadzi do fundamentalnych trudności z określeniem charakteru i skali ryzyka w odniesieniu do zjawisk i instytucji tworzących rynek usług medycznych i działających na nim. Ciekawe spostrzeżenia i informacje na ten temat, chociaż poczynione z typowo amerykańskiego punktu widzenia, można znaleźć w pracy T. Getzena [2].

Należy wyjaśnić przy tym, że w opracowaniu nie wspomina się o zarządzaniu ryzykiem w ochronie zdrowia w kontekście wyboru metod leczenia pacjentów, procedur bezpieczeństwa w szpitalach, zagospodarowania odpadów itp. Te bardzo ważne i istotne dla zapewnienia jakości leczenia działania i procedury nie są tematem opracowania, koncentrujemy się jedynie na aspektach ekonomicznych.

W związku z tym przedmiotem zainteresowania pozostaje tu:
- próba określenia, w jakich miejscach i w jaki sposób można zidentyfikować ryzyko w świadczeniu usług medycznych,
- którzy z uczestników rynku usług medycznych ponoszą największe ryzyko ekonomiczne a którzy z tego korzystają,
- jak należy wiązać wycenę poszczególnych usług z możliwym ryzykiem niepowodzenia pomimo zastosowania właściwych procedur.

Należy na początku przypomnieć, że na rynku świadczeń medycznych występują trzy grupy podmiotów:
- świadczeniodawcy (szpitale, sanatoria, przychodnie, apteki, prywatna praktyka lekarska itp.), reprezentujący podaż usług,
- świadczeniobiorcy (potencjalni pacjenci) jako strona popytowa,
- tzw. trzecia strona, czyli płatnicy za usługi (NFZ – ze składek pracowników i pracodawców, rząd – środki budżetowe, ubezpieczyciele, samorządy, prywatni pacjenci) oraz instytucje regulujące i kontrolujące funkcjonowanie tego rynku.

Każda z tych grup w inny sposób postrzega w dzisiejszych realiach ryzyko w sektorze ochrony zdrowia. Dla dostawców usług i produktów medycznych ma ono charakter rynkowy, jest to więc przede wszystkim kwestia ilości i opłacalności wykonywanych procedur medycznych czy rozmiarów sprzedaży. Pacjenci widzą w nim sytuację osobistą, intymną, a niekiedy osobistą tragedię, stąd ich zachowania są dalekie od normalnych reakcji rynkowych. Dla płatników instytucjonalnych najważniejszy jest bilans wpływów i wydatków, za który są odpowiedzialni, czy to przed udziałowcami czy też przed całym społeczeństwem. Zachowania na rynku utrzymania i poprawy zdrowia przypominają więc grę pomiędzy wymienionymi podmiotami, przy czym ogólnymi regułami gry i wyborem strategii w danym kraju w zależności od wyboru modelu działalności sektora kieruje państwo.

Specyfika rynku medycznego polega również na tym, że potencjalny popyt na usługi w zakresie poprawy zdrowia nie jest w żaden sposób naturalnie ograniczony. O ile istnieją naturalne bariery możliwej skali konsumpcji dla każdego potencjalnego klienta na innych rynkach, to praktycznie każdy świadomy obywatel będzie zainteresowany poprawą stanu własnego zdrowia, niezależnie od samopoczucia i wieku. Podstawowym ograniczeniem, jak na każdym innym rynku, pozostaje cena usługi, która skutecznie blokuje dostęp wielu zainteresowanym poprawą samopoczucia. To oraz ograniczone możliwości płatników skłaniają do konieczności wprowadzania regulacji funkcjonowania mechanizmów dystrybucji świadczeń medycznych w zależności od przyjętego w danym kraju modelu działalności tego sektora.

Wśród rozwiązań instytucjonalnych istniejących w większości krajów świata można w praktyce wyróżnić trzy typowe modele systemów opieki zdrowotnej: ubezpieczeniowy, usługowy i rezydualny [13, s. 36]. Wykształciły się w wyniku wieloletnich doświadczeń, sporów i modyfikacji. W rzeczywistości jednak większość funkcjonujących realnie jest mieszankami tych modeli o różnej strukturze wewnętrznej. Ma to oczywisty wpływ na kwestie postrzegania i oceny ryzyka tak przez instytucje organizujące świadczenia (dostawców usług, leków i urządzeń), płatników, jak i klientów, czyli potencjalnych pacjentów.

W warunkach polskich po wprowadzanych przez ostatnie 10 lat zmianach występuje mieszany model ubezpieczeniowo-usługowy, z jednej strony finansowanie usług w przeważającej mierze opiera się na funduszach zbieranych od pracowników, z drugiej zaś przywileje konsumentów i prawo do świadczeń są analogiczne jak w modelu usługowym. Wskazuje to na niekonsekwencję przyjętych rozwiązań, gdyż w takim schemacie płacący największe składki zdrowotne mają praktycznie takie same prawa, jak osoby w ogóle ich niepłacące.

Ta zasada może być akceptowalna, jeżeli ma uzasadnienie w podobnej strukturze i wielkości dochodów w społeczeństwie. Jest jednak mocno krytykowana w warunkach rozwarstwienia ekonomicznego, ponieważ pojawia się argument o braku związku między poziomem dostępu do świadczeń poszczególnych osób a wysokością płaconych przez nich składek. Szczególne zastrzeżenia budzą wtedy tzw. zachowania antyzdrowotne wśród osób płacących minimalne (lub żadne) składki.

2. Ryzyko uczestników rynku usług reprezentujących stronę podażową i płatników

W sytuacji realnego zróżnicowania dostępu do świadczeń zdrowotnych ważne jest ustalenie, kto (w sensie nie tylko osób, ale i instytucji) w warunkach polskich w największym stopniu jest narażony na ryzyko, a kto korzysta z ponoszonego przez innych ryzyka. W tym celu przeprowadzono identyfikację oraz określono rolę podmiotów na rynku usług medycznych. Wśród najważniejszych kreujących ten rynek można wymienić:

- instytucje organizujące ochronę zdrowia – szpitale, przychodnie, laboratoria, sanatoria, władze centralne i samorządowe,
- personel zatrudniony w instytucjach ochrony zdrowia – lekarzy, pielęgniarki, obsługę i konserwatorów sprzętu medycznego, personel porządkowy, administrację obiektów,
- producentów i dystrybutorów leków i preparatów laboratoryjnych,
- dostawców sprzętu medycznego, materiałów zaopatrzeniowych i wyposażenia,

- dostawców żywności i mediów niezbędnych jednostkom lecznictwa,
- płatników dystrybuujących środki finansowe, przeznaczone na opłaty za realizowane procedury medyczne i dostawy materiałów oraz leków (NFZ, ubezpieczyciele, samorządy),
- wreszcie pacjentów, dających się zdywersyfikować według różnych kryteriów.

Każda z wymienionych grup w różnym stopniu bierze udział w swoistej grze na rynku i próbuje uzyskać maksimum korzyści oraz zminimalizować ryzyko związane z realizowanymi zadaniami. Podstawową i specyficzna grupą są pacjenci, którzy dążą do zachowania i poprawy zdrowia, co ma rzecz jasna charakter niekomercyjny i znacząco różniący się od pozostałych uczestników tego rynku.

Zastanówmy się, jakie są oczekiwania i jakie ryzyko ponoszą poszczególne wyżej wymienione grupy instytucji i osób.

Instytucje opieki stacjonarnej w uproszczeniu dzielą się na niekomercyjne (non profit) i komercyjne. Dla jednostek komercyjnych istnienie na rynku i ryzyko związane z profilem działalności są pochodnymi atrakcyjności oferty i konkurencyjności. Zasadnicze znaczenie dla nich mają kontrakty zawierane z NFZ, pozostałą część oferty sprzedają na rynku. Ryzyko przetrwania i nieskrępowanej działalności jednostek opieki stacjonarnej działających w systemie non profit, które są finansowane praktycznie wyłącznie przez NFZ, jest określone w zasadniczym stopniu rolą szpitala w środowisku, w jakim prowadzi działalność terapeutyczną. Procesy specjalizacji i nasycenia techniką opieki szpitalnej tworzą swoiste ograniczenia dla małych i oferujących przyjazny pobyt lecznic. Familijna atmosfera oraz troskliwość personelu nie są w stanie zastąpić nowoczesnego sprzętu i specjalistycznych procedur. Podstawowe czynniki ryzyka funkcjonowania szpitala tkwią więc w jego strukturze i jej dopasowaniu do potrzeb środowiska potencjalnych pacjentów.

W polskich warunkach (inaczej niż np. w USA) lekarze są najczęściej pracownikami szpitali i ich wynagrodzenia są uzależnione od kontraktów, jakie posiada szpital, w którym pracują. Tym samym ich interesy związane są bezpośrednio z możliwościami, warunkami oraz wyposażeniem konkretnej jednostki leczącej. Ryzyko zawodowe, ponoszone przez nich i związane z możliwością utraty pracy, ma charakter pośredni i odnosi się głównie do braku dostatecznej ilości zakontraktowanych świadczeń, mniejszej liczby pacjentów lub złej polityki kierownictwa szpitala.

Należy wziąć pod uwagę, że lekarze, w przeciwieństwie do innych grup zatrudnionych w jednostkach ochrony zdrowia, mają wyjątkowo znaczący wpływ na ich kondycję finansową. Ponieważ są jedyną grupą osób podejmują-

cych samodzielnie decyzje o sposobie i przebiegu leczenia chorych, stąd istnieje z ich strony ryzyko sugerowania i wykorzystywania takich procedur, które nie są korzystne finansowo tak dla NFZ, ubezpieczycieli, jak i dla dyrekcji szpitali czy nawet pacjentów, zaś stwarzają możliwości uzyskiwania dodatkowych profitów przez nich samych. Kontrola tego jest bardzo trudna i jedynie częściowo możliwa poprzez wnikliwe analizowanie wykonywanych procedur medycznych dla konkretnych przypadków. W stosunku do pozostałych pracowników: pielęgniarek, personelu obsługowego itp. można stwierdzić, że ich zatrudnienie i wynagrodzenia są ściśle związane z wielkością i zawartością kontraktów na leczenie, realizowanych przez szpital. Ryzyko utraty pracy ponoszone przez nich jest przede wszystkim uzależnione od kondycji finansowej szpitala.

Ryzyko ponoszone przez NFZ nie jest ryzykiem w pełnym tego słowa znaczeniu, bo jedynym zagrożeniem dla tej instytucji mogą stać się ograniczone środki finansowe, napływające ze składek płaconych przez społeczeństwo.

Zupełnie inne problemy dotyczą producentów i dostawców leków. Zasadniczy dla nich problem sprowadza się do takiej organizacji rynku, aby przez jak najdłuższy czas pozostać jedynymi dostawcami ważnych dla życia i zdrowia preparatów i nie dopuścić konkurencji do głosu. Interesy uboższych społeczeństw i ich rządów są zupełnie przeciwstawne. Jest to wielki ogólnoświatowy problem, mający zarówno wymiar ekonomiczny i biznesowy, jak i moralny. Ochrona przed ryzykiem pogorszenia własnej pozycji finansowej przejawiana jest poprzez naciski na rządy celem nie dopuszczenia na rynek porównywalnych artykułów, oferowanych przez konkurencję, oraz ograniczenia dostaw leków generycznych.

Dostawcy sprzętu i wyposażenia medycznego ponoszą ryzyko nieuregulowania należności w podobnym stopniu jak pozostali dostawcy. Główne zagrożenia dotyczą tych kontrahentów, dla których jednostki służby zdrowia są podstawowymi i strategicznymi odbiorcami. Nie mają oni praktycznie możliwości zabezpieczenia się przed stratami, jeśli ich odbiorcy mają trudności finansowe czy rezygnują z zamówionych produktów. Dla mniejszych dostawców jest to często równoznaczne z zagrożeniem bankructwem.

W stosunkowo najlepszej sytuacji są dostawcy mediów (woda, ciepło, gaz, energia itp.) i żywności. Dla nich ryzyko niewypłacalności odbiorcy oznacza zaprzestanie dostaw, co w warunkach rozproszenia nie stwarza większych komplikacji. Podobnie jest z dostawami energii czy wody, przy czym odcięcie tego typu zaopatrzenia dla chorych ma nie tylko aspekt biznesowy, ale i z oczywistych względów moralny. Taka sytuacja wymusza na obu stronach konieczność negocjacji i polubownego oraz szybkiego załatwiania sporów.

3. Ryzyko dla konsumentów usług zdrowotnych i ich zachowania

Ze zrozumiałych względów problem ryzyka braku pomocy medycznej jest najistotniejszy dla potencjalnych pacjentów – konsumentów świadczeń zdrowotnych. Ma też specyficzny, dwoisty charakter. Z jednej strony pojawia się ryzyko przeżycia i wyleczenia z określonej choroby, a z drugiej sprowadza się do poczucia zagrożenia, wynikającego z niemożności sfinansowania w danym przypadku określonych procedur medycznych czy też terapii. Praktycznie każdy będąc w pełni świadomości bierze pod uwagę w przyszłości możliwość zachorowania, kalectwa czy też innych komplikacji zdrowotnych, a z tym wiąże się ocena szansy na uzyskanie niezbędnej pomocy.

Kwestia wyboru sposobu ochrony przed ryzykiem zależy od wielu czynników, które wiążą się nie tylko z osobistą sytuacją i wiekiem potencjalnego pacjenta, ale i z poziomem gospodarczym kraju, globalnymi wydatkami na zdrowie, stylem życia, a przede wszystkim organizacją ochrony zdrowia w jego ojczyźnie.

Żaden dostępny sposób ograniczania takiego ryzyka (np. oszczędzanie, pomoc rodziny, opieka społeczna) nie ma zasadniczego wpływu na zmianę sytuacji życiowej osoby dotkniętej chorobą lub wypadkiem.

System prywatnych, indywidualnych ubezpieczeń, który tak szeroko jest rozwinięty w USA i Europie Zachodniej, pozornie mógłby zabezpieczyć niemal wszystkie potrzeby bogatszej części przyszłych pacjentów. Zasadnicze znaczenie ma tu problem szacowania i wyceny ryzyka dla ubezpieczycieli, którzy powinni stanowić jedno z zasadniczych źródeł finansowania działalności sektora ochrony zdrowia.

Ograniczenia dla beneficjentów odnoszą się przede wszystkim do konieczności wczesnego wejścia do systemu ubezpieczeniowego tak, aby płacone składki pozwoliły na finansowanie większości zabiegów. W warunkach polskich jest to ciągle rzadkość, a większość osób zamiast wcześniej przygotowywać się do trudnych i kłopotliwych sytuacji, woli ubezpieczać się dopiero w momentach zagrożenia życia, gdy najczęściej nie posiada ani czasu ani środków na sfinansowanie niezbędnego leczenia.

Ten sposób rozumowania ma swe korzenie w wieloletnim przekonaniu o bezpłatności usług medycznych, które zapewnia państwo, i w razie konieczności o skuteczności ewentualnego dodatkowego opłacenia interwencji lekarza, najczęściej niezgodnie z prawem.

Pozostaje więc korzystanie z ubezpieczeń społecznych, które są najlepszą ze znanych form dywersyfikacji i w konsekwencji zmniejszania ryzyka w przypadku choroby dla pojedynczych osób. Powszechny charakter ubezpieczeń

społecznych to oprócz niewątpliwych zalet ryzyko wystąpienia niekorzystnych efektów, podwyższających koszty świadczeń i utrudniających funkcjonowanie systemu ochrony zdrowia. Efekty te są wynikiem asymetrii informacji oraz hazardu fizycznego, moralnego i duchowego [10, s. 25].

Z asymetrią informacji spotykamy się w momencie kontaktu pacjenta z lekarzem. Zdecydowana przewaga wiedzy i autorytetu lekarza może skutkować tym, że zastosowana terapia niekoniecznie jest najwłaściwsza dla pacjenta, ale za to najbardziej opłacalna dla szpitala czy lekarza. Powszechność systemu ubezpieczeniowego i brak bodźców prowadzących do kontroli prawidłowości stosowanej terapii stwarzają warunki do akceptacji wydatkowania zbędnych środków oraz ich marnotrawstwa.

O hazardzie fizycznym można powiedzieć wtedy, gdy osoba podlegająca ubezpieczeniu podejmuje pracę w oczywisty sposób związaną z zagrożeniami dla niej z powodu niewiedzy lub godzi się na nią bądź to pod wpływem atrakcyjności wynagrodzenia, bądź też przekonania, że jej stan zdrowia gwarantuje bezpieczne wykonywanie obowiązków. Hazard moralny przejawia się poprzez cechy charakterologiczne potencjalnego pacjenta (np. lekkomyślność), które mogą prowadzić w przyszłości do wysokich kosztów opieki medycznej. Jest on w praktyce najbardziej kłopotliwy dla ubezpieczyli prywatnych, bo najtrudniej go uwzględnić w kalkulacji składki. Hazard duchowy z kolei oznacza subiektywną reakcję ubezpieczonego, która wynika ze świadomości istnienia ochrony ubezpieczeniowej. Może mieć ona charakter pozytywny, ale i często negatywny, gdyż oznacza realną możliwość obniżki osobistych wydatków na ograniczenie ryzyka zachorowania, jeśli już została zagwarantowana opieka medyczna.

Konkluzja

Przedstawione rozważania nad identyfikacją i konsekwencjami występowania ryzyka w sektorze ochronie zdrowia są z konieczności fragmentaryczne i dalece niepełne. Wymagają dalszego pogłębienia i osobnego rozważenia możliwości identyfikacji i pomiaru ryzyka dla każdej grupy osób i instytucji działających w sektorze ochrony zdrowia. Pierwszoplanowym zadaniem pozostaje adaptacja znanych w literaturze metod szacowania i wyceny ryzyka lub też wypracowanie nowych na użytek omawianego sektora. Opracowanie miało na celu zwrócenie uwagi na kwestie, których praktyczna kwantyfikacja sprawiłaby wiele kłopotów.

Dyskutowany temat wymaga dużej wnikliwości, a jednocześnie podejścia syntetycznego, zaś w pierwszym rzędzie podjęcia sensownych analiz empirycznych i na tej podstawie formułowania wniosków i zaleceń odnośnie do pomiaru ryzyka udzielania świadczeń medycznych w warunkach polskich. Dotychczas

spotykane podejście do tej tematyki charakteryzowała fragmentaryczność i brak ujęcia dostosowanego do specyfiki zagadnienia. Wydaje się, że uświadomienie sobie wagi tego problemu przyczyni się do podjęcia kolejnych prób zbadania i kompleksowego określenia skali ryzyka w sektorze ochrony zdrowia w Polsce.

Literatura

1. Arrow K.J. (1979). Eseje z teorii ryzyka. PWN, Warszawa.
2. Getzen T. (2000). Ekonomika zdrowia. WN PWN, Warszawa.
3. Grzesiak S. (2007). Wielokryterialna analiza kosztów i korzyści usług zdrowotnych. W: Sterowanie kosztami w Zakładach Opieki Zdrowotnej. Zeszyty Naukowe. Szczecin, nr 478.
4. Iwanicz-Drozdowska M., Nowak A. (2001). Ryzyko banku. SGH, Warszawa.
5. Kaczmarek T.T. (2005). Ryzyko i zarządzanie ryzykiem – ujęcie interdyscyplinarne. Difin, Warszawa.
6. Osińska M. (2000). Ekonometryczne modelowanie oczekiwań gospodarczych. UMK, Toruń.
7. Przybylska-Kapuścińska W. (2001). Zarządzanie ryzykiem i płynnością banku komercyjnego. AE, Poznań.
8. Stiglitz J. (2004). Ekonomia sektora publicznego. WN PWN, Warszawa.
9. Tarczyński W., Łuniewska M. (2004). Dywersyfikacja ryzyka na polskim rynku kapitałowym. Placet, Warszawa.
10. Tarczyński W., Mojsiewicz M. (2001). Zarządzanie ryzykiem. PWE, Warszawa.
11. Zachorowska A. (2006). Ryzyko działalności inwestycyjnej przedsiębiorstw. PWE, Warszawa.
12. Zarządzanie ryzykiem (2007). Red. K. Jajuga. WN PWN, Warszawa.
13. Zarządzanie w opiece zdrowotnej (2001). Red. M. Kautsch, M. Whitfield, J. Klich. UJ, Kraków.

Michał Bernard Pietrzak
Uniwersytet Mikołaja Kopernika w Toruniu

ANALIZA DANYCH PRZESTRZENNYCH A JAKOŚĆ INFORMACJI

Wstęp

W ramach prowadzonych analiz danych przestrzennych należy podkreślić rolę jakości uzyskiwanych informacji. Wysoka jakość informacji powinna zagwarantować bowiem prawidłowy opis zjawiska oraz istniejących zależności. W przypadku realizacji przestrzennej analizy zjawisk ekonomicznych, proces wnioskowania może zostać zakłócony w wyniku niepoprawnej identyfikacji wewnętrznej struktury danych, co stanowi o ważności tej czynności badawczej. W związku z tym, głównym celem opracowania będzie rozpatrzenie zagadnienia prawidłowej identyfikacji wewnętrznej struktury danych przestrzennych, która powinna zapewnić sensowność otrzymanych wniosków zarówno w przypadku opisu badanego zjawiska, jak i w przypadku analizy istniejących zależności. Przedstawiony zostanie model trendu przestrzennego, za pomocą którego opisywana jest ogólna tendencja przestrzenna. Szczególna uwaga zostanie zwrócona na fakt, jak pominięcie ważnego składnika struktury danych, w postaci zjawiska autokorelacji przestrzennej, może prowadzić do błędnych interpretacji. W polskiej literaturze problematyka modelowania zjawisk przestrzennych nie była często podejmowana. Do najważniejszych pozycji należą [2; 5; 9; 10].

1. Identyfikacja wewnętrznej struktury zjawisk przestrzennych

W przypadku danych przestrzennych, badane zjawisko możemy potraktować jako dwuwymiarowe pole losowe $X(u)$, nazywane dalej procesem przestrzennym, gdzie $u = (u_1, u_2)$ zawiera współrzędne na płaszczyźnie [9, s. 88]. W analizach empirycznych, oprócz opisu badanego zjawiska, istotne jest również wykrycie zależności przyczynowo-skutkowych. W ramach wykonania

pełnej analizy przyczynowo-skutkowej, należy rozpocząć badanie od identyfikacji wewnętrznej struktury wybranych procesów przestrzennych, gdzie wyróżnić można dwa składniki. Pierwszy składnik $f(u)$ wyraża niejednorodność zjawiska w średniej. Chodzi tu o niejednorodność systematyczną, która wyraża zmienność w wartości średniej procesu przestrzennego [9, s. 102] i może być opisana modelem trendu przestrzennego. Drugi składnik, określony symbolem $e(u)$, wyraża zmienność zjawiska, posiadającego własność jednorodności[1] w średniej. Składnik jednorodny w średniej opisuje model szumu przestrzennego[2] albo model przestrzennego procesu autoregresyjnego, w przypadku stwierdzenia autozależności.

1.1. Identyfikacja własności niejednorodności w średniej

Niejednorodność systematyczną wyrazić można w postaci trendu przestrzennego, który definiowany jest jako ogólna tendencja przestrzenna[3]. W modelowaniu trendu przestrzennego wykorzystywane są funkcje wielomianowe. Ze względu na postać analityczną funkcji wielomianowej wyróżnić można przestrzenny trend liniowy, trend kwadratowy, trend sześcienny i trendy wyższych rzędów. W przypadku danych przestrzennych można wyróżnić dwie tendencje przestrzenne rozwoju zjawiska. Systematyczny rozwój zjawiska w określonym kierunku, nazywany dalej trendem kierunkowym, oraz rozwój zjawiska wokół dominującego centrum, określonym w pracy jako trend centralny. Trend kierunkowy opisywany jest modelem trendu liniowego, a w przypadku większej złożoności systematycznych zmian modelem trendu sześciennego lub wyższych rzędów. Natomiast trend centralny poprawnie opisuje model trendu kwadratowego, który może być zastąpiony modelami trendu wyższych rzędów w miarę wzrostu systematycznej złożoności.

---

[1] Zakłada się tutaj jednorodność w szerszym sensie (słabą jednorodność), w przypadku której spełnione są następujące założenia

$$E(X(u)) = \mu$$

$$\frac{1}{|N(h)|}\sum_{N(h)}\left(\left[X(u_i)-\mu\right]\left[X(\mu_i)-\mu\right]\right) = K(u_i,u_j{}') = K(h),$$

$$h = u_i - u_j,\ u_i = (u_{1i},u_{2i})$$

gdzie $N(h)$ oznacza zbiór par lokalizacji o różnicy h. Patrz Szulc (2007), s. 28, s. 88, s. 107.

[2] Szum przestrzenny (biały szum) stanowi jednorodne w szerszym sensie pole losowe. Dodatkowo czyni się założenia o zerowej wartości średniej oraz funkcji kowariancji postaci

$$K(u,u') = \begin{cases} \delta^2, & gdy\ \ u = u' \\ 0\ , & gdy\ \ u = u' \end{cases}$$ [9, s. 30].

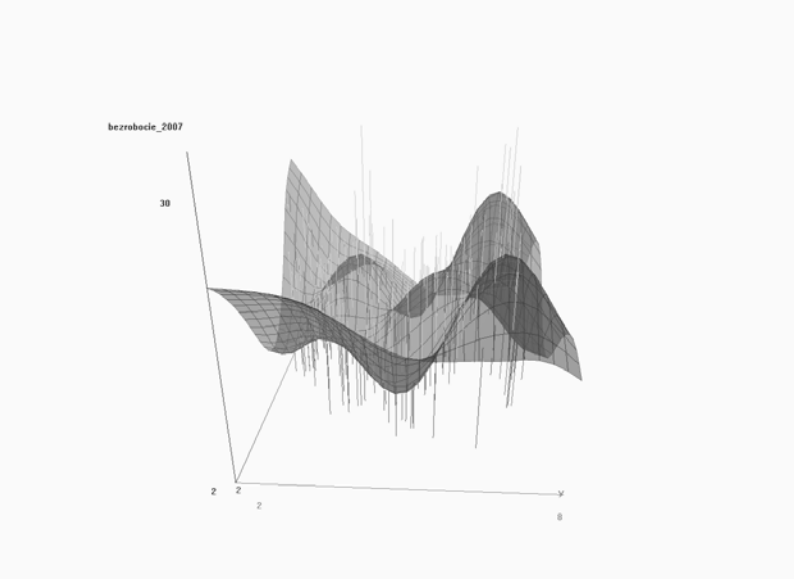
[3] Problematyka związana z trendem przestrzennym przedstawiona została kompleksowo w pracach [5; 9].

Model trendu liniowego określić można za pomocą wzoru

$$X(u) = \alpha_0 + \alpha_1 u_{1i} + \alpha_2 u_{2i} + e(u) \tag{1}$$

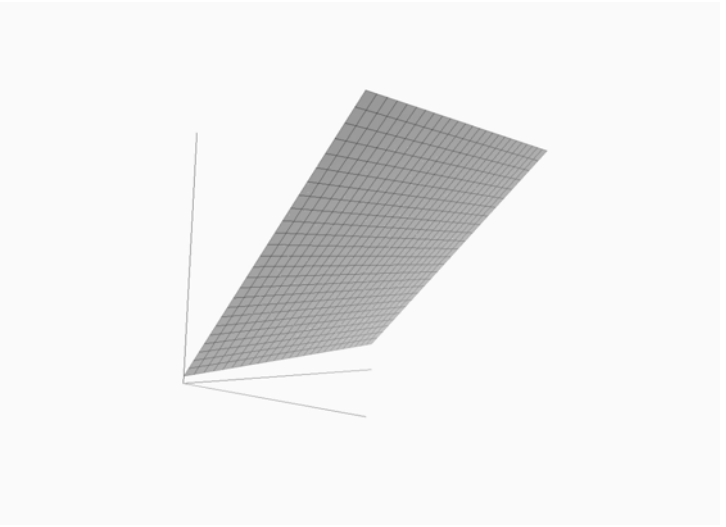
gdzie $\alpha_0, \alpha_1, \alpha_2$ są to parametry strukturalne, u_{1i}, u_{2i} są to współrzędne na płaszczyźnie, a $e(u)$ stanowi proces jednorodny w szerszym sensie.

Model trendu liniowego, jako model trendu kierunkowego, służy do określenia kierunku rozwoju przestrzennego badanego zjawiska. O kierunku i sile rozwoju świadczą znaki i wartości ocen parametrów strukturalnych stojących przy współrzędnych geograficznych. Na rys. 1 przedstawiono kształtowanie się stopy bezrobocia z podziałem na powiaty w 2007 roku. Natomiast na rys. 2 zaprezentowane zostało dopasowanie przestrzennego trendu liniowego, gdzie widoczna jest przestrzenna tendencja wzrostu stopy bezrobocia wraz z przechodzeniem w kierunku północno-wschodniej części Polski. Przy czym znacznie silniejszy wzrost stopy bezrobocia występuje przy założeniu kierunku zachód-wschód.



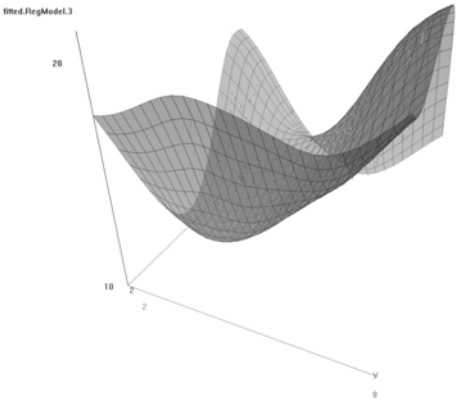
Rys. 1. Wartości stopy bezrobocia

Model trendu liniowego przedstawiony na rys. 2 uwzględnia główny kierunek przestrzennego rozwoju. Systematyczne zmiany najczęściej jednak charakteryzują się większą złożonością i w przypadku stopy bezrobocia oprócz ogólnej tendencji kierunkowej, występuje też tendencja wyższej stopy bezrobocia w powiatach przygranicznych województw, zachodnio-pomorskiego, dolnośląskiego i opolskiego.

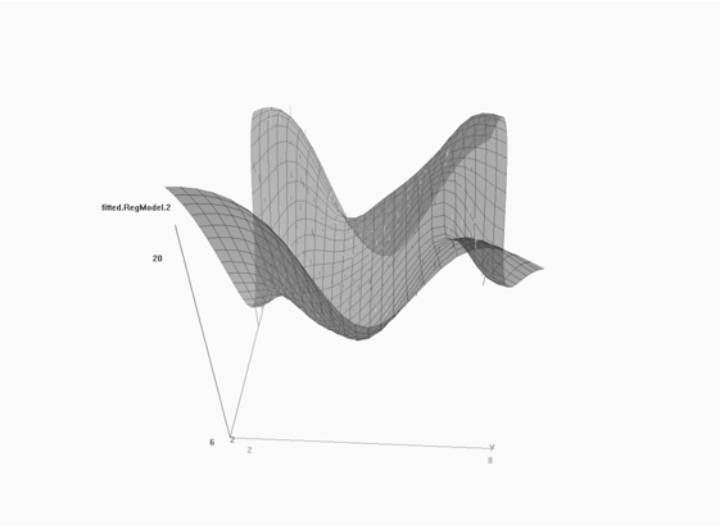


Rys. 2. Dopasowanie modelu trendu liniowego do wartości stopy bezrobocia

Wraz ze wzrostem stopnia modelu trendu uwzględniane jest coraz większe zróżnicowanie badanego szeregu. Widoczne jest to na rys. 3, 4, gdzie oprócz wymienionych powiatów przygranicznych, uwzględniony również został pas wyższych wartości stopy bezrobocia w województwach warmińsko-mazurskim i mazowieckim oraz pas utworzony od terenów byłego województwa Radomskiego, przechodzący przez województwo świętokrzyskie i sąsiadujące województwa małopolskie i podkarpackie.



Rys. 3. Dopasowanie modelu trendu sześciennego do wartości stopy bezrobocia



Rys. 4. Dopasowanie modelu trendu czwartego stopnia do wartości stopy bezrobocia

W przypadku dopasowania kilku modeli trendu przestrzennego pojawia się problem wyboru właściwego modelu. Podobnie jak w przypadku analizy szeregów czasowych, stosowany jest test F [5, s. 32], gdzie w hipotezie alternatywnej zakłada się istotny spadek wariancji resztowej. Ostatecznie jednak wybór stopnia modelu trendu należy do badacza, który powinien ustalić granicę, po przekroczeniu której opisywane jest dowolne zróżnicowanie procesu, a nie, zgodnie z definicją, jedynie zróżnicowanie systematyczne.

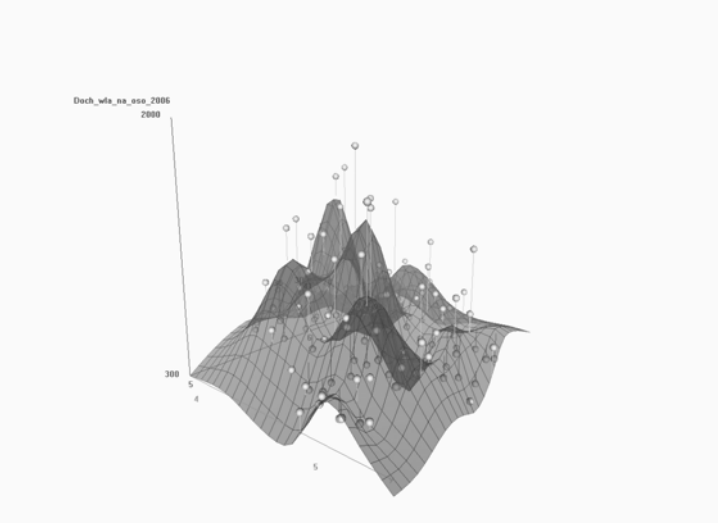
Model trendu kwadratowego[4] opisany jest wzorem

$$X(u) = \alpha_0 + \alpha_1 u_{1i} + \alpha_2 u_{2i} + \alpha_3 u_{1i}^2 + \alpha_4 u_{1i} u_{2i} + \alpha_5 u_{2i}^2 + e(u) \qquad (2)$$

gdzie $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ są to parametry strukturalne, u_{1i}, u_{2i} są to współrzędne na płaszczyźnie, a $e(u)$ stanowi proces jednorodny.

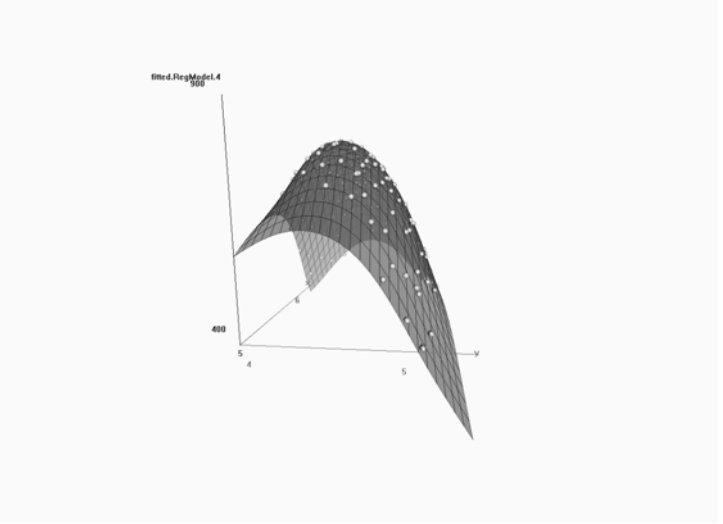
Model trendu kwadratowego poprawnie opisuje tendencją przestrzenną, nazwaną trendem centralnym. Na rys. 5 zaprezentowano kształtowanie się do dochodów własnych na osobę dla gmin województwa kujawsko-pomorskiego w 2006 roku.

[4] Analogicznie tworzone są modele trendu wyższych rzędów. Należy jednak pamiętać o założeniu, że łączny stopień iloczynu u_{1i}, u_{2i} nie może być większy niż stopień zakładanego trendu.



Rys. 5. Wartości dochodów własnych gmin na osobę

Natomiast na rys. 6 przedstawione zostało dopasowanie przestrzennego trendu kwadratowego, gdzie widoczna jest przestrzenna tendencja w postaci istnienia dominującego centrum tworzonego przez dwa, blisko siebie położone największe miasta województwa, Bydgoszcz i Toruń. Biorąc pod uwagę niewielką odległość między tymi miastami, przypuszczać można, że w przyszłości stworzą one aglomerację bydgosko-toruńską.



Rys. 6. Dopasowanie modelu trendu kwadratowego do wartości dochodów własnych

1.2. Identyfikacja własności autokorelacji przestrzennej

Składnik $e(u)$ opisuje własność jednorodności w średniej zjawiska. Po eliminacji trendu przestrzennego zjawisko przestrzenne powinno charakteryzować się jednorodnością w szerszym sensie. Własność ta nie wyklucza jednak występowania autozależności przestrzennych [2; 4; 5; 9; 10]. Zjawisko autokorelacji przestrzennej oznacza istnienie zależności między wartościami procesu w określonej jednostce, a wartościami tego procesu w jednostkach sąsiadujących. Dlatego kolejnym krokiem powinna być identyfikacja własności autokorelacji przestrzennej. Pominięcie zjawiska autokorelacji przestrzennej może stać się przyczyną błędnego wnioskowania, co pokazane zostanie w dalszej części pracy. Również pominięcie trendu przestrzennego w modelu regresji powoduje wzrost autozależności w resztach modelu.

Jedną z miar autokorelacji przestrzennej jest statystyka globalna *I* Morana [4; 7]. Wartości statystyki *I* Morana istotnie większe od zera wskazują na istnienie dodatniej autokorelacji przestrzennej, a wartości istotnie ujemne na autokorelację ujemną. Brak istotnej różnicy od zera świadczy o niewystępowaniu autozależności przestrzennych.

Miara ta określona jest wzorem[5]

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \cdot \frac{\sum_i \sum_j w_{ij} (x_i(u) - \bar{x}(u))(x_j(u) - \bar{x}(u))}{\sum_i (x_i(u) - \bar{x}(u))^2} \tag{3}$$

gdzie N jest liczbą jednostek przestrzennych, $x_i(u) = x(u_{1i}, u_{2i})$ oznacza wartość badanej zmiennej w *i*-tym regionie, $\bar{x}(u)$ jest średnią ze wszystkich regionów, natomiast w_{ij} jest elementem ustalonej wcześniej macierzy wag.

W przypadku istnienia autozależności przestrzennych zmienność składnika $e(u)$ można opisać za pomocą modelu autoregresyjnego procesu przestrzennego pierwszego rzędu [2, s. 13][6], co określamy wzorem

[5] Statystyka *I* Morana posiada rozkład normalny. Wyprowadzenie wzorów dla teoretycznej wartości oczekiwanej $E(I)$ oraz wariancji $Var(I)$ oraz problematyka związana z testowaniem statystyki poruszona została w pracy [4].

[6] W przypadku procesów stochastycznych zakłada się istnienie rzędu autoregresji dowolnego rzędu. Dla potrzeb analiz przestrzennych proces autoregresyjny pierwszego rzędu wydaje się w wystarczającym stopniu uwzględniać autozależności przestrzenne.

$$e(u) = qWe(u) + \eta(u) \tag{4}$$

gdzie q jest parametrem autoregresji, W[7] stanowi macierz wag przestrzennych, a $\eta(u)$ jest szumem przestrzennym.

W przypadku właściwego ustalenia struktury zarówno objaśnianego procesu przestrzennego, jak i procesów objaśniających można przejść do tworzenia zgodnego modelu przyczynowo-skutkowego[8]. W wyniku uwzględnienia właściwej struktury wszystkich procesów przestrzennych uzyskiwany jest pełny model zgodny, określony wzorem[9]

$$\begin{aligned} Y(u) = \alpha_0 + qWY(u) + f(u) + \alpha_1 X_1(u) + \alpha_2 WX_1(u) + ... + \\ + \alpha_{2n} X_n(u) + \alpha_{2n+1} WX_n(uu) + \eta(u) \end{aligned} \tag{5}$$

gdzie $\alpha_0, q, \alpha_1, ..., \alpha_n$ są parametrami strukturalnymi, $f(u)$ stanowi trend przestrzenny, W jest macierzą wag przestrzennych, a $\eta(u)$ jest szumem przestrzennych. Składniki nieistotne mogą następnie ulec redukcji, w wyniku czego otrzymany jest końcowy model zgodny.

2. Potencjalne skutki nieuwzględnienia zjawiska autokorelacji przestrzennej

Wnioskowanie statystyczne daje możliwość wydawania sądów o populacji na podstawie próby. Musi jednak zachodzić warunek niezależności obserwacji w próbie. Występowanie zjawiska autokorelacji prowadzi do obciążenia estymatorów, co z kolei przekłada się na jakość testów statystycznych i jakość wnioskowania statystycznego. W celu pokazania potencjalnych skutków nieuwzględnienia zjawiska autokorelacji przestrzennej wykonano trzy symulacje. Celem pierwszej symulacji była ocena jakości testu istotności dla wartości oczekiwanej w warunkach istnienia autokorelacji przestrzennej. Dla zwiększenia realności wykorzystania testu, symulacja została przeprowadzona w układzie przestrzennym Polski w podziale na powiaty. Rysunek 7 przed-

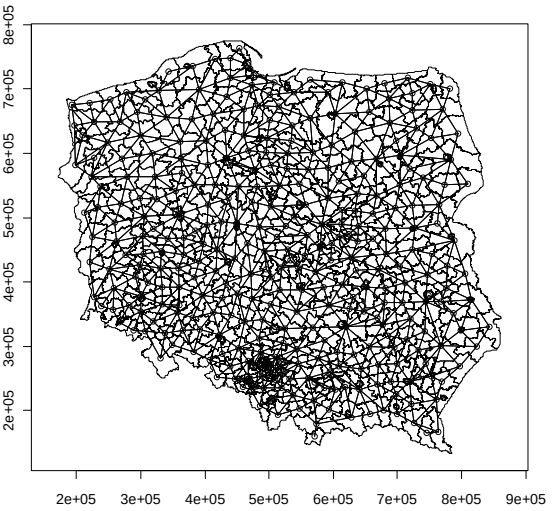
[7] W badaniach wykorzystuje się najbardziej intuicyjne kryterium sąsiedztwa zdefiniowane na podstawie posiadania wspólnej granicy. Tak określona macierz wag przestrzennych nazywana jest macierzą sąsiedztwa $W = (w_{i,j})_{i,j=1,...,n}$, która posiada zera na przekątnej, natomiast poza przekątną wartości macierzy są równe $w_{i,j} = \begin{cases} 1, & gdy\ obiekty\ i\ oraz\ j\ granicą\ ze\ sobą, \\ 0, & w\ przeciwnym\ przypadku. \end{cases}$

Dodatkowo standaryzuje się macierz wag poprzez wprowadzenie dla każdego wiersza zależności $\sum_j w_{ij} = 1$.

[8] Autor opiera się na pojęciach modelowania zgodnego, jak i modelu zgodnego wprowadzonych przez Z. Zielińskiego. Koncepcja ta w przypadku przestrzennych pól losowych została rozważona w pracy [9].

[9] Można rozważać, że zjawisko autokorelacji opisuje model autoregresyjny wyższego rzędu niż pierwszy. Należałoby wtedy odpowiednio rozbudować model przyczynowo-skutkowy.

stawia układ przestrzenny wykorzystany w ramach zrealizowanej symulacji, wraz z uwzględnieniem sąsiedztwa pierwszego rzędu. Test ten w założonym układzie przestrzennym może posłużyć do badania istotności wartości oczekiwanej wybranej cechy dla Polski, na podstawie posiadanych danych tej cechy w powiatach.



Rys. 7. Układ przestrzenny z uwzględnieniem sąsiedztwa I rzędu

Symulacja polegała na generowaniu kolejno szumu przestrzennego o określonej średniej oraz trzech kolejnych procesów autoregresyjnych o parametach $\alpha = 0.5$, $\alpha = 0.7$, $\alpha = 0.9$ [10]. Efektem było uzyskanie struktury przestrzennej charakteryzującej się własnością autokorelacji. W celu symulacji procesu autoregresyjnego skorzystano z zależności [2, s. 13].

$$X(u) = qWX(u) + e \Leftrightarrow (I - qW)X(u) = e(u)$$

$$X(u) = (I - qW)^{-1} e(u) \Leftrightarrow X = e(u) + qWe(u) + q^2W^2e(u) + ... \qquad (6)$$

gdzie q jest parametrem autoregresji, W stanowi macierz wag przestrzennych, a $e(u)$ jest szumem przestrzennym.

W wyniku przeprowadzonej symulacji otrzymano wartości krytyczne testu[11], którymi były odpowiednie kwantyle uzyskanych statystyk oraz policzono wartości prawdopodobieństwa popełnienia błędu pierwszego rodzaju, na podstawie liczby odrzuceń. Tabela 1 przedstawia wyniki przeprowadzonej

[10] W przypadku wszystkich symulacji proces szumu przestrzennego, jak i kolejnych procesów autoregresyjnych generowano 50 000 razy.

[11] Założono w symulacji test prawostronny. Wyniki dla testu lewostronnego lub dwustronnego można wywnioskować na podstawie wyników dla testu prawostronnego.

symulacji w przypadku szumu przestrzennego. Uzyskane wartości krytyczne, jak i błąd pierwszego rodzaju można uznać z prawidłowe. Oznacza to, że wnioskowanie statystyczne na podstawie tego testu, przy założeniu niezależności przestrzennej, prowadzi do właściwych wniosków[12].

Tabela 1

Wyniki symulacji przy założeniu generowania szumu przestrzennego

Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	2,33	1,64	1,27
Błąd pierwszego rodzaju	0,0104	0,04964	0,0994

Wyniki symulacji, w przypadku generowania przestrzennych procesów autoregresyjnych, przedstawione zostały w tabeli 2. Wraz ze wzrastającym poziomem autokorelacji przestrzennej, rosną wartości krytyczne. Oznacza to, że stosując wartości krytyczne z rozkładu normalnego i nie uwzględniając zjawiska autokorelacji, zwiększa się znacznie prawdopodobieństwa odrzucenia hipotezy zerowej. Przejawia się to w znacznie zwiększonym poziomie błędu pierwszego rodzaju, który rośnie wraz ze wzrostem autokorelacji przestrzennej. Prowadzi to do wniosku, że w przypadku istnienia zjawiska autokorelacji jakość testu znacznie spada.

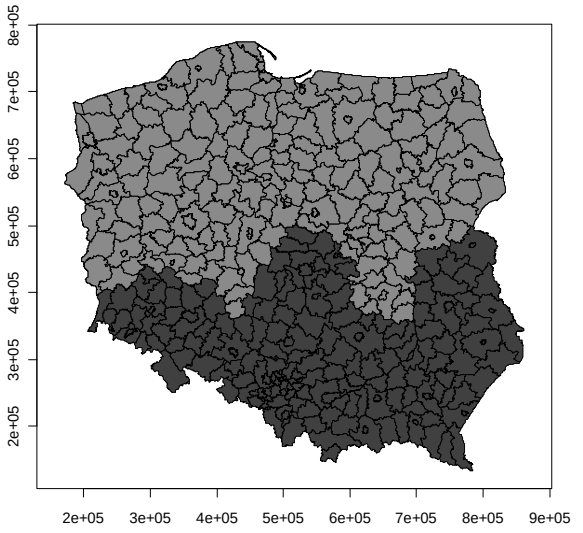
Tabela 2

Wyniki symulacji przy założeniu generowania autoregresyjnych procesów przestrzennych

Proces autoregresyjny $\alpha = 0{,}5$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	4,15	2,96	2,31
Błąd pierwszego rodzaju	0,09698	0,1804	0,2378
Proces autoregresyjny $\alpha = 0{,}7$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	5,5	3,82	3
Błąd pierwszego rodzaju	0,1564	0,2358	0,2892
Proces autoregresyjny $\alpha = 0{,}9$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	6,6	4,71	3,61
Błąd pierwszego rodzaju	0,2119	0,2881	0,3317

[12] Badacz powinien jednak pamiętać, że może popełnić błąd pierwszego rodzaju z prawdopodobieństwem równym założonemu poziomowi istotności α.

Zbadana została również jakość testu na istotność dwóch wartości oczekiwanych. W tym przypadku posłużono się układem administracyjnym Polski w podziale na powiaty. Wyodrębniono dwa obszary. Północny, który utworzyły powiaty województw zachodnio-pomorskiego, pomorskiego, warmińsko-mazurskiego, lubuskiego, wielkopolskiego, kujawsko-pomorskiego, mazowieckiego i podlaskiego. Południowy, powstały przez zsumowanie liczby powiatów województw dolnośląskiego, opolskiego, łódzkiego, śląskiego, świętokrzyskiego, lubelskiego, małopolskiego oraz podkarpackiego. Test ten można wykorzystać do wykazania istotnej różnicy lub jej braku w przypadku dwóch wartości oczekiwanych wybranej cechy, przy założeniu, że cecha pochodzi z rozkładów obserwowanych na dwóch obszarach. Rysunek 8 przedstawia dwa ustalone obszary, na podstawie których przeprowadzono kolejną symulację.



Rys. 8. Podział na dwa obszary przestrzenne

Dla obydwu obszarów generowano szumy przestrzenne oraz przestrzenne procesy autoregresyjne, przy założeniu tych samych współczynników autoregresji $\alpha_{1,2} = 0.5$, $\alpha_{1,2} = 0.7, \alpha_{1,2} = 0.9$. Następnie wyznaczono wartości krytyczne, jako kwantyle dla uzyskanych statystyk testu. Policzono również prawdopodobieństwo popełnienia błędu pierwszego rodzaju na podstawie liczby odrzuceń.

Otrzymane wyniki dla szumu przestrzennego zamieszczone zostały w tabeli 3.

Tabela 3

Wyniki symulacji przy założeniu generowania szumu przestrzennego

Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	2,34	1,65	1,29
Błąd pierwszego rodzaju	0,0104	0,0516	0,1016

Zawarte w tablicy wartości krytyczne oraz prawdopodobieństwa popełnienia błędu pierwszego rodzaju są prawidłowe, co oznacza poprawność testu w przypadku braku zależności przestrzennych.

Otrzymane wyniki jakości testu dla danych generowanych przez dwa przestrzenne procesy autoregresyjne zawarte zostały w tabeli 4.

Tabela 4

Wyniki symulacji przy założeniu generowania autoregresyjnych procesów przestrzennych

Proces autoregresyjny $\alpha = 0{,}5$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	3,61	3	2,31
Błąd pierwszego rodzaju	0,09794	0,1798	0,2382
Proces autoregresyjny $\alpha = 0{,}7$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	4,71	3,82	2,96
Błąd pierwszego rodzaju	0,1589	0,2398	0,2909
Proces autoregresyjny $\alpha = 0{,}9$			
Poziom istotności	0,01	0,05	0,1
Wartości krytyczne	6,6	5,4	4,15
Błąd pierwszego rodzaju	0,2068	0,2805	0,3246

Widoczne jest, że wraz ze wzrastającym poziomem autokorelacji przestrzennej, rosną również wartości krytyczne testu. Fakt ten powinien przejawić się w częstszym odrzucaniu hipotezy zerowej, niż to wynika z przyjętego poziomu istotności α. Prawdopodobieństwo popełnienia błędu pierwszego rodzaju znacznie wzrosło, co podobnie jak dla testu na istotność wartości oczekiwanej, świadczy o pogarszaniu się jakości testu.

Problemy przedstawione wcześniej dotyczyły wnioskowania co do podstawowych charakterystyk badanego zjawiska. Kolejnym krokiem będzie rozpatrzenie jakości estymacji modelu trendu przestrzennego, w przypadku nieuwzględnienia zjawiska przestrzennej autokorelacji. Pojawiają się tutaj dwa

aspekty. Aspekt pierwszy dotyczy istotności parametrów modelu. Istotna autokorelacja przestrzenna powoduje obciążenie wariancji składnika resztowego i w konsekwencji prowadzi do nieefektywnej estymacji parametrów strukturalnych [5, s. 39]. Oznacza to możliwość sytuacji pozornej istotności parametru w modelu lub sytuacji nieistotności parametru, podczas gdy jest on statystycznie istotny. Aspekt drugi dotyczy dopasowania modelu trendu przestrzennego, które przy nie uwzględnieniu autokorelacji przestrzennej pozornie rośnie. Ponownie wygenerowano szum przestrzenny oraz przestrzenne procesy autoregresyjne z parametrami $\alpha = 0.5, \alpha = 0.7, \alpha = 0.9$. Następnie szacowano kolejno przestrzenny model trendu liniowego, trendu kwadratowego oraz trendu sześciennego. Na podstawie uzyskanych wyników policzono kwantyle dla uzyskanych współczynników determinacji oraz prawdopodobieństwo istotności parametrów strukturalnych.

Tabele 5, 6 oraz 7 przedstawia wyliczone kwantyle dla uzyskanych wartości współczynników determinacji. We wszystkich trzech tabelach widoczny jest wzrost wartości kwantyli współczynnika determinacji w miarę wzrostu siły autozależności przestrzennych. Ponieważ w strukturze symulowanych procesów nie ma trendu przestrzennego, uzyskany wynik świadczy o możliwości pozornego dopasowania modelu. Błąd ten wynika z nieuwzględnienia zjawiska autokorelacji przestrzennej.

Tabela 5

Wartości kwantyli współczynnika determinacji dla estymacji trendu liniowego

Kwantyle*	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
0,9	0,0129	0,038	0,062	0,094
0,95	0,0154	0,048	0,079	0,118
0,99	0,0242	0,073	0,118	0,173

* Kolejne kwantyle oznaczają wartość, poniżej której znajduje się odpowiednio 90%. 95% oraz 99% wszystkich uzyskanych wartości współczynnika determinacji.

Tabela 6

Wartości kwantyli współczynnika determinacji dla estymacji trendu kwadratowego

Kwantyle	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
0,9	0,024	0,073	0,094	0,134
0,95	0,029	0,086	0,118	0,183
0,99	0,0398	0,114	0,173	0,234

Tabela 7

Wartości kwantyli współczynnika determinacji dla estymacji trendu sześciennego

Kwantyle	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
0,9	0,038	0,11	0,17	0,22
0,95	0,043	0,12	0,19	0,25
0,99	0,056	0,15	0,23	0,31

Tabele 8, 9 oraz 10 zawierają otrzymane wartości prawdopodobieństwa istotności parametrów strukturalnych. Analiza wyników wskazuje na fakt, że w przypadku istnienia autokorelacji znacznie zwiększa się prawdopodobieństwo istotności parametrów strukturalnych. W tabeli 8 przedstawiono dodatkowo w ostatnim wierszu prawdopodobieństwo, w przypadku istotności choć jednego z dwóch parametrów[13], prawdopodobieństwo uznania modelu za istotny wzrasta prawie dwukrotnie. Uwzględnienie ostatniego wiersza w tabeli 8 powoduje prawie dwukrotny wzrost prawdopodobieństwa do poziomu 46%, 63% oraz 74%, w zależności od siły autozależności przestrzennych.

Tabela 8

Prawdopodobieństwo istotności parametrów dla estymacji trendu liniowego

Parametry strukturalne	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
α_1	0,051	0,2742	0,39692	0,49974
α_2	0,0499	0,27412	0,3961	0,49984
Łącznie	0,0975	0,4692	0,63118	0,74564

Tabela 9

Prawdopodobieństwo istotności parametrów dla estymacji trendu kwadratowego*

Parametry strukturalne	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
α_3	0,05202	0,2642	0,38605	0,48721
α_4	0,05102	0,26765	0,3792	0,48317
α_5	0,0494	0,2659	0,387	0,47974

* W tabeli 9 i 10 podane zostały wyniki jedynie dla parametrów stojących przy iloczynach najwyższych stopni, dla trendu kwadratowego stopnia drugiego, a dla sześciennego trzeciego.

[13] Czynione jest założenie, że do uznania istotności trendu przestrzennego wystarczy istotność jednego z parametrów stojących przy iloczynach najwyższych stopni.

Tabela 10

Prawdopodobieństwo istotności parametrów dla estymacji trendu sześciennego

Parametry strukturalne	Szum przestrzenny	Proces autoregresyjny (0,5)	Proces autoregresyjny (0,7)	Proces autoregresyjny (0,9)
α_4	0,05066667	0,23715	0,33905	0,4311
α_5	0,0478	0,2366	0,33335	0,43915
α_6	0,0508	0,2526	0,35785	0,4675
α_7	0,0495333	0,25425	0,3579	0,4629

Zakończenie

Celem opracowania było przedstawienie problemu identyfikacji wewnętrznej struktury procesów przestrzennych. Ustalenie w sposób prawidłowy wewnętrznej struktury powinno stanowić podstawę każdej analizy ekonomicznej dla danych przestrzennych oraz zapewnić poprawność uzyskiwanych informacji. W artykule pokazana została możliwość opisu tej struktury poprzez wykorzystanie modelu trendu przestrzennego oraz przestrzennych modeli autoregresyjnych. Za pomocą przeprowadzonych symulacji wykazano, że nie uwzględnienie zjawiska autozależności przestrzennych znacznie obniża wartość poznawczą prowadzonych analiz. Istotna autokorelacja przestrzenna nie tylko pogarsza jakość testów w przypadku wnioskowania co do charakterystyk zjawiska, ale również wyklucza poprawne modelowanie ekonometryczne. Skutkiem zjawiska autokorelacji jest obciążenie estymatorów. Przekłada się to na jakość testów statystycznych, a w przypadku przestrzennych modeli ekonometrycznych na jakość estymacji i weryfikacji. Ostatecznie problem poprawnej identyfikacji struktury zjawisk przestrzennych, jako niezbędnej czynności badawczej w każdej ekonomicznej analizie przestrzennej, przekłada się na jakość uzyskiwanych informacji.

Literatura

1. Anselin L. (1988). Spatial Econometrics: Methods and Models. Kluwer Academic Press.
2. Bivand R. (1981). Modelowanie geograficznych układów czasowo-przestrzennych. PWN, Warszawa-Poznań.
3. Cliff A., Ord J.K. (1973). Spatial Autocorrelation. Pion, London.
4. Cliff A., Ord J.K. (1981). Spatial process: Models and Applications. Pion, London.

5. Czyż T. (1978). Metody generalizacji układów przestrzennych. PWN, Warszawa-Poznań.
6. Kopczewska K. (2006). Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN. CeDeWu, Warszawa.
7. Moran P. (1948). The Interpretation of Statistical Map. Journal of the Royal Statistical Society, Series B, 10, 243-251.
8. Paelick J.H.P., Klaassen L.H. (1979). Spatial Econometrics. Saxon Hause, Farnborough.
9. Szulc E. (2007). Ekonometryczna analiza wielowymiarowych procesów gospodarczych. UMK, Toruń.
10. Zeliaś A. (1991). Ekonometria przestrzenna. Wydawnictwo Ekonomiczne, Warszawa.

Elżbieta Aleksandra Studzińska
Uniwersytet Marii Curie-Skłodowskiej w Lublinie

RYZYKO W DZIAŁALNOŚCI BANKOWEJ

Wprowadzenie

Banki pod wpływem rosnącej konkurencji oraz oczekiwań coraz wyższej stopy zwrotu częściej skłonne są angażować się w ryzykowne przedsięwzięcia. Ryzyko jest ściśle związane z działalnością bankową i oznacza zagrożenie osiągnięcia zamierzonych zysków, co może wiązać się ze zmniejszeniem potencjalnych zysków, płynności, wiarygodności a także trudnościami finansowymi. Nie można wyeliminować ryzyka, ale banki starają się nim zarządzać, zarówno w procesie zwiększenia przychodów, jak i redukcji kosztów [9, s. 28].

Rozważania nad ryzykiem bankowym warto rozpocząć od określenia, czym właściwie jest ryzyko, w szczególności ryzyko bankowe. W literaturze można spotkać różne ujęcia ryzyka. W potocznym znaczeniu ryzyko związane jest z niepewnością, którą możemy zdefiniować jako brak pewności co do przyszłego przebiegu zdarzeń. Niepewność charakteryzuje się tym, że rozkład prawdopodobieństwa jest nieznany. W ekonomii i statystyce można spotkać się z odróżnieniem ryzyka i niepewności. Przez ryzyko rozumie się wówczas sytuację, w której są znane prawdopodobieństwa wystąpienia określonych zdarzeń, a więc znana jest wartość oczekiwana i wariancja zmiennej losowej. Niepewność charakteryzuje się tym, że rozkład prawdopodobieństwa jest nieznany. Zgodnie z tą koncepcją w działalności bankowej mamy do czynienia głównie z sytuacją niepewności, a nie ryzyka [13, s. 627]. W *Słowniku języka polskiego* ryzyko definiowane jest jako możliwość, prawdopodobieństwo, że coś się nie uda, przedsięwzięcie, którego wynik jest nieznany, niepewny, problematyczny [15].

Ryzyko zdefiniować można, kładąc akcent na jego różne aspekty:

- decyzyjny – niebezpieczeństwo podjęcia błędnych rozstrzygnięć (decyzji),
- działania – niebezpieczeństwo niepowodzenia podjętego działania,
- planistyczny – niezrealizowanie założonych wielkości celowych,
- straty – niebezpieczeństwo utraty majątku,
- celowy – niebezpieczeństwo negatywnego odchylenia od zamierzonego celu [8, s. 277].

Rozważania na temat istoty ryzyka można sprowadzić do dwóch zasadniczych nurtów pierwszy związany z teorią podejmowania decyzji, drugi odnoszący się do teorii zarządzania ryzykiem.

Pierwszy z nich rozwinął się na podstawie ujęcia problematyki ryzyka przez F.H. Knighta, który poszukiwał przyczynowego ujęcia ryzyka, przyporządkowując pojawienie się określonych zdarzeń rachunkowi prawdopodobieństwa. Zgodnie z jego poglądami ryzyko występuje wówczas, gdy można je określić za pomocą prawdopodobieństwa matematycznego, statystycznego i szacunkowego. W pozostałych przypadkach, gdy nie można określić ilościowo wyniku procesów czy decyzji, mamy do czynienia z niepewnością. Inną interpretację zaproponował J. Pfeffer, który stwierdził, że ryzyko jest kombinacją hazardu i mierzone jest prawdopodobieństwem, niepewność mierzona jest przez poziom wiary. W tym nurcie rozważań mieszczą się także poglądy O. Langego, określającego ryzyko jako niepewność przewidywanych zdarzeń z przyszłości, wynikającą z niepewności i niedokładności badań statystycznych na podstawie, których dokonuje się szacowania przyszłości.

Drugi nurt rozważań poświęconych ryzyku (w teorii zarządzania) koncentruje się przede wszystkim na skutkach ryzyka. W ramach tego nurtu wyróżnia się dwa sposoby interpretacji ryzyka:

- określanego jako możliwość uzyskania efektów niezgodnych (gorszych lub lepszych) z oczekiwaniami podejmującego decyzję,
- określanego jako możliwość poniesienia szkody lub straty (eksponującego negatywne skutki ryzyka) [11, s. 8].

Termin ryzyko stosowany w sektorze bankowym oznacza zagrożenie osiągnięcia zamierzonych celów. Dla banku oznacza to zmniejszenie potencjalnych zysków, utratę płynności i wiarygodności, zmniejszenie kapitału własnego, trudności finansowe, w krańcowym przypadku bankructwo. Istnieje, zatem ścisły związek między ryzykiem a stopniem realizacji celów [16, s. 23].

W literaturze przedmiotu istnieje wiele klasyfikacji ryzyka bankowego, każda z nich ma swoje zalety i wady Mimo wielu prób podejmowanych w tym zakresie, nie przedstawiono dotychczas jednolitej klasyfikacji ryzyka. Dla przejrzystej prezentacji tematyki ryzyka bankowego postaram się zaprezentować kilka spojrzeń na jego klasyfikację od prostego dychotomicznego podejścia, jakie prezentuje Z. Fedorowicz do specyficznego jakie zostało zaprezentowane w *Nowoczesnej bankowości* S. Heffernan.

Celem opracowania jest scharakteryzowanie ryzyka w działalności bankowej, które definiowane jest jako możliwość wystąpienia zdarzeń, które mogą negatywnie wpłynąć na wynik finansowy lub fundusze własne banku [10, s. 173]. W niniejszym opracowaniu skoncentrowano się na przedstawieniu poszczególnych definicji ryzyka występującego w działalności bankowej oraz próbie klasyfikacja rodzajów ryzyka typowo bankowego ze względu

na różne kryteria. Następnie podjęto próbę wskazania głównych przyczyny wystąpienia ryzyka bankowego. Zwrócono uwagę na kwestię dotyczącą metod pomiaru ryzyka oraz ocenie ryzyka bankowego. Podkreślono, wpływ zarządzania ryzykiem bankowym obejmującego wszystkie rodzaje ryzyka bankowego począwszy od błędnego oszacowania ryzyka, kończąc na niewłaściwych prognozach ekonomicznych.

1. Pojęcie i rodzaje ryzyka bankowego

Ryzyko bankowe jest ściśle powiązane z działalnością gospodarczą prowadzoną przez banki. Wynika ono głównie z funkcji pośredników finansowych, pełnionych przez nie w gospodarce narodowej. Interpretacja ryzyka w literaturze jest bardzo niejednorodna, o różnym stopniu szczegółowości, odmiennym zakresie czy źródłach powstawania oraz jego wielkości. Brak jest syntetycznej powszechnie uznawanej definicji ryzyka bankowego [11, s. 10]. Ryzyko bankowe dzieli się na poszczególne kategorie ryzyka osobno zdefiniowane z odpowiednimi miarami ryzyka. W określonych rodzajach działalności bankowej ujawniają się realne i potencjalne skutki ryzyka.

Ryzyko w działalności bankowej można podzielić według różnych kryteriów:

a) horyzontu czasowego,
b) obszaru ryzyka,
c) częstotliwości wystąpienia,
d) źródła ryzyka,
e) NUK.

1.1. Kryterium horyzontu czasowego

Ryzyko w działalności bankowej, w zależności od przyjętego horyzontu czasowego analizy, dzielmy na ryzyko strategiczne i ryzyko operacyjne. Ryzyko strategiczne wpływa na długookresową zdolność konkurencyjną banku. Zwraca się przy tym uwagę przede wszystkim na ryzyko związane ze strukturą właścicieli banku i jego zarządu. Ryzyko związane ze strukturą właścicieli polega na zagrożeniu, iż właściciele nie mogą lub nie chcą wyposażyć banku w kapitał niezbędny dla sprawnego funkcjonowania. Oprócz tego właściciele i zaangażowany przez nich zarząd podejmują wiele decyzji strategicznych, wywierających istotny wpływ na dalszą działalność banku i jego zagrożenie ryzykiem [14, s. 308]. Ryzyko operacyjne wynika z błędów powstałych przy dokonywaniu lub rozliczaniu transakcji, które są powodowane przez kadrę banku bądź urządzenia techniczne [5, s. 139]. Ryzyko operacyjne obejmuje analizę obszarów wystąpienia w banku potencjalnych strat spowodowanych przez

niekompetencję lub złą wolę pracowników, wadliwe działanie systemów informatycznych, niefunkcjonalne struktury organizacyjne i złą organizację pracy. Jest to ryzyko związane z poprawnym funkcjonowaniem banku zarówno w obszarze działalności ludzi, jak i sprawności techniki. Szeroko rozumiane ryzyko operacyjne obejmuje też część ryzyka prawnego, wpływa na ryzyko utraty reputacji [16, s. 25].

1.2. Kryterium obszaru ryzyka

Zgodnie z kryterium obszaru ryzyka, rozróżniamy ryzyko w obszarze finansowym i w obszarze techniczno-organizacyjnym. Ryzyko w obszarze finansowym jest ryzykiem typowo bankowym i ma znaczenie przy zarządzaniu ryzykiem w działalności bankowej. Do ryzyka w obszarze finansowym zaliczamy ryzyko płynności i wyniku. Ryzyko płynności oznacza zagrożenie przejściowej lub całkowitej płynności przez bank, czyli zdolności banku do terminowego regulowania zobowiązań. Płynność jest wypadkową wielkości aktywów i pasywów, które to wielkości są oceniane pod kątem okresu ich utrzymania, powinny ze sobą korespondować. Utratę płynności może spowodować masowe, wcześniejsze wycofywanie wkładów przez klientów lub brak spłaty kredytów albo opóźnienia w spłacie. W ekstremalnym stanie ryzyko płynności oznacza niemożność realizacji zobowiązań przez bank. Wtedy równa się też ryzyku niewypłacalności. W praktyce polega na braku wystarczającego dostosowania aktywów i pasywów w określonych terminach zapadalności i wymagalności, co niesie ze sobą zagrożenie brakiem zdolności do realizacji płatności. Zwykle przewidywane trudności z płynnością prowadzą do reakcji instytucji nadzoru bankowego, które zobowiązują bank do podjęcia działań zaradczych [16, s. 25]. W działalności pojedynczych instytucji finansowych ryzyko płynności może zostać wywołane poprzez spadek wzajemnego zaufania kontrahentów funkcjonujących na rynku międzybankowym, między innymi w skutek trudności w wiarygodnej ocenie ryzyka, w tym kredytowego, spowodowanych zmianą wyceny aktywów [12, s. 57-58]. Ryzyko wyniku to niebezpieczeństwo nieosiągnięcia przez bank założonego wyniku finansowego. Może ono oznaczać albo poniesienie straty, albo uzyskanie wyniku niższego od zaplanowanego [11, s. 37]. Wśród ryzyka wyniku występującego w obszarze finansowym można wyróżnić ryzyko związane z partnerem transakcji oraz ryzyko cenowe. Ryzyko związane z partnerem transakcji, czyli niebezpieczeństwo niezrealizowania zamierzonego wyniku na skutek straty wynikającej z nie wywiązania się ze zobowiązań partnera transakcji, np. ryzyko kredytowe, ryzyko start z tytułu spadku cen akcji, papierów dłużnych lub udziałów na skutek pogorszenia się standingu ich remitenta, ryzyko podmiotowe z tytułu operacji pozabilansowych. Może wystąpić przy tym ryzyko podmiotowe z tytułu transakcji bilansowych i pozabilansowych.

Na ryzyko z tytułu transakcji bilansowych narażony jest bank w przypadku:
- niespłacenia przez kredytobiorcę zaciągniętych kredytów,
- straty z powodu spadku kursu notowanych na giełdzie skryptów dłużnych, na skutek pogorszenia się standingu ich remitenta,
- utraty wartości posiadanych udziałów w innych firmach,
- niewypłacania dywidendy przez firmy, w których bank ma udziały,
- spadku kursu posiadanych akcji.

Ryzykiem z tytułu transakcji pozabilansowych tradycyjnych jest udzielenie gwarancji i pożyczek czy transakcji instrumentami pochodnymi (swapowych, terminowych, opcyjnych).

Ryzyko cenowe jest określane jako część ryzyka rynkowego, przyczyną jego wystąpienia jest niekorzystne dla banku kształtowanie się na rynku cen instrumentów finansowych, tj. stóp procentowych, kursów walut i kursów akcji.

Ryzyko cenowe z tytułu transakcji bilansowych to:
- ryzyko stopy procentowej, spowodowane niekorzystnymi tendencjami w zakresie rynkowych stóp procentowych,
- ryzyko spadku cen metali szlachetnych,
- ryzyko walutowe to niebezpieczeństwo pogorszenia się sytuacji finansowej banku na skutek niekorzystnych zmian kursu walutowego,
- ryzyko spadku kursu akcji, wynikające z ogólnego pogorszenia się indeksu giełdowego, a nie sytuacji danej firmy [6, s. 228].

Ryzykiem cenowym z tytułu transakcji pozabilansowych są operacje swapowe, terminowe i opcyjne.

Ryzyko w obszarze techniczno-organizacyjnym występuje w różnych rodzajach przedsiębiorstw, nie tylko w bankach. Można je podzielić na:
- ryzyko o charakterze personalnym, związane z kwalifikacjami zatrudnionego personelu,
- ryzyko o charakterze organizacyjnym, które mogą wynikać z niewłaściwej struktury organizacyjnej, zasad funkcjonowania kontroli wewnętrznej, planowania, księgowania, współdziałania pomiędzy poszczególnymi działami,
- ryzyko o charakterze rzeczowo-technicznym, wynikające z posiadanego przez bank wyposażenia technicznego i możliwości przeprowadzenia operacji niezbędnych do prawidłowego funkcjonowania, np. pojemności i rodzaju operacji możliwych do przeprowadzenia za pomocą systemu elektronicznego przetwarzania danych, niezawodności tego sytemu.

Do ryzyka w obszarze technicznym zaliczamy:
- ryzyko z tytułu odpowiedzialności banku, np. za błędy w doradzaniu klientom, w zarządzaniu majątkiem lub prospekcie emisyjnym,
- ryzyko związane z obrotem płatniczym, np. błędne wypełnienie zleceń płatniczych, realizacja sfałszowanych czeków,

- ryzyko zniszczenia majątku banku na skutek klęsk żywiołowych lub kradzieży,
- ryzyko zakłóceń w systemie elektronicznego przetwarzania danych.

1.3. Kryterium częstotliwości wystąpienia

Z punktu widzenia częstotliwości występowania ryzyka wyodrębnia się ryzyko płynności, kredytowe i rynkowe. Ryzyko płynności jest to składnik ryzyka struktury bilansu. Jest to zagrożenie zdolności banku do terminowego regulowania swych zobowiązań. Ryzyko kredytowe to zagrożenie, że płatności związane z obsługą kredytu (tj. raty kapitałowe i odsetki) nie zostaną uregulowane w całości bądź częściowo w terminie przewidzianym umową. Dzieli się na aktywne (związane z udzielaniem kredytu) i pasywne (związane z refinansowaniem). Ryzyko rynkowe (cenowe) jest to zagrożenie, że powstanie strata z tytułu niekorzystnej zmiany parametrów rynkowych, bądź instrumentów finansowych. Obejmuje ryzyko kursu walutowego, ryzyko handlowych papierów wartościowych oraz ryzyko surowcowe/towarowe [5, s. 22]. Coraz częściej wymienia się także ryzyko systemowe**,** czyli niebezpieczeństwo przenoszenia się kryzysu finansowego z jednego kraju do drugiego, spowodowane rosnącą współzależnością systemów finansowych.

1.4. Kryterium źródła ryzyka

Ryzyko bankowe możemy podzielić na wewnętrzne, zewnętrzne, krajowe i zagraniczne [6, s. 229]. Ze względu na źródła niepewności w działalności bankowej występują dwa rodzaje ryzyka wewnętrzne i zewnętrzne. Ryzyko występujące w działalności wewnętrznej banku może wpływać na działalność banku w postaci jakości zatrudnionego personelu, niezawodność systemów przetwarzania danych oraz zabezpieczeń majątku. Może być pochodną decyzji podejmowanych przez kierownictwo i zarząd banku. Ryzyko występujące w działalności zewnętrznej banku związane jest z pogarszaniem się wypłacalności kredytobiorców, wahaniami rynkowych stóp procentowych oraz wahaniami kursów walut. Ryzyko krajowe uzależnione jest od sytuacji polityczno-ekonomicznej, mającej wyraz w nastrojach politycznych, sytuacji makroekonomicznej danej gospodarki, poziomie inflacji, bezrobocia i stanu budżetu. Ryzyko zagraniczne wynika z międzynarodowych tendencji globalizacji i internacjonalizacji gospodarki, a także stopnia aktywności gospodarczej jednego lub grupy krajów, które łączy albo wspólne położenie geograficzne, albo zrzeszenie w międzynarodowym ugrupowaniu integracyjnym [11, s. 37].

1.5. Ryzyko według NUK

W literaturze przedmiotu można spotkać wiele podziałów ryzyka bankowego. Obecnie jednak podstawową klasyfikację ryzyka stworzył Bazylejski Komitet Nadzoru Bankowego – w Nowej Umowie Kapitałowej (Bazylea II), gdzie ryzyko podzielone jest na: kredytowe, rynkowe, stopy procentowej i inne rodzaje [5, s. 14].

Specyficzny rodzaj ryzyka bankowego, którym zarządzają banki został zaprezentowany w książce Shelagh Heffernan *Nowoczesna bankowość*. Występuje tutaj podział ryzyka na: kredytowe, kontrahenta, płynności, niewykonania zobowiązań lub rozliczeń. Ryzyko kredytowe określone zostało jako zagrożenie, że inwestycja w składnik aktywów lub udzielony kredyt nie zostaną odzyskane w razie niewypłacalności, a także ryzyko nieoczekiwanego spóźnienia w spłacie rat kredytowych. W sytuacji, w której dwie strony angażują się w umowę finansową mamy do czynienia z ryzykiem kontrahenta, które związane jest z niewywiązaniem się przez jedną ze stron z warunków umowy. Ryzyko płynności oznacza zagrożenie brakiem wystarczającej ilości środków pieniężnych na potrzeby działalności operacyjnej, tzn. niemożności uregulowania zobowiązań w momencie nadejścia terminu wymagalności. Ryzyko rozliczenia/płatności powstaje, gdy jedna ze stron nie uiszcza płatności lub dostarcza aktywa przed otrzymaniem należnej gotówki lub aktywów, narażając się na potencjalne straty [4, s. 129-130]. Ryzyko rynkowe związane jest najczęściej z obrotem papierów wartościowych. Banki mogą ponosić najczęściej ryzyko dotyczące akcji, papierów dłużnych, instrumentów pochodnych, towarów i walut. Ryzyko rynkowe obejmuje ryzyko walutowe i ryzyko stopy procentowej, ryzyko dźwigni finansowej oraz ryzyko operacyjne.

Reasumując, wyżej wspomniane rodzaje ryzyka funkcjonują we wszystkich elementach w działalności bankowej. Ogromne ilości rodzajów ryzyka bankowego i skutki przez nie implikowane powodują konieczność podjęcia kroków mających na celu ograniczanie ryzyk. Bardzo istotne jest zestawienie czynników wywołujących niebezpieczeństwa związane z funkcjonowaniem banku. Niżej zostały scharakteryzowane najbardziej istotne, występujące w polskim systemie bankowym.

2. Czynniki zewnętrzne wywołujące ryzyko bankowe

W działalności ekonomicznej, a zwłaszcza w działalności bankowej, nie można uniknąć ryzyka, gdyż w momencie podejmowania decyzji nie dysponuje się pełną informacją i nie zawsze można trafnie przewidzieć dalszy rozwój wydarzeń. Nie można też wykluczyć nieświadomych i świadomych zafałszowań

w informacjach oraz ich interpretacji. Wielkości ryzyka uzależnione są od czynników zewnętrznych oddziałujących na siebie. Do czynników zewnętrznych należą czynniki makroekonomiczne, społeczne, polityczne, demograficzne i techniczne. Do czynników makroekonomicznych zaliczamy politykę gospodarczą państwa, zadłużenie budżetu, stopę inflacji, wahania kursów walut, politykę banku centralnego i koniunkturę. Wśród czynników społecznych bardzo ważną rolę odgrywają postawy i zachowania klientów banku, skłonność do oszczędzania, rosnąca konkurencja. Czynniki polityczne silnie oddziaływają na wielkość ryzyka. Istotne znaczenie miał rozpad Związku Radzieckiego, zjednoczenie Niemiec, wojny, zmiany rządów, wypowiedzi polityków. Na czynniki demograficzne wpływ mają struktura ludności i stopa bezrobocia. Czynniki techniczne to w szczególności postęp w dziedzinie telekomunikacji i informatyki.

W książce Zofii Zawadzkiej i Władysława Jaworskiego *Bankowość podręcznik akademicki* najważniejsze czynniki wywołujące ryzyko bankowe zostały przedstawione jako:

- liberalizacja i deregulacja, czyli celowe znoszenie przez władze ograniczeń w funkcjonowaniu rynków finansowych, dopuszczając konkurencję na rynku krajowym banków z zagranicy,
- internacjonalizacja i globalizacja rynków finansowych powodujące większą zależność systemów finansowych poszczególnych krajów i niebezpieczeństwo przenoszenia kryzysów z jednego kraju do drugiego,
- rosnące deficyty budżetowe krajów rozwiniętych i konieczność ich finansowania na rynku finansowym,
- wzrastające zadłużenie krajów rozwijających się,
- sekurytyzacja, czyli zabezpieczenie należności papierami wartościowymi, w tym przede wszystkim sekurytyzacja bez udziału banku, polegająca na tym, że ryzyko i koszty są ponoszone wyłącznie przez depozytariusza i inwestora,
- zmiana zachowania i struktury klientów lokujących wkłady w banku, w tym przede wszystkim wzrastające znaczenie instytucjonalnych klientów banków, profesjonalnie zarządzających powierzonymi im funduszami,
- rozwój nowych rodzajów produktów bankowych, w tym pochodnych instrumentów finansowych,
- postęp techniczny, który poprzez rozwój informatyki i telekomunikacji zwiększył możliwości przetwarzania danych, przyśpieszył ich transfer i jednocześnie spowodował obniżkę kosztów jednostkowych transakcji. Szybki postęp techniczny przyczynił się także do skrócenia czasu życia produktów [14, s. 634-635],

– integracja ekonomiczna gospodarek, wynikająca z dążeń do integracji i międzynarodowego podziału pracy, także w sektorze finansowym,
– globalizacja rynków finansowych, polegająca na integracji narodowych i międzynarodowych rynków finansowych.

Inne przestawienie czynników ryzyka o charakterze zewnętrznym występuje w książce Zofii Zawadzkiej *Zarządzanie ryzykiem w banku komercyjnym*. Czynniki zewnętrzne można zakwalifikować do sześciu głównych grup; są to czynniki: makroekonomiczne, regulacyjne, globalizacyjne, popytowe, produktowe i inne zjawiska zewnętrzne. Do grupy czynników makroekonomicznych zaliczamy zmianę siły nabywczej pieniądza krajowego wynikającą z procesów inflacyjnych. Powoduje ona spadek realnej wartości aktywów banku. Spadek może być rekompensowany przez odpowiedni wzrost stóp procentowych. Zmiany wartości papierów rynkowych (np. kursy walutowe, rynkowe stopy procentowe) ich wahania mogą powodować spadek lub wzrost poziomu stóp procentowych lub wzrost wartości kontraktów zawartych z klientami. Zmiany parametrów rynkowych mogą prowadzić do obniżenia poziomu zysków, czego następstwem może być spadek wartości rynkowej kapitałów własnych banku. Rosnące deficyty budżetowe oraz zadłużenie krajów rozwijających się powoduje, iż banki finansujące te kraje stają w obliczu zagrożenia odzyskania zainwestowanych tam środków. Czynniki regulacyjne składają się ze zmian przepisów regulujących działalność banków odnoszących się do poziomu rezerwy obowiązkowej i regulacji ostrożnościowej oraz ze zmian przepisów podatkowych. Często spotykanym czynnikiem zewnętrznym zaliczanym do grupy czynników regulacyjnych jest deregulacja – polegająca na celowym znoszeniu przez władze legislacyjne ograniczeń administracyjnych w funkcjonowaniu rynków finansowych, co pozwala na rozszerzenie działalności banków i innych instytucji finansowych. Czynniki globalizacyjne umożliwiają podmiotom finansowym prowadzenie operacji na rynkach międzynarodowych. Zjawisko to może wywołać efekt domina, gdy pogorszenie sytuacji finansowej jednego banku może prowadzić do zakłóceń w funkcjonowaniu innych banków, a w skrajnym przypadku – do wywołania kryzysu o charakterze systemowym. Czynniki popytowe powodują wzrost znaczenia klientów indywidualnych. Wymaga to od banku większej troski o klienta, oferowanie usług oczekiwanych przez klientów, a czynniki produktowe wpływają na rozwój nowych produktów bankowych, instrumentów pochodnych. Inne zjawiska zewnętrzne przyczyniają się do wzrostu ryzyka, np. klęski żywiołowe, katastrofy czy kradzieże. Oddziałują nie tylko na działalność banków, ale towarzyszą wszystkim podmiotom prowadzącym działalność gospodarczą.

Wskazane wyżej czynniki ryzyka są w pełni odpowiednie i zasadne w kształtowaniu ryzyka bankowego. Poziom ryzyka z roku na rok stale wzrasta. Oprócz przyczyn wzrostu ryzyka bardzo ważną rolę odgrywają nowe tendencje

na rynku międzynarodowym. Współzależność pomiędzy poszczególnymi segmentami rynku finansowego sprawia, że impulsy pojawiające się na jednym rynku często przenoszone są na pozostałe segmenty.

3. Czynniki wewnętrzne wywołujące ryzyko bankowe

Aktualnie wśród wewnętrznych czynników należy wymienić czynnik ludzki i techniczny. Bardzo ważną rolę w czynniku ludzkim odgrywa poziom kwalifikacji pracowników banku, a także stosunek do obowiązków służbowych. Niski poziom kwalifikacji oraz niewłaściwe wykonywanie obowiązków służbowych stanowią poważny czynnik ryzyka w działalności bankowej. Na czynnik techniczny duży wpływ ma postęp techniczny. Poprzez rozwój informatyki i telekomunikacji, zwiększyła się szybkość przetwarzania danych i możliwa była obniżka kosztów jednostkowych transakcji bankowej, jednak spowodowało to osłabienie zabezpieczeń banku przed awariami technicznymi (np. awarią prądu), przestępcami włamującymi się do komputerowych baz danych czy niszczącym działaniem wirusów komputerowych [5, s. 11-13].

4. Przyczyny powstania ryzyka bankowego

Za najważniejsze przyczyny powstania ryzyka bankowego można uważać zmiany siły nabywczej pieniądza krajowego, jej zmniejszenie się w wyniku procesów inflacyjnych. Ubytek realnej wartości aktywów bankowych (lokat, kredytów), spłacanych przez dłużników w wysokości nominalnej niepodlegającej waloryzacji, oraz ubytek realnej wartości dochodów pieniężnych banków. Zmiany kursów walutowych powodują powstanie ryzyka kapitałowego w odniesieniu do składników majątkowych (aktywów bankowych) i zobowiązań banku, których wartość dla rozliczenia została ustalona w walutach obcych. Zmiany stóp procentowych na rynkach finansowych i w obrotach bankowych mogą wynikać ze (spadku lub wzrostu) wartości rynkowych papierów wartościowych. Wzrostowi stóp procentowych towarzyszy spadek wartości rynkowych papierów wartościowych, a spadek stóp procentowych powoduje wzrost wartości papierów i zmiany wysokości dochodów i wydatków banków, zmiany stawek obciążeń fiskalnych i parafiskalnych (stawek podatków opłat, składek z tytułu ubezpieczenia społecznego), zmiany te powodują ryzyko dochodowe, prowadząc do uszczuplenia nadwyżek finansowych banków, zmiany norm ostrożnościowych i innych norm regulujących działalność banków, ustalonych przez władze monetarne i stanowiące podstawę działania nadzoru bankowego, zmiany dotyczą wysokości stóp procentowych norm minimalnych rezerw płynności. Obowiązkowe rezerwy banków komercyjnych utrzymywane w Banku

Centralnym są oprocentowane. Występują także zmiany wskaźnika minimalnego wyposażenia kapitałowego banków oraz zmiany wskaźników ryzyka, ustalonych dla poszczególnych rodzajów aktywów bankowych [2, s. 15-16].

W literaturze naukowej zwraca się uwagę na zwiększenie się w kolejnych latach ryzyka w działalności bankowej. Oprócz ryzyka związanego z działalnością bankową, pojawiają się także nowe tendencje na międzynarodowych rynkach finansowych. Przyczyny wzrostu ryzyka w działalności bankowej powodują tendencję zwiększania niestabilności światowego systemu finansowego. Eksperci ostrzegają przed tzw. efektem domina, czyli sytuacją, w której załamanie jednego, nawet stosunkowo niewielkiego elementu w światowym systemie finansowym może spowodować zakłócenia w całym systemie [1, s. 26]. Wzajemnie powiązany międzynarodowy system finansowy staje się coraz bardziej wrażliwy na kryzys finansowy, który spowodował straty idące w miliardy dolarów oraz upadek czy nacjonalizację wielu banków globalnych. Kryzys ujawnił słabość obecnych metod, a jednocześnie uwypuklił różnice pomiędzy instytucjami, stosując różne podejścia do zarządzania ryzykiem. Kryzys pokazał także, że nadszedł czas na krytyczne spojrzenie na podstawy zarządzania ryzykiem. Zarządzanie ryzykiem powinno mieć charakter planowy i celowy. Działania powinny być podejmowane w sposób długofalowy i systematyczny. Aby prawidłowo zarządzać ryzykiem, należy je identyfikować, na bieżąco badać za pomocą różnych metod i zarządzać wskaźnikami [3, s. 50].

5. Identyfikacja i pomiar ryzyka

Potencjał zagrożeń istniejący w banku ustala się przez identyfikację ryzyka. Należy jednak pamiętać, że granica pomiędzy identyfikacją, analizą i oceną ryzyka oraz identyfikacją możliwych szans na sukces jest płynna. Zarządzający bankiem podejmują decyzję, czy ryzyko jest akceptowalne czy powinno być ograniczone lub przeniesione w całości na inne podmioty.

W ocenie ryzyka posługujemy się następującymi metodami:

- klasycznym zestawieniu niedopasowania pozycji (luka płynności, luka stopy procentowej, otwarta pozycja walutowa),
- metodą oceny ryzyka pojedynczego zaangażowania kredytowego (credit-scoring, analiza dyskryminacyjna, wewnętrzne ratingi ryzyka),
- modelami pomiaru określonego rodzaju ryzyka (rynkowego, kredytowego, operacyjnego).

W pierwszej grupie metod klasycznych można zestawić niedopasowanie pozycji, takie jak luka płynności (w tym urealniona o korekty terminów wpływów i wypływów określonych kwot pieniężnych), luka procentowa (gap), luka duration, niedopasowanie elastyczności czy otwarta pozycja walutowa. Metody

te pozwalają na oszacowanie poziomu różnych rodzajów ryzyka bankowego i w ograniczonym stopniu na aktywne zarządzanie nimi. Podstawową ich zaletą jest łatwość i interpretacji wyników, wadą jest to, że nie umożliwiają wyznaczania łącznej ekspozycji na ryzyko bankowe.

W drugiej grupie metod wewnętrzne ratingi ryzyka pozwalają na określenie ryzyka dla poszczególnych zaangażowań kredytowych banku w instytucje finansowe i podmioty gospodarcze na podstawie przyjętych kryteriów. Ratingi te służą do oceny stopnia ryzyka portfela kredytowego bądź jego poszczególnych segmentów oraz określenia wymaganych rezerw celowych. W systemie ratingów wewnętrznych może być również wykorzystywana metoda punktowa (credit scoring), jak też wiedza ekspercka, dzięki której koryguje się otrzymane wyniki.

Ostatnia grupa metod obejmuje modele do pomiaru określonego rodzaju ryzyka. Jako pierwszy pojawił się model ryzyka rynkowego – Risk Metrics, opublikowany przez amerykański bank inwestycyjny JP Mogran. Generalnie modele pomiaru ryzyka rynkowego szacują wartość zagrożoną (Daily earnings at risk-DEaR oraz value at risk-VaR), a więc pozwalają z określonym prawdopodobieństwem wyznaczyć wielkość maksymalnych strat banku z tytułu utrzymywania otwartych pozycji w ciągu najbliższych 24 godzin (DEaR) lub dla okresów dłuższych niż jeden dzień (VaR). Przy stosowaniu tych modeli zakłada się, że zmiany kursów i cen, które miały miejsce w przeszłości, jak również to, że pozycje, które znajdują się w momencie dokonywania szacunków, będą w portfelu przez określony czas.

Coraz popularniejsze stają się modele ryzyka kredytowego, które mają szacować ryzyko całego portfela kredytowego. Wyróżnia się dwa rodzaje modeli:

- default – służące szacowaniu strat kredytowych z określonym prawdopodobieństwem (np. KMV),
- mark-to-market – służące szacowaniu zmian wartości rynkowej portfela kredytowego (np. Credit Metrics).

Modele te nie są jeszcze powszechnie stosowane przez banki. Pomiar i zarządzanie ryzykiem to z jednej strony nauka, a z drugiej sztuka [7, s. 129-130]. Pomiar ryzyka jest dokonywany za pomocą różnych metod. W zależności od rodzaju ryzyka i wielkości zaangażowania banku, mogą być stosowane różne metody pomiaru, począwszy od prostych, opisowych, do skomplikowanych modeli ekonometrycznych. W wypadku ryzyka kredytowego do jego oceny można zastosować między innymi metody punktowe, credit scoring, wewnętrzne ratingi ryzyka. Do szacunków wielkości ryzyka stopy procentowej można wykorzystać między innymi metodę luki, duracje lub analizę elastyczności. Znacznie większe trudności nastręcza stosowanie ryzyka związanego ze stosowaniem przez banki instrumentów pochodnych, w szczególności instru-

mentów asymetrycznych. Zarządzanie tym ryzykiem wymaga posługiwania się rozbudowanymi modelami ekonometrycznymi. Oprócz miar zmienności (np. odchylenie standardowe), miar wrażliwości (np. zmodyfikowane duration), liter greckich (np. delta, gamma, rho) banki do zarządzania ryzykiem w coraz szerszym zakresie wykorzystują wewnętrzne modele ryzyka oparte na koncepcji Value at Risk. Bez stosowania rozbudowanych systemów pomiaru ryzyka trudno sobie wyobrazić działalność dużego banku. Na wynikach otrzymanych z modeli można polegać tylko o tyle, o ile prawidłowe są założenia i dane statystyczne przyjęte przy ich szacowaniu [6, s. 231-232].

Podsumowanie

Ryzyko bankowe może stwarzać zagrożenie dla prawidłowego funkcjonowania banku i powinno być ograniczane, kontrolowane oraz mierzone przez wewnętrzny sposób pomiaru ryzyka bankowego przy uwzględnieniu regulacji ostrożnościowych nadzoru bankowego. Ogromną rolę przy pomiarze banku stanowi doświadczenie własne i innych banków. Ryzyko bankowe jest bardzo trudne do oszacowania. Banki powinny stale monitorować zagrożenia. Rozpoznawanie i identyfikowanie zagrożeń ma istotny wpływ na funkcjonowanie banku, gdyż najczęściej banki odczuwają negatywne skutki dopiero po fakcie. Ignorowanie ryzyka może przyczynić się do kłopotów finansowych banku, a w najgorszych przypadkach do bankructwa. Należy pamiętać, że bank jest instytucją zaufania publicznego, dlatego agresywne podejście polegające na niwelowaniu ryzyka może narazić bank na zbyt duże ryzyko. Istnieje możliwość stosowania przez bank dynamicznego podejścia wykorzystanego do uzyskania dodatkowych korzyści z racji ponoszonego ryzyka. Działalność banków podlega regulacjom prawnym i nadzorowi ze strony uprawnionych do tego instytucji. Zakres regulowania działalności bankowej ulega i ewolucji. Podstawowe przyczyny tej ewolucji to występowanie nowych zjawisk w gospodarce, postęp techniczny, powstawanie nowych produktów i operacji bankowych, a także znaczny rozwój teorii finansów i bankowości. Jeszcze kilkanaście lat temu kierownictwo i zarząd banków kierowało swe działania na unikanie ryzyka. Obecnie standardem staje się akceptowanie ryzyka na pewnym poziomie, biznes przyjmuje na siebie ryzyko. Banki komercyjne podejmują obecnie ryzyko, oferując coraz to bardziej doskonałe produkty, odpowiadające potrzebom najbardziej wymagających klientów.

W Polsce funkcje organu nadzoru bankowego od 1.01.2008 pełni Komisja Nadzoru Finansowego, która sprawuje nadzór nad sektorem bankowym, rynkiem kapitałowym, ubezpieczeniowym, emerytalnym oraz instytucjami pieniądza elektronicznego. Celem Komisji Nadzoru Finansowego jest zapewnienie

bezpieczeństwa środków pieniężnych zgromadzonych na rachunkach bankowych oraz zgodność działalności banków z przepisami ustawy Prawo bankowe, ustawy o NBP, statutem oraz decyzją o wydaniu zezwolenia na utworzenie banku.

Regulacje ostrożnościowe nadzoru mają zapewnić systemowi bankowemu ochronę przed podejmowaniem zbyt dużego ryzyka, a tym samym – zwiększyć jego bezpieczeństwo. Ze względu na wrażliwość systemu bankowego, władze nadzorcze nie mogą przekazać całego procesu ograniczania ryzyka w ręce samych banków. Banki muszą opracować dodatkowo własne rozwiązania, które umożliwiają im kompleksową analizę i zarządzanie ryzykiem, wykraczające poza minimalne wymogi nadzoru bankowego. Rozwiązania te powinny obejmować zasady identyfikowania i pomiaru ryzyka, ograniczania eksplozji na ryzyko (wewnętrzne regulacje ostrożnościowe) [5, s. 18-19].

Czas kryzysu to najlepszy okres weryfikacji systemu zarządzania ryzykiem bankowym. Obecnie widoczny zastój na rynkach finansowych, dał wiele do myślenia bankowcom. Zastanawiali się nad błędami a także rozwiązaniami, w celu uniknięcia podobnych problemów w przyszłości. Zarządzanie ryzykiem jest najważniejszym zadaniem w działalności bankowej. Odnosi się do instrumentów, które lokalizują ryzyko i wypracowują rozwiązania mające na celu zminimalizowanie ryzyka. Zarządzanie ryzykiem w sposób niewłaściwy może przyczynić do zagrożenia wypłacalności banku. Zarządzanie ryzykiem bankowym obejmuje wszystkie rodzaje ryzyka bankowego, dlatego bardzo ważne jest rozpoznanie i jego klasyfikacja, a następnie zarządzanie nim.

Literatura

1. Dmowski A. (2006). Wpływ nowej umowy kapitałowej na pomiar i wymóg kapitałowy z tytułu ryzyka rynkowego w banku komercyjnym. PWSBiA, Warszawa.
2. Fedorowicz Z. (1996). Ryzyko bankowe. PWSBiA, Warszawa.
3. Graczyk B., Paskudzki G. (2009). Zarządzanie ryzykiem w czasach kryzysu. Bank, 7.
4. Heffernan S. (2007). Nowoczesna bankowość. PWN, Warszawa.
5. Iwanicz-Drozdowska M., Nowak A. (2002). Ryzyko bankowe. AGH, Warszawa.
6. Iwanicz-Drozdowska M., Zawadzka Z. (2006). Pojęcie i rodzaje ryzyka bankowego. W: Bankowość. Zagadnienia podstawowe. Poltex, Warszawa.
7. Iwanicz-Drozdowska M. (2005). Zarządzanie finansowe bankiem. PWE, Warszawa.

8. Jagiełło R. (2007). Charakterystyka, rodzaje i źródła ryzyka bankowego. W: Współczesna bankowość. T. I. Red. M. Zaleska. Difin, Warszawa.
9. Lenczewski-Martins C., Niedziółka P. (2005). Kwantyfikacja ryzyka operacyjnego w banku oraz jego wpływ na wymóg kapitałowy. Bank i Kredyt, 5.
10. Oleksiak K., Zawadzki A. (2002). Podstawy bankowości. WSUiB, Warszawa.
11. Przybylska-Kapuścińska W. (2001). Zarządzanie ryzykiem i płynnością banku komercyjnego. AE, Poznań.
12. Zaleska M. (2009). W dobie kryzysu. Bank, 3.
13. Zawadzka Z. (2004). Ryzyko bankowe. W: Bankowość podręcznik akademicki. Red. W.L. Jaworski, Z. Zawadzka. Poltex, Warszawa.
14. Zawadzka Z. (2002). Ryzyko bankowe – uwagi ogólne. W: Współczesny bank. Red. W.L. Jaworski. Poltex, Warszawa.
15. Słownik Języka Polskiego. (1989). PWN, Warszawa.
16. Żółtkowski W. (2007). Zarządzanie ryzykiem bankowym w praktyce. CeDeWu, Warszawa.

Jerzy Zemke
Uniwersytet Gdański

MODEL RYZYKA DECYZJI STRATEGICZNYCH ZARZĄDU ORGANIZACJI GOSPODARCZEJ

Zarządzanie strategiczne jest koniecznością wynikającą z postępującej koncentracji kapitału, dywersyfikacji organizacji, zauważalnej globalizacji działań, dynamiki procesów innowacyjnych czy rozbudowy baz danych. Cechy te są powodem wzrostu potencjału organizacji, ale jednocześnie wymuszają konieczność identyfikowania sygnałów napływających z otoczenia ale także monitorowania wnętrza organizacji [5, s. 28]. Te jakościowe zmiany wymagają także zmian jakości procesów decyzyjnych. Pożądane są decyzje trafne, natychmiastowe, elastyczne, o dużym zasięgu terytorialnym, wymagające dużych nakładów finansowych, jednocześnie wraz z tymi zmianami rośnie niepewność rezultatów realizowanych działań; stan ten określany jest jako zarządzanie w warunkach ryzyka.

Hipotezą główną pracy jest przypuszczenie, iż podjęte w procesach zarządzania ryzyko możemy ocenić monitorując skutki decyzji realizowanych działań. W konstrukcji dowodu hipotezy element centralny stanowi sformalizowany opis ryzyka – probabilistyczny model ryzyka do celów zarządzania organizacją gospodarczą. Opis powinien gwarantować istotną właściwość dla jakości procesów zarządzania, mianowicie możliwość pomiaru ryzyka. Rozwiązanie tego zadania wymagać będzie konstrukcji przestrzeni zdarzeń elementarnych, konstrukcji pewnego szczególnego σ ciała, które tworzą wszystkie podzbiory przestrzeni zdarzeń łącznie z całą przestrzenią oraz definicji odwzorowania o własnościach właściwych mierze prawdopodobieństwa.

1. Ryzyko w systemie zarządzania strategicznego organizacją

Zarządzanie organizacją jest układem współzależnych, zintegrowanych funkcji kierowniczych, podporządkowanych rozwiązywaniu istotnych jego zadań, decydujących o przetrwaniu na rynku sektora, na którym organizacja funkcjonuje, oraz dalszym jej rozwoju. Funkcje kierownicze generują podsystemy planowania, organizacji, kontroli i motywowania, wspomagane przez system informacyjny [5, s. 49].

Podsystem planowania porządkuje procesy formułowania strategii organizacji, na który składają się z fazy:

1) analizy makrootoczenia, otoczenia konkurencyjnego, analiza „wnętrza" organizacji,
2) formułowania strategii,
3) implementacji strategii,
4) kontroli i oceny [11].

Implementacja strategii łącznie z kontrolą i oceną ma wymiar wykonawczy, gdyż kojarzona jest z procesem podejmowania decyzji oraz ich realizacją. Sformułowaną strategię uzupełnia identyfikacja celów strategicznych, określonych w realizowanej strategii rozwoju. Przyjęte cele mają wzmacniać pozycję konkurencyjną organizacji na rynku sektora, w którym funkcjonuje.

Procesy decyzyjne związane są z realizacją celów strategicznych, które mogą dotyczyć zwiększenia udziału organizacji w sprzedaży rynku, celem może być także możliwość wpływu na poziom cen, ponadto, istotnym celem organizacji może być wpływ na poziom barier wejścia na rynek innych, konkurencyjnych podmiotów.

Czym jest zatem ryzyko towarzyszące procesem decyzyjnym? Większość definicji określa ryzyko jako uzyskanie wyniku innego aniżeli oczekiwany, definicja taka ogranicza zdecydowanie możliwości praktycznego pomiaru ryzyka. Jak zatem budować modele ryzyka, by możliwy był jego pomiar umożliwiający korekty realizowanych decyzji bądź wprost ich wstrzymanie wobec grożących nieodwracalnych negatywnych skutków dalszej kontynuacji procesów zarządzania? Odpowiedź zawarta jest w tym opracowaniu.

2. Procesy kontroli ryzyka. Zmienne kontrolowane skutków podejmowanych decyzji

Procesy zarządzania organizacją gospodarczą można postrzegać jako zbiór określonych decyzji i działań mających decyzje te zrealizować. Decyzje podejmowane są w określonych uwarunkowaniach, zarówno zewnętrznych w stosunku do organizacji, jak też wewnętrznych, a zatem już w momencie podjęcia decyzji podjęte zostaje także ryzyko tego, że w toku realizacji decyzji mogą wystąpić zmiany uwarunkowań, co spowoduje, iż planowane cele nie zostaną osiągnięte zarówno w kontekście jakościowym, jak też ilościowym. Można zatem przyjąć, że z procesem realizacji decyzji związane jest pewien stan ryzyka, który chcemy opisać bądź jakościowo, bądź ilościowo, chcemy w każdym momencie znać parametry tego stanu – miary ryzyka[1].

[1] Taka koncepcja ryzyka, łączy procesy decyzyjne z ryzykiem, wyróżnia pojęcie stanu ryzyka oraz miarę (miary) ryzyka, stąd miarę (miary) ryzyka postrzegać należy jako parametry stanu ryzyka.

Kontrola, obok planowania strategii rozwoju organizacji, stanowi drugi istotny element systemu zarządzania strategicznego. Definicje kontroli strategicznej eksponują zwykle jej podstawowe funkcje, którymi są:

- porównywanie przebiegów i wyników działań z przyjętymi standardami,
- identyfikacja i analiza odchyleń od oczekiwanych stanów (standardów) wraz z podejmowaniem działań korygujących nakierowanych na osiągnięcie planowanych wyników [5, s. 53].

Istotą kontroli strategicznej jest stały nadzór na procesami planowania strategicznego, sprawowany w kontekście ich wykonalności [7]. Kontekst to szczególny, gdyż dotyczy wczesnego rozpoznawania zagrożeń mogących powodować destabilizację warunków zarządzania strategicznego, stanowiąc potencjalne źródło zagrożeń realizacji strategii.

Kontrola strategiczna wyraża się poprzez analizę i ocenę wewnętrznych i zewnętrznych warunków realizacji strategii. Może to prowadzić do modyfikacji projektów strategicznych, co z kolei oznaczać będzie konieczność korekty przedsięwzięć służących implementacji strategii organizacji.

Kontrola strategiczna odbywa się, poprzez stałe monitorowanie realizowanej strategii i prowadzenie na podstawie danych monitoringu analiz oraz ocen wewnętrznych i zewnętrznych uwarunkowań strategii w wymiarze codziennej praktyki zarządzania. Kontrola dotyczy:

- kontroli zasobów rzeczowych; związana jest z zarządzaniem zapasami, zarządzaniem jakością oraz utrzymaniem majątku produktywnego,
- kontroli zasobów ludzkich; związana jest z polityką personalną organizacji, obejmuje kontrolę zasad przyjęć oraz obsady stanowisk pracy, kontrolę szkoleń i rozwoju kadry zarządzającej, kontrolę wyników pracy oraz progresji wynagrodzeń,
- kontroli zasobów informacyjnych, obejmuję analizę otoczenia strategicznego organizacji, kontrolę realizacji prognoz ekonomicznych, kontrolę zgodności procesów produkcji z harmonogramami, kontrolę postrzegania organizacji – public relations,
- kontroli finansowej, która obejmuje analizę stanu ryzyka związanego z zarządzaniem długoterminowym, analizuje dynamikę zmian relacji pomiędzy zobowiązaniami krótko i długoterminowymi, podobnie analizuje dynamikę relacji pomiędzy należnościami krótko i długoterminowymi (kontrolę finansową postrzeganą w kontekście ryzyk decyzji i działań o zasięgu strategicznym uzupełnia kontrola bieżącego zarządzania zasobami kapitału obrotowego – skutki ryzyka operacyjnego zarządzania zasobami finansowego stanowią same w sobie zagrożenia dla procesów długoterminowego zarządzania zasobami finansowymi),

– kontroli parametrów makrootoczenia (ekonomicznych, prawnych, finansowych, środowiskowych)[2].

Kontrola strategiczna nie jest procesem jednorodnym, wyraźnie rysuje się nim kilka odrębnych faz:

1) określenie wielkości oczekiwanych i standardów działań,

2) określenie metod pomiaru i monitoringu,

3) porównywanie stanów, wielkości rzeczywistych strategii i oczekiwanych zmiennych kontrolowanych,

4) ustalenie przyczyn odchyleń, ustalenie stopnia w jakim zarząd organizacji kontroluje źródła i przyczyny odchyleń,

5) projektowanie i wdrażanie projektów do działań korygujących [5, s. 87].

Określenie standardów działań w procesie kontroli strategicznej jest równoznaczne między innymi z wyborem zmiennych kontrolnych. Stopień realizacji celów jest kontrolą wyników, które pozwalają ocenić sprawność funkcjonowania organizacji. Proces kontroli w takim wymiarze odbywa się poprzez porównanie zmierzonych w procesie monitoringu wyników, z założeniami zdefiniowanymi w strategii organizacji [10, s. 239-255].

Cele organizacji zdefiniowane są w założeniach planu, w tej jego części, która identyfikuje cele strategiczne organizacji gospodarczej. Ich definicja powiązana jest z identyfikacją zmiennych kontrolowanych procesu kontroli realizacji planów.

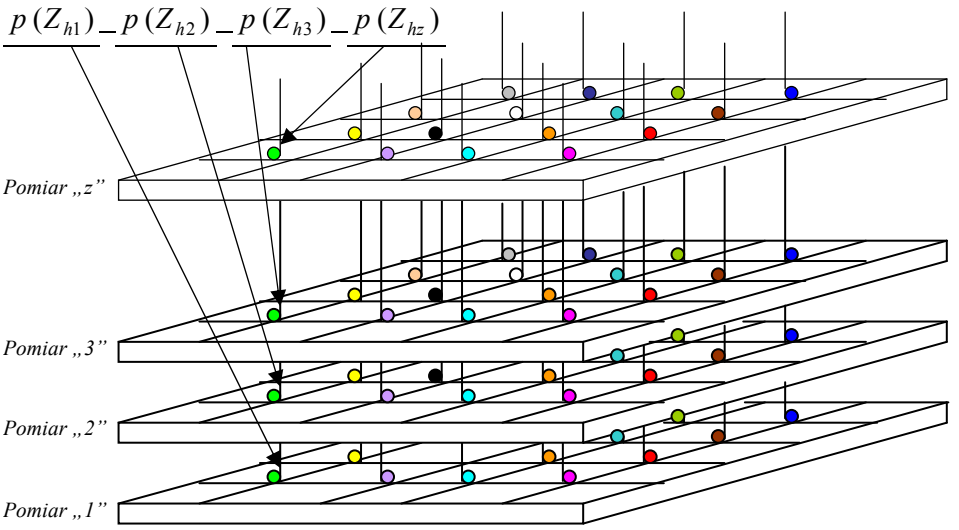
Niech $\boldsymbol{Z} = \{\boldsymbol{Z_1}, \boldsymbol{Z_2}, \ldots, \boldsymbol{Z_r}\}$ oznacza skończony zbiór zmiennych kontrolowanych zidentyfikowanych w planie strategicznym organizacji, moc zbioru *Z* jest równa $r \in \mathbb{C}$, *r* przyjmuje skończoną wartość całkowitą (rys. 1). Dla dowolnej wartości *h* takiej, że *h = 1,2, ... ,r*, $\boldsymbol{p}(\boldsymbol{Z_{hl}})$ jest miarą zmiennej kontrolowanej $\boldsymbol{Z_h}$ w *l- tym* cyklu procesu monitoringu, gdzie *l = 1,2, ... , z*[3].

Zbiór miar – punktów próbkowych – $\{\boldsymbol{p}(\boldsymbol{Z_{h1}}), \boldsymbol{p}(\boldsymbol{Z_{h2}}), \ldots \boldsymbol{p}(\boldsymbol{Z_{hz}})\}$ jest wynikiem pomiarów w procesach kontroli realizacji strategii *h-tej* zmiennej kontrolowanej, tym samym miary zmiennej $\boldsymbol{Z_h}$ identyfikują zdarzenie $\boldsymbol{Z_h} \subset \boldsymbol{\Omega}$.

---

[2] Struktura zbioru celów kontrolowanych jest zbiorem oczekiwań właścicieli, opartych na akceptowanych przez nich założeniach strategii, są związane z umacnianiem pozycji na rynku sektora, osiągnięciem oczekiwanego poziomu wskaźników finansowych, świadczących o wiarygodności organizacji a także dających satysfakcję finansową właścicielom.

[3] Teoretycznie $l \rightarrow \infty$, w praktyce zarządzania *l = 1,2, ... , z*, gdzie $z \in \mathbb{C}$, *z* przyjmuje skończoną wartość całkowitą. Rysunek 1 ilustruje zasady definicji przestrzeni zdarzeń elementarnych $\boldsymbol{\Omega}$, z oznaczoną w strategii częstotliwością, dokonywane są pomiary zmiennych kontrolowanych $\{\boldsymbol{Z_{hl}}\}$, w rezultacie po zrealizowaniu „*z*" sesji pomiaru, baza pomiarów zmiennych kontrolowanych definiowana jest poprzez ciągi zdarzeń $\boldsymbol{p}(\boldsymbol{Z_{hl}})$, tzn. $\boldsymbol{\Omega} = \{\boldsymbol{p}(\boldsymbol{Z_{hl}})\}$.



Rys. 1. Przestrzeń Ω zdarzeń elementarnych

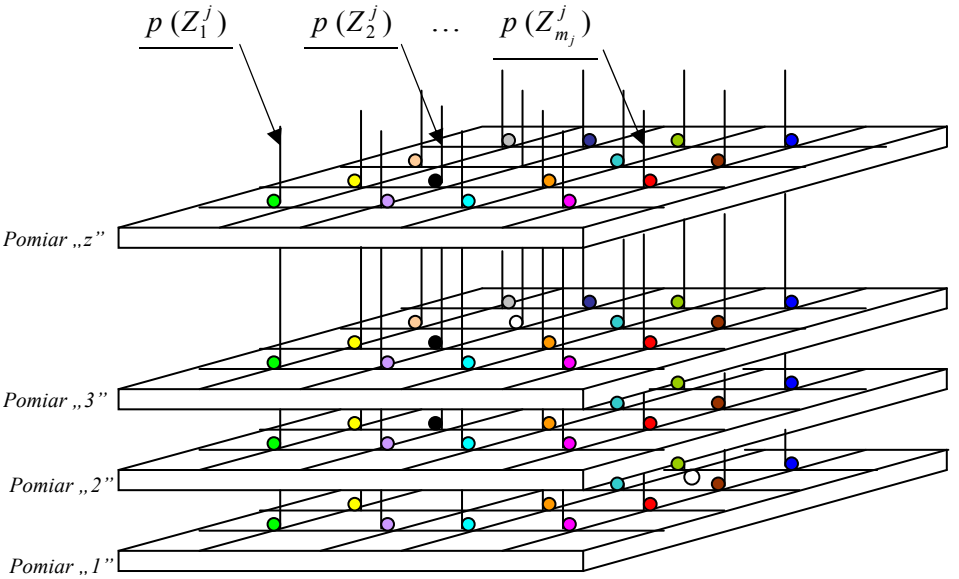
Przestrzeń zdarzeń Ω jest skończonym zbiorem zdarzeń, gdyż ich liczbę determinują zidentyfikowane w planach strategicznych organizacji zmienne kontrolowane. Jakakolwiek zmiana założeń strategii w trakcie jej realizacji, jeśli zakłada także zmianę zmiennych, wymagać będzie odpowiednio do dokonanych zmian, redefinicji przestrzeni zdarzeń.

3. Przestrzeń ryzyka, elementy przestrzeni, miara na elementach przestrzeni ryzyka

Stan ryzyka $\boldsymbol{R}^j$ decyzji i działań procesów zarządzania organizacją, gdzie $j = 1,2, \ldots,n$, w wielu przypadkach określa więcej niż jedna zmienna kontrolowana $\boldsymbol{Z}_k^j$, gdzie $k = 1,2,\ldots,m^j$, w rezultacie stan ryzyka $\boldsymbol{R}^j$ definiuje zbiór $\mathcal{F}^j$ zdarzeń $\{\boldsymbol{Z}_k^j\}$ (rys. 2)[4]. Do zbioru $\mathcal{F}^j$ należą zdarzenia przestrzeni $\boldsymbol{\Omega}$, zatem dokonując możliwego uogólnienia nie wykluczającego przypadku, iż wszystkie rodzaje ryzyka decyzji i działań są określane przez więcej niż jedną

[4] Zakładamy, że istnieje odwzorowanie $\boldsymbol{g}^j : \boldsymbol{R}^j \rightarrow \mathcal{F}$, które określa stan ryzyka $\boldsymbol{R}^j$ poprzez zbiór $\{\boldsymbol{p}(\boldsymbol{Z}_1^j), \boldsymbol{p}(\boldsymbol{Z}_2^j), \ldots \boldsymbol{p}(\boldsymbol{Z}_{m_j}^j)\} \subset \mathcal{F}$.

zmienną kontrolowaną, można wnioskować, że zbiór $\mathcal{F} = \bigcup_{j=1}^{n} \mathcal{F}^j$ jest przestrzenią wszystkich podzbiorów $\mathcal{F}^j = \{p(Z_k^j)\}$, gdzie $p(Z_k^j)$ jest miarą *k-tej* zmiennej odwzorowania g^j.



Rys. 2. Przestrzeń stanów ryzyk decyzji i działań w procesach zarządzania strategicznego

Definicja 1. Niech Ω jest niepustym zbiorem zdarzeń elementarnych, a zbiór $\mathcal{F}$ rodziną podzbiorów Ω spełniającą warunki:

1) $\Omega \in \mathcal{F}$,

2) $\emptyset \in \mathcal{F}$,

3) jeśli $A \in \mathcal{F}$, to $A' \in \mathcal{F}$ (A' – dopełnienie zbioru A w przestrzeni Ω, $A' = \Omega - A$),

4) jeśli $A_1, A_2, \ldots \in \mathcal{F}$, to $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Rodzinę $\mathcal{F}$ podzbiorów zbioru Ω nazywamy σ – ciałem zdarzeń, a elementy tej rodziny określane są zdarzeniami losowymi.

Niech Ω jest zbiorem zdarzeń elementarnych, a $\mathfrak{R}$ jest zbiorem liczb rzeczywistych, oznaczmy zbiór wszystkich odcinków $\mathcal{B}$, zbiór ten definiuje najmniejsze σ – ciało zawierające te odcinki.

Definicja 2. Jeżeli tożsamościowo równe są zbiory Ω oraz $\mathfrak{R} / \Omega \equiv \mathfrak{R}$ to zbiór $\mathfrak{B}$ jest σ – ciałem $\mathfrak{B}$ zbiorów borelowskich.

Jeśli funkcja P: $\mathfrak{F} \rightarrow \mathfrak{R}$ jest odwzorowaniem mającym pewne szczególne własności; przyjmuje wartości z określonego przedziału domkniętego, określona jest miara przestrzeni wszystkich zdarzeń elementarnych Ω oraz wartość funkcji P na sumie podzbiorów zbioru $\mathfrak{F}$ jest sumą wartości miar na składnikach sumy, to funkcja P jest pewną szczególną miarą określoną na zbiorze $\mathfrak{F}$.

Definicja 3. Jeżeli funkcja P: $\mathfrak{F} \rightarrow \mathfrak{R}$, spełnia poniższe trzy warunki:

1) dla dowolnego zbioru $A \in \mathfrak{F}$ zachodzi $0 \leq P(A) \leq 1$,
2) $P(\Omega) = 1$,
3) jeśli podzbiory zdarzeń A_i, gdzie $i = 1,2, \ldots$ spełniają warunek $A_i \cap A_j = 0$ dla $i \neq j$, gdzie $j = 1,2, \ldots$, to $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, to funkcja P

jest miarą określaną prawdopodobieństwem zajścia dowolnego zdarzenia losowego A z przestrzeni zdarzeń[5].

Definicja 4. Niech Ω oznacza przestrzeń zdarzeń procesów kontroli strategicznej[6]. Trójka elementów (Ω, $\mathfrak{F}$, P) definiuje przestrzeń probabilistyczną wtedy, tylko wtedy, gdy Ω jest zbiorem (przestrzenią) zdarzeń elementarnych – miar zmiennych kontrolowanych, $\mathfrak{F}$ jest pewnym szczególnym zbiorem (σ – ciałem zbiorów borelowskich rozpiętym nad przestrzenią Ω), natomiast P jest funkcją określoną na zbiorze $\mathfrak{F}$, przyjmującą wartości ze zbioru liczb rzeczywistych[7].

3.1. Wektor losowy

Niech (Ω, $\mathfrak{F}$, P) oznacza przestrzeń probabilistyczną a ($\mathfrak{R}^k$, $\mathfrak{B}^k$) przestrzeń mierzalną w k wymiarowej przestrzeni euklidesowej $\mathfrak{R}^k$, gdzie $\mathfrak{B}^k$ jest σ – ciałem podzbiorów borelowskich w tej przestrzeni, $k \geq 1$.

Wniosek: elementami przestrzeni stanów ryzyka procesów decyzyjnych są wektory postaci: $\{p(Z_1^j), p(Z_2^j), \ldots p(Z_{m_j}^j)\}$, gdzie j oznacza ryzyko R^j, a współrzędne $p(Z_k^j)$ są zmiennymi losowymi aproksymującymi zmienne kontrolowane ryzyka.

---

[5] Aksjomatyczna definicja prawdopodobieństwa.

[6] Gdy przestrzeń Ω ma skończoną lub przeliczalną liczbę elementów, wówczas każdy podzbiór Ω jest zdarzeniem losowym (teza ta nie jest zawsze prawdziwa, gdy Ω nie jest przeliczalna).

[7] Własności rodziny wszystkich odcinków prostej $\mathfrak{R}$ opisał jako pierwszy E. Borel.

Definicja 5. Wektorem losowym k-wymiarowym, nazywamy funkcję $X: \Omega \to \Re^k$ mierzalną względem σ-ciała $\mathcal{F}$ ($\mathcal{F}$ – mierzalną), tzn. taką, że $X^{-1}(\mathbf{B}) \in \mathcal{F}$ dla każdego $B \in \mathcal{B}^k$ [8].

Miary wektora losowego

Definicja przestrzeni zdarzeń Ω oraz σ – ciała zbiorów borelowskich $\mathcal{F}$ identyfikuje przestrzeń ryzyk decyzji i działań zarządu organizacji, natomiast funkcja P, określona na zbiorze $\mathcal{F}$, pozwala elementy tej przestrzeni mierzyć. Definicja miary P „otwiera" zasoby instrumentów opisu ilościowego elementów σ – ciała $\mathcal{F}$, pod warunkiem definicji takiego odwzorowania, które elementom przestrzeni zdarzeń Ω przyporządkowuje skończone ciągi zdarzeń elementarnych należących do $\Re^k$. Tak określona funkcja, jeśli jest tylko mierzalna względem σ – ciała $\mathcal{F}$, definiuje wektor losowy będący probabilistycznym modelem ryzyka. Funkcja gęstości prawdopodobieństwa składowych wektora losowego, pozwala zdefiniować system miar wektora[9]. Elementami systemu są funkcje charakterystyczne: wektor wartości oczekiwanych składowych, macierz wariancji i kowariancji składowych wektora, rozkłady warunkowe:

Definicja 6. Prawdopodobieństwo zdarzeń, że składowe losowe $\{X_j^i\}$ przyjmą wartości x_j^i należące do przedziału $[a_j^i, b_j^i]$, jest równe

$$P(a_1^i \le X_1^i \le b_1^i, \ldots, a_k^i \le X_k^i \le b_k^i) = \int_{a_1^i}^{b_1^i} \ldots \int_{a_k^i}^{b_k^i} f^i(x_1^i, \ldots, x_k^i)\, dx_1^i, \ldots, dx_k^i$$ [10]

---

[8] Mierzalność odwzorowania X oznacza, że X: $(\Omega, \mathcal{F}) \to (\Re^k, \mathcal{B}^k)$. Wektory losowe oznaczamy dużymi literami, np. X, Y, Z itd., podobnie zmienne losowe: X, Y, Z itd., natomiast wartości zmiennych losowych i wektorów losowych oznaczamy małymi literami, np. $x_i = X_i(\omega)$, $\omega \in \Omega$. Współrzędne $X_i(\omega)$, $gdzie$; $i = 1,\ldots,k$, są wartościami zmiennych losowych X_i : $(\Omega, \mathcal{F}) \to (\Re, \mathcal{B})$, są współrzędnymi wektora losowego $X = (X_1, \ldots, X_k)$.

[9] Rozkład gęstości prawdopodobieństwa jest odwzorowaniem, które zbiorom A przyporządkowuje pewną miarę $F\{A\} \ge 0$, spełniająca warunek przeliczalnej addytywności oraz warunek unormowania $F\{-\infty, \infty\} = 1$, gdzie zbiory A, to zbiory borelowskie, spełniające warunek $F\{A\} = \sum_k F\{A_k\}$, który oznacza, że dla każdego zbioru A składającego się ze skończenie lub przeliczalnie wielu rozdziałów rozłącznych A_k, miara F na zbiorze A jest przeliczalnie addytywna. Odwzorowanie F jest funkcją Baire'a, jest to klasa funkcji ciągłych $\mathfrak{A}$, która spełnia dwa warunki:

1) $\mathfrak{A}$ jest zbiorem funkcji ciągłych $\{F_n\}$,

2) istnieje $\lim_{n\to\infty} F_n(A) = F(A)$ dla dowolnego A oraz $F(A)$ należy także do $\mathfrak{A}$ [4, s. 102-124].

[10] f^i oznacza funkcję gęstości rozkładu prawdopodobieństwa wektora losowego.

Definicja 7. Wartość oczekiwana wektora losowego

$$E(X^i) = (E(X_1^i), E(X_2^i), ..., E(X_k^i))^{11}, \text{ gdzie } E(X_j^i) = \int_{a_j^i}^{b_j^i} x_j^i \, f^i(x_1^i, ..., x_k^i)\, dx_j^i ,$$

$$j = 1,2,..,k$$

Definicja 8. Wariancja wektora losowego: $Var(X^i) = (Var(X_1^i), Var(X_2^i), ..., Var(X_k^i))$,

gdzie $Var(X_j^i) = \int_{a_j^i}^{b_j^i} [x_j^i - E(X_j^i)]^2 \, f^i(x_1^i, x_2^i, ..., x_k^i)\, dx_j^i$, $j = 1,2,..,k$

Definicja 9. Kowariancje składowych losowych $\{X_l^i, X_j^i\}$ są równa

$$Cov(X_l^i, X_j^i) = \int_{a_l^i}^{b_l^i} \int_{a_j^i}^{b_j^i} [x_l^i - E(X_l^i)]\,[x_j^i - E(X_j^i)]\, f^i(x_1^i, x_2^i, ..., x_k^i) dx_l^i \, dx_j^i ,$$

$$l, j = 1,2,..,k \;\wedge\; l \neq j$$

Definicja 10. Jeżeli relacja $R^i(X_1^i, X_2^i, ..., X_k^i)$ jest ciągła, to prawdopodobieństwo warunkowego rozkładu wektora ryzyka R^i, względem składowej losowej $\{X_j^i\}$, przy warunku $X_j^i = x$ jest równa

$$P\{(a_1^i < x_1^i \le b_1^i ..., a_{j-1}^i < x_{j-1}^i \le b_{j-1}^i, a_{j+1}^i < x_{j+1}^i \le b_{j+1}^i, ..., a_k^i < x_k^i \le b_k^i) \,|\, (x_j^i = x)\} = H$$

$$H = \int_{a_1^i}^{b_1^i} ... \int_{a_{j-1}^i}^{b_{j-1}^i} \int_{a_{j+1}^i}^{b_{j+1}^i} ... \int_{a_k^i}^{b_k^i} \frac{f^i(x_1^i, x_2^i, ..., x_k^i)}{f^i(x_j^i)} dx_1^i ... dx_{j-1}^i \, dx_{j+1}^i ... dx_k^i , \; j = 1,2,..,k$$

Definicja 11. Warunkowa wartość oczekiwana wektora ryzyka R^i przy warunku $X_j^i = x$ jest wektorem

$$E(R^i \,|\, X_j^i = x) = (E(X_1^i \,|\, (X_j^i = x)), ..., E(X_{j-1}^i \,|\, (X_j^i = x)), ..., E(X_k^i \,|\, (X_j^i = x)))$$

gdzie

$$E(X_l^i \,|\, (X_j^i = x)) = \int_{a_j^i}^{b_j^i} x_j^i \frac{f^i(x_1^i, x_2^i, ..., x_k^i)}{f^i(x_j^i)} dx_j^i \; ; \; l = 1,2,\ldots,k \;\wedge\; l \neq j$$

[11] Wartość oczekiwana wielowymiarowej zmiennej losowej jest wektorem o składowych $E(X_j^i)$

Definicja 12. Warunkowa wariancja wektora ryzyka $\boldsymbol{R}^i$ przy warunku $\boldsymbol{X}_j^i = \boldsymbol{x}$ jest wektorem

$$\boldsymbol{Var}(\boldsymbol{R}^i \,\big|\, \boldsymbol{X}_j^i = \boldsymbol{x}) = (\boldsymbol{Var}(\boldsymbol{X}_1^i \,\big|\, (\boldsymbol{X}_j^i = \boldsymbol{x})),...,\boldsymbol{Var}(\boldsymbol{X}_{j-1}^i \,\big|\, (\boldsymbol{X}_j^i = \boldsymbol{x})),...,\boldsymbol{Var}(\boldsymbol{X}_k^i \,\big|\, (\boldsymbol{X}_j^i = \boldsymbol{x})))$$

gdzie

$$\boldsymbol{Var}\,(\boldsymbol{X}_l^i \,\big|\, (\boldsymbol{X}_j^i = \boldsymbol{x})) = \int_{a_l^i}^{b_l^i} [\boldsymbol{x}_l^i - \boldsymbol{E}(\boldsymbol{X}_l^i \,\big|\, \boldsymbol{X}_j^i = \boldsymbol{x})]^2 \frac{\boldsymbol{f}^i(\boldsymbol{x}_1^i, \boldsymbol{x}_2^i, ..., \boldsymbol{x}_k^i)}{\boldsymbol{f}^i(\boldsymbol{x}_j^i)} \boldsymbol{dx}_l^i \text{ , gdzie}$$

$$\boldsymbol{l = 1,\ldots,k} \;\wedge\; \boldsymbol{l \neq j}$$

Definicja 13. Warunkowe kowariancje składowych $\{\boldsymbol{X}_l^i, \boldsymbol{X}_h^i\}$ *i-tego* wektora ryzyka $\boldsymbol{R}^i$ przy warunku $\boldsymbol{X}_j^i = \boldsymbol{x}$ oraz dla $l,\ h = 1,\ldots,k \;\wedge\; \boldsymbol{l \neq j}, \boldsymbol{h \neq j}$ są równe

$$\boldsymbol{Cov}((\boldsymbol{X}_l^i, \boldsymbol{X}_h^i) \big| (\boldsymbol{X}_j^i = \boldsymbol{x})) = \int_{a_l^i}^{b_l^i} \int_{a_h^i}^{b_h^i} [\boldsymbol{x}_l^i - \boldsymbol{E}(\boldsymbol{X}_l^i \,\big|\, (\boldsymbol{X}_j^i = \boldsymbol{x}))]\,[\boldsymbol{x}_h^i - \boldsymbol{E}(\boldsymbol{X}_h^i \big| (\boldsymbol{X}_j^i = \boldsymbol{x}))] \frac{\boldsymbol{f}^i(\boldsymbol{x}_1^i, \boldsymbol{x}_2^i, ..., \boldsymbol{x}_k^i)}{\boldsymbol{f}^i(\boldsymbol{x}_j^i)} \boldsymbol{dx}_l^i \, \boldsymbol{dx}_h^i$$

Przykład. Ryzyko obsługi zobowiązań długoterminowych wynikłych z realizacji planów strategicznych wystawia organizację na niekorzystne dla jej funkcjonowania skutki, które mogą być spowodowane:

a) zbyt niskim zasobem wolnych środków finansowych,

b) niewłaściwą strukturą zadłużenia,

c) niekorzystną tendencją zmian rynkowych stóp procentowych.

Proces kontroli skutków ryzyka decyzji i działań związanych z finansowaniem działalności organizacji gospodarczej kapitałem pożyczonym zakłada monitorowanie stanu podjętego ryzyka. Teoria zarządzania wyznacza szczególną funkcję tzw. zmiennym kontrolowanym procesu monitoringu aproksymujące skutki podjętego ryzyka [15, s. 80].

Niech $\boldsymbol{F} = \{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \boldsymbol{Z}_3, \boldsymbol{Z}_4\}$, gdzie: $\boldsymbol{Z}_1 = \boldsymbol{WACC}$, $\boldsymbol{Z}_2 = \boldsymbol{OZ}$, $\boldsymbol{Z}_3 = \boldsymbol{D}$, $\boldsymbol{Z}_4 = \boldsymbol{W}_w$, oznacza zbiór zmiennych kontrolowanych procesu obsługi zobowiązań długoterminowych. Zmienna $\boldsymbol{WACC}$ określa średni ważony koszt kapitału własnego, a ponadto zmienna $\boldsymbol{OZ}$ oznacza okres zwrotu z inwestycji, zmienna $\boldsymbol{D} = \boldsymbol{S}_K - \boldsymbol{S}_{IRR}$ jest różnicą pomiędzy stopami kryterialną $\boldsymbol{S}_K$ i wewnętrzną stopą

zwrotu z inwestycji S_{IRR}, natomiast $W_w = \frac{WS_p}{Z}$ jest miarą relacji pomiędzy stanem wolnych środków pieniężnych do stanu zobowiązań Z (łącznie długoterminowych + bieżące)[12].

Niech f oznacza funkcję gęstości prawdopodobieństwa składowych wektora losowego[13]

$$f(z_1,z_2,z_3,z_4) = \frac{1}{(2\pi)^{\frac{4}{2}} * \sqrt{|M|} * \prod_{l=1}^{4} z_h} \exp\{-\frac{1}{2}\sum_{g=1}^{4}\sum_{h=1}^{4}\eta_{gh}(\ln z_g - \mu_g)(\ln z_h - \mu_h)\}$$

gdzie: $M = \begin{pmatrix} \sigma_1^2 & \eta_{12} & \eta_{13} & \eta_{14} \\ \eta_{21} & \sigma_2^2 & \eta_{23} & \eta_{24} \\ \eta_{31} & \eta_{32} & \sigma_3^2 & \eta_{34} \\ \eta_{41} & \eta_{42} & \eta_{43} & \sigma_4^2 \end{pmatrix}$, σ_h^2 oznacza wariancję zmiennej losowej z_h, a η_{gh} kowariancje zmiennych $\{z_g, z_h\}$, gdzie: $g, h = 1,2,3,4$, jak oszacować elementy macierzy M? Odpowiedzi na to istotne pytanie, przesądzające o znaczeniu praktycznym przedstawionego w tym opracowaniu rozwiązania, należy szukać w tej jego części, w której przedstawiona została konstrukcja przestrzeni zdarzeń elementarnych Ω. Realizacja założeń kontroli wyraża się, monitorowaniem zmiany zmiennych kontrolowanych stanów ryzyka, w rezultacie definiujemy zbiór Z, który tworzą zmienne kontrolowane zmian stanu ryzyka. Elementy tego zbioru można uporządkować tak, że dowolny element z_{ih} oznacza wartość zmiennej Z_h w i „sesji" pomiaru zmiennych kontrolowanych

$$Z = \begin{pmatrix} z_{11} & z_{12} & z_{13} & z_{14} \\ z_{21} & z_{22} & z_{23} & z_{24} \\ \dots & \dots & \dots & \dots \\ z_{n1} & z_{n2} & z_{n3} & z_{n4} \end{pmatrix}$$

gdzie n oznacza liczbę zrealizowanych „sesji" pomiaru zmiennych kontrolowanych w procesie kontroli podjętych decyzji i związanego z tym ryzyka. W każdej „sesji" kontroli procesów zarządzania, w oparciu o zbiór pomiarów $\{z_{ih}\}$, gdzie $h = 1,2,3,4,\ i = 1,2,\dots,n$ definiowane są granice zmienności $[a_h, b_h]$

---

[12] $WACC = \frac{K_w}{K_w + K_0} k_e + \frac{K_0}{K_w + K_0} k_d (1 - t_c)$ (Weighted Average Cost of Capital), gdzie: K_w – *kapitał własny,* K_o – *kapitał obcy,* k_e – *koszt kapitału własnego,* k_d – *koszt kapitału obcego,* t_c – *stopa podatku dochodowego.*

[13] Funkcja rozkładu logarytmiczno-normalnego.

zmiennej Z_h. Jednocześnie na podstawie pomiarów $\{z_{ih}\}$ tworzących bazę informacji o zmiennych kontrolowanych można oszacować elementy macierzy M oraz parametry μ_g i μ_h funkcji f [14].

Przyjęcie założenia o postaci funkcji gęstości rozkładu prawdopodobieństwa, pozwala zdefiniować zbiór miar wektora. Elementami tego zbioru są funkcje charakterystyczne: prawdopodobieństwo tego, że składowe przyjmą wartości z określonych przedziałów liczbowych, wektor wartości oczekiwanych składowych, macierz wariancji i kowariancji składowych wektora, rozkłady warunkowe:

1. Prawdopodobieństwo tego, że zmienne losowe $\{Z_h\}$ przyjmą wartości $z_h \in [a_h, b_h]$, tzn. $P(a_1 \le z_1 \le b_1, a_2 \le z_2 \le b_2, a_3 \le z_3 \le b_3, a_4 \le z_4 \le b_4) = P$, gdzie P prawdopodobieństwo zajścia takiego zdarzenia, obliczamy jako

$$P = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\int_{a_3}^{b_3}\int_{a_4}^{b_4} f(z_1, z_2, z_3, z_4)\, dz_1\, dz_2\, dz_3\, dz_4$$

2. Wartość oczekiwana wektora losowego

$E(Z) = (E(Z_1), E(Z_2), E(Z_3), E(Z_4))$ [15], gdzie $E(Z_h) = \int_{a_h}^{b_h} z_h\ f(z_1, z_2, z_3, z_4)\, dz_h$,

gdzie $h = 1,2,3,4$

3. Wariancja zmiennej losowego $\{Z_h\}$

$$Var(Z) = (Var(Z_1), Var(Z_2), Var(Z_3), Var(Z_4))$$

gdzie: $Var(Z_h) = \int_{a_h}^{b_h} [z_h - E(z_h)]^2\ f(z_1, z_2, z_3, z_4)\, dz_h$, $h = 1,2,3,4$

4. Kowariancja zmiennych losowych $\{Z_i, Z_j\}$

$$Cov(Z_i, Z_j) = \int_{a_i}^{b_i}\int_{a_j}^{b_j} [z_i - E(z_i)]\,[z_j - E(z_j)]\ f(z_1, z_2, z_3, z_4)\, dz_i\, dz_j, \quad i, j = 1,2,3,4$$

$$\wedge \quad i \neq j$$

[14] Są wartościami parametrów początkowych $\{\mu, \sigma, \eta\}$ funkcji f oszacowanymi na próbie [1, n], odpowiadającej zrealizowanej ***n***-tej „sesji" procesu kontroli rezultatów podjętych decyzji. W kolejnej ***n+1*** „sesji" wartości parametrów $\{\mu, \sigma, \eta\}$ funkcji gęstości f szacowane są już w oparciu o wyznaczone w „sesji" ***n*** wartości oczekiwane składowych $E(Z_h)$, wariancje $Var(Z_h)$ oraz kowariancje $Cov(Z_g, Z_h)$.

[15] Wartość oczekiwana wielowymiarowej zmiennej losowej jest wektorem o składowych $E(Z_h)$.

I. Prawdopodobieństwo warunkowego rozkładu wektora ryzyka $\boldsymbol{R}$, względem składowej losowej $\{\boldsymbol{Z_j}\}$, przy warunku $\boldsymbol{Z_j = z}$ jest równa

$$\boldsymbol{P\{(a_1 < z_1 \le b_1 \ldots, a_{j-1} < z_{j-1} \le b_{j-1}, a_{j+1} < z_{j+1} \le b_{j+1}, \ldots, a_4 < z_4 \le b_4) \,|\, (z_j = z)\} = H}$$

$$\boldsymbol{H = \int_{a_1}^{b_1} \ldots \int_{a_{j-1}}^{b_{j-1}} \int_{a_{j+1}}^{b_{j+1}} \ldots \int_{a_4}^{b_4} \frac{f(z_1, z_2, z_3, z_4)}{f(z_j)} dz_1 \ldots dx_{j-1}\, dx_{j+1} \ldots dz_4, \; j = 1,2,3,4}$$

II. Warunkowa wartość oczekiwana wektora ryzyka $\boldsymbol{R}$ przy warunku $\boldsymbol{Z_j = z}$ jest wektorem

$$\boldsymbol{E(R \,|\, Z_j = z) = (E(Z_1 \,|\, (Z_j = z)), \ldots, E(Z_{j-1} \,|\, (Z_j = z)), \ldots, E(Z_4 \,|\, (Z_j = z)))}$$

gdzie $\boldsymbol{E(Z_l \,|\, (Z_j = z)) = \int_{a_l}^{b_l} z_l \frac{f(z_1, z_2, z_3, z_4)}{f(z_j)} dz_l}$; $\boldsymbol{l = 1,2,3,4 \;\wedge\; l \neq j}$

III. Warunkowa wariancja wektora ryzyka $\boldsymbol{R}$ przy warunku $\boldsymbol{Z_l = z}$ jest wektorem

$$\boldsymbol{Var(R \,|\, Z_j = z) = (Var(Z_1 \,|\, (Z_j = z)), \ldots, Var(Z_{j-1} \,|\, (Z_j = z)), \ldots, Var(Z_4 \,|\, (Z_j = z)))}$$

gdzie $\boldsymbol{Var(Z_l \,|\, (Z_j = z)) = \int_{a_l}^{b_l} [z_l - E(Z_l = z)]^2 \frac{f(z_1, z_2, z_3, z_4)}{f(z_j)} dz_l}$, gdzie $\boldsymbol{l = 1,2,3,4 \;\wedge\; l \neq j}$

IV. Warunkowe kowariancje składowych $\{\boldsymbol{Z_l, Z_h}\}$ wektora ryzyka $\boldsymbol{R}$ przy warunku $\boldsymbol{Z_j = z}$ są równe

$$\boldsymbol{Cov((Z_l, Z_h) \,|\, (Z_j = z)) = \int_{a_l}^{b_l} \int_{a_h}^{b_h} [z_l - E(Z_l \,|\, (Z_j = z))]\,[z_h - E(Z_h \,|\, (Z_j = z))] \frac{f(z_1, z_2, z_3, z_4)}{f(z_j)} dz_l\, dz_h}$$

gdzie $\boldsymbol{l, h = 1,2,3,4 \;\wedge\; l \neq j, h \neq j}$

W rezultacie zrealizowania kolejnej ***i****-tej* „sesji" pomiaru zmiennych kontrolowanych, aktualizowane są granice $[\boldsymbol{a_h, b_h}]$ zmian zmiennych $\boldsymbol{Z_h}$, pozwala to oszacować ponownie miary zdefiniowane w tym przykładzie zależnościami 1-4 oraz I-IV. W wyniku wielokrotnych powtórzeń procedury, na którą składa się pomiar zmiennych kontrolowanych oraz wyznaczanie miar charakterystycznych wektora ryzyka, otrzymamy zbiory miar:

1. $\boldsymbol{< P(a \le z \le b), E(R), Var(R), Cov(Z_g, Z_h) >^{(i)}}$

I. $< P(a \le z \le b \mid Z = z), E(R \mid Z = z), Var(R \mid Z = z), Cov(Z_g, Z_h \mid Z = z) >^{(i)}$ [16], gdzie $^{(i)}$ wskazuje miary statystyczne wektora ryzyka wyznaczone w ***i**-tej* „sesji" procesu monitoringu zmiennych kontrolowanych.

Analiza zmian wartości miar wektora ryzyka w „sesji" (i +1) w relacji do oszacowanych miar „sesji" (i) dostarcza znaczących informacji o zmianach stanu ryzyka. Jeśli przedział czasu pomiędzy sąsiednimi „sesjami" procesu kontroli realizacji planowanych celów jest niewielki, zarządzający otrzymują niemalże natychmiast informacje o zmianach stanu ryzyka związanego nieodłącznie z procesami zarządzania organizacją gospodarczą.

Podsumowanie

Konstrukcja przestrzeni ryzyka przedstawiona w tym opracowaniu została oparta na dwu hipotezach, treść których odwołuje się do teorii zarządzania organizacją gospodarczą. Pierwsza z nich jest oczywista, ryzyko podjęte w procesie zarządzania można ocenić w po skutkach realizowanych decyzji i działań. Informacji dostarcza tu proces kontroli implementacji planów strategicznych. Monitoring w procesie kontroli realizacji planów jest działaniem celowym, skoncentrowanym na dostarczaniu informacji o zmianach, dynamice i trendzie zmian tzw. zmiennych kontrolowanych procesu zarządzania. Jest to uporządkowany zbiór zmiennych, według określonej relacji porządkującej, która identyfikuje zmienne kontrolowane: zasobów rzeczowych, zasobów ludzkich, zasobów informacyjnych, zasobów systemu finansowego oraz zasobów makrootoczenia. Ta identyfikacja zmiennych kontrolowanych stanowi treść drugiej hipotezy, która utożsamia ryzyka decyzji i działań ze zdefiniowanymi zmiennymi kontrolowanymi. Jak na podstawie obydwu hipotez rozwiązano zadanie budowy modelu ryzyka?

Skorzystano z definicji przestrzeni zdarzeń elementarnych do konstrukcji przestrzeni probabilistycznej, który jest zbiorem wektorów losowych, z określoną miarą *P* (spełnia aksjomaty prawdopodobieństwa). Wynik ten uzyskano w rezultacie zdefiniowania ciała zdarzeń $\mathcal{F}$, które pełni rolę „katalizatora" funkcji odwzorowującej elementy przestrzeni zdarzeń elementarnych Ω w przestrzeń wektorów losowych $(X: \Omega \to \mathcal{R}^k$ mierzalną względem

[16] Zapis $(a \le z \le b)$ jest skrótem $(a_1 \le z_1 \le b, a_2 \le z_2 \le b_2, a_3 \le z_3 \le b_3, a_4 \le z_4 \le b_4)$, $E(R)$, $Var(R)$ oznacza odpowiednio wektor wartości oczekiwanej oraz wektor wariancji wektora ryzyka R, $Cov(Z_g, Z_h)$ – kowariancja składowych (Z_g, Z_h) wektora ryzyka R, gdzie: $g, h = 1,2,3,4 \ \wedge \ g \neq h$. Podobnie $P(a \le z \le b \mid Z = z)$, $E(R \mid Z = z)$, $Var(R \mid Z = z)$, $Cov(Z_g, Z_h \mid Z = z)$ oznaczają miary warunkowe.

σ – ciała $\mathcal{F}$). Wynik takiej konstrukcji identyfikuje ryzyko jako wielowymiarowy wektor losowy, którego współrzędne są zmiennymi losowymi identyfikowanymi ze zmiennymi kontrolowanymi. Taka identyfikacja pozwala mierzyć zmiany stanu ryzyka w czasie, gdyż możliwa staje się definicja miar charakterystycznych tak zdefiniowanego wektora losowego.

Literatura

1. Arrow K.J. (1979). Esej z teorii ryzyka. PWN, Warszawa.
2. Bartoszewicz J. (1989). Wykłady ze statystyki matematycznej. PWN, Warszawa
3. Dobbins R., Frąckowiak W., Witt S.F. (1992). Praktyczne zarządzanie kapitałami firmy, PAANPOL, Poznań.
4. Feller W. (1977, 1978). Wstęp do rachunku prawdopodobieństwa. T.1 i 2. PWN, Warszawa.
5. Gołębiewski T. (2001). Zarządzanie strategiczne. Planowanie i kontrola. Difin, Warszawa.
6. Knight F.H. (1921). Risk, Uncertainty and Profit. University of Boston Press, Boston.
7. Zarządzanie strategiczne. Koncepcja, metody. Red. R. Krupski. (1998). AE, Wrocław.
8. Pawłowski Z. (1969). Wstęp do statystyki matematycznej. PWN, Warszawa.
9. Sangowski T. (1995). Gospodarka finansowa zakładu ubezpieczeń. AE, Poznań.
10. Modelowanie preferencji a ryzyko ’08. (2008). Red. T. Trzaskalik. AE, Katowice.
11. Wheelen T., Hunger D. (1992). Strategic Management and Business Policy. Addison-Wesley, Reading, MA.
12. Willett A.H. (1951). The Economic Theory of Risk and Insurance. University of Pensylwania Press, Philadelphia.
13. Zemke J. (2005). Ryzyko – wielowymiarowa zmienna losowa. Biznes, Bankowość i Finanse na Rynkach Wschodzących. Szczecin.
14. Zemke J. (2006). Determinant of the Risk Management Model in Bussines. Clute Institute for Academic Research Littleton USA.
15. Zemke J. (2009). Ryzyko zarządzania organizacją gospodarczą. UG, Gdańsk.

RYZYKO – EKONOMIA, MATEMATYKA I EKONOMETRIA

Agnieszka Przybylska-Mazur
Akademia Ekonomiczna w Katowicach

ZASTOSOWANIE METODY AUTOKOWARIANCYJNEJ DO PROGNOZOWANIA WSKAŹNIKA INFLACJI

Wstęp

Rozwiązanie problemu decyzyjnego jest decyzją optymalną wybraną ze zbioru decyzji dopuszczalnych. Ważnymi czynnikami wpływającymi na podjęcie optymalnej decyzji są prognozy gospodarcze, które pozwalają przewidzieć przyszłe trendy gospodarki. Jedną z istotnych determinant wyznaczającą przyszłe trendy gospodarki jest prognoza wskaźnika inflacji, gdyż wpływa ona między innymi na wysokości stóp procentowych, stopy bezrobocia oraz na wielkość PKB w przyszłości. Przy podejmowaniu decyzji monetarnych, czyli decyzji związanych z realizacją przyjętej polityki pieniężnej wiele banków centralnych, w tym również Narodowy Bank Polski, wybiera jako cel realizacji polityki pieniężnej stabilny poziom cen i dlatego prognozują one inflację.

W pracy została przedstawiona jedna z metod prognozowania wskaźnika inflacji – metoda autokowariancyjną.

1. Prognoza autokowariancyjna szeregów czasowych jednowymiarowych

Metoda autokowariancyjna prognozująca stacjonarne szeregi czasowe jednowymiarowe zaprezentowana została po raz pierwszy w [3].

Główną rolę w autokowariancyjnej metodzie prognozowania odgrywa funkcja autokowariancji, a właściwie jej estymator. Funkcja autokowariancji nieskończonego stacjonarnego ciągłego szeregu czasowego, którego wartość oczekiwana jest równa zero, jest określona następującym równaniem

$$c_X(k) = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t)\, X(t+k)\, dt \qquad (1)$$

gdzie k jest przesunięciem czasowym.

W praktyce szeregi czasowe są skończone i określone jedynie w punktach dyskretnych i zamiast funkcji autokowariancji posługujemy się jej estymatorem.

Obciążony estymator autokowariancji jest określony następującym wzorem

$$\hat{c}_k^{(N)} = \frac{v_k}{N-k} \sum_{t=1}^{N-k} x_t x_{t+k} \quad \text{dla} \quad k = 0, 1, \ldots, N-1 \tag{2}$$

gdzie:

x_t – jest stacjonarnym szeregiem czasowym określonym w równoodległych momentach czasu, o wartości oczekiwanej równej zero,
N – jest liczbą punktów szeregu czasowego,
v_k – jest dowolnym oknem wagowym.

Okno wagowe zmniejsza obciążenie wyznaczonych estymatorów autokowariancji, które jest spowodowane skończonym czasem obserwacji szeregu.

Natomiast nieobciążony estymator autokowariancji obliczamy korzystając z następującego wzoru

$$\hat{c}_k^{(N)} = \frac{1}{N-k} \sum_{t=1}^{N-k} x_t x_{t+k} \quad \text{dla} \quad k = 0, 1, \ldots, N-1 \tag{3}$$

Ideą autokowariancyjnej metody prognozowania jest znalezienie takiego estymatora pierwszego punktu prognozy x_{N+1}, aby był spełniony warunek

$$P = \sum_{k=0}^{N-1} (\hat{c}_k^{(N+1)} - \hat{c}_k^{(N)})^2 \to \min \tag{4}$$

Zatem idea autokowariancyjnej metody prognozowania polega na wyznaczeniu takiego estymatora pierwszego punktu prognozy x_{N+1}, aby suma kwadratów odchyleń pomiędzy obciążonym estymatorem autokowariancji $\hat{c}_k^N$ i obciążonym estymatorem autokowariancji $\hat{c}_k^{N+1}$ z dodanym pierwszym punktem prognozy była minimalna.

Ponieważ

$$\hat{c}_k^{(N)} = \frac{v_k}{N-k} \sum_{t=1}^{N-k} x_t x_{t+k}$$

$$\hat{c}_k^{(N+1)} = \frac{w_k}{N+1-k} \sum_{t=1}^{N+1-k} x_t x_{t+k} =$$

$$= \frac{w_k}{N+1-k} \sum_{t=1}^{N-k} x_t x_{t+k} + \frac{w_k}{N+1-k} x_{N+1-k}\, x_{N+1} =$$

$$= \frac{w_k}{v_k} \frac{N-k}{N+1-k} \hat{c}_k^{(N)} + \frac{w_k}{N+1-k} x_{N+1-k}\, x_{N+1}$$

funkcję P można zapisać jako funkcję tylko x_{N+1} w następujący sposób

$$P(x_{N+1}) = \sum_{k=0}^{N-1} \left[\left(\frac{w_k}{v_k} \frac{N-k}{N+1-k} - 1 \right) \hat{c}_k^{(N)} + \frac{w_k}{N+1-k} x_{N+1-k}\, x_{N+1} \right]^2$$

czyli

$$P(x_{N+1}) = \sum_{k=1}^{N-1} \left[\left(\frac{w_k}{v_k} \frac{N-k}{N+1-k} - 1 \right) \hat{c}_k^{(N)} + \frac{w_k}{N+1-k} x_{N+1-k}\, x_{N+1} \right]^2 +$$

$$+ \left[\left(\frac{w_0}{v_0} \frac{N}{N+1} - 1 \right) \hat{c}_0^{(N)} + \frac{w_0}{N+1} x_{N+1}^2 \right]^2$$

Funkcja P osiąga minimum, gdy $\frac{dP(x_{N+1})}{dx_{N+1}} = 0$. Ponieważ

$$\frac{dP(x_{N+1})}{dx_{N+1}} =$$

$$= 2 \sum_{k=1}^{N-1} \left[\left(\frac{w_k}{v_k} \frac{N-k}{N+1-k} - 1 \right) \hat{c}_k^{(N)} + \frac{w_k}{N+1-k} x_{N+1-k}\, x_{N+1} \right] \frac{w_k}{N+1-k} x_{N+1-k} +$$

$$+ 2 \left[\left(\frac{w_0}{v_0} \frac{N}{N+1} - 1 \right) \hat{c}_0^{(N)} + \frac{w_0}{N+1} x_{N+1}^2 \right] \frac{w_0}{N+1} 2 x_{N+1}$$

zatem, aby wyznaczyć estymator x_{N+1} pierwszego punktu prognozy należy rozwiązać następujące równanie trzeciego stopnia

$$x_{N+1}^3 + 3 p x_{N+1} + 2q = 0 \tag{5}$$

gdzie

$$p = \sum_{k=1}^{N-1} x_{k+1}^2 \left(\frac{w_{N-k}^2 (N+1)^2}{6 w_0^2 (k+1)^2} + \frac{N w_0 - (N+1) v_0}{3 w_0 N} \right) +$$

$$+ \frac{N w_0 - (N+1) v_0}{3 w_0 N} x_1^2 \tag{6}$$

$$q = \frac{(N+1)^2}{4w_0^2} \sum_{k=1}^{N-1} \frac{w_k \, \hat{c}_k^{(N)} x_{N-k+1} [(\frac{w_k}{v_k} - 1)(N-k) - 1]}{(N-k+1)^2} \tag{7}$$

Równania trzeciego stopnia ma jeden pierwiastek rzeczywisty i dwa zespolone, gdy okno wagowe v_k spełnia następujący warunek

$$v_k \geq \sqrt{\frac{2}{N-1}} \; (1 - \frac{k}{N})$$

W analizie szeregów czasowych wszystkie okna wagowe spełniają powyższy warunek.

Wzór Cardana na rzeczywisty pierwiastek równania $y^3 + p_1 y + q_1 = 0$ jest następujący

$$y = \sqrt[3]{\frac{-q_1 - \sqrt{q_1^2 - \frac{4p_1^3}{27}}}{2}} + \sqrt[3]{\frac{-q_1 + \sqrt{q_1^2 - \frac{4p_1^3}{27}}}{2}}$$

Podstawiając w powyższym wzorze $p_1 = 3p, q_1 = 2q$ otrzymujemy następujący wzór na szukany estymator x_{N+1} pierwszego punktu prognozy:

$$x_{N+1} = \sqrt[3]{-q + \sqrt{q^2 + p^3}} + \sqrt[3]{-q - \sqrt{q^2 + p^3}} \tag{8}$$

W autokowariancyjnej metodzie prognozowania każde okna wagowe v_k, w_k można zastosować. Jednak najbardziej odpowiednie jest okno wagowe Kaisera ze względu na fakt, że dla tego okna błąd estymatora autokowariancji wynikający ze skończonej długości szeregu czasowego jest najmniejszy. Zastosowanie okna wagowego Kaisera wydłuża jednak czas obliczeń, ze względu na występujące funkcje Bessela w równaniu tego okna wagowego.

Czas obliczeń estymator x_{N+1} pierwszego punktu prognozy można jednak skrócić w następujący sposób: przez zastosowanie metody rekurencyjnej do obliczania następnego nieobciążonego estymatora autokowariancji, gdy znany jest poprzedni

$$\hat{c}_k^{(n+1)} = \frac{\hat{c}_k^{(n)}(n-k) + x_{n+1} \cdot x_{n-k+1}}{n-k+1} \tag{9}$$

dla $k = 0, 1, \ldots, n, \quad n = 1, 2, ..., N$.

Drugim sposobem skrócenia obliczeń estymatora x_{N+1} pierwszego punktu prognozy, jednak z nieznacznym pogorszeniem dokładności prognozy jest zastosowanie okna wagowego Bartletta, czyli okna wagowego określonego wzorem

$$v_k = 1 - \frac{k}{N} \qquad \text{dla} \quad k = 0, 1, \ldots, N-1 \tag{10}$$

Podstawiając do wzorów (6) i (7) okno wagowe Bartletta $v_k = 1 - \frac{k}{N}$, $w_k = 1 - \frac{k}{N+1}$ otrzymujemy następujące uproszczone wzory na współczynniki p, q

$$p = \frac{1}{6} \sum_{k=1}^{N-1} x_{k+1}^2 - \frac{1}{3} \hat{c}_0^{(N)} \tag{11}$$

$$q = -\frac{1}{4} \sum_{k=1}^{N-1} \hat{c}_k^{(N)} x_{N-k+1} \tag{12}$$

Badania modelowe na szeregach czasowych [3] wykazały, że wyznaczona wartość estymatora x_{N+1} pierwszego punktu prognozy jest zwykle mniejsza od prawdziwej bezwzględnej wartości modelowego szeregu czasowego. Zatem, aby zmniejszyć błąd prognozy przyjmuje się, że pierwszy punkt prognozy $\hat{X}_{N+1}$ jest iloczynem estymatora pierwszego punktu prognozy x_{N+1} oraz pewnego stałego współczynnika m_α, który jest średnią obliczoną metodą estymacji szorstkiej. Wartość tego współczynnika uzyskuje się poprzez porównanie estymatorów pierwszych punktów prognozy z rzeczywistymi wartościami szeregu czasowego dla ostatnich $N - L + 1$ punktów szeregu czasowego. Średnia „odporna" („robust") m_α obliczana jest z następującego warunku

$$\sum_{j=0}^{N-L} \left| m_\alpha - \alpha_j \right| \rightarrow \min$$

gdzie parametry $\alpha_j = \frac{x_{L+j}}{\hat{x}_{L+j}}$ dla $j = 0, 1, \ldots, N-L$ są obliczane na podstawie znanych punktów szeregu czasowego x_{L+j} oraz ich pierwszych estymatorów prognozy $\hat{x}_{L+j} \neq 0$, gdzie L powinno być dostatecznie duże. Przyjmuje się najczęściej a priori $L = \frac{2}{5} N$. Jeżeli dla $j > 0$ $\hat{x}_{L+j} = 0$, wówczas parametr musi być zastąpiony poprzednim α_{j-1}.

Po obliczeniu pierwszego punktu prognozy $\hat{X}_{N+1}$ jest on dołączany do szeregu czasowego, co umożliwia wyznaczenie następnego punktu prognozy $\hat{X}_{N+2} = m_\alpha \cdot x_{N+2}$ itd.

2. Przykład empiryczny

Analizie poddano wskaźnik inflacji z okresu od lipca 1996 roku do grudnia 2008 – szereg 150 obserwacji.

Wykonując test DF stwierdzono, że szereg 150 obserwacji jest stacjonarny. Następnie szereg stacjonarny sprowadzono do szeregu o średniej równej zero.

Wykorzystując wzór (8) na pierwszy punkt prognozy oraz wzory (11), (12) na współczynniki p, q otrzymano rekurencyjnie wyniki zawarte w tabeli 1.

Tabela 1

Estymator pierwszego punktu prognozy

Numer okresu	Współczynnik p	Współczynnik q	Estymator pierwszego punktu prognozy
1	644,46	393,04	-0,41
2	645,08	368,22	-0,38
3	645,61	365,69	-0,38
4	646,13	352,10	-0,36

Następnie rekurencyjnie wyznaczono średnią „odporną” oraz wartość prognozy dla stacjonarnego szeregu wskaźnika inflacji. Wyniki obliczeń zaprezentowano w tabeli 2.

Tabela 2

Prognozy dla stacjonarnego szeregu wskaźnika inflacji

Numer okresu	Miesiąc	Estymator x_{N+1} pierwszego punktu prognozy	Średnia „odporna” m_α	Prognoza $\hat{X}_{N+1}$
1	styc-09	-0,41	4,76	-1,94
2	lut-09	-0,38	[4,73; 4,76]	-1,81
3	mar-09	-0,38	[4,67; 4,73]	-1,79
4	kwic-09	-0,36	4,70	-1,71

W tabeli 3 zestawiono wyznaczony rekurencyjnie prognozowany wskaźnik inflacji oraz rzeczywistą wartość wskaźnika inflacji.

Tabela 3

Prognozy wskaźnika inflacji

Numer okresu	Miesiąc	Prognozowany wskaźnik inflacji	Rzeczywista wartość wskaźnika inflacji	Błąd prognozy
1	styc-09	104,11	102,8	-1,31
2	lut-09	104,22	103,3	-0,92
3	mar-09	104,23	103,6	-0,63
4	kwic-09	104,30	104,0	-0,3

Zakończenie

W pracy zaprezentowano prognozy wskaźnika inflacji na podstawie wybranej metody prognozowania – metody autokowariancyjnej. W przykładzie empirycznym wykorzystano do obliczeń szereg stacjonarny 150 danych dotyczących wskaźnika inflacji. Na podstawie obliczeń stwierdzono, że kierunek zmian prognoz wskaźnika inflacji pokrywa się z kierunkiem zmian rzeczywistej wartości wskaźnika inflacji i daje coraz mniejsze błędy prognozy.

Literatura

1. Box G.E.P., Jenkins G.M. (1976). Time Series Analysis, Forecasting and Control. Holden Day, San Francisco.
2. Dittmann P. (2004). Prognozowanie w przedsiębiorstwie. Metody i ich zastosowania. Oficyna Ekonomiczna, Kraków.
3. Kosek W. Short Periodic Autoregressive Prediction of the Earth Rotation Parameters. Artificial Satellites, Planetary Geodesy, 27, 2.
4. Lutkepohl H. (2005). New Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.

Olena Sobotka
Uniwersytet Ekonomiczny we Wrocławiu

ZASTOSOWANIE KONCEPCJI DOMINACJI STOCHASTYCZNEJ DO WSPOMAGANIA DECYZJI W PROBLEMACH SFORMUŁOWANYCH W KATEGORIACH STRAT

Wstęp

Wykorzystanie matematycznych, statystycznych, ekonomicznych i innych metod analizy wariantów decyzyjnych w celu ich porównania i oceny względem ustalonych celów jest wspomaganiem podejmowania decyzji. Metody wspomagania decyzji, które opierają się na normatywnej teorii podejmowania decyzji, rekomendują tzw. racjonalne wybory w określonych sytuacjach decyzyjnych. Zgodnie z założeniem racjonalności, decydent zmierza do wyboru decyzji optymalnej, maksymalizującej jego funkcję użyteczności.

Klasyczne założenie racjonalności jest często podważane przez rezultaty deskryptywnej teorii decyzji. H. Simon pierwszy zwrócił uwagę na kilka ważnych ograniczeń, dotyczących racjonalności procesu podejmowania decyzji. Najważniejsze z nich to: 1) fakt, że decydent realizuje zwykle nie jeden cel, a kilka, często niezgodnych, celów, 2) brak pełnej informacji o wariantach decyzyjnych i ich cechach, 3) decydent, poszukując wariantu decyzyjnego, który spełniałby wszystkie jego wymagania, nie ma możliwości porównania wszystkich wariantów. H. Simon [9] wprowadził pojęcie racjonalności ograniczonej, które osłabia niektóre klasyczne założenia, dotyczące wyboru decyzji. Zgodnie z założeniem racjonalności ograniczonej, decydent poszukuje wariantu decyzyjnego, w dostatecznym stopniu spełniającego jego wymagania, decyzji zadowalającej.

Sytuacje decyzyjne, charakteryzujące się brakiem pewności co do wyników decyzji, pojawiają się we wszystkich dziedzinach współczesnego zarządzania. W analizie tego rodzaju sytuacji rozróżnia się pojęcia ryzyka i niepewności. Ryzyko odnosi się do sytuacji, w których decydent może opisać niepewność, z którą ma do czynienia, za pomocą jednoznacznie określonych rozkładów prawdopodobieństwa.

Racjonalne wybory w warunkach ryzyka opisują modele preferencji w postaci funkcji równoważnika pewności (certainty equivalent) (funkcji użyteczności, funkcji percepcji prawdopodobieństwa). Wspomaganie podejmowania decyzji zgodnie z modelem racjonalnych wyborów polega na wskazaniu wariantu decyzyjnego, optymalizującego odpowiednią funkcję preferencji.

Do porównania wariantów decyzyjnych w warunkach ryzyka w celu poszukiwania rozwiązania zadowalającego można wykorzystać koncepcję dominacji stochastycznej [11]. Wykorzystanie koncepcji dominacji stochastycznej polega na wyznaczeniu wariantów decyzyjnych o rozkładach prawdopodobieństwa wyników niezdominowanych w sensie relacji preferencji, odpowiadającej określonym, ogólnym założeniom o wyborach decydenta w warunkach ryzyka (np. awersji czy skłonności do podejmowania ryzyka).

Relacje dominacji stochastycznej opierają się na matematycznych porządkach stochastycznych i pozwalają porównywać rozkłady prawdopodobieństwa wyników decyzji według ogólnie określonej racjonalnej relacji preferencji w warunkach ryzyka. Najczęściej we wspomaganiu decyzji wykorzystuje się relację dominacji stochastycznej drugiego rzędu, która odpowiada awersji do ryzyka decydenta, opisującej racjonalne preferencje w problemach maksymalizacji zysku, oraz odwrotną relację dominacji stochastycznej drugiego rzędu, która odpowiada racjonalnym preferencjom w problemach minimalizacji strat.

Nieparametryczny charakter relacji dominacji stochastycznych, zarówno w odniesieniu do postaci rozkładu prawdopodobieństwa, jak i w odniesieniu do funkcji preferencji w warunkach ryzyka, pozwala na ich szerokie zastosowanie do wspomagania podejmowania decyzji w warunkach ryzyka [4]. Zastosowanie to jednak ogranicza się do problemów o ograniczonym i przeliczalnym zbiorze decyzji dopuszczalnych oraz sytuacji, w których nie uwzględnia się współzależności losowych parametrów, od których zależy wynik decyzji. Jest to spowodowane brakiem odpowiednich metod obliczeniowych w przypadku problemów z nieprzeliczalnym i nieskończonym zbiorem decyzji dopuszczalnych. Natomiast właśnie tego typu problemy decyzyjne pozwalają uwzględnić efekt rozproszenia ryzyka czy dywersyfikacji, tzn. otrzymania lepszych, z punktu widzenia racjonalnie postępującego decydenta, rozwiązań przez kombinacje wariantów decyzyjnych.

W ostatnich latach pojawiło się kilka propozycji wyznaczania zbioru decyzji efektywnych w sensie relacji dominacji stochastycznej pierwszego i drugiego rzędu w problemach z nieprzeliczalnym i nieskończonym zbiorem decyzji dopuszczalnych. Najbardziej interesującą z punktu widzenia zastosowań w rzeczywistych problemach decyzyjnych wydaje się propozycja W. Ogryczaka [6] przedstawiona w artykule *Multiple Criteria Optimization and Decisions under Risk*, polegająca na wykorzystaniu metod optymalizacji wielokryterialnej.

Zadania wielokryterialne, pozwalające wygenerować decyzje efektywne w sensie relacji dominacji stochastycznej drugiego rzędu, zostały sformułowane dla przypadku problemu maksymalizacji zysków. Natomiast, z punktu widzenia wyborów dokonywanych w warunkach ryzyka, ważne jest, czy problem jest sformułowany w kategoriach zysku czy w kategoriach strat [3]. W problemach minimalizacji strat, powstających np. w zarządzaniu w sferze ubezpieczeń lub w zarządzaniu zapasami, uwzględnienie wzajemnej zależności parametrów losowych jest bardzo ważne [2]. Dlatego celem tego opracowania jest pokazanie możliwości zastosowania metod optymalizacji wielokryterialnej do generowania decyzji efektywnych w sensie relacji dominacji stochastycznej w problemach minimalizacji strat.

1. Zastosowanie koncepcji odwrotnej dominacji stochastycznej do wspomagania decyzji w problemach sformułowanych w kategoriach strat

1.1. Odwrotna relacja dominacji stochastycznej

Przedmiotem tej pracy jest problem wyboru decyzji w warunkach ryzyka z nieprzeliczalnym i nieskończonym zbiorem decyzji dopuszczalnych, który możemy sformułować z wykorzystaniem parametrów losowych. Oznaczymy przez $\mathbf{x} = (x_1,x_2,...,x_n)$ wektor decyzji, gdzie x_j jest zmienną decyzyjną (j=1,...,n), a D – zbiór decyzji dopuszczalnych, na którym określona jest skalarna funkcja celu, wartość której zależy od realizacji niepewnych parametrów zewnętrznych: $f(\mathbf{x},\boldsymbol{\omega}) \rightarrow \text{opt}$, gdzie $\boldsymbol{\omega}=(\omega_1,...,\omega_n)$ jest wektorem parametrów losowych, których rozkłady prawdopodobieństwa są znane. Wtedy wynik każdej decyzji jest zmienną losową $Y=f(\mathbf{x},\boldsymbol{\omega})$, której rozkład prawdopodobieństwa możemy wyznaczyć.

Według teorii podejmowania decyzji w warunkach ryzyka, wybory decydentów są zależne od sposobu sformułowania problemu, tzn. ujęcia celów sytuacji decyzyjnej w kategoriach zysku lub w kategoriach strat Dlatego w pracy będziemy rozróżniać dwa rodzaje problemów wyboru decyzji w warunkach ryzyka: problemy sformułowane w kategoriach zysków i problemy sformułowane w kategoriach strat.

Problemem sformułowanym w kategoriach zysków będziemy nazywać problem decyzyjny, w którym funkcja celu jest maksymalizowana. W takich problemach rozkłady prawdopodobieństwa wyników decyzji zwykle opisuje się za pomocą funkcji dystrybuant: $F_Y(t) = \Pr\{Y \le t\}, t \in R$. Problemem sformułowanym w kategoriach strat będziemy nazywać problem, gdzie funkcja celu

dąży do minimum. W przypadku minimalizowania strat rozkłady prawdopodobieństwa wyników często opisuje się za pomocą funkcji dekumulacyjnych (survival function, tail function): $S_Y(t) = \Pr\{Y > t\} = 1 - F_Y(t), t \in R$.

Racjonalne wybory w warunkach ryzyka opisują modele preferencji w postaci funkcji równoważnika pewności (certainty equivalent) (funkcji użyteczności, funkcji percepcji prawdopodobieństwa). Wspomaganie podejmowania decyzji zgodnie z modelem racjonalnych wyborów polega na wskazaniu wariantu decyzyjnego, optymalizującego odpowiednią funkcję preferencji.

Decydenci mają skłonność do unikania ryzyka, gdy problem jest przedstawiony w kategoriach zysków i jednocześnie wykazują skłonność do podejmowania ryzyka, jeśli problem jest sformułowany w kategoriach strat.

Do porównania wariantów decyzyjnych w warunkach ryzyka, w celu poszukiwania rozwiązania zadowalającego, można wykorzystać koncepcję dominacji stochastycznej.

Niech Z_1 i Z_2 ($Z_1, Z_2 \geq 0$) będą losowymi wynikami problemu wyboru decyzji sformułowanego w kategoriach strat.

Odwrotna relacja dominacji stochastycznej drugiego rzędu pozwala wyodrębnić zbiór wariantów decyzyjnych problemu sformułowanego w kategoriach strat, których losowe wyniki będą preferowane przez decydentów ze skłonnością do ryzyka w przestrzeni strat, co odpowiada awersji do ryzyka w przestrzeni zysków [12].

Rozkład prawdopodobieństwa zmiennej losowej Z_2 będzie zdominowany przez rozkład prawdopodobieństwa zmiennej losowej Z_1 w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu, jeśli losowa strata Z_1 będzie nie gorsza niż losowa strata Z_2 w sensie wszystkich modeli preferencji w postaci funkcji równoważnika pewności, które opisują skłonność do ryzyka decydenta. Ustalić, czy zachodzi odwrotna relacja dominacji stochastycznej drugiego rzędu można porównując funkcje dekumulacyjne lub funkcji odwrotne do funkcji dekumulacyjnych rozkładów prawdopodobieństwa zmiennych losowych

$$Z_1 \succ_2^* Z_2 \text{ wtedy i tylko wtedy, gdy } \int_t^\infty S_{Z_1}(r)dr \leq \int_t^\infty S_{Z_2}(r)dr, \forall t$$

jeśli odpowiednie całki istnieją.

$$Z_1 \succ_2^* Z_2 \text{ wtedy i tylko wtedy, gdy } \int_0^q S_{Z_1}^{(-1)}(p)dp \leq \int_0^q S_{Z_2}^{(-1)}(p)dp, \forall 0 \leq q \leq 1$$

Relacje dominacji stochastycznych wykorzystuje się w analizie problemów decyzyjnych do uporządkowania zbiorów rozwiązań dopuszczalnych z wydzieleniem z nich zbiorów rozwiązań efektywnych w sensie odpowiedniej relacji dominacji stochastycznej. Rozwiązanie dopuszczalne należy do zbioru rozwiązań efektywnych w sensie odpowiedniej relacji dominacji stochastycznej,

jeśli w zbiorze rozwiązań dopuszczalnych nie istnieje inne rozwiązanie, którego rozkład prawdopodobieństwa wyników dominuje rozkład prawdopodobieństwa wyników danego rozwiązania w sensie odpowiedniej relacji dominacji stochastycznej.

1.2. Wykorzystanie odwrotnej relacji dominacji stochastycznej do porównania portfeli kontraktów ubezpieczeniowych

W sferze ubezpieczeń ryzyko jest związane z losowością zdarzeń, które powodują roszczenia finansowe, wynikające z kontraktu ubezpieczeniowego. Roszczenia modeluje się w postaci zmiennej losowej. Problemy decyzyjne, powstające w ubezpieczeniach, są formułowane w kategoriach strat.

Jednym z narzędzi zarządzania ryzykiem w sferze ubezpieczeń jest reasekuracja – rozłożenie ryzyka między ubezpieczycielem i reasekuratorem, ubezpieczenie ubezpieczyciela [8].

Rozróżnia się dwa rodzaje umów reasekuracyjnych: reasekurację proporcjonalną (np. kwotową) – kiedy udział reasekuratora w składce i w płatnościach roszczeń jest w stałej proporcji do jednostek ryzyka, biorących udział w reasekuracji, reasekurację proporcjonalną – stosuje się ją do poszczególnych kontraktów (lub grup kontraktów) ubezpieczeniowych.

Drugim rodzajem umów jest reasekuracja nieproporcjonalna (np. reasekuracja nadwyżki szkody (stop-loss reisurance) – zabezpieczenie przed skutkami nadmiernie wysokiego roszczenia całego portfela kontraktów (z uwzględnieniem reasekuracji proporcjonalnej poszczególnych kontraktów). Ustala się udział własny (wartość zatrzymaną lub priorytet – deductible lub priority – maksymalną wysokość roszczeń portfela, które będą pokrywane przez samego ubezpieczyciela; nadwyżkę roszczeń pokrywa reasekurator. Przy ustalaniu składki za reasekurację nieproporcjonalną bierze się pod uwagę wysokość oczekiwanych roszczeń, które będą pokrywane przez reasekuratora

$$\pi(d) = (1+\rho)E(Z-d)^+$$

gdzie:

$\pi(d)$ – składka dla wartości zatrzymanej d,
Z – całkowita wysokość roszczeń portfela,
ρ – składka dodatkowa.

Portfel kontraktów ubezpieczeniowych jest mniej ryzykowny, jeśli składka za reasekurację tego portfela jest mniejsza.

W naukach aktuarialnych do porównana portfeli ubezpieczeniowych stosuje się porządek zatrzymania straty [2; 8].

Zmienna losowa Z_1 poprzedza zmienną losową Z_2 według porządku zatrzymania straty, jeśli

$$Z_1 \leq_{SL} Z_2 \Rightarrow \pi_{Z_1}(d) \leq \pi_{Z_2}(d) \forall d \geq 0$$

Ponieważ funkcja $(Z - d)^+$ jest wypukła, relacja odwrotnej dominacji stochastycznej drugiego rzędu generuje porządki ekwiwalentne z porządkiem zatrzymania straty: jeśli zmienna losowa Z_1 poprzedza zmienną losową Z_2 według porządku zatrzymania straty, to rozkład prawdopodobieństwa zmiennej Z_2 jest zdominowany przez rozkład prawdopodobieństwa zmiennej Z_1 w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu

$$Z_1 \leq_{SL} Z_2 \Leftrightarrow Z_1 \succ_2^* Z_2$$

Oznaczmy przez $\mathbf{x} = (x_1, x_2, ..., x_n)$ udziały poszczególnych kontraktów ubezpieczeniowych w portfelu (zmienne decyzyjne), a przez $\tau = (\tau_1, ..., \tau_n)$ – losowe wartości roszczeń, związanych z odpowiednimi kontraktami. Wtedy $L(\mathbf{x}) = \sum_{i=1}^{n} \tau_i x_i$ będzie losową wartością roszczeń całego portfela. Problem decyzyjny wyznaczenia optymalnego składu portfela kontraktów ubezpieczeniowych będzie miał postać

$$E(L(\mathbf{x}) - d)^+ \xrightarrow[\mathbf{x}]{} \min$$

$$L(\mathbf{x}) = \sum_{i=1}^{n} \tau_i x_i$$

$$\sum_{i=1}^{n} x_i = 1$$

$$x_i \geq 0$$

Za pomocą odwrotnej relacji dominacji stochastycznej drugiego rzędu można wyznaczyć zbiór portfeli efektywnych – zbiór portfeli o niezdominowanych (w sensie odwrotnej relacji dominacji stochastycznej) rozkładach prawdopodobieństwa losowych roszczeń $L(\mathbf{x})$. Zbiór portfeli efektywnych w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu jest zbiorem portfeli ubezpieczeniowych, których składki za reasekurację nieproporcjonalną są najmniejsze przy wszystkich wartościach udziału własnego (zatrzymanej straty) $d \geq 0$.

1.3. Wykorzystanie odwrotnej relacji dominacji stochastycznej w problemach optymalizacji wielkości zapasów

W zarządzaniu przedsiębiorstwem w warunkach niepewności, związanej z wielkością przyszłego popytu, powstaje problem wyznaczenia optymalnej wielkości zapasów. Jeśli wielkość popytu modeluje się za pomocą zmiennej losowej, to jest to problem wyboru decyzji w warunkach ryzyka sformułowany w kategoriach strat.

Oznaczmy przez ξ losową wielkość popytu, przez s – wielkość zapasu; przez $h(z) = h(s-\xi)^+$ – funkcję kosztów z powodu nadwyżki zapasu; $p(z) = p(\xi - s)^+$ – funkcję kosztów z powodu niedoboru zapasu.

Problem wyznaczenia optymalnej wielkości zapasu w jednym okresie będzie miał postać (D – zbiór dopuszczalnych wielkości zapasu)

$$K^{\xi}(s) = E(h[(s-\xi)^+] + p[(\xi - s)^+]) \xrightarrow[s\in D]{} \min$$

Przy założeniu, że funkcje kosztów h(z) i p(z) są niemalejące i wypukłe, udowodnione zostało twierdzenie [1]: jeśli ξ_1 i ξ_2 są losowymi wielkościami popytu na dwa produkty takie, że $E\xi_1=E\xi_2$ i rozkład prawdopodobieństwa zmiennej losowej ξ_2 jest zdominowany przez rozkład prawdopodobieństwa zmiennej losowej ξ_1 w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu, to $K^{\xi_1}(s) \le K^{\xi_2}(s) \forall s \ge 0$.

Podobna zależność została udowodniona w cytowanej wyżej pracy dla przypadku wielookresowego dynamicznego problemu optymalizacji wielkości zapasów przy założeniu, że popyt we wszystkich okresach jest zmienną losową o takim samym rozkładzie prawdopodobieństwa.

Jeśli ξ_1 i ξ_2 są losowymi wielkościami popytu na dwa produkty takie, że $E\xi_1=E\xi_2$ i rozkład prawdopodobieństwa zmiennej losowej ξ_2 jest zdominowany przez rozkład prawdopodobieństwa zmiennej losowej ξ_1 w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu, to minimalny n-okresowy koszt związany z wielkością zapasu produktu 1 jest mniejszy niż minimalny n-okresowy koszt związany z wielkością zapasu produktu 2, przy każdej początkowej wielkości zapasu s

$$K_n^{\xi_1}(s) \le K_n^{\xi_2}(s) \forall s \ge 0$$

Istnieje wiele innych obszarów zastosowań relacji dominacji stochastycznej, przy tym w większości problemów można zauważyć rozważania o możliwościach zmniejszania ryzyka w sensie relacji dominacji stochastycznej przez dywersyfikację.

2. Zastosowanie metod optymalizacji wielokryterialnej do wyznaczania decyzji efektywnych w sensie odwrotnej relacji dominacji stochastycznej oraz finalnego rozwiązania zadowalającego

W pracach Ogryczaka [5; 6] zostały sformułowane zadania wielokryterialne, pozwalające wyznaczyć zbiory rozwiązań efektywnych w sensie relacji dominacji stochastycznej pierwszego i drugiego rzędu problemów decyzji sformułowanych w kategoriach zysków. Łączny rozkład prawdopodobieństwa wektora parametrów losowych problemu decyzyjnego $\boldsymbol{\omega}$ został uwzględniony przez wprowadzenie zbioru stanów otoczenia $\{s_k\}_{k=1,\ldots,m}$ i zbioru prawdopodobieństw stanów otoczenia $\{p_k\}_{k=1,\ldots,m}$ oraz zbioru realizacji wektora parametrów losowych w każdym stanie otoczenia $\boldsymbol{\omega}_k = (\omega_{1k},...,\omega_{nk})$.

Natomiast, z punktu widzenia wyborów dokonywanych w warunkach ryzyka, ważne jest, czy problem jest sformułowany w kategoriach zysku czy w kategoriach strat. W problemach minimalizacji strat, powstających np. w zarządzaniu w sferze ubezpieczeń lub w zarządzaniu zapasami, uwzględnienie wzajemnej zależności parametrów losowych jest bardzo ważne.

Wykorzystując modele przedstawione przez Ogryczaka i Zawadzkiego [7], można sformułować problemy optymalizacji wielokryterialnej, określające zbiór decyzji efektywnych w sensie odwrotnej dominacji stochastycznej pierwszego i drugiego rzędu, czyli zbiór decyzji efektywnych dla problemu minimalizacji strat w warunkach ryzyka.

Niech zbiór $\{\lambda_i\}_{i=1,\ldots,r}$ będzie zbiorem poziomów prawdopodobieństwa, wydzielonym na odcinku [0;1] takim, że $0 < \lambda_1 < \lambda_2,..., \lambda_r=1$. Oznaczmy przez $t_{\lambda_i} = S_Y^{(-1)}(\lambda_i)$ największą częściową średnią dla poziomu prawdopodobieństwa λ_i – wartość funkcji $S_Y^{(-1)}(q)$ wyniku decyzji problemu minimalizacji strat w punkcie λ_i. Oznaczmy przez $\overline{m}_{\lambda_i} = S_Y^{(-2)}(\lambda_i)$ wartość funkcji $S_Y^{(-2)}(q) = \int_0^q S_Y^{(-1)}(p)dp$ wyniku decyzji problemu minimalizacji losowych strat w punkcie λ_i. Wielokryterialny problem, którego zbiorem rozwiązań efektywnych będą decyzje o wynikach niezdominowanych w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu, będzie miał postać

$$\min\{\overline{m}_{\lambda_1}, \overline{m}_{\lambda_2}, ..., \overline{m}_{\lambda_r}\}$$

$$\overline{m}_{\lambda_i} = \lambda_i \cdot t_{\lambda_i} - \sum_{k=1}^{m} d_{ki}^{+} p_k, \ \mathrm{i} = 1,...,\mathrm{r}$$

$$d_{ki}^{+} \geq f(\mathbf{x}, \boldsymbol{\omega}_k) - t_{\lambda_i}, \ \mathrm{k} = 1,...,\mathrm{m}, \ \mathrm{i} = 1,...,\mathrm{r}$$

$$d_{ki}^{+} \geq 0, \mathrm{k} = 1,...,\mathrm{m}, \mathrm{i} = 1,...,\mathrm{r}$$

$$\mathbf{x} \in \mathbf{D}$$

gdzie:

$f(\mathbf{x}, \omega_k)$ – wynik decyzji $\mathbf{x}$ przy realizacji stanu $\mathrm{s_k}$,

λ_i – wybrany poziom prawdopodobieństwa,

t_{λ_i} – wartość większą od której wynik decyzji osiąga z prawdopodobieństwem λ_i,

$\overline{m}_{\lambda_i}$ – średnia wartości wyników, większych od t_{λ_i},

d_{ki}^{+} – zmienne pomocnicze.

Przykład 1

Wyznaczenie decyzji efektywnych w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu.

Rozpatrzmy problem wyboru portfela ryzykownych projektów, których realizacja oznacza ponoszenie strat. Brane są pod uwagę 4 projekty A, B, C, D. Dla każdego z projektów zostały oszacowane straty przy 5 możliwych stanach otoczenia, przedstawione w tabeli 1.

Tabela 1

Dane do przykładu 1

	Prawdopodobieństwo	Straty			
		projekt A	projekt B	projekt C	projekt D
	$\mathrm{p_k}$	$\mathrm{c_{1k}}$	$\mathrm{c_{2k}}$	$\mathrm{c_{3k}}$	$\mathrm{c_{4k}}$
Stan otoczenia 1	0,2	200	90	50	350
Stan otoczenia 2	0,1	100	90	40	150
Stan otoczenia 3	0,2	30	85	40	100
Stan otoczenia 4	0,2	80	40	170	120
Stan otoczenia 5	0,3	100	120	70	50

Zmienne decyzyjne x_j (j=1,...,4) będą oznaczały udziały odpowiednich projektów w portfelu; wtedy zbiór decyzji dopuszczalnych będzie określony ograniczeniami: $\sum_{j=1}^{4} x_j = 1;\ x_j \geq 0$.

Przy każdym ze stanów otoczenia jesteśmy zainteresowani mniejszą wartością strat: $f(\mathbf{x},\mathbf{c}) \xrightarrow[\mathbf{x}]{} \min$.

W celu wygenerowania portfeli efektywnych w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu został sformułowany wielokryterialny problem minimalizacji największych częściowych średnich strat (2) dla zbioru poziomów prawdopodobieństwa $\lambda = \{0{,}1;\ 0{,}4;\ 0{,}6;\ 0{,}9;\ 1\}$.

Do poszukiwania i wyboru rozwiązań efektywnych sformułowanego problemu wielokryterialnego została wykorzystana dwu-referencyjna procedura optymalizacji wielokryterialnej.

W tabeli 2 przedstawiono trzy rozwiązania znalezione za pomocą dwu-referencyjnej procedury.

Tabela 2

Rozwiązania efektywne

	Wartości zmiennych decyzyjnych			
	x_1 (projekt A)	x_2 (projekt B)	x_3 (projekt C)	x_4 (projekt D)
Rozwiązanie 1	0	0,45	0,55	0
Rozwiązanie 2	0	0,54	0,38	0,08
Rozwiązanie 3	0	0	1	0

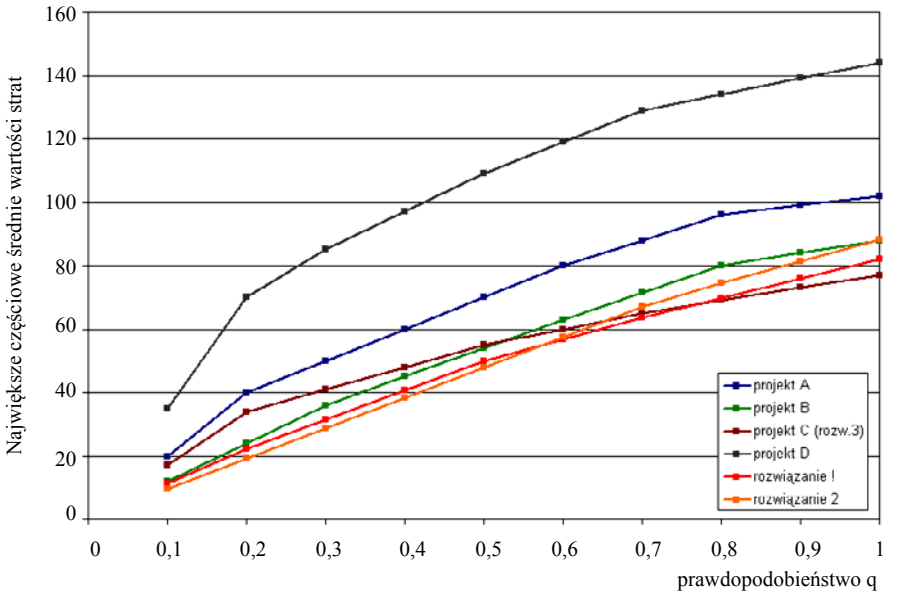
W tabeli 3 przedstawiono wartości funkcji $S_Y^{(-2)}(q)$ – największe częściowe średnie wartości funkcji f(**x,c**).

Tabela 3

Największe częściowe średnie wartości strat

	Prawdopodobieństwo q									
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
Projekt A	20	40	50	60	70	80	88	96	99	102
Projekt B	12	24	36	45	54	63	72	80	84	88
Projekt C	**17**	**34**	**41**	**48**	**55**	**60**	**65**	**69**	**73**	**77**
Projekt D	35	70	85	97	109	119	129	134	139	144
Rozwiązanie 1	**11**	**22**	**32**	**41**	**50**	**57**	**64**	**70**	**76**	**82**
Rozwiązanie 2	**10**	**19**	**29**	**38**	**48**	**57**	**67**	**75**	**81**	**88**

Na rys. 1 przedstawiono rozkłady największych częściowych średnich wartości strat projektów i wyznaczonych portfeli.



Rys. 1. Wartości największych częściowych średnich wartości strat

Jak można zauważyć w tabeli 3 i na rys. 1, wszystkie znalezione za pomocą problemu wielokryterialnego rozwiązania problemu minimalizacji strat są efektywne w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu. Rozkład prawdopodobieństwa strat projektu B, który nie jest zdominowany przez rozkłady prawdopodobieństwa strat innych projektów, jest zdominowany przez rozkłady prawdopodobieństwa strat obu znalezionych portfeli. Przedstawione wyniki pokazują jednocześnie możliwość polepszenia rozkładów prawdopodobieństwa strat w sensie odwrotnej dominacji stochastycznej przez dywersyfikację.

W przypadku problemów z nieprzeliczalnym i nieskończonym zbiorem decyzji dopuszczalnych, zbiór decyzji efektywnych też jest w ogólnym przypadku zbiorem nieprzeliczalnym i nieskończonym, dlatego ważne staje się wspomaganie wyboru decyzji zadowalającej, w ogólnym przypadku decyzji finalnej, ze zbioru decyzji efektywnych. Metody wspomagania decyzji wielokryterialnych pozwalają nie tylko wygenerować rozwiązania efektywne, ale i wspomagać wybór rozwiązania finalnego zgodnie z preferencjami decydenta.

Przeprowadzona w pracy [10] analiza zgodności struktury preferencji wobec kryteriów przedstawionych zadań wielokryterialnych z odpowiednimi modelami preferencji decydenta w warunkach ryzyka pozwala twierdzić, że zastosowanie wielokryterialnych metod wyboru decyzji do tych zadań umożliwia także wybór decyzji finalnej zgodnie z preferencjami decydenta w warunkach ryzyka.

Podsumowanie

W opracowaniu pokazano wykorzystanie relacji dominacji stochastycznej do analizy decyzji w problemach sformułowanych w kategoriach strat. Zademonstrowano zastosowanie odwrotnej relacji dominacji stochastycznej drugiego rzędu w problemach decyzyjnych z dziedziny ubezpieczeń i zarządzania zapasami. Opisano zastosowanie wielokryterialnego zadania maksymalizacji największych częściowych średnich wartości rozkładu (sformułowane przez W. Ogryczaka dla zadań optymalnej lokalizacji) do generowania decyzji efektywnych w sensie odwrotnej relacji dominacji stochastycznej drugiego rzędu.

Przedstawione w opracowaniu wyniki świadczą o tym, że zaproponowane przez W. Ogryczaka podejście do generowania decyzji efektywnych w sensie relacji dominacji stochastycznej drugiego rzędu, a polegające na wykorzystaniu metod analizy i optymalizacji wielokryterialnej, można zastosować nie tylko do problemów podejmowania decyzji w warunkach ryzyka sformułowanych w kategoriach zysku, ale i do problemów sformułowanych w kategoriach strat. W naukach aktuarialnych ostatnio można zauważyć wzrost zainteresowania możliwością uwzględnienia wzajemnej zależności rozkładów prawdopodobieństwa parametrów losowych przy analizie portfeli kontraktów ubezpieczeniowych, gdzie wcześniej powszechnie było stosowane założenie o niezależności tych parametrów. Możliwość uwzględnienia wzajemnej zależności parametrów losowych w procesie wyznaczenia decyzji efektywnych, w sensie relacji dominacji stochastycznej w problemach sformułowanych w kategoriach strat, otwiera nowe możliwości zastosowania analizy portfelowej w tego typu problemach.

Wykorzystanie metod optymalizacji wielokryterialnej pozwala wspomagać wybór decyzji w warunkach ryzyka, zgodnie z koncepcją ograniczonej racjonalności, i w otwartych problemach podejmowania decyzji. To pozwala wykorzystać podejścia do analizy problemów wielokryterialnych w warunkach ryzyka oparte na wykorzystaniu koncepcji dominacji stochastycznej także w problemach z nieskończonym i nieprzeliczalnym zbiorem decyzji dopuszczalnych.

Literatura

1. Bulinskaya E.V. (2004). Stochastic Orders and Inventory Problems. International Journal of Production Economics, 88, 125-135.
2. Denuit M., Dhaene J., Goovaerts M., Kaas R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. J. Wiley&Sons, West Sussex.
3. Kahneman D., Tversky A. (1979). Prospect Theory: An Analysis of Decision under Risk. Econometrica, 47, 263-291.
4. Levy H. (1992). Stochastic Dominance and Expected Utility: Survey and Analysis. Management Science, 33, 555-593.
5. Ogryczak W. (1997). Wielokryterialna optymalizacja liniowa i dyskretna. Uniwersytet Warszawski, Warszawa.
6. Ogryczak W. (2002a). Multiple Criteria Optimization and Decisions under Risk. Control and Cybernetics, 31.
7. Ogryczak W., Zawadzki M. (2002b). Zadania lokalizacyjne z kryterium minimalizacji najgorszej średniej warunkowej. Zeszyty Naukowe. Politechnika Śląska, Gliwice, 136.
8. Analiza i metody zmniejszania ryzyka w polskim systemie ubezpieczeń majątkowych. (1998). Red. W. Ronka-Chmielowiec. AE, Wrocław.
9. Simon H.A. (1979). Rational Decision Making in Business Organizations. The American Economic Review, 69, 493-513.
10. Sobotka O. (2008). Wspomaganie decyzji w warunkach ryzyka z wykorzystaniem koncepcji dominacji stochastycznej. Uniwersytet Ekonomiczny, Wrocław.
11. Trzaskalik T., Trzpiot G., Zaraś K. (1998). Modelowanie preferencji z wykorzystaniem dominacji stochastycznych. AE, Katowice.
12. Wang S., Young W. (1998). Ordering Risks: Expected Utility Theory Versus Yaari's Dual Theory of Risk. Insurance: Mathematics and Economics, 22, 145-161.