INNOWACJE W FINANSACH I UBEZPIECZENIACH – – METODY MATEMATYCZNE I EKONOMETRYCZNE

Studia Ekonomiczne

ZESZYTY NAUKOWE WYDZIAŁOWE

UNIWERSYTETU EKONOMICZNEGO

W KATOWICACH

INNOWACJE W FINANSACH I UBEZPIECZENIACH – – METODY MATEMATYCZNE I EKONOMETRYCZNE

Redakcja
Ewa Dziwok i Jerzy Mika

Katowice 2013

Komitet Redakcyjny
Krystyna Lisiecka (przewodnicząca), Anna Lebda-Wyborna (sekretarz),
Florian Kuźnik, Maria Michałowska, Antoni Niederliński, Irena Pyka,
Stanisław Swadźba, Tadeusz Trzaskalik, Janusz Wywiał, Teresa Żabińska

Komitet Redakcyjny Wydziału Finansów i Ubezpieczeń
Irena Pyka (redaktor naczelny), Daniel Iskra (sekretarz),
Halina Henzel, Halina Zadora, Maria Smejda

Rada Programowa
Lorenzo Fattorini, Mario Glowik, Gwo-Tzeng,
Zdeněk Mikolaš, Marian Noga, Bronisław Micherda, Miloš Král

Redaktor
Barbara Cebo

Skład
Grzegorz Bociek

© Copyright by Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach 2013

ISSN 2083-8611

Wszelkie prawa zastrzeżone. Każda reprodukcja lub adaptacja całości
bądź części niniejszej publikacji, niezależnie od zastosowanej
techniki reprodukcji, wymaga pisemnej zgody Wydawcy

WYDWNICTWO UNIWERSYTETU EKONOMICZNEGO W KATOWICACH
ul. 1 Maja 50, 40-287 Katowice, tel. 32 257-76-30, fax 32 257-76-43
www.wydawnictwo.ue.katowice.pl, e-mail: wydawnictwo@ue.katowice.pl

SPIS TREŚCI

Maria Forlicz
Uniwersytet Ekonomiczny we Wrocławiu

EFEKT SKALI W DYSKONTOWANIU SUBIEKTYWNYM

Wprowadzenie

Według teorii zdyskontowanej użyteczności, czyli klasycznie przyjmowanej teorii podejmowania decyzji w czasie, każda kwota/wielkość dyskontowana w jednakowych warunkach otoczenia powinna być dyskontowana przy użyciu jednakowej stopy dyskonta. Występowanie efektu skali, dość często opisywanego zjawiska, jest niezgodne z tą teorią. Efekt ten polega na tym, że większe wielkości są słabiej dyskontowane w czasie niż mniejsze. Subiektywna stopa dyskonta przy dyskontowaniu 100 zł jest wyższa od subiektywnej stopy dyskonta używanej przy dyskontowaniu 100 000 zł.

1. Wyniki dotychczasowych badań

Jedne z pierwszych i najczęściej opisywanych badań mających udowodnić występowanie lub też zaprzeczyć istnieniu efektu skali przeprowadzili Benzion, Rapoport i Yagil [1] w 1989 r. Przebadali oni 204 studentów ekonomii zadając każdemu 80 pytań (64 dotyczyło dyskontowania, 16 osoby wypełniającej kwestionariusz). W pytaniach zmieniały się scenariusze, kwoty i czasy odroczenia nagród. Scenariusz A polegał na odłożeniu w czasie wpływu pewnej kwoty, scenariusz B na odroczeniu zapłaty (wypływu), scenariusz C na przyspieszeniu wpływu, scenariusz D na przyspieszeniu zapłaty (wypływu). Za każdym razem ankietowany musiał podać kwotę, która spowoduje, że opcja natychmiastowa będzie dla niego tak samo interesująca, jak opcja późniejsza. Dyskontowano kwoty 40, 200, 1000 i 5000 dolarów, a odroczenia/przyspieszenia wynosiły pół roku, 1, 2

i 4 lata. Przeprowadzający badanie wyliczyli, że średnie stopy dla poszczególnych sum wyniosły: 29% dla 40 dolarów, 21,7% dla 200 dolarów, 20% dla 1000 dolarów, 14,3% dla 5000 dolarów. Dla wszystkich czterech scenariuszy wpływ wysokości sumy na wysokość stóp dyskonta był istotny ($p < 0{,}001$), stopy te zmieniały się odwrotnie do wysokości dyskontowanych kwot.

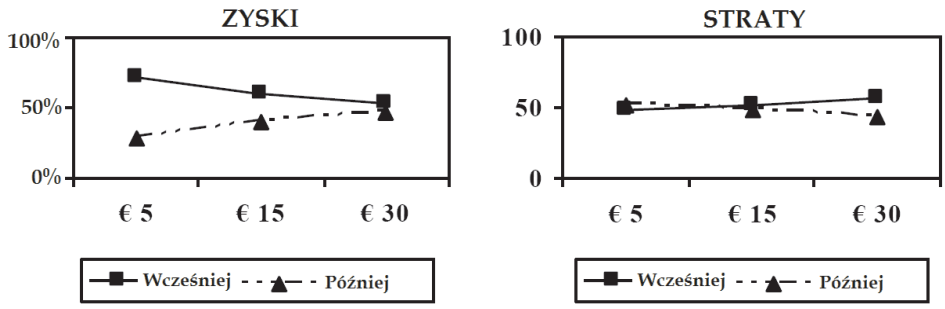
Meyerson i Green [5] przeprowadzili badanie, w którym brało udział jedynie 12 osób indagowanych indywidualnie. Dyskontowane były kwoty 1000 i 10 000 dolarów, a stopy dyskonta były określane na podstawie momentu, w którym badany zmieniał wybór z nagrody możliwej do otrzymania wcześniej na późniejszą lub na odwrót (w zależności od kolejności prezentowanych wariantów wyboru – zmiany kwot możliwych do otrzymania były prezentowane natychmiast lub w kolejności malejącej albo rosnącej). Po zestawieniu otrzymanych w ten sposób danych określono indywidualne preferencje każdego z badanych (oznacza to, że tym razem nie wnioskowano na średnich grupowych). Większość obiektów dyskontowała niskie kwoty silniej, chociaż z tego wzoru zachowania wyłamały się dwie osoby.

Kolejne badania pozwalające lepiej poznać efekt skali przeprowadzili Kirby i Maraković [4]. Wysłali oni listy do 2000 studentów swojego college'u z prośbą o wypełnienie kwestionariusza. Odpowiedziały im 672 osoby, z czego prawidłowo wypełnionych było 621 arkuszy. Kwestionariusz zawierał 21 pytań. Każde miało taką samą formę: „Co byś wolał: x dziś czy y za t dni?" ($y > x$). Wysokość odroczonej kwoty wahała się od 30 do 85 dolarów, a długość odroczenia od 10 do 75 dni. Ankietowani zostali powiadomieni, że po udzieleniu przez nich odpowiedzi na zawarte w kwestionariuszu pytania jedna z wylosowanych osób otrzyma rzeczywistą wypłatę kwoty, którą wybrała w losowo wybranym jednym pytaniu. Chciano w ten sposób zmotywować studentów do dogłębnego zastanowienia się nad swoim wyborem i własnymi preferencjami. Spośród badanych, 65 osób zawsze wybierało tylko natychmiastową nagrodę lub tylko odroczoną. Dodatkowo 28 osób zawsze wybierało tylko natychmiastową nagrodę lub tylko odroczoną w obrębie kwoty zakwalifikowanej do tych o danej skali (małej, średniej lub dużej), co uniemożliwiło wyliczenie dla tych osób subiektywnych stóp dyskonta. Jeśli chodzi o pozostałe 528 obiektów, to tak jak w poprzednich badaniach również i w tym średnie stopy dyskonta malały wraz ze wzrostem kwoty, a zależność ta była statystycznie istotna.

Johnson i Bickel [3] obiecali wszystkim swoim ankietowanym wypłatę pieniędzy zgodnie z dokonanymi przez nich wyborami. Potencjalne wypłaty nie dotyczyły wszystkich pytań, a jedynie ich części (autorzy porównywali zachowanie ankietowanych przy hipotetycznych i rzeczywistych wypłatach). Osoby biorące

udział w badaniu wiedziały, przy których pytaniach istnieje możliwość otrzymania realnych pieniędzy i o tym, że pieniądze zostaną im wypłacone po jednym razie na każdą dyskontowaną kwotę 10, 25, 100 i 250 dolarów. Do przeprowadzenia badania Jaohnson i Bickel użyli komputera, który według pewnego algorytmu tak ustawiał kwoty możliwe do otrzymania natychmiast, aby jak najlepiej przybliżyć wartość punktu obojętności. W ten sposób zostało przebadanych 6 osób. Dla 4 z nich stopy dyskonta malały zawsze wraz ze wzrostem wysokości dyskontowanej kwoty. Dla wszystkich obiektów stopa dyskontowa przy dyskontowaniu 10 dolarów była wyższa od stopy dyskonta przy dyskontowaniu 250 dolarów.

Faralla, Benuzzi, Nichelli i Dimitri [2] porównywali zachowanie się subiektywnych stóp dyskonta w dwóch różnych sytuacjach – ponoszenia straty i osiągania zysku. 25 osób wypełniło ankietę zawierającą 240 pytań. Badanych zapewniono, że po jednej z ich decyzji dotyczących zysków i dotyczących strat zostanie zrealizowanych (oprócz tego zapłacono im za udział w badaniu). Ankietowani musieli zawsze wybierać pomiędzy mniejszą kwotą dostępną wcześniej (choć niekoniecznie natychmiast) i większą dostępną później. Częstość wybierania poszczególnych opcji w zależności od dyskontowanej kwoty obrazuje rys. 1.



Rys. 1. Częstość wyboru poszczególnych opcji dla zysków i strat w zależności od dyskontowanej kwoty

Źródło: Opracowanie własne na podstawie: [2].

W 2006 r. trójka amerykańskich badaczy [6] sprawdziła, czy efekt skali dotyczy jedynie dyskontowania wartości pieniężnych czy także innych dóbr, a dokładniej pożywienia. Analizując otrzymane wyniki można powiedzieć, że pieniądze faktycznie są silniej dyskontowane, gdy jest ich mniej, jednak w przypadku pożywienia ciężko już się dopatrzyć takiej prawidłowości. Według ich wyliczeń, różnica w dyskontowaniu dla małych i dużych kwot, a także dla małych i dużych ilości pożywienia była w obydwu przypadkach nieistotna statystycznie.

2. Badania własne

W niniejszym opracowaniu analizie zostaną poddane wyniki uzyskane w trzech różnych badaniach z lat 2008, 2009 i 2011, przeprowadzonych w celu zbadania innych aspektów dyskontowania subiektywnego, ale pozwalających również na obserwowanie zachowania subiektywnych stóp dyskonta w zależności od dyskontowanej kwoty.

Badanie z 2008 r. przeprowadzono na 170 osobach. Była to ankieta w formie arkusza zawierająca 12 pytań umożliwiających wyliczenie subiektywnych stóp dyskonta* ankietowanych oraz 4 pytania dotyczące bezpośrednio ankietowanych (płeć, wiek, miejsce zamieszkania, sytuacja finansowa). Część pytań była otwarta, część zamknięta. Odpowiedzi, które będą brane pod uwagę w tej pracy dotyczą pytań otwartych.

W pierwszej kolejności porównane zostaną odpowiedzi na pytania najprostszego typu:

1. Zmarł twój bardzo daleki krewny, który jednak o tobie nie zapomniał i zapisał ci w spadku 20 tys. zł. Miną jednak 3 miesiące zanim otrzymasz pieniądze. Syn zmarłego proponuje ci jednak, że chętnie dziś odkupi od ciebie prawo do spadku? Ile musiałby ci zaoferować pieniędzy, żebyś się zgodził/zgodziła?
2. Za 3 miesiące kończy ci się 3-letnia lokata w banku, otrzymasz wtedy 1000 zł. Bank proponuje wypłatę pieniędzy już teraz, ale po potrąceniu części zysku. Ile musiałby ci wypłacić bank, żebyś się zgodził na wypłatę pieniędzy teraz, a nie za 3 miesiące?

W obydwu pytaniach ankietowany ma otrzymać pewną kwotę (większą lub mniejszą) za dokładnie 3 miesiące. Ponadto pytania zostały skonstruowane tak, aby ankietowany nie miał najmniejszych wątpliwości co do tego, że faktycznie pieniądze będą jego. Teoretycznie w jednakowych warunkach ankietowani powinni żądać takich samych stóp dyskonta dla kwoty 20 000 zł i dla 1000 zł. Obliczenia, których wyniki znajdują się w tab. 1 i 2 pokazują, że jest inaczej.

* Stopy dyskonta wyliczano według wzoru $r = \frac{FV}{PV} - 1$, gdzie FV to kwota otrzymywana później, PV to kwota otrzymywana wcześniej.

Tabela 1

Subiektywne 3-miesięczne stopy dyskonta dla małej i dużej kwoty w warunkach pewności

	Przy pełnych danych			Bez obserwacji odstających			Po usunięciu ujemnych stóp dyskonta		
Stopy dyskonta	średnia	mediana	n	średnia	mediana	n	średnia	mediana	n
Dla 20 000 zł	1034,44%	0%	136	-7,09%	0%	132	5,54%%	1,01%	90
Dla 1000 zł	9012,08%	5,26%	111	5,578%	5,26%	106	11,01%	11,11%	90

Tabela 2

Procent odpowiedzi o różnym stosunku stóp dyskonta dla małych i dużych kwot w warunkach pewności

	Przy pełnych danych	Bez obserwacji odstających	Po usunięciu ujemnych stóp dyskonta
Odpowiedzi dające wyższą stopę dyskonta dla małej kwoty	65,05%	65,98%	66,67%
Odpowiedzi dające wyższą stopę dyskonta dla dużej kwoty	22,33%	20,62%	14,28%
Odpowiedzi dające równe stopy dyskonta dla małej i dużej kwoty	12,62%	13,4%	19,05%
Liczba obserwacji	103	97	63

Jak widać, bez względu na to, czy weźmie się pod uwagę wszystkie odpowiedzi bez wyjątku czy tylko te, które mieszczą się w przedziale zmienności albo jedynie te, które dają dodatnie stopy dyskonta (ujemne subiektywne stopy dyskonta, z punktu widzenia czysto ekonomicznego, są przejawem nieracjonalności) zarówno średnie, jak i mediany stóp dyskonta są wyższe dla niższej kwoty. Ponadto około 66% ankietowanych przystała na niższe stopy dyskonta dla wyższej kwoty niż dla niskiej.

Podobnej analizie poddano odpowiedzi na pytania, w których ponownie odroczenie wynosiło 3 miesiące, jednak dodano (nie wprost) pewien element ryzyka. Tym razem dyskontowano kwoty 10 000 zł i 100 zł:

1. Prowadzisz mały biznes. Jeden z odbiorców twoich towarów ma ci oddać za 3 miesiące 10 000 zł. Pewna firma chce od ciebie odkupić ten dług. Ile pieniędzy musiałaby ci dać dzisiaj, abyś odsprzedał/odsprzedała ten dług?
2. Dostałeś na urodziny bilet na koncert pewnej grupy muzycznej, który odbędzie się za 3 miesiące. Nie lubisz tego typu muzyki, a chłopak kuzynki proponuje, że odkupi od ciebie bilet. Myślisz, że tuż przed koncertem mógłbyś go sprzedać za prawdopodobnie 100 zł. Za ile odsprzedałbyś/odsprzedałabyś go chłopakowi w tej chwili?

Tabela 3

Subiektywne trzymiesięczne stopy dyskonta dla małej i dużej kwoty w warunkach niepewności

	Przy pełnych danych			Bez obserwacji odstających			Po usunięciu ujemnych stóp dyskonta		
Stopy dyskonta	średnia	mediana	n	średnia	mediana	n	średnia	mediana	n
Dla 10 000 zł	10,33%	0%	132	-5,32%	0%	129	7,07%	5,26%	79
Dla 100 zł	20,85%	25%	123	20,85%	25%	123	33,73%	25%	100

Tabela 4

Procent odpowiedzi o różnym stosunku stóp dyskonta dla małych i dużych kwot w warunkach niepewności

	Przy pełnych danych	Bez obserwacji odstających	Po usunięciu ujemnych stóp dyskonta
Odpowiedzi dające wyższą stopę dyskonta dla małej kwoty	74,31%	76%	77,59%
Odpowiedzi dające wyższą stopę dyskonta dla dużej kwoty	15,60%	13%	6,90%
Odpowiedzi dające równe stopy dyskonta dla małej i dużej kwoty	10,09%	11%	15,52%
Liczba obserwacji	109	100	58

Tak samo, jak w przypadku poprzedniej pary pytań, również tu widać, że bez względu na to, którą wersję zestawu danych przyjąć, średnie i mediany stóp dyskonta były wyższe dla małej kwoty. Około 75% ankietowanych akceptowało niższe stopy dyskonta dla większych kwot.

Mniej jednoznaczne wyniki (tab. 5, 6) dały odpowiedzi na dwa kolejne pytania:

1. Jesteś pracownikiem biurowym w stoczni, która ostatnio przeżywa kłopoty finansowe. Teraz stocznię przejął nowy inwestor i obiecał, że za rok wypłaci pracownikom na twoim szczeblu premię w wysokości 32 000 zł. Istnieje możliwość zamiany tej kwoty na mniejszą (negocjowalną), ale wypłacaną już teraz. Jak uważasz, ile pieniędzy powinieneś dostać teraz, żeby opłacało ci się zrezygnować z tych 32 000 zł za rok?
2. Byłeś na wakacjach w Tajlandii. Z wyjazdu zostało ci 2100 batów (waluta tajlandzka). Polskie kantory nie wymieniają tej waluty. Pilot wycieczki proponuje, że kupi od ciebie całą pulę, ale ty wiesz, że za rok do Tajlandii leci twój kolega z pracy i z chęcią przed wylotem wszedłby w posiadanie batów. Podejrzewasz, że zgodziłby się je kupić za 150 zł. Ile musiałby ci dać pilot, żebyś oddał mu walutę?

Tabela 5

Subiektywne roczne stopy dyskonta dla małej i dużej kwoty w warunkach niepewności

	Przy pełnych danych			Bez obserwacji odstających			Po usunięciu ujemnych stóp dyskonta		
Stopy dyskonta	średnia	mediana	n	średnia	mediana	n	średnia	mediana	n
Dla 32 000 zł	875,22%	7%	126	8,60%	7%	112	11,65%	6,67%	99
Dla 150 zł	-3,08%	0%	123	-3,08%	0%	123	19,52%	15%	72

Tabela 6

Procent odpowiedzi o różnym stosunku stóp dyskonta dla małych i dużych kwot w warunkach niepewności

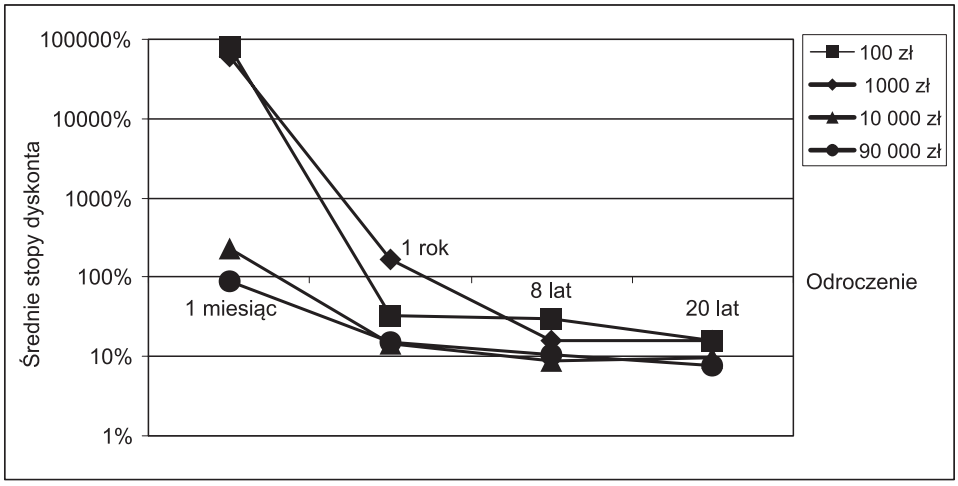
	Przy pełnych danych	Bez obserwacji odstających	Po usunięciu ujemnych stóp dyskonta
Odpowiedzi dające wyższą stopę dyskonta dla małej kwoty	35,96%	40%	62,90%
Odpowiedzi dające wyższą stopę dyskonta dla dużej kwoty	61,40%	58%	33,87%
Odpowiedzi dające równe stopy dyskonta dla małej i dużej kwoty	2,63%	2%	3,23%
Liczba obserwacji	114	100	62

W ostatniej analizowanej parze pytań zarówno średnie, jak i mediany były wyższe dla dużej kwoty, jednak tylko w przypadku, gdy brano pod uwagę pełne dane lub dane zredukowane o obserwacje odstające także procent osób akceptujących wyższe stopy dyskonta dla wyższej kwoty był dla tych danych większy. Gdy wzięto pod uwagę dane, z których wyeliminowano odpowiedzi dające ujemne stopy dyskonta oraz obserwacje odstające, sytuacja znowu zaczęła przypominać tę z poprzednich par pytań – zarówno średnia, jak i mediana stóp dyskonta były wyższe dla niższych kwot i ponad 62% par odpowiedzi stanowiły te o stopach dyskonta wyższych dla małych kwot.

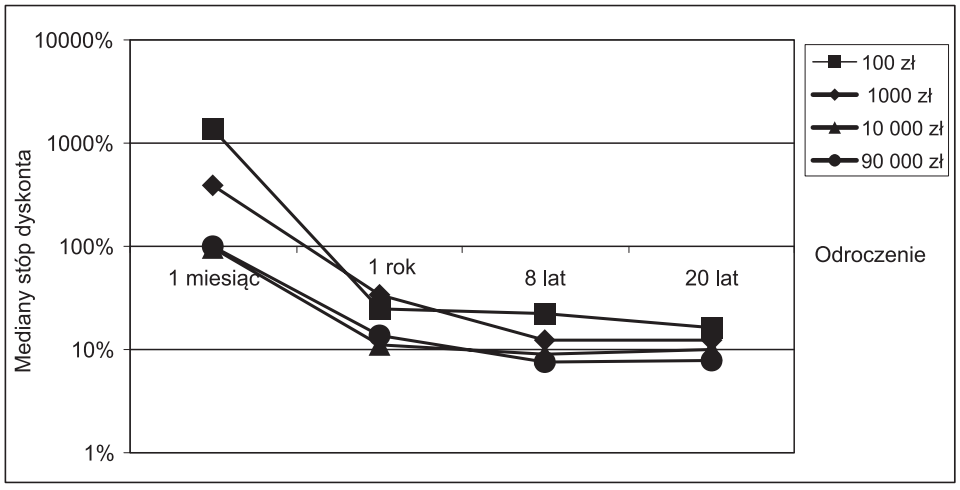
Ponieważ w opisywanym wyżej badaniu można by zakwestionować równość (w parach) prawdopodobieństwa zaistnienia określonych przepływów pieniężnych, jak również ich podobieństwo ze względu na aspekt psychologiczny postanowiono opisać badanie, w którym ta równość i podobieństwo nie mogą być negowane. W 2009 r. wśród 92 studentów przeprowadzono ankietę wyświetlając pytania na ekranach komputerów, tak że kiedy ankietowany odpowiedział na dane pytanie i przeszedł do następnego nie mógł się już cofnąć. Wszystkie pytania miały taką samą strukturę:

Ile byłbyś skłonny zapłacić teraz za zero-kuponową (czyli taką, która nie płaci odsetek) obligację Skarbu Państwa, której termin wygaśnięcia przypada za X lat, co oznacza, że otrzymasz za nią wtedy od Skarbu Państwa Y zł?

Dyskontowane kwoty (Y) to 100 zł, 1000 zł, 10 000 zł, 90 000 zł, a okresy odroczenia wypłaty (X) to 1 miesiąc, 1 rok, 8 lat i 20 lat. Wszystkie odpowiedzi poddano odpowiednim przekształceniom, aby otrzymać wartości stóp dyskonta wyrażone w skali roku. Następnie z zestawu danych usunięto obserwacje nietypowe, po czym wyliczono średnie i mediany stóp dyskonta dla każdej kwoty przy każdym odroczeniu (liczba obserwacji w zależności od kwoty i odroczenia waha się od 59 do 84). Wyniki tych obliczeń pokazują rys. 2 i 3.

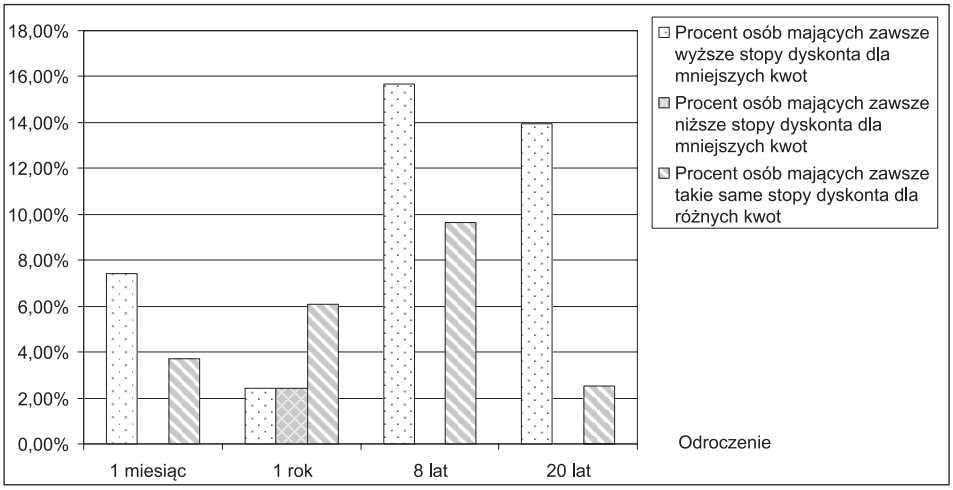


Rys. 2. Średnie stopy dyskonta dla różnych kwot i różnych odroczeń

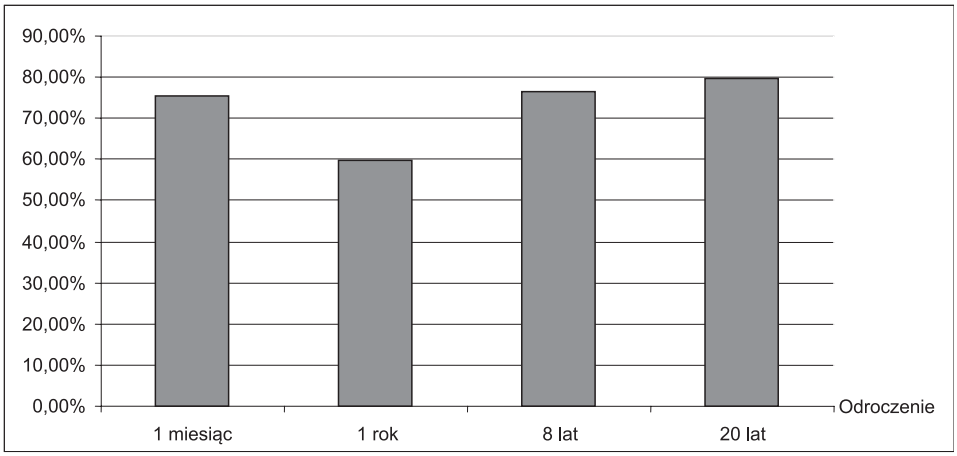


Rys. 3. Mediany stóp dyskonta dla różnych kwot i różnych odroczeń

Największe różnice w średnich i medianach stóp dyskonta dla różnych kwot występują przy jednomiesięcznym odroczeniu, natomiast najmniejsze dla dwudziestoletniego. Jak można zauważyć, średnie i mediany stóp dyskonta dla 100 zł (jak i dla 1000 zł) są zawsze wyższe niż dla kwoty 90 000. Do pewnego stopnia może to sugerować występowanie efektu skali, ale zdecydowano się również dokonać analizy na poziomie jednostki. Sprawdzono, dla jakiego odsetka osób stopy dyskonta są zawsze niższe dla wyższych kwot oraz dla ilu procent ankietowanych stopy dyskonta są wyższe dla 100 zł niż dla 90 000 zł. Wyniki tych obliczeń przedstawiają rys. 4 i 5.



Rys. 4. Procent osób o danej relacji stóp dyskonta dla małych i dużych kwot w zależności od okresu dyskontowania



Rys. 5. Procent osób o stopie dyskonta niższej dla kwoty 90 000 zł niż dla 100 zł

Mało który z ankietowanych wykazał się spójnością decyzji. Najwięcej, bo blisko 10% ankietowanych utrzymało równe stopy dyskonta dla wszystkich kwot przy 8-letnim odroczeniu. Przy tej samej długości odroczenia zanotowano też największy odsetek grup odpowiedzi o stopach dyskonta zawsze wyższych dla niższych kwot. W większości przypadków zachowanie się indywidualnych stóp dyskonta w zależności od kwoty było dość chaotyczne. Przyglądając się rys. 5 można jednak zauważyć, że zawsze (bez względu na odroczenie) odsetek osób o wyższych stopach dyskonta dla 100 zł niż dla 90 000 jest większy od 50% – ten odsetek waha się od 59,77% (dla 1 roku) do 80,23% (dla 20 lat). Dodatkowo odsetek osób o takich samych stopach dyskonta dla kwot 100 zł i 90 000 zł wahał się od 5,75% (dla 1 roku) do 11,49% (dla 8 lat).

Kolejne badanie przeprowadzono w październiku 2011 r. Ankietowani zostali podzieleni na dwie grupy – pytania dla pierwszej grupy były związane głównie z ujemnymi przepływami pieniężnymi, natomiast druga grupa miała do czynienia głównie z dodatnimi przepływami pieniężnymi. Łącznie rozdano 250 arkuszy ankiety (po równo w jednej i drugiej grupie). Niestety, dla sporej części badanych pytania okazały się za trudne, stąd liczba rzeczywistych obserwacji znacznie zmalała. Uzyskane dane podzielono na osiem części: analizowano roczne natychmiastowe (czyli dyskontujące z momentu za rok na moment obecny) stopy zwrotu przy dużych kwotach odpływających (np. sytuacja, w której ankietowany ma zapłacić za coś wcześniej lub później) od decydenta, małych kwotach odpływających od decydenta, dużych kwotach wpływających (np. sytuacja, w której ankietowany musi zdecydować czy sprzedać coś wcześniej czy później) do decydenta, małych kwotach wpływających do decydenta oraz roczne stopy dyskontujące z momentu za dwa lata na moment za rok przy dużych kwotach odpływających od decydenta, małych kwotach odpływających od decydenta, dużych kwotach wpływających do decydenta, małych kwotach wpływających do decydenta. Przedstawione poniżej obliczenia będą się odnosić do próbek o różnej liczebności – liczebność próbek zależy od liczby osób, które odpowiedziały na dane pytania oraz od liczby odpowiedzi, które trzeba było wyeliminować ze względu na ich nieracjonalność lub niedopuszczalność. Średnie i mediany stóp dyskonta dla każdej z ośmiu grup przedstawia tab. 7.

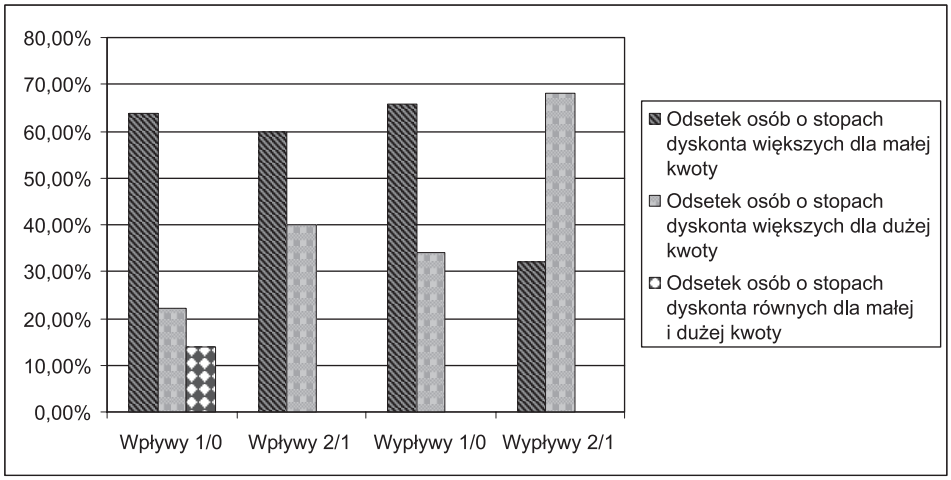
Tabela 7

Średnie i mediany rocznych stóp dyskonta dla małych i dużych kwot w zależności od kierunku przepływu pieniędzy i przesunięcia w czasie

		Małe kwoty			Duże kwoty		
	Stopa dyskonta	średnia	mediana	n	średnia	mediana	n
Wpływy	z momentu za rok na teraz	0,225029 (lub 0,077466 przy uwzględnieniu ujemnych stóp dyskonta)	0,111111 (lub 0,041667 przy uwzględnieniu ujemnych stóp dyskonta)	80 (110)	0,36702 (lub -0,01273 przy uwzględnieniu ujemnych stóp dyskonta)	0,115385 (lub -0,09375 przy uwzględnieniu ujemnych stóp dyskonta)	42 (114)
	z momentu za dwa lata na moment za rok	0,548698 (lub 0,526892 przy uwzględnieniu ujemnych stóp dyskonta)	0,14194 (lub 0,115384 przy uwzględnieniu ujemnych stóp dyskonta)	117 (120)	0,116505 (lub -0,16625 przy uwzględnieniu ujemnych stóp dyskonta)	0,068966 (lub -0,225 przy uwzględnieniu ujemnych stóp dyskonta)	36 (118)
Wypływy	z momentu za rok na teraz	1,371675	0,428571	118	13,27326 (lub 12,11515 przy uwzględnieniu ujemnych stóp dyskonta)	0,45 (lub 0,318182 przy uwzględnieniu ujemnych stóp dyskonta)	85 (93)
	z momentu za dwa lata na moment za rok	0,555557	0,199038	114	14,46287 (lub 13,91178 przy uwzględnieniu ujemnych stóp dyskonta)	0,347826 (lub 0,291667 przy uwzględnieniu ujemnych stóp dyskonta)	105 (109)

Jak można zauważyć, w przeciwieństwie do wyników poprzednich ankiet zarówno średnie, jak i mediany w przeważającej części są wyższe dla większych kwot (chyba że wzięłoby się pod uwagę ujemne stopy dyskonta, ale, jak już wspomniano, zgodnie z teorią stopy dyskonta nie powinny być ujemne). Ponadto widać także coś, co już jakiś czas temu zwróciło uwagę badaczy – ludzie silniej dyskontują w momencie, gdy ruch pieniędzy odbywa się od nich, a nie do nich.

Po wyeliminowaniu ze wszystkich par pytań (parę stanowiły pytania o takim samym kierunku przepływu pieniędzy i takim samym okresie i momencie, na który się dyskontowało, ale o różnych wielkościach kwot), w których przynajmniej raz pojawiła się ujemna stopa dyskonta obliczono dla każdej pary odsetek osób, które miały wyższe stopy dyskonta dla małych kwot, wyższe stopy dyskonta dla dużych kwot, równe stopy dyskonta bez względu na kwotę. Wyniki tych obliczeń przedstawia rys. 6.



Rys. 6. Procent osób o danej relacji stóp dyskonta dla małych i dużych kwot w zależności od kierunku przepływu pieniędzy i momentów, między którymi odbywa się dyskontowanie

Jak można zaobserwować na rys. 6, zazwyczaj odsetek osób o stopach dyskonta większych dla małej kwoty jest większy niż odsetek osób o stopach dyskonta wyższych dla dużych kwot (rzadko kiedy zdarza się sytuacja, że stopy te są równe). Sytuacja ulega jedynie odwróceniu, gdy mamy do czynienia z dyskontowaniem z roku drugiego na pierwszy, gdy pieniądze odpływają od ankietowanych. Trudno w tym momencie określić, jaka jest tego przyczyna.

Podsumowanie

Podobnie jak w badaniach opisanych w opracowaniach innych autorów, da się zauważyć, że efekt skali może faktycznie występować, jednak nie dotyka on całej populacji, a jedynie jej części.

Literatura

1. Benzion U., Rapoport A., Yagil J., Discount Rates Inferred from Decisions: An Experimental Study, „Management Science" 1989, Vol. 35.
2. Faralla V., Benuzzi F., Nichelli P., Dimitri N., Gain and Losses in Intertemporal Preferences: A Behavioural Study, Labsi Working Paper, 2010, www.labsi.org/wp/labsi33.pdf (23.09.2011).

3. Johnson M.W., Bickel W.K., Within-subject comparison of real and hypothetical money rewards in delay discounting, „Journal of the Experimental analysis of behavior" 2002, Vol. 77.
4. Kirby K.N., Maraković N.N., Delay-discounting probabilistic rewards: Rates decrease as amounts increase, „Psychonomic Bulletin & Review" 1996, Vol. 3.
5. Myerson J., Green L., Discounting of delayed rewards: Models of individual choice, „Joutnal of the Experimental Analysis of Behavior" 1995, Vol. 64.
6. Odum A.L., Baumann A.A.L., Rimington D.D., Discounting of delayed hypothetical money and food: Effects of amount, „Behavioural Processes" 2006, Vol. 73.

THE MAGNITUDE EFFECT IN SUBJECTIVE DISCOUNTING

Summary

According to discounted utility theory, which is considered normative theory of intertemporal choice, every amount of money or other asset discounted in the same circumstances should be discounted with the same discount rate. A lot of scientists claim that in reality subjective discount rates are not constant in constant circumstances. Magnitude effect, quite often described, may cause that bigger amounts are discounted less steeply than smaller ones. In this article results of previous research are described and compared to the results obtained by the author.

Aleksandra Iwanicka
Uniwersytet Ekonomiczny we Wrocławiu

WPŁYW STOPNIA ZALEŻNOŚCI POMIĘDZY WYSOKOŚCIĄ WYPŁAT NA PRAWDOPODOBIEŃSTWO RUINY W DWUKLASOWYM MODELU RYZYKA*

Wprowadzenie

W firmie ubezpieczeniowej zarządza się wieloma klasami ryzyka, w których część wpłat może być powodowana przez te same czynniki ryzyka. Czynniki te możemy traktować jako zewnętrzne czynniki ryzyka. Natomiast czynniki ryzyka powodujące wypłaty tylko w jednej klasie ryzyka można traktować jako wewnętrzne czynniki ryzyka. Jednoczesne oddziaływanie zewnętrznych czynników ryzyka na różne klasy ryzyka może skutkować jednoczesnym pojawianiem się wpłat w tych klasach, których wysokość może być zależna od siebie. Do tej pory w literaturze zależność ta nie była uwzględniana w modelach ryzyka dla kilku klas ryzyka, tzw. wieloklasowych modelach ryzyka. Celem niniejszej pracy jest przedstawienie wyników numerycznej analizy wpływu stopnia zależności pomiędzy jednocześnie pojawiającymi się wypłatami na prawdopodobieństwo ruiny w skończonym horyzoncie czasowym w modelu ryzyka dla dwóch klas ubezpieczeń.

Na początku zdefiniujmy model ryzyka dla jednej klasy ubezpieczeń w chwili t jako [2]:

$$R(t) = u + ct - S(t) \tag{1}$$

gdzie:
u – kapitał początkowy,

* Praca naukowa finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy nr 3361/B/H03/2010/38.

c – stała w jednostce czasu intensywność napływu składki z jednej klasy ryzyka,

$S(t)$ – suma zagregowanych wypłat do moment t włącznie, tj. $S(t) = \sum_{i=1}^{N(t)} X_i$,

gdzie $\{X_i\}_{i=1}^{\infty}$ jest ciągiem kolejnych niezależnych od siebie wypłat o tym samym rozkładzie,

$\{N(t)\}_{t\geq 0}$ – punktowy proces zliczający wpłaty.

Ponadto przyjmuje się założenie, że $\{X_i\}_{i=1}^{\infty}$ i $\{N(t)\}_{t\geq 0}$ są niezależne od siebie. Jeśli o procesie $\{N(t)\}_{t\geq 0}$ założymy, że jest jednorodnym procesem Poissona, to wówczas model ryzyka (1) nazywamy klasycznym modelem ryzyka. Aby zapewnić wypłacalność ubezpieczyciela o stałej c zakłada się, że musi spełniać warunek:

$$c > E(S(1))$$

który można zapisać w postaci:

$$c = (1+\theta)\, E(S(1))$$

gdzie:

θ – stała dodatnia wartość, którą nazywa się względnym współczynnikiem narzutu na bezpieczeństwo.

Przez ruinę w jednej klasie ryzyka rozumiemy spadek po raz pierwszy procesu ryzyka (1) poniżej zera w pewnej chwili t. W celu zdefiniowania prawdopodobieństwa ruiny oznaczmy czas ruiny jako:

$$T(u) = \inf\{t \geq 0 : R(t) < 0\}$$

Wówczas prawdopodobieństwo ruiny w skończonym horyzoncie czasowym H definiujemy następująco:

$$\Psi(u, H) = P(T(u) \leq H)$$

Konstrukcja modeli ryzyka dla kilku klas ryzyka jest uzależniona od pojęcia „ruiny" dla kilku klas ryzyka. W literaturze spotykane są dwa różne pojęcia „ruiny" dla kilku klas ubezpieczeń. Załóżmy, że mamy n klas ryzyka. Niech $\{R_i(t)\}_{t\geq 0}$ oznacza proces ryzyka dla i-tej klasy ryzyka, gdzie $i = 1,\ldots, n$. Wówczas przez ruinę możemy rozumieć spadek po raz pierwszy procesu ryzyka w co najmniej jednej z klas ryzyka, tj. $R_i(t) < 0$ po raz pierwszy w chwili t dla co najmniej jednego $i \in \{1,\ldots, n\}$. Wówczas model ryzyka dla n klas ryzyka definiuje się poprzez wielowymiarowy model ryzyka postaci:

$$\mathbf{R}(t) = \begin{pmatrix} R_1(t) \\ R_2(t) \\ \vdots \\ R_n(t) \end{pmatrix}$$

Poprzez ruinę można również rozumieć spadek po raz pierwszy w pewnej chwili t agregacji procesów ryzyka dla wszystkich klas ryzyka, tj. $R_1(t) + +\ldots + R_n(t) < 0$. Wówczas proces ryzyka dla n klas ryzyka definiujemy poprzez agregację procesów ryzyka dla tych klas, którą krótko nazywamy agregacją kilku klas ryzyka i która w chwili t ma postać:

$$R(t) = R_1(t) + R_2(t) + \ldots + R_n(t)$$

1. Analiza w dwuwymiarowym modelu ryzyka

Dwuwymiarowy model ryzyka uwzględniający zależność pomiędzy jednocześnie pojawiającymi wypłatami definiujemy w następujący sposób*:

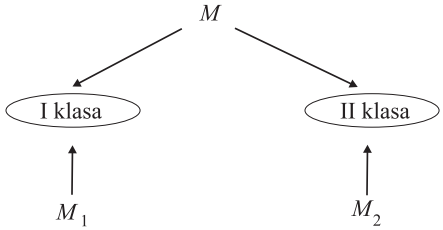
$$\begin{aligned}\mathbf{R}(t) &= \begin{pmatrix} R_1(t) \\ R_2(t) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} t - \begin{pmatrix} S_1(t) \\ S_2(t) \end{pmatrix} = \\ &= \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} t - \begin{pmatrix} \sum_{i=1}^{M_1(t)} X_{1i} + \sum_{i=1}^{M(t)} X'_{1i} \\ \sum_{i=1}^{M_2(t)} X_{2i} + \sum_{i=1}^{M(t)} X'_{2i} \end{pmatrix}\end{aligned} \tag{2}$$

gdzie:
u_i $(i = 1,2)$ – kapitał początkowy dla i-tej klasy ryzyka,
c_i $(i = 1,2)$ – stała dodatnia intensywność napływu składki w jednostce czasu dla i-tej klasy ryzyka,
$S_1(t) = \sum_{i=1}^{M_1(t)} X_{1i} + \sum_{i=1}^{M'(t)} X'_{1i}$ – suma zagregowanych wypłat do momentu t włącznie z pierwszej klasy ryzyka,
$S_2(t) = \sum_{i=1}^{M_2(t)} X_{2i} + \sum_{i=1}^{M'(t)} X'_{2i}$ – suma zagregowanych wypłat do momentu t włącznie z drugiej klasy ryzyka, gdzie $\{X_{1i}\}_{i=1}^{\infty}$ i $\{X'_{1i}\}_{i=1}^{\infty}$ są ciągami niezależnych od siebie wypłat (dodatnich zmiennych losowych) w pierwszej klasie ryzyka o tym samym rozkładzie z gęstością $f_1(x)$ i dystrybuantą $F_1(x)$,

* Podobny dwuwymiarowy model ryzyka jednak bez uwzględnienia zależności pomiędzy jednocześnie pojawiającymi się wypłatami został wprowadzony przez K.C. Yuen, J. Guo, X. Wu w [8].

$\{X_{2i}\}_{i=1}^{\infty}$ i $\{X'_{2i}\}_{i=1}^{\infty}$ – ciągi niezależnych wypłat (dodatnich zmiennych losowych) w drugiej klasie ryzyka o tym samym rozkładzie z gęstością $f_2(x)$ i dystrybuantą $F_2(x)$.

Ponadto zakładamy, że $\{X_{1i}\}_{i=1}^{\infty}$ i $\{X_{2i}\}_{i=1}^{\infty}$ są niezależnymi od siebie ciągami i niezależnymi od ciągów $\{X'_{1i}\}_{i=1}^{\infty}$ i $\{X'_{2i}\}_{i=1}^{\infty}$. Natomiast ciągi $\{X'_{1i}\}_{i=1}^{\infty}$ i $\{X'_{2i}\}_{i=1}^{\infty}$ są zależne od siebie w taki sposób, że tylko jednocześnie pojawiające się wypłaty w obu klasach są zależne od siebie, tzn. dla dowolnie wybranego $i \in N$ wypłaty X'_i i Y'_i są zależne od siebie, a struktura zależności pomiędzy nimi opisana jest funkcją łączącą $C(u, v)$, tj. dystrybuantę $F(x, y)$ łącznego rozkładu pary tych zmiennych losowych (X'_{1i}, X'_{2i}) można zapisać jako $F(x,y) = C(F_1(x), F_2(y))$. Funkcje łączące są doskonałym narzędziem do modelowania struktury zależności pomiędzy zmiennymi losowymi [3]. Ponadto $\{M_1(t)\}_{t\geq 0}$ ($\{M_2(t)\}_{t\geq 0}$) jest procesem punktowym zliczającym wypłaty powodowane przez wewnętrzne czynniki ryzyka w pierwszej (drugiej) klasie ryzyka, tj. wypłaty X_{1i} (X_{2i}), które są niezależne od wypłat z drugiej (pierwszej) klasy ryzyka (zob. rys. 1). Natomiast proces $\{M'(t)\}_{t\geq 0}$ jest procesem punktowym zliczającym zależne wypłaty X'_{1i} i X'_{2i} osobno w pierwszej i drugiej klasie ryzyka, które powodowane są przez zewnętrzne czynniki ryzyka, tj. czynniki ryzyka oddziałujące jednocześnie na obie klasy ryzyka (zob. rys. 1). Wszystkie procesy zliczające wypłaty są niezależne od siebie oraz od wszystkich ciągów wypłat. Poza tym, aby zapewnić wypłacalność ubezpieczyciela, stałe dodatnie intensywności napływu składki w obu klasach ryzyka muszą spełniać warunek: $c_i > ES_i(1)$ dla $i = 1,2$, który można zapisać w postaci: $c_i = (1+\theta_i)ES_i(1)$ dla $i = 1,2$, gdzie $\theta_i > 0$.



Rys. 1. Oddziaływanie zewnętrznych i wewnętrznych czynników ryzyka na dwie klasy ryzyka

W celu formalnego zdefiniowania prawdopodobieństwa ruiny w modelu ryzyka (2) oznaczmy czas ruiny jako:

$$T(u_1, u_2) = \inf\{t : R_1(t) < 0 \vee R_2(t) < 0\}$$

Wówczas prawdopodobieństwo ruiny w nieskończonym horyzoncie czasowym H definiujemy następująco:

$$\Psi(u_1, u_2; H) = P(T(u_1, u_2) < H) \tag{3}$$

Zajmijmy się teraz numeryczną analizą wpływu stopnia zależności pomiędzy wpłatami powodowanymi przez zewnętrzne czynniki ryzyka w obu klasach na prawdopodobieństwo ruiny na przykładzie wybranego modelu ryzyka (2). Przyjmijmy zatem model ryzyka (2), w którym o procesach zliczających wypłaty $\{M_1(t)\}_{t\geq 0}$, $\{M_2(t)\}_{t\geq 0}$ i $\{M'(t)\}_{t\geq 0}$ klasycznie zakładamy, że są jednorodnymi procesami Poissona z tą samą intensywnością równą 1. Niech $\theta_1 = \theta_2 = 0{,}05$ i $u_1 = u_2 = u$. Przyjmijmy, że funkcja łącząca opisująca strukturę zależności par (X'_{1i}, X'_{2i}) jest funkcją Claytona, która dobrze modeluje dodatnią zależność pomiędzy wypłatami. Analiza jest przeprowadzona osobno w dwóch przypadkach rozkładów wypłat: lekko-ogonowych i ciężko-ogonowych. W tab. 1 i 2 zawarte są wyniki symulacji prawdopodobieństwa ruiny (3) metodą Monte Carlo na podstawie 100 000 trajektorii przyjętego procesu ryzyka dla różnych horyzontów czasowych H i różnych wartości kapitału początkowego u i różnego stopnia zależności pomiędzy wpłatami X'_{1i} i X'_{2i}, wyrażonego współczynnikiem τ-Kendalla. Wyniki zawarte w tab. 1 otrzymano w przypadku, gdy wypłaty mają lekko-ogonowe rozkłady wypłat: w pierwszej klasie mają rozkład wykładniczy z parametrem równym 1, a w drugiej klasie mają rozkład Weibulla z parametrem skali równym 1 i z parametrem kształtu równym 10. Wyniki zestawione w tab. 2 otrzymano w przypadku, gdy wypłaty mają rozkłady ciężko-ogonowe: w pierwszej klasie mają rozkład logarytmiczno-normalny z parametrem log-skali równym 0,25 i z parametrem kształtu równym 0,125, natomiast w drugiej klasie mają rozkład Weibulla z parametrem skali równym 1 i z parametrem kształtu równym 0,8.

Tabela 1

Wyniki symulacji prawdopodobieństwa ruiny w przypadku lekko-ogonowych rozkładów wypłat

u/H	5	10	15	20
5				
$\tau = 0{,}1$	0,2595	0,4215	0,5091	0,5709
$\tau = 0{,}5$	0,2570	0,4214	0,5083	0,5698
$\tau = 0{,}9$	0,2556	0,4188	0,5072	0,5662
10				
$\tau = 0{,}1$	0,0387	0,1111	0,1781	0,2349
$\tau = 0{,}5$	0,0384	0,1108	0,1770	0,2339
$\tau = 0{,}9$	0,0384	0,1093	0,1761	0,2325
15				
$\tau = 0{,}1$	0,0052	0,0238	0,0520	0,0814
$\tau = 0{,}5$	0,0050	0,0237	0,0515	0,0797
$\tau = 0{,}9$	0,0051	0,0239	0,0507	0,0791

Tabela 2

Wyniki symulacji prawdopodobieństwa ruiny w przypadku ciężko-ogonowych rozkładów wypłat

u/H	5	10	15	20
5				
$\tau = 0,1$	0,3822	0,5438	0,6314	0,6814
$\tau = 0,5$	0,3816	0,5443	0,6275	0,6801
$\tau = 0,9$	0,3777	0,5436	0,6217	0,6750
10				
$\tau = 0,1$	0,1003	0,2199	0,3076	0,3762
$\tau = 0,5$	0,0989	0,2177	0,3058	0,3727
$\tau = 0,9$	0,0988	0,2169	0,3042	0,3698
15				
$\tau = 0,1$	0,0239	0,0748	0,1272	0,1813
$\tau = 0,5$	0,0238	0,0751	0,1274	0,1787
$\tau = 0,9$	0,0241	0,0747	0,1281	0,1777

Na podstawie analizy wyników zestawionych w tab. 1 i 2 można zauważyć, że przy ustalonym kapitale początkowym u i przy ustalonym horyzoncie czasowym H w większości przypadków obserwujemy niewielki spadek prawdopodobieństwa ruiny (3) wraz ze wzrostem stopnia zależności pomiędzy wypłatami X'_{1i} i X'_{2i}, jednak różnice te są nieznaczne.

2. Analiza w agregacji dwóch klas ryzyka

Agregację dwóch klas ryzyka uwzględniającą zależność pomiędzy jednocześnie pojawiającymi się wypłatami definiujemy w następujący sposób*:

$$R(t) = R_1(t) + R_2(t) = u + ct - S_1(t) - S_2(t) = \\ = u + ct - \left(\sum_{i=1}^{M_1(t)} X_{1i} + \sum_{i=1}^{M(t)} X'_{1i} \right) - \left(\sum_{i=1}^{M_2(t)} X_{2i} + \sum_{i=1}^{M(t)} X'_{2i} \right) \qquad (4)$$

gdzie:
u – kapitał początkowy wspólny dla obu klas ryzyka,
c – stała dodatnia intensywność napływu składki w jednostce czasu z obu klas ryzyka, która w celu zapewnienia wypłacalności ubezpieczyciela musi spełnić warunek:

* Podobny model ryzyka, jednak bez uwzględnienia zależności pomiędzy jednocześnie pojawiającymi się wypłatami był rozpatrywane m.in. przez S. Li i J. Garrido [7].

$c > E(S_1(1) + (1) + S_2(1))$, który można zapisać w postaci $c = (1+\theta)\, E(S_1(1) + (1) + S_2(1))$, gdzie θ jest dodatnią stałą.

Pozostałe oznaczenia i założenia pozostają takie same, jak w modelu ryzyka (2).

Na potrzeby przeprowadzenia analizy wpływu stopnia zależności pomiędzy jednocześnie pojawiającymi się wypłatami na prawdopodobieństwo ruiny w modelu ryzyka (4) przyjmiemy założenie z klasycznej teorii ruiny, że procesy $\{M_1(t)\}_{t\geq 0}$, $\{M_2(t)\}_{t\geq 0}$ i $\{M'(t)\}_{t\geq 0}$ są jednorodnymi procesami Poissona z intensywnościami odpowiednio λ_1, λ_2 i λ. Wówczas proces ryzyka (4) można przekształcić do klasycznego modelu ryzyka postaci [1]:

$$R'(t) = u + ct - \sum_{i=1}^{N(t)} Z_i$$

gdzie:

$\{N(t)\}_{t\geq 0}$ – jednorodny proces Poissona z intensywnością $\lambda_1 + \lambda_2 + \lambda$,

$\{Z_i\}_{i=1}^{\infty}$ – ciąg kolejnych dodatnich i niezależnych zmiennych losowych o tym samym rozkładzie z gęstością:

$$f_Z(x) = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda} f_1(x) + \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda} f_2(x) + \frac{\lambda}{\lambda_1 + \lambda_2 + \lambda} f_3(x) \qquad (5)$$

gdzie $f_3(x)$ jest gęstością zmiennej losowej $X'_{1i} + X'_{2i}$, pozostałe oznaczenia są takie same, jak w modelu (4).

Wówczas można próbować wykorzystać znane metody wyznaczania prawdopodobieństwa ruiny w klasycznym modelu ryzyka, które znajdują się w wielu pozycjach dotyczących klasycznej teorii ruiny*, jednak problem może stanowić wyznaczenie w (6) gęstości $f_3(x)$ sumy zależnych zmiennych losowych $X'_{1i} + X'_{2i}$.

Załóżmy, że wszystkie intensywności procesów Poissona $\{M_1(t)\}_{t\geq 0}$, $\{M_2(t)\}_{t\geq 0}$ i $\{M'(t)\}_{t\geq 0}$ są równe 1. Niech ponadto $\theta + 0{,}05$. W tab. 3 i 4 zawarte są wyniki symulacji prawdopodobieństwa ruiny otrzymane metodą Monte Carlo na podstawie 100 000 trajektorii procesu ryzyka (4) dla różnych wartości kapitału początkowego u i dla różnych horyzontów czasowych H, przy różnym stopniu zależności pomiędzy jednocześnie pojawiającymi się wpłatami X'_{1i} i X'_{2i} w obu klasach ryzyka, który mierzony jest współczynnikiem τ-Kendalla. W tab. 3 wyniki symulacji otrzymane są przy założeniu lekko-ogonowych rozkładów wypłat, natomiast w tab. 4 przy założeniu ciężko-ogonowych rozkładów wypłat. Przyjęte są takie same rozkłady, jak w analizie przeprowadzonej w drugim punkcie pracy.

* Na przykład w monografii S. Asmussena [2].

Tabela 3

Wyniki symulacji prawdopodobieństwa ruiny w przypadku lekko-ogonowych rozkładów wypłat

u/H	5	10	15	20
5				
$\tau = 0,1$	0,3231	0,4511	0,5129	0,5526
$\tau = 0,5$	0,3296	0,4531	0,5134	0,5536
$\tau = 0,9$	0,3311	0,4558	0,5141	0,5547
10				
$\tau = 0,1$	0,0964	0,1912	0,2605	0,3068
$\tau = 0,5$	0,0970	0,1950	0,2610	0,3073
$\tau = 0,9$	0,0976	0,1973	0,2617	0,3079
15				
$\tau = 0,1$	0,0214	0,0721	0,1165	0,1549
$\tau = 0,5$	0,0226	0,0724	0,1186	0,1570
$\tau = 0,9$	0,0231	0,0738	0,1215	0,1577

Tabela 4

Wyniki symulacji prawdopodobieństwa ruiny w przypadku ciężko-ogonowych rozkładów wypłat

u/H	5	10	15	20
5				
$\tau = 0,1$	0,4106	0,5233	0,5827	0,6170
$\tau = 0,5$	0,4102	0,5280	0,5826	0,6210
$\tau = 0,9$	0,4107	0,5284	0,5848	0,6198
10				
$\tau = 0,1$	0,1695	0,2835	0,3552	0,4015
$\tau = 0,5$	0,1725	0,2883	0,3615	0,4076
$\tau = 0,9$	0,1760	0,2903	0,3580	0,4072
15				
$\tau = 0,1$	0,0630	0,1420	0,2018	0,2509
$\tau = 0,5$	0,0645	0,1456	0,2078	0,2531
$\tau = 0,9$	0,0659	0,1485	0,2097	0,2575

Analizując wyniki zawarte w tab. 3 i 4 można zauważyć, że w większości przypadków przy ustalonym kapitale początkowym u i jednocześnie przy ustalonym horyzoncie czasowym H wraz ze wzrostem stopnia zależności pomiędzy wypłatami X'_{1i} i X'_{2i} następuje nieznaczny wzrost prawdopodobieństwa ruiny. Różnice te jednak są nieznaczne.

Podsumowanie

W pracy pokazane jest, że konstrukcja wieloklasowych modeli ryzyka jest ściśle uzależniona od pojmowania ruiny w kilku klasach ryzyka. W związku z tym przede wszystkim wyróżnia się wielowymiarowe modele ryzyka i agregacje kilku klas ryzyka. W obu tych rodzajach modeli przy ograniczeniu do dwóch klas ryzyka przeprowadzana jest numeryczna analiza wpływu stopnia zależności pomiędzy jednocześnie pojawiającymi się wpłatami w obu klasach ryzyka na prawdopodobieństwo ruiny.

W większości przypadków wyborów ustalonych kapitałów początkowych i horyzontów czasowych wraz ze wzrostem stopnia zależności pomiędzy jednocześnie pojawiającymi się wypłatami w obu klasach ryzyka po przyjęciu dwuwymiarowego modelu ryzyka można obserwować nieznaczny spadek prawdopodobieństwa ruiny, natomiast przy założeniu agregacji dwóch klas ryzyka można obserwować nieznaczny wzrost prawdopodobieństwa ruiny, jednak rząd tych różnic osobno w obu modelach jest praktycznie nieistotny. Wyniki tej krótkiej analizy mogłyby zatem wskazywać, że w praktyce można pomijać ten rodzaj zależności, co będzie przedmiotem dalszych badań.

Nie zawsze można pomijać fakt oddziaływania zewnętrznych czynników ryzyka, które mogą powodować pojawianie się różnorodnych szkód skutkujących jednoczesnym pojawianiem się wypłat w różnych klasach ryzyka. Przy klasycznym założeniu o procesach zaliczających wypłaty, tj. przyjmując, że są one jednorodnymi procesami Poissona, wyniki numerycznych analiz pokazują, że wzrost siły oddziaływania zewnętrznych czynników ryzyka, przy jednoczesnym proporcjonalnym zmniejszaniu siły oddziaływania wewnętrznych czynników ryzyka, powoduje wyraźny wzrost prawdopodobieństwa ruiny w przypadku agregacji kilku klas ryzyka [4; 5] i dość wyraźny spadek prawdopodobieństwa ruiny w przypadku dwuwymiarowego modelu ryzyka [6].

Literatura

1. Ambagaspitiya R.S., On the distribution of a sum of correlated aggregate claims, “Insurance: Mathematics and Economics” 1998, No. 23.
2. Asmussen S., Ruin Probabilities, “Advanced Series on Statistical Science & Applied Probability” 2000.
3. Heilpern S., Funkcje łączące, AE, Wrocław 2007.

4. Iwanicka A., Wpływ zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w nieskończonym horyzoncie czasowym w wieloklasowym modelu ryzyka, Zeszyt Naukowy „Ekonometria" XXIII, AE, Wrocław 2009.
5. Iwanicka A., Wpływ zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w skończonym horyzoncie czasowym w wieloklasowym modelu ryzyka, Zeszyt Naukowy „Ekonometria" XXVI, AE, Wrocław 2009.
6. Iwanicka A., Oddziaływanie zewnętrznych czynników ryzyka na prawdopodobieństwo ruiny w dwuwymiarowym modelu ryzyka, Zeszyt Naukowy „Statystyka aktuarialna – teoria i praktyka", UE, Wrocław (w druku).
7. Li S., Garrido J., Ruin probabilities for two classes of risk processes, "Astin Bulletin" 2005, No. 35.
8. Yuen K.C., Guo J., Wu X., On the first time of ruin in the bivariate compound Poissona model, "Insurance: Mathematics and Economics" 2006, No. 38.

THE IMPACT OF DEPENDENCE OF CLAIMS SIZES ON RUIN PROBABILITY IN TWO CLASSES RISK MODEL

Summary

These paper considers a risk model for two dependant classes of insurance business. The dependence between these classes is caused by appearing of some claims at the same time in both classes and additionally the sizes of these claims are dependant. The structure of the dependence between these claims sizes is described by copulas. The main aim of the paper is to investigate the impact of the level of dependence between these claims sizes on the finite-time ruin probability in considered risk model. short numerical analysis.

Tomasz Karczyński
Edward Radosiński
Politechnika Wrocławska

IMPLEMENTACJA I OCENA SYSTEMU EKSPERCKIEGO SIECI NEURONOWYCH W ANALIZIE RYNKU AKCJI

Wprowadzenie

Specyfika rynków kapitałowych, a dokładniej ich zależność od ogromnej liczby czynników, zarówno ekonomicznych, jak i psychologicznych, nieustanie zmieniających się w czasie, stwarza bardzo poważne problemy w jednoznacznej analizie [3; 5; 9; 13]. Mówiąc o metodach analiz rynków giełdowych warto wspomnieć o dwóch dominujących nurtach. Pierwszy, zwany fundamentalnym, jest oparty na modelach matematycznych. Modele te uwzględniają tylko wybrane wycinki rzeczywistości, w jakiej działają rynki kapitałowe. Kolejny nurt, zwany technicznym, jest oparty na założeniu, że rynki kapitałowe dyskontują wszelkie informacje w cenach aktywów [5]. Wadą tego nurtu jest mnogość metodyk analiz oraz niejednoznaczność w ich interpretacji. Często skłaniają one inwestorów do podejmowania pochopnych decyzji [7]. Rozwój technik komputerowych wpłynął w znaczącym stopniu na szybkość i jakość pozyskiwanych analiz ekonomicznych [4]. Powszechne stało się wykorzystywanie komputerowych systemów giełdowych, zintegrowanych z systemami eksperckimi (tzw. HFT), które pozwalają na bardzo szybkie reagowanie na zmiany na rynkach akcji.

Do jednego z nurtów współczesnej informatyki, którego potencjał wydaje się nie w pełni wykorzystywany w rozwiązywaniu problemów ekonomicznych należy obszerna tematyka sztucznej inteligencji [4]. Moce obliczeniowe komputerów, zdolności gromadzenia, przetwarzania i analizowania danych za pomocą technik sztucznej inteligencji pozwalają przypuszczać, że wspomniany problem wiel-

kowymiarowości i zmienności w czasie czynników decydujących o zachowaniu rynków kapitałowych w końcu może być rozwiązany. Jakość analiz wygenerowanych przez systemy eksperckie powinna w sposób znaczący przewyższać jakość analiz wykonywanych przez człowieka [15].

W niniejszej pracy zaprezentowany zostanie system zbudowany na potrzeby analizy rynków akcji. Jego architektura została oparta na rozwiązaniach sztucznej inteligencji, a dokładnie na dwóch modelach sieci neuronowych: sieci wielowarstwowej, uczonej algorytmem wstecznej propagacji błędów oraz sieci typu SOM. Opracowany system bazuje na założeniu analizy technicznej, że ceny aktywów giełdowych zawierają wszelkie informacje rynkowe. Rynek dyskontuje informacje i definiuje cenę aktywa. Założenie to pozwoliło na modelowanie wiedzy na podstawie danych opisujących przebieg sesji giełdowej.

Przeprowadzone badania mają na celu uzyskanie odpowiedzi na pytanie: czy wyniki wygenerowane przez opracowany system ekspercki istotnie różnią się od wyników klasycznych metod analizy rynków akcji.

1. System analiz danych giełdowych – budowa i architektura

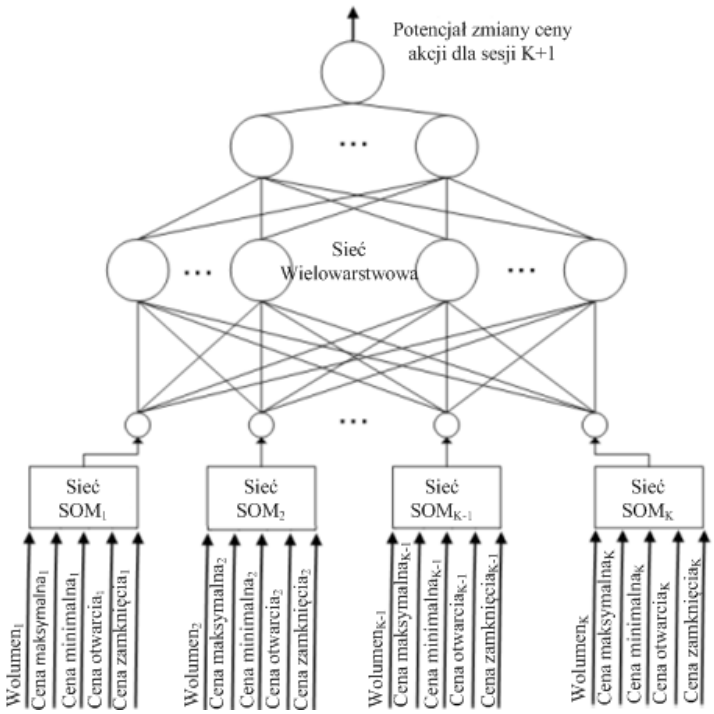
Architektura sieci neuronowych użytych do budowy systemu eksperckiego analiz giełdowych bazuje na pięciu fundamentalnych zbiorach danych, opisujących zachowania aktywa podczas sesji. Są to odpowiednio informacje o:

– wolumenie obrotów aktywa na koniec sesji,
– cenie minimalnej dla sesji, jaką należało zapłacić za walor,
– cenie maksymalnej dla sesji, jaką należało zapłacić za walor,
– cenie za walor na otwarciu sesji,
– cenie za walor na zamknięciu sesji.

Dla celów badawczych, na wejście sieci neuronowej zaprezentowane zostaną dane opisujące sesje kolejno następujące po sobie. Aby zmniejszyć rozmiar danych dokonano redukcji wektora danych giełdowych, prezentowanego na wejściu systemu. Było to możliwe dzięki wykorzystaniu modelu sieci neuronowej opartego na architekturze SOM. Sieci tego typu nadają się idealnie do rozwiązywania problemu klasyfikacji i kompresji danych [10; 14]. Te cechy sprawiły, że oprócz wspomnianej redukcji wymiaru danych wyeliminowane zostały dane sesji bardzo podobnych do siebie. Tak zredukowany ciąg danych został zaprezentowany na wejściu sieci wielowarstwowej, uczonej algorytmem wstecznej propagacji błędów [10; 14]. Na wyjściu sieci oczekiwana jest wartość opisująca potencjał zmia-

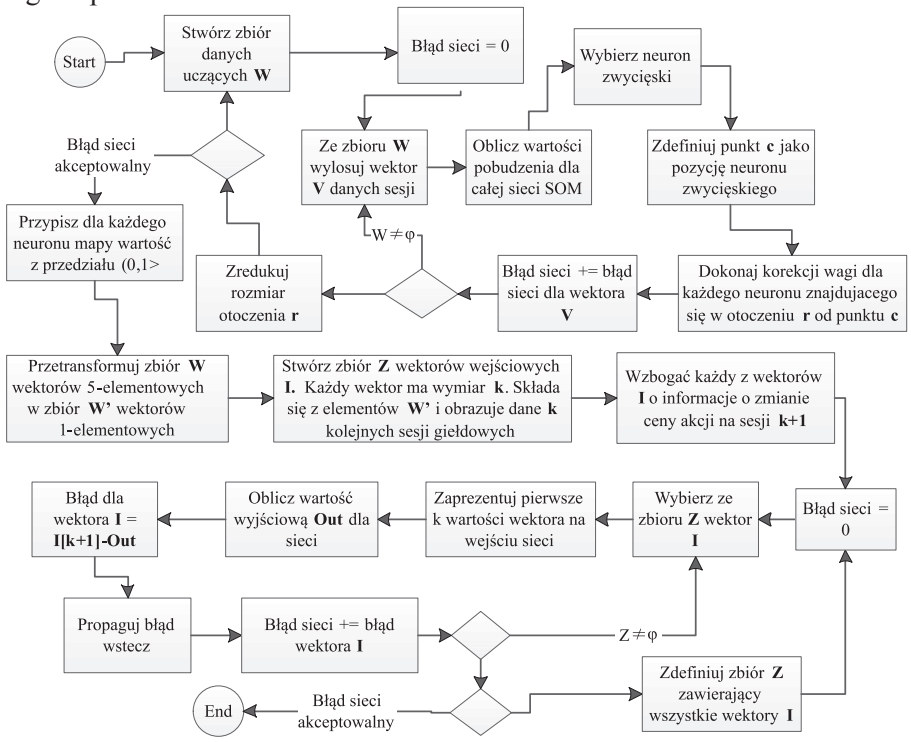
ny ceny waloru na kolejnej sesji giełdowej. Wartość ta stanowi informację analityczną, wygenerowaną przez system ekspercki.

Na rys. 1 zaprezentowano ideę budowy systemu. Bazując na danych spółki giełdowej, na wejściu systemu przedstawione są wektory danych obrazujące przebieg *k* następujących po sobie sesji giełdowych. W pierwszym etapie, wykorzystując sieci typu SOM, dla każdej z *k* sesji dokonana zastanie transformacja wektorów o wymiarze 5 na wektory jednoelementowe [8]. Przetworzone dane zostaną przekazane na wejście sieci wielowarstwowej. Ta część systemu generuje ostateczną informację systemu eksperckiego o potencjale zmiany ceny waloru dla kolejnej sesji, tj. *k + 1*. Najistotniejszym etapem budowy opisywanego systemu jest proces uczenia sieci neuronowych, mający szczególne znaczenie dla działania systemu. System, bazując na olbrzymiej ilości danych z sesji giełdowych dla spółki, jakie odbyły się pomiędzy 2005 r. a 2008 r. stara się odnaleźć wzorce pozwalające na dokonanie trafnej analizy kursu akcji w przyszłości. Ciąg danych zwany jest ciągiem uczącym. W pierwszym etapie procesu uczenia następuje optymalizacja wartości połączeń pomiędzy neuronami w sieci SOM. Jest to tzw. nauczanie bez nauczyciela.



Rys. 1. Model budowy systemu analiz danych giełdowych

W toku tego procesu sieć stara się tak dostosować wartości wag połączeń neuronów, aby w sposób jednoznaczny dokonywać klasyfikacji danych dla sesji giełdowych [10; 14; 16]. Dzięki temu zabiegowi sesje, które cechują się dużym stopniem zbieżności względem siebie zostaną zaklasyfikowane do tej samej grupy. Po ich transformacji w wektor jednoelementowy zachowana zostanie informacja o podobieństwie – z podobną siłą będą wzbudzały sygnał na wejściu kolejnej sieci. Po zakończeniu procesu uczenia sieci SOM i dokonaniu transformacji danych następuje proces uczenia sieci wielowarstwowej. Jest to tzw. nauczanie z nauczycielem. Każdy wektor danych zostaje wzbogacony o informację o prawdziwej, maksymalnej cenie kursu akcji na kolejnej sesji giełdowej. Wartość ta nie będzie nigdy podawana na wejściu sieci. W toku nauczania oczekiwane jest otrzymywanie wyników systemu zgodnych z nowo dodaną informacją [8]. Warto nadmienić, że proces uczenia sieci SOM odbywał się według algorytmu „meksykańskiego kapelusza".



Rys. 2. Przebieg procesu nauki w systemie analiz danych giełdowych

Metoda rozróżnia siłę, z jaką mają zostać skorygowane wartości wag neuronów sąsiadujących z neuronem zwycięskim. Im „bliższy sąsiad zwycięzcy", tym większa siła zmian jego wag [10]. Ideę wyraża się wzorem:

$$h(\rho, r) = \exp\left[-\frac{\|\rho\|^2}{2\sigma^2(t)}\right]$$

gdzie:
σ^2 (t) – współczynnik zmiany siły korekcji wag w zależności od kolejności cyklu uczenia,
ρ – dystans do neuronu zwycięskiego.

Każda kolejna epoka uczenia charakteryzuje się spadkiem siły korekcji wag. Po zakończeniu procesu uczenia sieci SOM następuje proces transformacji danych wejściowych z wektora pięcioelementowego na wektor jednoelementowy. Każdemu neuronowi w warstwie mapy sieci SOM przypisana jest unikalna wartość z przedziału (0,1). Zachowana zostaje zasada, że różnica wartości nadana neuronom zawierającym skrajnie różne dane wynosi 1. Po procesie transformacji tworzony jest zbiór wektorów o długości *k*, zawierających wartości wyliczone w poprzednim kroku. Wektor ten modeluje skompresowane dane z *k* następujących po sobie sesji giełdowych. Każdy tak utworzony wektor wzbogacony zostaje o informacje o maksymalnej cenie akcji w kolejnej sesji, tj. *k+1*. Podczas procesu nauczania sieci wielowarstwowej, wszystkie wartości do indeksu *k* traktowane są jako dane wejściowe. Wartość na pozycji *k+1* stanowi oczekiwaną wartość, jaką powinien zwrócić system. Proces nauki w wykorzystanej sieci wielowarstwowej, bazując na algorytmie wstecznej propagacji błędu, minimalizuje funkcję błędu sieci [10; 14]. Samą funkcję można zdefiniować za pomocą wzoru:

$$Q(n) - \sum_{i=1}^{N} \varepsilon_i^2(n) = \sum_{i=1}^{N} [d_i(n) - y_i(n)]^2$$

gdzie:
Q – funkcja błędu sieci dla *n*-tego wektora danych wejściowych,
N – liczba warstw sieci wielowarstwowej,
ε_i – błąd sieci w warstwie *i*,
d – oczekiwana wartość wyjścia sieci,
y – wartość zwrócona przez sieć.

Proces korekcji wag jest zależny od współczynnika nauczania oraz składowej gradientu funkcji błędu [10; 14] i wyrażony jest wzorem:

$$w_{ij}^k(n+1) = w_{ij}^k(n) + 2\omega\delta_i^k(n)x_j^k(n) + \alpha[w_{ij}^k(n) - w_{ij}^k(n-1)]$$

gdzie:

w_{ij}^k – wartość wagi *i*-tego neuron w warstwie *k*-tej dla *j*-tego wektora danych,

δ_i^k – gradient funkcji błędu dla *k*-tej warstwy,
ω – współczynnik nauki,

x_j^k – j-ty sygnał wejściowy przekazany do warstwy k-tej,
α – współczynnik momentum propagujący kierunek zmiany gradientu funkcji błędu.

2. Idea procesu oceny system eksperckiego

Na potrzeby procesu oceny systemu analiz danych giełdowych wykorzystano dane z sesji giełdowych pomiędzy 01.01.2009 r. a 31.12.2009 r. Dane te obrazowały zachowania akcji polskich spółek należących do różnych sektorów gospodarki:

- PKO Bank Polski S.A. – jeden z największych banków komercyjnych w Europie Środkowowschodniej,
- KGHM Polska Miedź S.A. – dziewiąty pod względem wielkości producent miedzi na świecie i trzeci na świecie producent srebra,
- PKN Orlen S.A. – jedna z największych rafinerii ropy naftowej oraz sieci dystrybucji paliw płynnych w Europie Centralnej,
- Grupa LOTOS S.A. – koncern rafineryjno-wydobywczy prowadzący sprzedaż detaliczną i hurtową produktów ropopochodnych.

Przedsiębiorstwa notowane są na Warszawskiej Giełdzie Papierów Wartościowych należą do indeksu WIG20, który obrazuje stan notowań dwudziestu najbardziej wartościowych spółek na GPW.

Na potrzeby oceny analiz wygenerowanych przez opisywany system ekspercki wyniki zostaną skonfrontowane z wynikami uzyskanymi za pomocą klasycznych technik analizy danych giełdowych. Są to:

- Metoda portfelowa Markowitza – metoda inwestowania długoterminowego. Na potrzeby oceny ustalono minimalny poziom zysku, na poziome rentowności dziesięcioletnich obligacji Skarbu Państwa. Parametry portfela zostały ustalone na podstawie danych giełdowych spółek z sesji w latach 2005-2008 [3; 5; 9].
- Bootstrap – metoda oparta na idei „Monte Carlo", która pozwala szacować kierunek zmian wszędzie tam, gdzie nieznany jest rozkład prawdopodobieństwa wystąpienia pewnych wartości. W przypadku analiz giełdowych rozkład prawdopodobieństwa wartości zmian kursu akcji jest szacowany na podstawie danych historycznych [2; 12]. Metoda ta do celów badawczych została użyta jako krótkoterminowa, tzn. dopuszczalne jest sprzedawanie bądź kupowanie składników portfela akcji na koniec każdej sesji [2; 6].

– MACD – metoda średnich kroczących. Bazuje na relacji między dwoma średnimi (tzw. EMA): z 26 i 12 sesji. Średnia z 12 sesji „wygładza" średnią z 26 sesji. Za linię sygnału wykorzystano średnią 9 kolejnych sesji. Każde przecięcie przez nią średnich z 26 i 12 sesji może wygenerować sygnał zakupu bądź sprzedaży akcji [1; 5]. Metoda ta ma charakter krótkoterminowy.

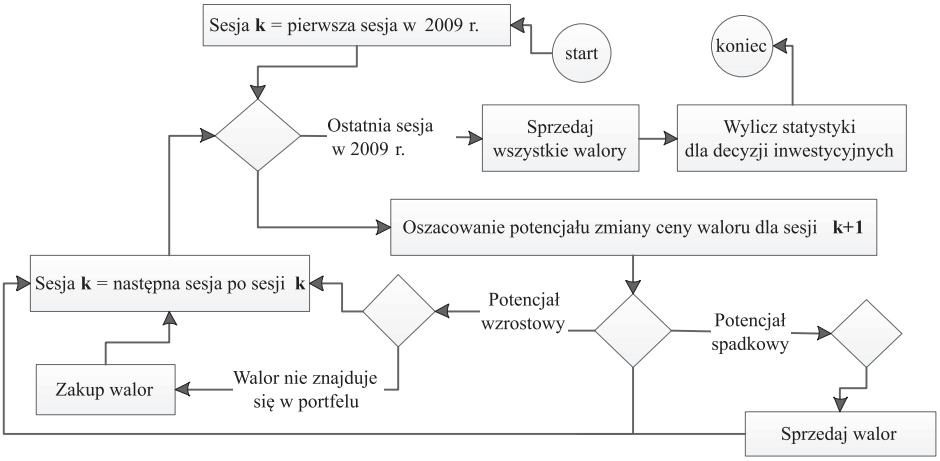
Idea oceny skuteczności analiz poszczególnych metod opiera się na modelu portfelowym. Jest to podyktowane faktem, że na potrzeby badawcze użyty został również model Markowitza. W związku z tym, oceniając skuteczność analiz wybranych metod należy brać pod uwagę nie precyzję szacowań wartości zmian poszczególnych walorów, ale precyzję szacunków dla całego portfela [17]. Przyjęta została zasada, że dla metod innych niż metoda portfela Markowitza procentowy udział walorów w portfelu nie będzie w czasie procesu badawczego optymalizowany.

Tabela 1

Udział walorów w portfelu

Metoda	Spółka	Udział w portfelu
Markowitz	PKO Bank Polski	13,3%
	KGHM Polska Miedź SA	1,9%
	PKN Orlen	0,0%
	Grupa LOTOS S.A	84,8%
Bootstrap MACD System sieci neuronowych	PKO Bank Polski	25,0%
	KGHM Polska Miedź SA	25,0%
	PKN Orlen	25,0%
	Grupa LOTOS S.A	25,0%

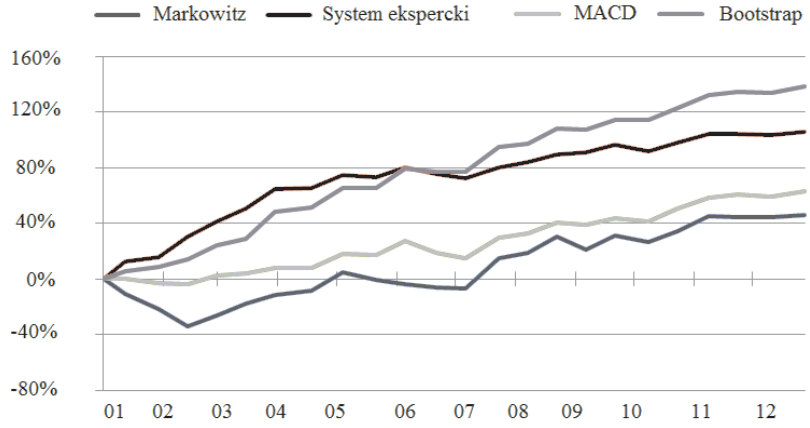
Na potrzeby oceny systemu analiz danych giełdowych wykorzystano informacje z sesji giełdowych pomiędzy 01.01.2009 r. a 31.12.2009 r. W oparciu o sygnały wygenerowane przez poszczególne metody, walory wchodzące w skład portfela mogą zostać sprzedane bądź zakupione. Rysunek 3 prezentuje ideę algorytmu inwestycyjnego.



Rys. 3. Idea algorytmu realizacji procesu inwestycyjnego

3. Wyniki oceny analiz systemu eksperckiego

Na rys. 4 zaprezentowano zyskowność inwestycji porównywanych metod analizy rynków akcji. Każda z inwestycji rozpoczyna się na pierwszej sesji 2009 r. Na kolejnych sesjach giełdowych, w oparciu o wyniki analiz danych giełdowych, skład portfeli badanych metod analiz ulegał modyfikacji. Wyjątkiem jest metoda Markowitza traktowana jako inwestycja długoterminowa [17]. Na ostatniej sesji 2009 r. wszystkie walory są sprzedawane. Po tej operacji następuje wyliczenie podstawowych miar cechujących przebieg inwestycji.



Rys. 4. Przebieg inwestycji dla porównywanych modeli analiz

W celu poprawy jakości oceny decyzji inwestycyjnych wzbogacono zbiór badanych modeli o trzy quasi-modele stanowiące punkty odniesienia w procesie porównawczym:

– Maksimum – model idealistyczny, składający się tylko z tych decyzji inwestycyjnych, które maksymalizują zyskowność portfela.
– Minimum – model składający się tylko z tych decyzji inwestycyjnych, które minimalizują zyskowność portfela.
– Rzeczywisty – koncepcja "kup i czekaj"; brak decyzji inwestycyjnych zmieniających skład portfela.

Dla celów oceny decyzji inwestycyjnych użyto siedem wskaźników:

– zyskowność – finalna zyskowność portfela na koniec 2009 r.,
– odchylenie standardowe – tożsame z ryzykiem inwestycyjnym,
– dzienny zysk – średni dzienny zysk portfela inwestycyjnego,
– wzrosty – liczba decyzji zwiększających zyskowność portfela,
– spadki – liczba błędnych decyzji inwestycyjnych,
– trafność decyzji – miernik przyjmujący wartość z zakresu <0,1>; wartość 1 oznacza, że portfel generował wyłącznie zyski; wartość 0 sugeruje, że portfel inwestycyjny generował wyłącznie straty,
– poziom ufności – wraz ze średnią wartością wyznacza przedziały mogące dostarczyć informacji o potencjalnej zbieżności pomiędzy wynikami portfeli.

Wartości finalne zyskowności i odchylenia standardowego stanowią kryteria oceny przebiegu procesu inwestycyjnego dla badanych metod.

Tabela 2

Wyniki porównywanych modeli inwestycyjnych

	Zyskowność	Odchylenie standardowe	Dzienny zysk	Wzrosty	Spadki	Trafność decyzji	Poziom ufności
Minimum	-425,21%	2,76%	-1,69%	64	184	0,26	0,0034
Maksimum	597,96%	3,30%	2,37%	199	48	0,81	0,0041
Markowitz	36,41%	2,93%	0,14%	136	116	0,54	0,0036
System ekspercki	107,28%	1,72%	0,43%	138	95	0,59	0,0021
MACD	61,60%	1,98%	0,24%	127	110	0,54	0,0024
Bootstrap	140,44%	2,11%	0,56%	142	95	0,60	0,0026
Rzeczywisty	79,76%	2,69%	0,32%	134	118	0,53	0,0033

Ostatni etap oceny modeli inwestycyjnych oparto na wnioskowaniu statystycznym, tj. na teście T-Studenta. Test ten mierzy, czy wyniki decyzji inwestycyjnych dowolnych dwóch badanych metod są istotnie różne od siebie w znaczeniu statystycznym, tj. czy zyski generowane przez portfele różnią się w istotny sposób między sobą pod względem wartości średniej. Za hipotezę zerową przyjęto rów-

ność średnich zysków dwóch portfeli, za poziom istotności przyjęto wartość 0,05. Za metodę najlepszą zostanie uznana ta, która wykaże największą statystyczną zbieżność z wynikami quasi-metody „maksimum". Im większa zbieżność, tym lepsza ocena wyników analiz wygenerowanych przez badaną metodę.

Tabela 3

Test t-studenta dla decyzji inwestycyjnych porównywanych metod

	Minimum	Maksimum	Markowitz	System ekspercki	MACD	Bootstrap	Rzeczywisty
Minimum	1	0	0	0	0	0	0
Maksimum	0	1	0	0	0	0	0
Markowitz	0	0	1	0,1900	0,6543	0,0705	0,4928
System	0	0	0,1900	1	0,2731	0,4430	0,5870
MACD	0	0	0,6543	0,2731	1	0,0868	0,7319
Bootstrap	0	0	0,0705	0,4430	0,0868	1	0,2636
Rzeczywisty	0	0	0,4928	0,5870	0,7319	0,2636	1

Duża zbieżność wyników inwestycji badanej metody z quasi-metodą „rzeczywisty" może być interpretowana jako sygnał silnej zależności wartości analiz od sytuacji panującej na rynku. Takie zachowanie pozwala przypuszczać, że podczas bessy na rynku akcji decyzje wygenerowane przez badaną metodę analizy rynku będą błędne.

Podsumowanie

Na podstawie przeprowadzonych badań wykazano, że analizy wygenerowane przez system ekspercki sieci neuronowych mogą być uważane za efektywną metodę podejmowania decyzji inwestycyjnych. Ocena statystyczna potwierdza zdolność do generowania decyzji charakteryzujących się niskim poziomem ryzyka inwestycyjnego, przy jednoczesnym wysokim poziomie zwrotu z inwestycji. Wyniki decyzji systemu eksperckiego nie mogą być uważane za zbieżne z wynikami decyzji optymalnych. Istotny jest również fakt, że żadna spośród badanych metod nie wykazała statystycznej zbieżności z modelem idealnym.

W kolejnych pracach badawczych planowane jest zwiększenie liczby metod uwzględnionych w procesie porównawczym, odejście od założenia braku modyfikacji procentowego udziału danego waloru w portfelu inwestycyjnym. Sam proces badawczy zostanie przeprowadzony w trzech stadiach rynku: „rynku niedźwiedzia", „rynku byka" i rynku w stanie konsolidacji. System ekspercki zostanie rozbudowany o bazę wiedzy opartą na modelu logiki rozmytej [14]. Baza ta umoż-

liwi interpretowanie, w oparciu o techniki rozmyte, danych makro- i mikroekonomicznych. Tak zbudowany system będzie miał za zadanie analizowanie możliwie najszerzej zbioru czynników decydujących o zachowaniu rynków giełdowych. Analizowanie szerokiego spektrum danych powinno w sposób istotny poprawić jakość generowanych rezultatów.

Literatura

1. Appel G., Dobson E., Understanding MACD, Traders Press 2008.
2. Boyle P., Seng Tan K., Quasi-Monte Carlo Methods in Numerical Finance, "Management Science" 1996.
3. Brock W., Lakonishok J., Lebaron B., Simple Technical Trading Rules and the Stochastic Properties of Stock Returns, "The Journal of Finance" 1992.
4. Chrzan P., Timofiejczuk G., Porównanie zastosowania sieci neuronowych i modeli klasy GARCH w prognozowaniu stóp zwrotu. Część 1, s. 147-156, w: Rynek Kapitałowy. Skuteczne inwestowanie, red. W. Tarczyński, Uniwersytet Szczeciński, Szczecin 2002.
5. Czekała M., Analiza fundamentalna i techniczna, Akademia Ekonomiczna, Wrocław 2007.
6. Davison, A., Hinkley V., Bootstrap Methods and their application, Cambridge University Press 1997.
7. Edwards R., Magee J.,. Bassetti W.H.C,. Technical Analysis of Stock Trends, Ninth ed. Boca Raton: CRC Press 2007.
8. Karczyński T., Techniki sztucznej inteligencji w zarządzaniu – sieci neuronowe w analizie cen akcji, Politechnika Wrocławska, Wrocław 2010.
9. Mamaysky H., Wang J., Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, "The Journal of Finance" 2000.
10. Markowska-Kaczmar U., Sieci neuronowe w zastosowaniach, Oficyna EXIT, Warszawa 2003.
11. Maryański W., Podstawy analizy technicznej, wypracowanie: 2005.04.05.
12. Mielczarek B., Metody próbkowania w symulacji Monte Carlo, Wydział Informatyki i Zarządzania PWR, Wrocław 2007.
13. Park C., Irwin S., The Profitability of Technical Analysis: A Review, AgMAS Project Research Report 2004-04, University of Illinois at Urbana-Champaign.

14. Pliński M., Rudkowska D., Rutkowski L., Sieci neuronowe, algorytmy genetyczne i systemy rozmyte, PWN, Łódź-Warszawa 1997.
15. Radosiński E., Systemy informatyczne w dynamicznej analizie decyzyjnej, PWN, Warszawa 2005.
16. Rutkowski L., Metody i techniki sztucznej inteligencji. Inteligencja obliczeniowa, PWN, Warszawa 2005.
17. Tarczyński W., Analiza portfelowa na giełdzie papierów wartościowych, PTE, Szczecin 1996.
18. Cesari R., Cremonini D., Benchmarking, Portfolio Insurance and Technical Analysis: A Monte Carlo Comparison of Dynamic Strategies of Asset Allocation. "Journal of Economic Dynamics and Control" 2003.

IMPLEMENTATION AND EVALUATION OF THE NEURAL NETWORK SYSTEM FOR STOCK MARKET DATA ANALYSIS

Summary

The application of neural network system for multi-dimensional stock market data analysis is presented in the paper. Developed system predicts stock price movements based on daily quotation data like: volume, minimum and maximum session price, opening and closing price. Several studies were carried out, to compare systems investment decisions, with decisions that were made on the basis of some commonly used methods of stock market analysis. These methods are: MACD, Bootstrap, Markowitz Portfolio. For valuation purpose, the real stock market data of the four largest Polish companies were used. All companies are quoted on the Warsaw Stock Exchange and belong to the WIG 20 index. For the benchmarking, only stock data from the year 2009 were used. In order to enrich the benchmarking tests, three investment scenarios were added. First known as the skeptical assume that only incorrect investment decisions were made. Second known as the optimistic assume that only correct investment decisions were made. Last one known as passive assume that no investment decision were made – it is so called "buy and hold" conception.

The benchmarking results confirmed, that the neural network system is able to make investment decisions, that significantly increase the profitability of the investment portfolio. Neural network system provide investment suggestions, that can be considered as an alternative to other commonly used methods of stock market analysis. However statistical tests proved a high correlation between quality of systems investment decisions and

market trend and lack of correlation to the “optimistic” scenario. Neural network systems may help in investment process, but cannot be considered as fully reliable way of investment process automation.

Andrzej Karpio
Dorota Żebrowska-Suchodolska
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

PRZYPADKOWOŚĆ WYNIKÓW INWESTYCYJNYCH FIO FUNKCJONUJĄCYCH NA POLSKIM RYNKU KAPITAŁOWYM*

Wprowadzenie

Od ponad dziesięciu lat bardzo popularną formą oszczędzania osób indywidualnych jest zakup jednostek uczestnictwa otwartych funduszy inwestycyjnych**. Inwestorzy indywidualni nie mając należytej wiedzy dotyczącej możliwości, jakie daje rynek kapitałowy, czy też chociażby z braku czasu powierzają swoje oszczędności instytucjom powołanym do profesjonalnego zarządzania kapitałem. W tym miejscu trzeba sobie odpowiedzieć na podstawowe pytanie: Który fundusz wybrać? Ale nawet po dokonaniu wyboru jeszcze ważniejsze jest to, czy ten fundusz ma szanse pozostać liderem przez dłuższy okres. Przy wykorzystaniu testów statystycznych w pracy próbowano uzyskać odpowiedzi na następujące pytania: Czy procentowe zmiany jednostek uczestnictwa poszczególnych funduszy miały charakter losowy? Czy odchylenia standardowe (ryzyka) badanych funduszy różniły się od siebie? Czy osiągane średnie stopy zwrotu różnych funduszy były statystycznie różne? Czy pojawienie się funduszy w kolejnych okresach na wysokich pozycjach rankingowych, tworzonych jedynie na podstawie stóp zwrotu, mają charakter trwały?

* Praca naukowa finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy N N111 277638.

** Wartość aktywów netto na koniec 2000 r. wynosiła 7 mln PLN, a na koniec 2009 r. – 93,1 mln PLN (na podstawie raportu IZFiA).

1. Metodologia badań

Badania dotyczyły funduszy akcyjnych i zrównoważonych funkcjonujących na polskim rynku w okresie 1.02.2007-31.08.2011 r. Wybranie takiego okresu badań wiązało się z tym, że wyniki uzyskane w niniejszej pracy autorzy chcieli porównać z okresem od 31.01.2003 r.-31.01.2007 r. [3]. Pierwszy okres przypadał na hossę, a obecnie rozważany obejmuje kryzys, zatem złą koniunkturę panującą na rynku kapitałowym. W konsekwencji, badania objęły 16 funduszy akcyjnych i 12 funduszy zrównoważonych (tab. 1). Taka liczba funduszy pojawiła się we wcześniejszych badaniach, w których kryterium wyboru funduszu było jego funkcjonowanie na rynku przez cały okres. Badania zostały przeprowadzone na podstawie miesięcznych i tygodniowych zmian jednostek uczestnictwa. Szczegółowe wyniki zostały przedstawione dla danych miesięcznych, a potraktowane informacyjnie dla danych tygodniowych. Dodatkowo autorzy dochodzą do wniosku, że okres tygodniowy jest zbyt krótki, aby prowadził do wniosków istotnie wpływających na decyzje inwestycyjne zarządzających portfelami funduszy.

W tab. 1 znalazły się fundusze objęte badaniami, pogrupowane według towarzystw inwestycyjnych. Z tego powodu braki w drugiej kolumnie oznaczają, że w ramach danego towarzystwa znalazły się dwa lub trzy fundusze akcyjne, a jeden zrównoważony (inne nie funkcjonowały przez cały okres brany pod uwagę w badaniach)[*].

Tabela 1

Lista funduszy akcyjnych i zrównoważonych objętych badaniem (1.02.2007-31.08.2011)

Fundusze akcyjne	Fundusze zrównoważone
Legg Mason Akcji	Legg Mason Zrównoważony
Unikorona Akcje	Unikorona Zrównoważony
BPH A Dynam	BPH Aktywnego Zarządzania
BPH Akcji	
DWS Top 25	DWS Zrównoważony
DWS Akcji Plus	
DWS Akcji	
Arka BZWBK Akcji	Arka BZWBK Zrównoważony
Millenium Akcji	Millenninum Zrównoważony
PKO/CS Akcji	PKO/CS Zrównoważony
Pioneer Akcji Polskich	Pioneer Zrównoważony
Skarbiec Akcja	Skarbiec Waga

[*] Metodologia badań szczegółowo została opisana w pracy [4], dotyczącej okresu hossy panującej na rynku kapitałowym. Metody statystyczne i ich interpretacja zostały zaczerpnięte z prac [1; 2; 3; 5; 6].

cd. tabeli 1

ING Akcji	ING Zrównoważony
SEB 3	SEB1
Cu Akcji Polskich	
PZU Akcji Krakowiak	
	KBC Aktywny

Do odpowiedzi na poszczególne pytania użyto testów:
- normalności rozkładu: Jarque-Bery, Shapiro-Wilka, Kołmogorova-Smirnowa z poprawką Lillieforse'a,
- jednorodności wariancji: Hartleya, Cochrana, Bartletta,
- Tukeya, Bonferroniego,
- kwantylowego.

Wszędzie, gdzie było to konieczne przyjęto poziom ufności równy 0,05.

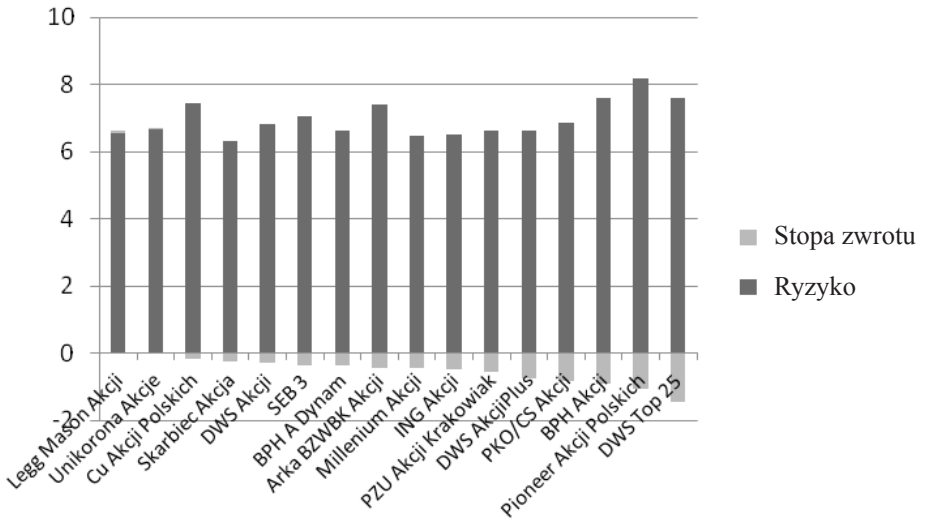
2. Wyniki badań

W badanym okresie (1.02.2007-31.08.2011) średnie miesięczne stopy zwrotu funduszy akcyjnych przyjmowały wartości widoczne w tab. 2 (uszeregowane od najniższej do najwyższej stopy zwrotu) oraz dodatkowo na rys. 1.

Tabela 2

Średnie miesięczne wartości stopy zwrotu i ryzyka funduszy akcyjnych

Fundusz akcyjny	Stopa zwrotu (%)	Ryzyko (%)
Legg Mason Akcji	0,074	6,537
Unikorona Akcje	0,003	6,650
Cu Akcji Polskich	-0,160	7,423
Skarbiec Akcja	-0,239	6,298
DWS Akcji	-0,277	6,832
SEB 3	-0,354	7,056
BPH A Dynam	-0,366	6,634
Arka BZWBK Akcji	-0,427	7,403
Millenium Akcji	-0,438	6,472
ING Akcji	-0,491	6,514
PZU Akcji Krakowiak	-0,546	6,623
DWS AkcjiPlus	-0,741	6,619
PKO/CS Akcji	-0,824	6,869
BPH Akcji	-0,918	7,616
Pioneer Akcji Polskich	-1,066	8,195
DWS Top 25	-1,448	7,589



Rys. 1. Średnie miesięczne wartości stopy zwrotu i ryzyka funduszy akcyjnych

Jak można było oczekiwać, praktycznie wszystkie fundusze miały ujemne stopy zwrotu, zmieniały się one w stosunkowo dużym zakresie, czemu nie towarzyszyło istotne zróżnicowanie ryzyka. Są to tylko spostrzeżenia „wizualne", a ich potwierdzenie było przedmiotem dalszych badań. W szczególności badanie normalności wyeliminowało 6 spośród 16 funduszy, dla których przynajmniej dwa spośród trzech testów prowadziły do odrzucenia hipotezy zerowej o normalności rozkładu miesięcznych stóp zwrotu (w tabeli zaznaczono je pismem pochyłym).

Dla pozostałych funduszy testy równości wariancji oraz równości średnich stóp zwrotu nie dały podstaw do odrzucenia hipotezy zerowej. W konsekwencji można stwierdzić, że w badanym okresie fundusze akcji o normalnym rozkładzie stóp zwrotu jednostek uczestnictwa charakteryzowały się jednakowymi (ze statystycznego punktu widzenia) miesięcznymi stopami zwrotu oraz jednakowym ryzykiem.

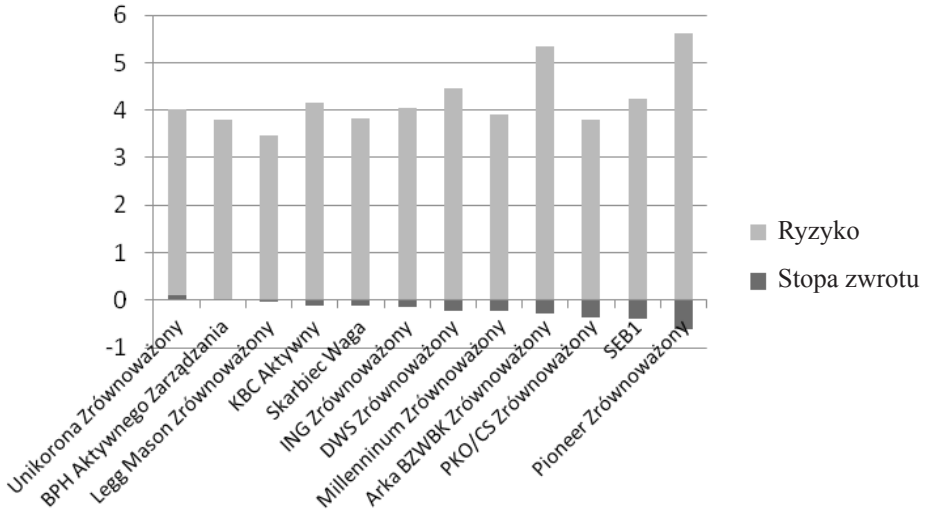
W przypadku danych tygodniowych wszystkie zmiany jednostek uczestnictwa nie miały rozkładu normalnego, co wydaje się zrozumiałe, gdyż jest mało prawdopodobne, aby zmiany portfeli inwestycyjnych funduszy następowały tak często. Można zatem przyjąć, że tydzień jest okresem nieodpowiednim do badań, jest zbyt krótki.

W badanym okresie (1.02.2007-31.08.2011), średnie miesięczne stopy zwrotu funduszy zrównoważonych przyjmowały wartości widoczne w tab. 3 (uszeregowanie od najniższej do najwyższej stopy zwrotu) oraz zilustrowane na rys. 2.

Tabela 3

Średnie miesięczne wartości stopy zwrotu i ryzyka funduszy zrównoważonych

Fundusz	Stopa zwrotu (%)	Ryzyko (%)
Unikorona Zrównoważony	0,113	3,916
BPH Aktywnego Zarządzania	0,009	3,802
Legg Mason Zrównoważony	-0,024	3,463
KBC Aktywny	-0,119	4,146
Skarbiec Waga	-0,129	3,822
ING Zrównoważony	-0,148	4,049
DWS Zrównoważony	-0,233	4,449
Millenninum Zrównoważony	-0,236	3,913
Arka BZWBK Zrównoważony	-0,273	5,338
PKO/CS Zrównoważony	-0,357	3,805
SEB1	-0,402	4,252
Pioneer Zrównoważony	-0,617	5,630



Rys. 2. Średnie miesięczne wartości stopy zwrotu i ryzyka funduszy zrównoważonych

W tym przypadku, podobnie jak dla funduszy akcji, tylko dwa fundusze miały dodatnie stopy zwrotu, a pozostałe ujemne. Wykres na rys. 2 przypomina poprzedni, jednak precyzyjne rachunki pokazują pewną różnicę. W teście normalności brak podstaw do odrzucenia hipotezy zerowej dotyczył czterech funduszy: BPH Aktywnego Zarządzania, KBC Aktywny, Millenium Zrównoważony, Unikorona Zrównoważony. Testy równości wariancji oraz równości średnich stóp zwrotu nie dały podstaw do odrzucenia hipotez zerowych.

Analiza tygodniowych stóp zwrotu pokazała, że zmiany jednostek uczestnictwa wszystkich funduszy nie miały rozkładu normalnego. Można zauważyć, że w przypadku funduszy zrównoważonych testy przeprowadzane na danych miesięcznych i tygodniowych nieznacznie się od siebie różniły.

Stabilność pozycji rankingowych funduszy badano testem kwantylowym (w pracy kwantylem była mediana). Jego podstawą było uporządkowanie funduszy malejąco według osiąganej stopy zwrotu w kolejnych miesiącach. Następnie zliczano ile razy na kolejnych miejscach pojawiają się poszczególne fundusze. Otrzymane dane ponownie szeregowano malejąco. W fazie końcowej testowano hipotezę o losowym pojawieniu się funduszy na pierwszym, drugim i trzecim miejscu ze względu na stopę zwrotu. Hipoteza alternatywna była bezpośrednim zaprzeczeniem hipotezy zerowej. Uzyskano następujące wyniki (pominięto wyniki testu normalności stóp zwrotu).

Tabela 4

Wyniki testu kwantylowego

	Fundusze akcyjne		Fundusze zrównoważone	
	stopy miesięczne	stopy tygodniowe	stopy miesięczne	stopy tygodniowe
I miejsce	nielosowy	losowy	nielosowy	losowy
II miejsce	losowy	losowy	losowy	nierozstrzygnięte
III miejsce	losowy	losowy	losowy	losowy

Widać, że pojawianie się funduszy na pierwszej pozycji rankingowej tworzonej na podstawie stóp miesięcznych nie ma charakteru losowego zarówno dla funduszy akcyjnych, jak i zrównoważonych. Mając ciągle na uwadze stopy miesięczne należy stwierdzić, że identyczne wyniki uzyskano dla okresu hossy.

Podsumowanie

Odpowiadając na pytania postawione w części wstępnej pracy można stwierdzić, że miesięczne stopy zwrotu w zdecydowanej większości funduszy akcyjnych charakteryzują się normalnością rozkładu. Świadczy to o braku jakiejkolwiek tendencji w wynikach osiąganych przez zarządzających. Klienci funduszy raczej oczekiwaliby, że stopy zwrotu wybranego przez nich funduszu nie będą przypadkowe, ale będą wykazywały tendencję do wzrostu. Uzyskany wynik wydaje się uzasadniać niezbyt pochlebną opinię o zarządzających, co gorsza, dotyczy to praktycznie wszystkich funduszy. Potwierdzeniem takiego wniosku jest fakt, że z punktu widzenia statystycznego fundusze akcyjne osiągały jednako-

we miesięczne stopy zwrotu i ryzyka. Niestety, odwołując się do badań z okresu hossy (31.01.2003-31.01.2007) wnioski są niemal identyczne, zatem koniunktura panująca na rynku kapitałowym nie ma wpływu na podejście zarządzających do powierzonych im oszczędności. Pojawia się więc pytanie: Dlaczego fundusz pobiera opłatę za zarządzanie, skoro w uzyskanych wynikach „brak" aktywnego zarządzania? Zarządzający portfelami różnych funduszy nie są w stanie wygenerować istotnie lepszych stóp zwrotu niż ich konkurenci w czasie hossy, ani ustrzec się przed większymi spadkami w czasie bessy.

W przypadku funduszy zrównoważonych w zdecydowanej większości przypadków stopy miesięczne nie mają rozkładu normalnego, natomiast dla stóp tygodniowych hipoteza zerowa o normalności rozkładu stóp zwrotu została odrzucona dla wszystkich funduszy. Te fundusze zrównoważone, których jednostki uczestnictwa charakteryzowały się normalnością rozkładu (4 fundusze) osiągały statystycznie te same średnie stopy zwrotu i jednakowe ryzyko – podobnie jak w przypadku funduszy akcyjnych. Wyniki badania funduszy zrównoważonych dla okresu hossy (badania w pracy [4]) i bessy dają przeciwne wnioski. Podczas hossy zarówno stopy miesięczne, jak i tygodniowe w większości przypadków charakteryzowały się normalnością rozkładu.

W konsekwencji można stwierdzić, że ewentualna normalność stóp zwrotu jednostek uczestnictwa funduszy zrównoważonych zależy od koniunktury giełdowej i można mieć wątpliwości, czy zaprezentowana metodologia może być stosowana do funduszy z tej klasy ryzyka, ponieważ wartości jednostek uczestnictwa tych funduszy zależą nie tylko od akcji notowanych na giełdzie, ale i od mniej płynnych instrumentów dłużnych.

Badanie stabilności pozycji rankingowej prowadzi do wniosku, że zajmowanie wysokiej pozycji rankingowej nie ma charakteru losowego, natomiast pozycje II i III mają charakter losowy. Wynika stąd, że „liderzy" mają tendencję do pozostawania liderami, bez względu na koniunkturę rynkową (hossa, bessa).

Warto podkreślić, że uzyskane wyniki nie najlepiej świadczą o umiejętnościach zarządzających portfelami funduszy, czy też o braku należytej motywacji do uzyskiwania wyników lepszych od konkurencji. Być może rozwiązaniem byłaby zmiana prawa, aby wynagrodzenie funduszy ściślej zależało od wyników inwestycyjnych. Uczyniono tak w przypadku funduszy emerytalnych i wydaje się, że skutki już widać. Z badań portalu Analizy Online wynika, że w ostatnim roku stopy zwrotu funduszy emerytalnych okazały się lepsze niż funduszy stabilnego wzrostu, co można przypisać zmianie ustawy o funkcjonowaniu OFE.

Literatura

1. Domański Cz., Statystyczne testy nieparametryczne, PWE, Warszawa 1979.
2. Domański Cz., Pruska K., Nieklasyczne metody statystyczne, PWE, Warszawa 2000.
3. Dobosz M., Wspomagana komputerowo statystyczna analiza wyników badań, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2004.
4. Karpio A., Żebrowska-Suchodolska D., Losowość wyników inwestycyjnych osiąganych przez FIO funkcjonujące na polskim rynku kapitałowym, w: Matematyczne aspekty ekonomii, ryzyko-reasekuracja-równowaga, red. W. Kulpa, Uniwersytet Kardynała Stefana Wyszyńskiego, Warszawa 2008.
5. Luszniewicz A., Słaby T., Statystyka z pakietem komputerowym STATISTICA PL, C.H. Beck, Warszawa 2003.
6. Witkowska D., Podstawy ekonometrii i teorii prognozowania, Oficyna Ekonomiczna, Kraków 2005.

RANDOMNESS OF THE INVESTMENT RESULTS OF THE FIO OPERATING ON POLISH CAPITAL MARKET

Summary

The dynamic development of the investment funds market in Poland lets us carry out research regarding different aspects of their functioning. A potential participant of the funds must especially answer such questions as: Which fund should he choose?, Has already chosen fund got any chances to become the leader for a longer period of time?

The work is devoted to one of the variants of answering the problem that is being discussed, especially it analyses the repetitiveness of the FIO investment results achieved in 2003-2007 and 2007-2011. From the funds' customers' points of view, this is a fundamental issue because entrusting their savings with a "good" fund, one wants to know whether it still will be achieving satisfactory investment results. Maybe a good ranking position at a certain moment is accidental? The research was done on the funds from different risk categories such as equity funds, balanced ones.

Adam Krzemienowski
Politechnika Warszawska

WIELOWYMIAROWA WARUNKOWA WARTOŚĆ ZAGROŻONA JAKO MIARA RYZYKA*

Wprowadzenie

Praca przedstawia nową miarę ryzyka nazwaną wielowymiarową warunkową wartością zagrożoną (Multivariate Conditional Value-at-Risk – MCVaR) jako narzędzie wyboru w warunkach ryzyka. Zakłada się, że miara ma służyć do pomiaru ryzyka wielowymiarowego, definiowanego jako wielowymiarowa (wektorowa) zmienna losowa, której elementy (współrzędne) reprezentują realizacje jednowymiarowych zmiennych losowych modelujących rozważane czynniki ryzyka. Przyjmuje się, że jednowymiarowe zmienne losowe są zależne w sensie stochastycznym, a ich struktura zależności jest dana funkcją połączenia (copula function). Ponadto czynniki ryzyka są w pełni substytucyjne, tzn. odpowiednie zmienne losowe są wyrażone w tych samych jednostkach (np. monetarnych).

W celu zdefiniowania miary wprowadza się pojęcie „kwantyla wielowymiarowego", określonego jako hiperprostopadłościan obejmujący najgorsze realizacje wielowymiarowej zmiennej losowej o łącznym prawdopodobieństwie równym rzędowi kwantyla. Warto zauważyć, że tak zdefiniowanych kwantyli danego rzędu jest nieskończenie wiele. MCVaR jest definiowana jako najgorsza oczekiwana realizacja w ramach kwantyla danego rzędu (tzn. najmniejsza, jeśli większe wartości są preferowane).

MCVaR jest skalarną miarą ryzyka wielowymiarowego, pozwalającą na parametryzowanie poziomu awersji do ryzyka od skrajnego pesymizmu po neutralność względem ryzyka poprzez zmianę rzędu kwantyla. Miara jest typem pesymistycznej miary ryzyka i definiuje niemalże ten sam rodzaj ryzyka co popularna i dobrze zbadana warunkowa wartość zagrożona (Conditional Value-at-Risk – CVaR) [8].

* Praca wykonana w ramach grantu NCN N N111 4534400 „Wielowymiarowa warunkowa wartość zagrożona jako miara ryzyka".

Ściślej, dla ryzyka jednowymiarowego miary są tożsame, natomiast dla ryzyka wielowymiarowego różnią się definicją kwantyli – definicja CVaR prowadzi do jednoznacznie określonego kwantyla wielowymiarowego danego rzędu, którym jest hiperostrosłup foremny z wierzchołkiem zaczepionym w najgorszej realizacji wielowymiarowej zmiennej losowej.

Istotną zaletą MCVaR jest jej model optymalizacyjny, który pozwala na efektywne rozwiązywanie rzeczywistych problemów decyzyjnych w warunkach ryzyka, opierając się na pełnej informacji niesionej przez wielowymiarową zmienną losową. Ze wstępnej analizy wynika, że model optymalizacyjny wielowymiarowej warunkowej wartości zagrożonej jest modelem programowania liniowego z nieskończoną liczbą ograniczeń (wynika to z niejednoznaczności kwantyla wielowymiarowego). Niemniej jednak dualna postać tego modelu może być efektywnie rozwiązywana przy użyciu techniki generacji kolumn opartej na dekompozycji Dantziga-Wolfe'a [4], dla której zadanie nadrzędne jest problemem liniowym, a podrzędne – nieliniowym, niewypukłym (z uwagi na funkcję połączenia określającą strukturę zależności pomiędzy zmiennymi losowymi).

MCVaR, w odróżnieniu od CVaR, pozwala na zastosowanie techniki generacji kolumn, która ma własność dostępu do potencjalnie nieskończonej liczby realizacji wielowymiarowej zmiennej losowej, gdyż są one generowane na bieżąco i nie muszą być przechowywane w pamięci komputera. Umożliwia to wyznaczenie rozwiązania optymalnego problemu wyboru w warunkach ryzyka, opierając się na wszystkich realizacjach wielowymiarowej zmiennej losowej, których liczba może być przeogromna – innymi słowy, na pełnej informacji niesionej przez wielowymiarową zmienną losową. Nie jest to możliwe w przypadku CVaR, której najbardziej efektywny obliczeniowo model optymalizacyjny [6] wymaga, aby wszystkie realizacje były przechowywane w pamięci komputera, co przekłada się na niewielką ilość informacji, na podstawie których wyznaczane jest rozwiązanie optymalne problemu.

1. Definicja wielowymiarowej warunkowej wartości zagrożonej

Rozważmy n-wymiarowy wektor losowy $\mathbf{R} = (R_1, \ldots, R_n)^T$, którego każda składowa reprezentuje czynnik ryzyka. Ograniczymy przestrzeń ryzyka do wektorów $\mathbf{R} \in L_n^1(\Omega, F, \mathbf{P})$ o wartościach rzeczywistych oraz załóżmy, że czynniki ryzyka są w pełni substytucyjne, tj. wyrażone w tych samych jednostkach. Niech F_i będzie dystrybuantą zmiennej losowej R_i, $i = 1, \ldots, n$, tj. $F_{R_i}(\eta) = \mathrm{P}(R_i \leq \eta)$. Zmienne losowe R_i zależą od siebie w sensie stochastycz-

nym i ich struktura zależności jest dana funkcją połączenia C. W szczególności $H(\xi, \ldots \zeta) = C(F_{R_1}(\xi), \ldots F_{R_n}(\zeta))$, gdzie H jest dystrybuantą łącznego rozkładu wektora losowego $\mathbf{R}$. Dalej, niech $F_{R_i}^{(-1)}$ będzie lewostronnie ciągłą odwrotnością F_{R_i} (funkcją kwantlową), tj. $F_{R_i}^{(-1)}(p) = \inf\{\eta : F_{R_i}(\eta) \geq p\}$.

Dla ustalonego poziomu tolerancji $\beta \in (0,1]$ definiujemy wielowymiarową warunkową wartość zagrożoną ($\mathbf{MCVaR}_\beta$) jako:

$$\mathrm{MCVaR}_\beta(\mathbf{R}) = \frac{1}{\beta} \min_{u_i \in [\beta,1]} \int_0^{u_1} \cdots \int_0^{u_n} \sum_{i=1}^{n} F_{R_i}^{(-1)}(p_i)\ dC(p_1, \ldots, p_n)$$

$$\text{p. w. } C(u_1, \ldots, u_n) = \beta. \tag{1}$$

Dla ustalonych poziomów $u_i \in (0,1]$ zdefiniujmy wielowymiarowy kwantyl rozkładu wektora losowego $\mathbf{R}$:

$$\mathbf{Q}(u_1, \ldots, u_n) = \{(q_1, \ldots, q_n)^T \ : \ q_i = F_{R_i}^{(-1)}(u_i) \ \text{ dla } \ i = 1, \ldots, n\}$$

Wprowadźmy zbiór wielowymiarowych kwantyli rzędu β:

$$\mathcal{S}_\beta = \{\mathbf{Q}(u_1, \ldots, u_n) \ : \ C(u_1, \ldots, u_n) = \beta\}$$

Jeśli nie ma skoku w optymalnym wielowymarowym β-kwantylu, wartość MCVaR równa się minimalnej warunkowej wartości oczekiwanej sumy czynników ryzyka, przy warunku $\mathbf{R} \leq \mathbf{Q}$ dla wszystkich $\mathbf{Q} \in \mathcal{S}_\beta$, tj.:

$$\mathrm{MCVaR}_\beta(\mathbf{R}) = \min_{\mathbf{Q} \in \mathcal{S}_\beta} \mathbb{E}(\mathbf{1}^{\mathrm{T}}\mathbf{R} | \mathbf{R} \leq \mathbf{Q})$$

Warto zauważyć, że $\mathbf{MCVaR}_\beta$ dąży do $\min_{\omega \in \Omega}(\mathbf{1}^{\mathrm{T}}\mathbf{R})$ dla $\beta \to 0$ i $\mathbf{R}$ ograniczonego z dołu oraz równa się $\mathbb{E}(\mathbf{1}^{\mathrm{T}}\mathbf{R})$, gdy $\beta = 1$. W związku z tym miara obejmuje całe spektrum ostrożnych postaw względem ryzyka, począwszy od skrajnej awersji, a skończywszy na neutralności względem ryzyka.

2. Związek miary z warunkową wartością zagrożoną

Warunkowa wartość zagrożona CVaR jest miarą ryzyka jednowymiarowego. Dla ustalonego poziomu $\beta \in (0,1]$ definiujemy CVaR_β jako średnią w ramach β-kwantyla, tj.:

$$\mathrm{CVaR}_\beta(R) = \frac{1}{\beta} \int_0^{\beta} F_R^{(-1)}(p)\ dp$$

CVaR jest miarą koherentną (zob. np. [7]) i liczne badania empiryczne (zob. np. [1; 5; 8]) potwierdziły jej przydatność w różnych zagadnieniach finansowych. Miara ta jest również stosowana do pomiaru ryzyka wielowymiarowego – do miary przekazywana jest suma zmiennych losowych reprezentujących czynniki ryzyka, tj.:

$$\mathrm{CVaR}_{\beta}\left(\mathbf{1}^{\mathrm{T}}\mathbf{R}\right) = \frac{1}{\beta}\int_{0}^{\beta} F_{\mathbf{1}^{T}\mathbf{R}}^{(-1)}(p)\, dp \tag{2}$$

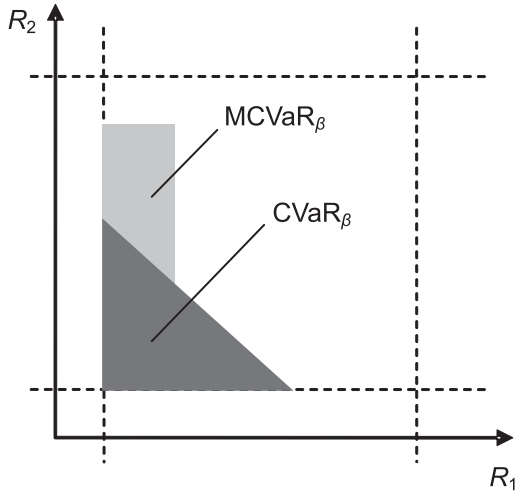
Warto zauważyć, że chociaż obie miary wyznaczają średnie w ramach kwantyli, w przypadku wielowymiarowym CVaR jest miarą bardziej ostrożną niż MCVaR. Ze wzoru (2) wynika, że CVaR używa półpłaszczyzny jako wielowymiarowego kwantyla i w związku z tym obejmuje wszystkie najgorsze przypadki (najmniejsze wartości) dla zadanego poziomu tolerancji β. Nie jest to prawdą w przypadku miary MCVaR, która obejmuje tylko pewną część najgorszych przypadków w ramach zadanego poziomu tolerancji β, ponieważ zgodnie z definicją (1) używa stożka jako kwantyla. W związku z tym dla dowolnego wektora losowego $\mathbf{R}$ i $\beta \in (0,1]$ zachodzi:

$$\mathrm{MCVaR}_{\beta}(\mathbf{R}) \geq \mathrm{CVaR}_{\beta}(\mathbf{1}^{\mathrm{T}}\mathbf{R}) \tag{3}$$

W przypadku jednowymiarowym obie miary się pokrywają, tj. dla dowolnej zmiennej losowej R i $\beta \in (0,1]$:

$$\mathrm{MCVaR}_{\beta}(R) = \mathrm{CVaR}_{\beta}(R) \tag{4}$$

Rysunek 1 przedstawia graficzne porównanie miar $\mathbf{CVaR}_{\beta}$ i $\mathbf{MCVaR}_{\beta}$ dla dwuwymiarowego wektora losowego $\mathbf{R}$ ograniczonego z dołu i z góry.



Rys. 1. CVaR_{β} i MCVaR_{β} dla dwuwymiarowego wektora losowego

3. Własności miary

Wielowymiarowa warunkowa wartość zagrożona $MCVaR_\beta$ ma następujące własności:

(i) $MCVaR_\beta$ jest translacyjnie ekwiwariantna (translation-equivariant), tj.:

$$MCVaR_\beta\ (R + c) = MCVaR_\beta\ (R) + 1^T c$$

dla stałego wektora c,

(ii) $MCVaR_\beta$ jest dodatnio homogeniczna (positively homogenous), tj.:

$$MCVaR_\beta\ (\lambda R) = \lambda\ MCVaR_\beta\ (R)$$

gdy $\lambda > 0$.

(iii) $MCVaR_\beta$ w ogólności nie jest monotoniczna, tj. jeśli:

$$R_1(\omega) \geq R_2(\omega) \text{ dla każdego } \omega \in \Omega$$

to nie zawsze (iv) $MCVaR_\beta$ jest nadaddytywna (superadditive) w następującym sensie:

$$MCVaR_\beta(\mathbf{R}) \geq \sum_{i=1}^{n} MCVaR_\beta(R_i)$$

Dowód

Własności (i) i (ii) są oczywiste z definicji $MCVaR_\beta$.

Pokażmy kontrprzykład jako dowód (iii). Rozważmy dwa losowe wektory $\mathbf{R_1}$ i $\mathbf{R_2}$ z następującymi rozkładami:

i	$\mathbf{R_1}\ (\omega_i)$	$\mathbf{R_2}\ (\omega_i)$	$\mathbf{P}\ (\omega_i)$
1	$(1;0)^T$	$(1;0)^T$	0,1
2	$(0;1)^T$	$(0;1)^T$	0,1
3	$(2;2)^T$	$(1;1)^T$	0,8

Zachodzi relacja $\mathbf{R_1}\ (\omega) \geq \mathbf{R_2}\ (\omega)$ dla każdego $\omega \in \Omega$. Obliczmy wartość $\mathbf{MCVaR_{0,2}}\ (\mathbf{R_1})$:

$$MCVaR_{0,2}(\mathbf{R}_1) = \frac{1}{0,2}((1+0)\cdot 0,1 + (0+1)\cdot 0,1) = 1$$

W przypadku $\mathbf{MCVaR_{0,2}}\ (\mathbf{R_2})$ atomy prawdopodobieństwa zostaną podzielone z uwagi na fakt, że wektory $(1;0)^T$, $(0;1)^T$ i $(1;1)^T$ leżą na obwodzie stożka (kwantyla). W związku z tym wartość $\mathbf{MCVaR_{0,2}}\ (\mathbf{R_2})$ może być wyznaczona przez rozwiązanie następującego zadania optymalizacji:

$$\min_{w_1,w_2} \frac{1}{0,2}((1+0)\cdot 0,1w_1 + (0+1)\cdot 0,1w_2 + (1+1)\cdot 0,8w_1w_2)$$

$$\text{p.w. } 0,1w_1 + 0,1w_2 + 0,8w_1w_2 = 0,2$$

$$0 \leq \mathbf{w}_i \leq 1 \text{ dla } i = 1,2$$

Po rozwiązaniu dostajemy $\mathrm{MCVaR}_{0,2}(\mathbf{R}_2) = 13/9 > 1 = \mathrm{MCVaR}_{0,2}(\mathbf{R}_1)$. W związku z tym MCVaR nie jest monotoniczna. W celu udowodnienia własności (iv) użyjemy zależności (3) i (4) oraz faktu, że CVaR jest nadaddytywna (zob. [7]). Wykorzystując powyższe dostajemy:

$$\mathrm{MCVaR}_{\beta}(\mathbf{R}) \geq \mathrm{CVaR}_{\beta}\left(\mathbf{1}^{\mathrm{T}}\mathbf{R}\right) \geq \sum_{i=1}^{n} \mathrm{CVaR}_{\beta}(R_i) = \sum_{i=1}^{n} \mathrm{MCVaR}_{\beta}(R_i)$$

Artzner i in. [2; 3] nazywają miarę ryzyka koherentą, jeżeli jest translacyjnie ekwiwariantna, dodatnio homogeniczna, monotoniczna i nadaddytywna. MCVaR_{β} nie jest koherentna w tym sensie, ponieważ nie jest monotoniczna [9].

Rozważmy następujący przypadek szczególny związany z optymalizacją portfela czynników ryzyka: Jesteśmy zainteresowani wyborem optymalnych proporcji czynników ryzyka $\boldsymbol{R_i}$ w sensie miary, skalowanych wagami portfela $\boldsymbol{x_i}$, tj. maksymalizacją $\mathrm{MCVaR}_{\beta}(x \circ R)$, gdzie $\circ$ jest operatorem iloczynu Hadamarda. Dla liniowo przekształconych wektorów losowych MCVaR_{β} zachowuje własność monotoniczności, tj. jeśli:

$$x \circ R(\omega) \geq R(\omega) \text{ dla każdego } \omega \in \Omega$$

to:

$$\mathrm{MCVaR}_{\beta}(x \circ R) \geq \mathrm{MCVaR}_{\beta}(R)$$

Dowód

Założenie twierdzenia zachodzi tylko dla $x = (x_P, x_N)^T$, $\mathbf{R} = (R_P, R_N)^T$, $x_P \geq 1$ i $x_N \leq 1$, gdzie P i N są zbiorami indeksów zdefiniowanymi następująco: $P = \{p : R_p(\omega) \geq 0 \text{ dla każdego } \omega \in \Omega\}$, $N = \{n : R_n(\omega) \leq 0 \text{ dla każdego } \omega \in \Omega\}$.

Z definicji $\mathbf{MCVaR}_{\beta}$ wynika, że:

$$\mathrm{MCVaR}_{\beta}((\mathbf{x}_P, \mathbf{x}_N)^{\mathrm{T}} \circ (\mathbf{R}_P, \mathbf{R}_N)^{\mathrm{T}}) \geq \mathrm{MCVaR}_{\beta}((\mathbf{R}_P, \mathbf{R}_N)^{\mathrm{T}})$$

dla dowolnych $x_p \geq 1$ i $x_N \leq 1$.

W związku z tym MCVaR zachowuje koherentność w problemach, gdzie rozważona jest kombinacja liniowa czynników ryzyka.

Podsumowanie

Wielowymiarowa warunkowa wartość zagrożona jest skalarną, kwantylową miarą ryzyka modelowanego wielowymiarową zmienną losową. Zakłada się, że czynniki ryzyka są w pełni substytucyjne, tzn. wyrażone w tych samych jednostkach. Wielowymiarowa warunkowa wartość zagrożona jest miarą ekstremal-

ną, umożliwiającą parametryzowanie poziomu awersji do ryzyka, począwszy od skrajnego pesymizmu, po neutralność względem ryzyka poprzez zmianę rzędu kwantyla. Miara w ogólności nie jest miarą koherentną, ale dla problemu wyboru portfela czynników ryzyka zachowuje własność koherentności. Istotną zaletą wielowymiarowej warunkowej wartości zagrożonej jest możliwość wyznaczenia optymalnego portfela czynników ryzyka z wykorzystaniem techniki generacji kolumn w oparciu o pełną informację niesioną przez wielowymiarową zmienną losową. Nie jest to możliwe przy użyciu innych typowych kwantylowych miar ryzyka, jak wartość zagrożona czy warunkowa wartość zagrożona.

Literatura

1. Andersson F., Mausser H., Rosen D., Uryasev S., Credit risk optimization with Conditional Value-at-Risk criterion. „Mathematical Programming" 2001, No. 89.
2. Artzner P., Delbaen F., Eber J.-M., Heath D., Thinking coherently, „Risk" 1997, No. 10.
3. Artzner P., Delbaen F., Eber J.-M., Heath D., Coherent measures of risk. „Mathematical Finance" 1999, No. 9.
4. Dantzig G.B., Wolfe P., The decomposition algorithm for linear programs, „Econometrica" 1961, No. 29(4).
5. Mansini R., Ogryczak W., Speranza M.G., LP solvable models for portfolio optimization: A classification and computational comparison, „IMA Journal of Management Mathematics" 2003, No. 14.
6. Ogryczak W., Śliwiński T., On solving the dual for portfolio selection by optimizing Conditional Value at Risk, „Computational Optimization and Applications" 2011, No. 50.
7. Pflug G.Ch., Some remarks on the Value-at-Risk and the Conditional Value-at-Risk, in: Probabilistic Constrained Optimization: Methodology and Applications, ed. S. Uryasev, Kluwer A.P., Dordrecht 2000.
8. Rockafellar R.T., Uryasev S., Optimization of Conditional Value-at-Risk. „Journal of Risk" 2000, No. 2.
9. Rockafellar R.T., Uryasev S., Conditional Value-at-Risk for general distributions. „Journal of Banking and Finance" 2002, No. 26.

THE MULTIVARIATE CONDITIONAL VALUE-AT-RISK AS A MEASURE OF RISK

Summary

The Multivariate Conditional Value-at-Risk (MCVaR) is a scalar risk measure for multivariate risks modeled by multivariate random variables. It is assumed that the univariate risk components are perfect substitutes, i.e., they are expressed in the same units. MCVaR is a quantile risk measure that allows one to emphasize the consequences of more pessimistic scenarios. By changing the level of the quantile, the measure permits to parameterize prudent attitudes toward risk ranging from extreme risk aversion to risk neutrality. In terms of definition, MCVaR is slightly different from the popular and well-researched Conditional Value-at-Risk (CVaR). Nevertheless, this small difference allows one to efficiently solve MCVaR portfolio optimization problems based on the full information carried by a multivariate random variable using column generation technique, which is not possible in the case of CVaR.

Stanisław Lewiński vel Iwański
Sebastian Tomczak
Zofia Wilimowska
Politechnika Wrocławska

SYTUACJA FINANSOWA POLSKICH SPÓŁEK A MODELE KSZTAŁTOWANIA STRUKTURY KAPITAŁOWEJ*

Wprowadzenie

Ustalenie racjonalnej struktury finansowej jest warunkiem koniecznym prawidłowego funkcjonowania przedsiębiorstwa. Odpowiednia struktura finansowa (kapitałowa) może stanowić o jego przewadze konkurencyjnej, natomiast nieprawidłowa może powodować takie problemy, jak: wysokie ryzyko finansowe, niewypłacalność, niska zdolność kredytowa, obniżona wartość rynkowa, zagrożenie bankructwem.

Zagadnienia związane ze strukturą finansową są często kanwą ekonomicznych rozważań. Prace teoretyczne i empiryczne wielu ekonomistów (w szczególności zachodnich) zmierzają do udzielenia odpowiedzi na pytanie: Czy dla każdego przedsiębiorstwa istnieje optymalna struktura finansowa, jak również, w jakim zakresie ma ona wpływ na jego wartość? Mimo że większość naukowców skłania się do stwierdzenia, iż w zasadzie dla każdego przedsiębiorstwa można znaleźć najlepszą w danych warunkach strukturę finansową zwiększającą jego wartość, to jednak nadal istnieje rozbieżność w sposobie podejścia do problemu.

Próbując odpowiedzieć na powyższe pytanie należy się odnieść do rynkowej zależności pomiędzy strukturą finansową a wartością przedsiębiorstw, jak również wnikliwiej przyjrzeć się stosowanym modelom dyskryminacyjnym, w tym wykorzystywanym w nich wskaźnikom.

* Publikacja współfinansowana przez Unię Europejską w ramach Europejskiego Funduszu Społecznego.

1. Modele zarządzania strukturą finansową

Zajmując się zagadnieniem struktury finansowej nie można zapominać o istniejącej teorii i stosowanych modelach. Powszechnie w teorii i praktyce przyjmuje się (zgodnie z podejściem tradycyjnym), że zwiększając zadłużenie przedsiębiorstwa można zwiększać jego wartość (wykorzystując fakt relatywnie niższego kosztu kapitałów obcych względem kapitałów własnych). Nie może się to jednak odbywać w nieskończoność, ponieważ po przekroczeniu pewnego punktu granicznego wycena firmy będzie spadać (zamiast wzrastać). Związane jest to z ponadproporcjonalnym wzrostem kosztów wynikających z zadłużenia, przewyższających pozytywne efekty tarczy podatkowej. W ujęciu ogólnym (w szczególności statycznym) [2; 12] można stwierdzić, że w sytuacji przekroczenia przez przedsiębiorstwo pewnego granicznego (optymalnego) punktu zadłużenia nastąpi odwrotne działanie efektu dźwigni finansowej (tzw. maczuga finansowa). Firma zamiast uzyskiwać relatywnie niższe koszty kapitałów obcych de facto podwyższy swoje pozostałe koszty związane z zadłużeniem. W tej sytuacji wystąpią zwiększone wymagania inwestorów odnośnie do udziału w zysku (w wyniku rosnącego ryzyka finansowego), banki będą wymagały większych zabezpieczeń dla długu oraz wystąpią koszty związane z zagrożeniem bankructwem.

W ujęciu tradycyjnym istotne jest ujmowanie granicznego poziomu zadłużenia jako stałego w czasie. Zgodnie z tym założeniem menedżerowie przedsiębiorstw dążą do osiągnięcia niezmiennego w czasie, optymalnego poziomu zadłużenia w drodze tzw. procesu „wymiany" pomiędzy korzyściami z większego zadłużania się a kosztami zaburzeń finansowych wynikających z nadmiernego zaciągania długu.

Dynamiczne podejście do zarządzania strukturą finansową różni się od statycznego głównie ujęciem zmienności struktury finansowej w czasie. W praktyce obszaru dynamicznych modeli (m.in. E. Fischer, R. Heinkel i J. Zechner [3], R. Goldstein, N. Ju i H. Leland [4] oraz H. DeAngelo, L. DeAngelo i T.M. Whited [1]) przyjmuje się, że przedsiębiorstwa mogą oddalać się od swoich docelowych długoterminowych „optymalnych" struktur finansowania w zależności od bieżących warunków rynkowych. Jednocześnie przedsiębiorstwa te będą konsekwentnie szukać sposobów na powrót do docelowej struktury finansowej poprzez redukcję długu zaciągniętego (tymczasowo) w związku z bieżącymi możliwościami inwestycyjnymi lub koniecznością realizacji dochodów (np. wypłaty dywidend). We wskazanym obszarze (w szczególności według Harry'ego DeAngelo, Lindy DeAngelo i Toniego M. Whiteda) ważne znaczenie ma fakt krótkoterminowości zaciąganego zadłużenia [1], o której nie przesądza czas wymagalności, lecz okres

wykorzystania długu przez rozważane przedsiębiorstwo. W proponowanym modelu charakter docelowej struktury finansowej jest uzależniony od natury możliwości inwestycyjnych. Docelowa struktura finansowa może być tymczasowo zmieniana w zależności od bieżących i potencjalnych zaburzeń w procesie inwestycyjnym*. W modelu dług jest traktowany jako rzadki zasób, którego optymalna użyteczność zależy od bieżących i potencjalnych zaburzeń/potrzeb inwestycyjnych.

Na uwagę zasługuje również fakt, że statystyczna sprawdzalność wskazanego modelu [1, s. 241], jak i innych z obszaru dynamicznego jest bardzo wysoka dla niefinansowych przedsiębiorstw amerykańskich**, jednak modelu tego (podobnie jak i innych modeli amerykańskich i zachodnich) raczej nie można w pełni transponować do warunków polskich. Jest to związane głównie z odmienną specyfiką polskiego rynku kapitałowego, który jest mniej stabilny i mniej dojrzały niż rynki zagraniczne. Polskie przedsiębiorstwa mają także odmienną strukturę terminową zadłużenia niż spółki zachodnie.

2. Modele dyskryminacyjne a finansowanie

Oprócz zagadnienia kształtowania struktury finansowej i jej wpływu na wartość przedsiębiorstwa, istotna jest również problematyka oceny prawidłowości już zastanej struktury finansowej przedsiębiorstw. Ocenie tej dedykowane są modele dyskryminacyjne. Modele analizy są najpopularniejszą grupą metod prognozowania niewypłacalności przedsiębiorstw. Istota tych metod ogranicza się do rozdzielania zbioru danych na dwie bądź więcej klas, co ma ważne znaczenie informacyjne dla badacza.

W analizie dyskryminacyjnej wykorzystuje się analizę wskaźnikową oraz sformalizowane narzędzia wnioskowania (metody analizy statystycznej). Wskaźnik syntetyczny (agregatowy) konstruowany jest na podstawie danych empirycznych. Na wskaźnik ten składa się kilka wskaźników cząstkowych, którym przypisuje się współczynniki tzw. wagi (rozróżniają one znaczenie poszczególnych wskaźników). Zastosowanie analizy dyskryminacyjnej dąży do redukcji wielowymiarowej przestrzeni zmiennych do jednej zmiennej syntetycznej, agregatowej [6].

---

* Zaburzenia są tu rozumiane jako nieprzewidziane i niekorzystne czynniki mające wpływ na wysokość kosztów procesu inwestycyjnego.

** Dane dotyczą przedsiębiorstw funkcjonujących w Stanach Zjednoczonych w latach 1988-2001.

W opracowaniu zostaną przedstawione trzy modele analizy dyskryminacyjnej, w których wykorzystywany jest minimum jeden wskaźnik struktury finansowej (mający jednocześnie najwyższą wagę w modelu). W dalszej części modele te zostaną zweryfikowane pod kątem empirycznym w celu oceny ich skuteczności w prognozowaniu upadłości przedsiębiorstw.

Model Legaulta

Pierwszym z omawianych modeli dyskryminacyjnych jest model Legaulta [9] opisywany według wzoru:

$$\text{CA-Score} = 4.5913\ X1 + 4.5080\ X2 + 0.3936\ X3 - 2.7616 \quad (1)$$

gdzie:
X1 – kapitał własny/aktywa ogółem,
X2 – zysk brutto + pozycje nadzwyczajne + koszty finansowe/aktywa ogółem,
X3 – przychody ze sprzedaży/aktywa ogółem.

W modelu najwyższą wagę ma wskaźnik struktury X1.

Model Mączyńskiej i Zawadzkiego [11, s. 21]

Kolejny model z omawianej grupy został skonstruowany przez E. Mączyńską oraz M. Zawadzkiego. Opisywany jest według wzoru:

$$Z = -1{,}498 + 9{,}498X1 + 3{,}566X2 + 2{,}903X3 + 0{,}452X4 \quad (2)$$

gdzie:
X1 – zysk operacyjny/aktywa ogółem,
X2 – kapitał własny/aktywa ogółem,
X3 – zysk netto + amortyzacja/zobowiązania ogółem,
X4 – aktywa obrotowe/zobowiązania krótkoterminowe.

W modelu najwyższą wagę ma wskaźnik struktury finansowej X2.

Model Hadasik [5, s. 154]

Ostatni z omawianych modeli dyskryminacyjnych został skonstruowany przez D. Hadasik. Model ten opisuje wzór:

$$PG2 = 0{,}703585X1 - 1{,}2966X2 - 2{,}21854X3 + 1{,}52891X4 + \\ + 0{,}00254294X5 - 0{,}0140733X6 + 0{,}0186057X7 + 2{,}36261 \quad (3)$$

gdzie:
X1 – aktywa obrotowe/zobowiązania krótkoterminowe,
X2 – aktywa obrotowe – zapasy/zobowiązania krótkoterminowe,
X3 – zobowiązania ogółem/aktywa ogółem,
X4 – kapitał obrotowy/aktywa ogółem,
X5 – należności * 365/przychody ze sprzedaży,
X6 – zapasy * 365/przychody ze sprzedaży,
X7 – zysk netto/zapasy.

W modelu najwyższą wagę ma wskaźnik struktury finansowej X3.

Wskazane modele składają się z różnych wskaźników. Dla każdego z nich przypisana jest inna waga. W tym miejscu przedstawione zostaną wagi wskaźników struktury finansowej w modelach. Wyszczególnione do dalszej analizy wskaźniki struktury finansowej są następujące:
– wskaźnik udziału kapitału własnego w finansowaniu majątku (X1 w modelu Legaulta oraz X2 w modelu Mączyńskiej),
– wskaźnik ogólnego zadłużenia (X3 w modelu Hadasik),
– wskaźnik pokrycia zobowiązań nadwyżką finansową (X3 w modelu Mączyńskiej i Zawadzkiego).

Każdy z wymienionych wskaźników struktury finansowej charakteryzuje się przypisaną najwyższą wagą w analizowanych modelach, zgodnie z tab. 1.

Tabela 1

Składowe wybranych modeli dyskryminacyjnych

Model	Liczba wskaźników	Waga	Suma wag	Udział % wskaźników struktury
Legaulta	3	4,591	9,4929	48,38
Mączyńskiej i Zawadzkiego	4	6,469	16,419	39,40
Hadasik	7	2,21854	5,0792	43,68

Analizując tab. 1 według wagi wskaźnika można zauważyć, że to właśnie w pierwszym modelu najwyższy jest udział wskaźnika struktury finansowej – 48,38% (mowa o wskaźniku udziału kapitału własnego w finansowaniu majątku). Nieco niższy jest udział procentowy wskaźnika struktury w modelu Hadasik (mowa o wskaźniku ogólnego zadłużenia). Najniższy udział procentowy (według wagi wskaźnika) mają wskaźniki struktury finansowej w modelu Mączyńskiej, są to wskaźniki: kapitału własnego oraz pokrycia zobowiązań nadwyżką finansową. Wysoki procentowy udział określonego wskaźnika struktury finansowej w modelu może wskazywać na jego predykcyjność.

3. Badania struktury finansowej przedsiębiorstw

Na początku zaprezentowane zostaną wyniki analiz dla spółek, które ogłosiły upadłość. Badania te przedstawią możliwości wybranych modeli dyskryminacyjnych w zakresie predykcji wystąpienia bankructwa przedsiębiorstw posiadających określoną strukturę finansową. Następnie przedstawione zostaną rezultaty analiz dla spółek notowanych na GPW, funkcjonujących w dalszym ciągu na rynku. Zaprezentowane badania dotyczą wpływu struktury kapitałowej przedsiębiorstw na ich wartość rynkową.

3.1. Modele dyskryminacyjne w ocenie struktury kapitału

Analizie empirycznej zostały poddane wyselekcjonowane i opisane w poprzednim rozdziale modele dyskryminacyjne pod kątem ich predykcyjności, czyli zdolności do przewidywania upadłości przedsiębiorstw. Do badań wykorzystano uproszczone raporty finansowe poszczególnych upadłych spółek. Źródłem uproszczonych raportów finansowych jest baza EMIS. Analizę przeprowadzono dla 100 spółek, które ogłosiły upadłość w okresie obejmującym lata 2005-2010. Należy zaznaczyć, że w badanym okresie liczba upadłych spółek była znacznie wyższa, jednak nie dla wszystkich upadłych spółek dostępne były niezbędne dane finansowe. Dla zgromadzonych sprawozdań finansowych przeprowadzono obliczenia wybranych modeli na rok, dwa i trzy lata przed bankructwem. Wyniki dla modeli przedstawiono w tab. 2.

Tabela 2

Zdolność modeli w przewidywaniu upadłości

Model	Rok przed bankructwem	Dwa lata przed bankructwem	Trzy lata przed bankructwem
Legaulta	87,00%	65,00%	67,00%
Mączyńskiej i Zawadzkiego	75,00%	44,00%	43,00%
Hadasik	69,00%	45,00%	41,00%

Źródło: Opracowanie na podstawie wyników analizy empirycznej.

Analizując modele przedstawione w tab. 2 można powiedzieć, że model Legaulta uzyskał najwyższy poziom przewidywania upadłości przedsiębiorstw w badanym okresie. Jak już wcześniej wspomniano, model ten składa się z trzech wskaźników, wśród których najwyższą wagę ma wskaźnik udziału kapitału

własnego w finansowaniu majątku. Poziom tego wskaźnika w modelu Legaulta w znacznej mierze wpływa na zdolność przewidywania upadłości przedsiębiorstw.

Oceniając pozostałe dwa modele można zauważyć, że znacznie słabiej przewidują upadłość firm na dwa oraz na trzy lata przed upadłością. Możliwą przyczynę można wiązać z rozmytością wpływu wskaźnika struktury finansowej na zachowanie modelu.

3.2. Wpływ struktury finansowej na wartość rynkową przedsiębiorstw

W ramach analizy empirycznej, obok analizy oceny struktury finansowej spółek upadłych przeprowadzono także badania weryfikacji założeń istniejącej teorii kształtowania struktury finansowej (teoria „wymiany" oraz dynamiczne modele struktury finansowej). W 2010 r. dla próby 100 niefinansowych spółek notowanych na GPW* przeprowadzono analizy dotyczące wpływu struktury finansowej na ich wartość rynkową. Omawiane badania dotyczą lat 2004-2009.

Do zobrazowania badanych zależności wykorzystano wskaźnik udziału zadłużenia w strukturze finansowej (relacja długu do kapitałów własnych – D/E) [7] oraz miarę rynkowej wartości dodanej (MVA) [14].

Opisywana analiza opierała się głównie na badaniu współczynników korelacji liniowej pomiędzy miernikami, z uwzględnieniem ich statystycznej istotności (zbadanej testem statystycznym t-Studenta) [10].

W wyniku przeprowadzonych analiz zidentyfikowano istotną statystycznie korelację pomiędzy strukturą finansową (D/E) a rynkową wartością dodaną (MVA) dla 62 spółek spośród analizowanej próby. Poniżej zaprezentowane zostaną wyniki badań cząstkowych dla podgrup, dla których wykryto powyższą zależność.

Pierwsza z podgrup to 28 spółek przemysłu spożywczego, lekkiego i drzewno-papierniczego oraz z zakresu usług deweloperskich. Dla poszczególnych lat otrzymano następujące wyniki korelacji zbioru danych mierników oraz wartości rynkowej dla 28 spółek:

* Notowania giełdowe oraz dane finansowe pochodzą z bazy Money.pl: www.money.pl

Tabela 3

Wyniki analizy dla pierwszej analizowanej podgrupy 28 spółek GPW

Współczynnik korelacji między D/E a MVA	Lata					
	2009	2008	2007	2006	2005	2004
	0,0731	**0,4577**	**0,4884**	**0,2703**	**0,1793**	**-0,0739**
Weryfikacja istotności współczynnika korelacji przy użyciu statystyki t-Studenta						
Liczba spółek (n)	28	28	28	28	28	28
t obliczeniowe	0,374	2,625	2,854	1,431	0,929	-0,378
t tablicowe (1-1/2α,,v)	1,706	1,706	1,706	1,706	1,706	1,706
. - t (1-1/2α,v)	-1,706	-1,706	-1,706	-1,706	-1,706	-1,706
t (1-1/2α,v)	1,706	1,706	1,706	1,706	1,706	1,706
Hipoteza	H0	H1	H1	H0	H0	H0

Tylko dla lat 2007-2008 wystąpiła istotna statystycznie korelacja pomiędzy badanymi cechami. W przypadku pozostałych lat współczynniki korelacji liniowej były znacznie niższe i poddane weryfikacji statystycznej nie okazały się istotne.

Druga podgrupa spółek, dla których zidentyfikowano istotną statystycznie zależność pomiędzy strukturą finansową a wartością rynkową to 34 spółki przemysłu ciężkiego, z zakresu: budownictwa, przemysłu elektromaszynowego i energetycznego. Tabela 4 przedstawia wyniki analizy.

Tabela 4

Wyniki analizy dla drugiej analizowanej podgrupy 34 spółek GPW

Współczynnik korelacji między D/E a MVA	Lata					
	2009	2008	2007	2006	2005	2004
	0,6605	**0,6622**	**0,2114**	**0,2455**	**-0,0955**	**0,0698**
Weryfikacja istotności współczynnika korelacji przy użyciu statystyki t-Studenta						
Liczba spółek (n)	34	34	34	34	34	34
t obliczeniowe	4,976	4,999	1,224	1,432	-0,543	0,396
t tablicowe (1-1/2α,,v)	1,694	1,694	1,694	1,694	1,694	1,694
. - t (1-1/2α,v)	-1,694	-1,694	-1,694	-1,694	-1,694	-1,694
t (1-1/2α,v)	1,694	1,694	1,694	1,694	1,694	1,694
Hipoteza	***H1***	***H1***	H0	H0	H0	H0

Tylko w latach 2008-2009 wystąpiła istotna statystycznie korelacja pomiędzy badanymi cechami. W przypadku pozostałych lat współczynniki korelacji były znacznie niższe i statystycznie nieistotne.

Podsumowując tę część analizy należy stwierdzić, że dla większości analizowanych spółek obserwuje się występowanie zależności pomiędzy ich strukturą finansową a wartością rynkową. Aż dla 62 zanotowano istotną statystycznie korelację liniową pomiędzy badanymi miarami analizowanych zależności. Natomiast

dla pozostałych spółek obserwowano także związek pomiędzy badanymi cechami, niestety nieistotny pod kątem statystycznym (jeżeli bierzemy pod uwagę prostą korelację liniową).

Podsumowanie

Na podstawie dokonanych analiz sprawdzalności wybranych modeli dyskryminacyjnych łatwo zauważyć, że dzięki tym modelom z dużym prawdopodobieństwem można przewidywać upadłość przedsiębiorstw. Istotną kwestią jest to, że w wybranych do analizy modelach wskaźniki struktury finansowej wykorzystywane są jako główne (według wagi). Może to oznaczać, że mierniki te mogą być także przydatne w analizie wpływu źródeł finansowania na wartość rynkową przedsiębiorstw. Z analizowanych modeli najbardziej „sprawdzalny" okazał się model Legaulta, w którym wykorzystywany był wskaźnik udziału kapitału własnego w aktywach ogółem. Ze względu na powyższe, zasadne byłoby uwzględnienie tego wskaźnika w analizach dotyczących wpływu struktury finansowej na wartość rynkową przedsiębiorstw.

Należy podkreślić, że w przypadku polskich spółek giełdowych oddziaływanie źródeł finansowania na wartość przedsiębiorstw jest wyraźnie zauważalne. Aż dla 62 spośród 100 wybranych firm zanotowano istotną statystycznie korelację pomiędzy ich strukturą finansowania (D/E) a wartością rynkową (MVA). Dla pozostałych zauważalne są zależności, których nie „wykrywa" prosta korelacja liniowa.

Ze względu na wskazane wyniki analizy dla modeli dyskryminacyjnych zasadne wydaje się włączenie do dalszych badań także innych wskaźników struktury finansowej (m.in. udziału kapitału własnego w aktywach ogółem), w celu identyfikacji „wyraźniejszego" wpływu na wartość rynkową przedsiębiorstw. Dodatkowe wskaźniki mogą być szczególnie przydatne w dalszej analizie nakierunkowanej na budowę całościowego modelu dla polskich firm.

Literatura

1. DeAngelo H., DeAngelo L., Whited T. M., Capital structure dynamics and transitory debt, „Journal of Financial Economics" 2011, nr 99.
2. Durand D., Cost of Debt and Equity Funds for Business. Trends and Problems of Mea-surement. Conference on Research in Business Finance, National Beaureau of Economic Research, 1952.

3. Fischer E., Heinkel R., Zechner J., Dynamic capital structure choice: theory and tests, „Journal of Finance" 1989, No. 44.
4. Goldstein R., Ju N., Leland H., An EBIT – based model of dynamic capital structure, „Journal of Business" 2001, No. 74(4).
5. Hadasik D., Upadłość przedsiębiorstw w Polsce i metody jej prognozowania, Zeszyty Naukowe, Seria II, nr 153, Akademia Ekonomiczna, Poznań 1998.
6. Hamrol M., Chodakowski J., Prognozowanie zagrożenia finansowego przedsiębiorstwa. Wartość predykcyjna polskich modeli analizy dyskryminacyjnej, "Badania Operacyjne i Decyzje" 2008, nr 3.
7. Jerzemowska M., Kształtowanie struktury kapitału w spółkach akcyjnych, PWN, Warszawa 1999.
8. Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M., Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II. Statystyka matematyczna, Wydawnictwo Naukowe PWN, Warszawa 1994.
9. Legault J.C.A., Score A., Warning System for Small Business Failures, „Bilanas" 1987, June.
10. Maliński M., Weryfikacja hipotez statystycznych wspomagana komputerowo, Politechnika Śląska, Gliwice 2004.
11. Mączyńska E., Zawadzki M., Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw, "Ekonomista" 2006, nr 2.
12. Modigliani F., Miller M., Koszt kapitału, finanse przedsiębiorstw i teoria inwestycji, "The American Economic Rewiev" 1958, No. 48.
13. Ostaszewski J., Źródła pozyskiwania kapitału przez spółkę akcyjną, Difin, Warszawa 2000.
14. Przedsiębiorstwo. Wartość. Zarządzanie, red. C. Suszyński, PWE, Warszawa 2007.
15. Shyam-Sunder L., Myers S.C., Testing static tradeoff against pecking order models of capital structure, „Journal of Financial Economics" 1999, No. 51.

FINANCIAL CONDITION OF POLISH COMPANIES AND MODELS OF CAPITAL STRUCTURE FORMATION

Summary

Running a business is not possible without suitable resources of financing. Determination of rational financial structure is a necessary condition for efficient operation of

a company. Financial resources as well as their configuration may lead either to development or to bankruptcy of a company. Thus, the decision about financial structure is a matter of fundamental importance to subsequent operation of organization. Moreover, this decision is made by the management during continuous company's activity. Unfortunately, both the description of the very process of reaching the decision and its validation seem to be really hard to formalize and to operationalize. Despite the fact that there have been carried out analyses, the issue of managing the financial structure remains to be resolved. Above all, the difficulties reside in the process of devising a suitable and universal model which would be widely accepted, particularly in the context of Polish capital market.

This paper brings up modern models for constructing the capital structure as well as models for analyzing the efficiency of managing such a structure. Polish capital market issue, given in question is also illustrated in the context of administering financial structure in the Polish companies (including the ones which went out of business on stoc.

Maciej Malaczewski
Uniwersytet Łódzki

POSTĘP TECHNICZNY A ROLA ZASOBÓW NATURALNYCH W PROCESIE PRODUKCJI*

Wprowadzenie

Postępujący rozwój technologii produkcyjnych powoduje powstawanie nowych możliwości wykorzystywania różnych form zasobów naturalnych, a istniejące prace teoretyczne nie wykluczają możliwości pojawienia się kolejnych. Trudno jednak jest teraz udzielić odpowiedzi na pytanie o rolę zasobów naturalnych w procesie produkcyjnym, nie jest bowiem jasne, czy są one komplementarne czy substytucyjne z innymi czynnikami. Opinie badaczy nie są jednoznaczne. Z jednej strony dostrzega się komplementarność zasobów naturalnych z kapitałem fizycznym, z drugiej istniejące nowe formy kapitału fizycznego pozwalają zastąpić lub zmniejszyć zużycie różnych rodzajów zasobów naturalnych w procesach produkcyjnych, co może świadczyć o przynajmniej częściowej substytucyjności tych dwóch form kapitału, a z pewnością o substytucyjności różnych rodzajów zasobów naturalnych między sobą. Powstaje pytanie o wpływ, jaki te fakty mają na długookresowy wzrost gospodarczy. Wiadomo bowiem, że wiele produkowanych obecnie dóbr wymaga zużycia nieodnawialnych zasobów naturalnych. Istnieją opinie, iż postęp techniczny rozwiąże wszelkie problemy związane z wyczerpywaniem się zasobów naturalnych.

Celem opracowania jest próba odpowiedzi na pytania o teoretyczne zależności pomiędzy długookresowym wzrostem gospodarczym, postępem technicznym a zużyciem zasobów naturalnych, uwzględniająca postulaty zarówno nowoczesnej teorii wzrostu, jak i ekonomii ekologicznej. Przedmiotem badań jest wpływ, jaki odpowiednia rola zasobów naturalnych (komplementarna lub substytucyjna) wobec pozostałych czynników produkcji oraz ich zużywanie się może mieć na

* Praca jest finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy własny nr N N112 553138.

tempo wzrostu gospodarczego. Dodatkowym zagadnieniem jest rola, jaką w tym wpływie odgrywa postęp techniczny generowany zewnętrznie (tzw. egzogeniczny postęp techniczny). Praca ma charakter teoretyczny.

1. Przegląd literatury

Autorzy licznych prac podkreślają, że problem wyczerpywania się złóż zasobów naturalnych może zostać rozwiązany przez postęp techniczny. Dasgupta i Heal [4] stwierdzają, że wobec zagrożenia zerową produkcją istotna staje się możliwość substytucji w produkcji pomiędzy zasobami naturalnymi a odnawialnymi czynnikami produkcji. Stiglitz [9] wskazuje, że rozwój technologii może na trzy różne sposoby pomóc ludzkości w problemie niedostatku zasobów naturalnych:

1) zwiększenie TFP pozwoli na utrzymanie tego samego poziomu produkcji przy obniżeniu wysokości tych nakładów,
2) nowo powstała technologia może umożliwić substytucję czynników produkcji w większym stopniu niż obecnie,
3) użytkowanie kapitału ludzkiego jako kolejnego czynnika produkcji umożliwi wykorzystanie rosnących korzyści skali.

Solow [8] zauważa, że wyczerpujące się zasoby naturalne można zastąpić przez zwiększenie wielkości odnawialnego kapitału, krytykuje także ideę nieograniczonego postępu technicznego. Sugeruje zatem, że całkowite wyeliminowanie zasobów naturalnych z procesu produkcyjnego może nie być możliwe. Kontynuując ten wątek, Dasgupta [3] stwierdza, że produkcja wszelkich dóbr wymaga składników pochodzących z natury, czy to w formie fizycznej czy też energii do jego wyprodukowania. Zauważa również, że popyt na zasoby naturalne wynika wyłącznie z zapotrzebowania na produkty, które są tworzone przy ich udziale. Constanza i Daly [2] zauważają, że gdyby kapitał fizyczny był doskonałym substytutem kapitału naturalnego, to nigdy by nie powstał. Zamiana bowiem jednej z form kapitału na inną, doskonale do niej substytucyjną, jest nieopłacalna – gdyż przy tej zamianie konieczny jest pewien dodatkowy nakład pracy. Kapitał fizyczny powstał zatem jako komplementarny do kapitału naturalnego, a ich wzajemna substytucyjność musi być niska. Cleveland i Ruth [1] twierdzą, że kapitał fizyczny i naturalny muszą być do pewnego stopnia jednocześnie komplementarne i substytucyjne. W dotychczasowej literaturze jedni autorzy eksponują bardziej substytucyjność tych dwóch form kapitału, inni twierdzą, że są one komplementarne, a o substytucyjności w ogóle nie może być mowy. Do tych pierwszych

można zaliczyć „technologicznych optymistów", którzy twierdzą, iż ze względu na substytucyjność kapitału naturalnego i kapitału fizycznego wystarczająco duża akumulacja tego drugiego, powiązana z rozwojem technologii, rozwiąże problem wyczerpujących się źródeł zasobów naturalnych. „Technologiczni pesymiści", akcentujący przede wszystkim komplementarność kapitału naturalnego i innych form kapitału wskazują, że wraz z wyczerpaniem się złóż zasobów naturalnych nastąpi drastyczne zmniejszenie się produkcji pewnych dóbr. Sugerują oni oszczędzanie zużycia nieodnawialnych źródeł energii, aby ten moment opóźnić. Georgescu-Roegen [5] uważa, że substytucyjność pomiędzy pewnymi grupami kapitału naturalnego jest pewna, lecz teorie głoszące substytucję pomiędzy kapitałem fizycznym a naturalnym są nieprawdziwe. Podaje wiele argumentów, m.in. takie, że większość istniejącego kapitału fizycznego używanego w procesie produkcyjnym wymaga kapitału naturalnego jako źródła energii. Płynie stąd prosty wniosek, że olbrzymi nakład kapitału fizycznego wymaga odpowiednio dużego nakładu energii czerpanej właśnie z kapitału naturalnego.

Jedną z głównych inspiracji niniejszej pracy stanowi artykuł [6]. Autorzy twierdzą, iż postęp techniczny może być jednym z rozwiązań problemu nieodnawialności zasobów naturalnych. Przedstawili teoretyczny model wzrostu gospodarczego. Funkcja produkcji typu CES zawiera trzy różne składniki ulegające efektom postępu technicznego, jednym z nich jest zmiana w czasie wielkości elastyczności substytucji pomiędzy odnawialnym i nieodnawialnym zasobem. Gdy elastyczność tej substytucji dąży do zera oba zasoby są doskonale komplementarne. Zwiększanie się stopnia substytucyjności pomiędzy dwoma czynnikami produkcji, będące efektem postępu technicznego, stwarza możliwości substytucji obu czynników, w granicy czyniąc te czynniki doskonale substytucyjnymi. Autorzy obserwują zmiany zachodzące w modelowanej gospodarce pod wpływem zmieniających się egzogenicznie w czasie parametrów odzwierciedlających postęp techniczny, prowadzący do zmiany komplementarności zasobów odnawialnych i nieodnawialnych na ich doskonałą substytucyjność.

2. Model

Podstawą konstrukcji modelu jest praca [6]. Rozważamy gospodarkę zamkniętą, bez widocznego udziału państwa. Gospodarka ta zawiera L nieskończenie długo żyjących jednostek. Dla uproszczenia zakładamy, że $L = 1$ oraz że ilość jednostek jest stała w czasie. Gospodarstwa domowe składające się z tych jed-

nostek podejmują racjonalne decyzje w warunkach konkurencyjnej gospodarki, maksymalizując swoją łączną użyteczność konsumpcji:

$$U(C) = \int_0^{+\infty} \frac{C^{1-\gamma}}{1-\gamma} e^{-\rho t} dt$$

gdzie:
U – wielkość użyteczności płynąca ze zrealizowanej konsumpcji C,
ρ – współczynnik dyskonta konsumpcji, informujący o tym, jaka jest relacja użyteczności płynącej z konsumpcji w okresie bieżącym do użyteczności konsumpcji realizowanej w okresie przyszłym $\rho \in (0,1)$,
parametr γ – skłonność gospodarstw domowych do wygładzania konsumpcji, $\gamma \geq 0$.

Złoża nieodnawialnych zasobów naturalnych występują w modelowanej gospodarce w ilości S_{N0}, a ich zmniejszanie się postępuje w tempie zgodnym z wielkością ich wydobycia:

$$S_N = -R_N$$

gdzie:
S_N – wielkość złóż zasobów nieodnawialnych,
R_N – wielkość ich zużycia, używana w procesie produkcyjnym,

$$S_{N0} \geq \int_0^{+\infty} R_N dt.$$

Równocześnie w modelowanej gospodarce występuje także zasób S_{O0} odnawialnych zasobów, które mogą być utożsamiane z odnawialnymi zasobami naturalnymi lub z zasobami kapitału fizycznego. Zmniejszanie się zasobów odnawialnych postępuje w tempie zgodnym z wielkością ich spożytkowania. Odnawiają się one w tempie μ:

$$S_O = -R_O + \mu S_O$$

gdzie:
S_O – wielkość zasobów odnawialnych,
R_O – wielkość jego zużycia wykorzystywana w procesie produkcyjnym. Wymagamy, aby $\forall_t S_O \geq 0$.

Wielkość produkcji dana jest przez funkcję produkcji:

$$Y = AR^{\beta} L^{1-\beta}$$

gdzie:
R – łączna wielkość zasobów zużywanych w procesie produkcyjnym,
β – elastyczność produkcji względem wielkości R,
A – pewny stały współczynnik mający interpretację TFP.

Cały postęp techniczny ma źródło zewnętrzne, dopuszczamy zatem wzrost A w czasie, w stałym tempie g:

$$A = gA$$

Łączna wielkość zasobów zużywanych w procesie produkcyjnym R składa się po części z zasobów odnawialnych i nieodnawialnych. Agregacja tych dwóch różnych form w jedną ma postać funkcji CES:

$$R = [\alpha(aR_N)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(bR_O)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}$$

gdzie parametry a i b stanowią o produktywności poszczególnych form zasobów naturalnych, parametr $\alpha \in (0,1)$ waży udział R_N i R_O w zasobie R, σ zaś jest elastycznością substytucji pomiędzy zasobami. W przypadku, gdy $\sigma = 0$ mamy do czynienia z doskonałą komplementarnością obu form zasobów. Powyższa agregacja jest wtedy postaci funkcji Leontiewa, $R = \min\{a R_N, b R_O\}$. Jeżeli $\sigma \in (0,1)$ oba rodzaje zasobów są komplementarne, aczkolwiek w różnym stopniu. Gdy $\sigma = 1$ agregacja staje się agregacją typu Cobba-Douglasa postaci $R = (a R_N)^{\alpha} (b R_O)^{1-\alpha}$. Jeżeli parametr elastyczności substytucji obu form zasobów jest większy od jedności, oba rodzaje zasobów są wobec siebie wzajemnie substytucyjne, jej stopień rośnie wraz z σ. Gdy σ dąży do nieskończoności oba rodzaje zasobów stają się doskonale substytucyjne, a ich agregacja jest postaci $R = \alpha a RN + (1-\alpha) b R_O$.

Zmiany technologiczne objawiają się zatem przez wzrost σ, a, b, μ oraz zmniejszanie się α. Ponieważ abstrahujemy od inwestycji w kapitał fizyczny, cała wielkość produkcji przeznaczana jest na konsumpcję:

$$Y = C$$

Zmienne R_N i R_O są zatem w naszym problemie zmiennymi decyzyjnymi, a zmienne S_N i S_O zmiennymi stanu.

3. Rozwiązanie modelu i analiza wyników

Hamiltonian dany jest wzorem:

$$H(Y, R_N, R_O, S_N, S_O, \lambda_1, \lambda_2) = \frac{Y^{1-\gamma}}{1-\gamma} e^{-\rho\tau} + \lambda_1(-R_N) + \lambda_2(-R_O + \mu S_O)$$

Oznaczmy przez *gx* stopę wzrostu w czasie zmiennej *X*. Uzyskany z warunków pierwszego rzędu układ równań, wraz z uprzednio omówionymi równaniami ruchu dla zmiennych stanu (czyli dla S_N i S_O), daje następujące rozwiązanie*:

$$g_{R_N} = \frac{1}{\beta}\left(\frac{1}{(1-\beta)+\gamma\beta} - 1\right) g + \frac{\rho}{(\beta-1)-\gamma\beta} - (1-\alpha)\sigma\mu \left(\frac{bR_O}{\left(\frac{Y}{A}\right)^{\frac{1}{\beta}}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{1}{(\sigma(\beta-1)-\gamma\beta\sigma)} + 1\right)$$

$$g_{R_O} = \frac{1}{\beta}\left(\frac{1}{(1-\beta)+\gamma\beta} - 1\right) g + \frac{\rho}{(\beta-1)-\gamma\beta} - (1-\alpha)\sigma\mu \left(\frac{bR_O}{\left(\frac{Y}{A}\right)^{\frac{1}{\beta}}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{1}{(\sigma(\beta-1)-\gamma\beta\sigma)} + 1\right) + \sigma\mu$$

$$g_Y = g + \frac{\beta\left(\alpha\,(aR_N)^{\frac{\sigma-1}{\sigma}} g_{R_N} + (1-\alpha)(bR_O)^{\frac{\sigma-1}{\sigma}} g_{R_O}\right)}{\alpha\,(aR_N)^{\frac{\sigma-1}{\sigma}} g_{R_N} + (1-\alpha)(bR_O)^{\frac{\sigma-1}{\sigma}}}$$

$$g_Y = \frac{1}{(1-\beta)+\gamma\beta} g + \frac{-\beta\rho}{(1-\beta)+\gamma\beta} + \frac{\beta(1-\alpha)\mu}{\left((1-\beta)+\gamma\beta\right)} \left(\frac{bR_O}{\left(\frac{Y}{A}\right)^{\frac{1}{\beta}}}\right)^{\frac{\sigma-1}{\sigma}}$$

Zauważmy, że stopy wzrostu poszczególnych zmiennych zależą nie tylko od parametrów, ale też od wielkości zasobów. Stopy te zmieniają się zatem w czasie wraz ze zmianą udziału zasobów odnawialnych w łącznej wielkości zasobów używanych w procesie produkcyjnym. Dla gospodarki w stanie równowagi możemy dokonać analiz zmian tych stóp wzrostu w przypadku zmodyfikowania wielkości poszczególnych parametrów, obserwując tym samym różnice w stopach wzrostu, jakie mogą wystąpić pomiędzy dwoma identycznymi gospodarkami, różniącymi się tylko jednym aspektem. Ponieważ związki pomiędzy tymi parametrami a obliczonymi wartościami stóp wzrostu są silnie nieliniowe, znaki odpowiednich pochodnych będą podane dla konkretnych wartości pozostałych parametrów. Załóżmy, że: $A = R_O = R_N = a = b = 1$, $\gamma = 10$, $\rho = 0{,}04$, $\alpha = \beta = 0{,}5$, $\mu = 0{,}01$, $g = 0{,}03$, $\sigma \in \{0{,}5;\ 1;\ 2;\ 10\}$. Tabele 1 i 2 prezentują znaki kolejnych pochodnych po wybranych parametrach.

* Uzyskanie tego rozwiązania wymaga dość skomplikowanych obliczeń, które dla wygody w niniejszej pracy pomijamy. Obliczenia są jednak dostępne u autora pracy na życzenie.

Tabela 1

Znaki odpowiednich pochodnych

	$x = \sigma$	$x = A$	$x = \alpha$	$x = b$
$\frac{\partial g_{R_N}}{\partial_x}$	<0	<0	$\begin{cases} <0 & \sigma < 1 \\ =0 & \sigma = 1 \\ >0 & \sigma > 1 \end{cases}$	$\begin{cases} >0 & \sigma < 1 \\ =0 & \sigma = 1 \\ <0 & \sigma > 1 \end{cases}$
$\frac{\partial g_{R_O}}{\partial_x}$	>0	<0	$\begin{cases} <0 & \sigma < 1 \\ =0 & \sigma = 1 \\ >0 & \sigma > 1 \end{cases}$	$\begin{cases} >0 & \sigma < 1 \\ =0 & \sigma = 1 \\ <0 & \sigma > 1 \end{cases}$
$\frac{\partial g_Y}{\partial_x}$	=0	=0	$\begin{cases} <0 & \sigma < 1 \\ =0 & \sigma = 1 \\ >0 & \sigma > 1 \end{cases}$	$\begin{cases} <0 & \sigma < 1 \\ =0 & \sigma = 1 \\ >0 & \sigma > 1 \end{cases}$

Tabela 2

Znaki odpowiednich pochodnych

	$x = \mu$	$x = \alpha$	$x = \beta$	$x = g$
$\frac{\partial g_{R_N}}{\partial_x}$	<0	>0	>0	<0
$\frac{\partial g_{R_O}}{\partial_x}$	<0	>0	>0	<0
$\frac{\partial g_Y}{\partial_x}$	<0	<0	<0	>0

Z analizy wpływu zmian wartości poszczególnych parametrów na stopy wzrostu głównych zmiennych modelowanej gospodarki można wysnuć kilka wniosków:

– Zwiększenie stopnia komplementarności/substytucyjności obu rodzajów zasobów nie ma wpływu na stopę wzrostu gospodarczego. Zmiana tego parametru jest głównym, pożądanym tutaj efektem postępu technicznego. Zmniejszanie się stopnia komplementarności w kierunku substytucyjności, a następnie zwiększanie się stopnia substytucyjności tych dwóch czynników produkcji powoduje jedynie wymianę wykorzystania zasobów nieodnawialnych na wykorzystywanie zasobów odnawialnych. Efekt ten jest zgodny z teorią.
– Zwiększenie się TFP (A) nie ma bezpośredniego wpływu na stopę wzrostu gospodarczego. Jeżeli mamy zatem dwie gospodarki, mające identyczne parametry mikro- i makroekonomiczne, z których jedna ma wyższą łączną produktywność czynników produkcji, to obie będą się rozwijać w tym samym tempie. Jedyną korzyścią z posiadania wyższego A jest zużywanie mniejszych ilości obu form zasobów, osiągając tym samym taki sam produkt, jak gospo-

darka o niższym poziomie rozwoju technologii, wykorzystująca proporcjonalnie większe ilości zasobów.

– Zmiana produktywności obu form zasobów wywołuje różne efekty na stopach wzrostu podstawowych zmiennych, zależne od stopnia komplementarności obu czynników produkcji. Na przykład wzrost produktywności zasobu nieodnawialnego (a) wywołuje spadek stopy zużycia tego zasobu, jeśli obie formy zasobów są wzajemnie komplementarne oraz wzrost stopy zużycia, jeśli te czynniki produkcji są substytucyjne. Identyczny efekt wywołany jest w zużyciu zasobu odnawialnego, odwrotny w stopie wzrostu gospodarczego. Jest to spowodowane faktem, iż jeżeli np. zasoby są komplementarne, to podniesienie produktywności jednego z nich daje możliwość zmniejszenia jego wykorzystywania. Zasób ów jest bowiem w procesie produkcyjnym niezbędny, a zatem należy obniżyć jego zużycie, możliwie go oszczędzając. Komplementarność drugiego zasobu także zmusza do stosownego dopasowania używanej wielkości zasobu. Odwrotny efekt występuje, gdy zasoby są substytucyjne.
– W przypadku uzyskania wyższej wartości parametru μ dostajemy obniżenie stopy zużycia obu form zasobów oraz stopy wzrostu gospodarczego. Szybsze odnawianie się zasobu odnawialnego pozwala na dokonywanie jego wyższego zużycia dziś, co umożliwia zmniejszenie tempa zmian jego zużycia w przyszłości. Wiąże się to z niższym tempem wzrostu gospodarczego.
– Przesunięcie udziałów obu form zasobów w stronę zasobów odnawialnych obniża stopy ich zużycia. Zauważmy, że zmiana tego parametru nie zmienia nakładu R_N ani R_O, zwiększa jedynie udział jednego z nich w agregacji. Obniżenie roli zasobu nieodnawialnego podwyższa także stopę wzrostu gospodarczego.
– Zwiększenie udziału zagregowanych zasobów w łącznej produkcji powoduje podniesienie stóp ich zużycia oraz obniżenie stopy wzrostu produkcji. Jest to spowodowane oparciem gospodarki na czynniku produkcji, jakim są zasoby naturalne, m.in. na zasobach nieodnawialnych, których oszczędzanie jest wskazane.
– Zwiększanie się stopy egzogenicznego postępu technicznego g obniża stopy zużycia obu form zasobów oraz podnosi stopę wzrostu gospodarczego. Jest to zgodne z klasycznymi wynikami – głównym mechanizmem wzrostu okazuje się postęp techniczny. Stopa jego wzrostu powoduje obniżanie zużycia obu form zasobów w czasie, w szczególności zasobów nieodnawialnych.

Podsumowanie

Skonstruowany model, będący rozszerzoną wersją modelu z pracy [6], uwzględnia proces zużywania się i odnawiania jednej z form zasobów. W stosunku do oryginalnego modelu wprowadzenie kolejnego parametru pozwala na uzyskanie dalszych wyników uogólniających dotychczas uzyskane, aczkolwiek należy wprost przyznać, iż analiza rozszerzonego modelu nie jest jeszcze zakończona, ani nie jest tak bogata, jak ta przeprowadzona w modelu oryginalnym. W dalszym ciągu wymagana jest pogłębiona refleksja nad wynikami. Tym niemniej należy zauważyć, że jakościowo wyniki są podobne. Postęp techniczny, odzwierciedlający się w modelu na kilka sposobów, za każdym razem prowadzi do obniżenia zużycia obu form zasobów. Innowacje technologiczne zatem, nie tylko nakierowane bezpośrednio na zmianę roli zasobów naturalnych nieodnawialnych w procesie produkcyjnym, ale także na zwykłe podniesienie efektywności wykorzystania wszelkich czynników produkcji, będą prowadzić do zmniejszenia zużycia wyczerpujących się zasobów. Może to oznaczać, że skonstruowany model ogólnie ma przewidywania technologicznie optymistyczne. W dalszym ciągu jednak wiele nierealistycznych szczegółów wymaga dopracowania. Zauważmy bowiem, że wprowadzenie ujemnej stopy zmian wydobycia i wykorzystania zasobów w procesie produkcyjnym powoduje wykładniczy spadek R_N do zera w nieskończonym czasie. To bezpośrednio prowadzi do sytuacji, gdy w pewnych, dalszych momentach czasu wydobywane są śladowe ilości tego zasobu, np. jedna kropla ropy naftowej, a nawet jej część. Oznacza to, że ropa naftowa skończy się raczej wcześniej niż w nieskończonym czasie, co przewidują inne prace, np. [7]. Wydaje się także, iż należałoby w modelach tego typu wprowadzić zagadnienie akumulacji kapitału fizycznego, a nawet inwestycje w odnawianie zasobu odnawialnego.

Literatura

1. Cleveland C.J., Ruth M., When, where and by how much do biophysical limits constrain the economic process? A survey of Nicholas Georgescu-Roegen's contribution to ecological economics, "Ecological Economics" 1997, No. 22.
2. Constanza R., Daly H.E., Natural Capital and Sustainable Development, "Conservation Biology" 1992, Vol. 6, No. 1.
3. Dasgupta P., Natural Resources in the age of substitutability, in: Handbook of Natural Resource and Energy Economics, eds. A.V. Kneese, J.L. Sweeney, ESP B.V. 1993.

4. Dasgupta P., Heal G., The Optimal Depletion of Exhaustible Resources, "Review of Economic Studies" 1974, Symposium volume.
5. Georgescu-Roegen, The Entropy Law and the Economic Process, Harvard University Press, Cambridge, MA, 1971.
6. Growiec J., Schumacher I., On technical change in the elasticities of resource inputs, "Resources Policy" 2008, No. 33.
7. Lin C., Meng H., Ngai T., Oscherov V., Zhu Y., Hotelling revisited: Oil prices and Endogenous Technological Progress, "Natural Resources Research" 2009, Vol. 18, No. 1.
8. Solow R., Intergenerational Equity and Exhaustible Resources, "Review of Economic Studies" 1974, Symposium volume.
9. Stiglitz J., Growth with Exhaustible Resources: Efficient and Optimal Growth Paths, "Review of Economic Studies" 1974, Symposium volume.

TECHNOLOGICAL PROGRESS AND NATURAL RESOURCES IN PRODUCTION PROCESS

Summary

The ongoing technological progress creates many new possibilities for the use of natural resources. Economic theory does not give a straight answer to questions on the role of resources in the production process – it is unclear if natural resources are complementary or substitutive with other factors. A natural question on the impact of these facts on long-term economic growth arise. There are opinions that technological progress will provide solution to all the problems associated with the depletion of the resources.

The purpose of this paper is to attempt to answer questions about the theoretical relationship between long term economic growth, technological progress and the consumption of natural resources, taking into account the modern theory of growth and ecological economics. The main subject of research is the impact that role of natural resources (complementary or substitution) to the other factors of production can have on growth. An additional question is about role of exogenous technological progress.

Krzysztof Piasecki
Uniwersytet Ekonomiczny w Poznaniu

O ISTOCIE WARTOŚCI BIEŻĄCEJ

Wprowadzenie

Fundamentalnym założeniem arytmetyki finansowej jest pewnik, że wartość pieniądza rośnie wraz z upływem czasu, po jakim będzie on spożytkowany. Założenie to jest uzasadniane na ogół poprzez analizę równania wymiany pieniądza* zaproponowanego przez Irvinga Fishera. W analizie tej korzysta się z dodatkowego założenia o stałej ilości pieniądza. Jest to typowo normatywne założenie i z tego względu rozpatrywaną w arytmetyce finansowej wartość pieniądza będziemy nazywać wartością normatywną pieniądza. Proces przyrostu wartości normatywnej nazywamy procesem aprecjacji kapitału. Wzrost względnej aprecjacji wywołany przez przyrost wartości nominalnej pieniądza nazywamy efektem synergii kapitału. Jeśli względna aprecjacja przebiega w sposób niezależny od wartości nominalnej pieniądza, to stwierdzamy niewystępowanie efektu synergii. Z drugiej strony na ogół stosowana praktyka gospodarczo-finansowa powoduje przyrost ilości pieniądza szybszy od przyrostu wolumenu produkcji. Obserwujemy wtedy spadek wartości realnej pieniądza. Oznacza to, że wartości normatywnej pieniądza nie możemy identyfikować z jego wartością realną. Rodzi to pytanie o istotę pojęcia „wartości normatywnej". Konsekwencją tego pytania jest kolejne pytanie o istotę podstawowych funkcji arytmetyki finansowej.

W ostatnich latach w arytmetyce finansowej wyodrębnił się nurt badawczy eksponujący rolę odgrywaną przez pojęcie „użyteczności strumienia finansowego". Do tego nurtu należą m.in. prace [5; 6; 7; 8; 9; 10; 18]. Stosując to podejście możemy przedstawić pojęcie „wartości normatywnej" w kontekście funkcji użyteczności. Takie podejście rzuca nowe światło na podstawowe zmienne arytmetyki finansowej.

Celem rozważań prezentowanych w tym opracowaniu jest objaśnienie na gruncie teorii użyteczności pojęcia „wartości bieżącej". Cel ten zostanie osiągnię-

* Taka analiza została opisana np. w [15].

ty poprzez budowę modelu formalnego o możliwie niskim stopniu złożoności logicznej. Dzięki temu będzie można zdefiniować wartość bieżącą jako funkcję spełniającą pewien prosty układ aksjomatów. Uzyskana tą drogą definicja zostanie porównana z aksjomatyczną definicją wartości bieżącej przedstawioną w [11]. W celu pokazania przydatności budowanej teorii dyskutowany tutaj też będzie efekt synergii kapitału.

1. Uporządkowana przestrzeń strumieni finansowych

Niech będzie dany zbiór momentów czasowych $\Theta \subseteq [0, +\infty[$. W szczególnym przypadku może to być zbiór momentów kapitalizacji lub nieujemna półprosta czasu.

W analizie rynków finansowych każda z płatności jest reprezentowana przez instrument finansowy opisany jako strumień finansowy (t, C), gdzie symbol $t \in \theta$ oznacza moment przepływu strumienia, natomiast symbol $C \in \mathbf{R}$ opisuje wartość nominalną tego przepływu. Każdy z tych strumieni finansowych może być realizowaną należnością lub wymaganym zobowiązaniem. Wartość nominalna każdej należności jest nieujemna. Zobowiązania obciążające dłużnika stanowią zawsze należność wierzyciela. W tej sytuacji wartość zobowiązania jest równa wziętej ze znakiem minus wartości należności odpowiadającej temu zobowiązaniu.

W pierwszym kroku nasze rozważania ograniczymy do zbioru $\Phi^+ = \Theta \times [0, +\infty[$ wszystkich należności (t, C). Na zbiorze tych należności każdy z inwestorów określa swoje preferencje. Preferencje te mają pewne wspólne cechy.

Referując podstawy teorii kapitału, de Soto [16] przedstawił regułę preferencji czasowej. Reguła ta głosi, że przy uwzględnieniu zasady ceteris paribus podmiot ekonomiczny woli zaspokoić swoje potrzeby bądź osiągać postawione cele możliwie jak najszybciej. Inaczej mówiąc, kiedy podmiot ma przed sobą dwa cele o subiektywnie jednakowej wartości, to wyżej sobie ceni ten, który może osiągnąć w krótszym czasie. W szczególnym przypadku oznacza to, że inwestor porównując dwie wpłaty o równej wartości nominalnej zawsze preferuje wpłatę szybciej dostępną. Relację tę opisujemy za pomocą preporządku $\succcurlyeq_T$ zdefiniowanego następująco:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi^+ : (t_1, C) \succcurlyeq_T (t_2, C) \Leftrightarrow t_1 \leq t_2 \qquad (1)$$

Z drugiej strony jest oczywiste, że każdy podmiot ekonomiczny w swoim działaniu kieruje się regułą preferencji majątkowej. Reguła ta oznacza, że przy uwzględnieniu zasady ceteris paribus podmiot ekonomiczny woli wchodzić we

władanie możliwie jak najbardziej wartościowych przedmiotów ekonomicznych, czyli kiedy ma przed sobą dwa przedmioty ekonomiczne równocześnie dostępne, to wybiera ten, który charakteryzuje się większą subiektywną wartością. W szczególnym przypadku oznacza to, że inwestor porównując dwie równocześnie dostępne wpłaty zawsze wybiera wpłatę o wyższej wartości. Relację tę opisujemy za pomocą preporządku $\succcurlyeq_C$ zdefiniowanego następująco:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^+ : (t, C_1) \succcurlyeq_C (t, C_2) \Leftrightarrow C_1 \geq C_2 \tag{2}$$

Równoczesne uwzględnienie obu tych preporządków prowadzi do ostatecznego określenia relacji preferencji $\succcurlyeq$ na zbiorze Φ^+ należności, jako porównania wielokryterialnego:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^+ : (t_1, C_1) \succcurlyeq (t_2, C_2) \Leftrightarrow t_1 \leq t_2 \wedge C_1 \geq C_2 \tag{3}$$

Istnieje wtedy funkcja użyteczności U:$\Phi^+ \rightarrow [0, +\infty[$ spełniająca warunek:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^+ : (t_1, C_1) \succcurlyeq (t_2, C_2) \Rightarrow U(t_1, C_1) \geq U(t_2, C_2) \tag{4}$$

W kolejnym kroku nasze rozważania poświęcimy zbiorowi $\Phi^- = \Theta \times]-\infty, 0]$ wszystkich zobowiązań (t, C). Należność zawsze stanowi kapitał wierzyciela. Zobowiązanie dłużnika powstaje na skutek udostępnienia tego kapitału dłużnikowi przez wierzyciela. Zależność ta powoduje, że korzyści osiągane przez wierzyciela stanowią koszt dłużnika. Oznacza to m.in., że relacja preferencji określona na zbiorze zobowiązań jest relacją odwrotną do relacji preferencji określonej na zbiorze należności. W tej sytuacji relacja relacji preferencji $\succcurlyeq$ na zbiorze Φ^- zobowiązań jest określona za pomocą równoważności:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^- : (t_1, C_1) \succcurlyeq (t_2, C_2) \Leftrightarrow (t_2, -C_2) \succcurlyeq (t_1, -C_1) \tag{5}$$

Porównanie zależności (3) i (5) prowadzi do ostatecznego określenia relacji preferencji $\succcurlyeq$ na zbiorze Φ^- zobowiązań jako porównania wielokryterialnego:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^- : (t_1, C_1) \succcurlyeq (t_2, C_2) \Leftrightarrow t_1 \geq t_2 \wedge C_1 \geq C_2 \tag{6}$$

Istnieje wtedy funkcja określająca użyteczność poszczególnych zobowiązań. Z finansowego punktu widzenia każda należność jest bardziej użyteczna niż dowolne zobowiązanie, co zapisujemy:

$$\forall\big((t_1, C_1), (t_2, C_2)\big) \in \Phi^+ \times \Phi^- : U(t_1, C_1) \geq U(t_2, C_2) \tag{7}$$

Spełnienie tego warunku możemy uzyskać poprzez przyjęcie założenia, że użyteczność dowolnego zobowiązania jest liczbą nieujemną*. Możemy zatem stwierdzić, że istnieje funkcja użyteczności U;$\Phi^- \rightarrow]-\infty, 0]$ spełniająca warunek:

$$\forall(t_1, C_1), (t_2, C_2) \in \Phi^- : (t_1, C_1) \succcurlyeq (t_2, C_2) \Rightarrow U(t_1, C_1) \geq U(t_2, C_2) \tag{8}$$

* Pojęcie „ujemnej użyteczności" zostało zaproponowane i dyskutowane w pracach [2; 3; 15].

Podsumowując dotychczasowe rozważania możemy stwierdzić, że na zbiorze $\Phi = \Theta \times \mathbf{R}$ wszystkich strumieni finansowych została określona relacja preferencji $\succcurlyeq$. Relacja ta jest określona za pomocą alternatywy porównań wielokryterialnych w następujący sposób:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1) \succcurlyeq (t_2, C_2) \Leftrightarrow$$
$$\Leftrightarrow (t_1 \geq t_2 \wedge 0 \geq C_1 \geq C_2) \vee (t_1 \leq t_2 \wedge C_1 \geq C_2 \geq 0) \tag{9}$$

Preporządek ten nie jest liniowy. Istnieje tutaj funkcja użyteczności $U\colon \Phi \rightarrow \mathbf{R}$ spełniająca warunek:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1) \succcurlyeq (t_2, C_2) \Rightarrow U(t_1, C_1) \geq U(t_2, C_2) \tag{10}$$

Określona w ten sposób funkcja użyteczności może mieć subiektywny charakter [4].

2. Użyteczność strumienia finansowego

Zbadajmy teraz właściwości wyznaczonej powyżej funkcji użyteczności. Porównanie dziedzin funkcji użyteczności opisanej w (4) i (7) pozwala zapisać:

$$\forall t \in \Theta\colon \quad U(t, 0) = 0 \tag{11}$$

Preporządek $\succcurlyeq$ określa na zbiorze Φ ostry porządek $\succ$ zdefiniowany za pomocą równoważności:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1) \succ (t_2, C_2) \Leftrightarrow$$
$$\Leftrightarrow (t_1, C_1) \succcurlyeq (t_2, C_2) \wedge \sim \big((t_2, C_2) \succcurlyeq (t_1, C_1)\big) \tag{12}$$

Ostry porządek jest określony przez alternatywę porównań wielokryterialnych:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1) \succcurlyeq (t_2, C_2) \Leftrightarrow$$
$$\Leftrightarrow (t_1 \geq t_2 \wedge 0 \geq C_1 > C_2) \vee (t_1 > t_2 \wedge 0 \geq C_1 \geq C_2) \vee$$
$$\vee (t_1 \leq t_2 \wedge C_1 > C_2 \geq 0) \vee (t_1 < t_2 \wedge C_1 \geq C_2 \geq 0) \tag{13}$$

Z drugiej strony mamy tutaj:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\ (t_1, C_1) \succ (t_2, C_2) \Rightarrow U(t_1, C_1) > U(t_2, C_2) \tag{14}$$

Porównując (12), (13) i (14) otrzymujemy:

$$\forall t \in \Theta\ \forall C_1, C_2 \in \mathbb{R}\colon\ C_1 > C_2 \geq 0 \Rightarrow U(t, C_1) > U(t, C_2) \geq 0 \tag{15}$$

$$\forall t \in \Theta\ \forall C_1, C_2 \in \mathbb{R}\colon\ 0 \geq C_1 > C_2 \Rightarrow 0 \geq U(t, C_1) > U(t, C_2) \tag{16}$$

Wszystko to razem oznacza, że użyteczność jest funkcją rosnącą wartości nominalnej przepływu, co zapisujemy:

$$\forall t \in \Theta\ \forall C_1, C_2 \in \mathbb{R}\colon\ \ C_1 > C_2 \Rightarrow U(t, C_1) > U(t, C_2) \tag{17}$$

Następnie porównując (13) i (14) dostajemy:

$$\forall C > 0\ \forall t_1, t_2 \in \Theta\colon t_1 < t_2 \Rightarrow U(t_1, C) > U(t_2, C) \tag{18}$$

$$\forall C < 0\ \forall t_1, t_2 \in \Theta\colon\ t_1 < t_2 \Rightarrow U(t_1, C) < U(t_2, C) \tag{19}$$

Kwestią umowna jest wyskalowanie wartości funkcji użyteczności. Przyjmujemy tutaj, że użyteczność natychmiastowego przepływu finansowego jest równa wartości nominalnej tego przepływu. Założenie to zapisujemy jako warunek brzegowy:

$$\forall C \in \mathbb{R}:\ \ U(0, C) = C \tag{20}$$

Wszystkie te właściwości funkcji użyteczności zostaną dalej wykorzystane do badania właściwości podstawowych modeli arytmetyki finansowej.

3. Wartość bieżąca

Dla preporządku $\succcurlyeq$ określonego przez równoważność (9) wyznaczamy jego liniowe domknięcie $\sqsupseteq$. Preporządek $\sqsupseteq$ jest określony następująco:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1)\ \sqsupseteq (t_2, C_2) \Leftrightarrow U(t_1, C_1)\ \geq U(t_2, C_2) \tag{21}$$

Powyższy preporządek wyznacza następującą relację $\equiv$ równoważności strumieni finansowych:

$$\forall (t_1, C_1), (t_2, C_2) \in \Phi\colon (t_1, C_1) \equiv\ (t_2, C_2) \Leftrightarrow U(t_1, C_1) = U(t_2, C_2) \tag{22}$$

Jeśli dwa strumienie finansowe są jednakowo użyteczne, to uważamy je za równoważne. O strumieniu finansowym równoważnym do danego mówimy, że jest ekwiwalentem tego ostatniego. Wartość nominalną dowolnego ekwiwalentu danego strumienia finansowego identyfikujemy jako wartość normatywną tego strumienia.

Analiza warunków (17) i (18) prowadzi nas do sformułowania zasady aprecjacji. Zasada ta głosi, że wartość normatywna należności rośnie wraz z czasem, po jakim należność ta będzie płatna. W ten sposób teoria użyteczności potwierdza fundamentalny pewnik arytmetyki finansowej głoszący, że wartość pieniądza rośnie wraz z upływem czasu.

Opisane powyżej pojęcie „wartości normatywnej" może być ujęte w karby modelu formalnego. Niech będzie dany natychmiastowy przepływ finansowy o wartości nominalnej $C \in \mathbf{R}$. Przepływ ten jest jednoznacznie przypisany strumieniowi finansowemu (0, C). W dowolnym momencie $\forall t \in \Theta$ wartość normatywna dyskutowanego przepływu jest równa C_t. Zgodnie z definicją (22), relacji równoważności strumieni i warunkiem brzegowym (20) mamy tutaj tożsamość:

$$C = U(0, C) = U(\mathrm{t}, C_t) \tag{23}$$

Zgodnie z warunkiem (17), dla ustalonego momentu $\forall t \in \Theta$ jednoznacznie wyznaczamy wartość normatywną:

$$C_t = FV(t, C) = U^{-1}(t, C) \tag{24}$$

Określoną w ten sposób funkcję FV: $\Phi \rightarrow [0,+\infty[$ w arytmetyce finansowej nazywamy wartością przyszłą. Funkcję wartości przyszłej można przedstawić za pomocą tożsamości:

$$FV(t, C) = C \cdot \tilde{s}(t, C) \tag{25}$$

gdzie funkcja $\tilde{s}$: $\Phi \longrightarrow [0, +\infty[$ jest nazywana czynnikiem aprecjacji kapitału. Czynnik aprecjacji kapitału jest niemalejącą funkcją czasu spełniającą warunek brzegowy:

$$\tilde{s}(0, C) = 1 \tag{26}$$

Czynnik aprecjacji kapitału opisuje przebieg procesu względnej aprecjacji kapitału. Jeśli ten czynnik jest funkcją rosnącą dodatniej wartości kapitału, to wtedy mamy do czynienia z efektem synergii kapitału.

Głównym przedmiotem naszych dociekań będzie dowolny strumień finansowy (t, C). Dla tego strumienia możemy określić jego ekwiwalent. Wartość nominalną C_0 tego ekwiwalentu nazywamy wartością bieżącą i oznaczamy symbolem $PV(t, C)$. Zgodnie z definicją (22) relacji równoważności strumieni i warunkiem brzegowym (20) mamy tutaj tożsamość:

$$C_0 = PV(t, C) = U(0, C_0) = U(t, C) \tag{27}$$

Wartość bieżąca dowolnego strumienia finansowego jest identyczna z jego użytecznością. Stwierdzenie to w pełni wyjaśnia istotę pojęcia „wartości bieżącej". Z drugiej strony taka identyfikacja wartości bieżącej rodzi pewne problemy formalne, o których będzie mowa w kolejnej części. Teraz naszą uwagę skupimy na formalnych własnościach funkcji PV:$\Phi \rightarrow \mathbf{R}$ określonej przez tożsamość (27). Mamy tutaj:

$$\forall C \in \mathbb{R}: \quad PV(0, C) = C \tag{28}$$

$$\forall t \in \Theta: \quad PV(t, 0) = 0 \tag{29}$$

$$\forall (t_1, C), (t_2, C) \in \Phi^+: \quad t_1 < t_2 \Rightarrow PV(t_1, C) > PV(t_2, C) \tag{30}$$

$$\forall (t_1, C), (t_2, C) \in \Phi^-: \quad t_1 < t_2 \Rightarrow PV(t_1, C) < PV(t_2, C) \tag{31}$$

$$\forall (t, C_1), (t, C_2) \in \Phi: C_1 < C_2 \Rightarrow PV(t, C_1) < PV(t, C_2) \tag{32}$$

Funkcję wartości bieżącej można przedstawić za pomocą tożsamości:

$$PV(t, C) = C \cdot \tilde{v}(t, C) = \frac{C}{\tilde{s}\big(t, FV(t, C)\big)} \tag{33}$$

gdzie czynnik dyskontujący $\tilde{v}: \Phi \to]0; 1]$ jest nierosnącą funkcją czasu spełniającą warunek brzegowy:

$$\tilde{v}(0, C) = 1 \tag{34}$$

Zauważmy, że w [11] wartość bieżąca została zdefiniowana jako dowolna funkcja PV:$\Phi \to \mathbf{R}$ spełniająca warunki (28), (30) i dodatkowo warunek addytywności wartości bieżącej:

$$\forall (t, C_1), (t, C_2) \in \Phi: \; PV(t, C_1 + C_2) = PV(t, C_1) + PV(t, C_2) \tag{35}$$

W tej pracy warunek (35) braku efektu synergii został zastąpiony przez znacznie ogólniejszą koniunkcję warunków (29), (31) i (32). Z drugiej strony w [11] wykazano, że warunek (35) jest warunkiem koniecznym i dostatecznym na to, aby wartość bieżąca PV:$\Phi \to \mathbf{R}$ spełniająca warunki (28) i (30) stanowiła funkcję daną zależnością:

$$PV(t, C) = C \cdot v(t) \tag{36}$$

gdzie czynnik dyskontujący $v: \Theta \to]0; 1]$ jest nierosnącą funkcją czasu spełniającą tożsamość:

$$v(t) = \tilde{v}(t, C) \tag{37}$$

Oznacza to, że warunek (35) addytywności wartości bieżącej jest warunkiem koniecznym i dostatecznym do niewystępowania efektu synergii. Dowolna funkcja wartości przyszłej określona za pomocą tożsamości (27) może nie spełniać warunku (35). Obserwujemy wtedy zjawisko interakcji pomiędzy wartością nominalną i terminem przepływu strumienia finansowego. Stwierdzone do tej pory właściwości funkcji użyteczności U:$\Phi \to \mathbf{R}$ nie pozwalają stwierdzić, czy interakcja ta ma charakter efektu synergii kapitału.

4. Efekt synergii kapitału

Pierwsze prawo Gossena informuje o tym, że krańcowa użyteczność bogactwa maleje. Zbadajmy teraz konsekwencję przyjęcia założenia głoszącego, że funkcja użyteczności U:$\Phi^+ \to [0, +\infty[$ określona przez warunek (4) spełnia warunek malejącej krańcowej użyteczności bogactwa. Dla dowolnej funkcji wartości bieżącej możemy wtedy zapisać:

$$\forall (t, C_1), (t, C_2) \in \Phi^+ \; \forall \alpha \in]0; 1[:$$
$$\alpha \cdot PV(t, C_1) + (1 - \alpha) \cdot PV(t, C_2) < PV(t, \alpha \cdot C_1 + (1 - \alpha) \cdot C_2) \tag{38}$$

W nierówności (38) podstawiamy $C_2 = 0$. Dla dodatniej wartości $C_3 < C_1$ mamy wtedy:

$$\frac{C_3}{C_1}\cdot PV(t,C_1)+\left(1-\frac{C_3}{C_1}\right)\cdot PV(t,0)<PV\left(t,\frac{C_3}{C_1}\cdot C_1+\left(1-\frac{C_3}{C_1}\right)\cdot 0\right)$$

co razem z (29), po prostych przekształceniach prowadzi do:

$$\frac{PV(t,C_1)}{C_1}<\frac{PV(t,C_3)}{C_3}$$

Następnie, korzystając z (33) otrzymujemy:

$$\tilde{s}(t,C_3)<\tilde{s}(t,C_1) \tag{39}$$

Czynnik aprecjacji kapitału okazał się funkcją rosnącą wartości nominalnej. Oznacza to, że pierwsze prawo Gossena jest warunkiem dostatecznym dla ujawnienia się efektu synergii kapitału. Obiektywnie obserwowanym faktem jest też stosowanie w arytmetyce finansowej liniowej funkcji wartości bieżącej danej w postaci (36). Pozwala to na postawienie hipotezy, że dowolną funkcję wartości bieżącej możemy zapisać:

$$\forall(t,C_1),(t,C_2)\in\Phi^+\ \forall\alpha\in[0;1]:$$
$$\alpha\cdot PV(t,C_1)+(1-\alpha)\cdot PV(t,C_2)\le PV(t,\alpha\cdot C_1+(1-\alpha)\cdot C_2) \tag{40}$$

Nierówność (40) opisuje słabą wersję pierwszego prawa Gossena stwierdzającą, że krańcowa użyteczność bogactwa nie rośnie. W celu weryfikacji tej hipotezy zbadajmy konsekwencje zaprzeczenia warunku (40). Istnieją wtedy dodatnie wartości $C_1 < C_2$ i $\alpha \in]0; 1[$ takie, że dla pewnego momentu $t^* \in \Theta$ spełniona jest nierówność:

$$\alpha\cdot PV(t^*,C_1)+(1-\alpha)\cdot PV(t^*,C_2)>PV(t^*,\alpha\cdot C_1+(1-\alpha)\cdot C_2)$$

Oznacza to m.in., że istnieją wartości $\chi_1,\chi_2\in[C_1,C_2]$ takie, że spełnione są warunki:

$$\chi_1<\alpha\cdot C_1+(1-\alpha)\cdot C_2<\chi_2$$
$$\tilde{v}(t^*,\chi_1)<\frac{PV(t^*,C_2)-PV(t^*,C_1)}{C_2-C_1}<\tilde{v}(t^*,\chi_2)$$

Wtedy, korzystając z (33) otrzymujemy:

$$\tilde{s}(t^*,\chi_1)>\tilde{s}(t^*,\chi_2) \tag{41}$$

co pozostaje w sprzeczności z efektem synergii kapitału. Oznacza to, że pierwsze prawo Gossena jest warunkiem koniecznym i dostatecznym dla ujawnienia się efektu synergii kapitału.

Ponadto w praktyce finansów nie jest znany incydent zmniejszania się względnej prędkości aprecjacji równocześnie ze wzrostem wartości nominalnej. Fakt ten oznacza empiryczne zaprzeczenie nierówności (41). W ten sposób została pozytywnie zweryfikowana hipoteza stwierdzająca, że dowolna funkcja wartości przyszłej PV:$\Phi \rightarrow \mathbf{R}$ spełnia warunek (40).

Podsumowanie

Przedstawione w pracy rozważania na temat wzajemnych relacji pomiędzy użytecznością bogactwa a wartością bieżącą wykazują logiczną spójność formalnych modeli ekonomii i finansów.

Wykazanie, że wartość bieżąca danego strumienia finansowego jest identyczna z użytecznością tego strumienia wskazuje na subiektywny charakter pojęcia „wartości bieżącej". W tej sytuacji otrzymujemy podwaliny teoretyczne pod budowę modeli finansów behawioralnych wykorzystujących subiektywne oceny wartości bieżącej. Przykłady takich modeli można znaleźć w pracach [12; 13].

Na marginesie tej pracy warto też dostrzec, że dziedzina badań arytmetyki finansowej w coraz większym stopniu wykracza poza domenę teorii procentu. W tej sytuacji arytmetykę finansową należy traktować jako rozszerzenie opartej na obiektywnych przesłankach teorii procentu. Wobec subiektywnych aspektów podejmowanej problematyki dynamicznej oceny pieniądza jest to rozszerzenie istotne.

Literatura

1. Arrow K.J., Essays in the theory of risk-bearing, Elsevier, Amsterdam 1971.
2. Becker G.S., An Economic Analysis of Fertility, w: Demographic and Economic Change in Developed Countries, New York 1960.
3. Cooper B., Garcia Peñaloza C., Funk P., Status effect and negative utility growth, "The Economic Journal" 2001, No. 111.
4. Dacey R., Zielonka P., A detailed prospect theory explanation of the disposition effect, "Journal of Behavioral Finance" 2005, No. 2/4.
5. Doyle J.R., Survey of time preference, delay discounting models, Working Paper, Cardiff Business School, Cardiff University 2010, http://ssrn.com/abstract=1685861 (12.01.2011).
6. Epper T., Fehr-Duda H., Bruhin A., Uncertainty Breeds Decreasing Impatience: The Role of Risk Preferences in Time Discounting. Working Paper No. 412. Institute for Empirical Research in Economics, University of Zuerich, 2009, http://ssrn.com/abstract=1416007 (15.02.2011).
7. Frederick S., Loewenstein G., O'Donoghue T., Time Discounting and Time Preference: A critical Review, "Journal of Economic Literature" 2002, Vol. 40.
8. Killeen P.R., An additive-utility model of delay discounting, "Psychological Review" 2009, No. 116.

9. Kim B.K., Zauberman G., Perception of Anticipatory Time in Temporal Discounting, "Journal of Neuroscience, Psychology, and Economics" 2009, Vol. 2.
10. Kontek K., Decision Utility Theory: Back to von Neumann, Morgenstern, and Markowitz. Working Paper, 2010, http://ssrn.com/abstract=1718424 (03.01.2011).
11. Piasecki K., Od arytmetyki handlowej do inżynierii finansowej, Poznań 2005.
12. Piasecki K., Behavioural Present Value, "Behavioral & Experimental Finance eJournal" 2011/4, http://hq.ssrn.com/Journals/IssueProof.cfm?abstractid=1729351&journalid=1504395&issue_number=4&volume=3&journal_type=CMBO&function=showissue (22.02.2011).
13. Piasecki K., Simple Return Rate at Imprecision Risk, Capital Markets: Asset Pricing & Valuation eJournal 2011/115, http://hq.ssrn.com/Journals/IssueProof.cfm?abstractid=1885771&journalid=1508951&issue_number=115&volume=3&journal_type=CMBO&function=showissue (10.11.2011).
14. Piasecki K., Ronka-Chmielowiec W., Matematyka finansowa, C.H. Beck, Warszawa 2011.
15. Rabin M., Incorporating Fairness into Game Theory and Economics, "The American Economic Review" 1993, Vol. 83, No. 5.
16. Soto de J.H., Pieniądz, kredyt bankowy i cykle koniunkturalne, Warszawa 2009.
17. Zauberman G., Kyu Kim B., Malkoc S., Bettman J.R., Discounting Time and Time Discounting: Subjective Time Perception and Intertemporal Preferences, "Journal of Marketing Research" 2009, Vol. XLVI.

THE ESSENCE OF PRESENT VALUE

Summary

Theoretical basis for discussion is the natural nonlinear preorder defined on the set of financial flows. Then there exists a utility function consistent with this preorder. Two financial flows are equivalent iff their utilities are equal. Then we can show that the present value of financial flow is equal to its utility. Lack of capital synergy is not a necessary condition for this thesis.

Michał Stachura
Barbara Wodecka
Uniwersytet Jana Kochanowskiego w Kielcach

ALTERNATYWNE WZGLĘDEM UJĘCIA MARKOWITZA PODEJŚCIE DO SZACOWANIA STOPY ZWROTU Z PORTFELA

Wprowadzenie

W analizach portfelowych jednym z najważniejszych rozważanych parametrów jest oczekiwana stopa zwrotu. Szacować ją można – w sensie użytego wzoru – przy użyciu wielu różnych metod. Warto zwrócić uwagę, że po przyjęciu określonej metody możliwe do zastosowania są dwa zgoła odmienne ujęcia. Pierwsze (ujęcie niejednolite – UN) polega na osobnym szacowaniu oczekiwanych stóp zwrotu z pojedynczych walorów (zgodnie z przyjętą metodą), a następnie na swego rodzaju interpolacji – wymagającej użycia kolejnej metody – uzyskanych oszacowań na portfele o dowolnych udziałach procentowych walorów (tak jest np. w klasycznym modelu Markowitza). Drugie (ujęcie jednolite – UJ) – którego prezentacja oraz pokazanie pewnych jego przewag jest zasadniczym celem opracowania – polega na jednolitym traktowaniu wszystkich bez wyjątku portfeli (o różnych udziałach procentowych walorów, w tym także pojedynczych). Skutkiem tego uzyskuje się spójność merytoryczną metodologii szacowania oczekiwanych stóp zwrotu.

Wśród metod szacowania oczekiwanej stopy zwrotu, które są powszechnie akceptowane i stosowane w praktyce należy wymienić tzw. metody historyczne, bazujące na założeniu niezmienności stóp zwrotu w czasie i odwołujące się do ich przeszłych realizacji.

Niech zatem rozważony będzie pewien indeks $\mathbf{Y}$ (np. wartość pojedynczego waloru lub wartość portfela), który w chwili t ($t \in \{1, \ldots, T\}$) przyjął wartość Y_t, zaś jego stopa zwrotu była równa R_t. Wówczas jako oszacowanie oczekiwanej

stopy zwrotu r_E na chwilę $T+1$ można przyjąć: a) średnią arytmetyczną realizacji stopy zwrotu, b) średnią ważoną realizacji stopy zwrotu, c) zrealizowaną stopę zwrotu za cały okres inwestycji w przeliczeniu na okres jednostkowy, które dane są odpowiednio wzorami:

$$r_E = \frac{1}{T} \cdot \sum_{t=1}^{T} R_t \tag{1}$$

$$r_E = \sum_{t=1}^{T} c_t R_t \tag{2}$$

$$r_E = (Y_T / Y_1)^{1/T} - 1 \tag{3}$$

gdzie wagi c_t spełniają warunki: $c_t \geq_0$, $\sum_{t=1}^{T} c_t = 1$.

Aby uszczegółowić rozumienie obu ujęć (UN, UJ) szacowania stopy zwrotu w przypadku portfela, przyjmijmy kolejne oznaczenia. Niech $\mathbf{X}_1, \ldots, \mathbf{X}_n$ oznaczają n walorów składających się na portfel $\mathbf{Z}(\boldsymbol{w})$, gdzie $\boldsymbol{w} = (w_1, \ldots, w_n)$ określa udziały (pod względem wartości) poszczególnych walorów.

Wówczas w sensie UJ, po uprzednim wyborze metody (wzoru) szacowania oczekiwanej stopy zwrotu, stosuje się tę metodę dla każdego wyboru $\boldsymbol{w}$ udziałów walorów w portfelu. W szczególności więc również w odniesieniu do każdej ze składowych portfela użyta jest dokładnie ta sama metoda – dla $\mathbf{X}_i$ mamy bowiem $\boldsymbol{w} = (0, \ldots, 0, 1, 0, \ldots, 0)$.

W sensie UN, po uprzednim wyborze metody (wzoru) szacowania oczekiwanej stopy zwrotu, stosuje się tę metodę tylko względem walorów $\mathbf{X}_i$, a następnie z wykorzystaniem jakiejś metodologii szacuje się (interpoluje) oczekiwaną stopę zwrotu z portfela w oparciu o oszacowania wyznaczone dla walorów. W tym kontekście, w praktyce powszechnie stosowane jest podejście zaczerpnięte z klasycznej markowitzowskiej teorii optymalizacji portfela, tzn. za oczekiwaną stopę zwrotu z portfela przyjmuje się średnią ważoną stóp zwrotu ze wszystkich walorów (wagami są udziały $\boldsymbol{w}$). Właśnie ze względu na tę powszechność porównanie istoty UJ i UN zostanie poczynione w dalszym ciągu opracowania przez stałe odniesienie się w przypadku UN do liniowej, markowitzowskiej interpolacji oszacowań oczekiwanej stopy zwrotu z portfela.

Abstrahując od faktu, czy model Markowitza jest właściwym narzędziem do optymalizacji portfela* należy podkreślić, że bezrefleksyjne przeniesie jedy-

* Wśród wielu prac krytykujących metodę Markowitza można wymienić opracowanie [1], w którym pada nawet stwierdzenie, że „jest ona w praktyce nieużyteczna".

nie jednego elementu tego modelu – czyli metody szacowania oczekiwanej stopy zwrotu z portfela – w inne realia może prowadzić do pewnych przekłamań.

1. Rozbieżność między ujęciami szacowania oczekiwanej stopy zwrotu*

Aby uzmysłowić sobie wagę nakreślonego problemu dokonamy analizy prostego przykładu. Niech dane będą dwa walory $\mathbf{X}_1$ i $\mathbf{X}_2$, które rozważamy w dwóch następujących po sobie momentach $t = 0$ i $t = 1$. Przyjmujemy, że wartości walorów kształtują się następująco: $X_{1,0} = 80$, $X_{1,1} = 90$ oraz $X_{2,0} = 60$, $X_{2,1} = 55$. Stąd zrealizowane stopy zwrotu tych walorów wynoszą odpowiednio $\frac{1}{8}$, $\frac{1}{12}$ i są one przyjęte za oczekiwane stopy zwrotu z walorów. Następnie budujemy portfel złożony z $n_1 = 4$ jednostek waloru $\mathbf{X}_1$ i $n_2 = 6$ jednostek waloru $\mathbf{X}_2$ (tym samym udziały pod względem ilości wynoszą odpowiednio $v_1 = 0{,}4$, a $v_2 = 0{,}6$). Wartość portfela w momentach $t = 0$ i $t = 1$ jest więc odpowiednio równa $Z_0 = 680$ i $Z_1 = 690$. Oznacza to, że zrealizowana (i przyjęta w sensie UJ za oczekiwaną) stopa zwrotu z tego portfela wynosi $\frac{1}{68}$. Z drugiej strony stopa zwrotu w sensie UN z portfela wynosi $\frac{7}{276}$ $\left(= \frac{12}{23} \cdot \frac{1}{8} - \frac{11}{23} \cdot \frac{1}{12}\right)$ i jest około 1,725 razy większa od poprzedniego oszacowania. Widoczne staje się zatem, że ujęcia prowadzą do rozbieżnych oszacowań.

Zastawienie omówionych wartości znajduje się w tab. 1, w której warto zwrócić uwagę na udziały walorów w portfelu pod względem wartości (2 ostatnie kolumny), ponieważ udziały te są różne w momentach $t = 0$ i $t = 1$ (w przeciwieństwie do stałych udziałów ilościowych v_i).

Tabela 1

Zestawienie parametrów portfela

	$t = 0$	$t = 1$	r_E	n_i	v_i	$w_i\,(t = 0)$	$w_i\,(t = 1)$
$\mathbf{X}_1$	80	90	1/8	4	0,4	8/17	12/23
$\mathbf{X}_2$	60	55	–1/12	6	0,6	9/17	11/23
Z	680	690	1/68				

Analogiczne rozważania można przeprowadzić dla dowolnych udziałów (zarówno względem wartości $\boldsymbol{w} = (w, 1 - w)$ na moment $t = 1$, jak i ilości $\boldsymbol{v} = (v, 1 - v)$ walorów w portfelu. W efekcie uzyskuje się funkcyjną zależność wartości oczekiwanej stopy zwrotu z portfela od udziału w – lub odpowiednio udziału v – waloru $\mathbf{X}_1$. Zależności te prezentuje rys. 1, a jak łatwo policzyć są one dane wzorami:

* Wszystkie obliczenia i wykresy prezentowane w opracowaniu zostały wykonane w środowisku obliczeniowym R.

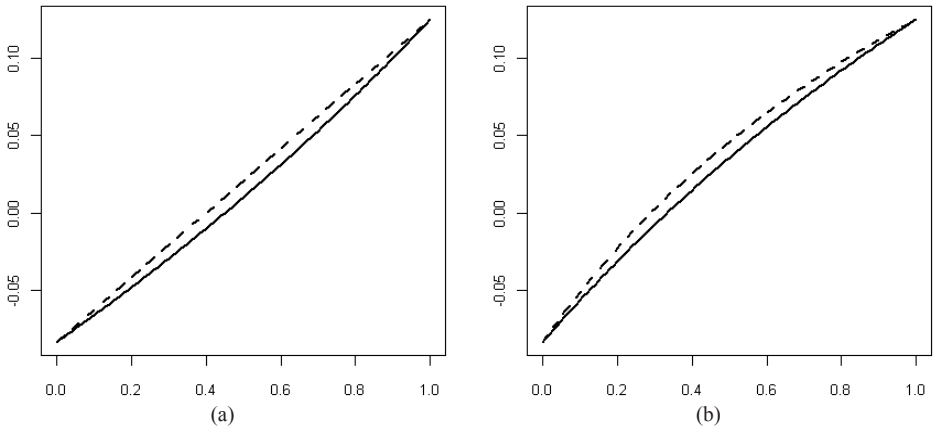
$$r_E^{UN}(w) = \frac{5w-2}{24} \tag{4}$$

$$r_E^{UN}(v) = \frac{38v-11}{12(7v+11)} \tag{5}$$

$$r_E^{UJ}(w) = \frac{20w-9}{4(27-5w)} \tag{6}$$

$$r_E^{UJ}(v) = \frac{3v-1}{4(v+3)} \tag{7}$$

gdzie r_E^{UN}, r_E^{UJ} oznaczają oczekiwane stopy zwrotu odpowiednio w rozumieniu UN i UJ.



Rys. 1. Oszacowania oczekiwanej stopy zwrotu w zależności od udziału waloru pierwszego w portfelu (wykres (a) prezentuje udział w sensie wartości, wykres (b) – udział w sensie ilości; linie ciągłe odpowiadają UJ, a przerywane – UN)

Widać więc, że oba ujęcia są zgodne jedynie dla portfeli złożonych z pojedynczych walorów, a poza tym UN daje zawsze większe oszacowania niż UJ. Warto więc odnotować, że może być też i tak, że oszacowania oczekiwanych stóp zwrotu – w sensie obu ujęć – są różnych znaków.

Wykresy (a) i (b) na rys. 1 nie przedstawiają różnych jakościowo przypadków, lecz stanowią spojrzenia z dwóch różnych perspektyw na to samo zagadnienie. Wynika to z faktu, że przy znanych wartościach walorów w portfelu mamy wzajemnie jednoznaczną odpowiedniość między udziałami pod względem wartości i udziałami pod względem ilości tych walorów w portfelu. Wobec tego wy-

bór typu udziałów jest zależny od potrzeb konkretnych analiz lub przyzwyczajeń badacza*.

W przypadku prezentowanego przykładu zależności między udziałami prezentują funkcje:

$$w = \frac{18v}{7v + 11} \tag{8}$$

$$v = \frac{11w}{18 - 7w} \tag{9}$$

2. Ilustracja empiryczna

Nakreślone na podstawie danych umownych problematyczne kwestie warto zobrazować również na podstawie przykładowych danych empirycznych. W tym celu analizom poddano portfele tworzone z akcji KGHM i PKN Orlen na 30.09.2011 r. z horyzontem czasowym wynoszącym 1 dzień roboczy. Za metodę szacowania oczekiwanej stopy zwrotu przyjęto średnią arytmetyczną z 500 ostatnich realizacji stopy zwrotu zgodnie ze wzorem (1)**. Tak rozumiane oszacowania oczekiwanych stóp zwrotu z walorów są w dalszym ciągu oznaczone jako $r_E^{(1)}$, $r_E^{(2)}$. Z kolei niech n_1 i n_2 oznaczają odpowiednio liczbę akcji KGHM i PKN Orlen w portfelu. Wówczas:

$$v_1 = \frac{n_1}{n_1 + n_2}, \quad v_2 = \frac{n_2}{n_1 + n_2} \quad (v_1 + v_2 = 1) \tag{10}$$

są udziałami akcji w portfelu pod względem ilości, natomiast:

$$w_1 = \frac{n_1 \cdot X_{1,T}}{Z_T}, \quad w_2 = \frac{n_2 \cdot X_{2,T}}{Z_T} \quad (w_1 + w_2 = 1) \tag{11}$$

udziałami pod względem wartości, gdzie $X_{1,T}$, $X_{2,T}$ są notowaniami akcji, a $Z_T = n_1 \cdot X_{1,T} + n_2 \cdot X_{2,T}$ – wartością portfela na dzień $T = 500$.

---

* Z praktycznego punktu widzenia bardziej użyteczne zdają się być udziały w sensie ilości, ponieważ częściej zdarzają się sytuacje, że inwestor w pewnym horyzoncie czasowym utrzymuje stałe udziały w portfelu w sensie ilości niż w sensie wartości. Wówczas, jeśli w pewnym horyzoncie czasowym udziały w sensie ilości są stałe, to udziały w sensie wartości zmieniają się przy każdorazowej zmianie wartości walorów.

** Wobec tego podstawą wszelkich obliczeń są dzienne notowania na zamknięcie kursów akcji obu spółek z okresu 8.10.2010-30.09.2011 (po 500 obserwacji) z GPW w Warszawie.

Przy przyjętych oznaczeniach i założeniach można wyprowadzić następujące wzory na oczekiwaną stopę zwrotu z portfela w rozumieniu UN i UJ w zależności od udziałów walorów:

$$r_E^{UN}(w) = w \cdot r_E^{(1)} + (1-w) \cdot r_E^{(2)} \tag{12}$$

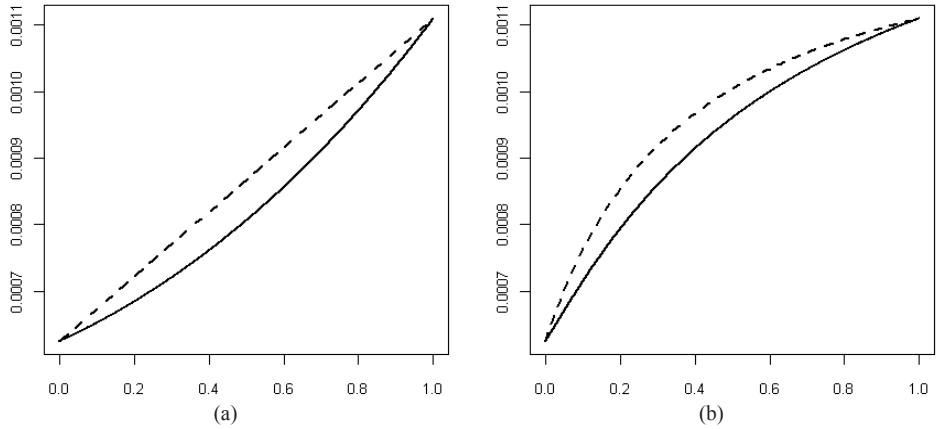
$$r_E^{UJ}(w) = \frac{1}{T} \sum_{t=1}^{T} \frac{wX_{2,T}(X_{1,t} - X_{1,t-1}) + (1-w)X_{1,T}(X_{2,t} - X_{2,t-1})}{wX_{2,T}X_{1,t-1} + (1-w)X_{1,T}X_{2,t-1}} \tag{13}$$

$$r_E^{UN}(v) = \frac{v \cdot r_E^{(1)} \cdot X_{1,T} + (1-v) \cdot r_E^{(2)} \cdot X_{2,T}}{v \cdot X_{1,T} + (1-v) \cdot X_{2,T}} \tag{14}$$

$$r_E^{UJ}(v) = \frac{1}{T} \cdot \sum_{t=1}^{T} \frac{v \cdot (X_{1,t} - X_{1,t-1}) + (1-v) \cdot (X_{2,t} - X_{2,t-1})}{v \cdot X_{1,t-1} + (1-v) \cdot X_{2,t-1}} \tag{15}$$

gdzie:
$w = w_1$, $v = v_1$ – stosownego typu udziały KGHM w portfelu,
$X_{1,T}$, $X_{2,T}$ – notowania akcji w dniu nr t.

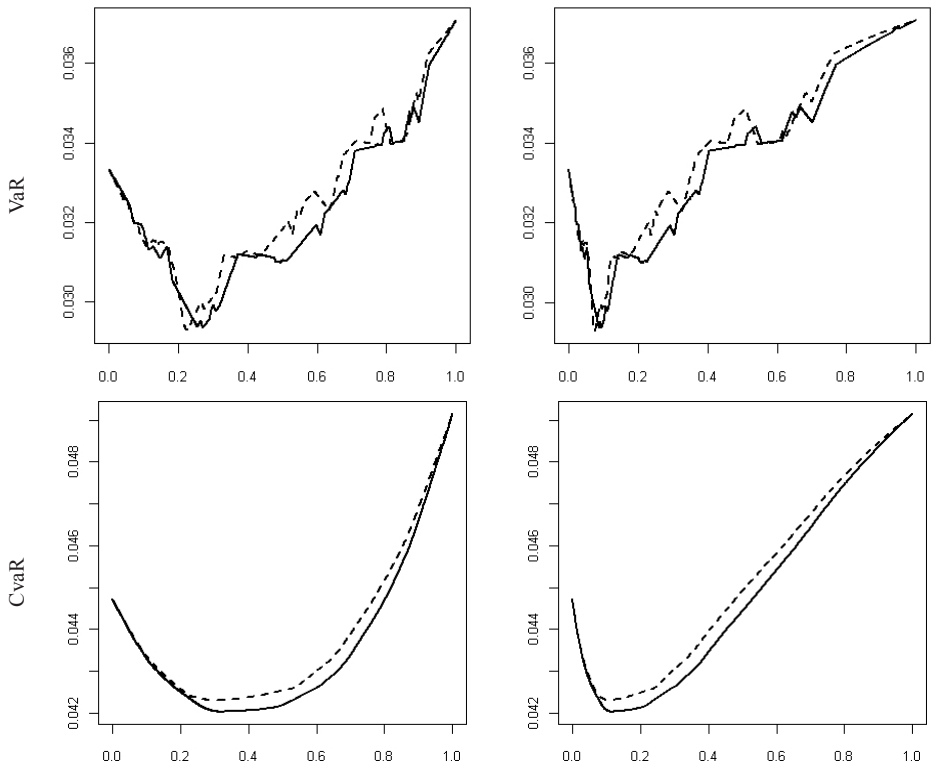
Przedstawione wzory dobitnie pokazują, że oszacowania oczekiwanej stopy zwrotu z portfela są ewidentnie zależne od przyjętego ujęcia, co dobrze uwidacznia rys. 2. Tym razem także okazuje się, że oba ujęcia są zgodne jedynie dla pojedynczych walorów i że poza tym UN daje zawsze większe oszacowania niż UJ.



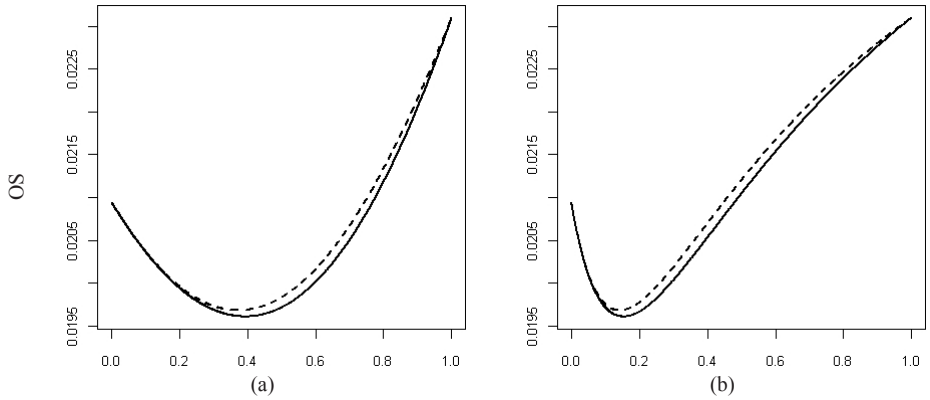
Rys. 2. Oszacowania oczekiwanej stopy zwrotu w zależności od udziału akcji KGHM w portfelu (wykres (a) prezentuje udział w sensie wartości, wykres (b) – udział w sensie ilości; linie ciągłe odpowiadają UJ, a przerywane – UN)

Stwierdzone rozbieżności w oszacowaniach oczekiwanej stopy zwrotu mogą mieć dalsze konsekwencje. Wyobraźmy sobie, że inwestor dysponując takimi oszacowaniami dokonuje optymalizacji portfela posługując się jedną z trzech miar ryzyka: wartością zagrożoną (VaR), warunkową wartością zagrożoną (CVaR) i odchyleniem standardowym (OS)*. Załóżmy ponadto, że wymienione miary dla ustalonego składu portfela szacowane są na podstawie zrealizowanych stóp zwrotu tego portfela, które z kolei odtwarzane są za pomocą takich samych metod i ujęcia, jak szacowana była odpowiadająca im oczekiwana stopa zwrotu.

Wyznaczone w ten sposób empiryczne wartości miar ryzyka, w zależności od udziału pierwszego waloru, zilustrowane są na rys. 3.



* Miary te dla prostoty rozważane są jako stosowne parametry rozkładu stóp zwrotu z portfela, w szczególności VaR jest więc rozumiane jako kwantyl tego rozkładu, a nie odpowiadająca temu kwantylowi strata na wartości portfela. Ponadto VaR i CVaR są wyznaczane z perspektywy krótkiej pozycji – odnoszą się zatem do prawych ogonów rozkładu stóp zwrotu.



Rys. 3. Wartości empirycznych miar ryzyka portfela w z zależności od udziału akcji KGHM (wykresy w kolumnie (a) prezentują udział w sensie wartości, wykresy w kolumnie (b) – udział w sensie ilości; linie ciągłe odpowiadają UJ, a przerywane – UN)

Wzrokowa analiza wykresów z rys. 3 wskazuje na ewidentną zależność oszacowań miar ryzyka od przyjętej metody szacowania oczekiwanej stopy zwrotu. Mimo że z pozoru wartości oszacowań miar ryzyka różnią się (w sensie konfrontacji UN z UJ) między sobą w nieznacznym stopniu, to jednak po przełożeniu tych rozbieżności na straty na wartości portfela można uzyskać rozbieżności, których nie da się bagatelizować.

Podsumowanie

Nakreślone w opracowaniu problemy, dobitnie widoczne w świetle zaprezentowanych prostych przykładów, pozwalają na wyszczególnienie kilku ogólniejszych kwestii:

– Wartość oszacowania oczekiwanej stopy zwrotu z portfela jest istotnie zależna nie tylko od przyjętej metody szacowania (fakt ten jest powszechnie znany), ale (co ważne, a niestety nie zawsze uświadomione przez praktyków) także od rodzaju przyjmowanego ujęcia (UN, UJ).
– Zaniedbywanie istnienia zależności oszacowania oczekiwanej stopy zwrotu od rodzaju stosowanego ujęcia, przejawiające się w bezrefleksyjnym „przeszczepianiu" ujęcia Markowitza bez sprawdzenia założeń czy model Markowitza można stosować, może prowadzić do przekłamań.
– Stosowanie ujęcia jednolitego (UJ) wydaje się o wiele zasadniejsze, ponieważ wśród zalet tego ujęcia należy wymienić przynajmniej dwie: UJ cechuje spójność merytoryczna, gdyż zawsze stosuje się tę samą metodę szacowania

oczekiwanej stopy zwrotu – tak dla pojedynczych walorów jak i dla portfela o dowolnych proporcjach udziałów tych walorów; UJ odwołuje się tylko do jednego – notabene niebudzącego żadnych kontrowersji – założenia, które dotyczy wyboru metody (w sensie wzoru) szacowania oczekiwanej stopy zwrotu. Nie ma więc potrzeby czynić kolejnego, dyskusyjnego założenia związanego z interpolacją oczekiwanych stóp zwrotu z walorów na portfel.

– Pewną niedogodnością UJ jest zdecydowanie większa złożoność obliczeniowa. Szczęśliwie jednak w dobie komputerów niedogodność ta jest wręcz zaniedbywana i z pewnością nie może być uważana za wadę ujęcia.

Wobec nakreślonych powyższej kwestii zasadne jest postawienie postulatu, aby w praktyce szacować oczekiwaną stopę zwrotu z portfela w sensie ujęcia jednolitego, zaś ujęcie niejednolite – w szczególności zaczerpnięte z modelu Markowitza – ograniczyć tylko przypadków, w których istnieją zasadne przesłanki, że spełnione są założenia co do stosowalności całego modelu.

Literatura

1. Galus S., Sokołowska K., Optymalizacja portfela na polskim rynku papierów wartościowych, Prace Naukowe WSB w Gdańsku, t. 1, Gdańsk 2008.
2. Jajuga K., Jajuga T., Inwestycje. Instrumenty finansowe, aktywa finansowe, ryzyko finansowe, inżynieria finansowa, Wydawnictwo Naukowe PWN, Warszawa 2006.
3. Markowitz H., Portfolio selection, „The Journal of Finance" 1952, Vol. 7, No. 1.
4. Markowitz H., Portfolio selection. Efficient diversification of investments, John Wiley & Sons, New York 1959.
5. Tarczyński W., Rynki kapitałowe. Metody ilościowe, t. 2, Placet, Warszawa 1997.
6. R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

ALTERNATIVE FORMULATION OF RATE-OF-RETURN ESTIMATION IN COMPARISON WITH MARKOWITZ APPROACH

Summary

In the study, two approaches of rate-of-return estimation are compared. One of them, that predominates in practice and that is called by the authors heterogeneous, refers to a separate rate-of-return estimation for every individual asset, and then to an interpolation of obtained values in order to asses rate of return for any portfolio with priorly given proportions of assets. The heterogeneous approach is based on premises concerning a proper method of rate-of-return estimation for individual assets, and a specific method of interpolating estimates for any portfolio as well. In contrast, the other approach, called homogeneous, refers to uniform treatment of all portfolios without exceptions, which leads to a direct rate-of-return estimation for any portfolio with priorly given proportions of assets. The essence of both approaches and discrepancies between them are illustrated with use of properly chosen examples (arbitrary and empirical). Examples' analysis indicates some advantage of the homogeneous approach over the heterogeneous one.

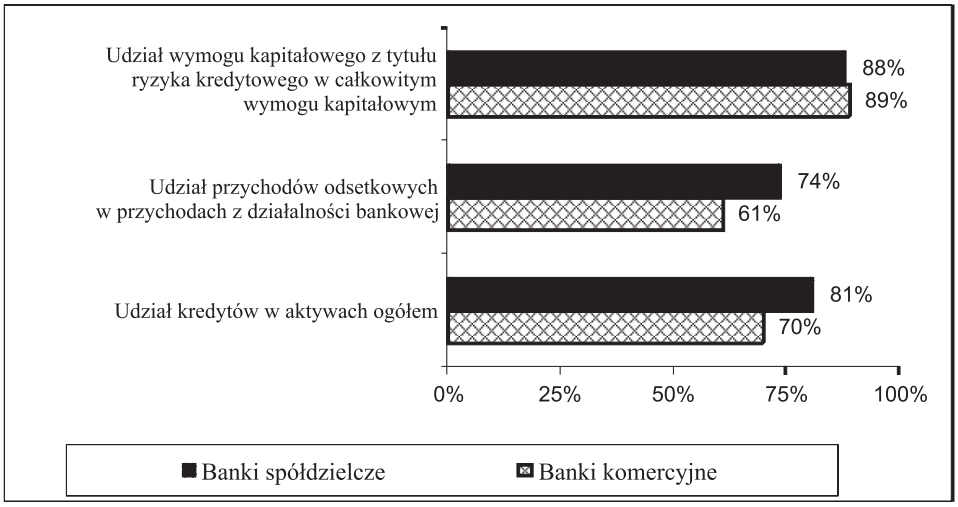
Przemysław Stodulny
Uniwersytet Ekonomiczny w Poznaniu

MODEL REFERENCYJNEJ STOPY PROCENTOWEJ W DZIAŁALNOŚCI KREDYTOWEJ BANKU

Wprowadzenie

Działalność kredytowa należy do najbardziej istotnych obszarów działalności każdego banku. Kredyty stanowią bowiem główny składnik aktywów oraz są podstawowym źródłem ryzyka bankowego (rys. 1). Wśród regulacji nadzorczych obowiązujących podmioty działające na polskim rynku większość stanowią te, które odnoszą się bezpośrednio lub pośrednio do aktywności kredytowej. Kredyty stanowią podstawowe źródło wyniku odsetkowego i zysku zatrzymanego, który warunkuje trwanie i stabilny rozwój banku. Od strony marketingowej produkty kredytowe są narzędziem wzmacniania lojalności klientów, gdyż z reguły wiążą kredytobiorcę z bankiem na dłuższy czas. Jak pokazują badania, większość klientów wykazuje mniejszą wrażliwość na zmiany oprocentowania kredytów niż depozytów [3]. Wynika to z faktu, że dostępność kredytu dla danego klienta jest ostatecznie uzależniona od decyzji banku, a zmiana kredytodawcy w trakcie trwania umowy wiąże się z dodatkowymi kosztami i znacznym nakładem czasu. Z wymienionych przesłanek wynika waga decyzji podejmowanych przez banki w zakresie ustalania ceny (stopy procentowej) dla produktów kredytowych.

W praktyce bankowej w odniesieniu do określonych produktów kredytowych powszechnie stosuje się oprocentowanie zmienne oparte na zewnętrznych stopach referencyjnych. Powoduje to poważne ryzyko bazowe stóp procentowych, zwłaszcza w sytuacji turbulencji na rynkach finansowych, co może się przekładać na pogarszanie się sytuacji płynnościowej banku. Prezentowana w opracowaniu metoda wyznaczania wewnętrznej stopy referencyjnej dla kredytów udzielanych przez banki może znaleźć zastosowanie szczególnie w przypadku, gdy bank realizuje strategię zmiennych stóp procentowych zależnych od wewnętrznych decyzji zarządu w odniesieniu do aktywów, jak i pasywów odsetkowych. Strategia ta jest powszechnie stosowana przez banki spółdzielcze.



Rys. 1. Wskaźniki dotyczące skali działalności kredytowej polskich banków (09.2011)

Źródło: Dane KNF, http:\www.knf.gov.pl

1. Praktyczne problemy ustalania cen na produkty bankowe

Wśród podstawowych metod ustalania cen na produkty bankowe trzeba wymienić metodę popytową, metodę opartą na kosztach oraz podejście konkurencyjne [1, s. 52]. Strategia cen zorientowanych na popyt opiera się na założeniu, że istnieje ujemna zależność popytu od ceny. Stąd obecny i prognozowany popyt na usługi bankowe w zestawieniu z zakładanym planem rozwoju aktywów i pasywów powinien być czynnikiem decydującym o wysokości oprocentowania dla oferowanych produktów kredytowych. Podejście konkurencyjne oznacza, że bank ustala oprocentowanie kredytów obserwując i analizując oferty konkurencji. Polski rynek bankowy ma strukturę oligopolistyczną ze względu na ograniczoną liczbę banków, wysokie bariery wejścia dotyczące wymaganych kapitałów i uregulowań prawnych, istnienie silnej współzależności między podmiotami. Zauważalny jest efekt adaptacyjny wielu średnich i mniejszych banków w zakresie polityki cenowej (wyczekiwanie na ruch ze strony liderów rynku). Stosowanie metody cen konkurencyjnych jest znacząco utrudnione przez ograniczoną informację, która jest udostępniana przez banki, zwłaszcza jeśli chodzi o rynek kredytów. W odniesieniu do oprocentowania banki podają często tylko ogólną formułę (określona stopa referencyjna powiększona o nieznaną marżę), możliwy zakres oprocento-

wania (np. od 13,99% do 17,9%) lub dolny jego przedział (od 13,99%). Faktyczne oprocentowanie udzielanego kredytu jest ustalane indywidualnie podczas rozpatrywania wniosku kredytowego, nierzadko na drodze negocjacji z klientem w powiązaniu z innymi parametrami cenowymi kredytu (prowizje, opłaty, ubezpieczenia). Metoda oparta na kosztach opiera się na założeniu, że da się zidentyfikować i przypisać do określonego produktu koszty związane z jego sprzedażą, sfinansowaniem, obsługą. W usługach bankowych faktyczne koszty związane z daną usługą są trudne do identyfikacji i rozdzielenia z uwagi na to, że zarówno personel, majątek trwały, jak i wyposażenie (w tym systemy informatyczne) są użytkowane na potrzeby wielu linii produktowych.

Najczęściej używaną formułą na ustalanie zmiennego oprocentowania kredytów jest formuła oparta na zewnętrznej stopie referencyjnej:

stopa oprocentowania kredytu = zewnętrzna stopa referencyjna + marża

W praktyce bankowej rolę stopy referencyjnej pełnią stopy rynku międzybankowego (WIBOR, WIBID LIBOR, EURIBOR) lub stopy procentowe banku centralnego (stopa redyskonta weksli, stopa lombardowa). Marża banku jest z reguły stała, nieujemna, ustalana indywidualnie lub ujednolicona w przypadku określonego typu produktów. W powyższej formule stopa oprocentowania kredytu zmienia się proporcjonalnie do stopy referencyjnej (przyrost krańcowy wynosi 1). W przypadku, gdy pozycje pasywów (źródła finansowania) są oparte na tej samej lub powiązanej stopie referencyjnej nie występują wahania spreadu odsetkowego. Niwelowane jest też ryzyko bazowe stóp procentowych. W przytoczonej formule elastyczność stopy oprocentowania kredytów względem stopy referencyjnej jest mniejsza niż 1, co oznacza, że w przypadku wzrostu stóp procentowych osiągana marża jest relatywnie niższa w stosunku do bazy.

W ofercie banków komercyjnych działających na polskim rynku przeważają produkty kredytowe z oprocentowaniem zmiennym opartym na zewnętrznej stopie referencyjnej, zwłaszcza w odniesieniu do kredytów średnio- i długoterminowych. Po stronie pasywnej przeważają pozycje zależne od suwerennych decyzji zarządów banków, które mogą się kierować określonymi w regulacjach przesłankami. Taka struktura pozycji odsetkowych stwarza ryzyko bazowe stóp procentowych, które może się ujawnić zwłaszcza w przypadku kryzysu płynnościowego. Z opisywaną sytuacją mieliśmy do czynienia na polskim rynku w roku 2007-2008. Załamanie rynku pożyczek subprime w USA doprowadziło do upadku kilku podmiotów i spowodowało utratę wzajemnego zaufania banków. Skutkiem tego było wyraźne ograniczenie aktywności na rynku międzybankowym, również w Polsce. Zawierane były nieliczne transakcje, i to z reguły na krótkie terminy. Funkcjonujące stawki rynku międzybankowego WIBOR i WIBID mia-

ły znaczenie raczej symboliczne, gdyż faktyczny koszt finansowania w przypadku pozyskiwanych depozytów był znacznie wyższy (tzw. wojna depozytowa). Spowodowało to spadek marży odsetkowej realizowanej przez banki komercyjne z 3,3% w 2006 r. do 3,0% w 2008 r.

2. Specyfika działalności banku spółdzielczego

W odróżnieniu od banków komercyjnych, w bankach spółdzielczych funkcjonuje w zdecydowanej większości koncepcja ustalania oprocentowania kredytów decyzją zarządu banku*. Uwzględniając specyfikę działalności banków spółdzielczych jest to strategia redukująca ryzyko bazowe stóp procentowych. Podmioty te finansują akcję kredytową głównie depozytami sektora niefinansowego i budżetowego, których oprocentowanie jest ustalane przez zarząd banku. Banki spółdzielcze w niewielkim stopniu wykorzystują finansowanie z rynku międzybankowego (m.in. za pośrednictwem banków zrzeszających)**. Stąd koszt finansowania jest w mniejszym stopniu uzależniony od rynkowych stóp referencyjnych. Daje to większą elastyczność w zmianach stóp procentowych dla aktywów i pasywów***. Wiąże się jednak z koniecznością posiadania mechanizmu podejmowania decyzji w sytuacji zmian w otoczeniu makroekonomicznym czy w sytuacji samego podmiotu. Narzędziem wspomagającym ten proces może być prezentowana w tym opracowaniu koncepcja wewnętrznej stopy referencyjnej dla ustalania oprocentowania kredytów. Opis modelu ma zastosowanie do banku spółdzielczego, a po koniecznych uzupełnieniach i modyfikacjach może być użyteczny również dla uniwersalnego banku działającego w formie spółki akcyjnej.

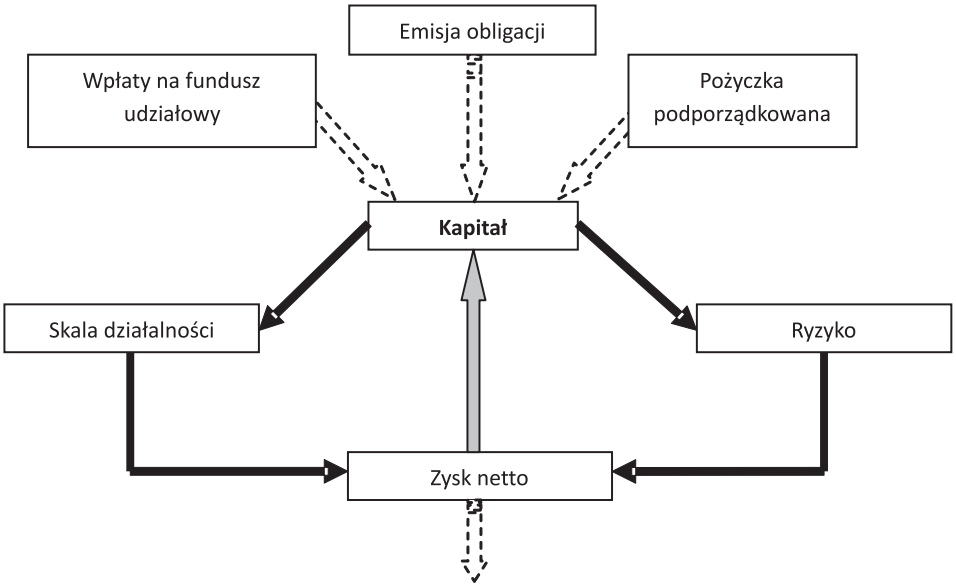
Bank powinien utrzymywać fundusze własne na poziomie nie niższym niż suma wymogów kapitałowych z tytułu poszczególnych rodzajów ryzyka [5, art. 128]. Oznacza to, że fundusze własne banku muszą uwzględniać kapitał regulacyjny na pokrycie ryzyka uwzględnionego w I filarze (ryzyko kredytowe, ryzyko operacyjne, ryzyko rynkowe) oraz ewentualne zwiększenia wynikające z II filaru (kapitał wewnętrzny). Podstawowymi czynnikami kształtującymi zapotrzebo-

* Wyjątek stanowią m.in. kredyty preferencyjne, których oprocentowanie jest uzależnione od stopy redyskonta weksli NBP.

** We wrześniu 2011 r. udział zobowiązań wobec sektora finansowego w sumie bilansowej wyniósł blisko 2,8% w przypadku banków spółdzielczych, podczas gdy dla banków komercyjnych było to 28%.

*** W okresie kryzysu finansowego sektor banków spółdzielczych zanotował wzrost marzy odsetkowej z 4,8% w 2006 r. do 5,2% w 2008 r.

wanie na kapitał jest aktualna i planowana skala działalności oraz ekspozycja na ryzyko (rys. 2).



Rys. 2. Determinanty kształtowania funduszy własnych banku spółdzielczego i podstawowe źródła ich zasilania

Planując rozwój działalności bank musi brać pod uwagę pozyskanie dodatkowych funduszy ze źródeł wewnętrznych, jak i zewnętrznych. Wzrost ten musi uwzględniać planowaną dynamikę zwiększania aktywów (w tym aktywów trwałych), obciążające je ryzyko kredytowe, zakładany poziom współczynnika wypłacalności, jak i zmiany wymogów kapitałowych na pozostałe rodzaje ryzyka. Podstawowym źródłem zwiększania kapitału własnego banku spółdzielczego jest zatrzymany zysk netto. Wśród zewnętrznych źródeł pozyskiwania kapitału można wymienić wpłaty na fundusz udziałowy, pożyczki podporządkowane oraz obligacje długoterminowe.

3. Model wewnętrznej stopy referencyjnej dla działalności kredytowej banku spółdzielczego

Zakładamy, że na oligopolistycznym rynku usług bankowych przeciętny koszt depozytów jest ustalany przez dominujące podmioty. Każdy bank realizując założony plan rozwoju za pomocą stopy oprocentowania może regulować wielkość bazy depozytowej niezbędnej do finansowania akcji kredytowej. Przy usta-

laniu oprocentowania przyjmowanych depozytów bank będzie stosował metodę popytową połączoną z monitorowaniem i analizą ofert konkurencji. Przyjmijmy, że bank wypracowuje zysk netto *ZN*, który jest przeznaczany częściowo na wynagrodzenie kapitału obcego *D* (dywidenda) oraz na wzrost funduszy własnych *ZND* (zysk zatrzymany). Tak więc:

$$ZN_t = ZND_t + D_t \tag{1}$$

Kwota ZND_t powiększa fundusze własne, a więc stanowi bazę potencjalnego rozwoju banku oraz bufor dla absorpcji podejmowanego ryzyka. Wysokość dywidendy jest proporcjonalna do udziału funduszu udziałowego *fu* w kapitałach banku *C* oraz akceptowalnej stopy dywidendy *d*. Wówczas:

$$ZN_t = ZND_t + d_t \cdot fu_t \cdot C_t \tag{2}$$

Wymagany poziom kapitałów banku określony jest przez zakładany poziom współczynnika wypłacalności *w* oraz wymogi kapitałowe na poszczególne rodzaje ryzyka zidentyfikowane przez bank*:

$$C_t = (K_t \cdot wk_t + AP_t \cdot wa_t + PW_t \cdot 12{,}5) \cdot w_t \tag{3}$$

gdzie:
K_t – kredyty i pożyczki,
wk_t – średnia waga ryzyka obciążająca obligo kredytowe,
AP_t – pozostałe aktywa,
wa_t – średnia waga ryzyka obciążająca pozostałe aktywa,
PW_t – pozostałe wymogi kapitałowe.

Poziom zysku brutto *ZB*, który bank będzie chciał zrealizować w okresie t można oszacować następująco:

$$\overline{ZB_t} = \frac{ZND_t + d_t \cdot fu_t \cdot \left(K_t \cdot wk_t + AP_t \cdot wa_t + PW_t \cdot 12{,}5\right) \cdot w_t}{1 - tx} \tag{4}$$

gdzie:
tx – stawka podatku dochodowego.

Załóżmy, że na zysk brutto składają się wynik odsetkowy *WO*, wynik prowizyjny *WP*, wynik na rezerwach *WR*, pozostałe pozycje wyniku *WOP* i koszty funkcjonowania banku (z amortyzacją) *KB*:

$$ZB_t = WO_t - KB_t + WP_t + WR_t + WOP_t \tag{5}$$

Przyjmijmy dalej, że wynik odsetkowy w okresie t można obliczyć jako różnicę przychodów od kredytów, rezerwy obowiązkowej *RO*, funduszu BFG ochro-

* Zakładamy, że bank stosuje metodę standardową obliczania wymogu kapitałowego z tytułu ryzyka kredytowego.

ny świadczeń gwarantowanych *FG*, pozostałych aktywów odsetkowych *PAO* oraz kosztów odsetkowych od zgromadzonych depozytów *BD*, wyemitowanych papierów wartościowych zaliczonych do funduszy *PF*, zaciągniętych kredytów i pożyczek *KR* i pozostałych pasywów odsetkowych *PPO*:

$$WO_t = K_t \cdot kp_t (x_t + sp_t) + RO_t \cdot ir_t + FG_t \cdot if_t + PAO_t \cdot iq_t - \\ - BD_t \cdot id_t - PF_t \cdot ip_t - KR_t \cdot ik_t - PPO_t \cdot iz_t \tag{6}$$

gdzie:
kp_t – wskaźnik kredytów pracujących,
x_t – średnia stopa oprocentowania dla obliga kredytowego,
sp_t – stawka prowizji od kredytów rozliczana w czasie,
ir_t – średnia stopa oprocentowania rezerwy obowiązkowej,
if_t – średnia stopa oprocentowania funduszu ochrony świadczeń gwarantowanych,
iq_t – średnia stopa oprocentowania pozostałych aktywów odsetkowych,
id_t – średnia stopa oprocentowania bazy depozytowej,
ip_t – średnia stopa oprocentowania papierów wartościowych zaliczonych do funduszy,
ik_t – średnia stopa oprocentowania zaciągniętych kredytów,
iz_t – średnia stopa oprocentowania pozostałych pasywów odsetkowych.

Zakładamy, że pomiędzy elementami aktywów i pasywów odsetkowych zachodzą następujące związki:

$$K_t + RO_t + FG_t + PAO_t = BD_t + PF_t + KR_t + PPO_t \tag{7}$$

$$RO_t = \alpha \cdot (BD_t + PF_t + PPO_t) - h \cdot eu_t \tag{8}$$

$$FG_t = \beta \cdot (BD_t + PF + PPO_t) \tag{9}$$

$$PAO_t = \gamma_t \cdot BD_t \tag{10}$$

gdzie:
α, h – parametry ustalane przez Radę Polityki Pieniężnej,
eu – kurs euro,
β – wskaźnik każdorazowo ustalany przez Radę BFG,
γ – parametr związany z wielkością bufora płynnościowego utrzymywanego przez bank.

Korzystając ze związków (7)-(10) można zapisać:

$$BD_t = \frac{K_t - KR_t + (PF_t + PPO_t) \cdot (\alpha + \beta - 1) - h \cdot eu_t}{(1 - \alpha - \beta - \gamma)} \tag{11}$$

Po odpowiednich przekształceniach otrzymujemy zależność opisującą wynik odsetkowy:

$$\begin{aligned} WO_t &= K_t \cdot \left(kp_t \cdot \left(x_t + sp_t\right) - g_t\right) + RO_t \cdot ir_t + FG_t \cdot if_t + \\ &+ PAO_t \cdot iq_t + KR_t\left(g_t - ik_t\right) + PF_t\left(id_t + \gamma \cdot g_t - ip_t\right) + \\ &+ PPO_t\left(id_t + \gamma \cdot g_t - iz_t\right) + g_t \cdot h \cdot eu \end{aligned} \tag{12}$$

gdzie:

$$g_t = \frac{id_t}{1 - \alpha - \beta - \gamma} \tag{13}$$

Wyrażenie (13) opisuje rzeczywisty koszt pozyskania finansowania dla udzielanych przez bank kredytów.

Z zestawienia zysku brutto osiąganego przez bank w okresie t z zyskiem zakładanym z równania (4) otrzymujemy:

$$\begin{aligned} &\frac{ZND_t + d_t \cdot fu_t \cdot \left(K_t \cdot wk_t + AP_t \cdot wa_t + PW_t \cdot 12{,}5\right) \cdot w_t}{1 - tx} = \\ &K_t \cdot \left(kp_t \cdot \left(x_t + sp_t\right) - g_t\right) + RO_t \cdot ir_t + FG_t \cdot if_t + \\ &+ PAO_t \cdot iq_t + KR_t\left(g_t - ik_t\right) + PF_t\left(id_t + \gamma \cdot g_t - ip_t\right) + \\ &+ PPO_t\left(id_t + \gamma \cdot g_t - iz_t\right) + g_t \cdot h \cdot eu - KB_t + WP_t + WR_t + WOP_t \end{aligned} \tag{14}$$

Odpowiednie przekształcenia prowadzą do uzyskania równania opisującego wewnętrzną stopę referencyjną x_t:

$$\begin{aligned} x_t = {} & \frac{\overline{ZB}_t}{K_t \cdot kp_t} + \frac{g_t}{kp_t} + \frac{KB_t - WP_t - WR_t - WOP_t}{K_t \cdot kp_t} && \text{a, b, c} \\ & - \frac{RO_t \cdot ir_t + FG_t \cdot if_t + PAO_t \cdot iq_t + KR_t\left(g_t - ik_t\right)}{K_t \cdot kp_t} - sp_t && \text{d} \\ & - \frac{PF_t\left(id_t + \gamma \cdot g_t - ip_t\right) + PPO_t\left(id_t + \gamma \cdot g_t - iz_t\right) + g_t \cdot h \cdot eu}{K_t \cdot kp_t} && \text{e} \end{aligned} \tag{15}$$

Pierwsza część równania (15) wyraża obciążenie pracującego obliga kredytowego założonym zyskiem brutto, druga (b) zakłada pokrycie kosztów finansowania (baza depozytowa). W trzeciej części (15c) zobrazowano obciążenie pracującego obliga kredytowego kosztami działalności banku skorygowanymi o pozostałe pozycje wyniku. Pozostałe wyrażenia (15d i 15e) obrazują korektę obciążenia pracującego obliga kredytowego z tytułu pozostałych pozycji wyniku odsetkowego.

Problem wyznaczenia wartości ZND_t sprowadza się do pytania o punkt odniesienia wyników osiąganych przez bank. Proponowane dwa podejścia odnoszą się do szacowanego zapotrzebowania na kapitał dla utrzymania/zwiększenia tempa wzrostu skali działania oraz do osiągnięcia określonych wskaźników efektywności działania. W tym drugim przypadku punktem odniesienia mogą być wskaźniki ROA* i ROE* dla grupy rówieśniczej banków. Wówczas:

$$ZND_t = ROA^* \cdot (K_t + AP_t) - D_t \tag{16}$$

gdzie:
D_t – dywidenda za okres t.

Wypracowany zysk netto jest dla banku podstawą do zwiększania kapitałów, co z kolei umożliwia zwiększanie skali działalności i skali podejmowanego ryzyka. W takim podejściu wielkość ZND_t można oszacować następująco:

$$\begin{aligned} ZND_t &= K_t \cdot (a \cdot w_{t+1} \cdot wk_{t+1} - w_t \cdot wk_t) + \\ &+ AP_t \cdot (b \cdot w_{t+1} \cdot wa_{t+1} - w_t \cdot wa_t) + \\ &+ 12{,}5 \cdot (w_{t+1} \cdot PW_{t+1} - w_t \cdot PW_t) - (FU_{t+1} - FU_t) - (PF_{t+1} - PF_t) \end{aligned} \tag{17}$$

gdzie:
a – zakładany wskaźnik wzrostu obliga kredytowego w okresie t + 1,
b – zakładany wskaźnik wzrostu pozostałych aktywów w okresie t + 1,
FU_t – średni fundusz udziałowy w okresie t.

W prezentowanej koncepcji wewnętrznej stopy referencyjnej nominalna stopa oprocentowania kredytu będzie wyznaczana według formuły:

stopa oprocentowania kredytu = wewnętrzna stopa referencyjna + korekta

Wewnętrzna stopa referencyjna jest zmienna i zależna od decyzji zarządu banku. Może ona dotyczyć tylko kredytów nowo udzielanych lub całego obliga kredytów w trakcie spłaty. W tym drugim przypadku będzie rozumiana jako minimalna dochodowość obliga kredytów zapewniająca pokrycie kosztów oraz realizację założonych celów. Znaczne odchylenie realizowanej dochodowości obliga kredytowego od zakładanej może być impulsem do podjęcia przez zarząd decyzji o zmianie stóp oprocentowania kredytów. Musi to być jednak decyzja uwzględniająca realia rynkowe (stopy rynkowe, zachowania konkurencji) oraz czynnik popytowy w zestawieniu z zakładanym wzrostem akcji kredytowej.

Korekta jest wartością z przedziału $\langle g_t - x_t; 4L_t - x_t \rangle$, gdzie L jest stopą lombardową NBP. Wartość korekty pozwala na różnicowanie oprocentowania kredytów ze względu na produkty, grupy klientów, kanały dystrybucji, jakość zabezpieczeń, zdolność kredytową klienta itp., jak i uwzględnianie indywidualnie wynegocjowanych warunków kredytowych.

Podsumowanie

Z zaprezentowanych zależności wynika, że:

– Gorsza jakość portfela wpływa na wzrost stopy referencyjnej poprzez: wzrost rezerw celowych obciążających wynik; zwiększenie wymogów kapitałowych; zmniejszenie wolumenu kredytów pracujących.
– Wzrost skali działalności banku (obliga kredytowego) wpływa na asymptotyczne obniżenie stopy referencyjnej o ile nie pociąga za sobą proporcjonalnego obciążenia kosztami funkcjonowania.
– Obniżenie wkładu funduszu udziałowego kosztem własnych źródeł kapitału wpływa na obniżenie wewnętrznej stopy referencyjnej.
– Zastępowanie funduszu udziałowego innymi formami kapitału obcego ujawnia efekt kosztowy (różnica między stopą dywidendy a kosztem alternatywnym przekłada się na obciążenie wyniku netto) i podatkowy (dywidenda jest wypłacana z zysku netto, odsetki od kapitału obcego stanowią koszt uzyskania przychodu).
– Konieczność utrzymywania większego bufora płynnościowego podnosi wewnętrzną stopę referencyjną poprzez zmianę struktury aktywów.

Zaprezentowany model matematyczny wewnętrznej stopy referencyjnej jest uproszczonym opisem działalności banku, poszczególne opisane pozycje mogą być w razie potrzeby uszczegółowione. Model zawiera parametry zewnętrzne i zmienne endogeniczne, które są wynikiem decyzji i działań zarządu banku. Praktyczne zastosowanie może się opierać na danych zrealizowanych (controlling) bądź na planach finansowych banku. W obu przypadkach celem może być przyjęcie określonej polityki cenowej w odniesieniu do działalności kredytowej oraz bieżące jej modyfikowanie dla realizacji przyjętych celów. Podobną kalkulację do zaprezentowanej można zastosować w odniesieniu do marży banku przy stosowaniu formuły zewnętrznej stopy referencyjnej.

Literatura

1. Grzywacz J., Marketing w działalności banku, Difin, Warszawa 2006.
2. Harrisom T., Financial Services Marketing, Pearson Education, Harlow 2000.
3. Wdowiński P., Ekonometryczna analiza popytu w polskiej gospodarce, UKNF, Warszawa 2011.

4. Witryna internetowa Komisji Nadzoru Finansowego www.knf.gov.pl
5. Ustawa Prawo Bankowe z 29 sierpnia 1997 r., Dz.U. 2002, nr 72, poz. 665.

REFERENCE RATE MODEL IN BANK LENDING ACTIVITIES

Summary

The paper presents the conception and method of estimating the internal reference rate for loans granted by banks. It's based on the structure of the bank's assets and liabilities, sources of financing and the cost of raising funds, the bank's capital needs, capital costs, and other items of the financial result. Presented method can be used particularly when the bank realizes strategy of variable interest rates depending on the internal decision of the Board in respect of the interest assets and liabilities. This strategy is commonly used by the co-operative banks.

The paper contains the model of the cooperative bank, described by mathematical equations that indicate relationships between the parts of the balance sheet, income statement and capital adequacy calculation. The resulting formula for the calculation of the internal reference rate for loans may be used for ongoing monitoring of the bank's operations and strategic planning.

Joanna Utkin
Szkoła Główna Handlowa

OPTYMALNY ZRANDOMIZOWANY TEST NA SKOŃCZONYM RYNKU ZUPEŁNYM

Wprowadzenie

Problematyka osłony dotyczy zabezpieczenia przyszłego zobowiązania na rynku kapitałowym. Klasyczny problem osłony kwantylowej polega na maksymalizacji prawdopodobieństwa zabezpieczenia na rynku zupełnym i pozbawionym możliwości arbitrażu. Znany jest warunek dostateczny takiej osłony przy danym ograniczeniu budżetowym [por. 1; 2]. Uogólniony problem osłony kwantylowej jest formułowany na rynku bez możliwości arbitrażu, który nie musi być zupełny. Zagadnienie to dotyczy maksymalizacji wartości oczekiwanej tzw. współczynnika sukcesu. Follmer, Leukert i Schied zaproponowali dwustopniowe rozwiązanie problemu: najpierw maksymalizuje się wartość oczekiwaną zrandomizowanego testu, a następnie za pomocą zrandomizowanego testu przedstawia się optymalny współczynnik sukcesu. Ponadto na rynku zupełnym przy spełnionym warunku dostatecznym klasycznej osłony kwantylowej otrzymano tę samą wartość końcową optymalnej wypłaty [por. 1; 2]. Maksymalizacja wartości średniej zrandomizowanego testu została oparta na uogólnionym lemacie Neymana-Pearsona.

Niniejsze opracowanie jest poświęcone badaniu własności optymalnego zrandomizowanego testu na skończonym rynku zupełnym. Własności te uzyskuje się dzięki rozwiązaniu zadania optymalizacji odpowiednimi metodami programowania matematycznego. Na ich podstawie wykazuje się, że minimum uogólnionej gęstości jest odpowiednim dolnym kwantylem. Umożliwia to porównanie wyników ze znanymi warunkami klasycznej osłony kwantylowej.

1. Problematyka osłony kwantylowej

1.1. Model rynku i zobowiązania

Rozważamy rynek kapitałowy o skończonej liczbie stanów. Występują na nim okresowo bezpieczne konto bankowe oraz ryzykowne instrumenty finansowe. Model rynku jest zupełny i pozbawiony możliwości arbitrażu. Transakcje na rynku odbywają się w chwilach $t = 0,\ldots, T$, $\Omega = \{\omega_1,\ldots, \omega_N\}$ jest zbiorem stanów rynku w chwili $t = T$, P jest funkcją prawdopodobieństwa rzeczywistego, zaś Q – funkcją prawdopodobieństwa martyngałowego na zbiorze Ω. Przez K oznaczamy proces konta bankowego. Procesowi wartości portfela V przyporządkowany jest zdyskontowany proces wartości $Y = V / K$. Przez V_T oznaczamy końcową wartość portfela, zaś przez V_0 jej wycenę bezarbitrażową na chwilę początkową.

Na rynku rozważamy dane zobowiązanie typu europejskiego płatne w terminie $t = T$. Jest ono reprezentowane przez zmienną losową L spełniającą założenie $0 \neq L \geq 0$. Zobowiązaniu L odpowiada zdyskontowane zobowiązanie $H = L / K_T$. Przez L_0 oznaczamy bezarbitrażową wycenę zobowiązania na moment $t = 0$. Zakładamy, że na początek mamy do dyspozycji budżet v spełniający ograniczenie:

$$0 < v < L_0$$

Wprowadzamy oznaczenie $\alpha = \dfrac{v}{L_0}$. Zatem $\alpha \in (0,1)$. Wówczas różnica:

$$1-\alpha = \frac{L_0 - v}{L_0} \tag{1}$$

reprezentuje względną cenę niedoboru dla danych: zobowiązania i budżetu.

Problemy osłony kwantylowej przytoczone poniżej będą formułowane w przypadku rynku skończonego. Zmienne losowe X na zbiorze Ω będą traktowane jak wektory z R^N o współrzędnych $X_i = X(\omega_i)$, $i = 1,\ldots, N$. Iloczyn zmiennych losowych $X \cdot Y$ potraktujemy jako wektor o współrzędnych X_iY_i, $i = 1,\ldots, N$, iloraz zmiennych losowych X / Y jako wektor o współrzędnych X_i / Y_i, $i = 1,\ldots, N$. Ponadto z uwagi na kontekst rynku zupełnego w sformułowaniach cytowanych twierdzeń uwzględnimy jedynie wypłaty końcowe, gdyż kwestia replikacji jest w tym przypadku oczywista.

1.2. Klasyczny problem osłony kwantylowej

W klasycznym zagadnieniu osłony kwantylowej szuka się wypłaty końcowej V_T^* maksymalizującej prawdopodobieństwo osłony przy ograniczeniu budżetowym, czyli:

$$P^T 1_{V_T^* \geq L} = \max\left\{P^T 1_{V_T \geq L} : V_0 \leq v \wedge V_T \geq 0\right\} \tag{2}$$

Pierwsza z nierówności stanowi początkowe ograniczenie budżetowe, druga zakaz końcowego zadłużenia w każdym stanie. W stanie ω mamy do czynienia z osłoną wówczas, gdy $V_T^*(\omega) \geq L(\omega)$.

Narzędzi do rozwiązywania problemu (2) dostarczają dwa twierdzenia: o wypłacie optymalnej oraz o warunku dostatecznym optymalności.

Twierdzenie o wypłacie optymalnej [por. 2, tw. 8.2]

Jeśli zdarzenie $A^* \subset \Omega$ jest rozwiązaniem zadania:

$$P^T 1_{A^*} = \max\left\{P^T 1_A : A \subset \Omega \wedge Q^T\left(H \cdot 1_A\right) \leq v\right\} \tag{3}$$

to wypłata końcowa $V_T^* = L \cdot 1_{A^*}$ jest rozwiązaniem zadania (2).

Przytoczenie następnego twierdzenia poprzedzają definicje dwóch pomocniczych pojęć.

Definicja 1

Funkcję $F : \Omega \to \langle 0,1 \rangle$ o wartościach $F(\omega) = Q(\omega)\, H(\omega) / L_0$ nazywamy funkcją prawdopodobieństwa pomocniczego. W przypadku rynku o N stanach traktujemy F jako wektor o współrzędnych:

$$F_i = \frac{Q_i H_i}{L_0}, i = 1,\ldots,N \tag{4}$$

Wniosek 1

Jeśli $A \subset \Omega$, to $F^T 1_A = Q^T\left(H \cdot 1_A\right) / L_0$.

Wniosek 2

F jest funkcją prawdopodobieństwa na zbiorze $\Omega - F^{-1}\left\{0\right\}$.

Definicja 2

Mówimy, że P / F jest uogólnioną gęstością P względem F na zbiorze Ω, jeśli $(P/F)(\omega) = P(\omega)/F(\omega)$ przy $F(\omega) \neq 0$, zaś $(P/F)(\omega) = +\infty$ przy $F(\omega) = 0$. Zgodnie z tą konwencją, na rynku skończonym wektor P/F ma współrzędne:

$$(P/F)_i = \begin{cases} P_i / F_i, F_i \neq 0 \\ +\infty, F_i = 0 \end{cases} \tag{5}$$

Twierdzenie o warunku dostatecznym optymalności [por. 2, tw. 8.3]

Jeśli zdarzenie:

$$A^* = \left\{\omega \in \Omega : P(\omega) / F(\omega) >_F q_{1-\alpha}(P/F)\right\}^* \tag{6}$$

spełnia równanie $F^T 1_{A^*} = \alpha$, to A^* jest zdarzeniem optymalnym zadania (3) i wypłata końcowa $V_T^* = L \cdot 1_{A^*}$ jest rozwiązaniem optymalnym zadania (2). Ponadto A^* jest zdarzeniem osłony.

Uwaga 1

Wymieniony w [1; 2] warunek wystarczający do spełnienia założenia twierdzenia o warunku dostatecznym, mianowicie $P^T 1_{P/Q=d\,H/L_0} = 0$ nie jest nigdy spełniony na rynku skończonym. Jeżeli $\Omega = N$, to istnieje $i \in \{1,\ldots N\}$, dla którego dolny kwantyl jest równy właściwej wartości zmiennej losowej ${}_F q_{1-a}(P/F) = P_i / F_i$, więc wcześniej wymieniony warunek nie zachodzi.

1.3. Uogólniony problem osłony kwantylowej

Dla danego zobowiązania Follmer i in. [1; 2] określili „współczynnik sukcesu". Sukces polega na osłonie zobowiązania przez wartość końcową portfela. Z uwagi na to, że martyngałami są zdyskontowane płatności losowe, w sformułowaniach pojęć wykorzystuje się zdyskontowane zobowiązanie H oraz zdyskontowaną wartość końcową portfela Y_T.

* ${}_F q_{1-\alpha}(P/F)$ oznacza dolny 1–α – kwantyl uogólnionej gęstości względem F.

Definicja 3

Współczynnikiem sukcesu nieujemnej zdyskontowanej wypłaty Y_T nazywamy zmienną losową określoną wzorem:

$$\Psi_Y = 1_{Y_T \geq H} + 1_{Y_T < H} \cdot Y_T / H \tag{7}$$

W uogólnionym zagadnieniu osłony kwantylowej szuka się zdyskontowanej wypłaty końcowej Y_T^* maksymalizującej wartość oczekiwaną współczynnika sukcesu przy ograniczeniu budżetowym:

$$P^T \Psi_{Y*} = \max \left\{ P^T \Psi_Y : Q^T Y_T \leq v \right\} \tag{8}$$

Metoda wyznaczania optymalnej wypłaty końcowej, przedstawiona w pracach [1; 2] opiera się na optymalizacji wartości oczekiwanej zrandomizowanego testu Ψ. Ten pomocniczy problem optymalizacyjny ma na rynku o N stanach końcowych postać:

$$P^T \Psi_{Y*} = \max \left\{ P^T \Psi_Y : Q^T Y_T \leq v \right\} \tag{9}$$

Związek problemów (8) i (9) wynika z następującego twierdzenia:

Twierdzenie o maksymalizacji średniego współczynnika sukcesu [por. 2, tw. 8.6]

Jeśli Ψ^* jest rozwiązaniem optymalnym problemu (9), to zdyskontowana wypłata $Y_T^* = H \cdot \Psi^*$ ma współczynnik sukcesu Ψ_{Y^*} optymalny ze względu na problem (8), przy czym $\Psi_{Y^*} = \Psi^*$ oraz $Q^T Y_T^* = v$.

Postać optymalnego zrandomizowanego testu wynika z tzw. uogólnionego lematu Neymana-Pearsona. W cytowanym wniosku z uogólnionego lematu Neymana-Pearsona uwzględniona jest uwaga 1.

Wniosek z uogólnionego lematu Neymana-Pearsona [por. 2, tw. A30]

Współrzędne zrandomizowanego testu będącego rozwiązaniem problemu (9) są następujące:

$$\Psi_i^0 = \begin{cases} 0 \quad \text{przy} \quad P_i / F_i < c & (10) \\ \dfrac{\alpha - F^T 1_{P/F>c}}{F^T 1_{P/F=c}} \quad \text{przy } P_i / F_i = c & (11) \\ 1 \quad \text{przy} \quad P_i / F_i > c & (12) \end{cases}$$

gdzie:

$$c = {}_F q_{1-\alpha} (P/F) \tag{13}$$

Numeryczne rozwiązania licznych przykładów zadania (9) nie zawierały przypadku (10) rozwiązania optymalnego. Stało się to motywacją do podjęcia badań zadania programowania liniowego (9) pod kątem własności rozwiązania optymalnego.

2. Optymalny zrandomizowany test jako rozwiązanie zadania programowania matematycznego

Podjęcie badania rozwiązania optymalnego problemu (9) za pomocą narzędzi programowania matematycznego dostarcza pełniejszej wiedzy o optymalnym zrandomizowanym teście niż formuły (10)-(13) oparte na uogólnionym lemacie Neymana-Pearsona.

Dane są skalar $\alpha \in (0,1)$ oraz dwa wektory P i F spełniające założenia $P \in R^N_{++}, P^T 1 = 1, F \in R^N_{+}, F^T 1 = 1$. Wprowadzamy następujący zbiór rozwiązań dopuszczalnych:

$$D = \left\{ \Psi \in R^N : \Psi \in \langle 0,1 \rangle^N \wedge F^T \Psi \leq \alpha \right\} \tag{14}$$

Rozważamy zadanie programowania liniowego:

$$\max_{\Psi \in D} P^T \Psi \tag{15}$$

Z założeń o wektorze F wiemy, że $F \in \langle 0,1 \rangle^N$ i $F^T 1 = 1$. Stąd i z założenia $\alpha \in (0,1)$ wynikają następujące związki położenia jednostkowej kostki $\langle 0,1 \rangle^N$ i hiperpłaszczyzny o równaniu $F^T \Psi = \alpha$:

$$\langle 0,1 \rangle^N \cap \left\{ \Psi \in R^N : F^T \Psi = \alpha \right\} \neq O$$

$$i \in \{1,...,N\} \Rightarrow \left\{ \Psi \in R^N : e_i^T \Psi = 0 \right\} \neq \left\{ \Psi \in R^N : F^T \Psi = \alpha \right\}$$

$$i \in \{1,...,N\} \Rightarrow \left\{ \Psi \in R^N : e_i^T \Psi = 1 \right\} \neq \left\{ \Psi \in R^N : F^T \Psi = \alpha \right\}$$

D jest zatem wielościanem wypukłym, skąd wynika istnienie rozwiązania optymalnego zadania (15). Ponadto $IntD \neq O$. W pewnych $\Psi \in IntD$ wektor F jest dopuszczalnym kierunkiem poprawiającym, gdyż $F^T P > 0$ (por. [3]). Wobec tego otrzymujemy wniosek:

Wniosek 3

Zadanie (15) ma rozwiązanie optymalne Ψ^0 spełniające równanie $F^T \Psi = \alpha$.

Pierwszy etap optymalizacji funkcji celu doprowadził do lokalizacji rozwiązania optymalnego na ścianie zbioru D:

$$\left\{ \Psi \in R^N : \Psi \in \langle 0,1 \rangle^N \wedge F^T \Psi = \alpha \right\} \tag{16}$$

Drugi etap będzie polegał na poprawianiu wartości funkcji celu ze względu na odpowiednio wybrane pary współrzędnych, co zaowocuje wyznaczeniem podzbioru rozwiązań optymalnych w zbiorze (16).

Wyszczególniamy najmniejszą wartość uogólnionej gęstości (por. definicja 2):

$$g_{\min} = \min\left\{P_i / F_i : i = 1,...,N\right\} \tag{17}$$

Wprowadzamy podział współrzędnych:

$$i \in I_0 \Leftrightarrow P_i / F_i = g_{\min} \tag{18}$$

$$I_1 = \left\{1,...,N\right\} - I_0 \tag{19}$$

Jeżeli $j \in I_0$ oraz $i \in I_1$, to $F_j > 0$ i nierówność $P_j / F_j < P_i / F_i$ wynikająca z (17), (18), (19) jest równoważna nierówności:

$$F_i / F_j < P_i / P_j \tag{20}$$

Ustalamy dowolnie $j \in I_0$. Dla każdego $i \in I_1$ dwie współrzędne Ψ_j, Ψ_i traktujemy jako zmienne, a pozostałe współrzędne (jeśli takie występują) jako stałe. Zadanie optymalizacyjne przybiera wówczas postać:

$$\begin{gathered} \max\left(P_j\Psi_j + P_i\Psi_i + const_1\right) \\ \begin{cases} F_j\Psi_j + F_i\Psi_i + const_2 = \alpha \\ \Psi_j, \Psi_i \in \langle 0,1 \rangle \end{cases} \end{gathered} \tag{21}$$

Z nierówności (20) wynika, że funkcja celu zadania (21) szybciej rośnie w kierunku e_i niż w kierunku e_j, a więc rozwiązanie optymalne zadania (21) ma współrzędną $\Psi_i^0 = 1$ Rezultat ten nie zależy od ustalonych wartości pozostałych współrzędnych wektora Ψ ze zbioru (16). Optymalizacja odbywa się mianowicie na części jednostkowego kwadratu leżącej pod prostą o danym nachyleniu, zaś funkcja celu jest liniowa. Wybór jednostkowy współrzędnej wierzchołka zbioru dopuszczalnego zależy jedynie od porównania tangensów kątów (16). Rozwiązując zadanie (21) dla każdego $i \in I_1$ otrzymujemy wniosek:

Wniosek 4

Jeżeli $i \in I_1$, to i współrzędna rozwiązania optymalnego problemu (15) jest równa $\Psi_i^0 = 1$.

Poszukiwanie rozwiązania optymalnego zadania (15) zostało więc zawężone do zbioru wektorów z R^N dla których:

$$\Psi_i^0 = 1 \quad \text{dla} \quad i \in I_1 \tag{22}$$

zaś współrzędne Ψ_j^0 dla $j \in I_0$ są związane z następującymi warunkami wynikającymi z wniosków 3 i 4:

$$\sum_{j \in I_0} F_j \Psi_j^0 = \alpha - \sum_{i \in I_1} F_i \quad \text{i} \quad \Psi_j^0 \in \langle 0,1 \rangle \tag{23}$$

Z określenia wskaźników (18) wiemy, że:

$$P_j = g_{\min} F_j \quad \text{dla} \quad j \in I_0 \tag{24}$$

Po podstawieniu (22) i (23) do wzoru na wartość funkcji celu zadania (15) i po skorzystaniu z (24) otrzymujemy:

$$P^T \Psi^0 = \alpha g_{\min} + \sum_{i \in I_1} \left(P_i - g_{\min} F_i \right) \tag{25}$$

Z wniosku 3 wiadomo, że zadanie (15) ma rozwiązanie optymalne. Jeśli I_0 jest zbiorem jednoelementowym, to jedyna współrzędna $\Psi_j^0, j \in I_0$ jest równa:

$$\Psi_j^0 = \left(\alpha - \sum_{i=I_1} F_i \right) / F_j \tag{26}$$

Jeżeli I_0 ma więcej niż jeden element, to wartość optymalna (25) jest przyjęta na zbiorze wektorów spełniających warunki (22) i (23). Niezależność wartości optymalnej od $\Psi_j^0, j \in I_0$ jest konsekwencją równoległości warstwicy funkcji celu jako funkcji zmiennych $\Psi_j^0, j \in I_0$, do zbioru określonego przez (23). Dowolnie ustalając $j \in I_0$ możemy przedstawić Ψ_j^0 w postaci parametrycznej:

$$\Psi_j^0 = \left(\alpha - \sum_{i=I_1} F_i - \sum_{k \in I_0 - \{j\}} F_k \Psi_k^0 \right) / F_j \tag{27}$$

gdzie $\Psi_k^0 \in \langle 0,1 \rangle, k \in I_0 - \{j\}$.

Wniosek 5

Jeżeli $\overline{\overline{I}}_0 = 1$, to współrzędne jedynego rozwiązania optymalnego Ψ_0 są dane za pomocą wzorów (22) i (26). Jeżeli $\overline{\overline{I}}_0 > 1$, to zbiór rozwiązań optymalnych składa się z wektorów Ψ^0 o współrzędnych spełniających warunki (22) i (27).

Poszukamy takiego rozwiązania optymalnego, które ma stałe współrzędne $\Psi^0 = \beta$ dla $j \in I_0$. Z (23) otrzymujemy:

$$\beta = \left(\alpha - \sum_{i=I_1} F_i \right) / \sum_{j \in I_0} F_j \tag{28}$$

Z istnienia rozwiązania optymalnego spełniającego (23) wiadomo, że $0 \le \beta \le 1$.

Wniosek 6

Jednym z rozwiązań optymalnych zadania (15) jest wektor $\Psi^0 \in R^N$ o współrzędnych:

$$\Psi_i^0 = \begin{cases} \beta \ , \ P_i / F_i = g_{\min} \\ 1 \ \ , P_i/F_i > g_{\min} \end{cases} \tag{29}$$

gdzie β jest dane wzorem (28). W przypadku, gdy wartość minimalna uogólnionej gęstości jest osiągnięta tylko w jednym stanie, układ (29) określa jedynie rozwiązanie optymalne.

Porównując wzory (10), (11) i (29) można zauważyć eliminację zerowych wartości współrzędnych, poza ewentualnym przypadkiem $\beta = 0$.

3. Rozwiązanie optymalne a dolny kwantyl uogólnionej gęstości

We wniosku z uogólnionego lematu Neymana-Pearsona, cytowanego w rozdziale 1.3, wartości współrzędnych rozwiązania optymalnego zależą od porównania wartości uogólnionej gęstości z jej dolnym $1 - \alpha$ kwantylem. Poniżej zbadamy prawdopodobieństwa pomocnicze odpowiednich zdarzeń.

Z charakterystyki optymalnego testu (23) wynika, że:

$$\sum_{i=I_1} F_i \le \alpha \tag{30}$$

Ponieważ zachodzi $\sum_{j\in I_0} F_j + \sum_{i\in I_1} F_i = 1,$ to na podstawie (30) otrzymujemy oszacowanie:

$$\sum_{j\in I_0} F_j \ge 1-\alpha \tag{31}$$

Biorąc pod uwagę (18) i (19), nierówności (30) i (31) możemy odpowiednio zapisać w postaci:

$$F^T 1_{P/F>g_{\min}} \le \alpha \tag{32}$$

$$F^T 1_{P/F=g_{\min}} \ge 1-\alpha \tag{33}$$

Z nierówności (33) i z definicji dolnego kwantyla zastosowanej do skokowego rozkładu prawdopodobieństwa pomocniczego F wynika, że:

$$_F q_{1-a}\left(P/F\right) = g_{\min} \tag{34}$$

zatem w układzie wzorów (10)-(12) zachodzi $g_{\min} = c$.

Po uzyskaniu równości (34) możemy zbadać warunek dostateczny klasycznej kwantylowej osłony zobowiązania. Zdarzenie osłony jest, jak wiadomo z (6) i (34), określone przez nierówność $P/F > g_{\min}$, więc jego prawdopodobieństwo

pomocnicze jest równe $F^T 1_{P/F > g_{\min}}$. Warunek dostateczny osłony kwantylowej, przytoczony w rozdziale 1.2, ma postać:

$$F^T 1_{P/F>g_{\min}} = \alpha \tag{35}$$

lub inaczej:

$$F^T 1_{P/F=g_{\min}} = 1-\alpha \tag{36}$$

Wyrażając (36) za pomocą zbioru wskaźników (18) otrzymujemy następujący wniosek:

Wniosek 7

W celu spełnienia warunku dostatecznego klasycznej osłony kwantylowej na rynku skończonym potrzeba i wystarcza, aby:

$$\sum_{j\in I_0} F_j = 1-\alpha \tag{37}$$

Zdarzenie osłony jest wówczas zbiorem $\{\omega_i: i \in I_1\}$. Ponadto w tym przypadku $\beta = 0$.

Na rynku zupełnym wartość początkową zobowiązania możemy, zgodnie z (18) i (19), rozłożyć na dwa składniki:

$$L_0 = \sum_{j\in I_0} Q_j \cdot H_j + \sum_{i\in I_1} Q_i \cdot H_i \tag{38}$$

Jeśli spełniony jest warunek klasycznej osłony kwantylowej (37), to z (5) wynika, że składnik wyceny bezarbitrażowej zobowiązań w osłoniętych stanach (19) jest największy i równy kwocie budżetu:

$$\sum_{i\in I_1} Q_i \cdot H_i = v \tag{39}$$

a w konsekwencji składnik wyceny zobowiązania w nieosłoniętych stanach (18) jest najmniejszy i równy wartości niedoboru:

$$\sum_{j\in I_0} Q_j \cdot H_j = L_0 - v \tag{40}$$

Na ogół przy uogólnionej osłonie kwantylowej zachodzą jedynie nierówności wynikające z (31) i (30), mianowicie:

$$\sum_{i\in I_1} Q_i \cdot H_i \le v \tag{41}$$

Tak więc wycena bezarbitrażowa zobowiązań w stanach minimalnej gęstości przy uogólnionej osłonie kwantylowej jest nie mniejsza od wartości niedoboru:

$$\sum_{j\in I_0} Q_j \cdot H_j \ge L_0 - v \tag{42}$$

Podsumowanie

Znajomość skończonego zbioru wartości uogólnionej gęstości *P/F* pozwala wskazać gęstość minimalną i porównać ją ze względną ceną niedoboru (1). Jeśli względna cena niedoboru jest równa minimalnej gęstości, to możemy zagwarantować klasyczną osłonę kwantylową zobowiązania wyznaczoną za pomocą zrandomizowanego testu (29) o współrzędnych z {0,1}. Gdy zaś względna cena niedoboru różni się od minimalnej gęstości, to musimy się zadowolić uogólnioną osłoną kwantylową zobowiązania, otrzymaną na podstawie zrandomizowanego testu (29) o współrzędnych z {β, 1}, gdzie $\beta > 0$ (por. (28)).

Literatura

1. Follmer H., Leukert P., Quantile hedging. "Finance Stochastics" 1999, No. 3.
2. Follmer H., Schied A., Stochastic finance, de Gruyter, Berlin 2004.
3. Zangwill W., Programowanie nieliniowe, WNT, Warszawa 1974.

OPTIMAL RANDOMIZED TEST IN THE FINITE COMPLETE MARKET

Summary

We deal with the finite complete arbitrage free financial market model. There are given a liability (as selling an european option) and an initial amount lower than the initial value of the liability. The quantile hedging is based upon the generalized Neyman- Pearson lemma, but this approach don't give all information on the optimal solution in the considered case. In the present paper the optimal randomized test is analysed with some methods of the mathematical programming. It is showed that the minimal generalized density of probabilities equals the needed lower quantil. Moreover we construct the optimal solutions set. It is the basis to formulate the sufficient condition of the classical quantile hedging.

Stanisław Wieteska
Uniwersytet Jana Kochanowskiego w Kielcach
Filia w Piotrkowie Trybunalskim

PRZYDATNOŚĆ PRAWDOPODOBIEŃSTWA SUBIEKTYWNEGO DO KALKULACJI STÓP SKŁADEK W UBEZPIECZENIACH OBOWIĄZKOWYCH ODPOWIEDZIALNOŚCI CYWILNEJ ZAWODOWEJ

Wprowadzenie

Ubezpieczenia majątkowe wkraczają w coraz to nowe nieznane w literaturze ubezpieczeniowej obszary działalności człowieka. Jednym z takich obszarów są ubezpieczenia odpowiedzialności cywilnej. W latach 2004-2010 ukazało się kilkanaście ogólnych warunków obowiązkowych ubezpieczeń odpowiedzialności cywilnej (OC) dla różnych grup zawodowych, a także różnych instytucji w postaci Rozporządzeń Ministra RP. Pomijamy tutaj ubezpieczenia obowiązkowe wynikające z mocy ustawy o działalności ubezpieczeniowej. Treść merytoryczna tych warunków jest bardzo skromna i ogólnikowa. Obszerną analizę zakresu odpowiedzialności oraz wyłączeń odpowiedzialności cywilnej w tych ubezpieczeniach spotykamy dość często w literaturze ubezpieczeniowej[*]. Ponieważ są to ubezpieczenia obowiązkowe, każdy zakład ubezpieczeń nie ma prawa odmówić zawarcia takiego ubezpieczenia, jednak przed zawarciem takiego ubezpieczenia konieczne jest obliczenie stóp składek dla danych ubezpieczeń. Aby z kolei obliczyć stopę składki muszą być obiektywne dane o częstości szkód (prawdopodobieństwa szkody), a także rozmiarze szkód możliwych do pokrycia przez zakład ubezpieczeń. Tymczasem osiągnięcie takich danych statystycznych nastręcza sporo trudności. Jak dotychczas, przed wprowadzeniem obowiązkowych ubezpieczeń OC

[*] Por. m.in. [5, s. 53-66].

zakłady ubezpieczeń nie dysponowały bazami danych o skali zjawisk szkodowych, będących przedmiotem tych ubezpieczeń.

Dostępność danych do obliczeń stóp składek utrudniają:
- ustawa o ochronie danych osobowych,
- ustawa o działalności ubezpieczeniowej z 22.05.2003 r. (Dz.U. nr 124),
- trudno dostępne dane z wyroków sądowych, dotyczących spraw odpowiedzialności cywilnej, zawodowej,
- praktycznie całkowicie niedostępne dane o szkodach danych grup zawodowych, będące w ich posiadaniu lub w ich stowarzyszeniach.

W sumie zakłady ubezpieczeń, a w szczególności ich działy aktuarialne są w bardzo trudnej sytuacji, aby policzyć stopę składki i dźwignąć ciężar odpowiedzialności za jej wysokość i skutki. Niestety, żadne z rozporządzeń nie pozwala na uzyskanie niezbędnych danych od określonych grup zawodowych podlegających obowiązkowemu ubezpieczeniu OC. Ponadto każde z nich ma ściśle określoną datę obowiązywania. Stąd presja czasu powoduje, że nie jesteśmy w stanie przeprowadzić ani obiektywnych badań, ani ocenić ryzyka ubezpieczanego przedmiotu w relatywnie krótkim czasie po wprowadzeniu rozporządzeń. W tej sytuacji możemy skorzystać jedynie z tzw. prawdopodobieństwa subiektywnego (PS).

Celem tego opracowania jest uzyskanie odpowiedzi na pytania: Na ile jesteśmy w stanie wykorzystać prawdopodobieństwo subiektywne przy kalkulacji stóp składek? Jakie zagrożenia mogą się pojawić dla zakładu ubezpieczeń przy stosowaniu prawdopodobieństwa subiektywnego w kalkulacji stóp składek?

1. Sposoby określenia prawdopodobieństwa subiektywnego (PS)

Rozkwit badań nad zastosowaniem prawdopodobieństwa subiektywnego nastąpił w latach 60-80 XX w. Liczne badania przeprowadzili wówczas m.in. L.J. Savoge, G.F. Pitz, D. Kahneman, Z. Shapira, D. Winterfeldt, A.M. Murphy [cyt. za 10, s. 269]. Badania prawdopodobieństwa subiektywnego ukierunkowano na postrzeganie i pomiar niepewności.

Definicja

Twórcą prawdopodobieństwa subiektywnego był de Finetti (1980), który zdefiniował je następująco: „Przypuśćmy, że pewien osobnik jest zobligowany do oceny stawki p, za którą byłby gotów zapłacić dowolną sumę S („dodatnia lub

ujemna"), w zależności od wystąpienia zdarzenia E za wygranie sumy $p \cdot S$, liczbę p ustalamy jako miarę prawdopodobieństwa przypisaną przez tego osobnika wystąpienia zdarzenia E" [3, s. 152]. Jak łatwo zauważyć, Finetti wyraża przekonanie przez typowanie zakładów niezależnie od rodzaju zdarzeń, które dotyczą.

T. Cyrul mówi, że „[…] głównym celem prawdopodobieństwa subiektywnego jest przywrócenie intuicyjnego znaczenia pojęcia prawdopodobieństwa jako stopnia przekonania (ang. degree of belief)" [3, s. 151]. Jego zdaniem jest to prawdopodobieństwo warunkowe, które można zapisać symbolicznie jako $P[E(C, I)]$, co oznacza, że zdarzenie E wystąpi pod warunkiem C, gdy nasza wiedza o tym fakcie wynosi I. E, C, I są atrybutami otaczającego nas świata, przy czym odnosi się to do stopnia przekonania konkretnego osobnika.

Zdaniem M. Szredera, przez prawdopodobieństwo tego, że jakiś sąd na temat zdarzenia A jest prawdziwy rozumie się stopień pewności lub przekonania danej osoby o prawdziwości tego sądu [13]. Prawdopodobieństwo subiektywne jest traktowane jako personalistyczne i warunkowe. Prawdopodobieństwo personalistyczne ma tę zaletę, że może być stosowane do zdarzeń częstych i rzadkich. Prawdopodobieństwo warunkowe jest obliczane przy przyjętych z góry założeniach.

De Finetti i Ramuscy wskazali na podstawową wadę prawdopodobieństwa subiektywnego, tj. aksjomat przeliczalnej addytywności. Dowodzi się jednak, że „[…] aksjomat przeliczalnej addytywności często jest utożsamiany z matematyczną konwencją przydatną w dowodzeniu ważnych twierdzeń rachunku prawdopodobieństwa i znajduje uzasadnienie na gruncie subiektywnej interpretacji prawdopodobieństwa" [4, s. 134].

Empiryczne badania Marksa (1995) i Irwina (1953) dowiodły, że przy dokonywaniu wyborów ludzie zawyżają prawdopodobieństwa wyników pozytywnych i zaniżają prawdopodobieństwa wyników negatywnych. Prawdopodobieństwo subiektywne najczęściej odnosi się do przekonania konkretnego osobnika.

A. Tverskij i Kahneman PS łączą z pojęciem „dostępności" (availability) „zdarzenia, dla których łatwo można podać podobne przypadki jest oceniane jako bardziej prawdopodobne niż zdarzenie, dla którego jest to trudne"* [cyt. za 11, s. 66].

Wise'a przyjmuje, że badani podają PS na podstawie podobieństw między zdarzeniami.

G. De Zeeuw oraz W.A. Wagenaar wyróżniają prawdopodobieństwa osobiste (PO) jako stopień pewności idealnego, racjonalnego człowieka spełniającego warunek wewnętrznej zgodności, czyli konsystencji. Autorzy ci wskazują na różnice między PS i PO.

* Zjawisko to jest opisywane w literaturze jako optymizm, myślenie życzeniowe, a także określane jako zależność między atrakcyjnością wyniku a oceną jego prawdopodobieństwa.

Punktem wyjścia w określaniu prawdopodobieństw subiektywnych są próby radzenia sobie z niepewnością wobec braku danych obiektywnych. Mamy tu do czynienia z niepewnością możliwych zdarzeń co do miejsca, czasu i okoliczności. Poczucie niepewności rejestrowane jest wewnątrz osoby i uzewnętrzniane w postaci przekonań i wygłaszanych opinii. Prawdopodobieństwo subiektywne nie ma jednoznacznej interpretacji. Interpretacje oparte są na sądach określonych postaci: „stopień wiary", „stopień przekonania", „szansa sukcesu", „szansa straty". Są to określenia powszechnie używane w mowie potocznej, psychologii, socjologii i podejmowaniu decyzji. Prawdopodobieństwo subiektywne jest to ocena zwykle słowna, lecz niekiedy w postaci liczby na skali 0-1 lub 0-100 wyrażająca stopień pewności odczuwalnej przez realne, żywe osoby" [14, s. 231].

2. Metody szacowania prawdopodobieństwa subiektywnego

Według J. Peno, można wyróżnić bezpośrednie i pośrednie metody pomiaru prawdopodobieństwa subiektywnego [10]. Do bezpośrednich metod można zaliczyć m.in.: metodę kolejnych przybliżeń, metodę kwantylową, metodę wiarygodności, metodę Shapiny. Metody te są szeroko omawiane i stosowane w literaturze i praktyce. Do metod pośrednich należy zaliczyć metody: koła prawdopodobieństwa, urnowa, drzewa zdarzeń i innych.

Dotychczasowa praktyka funkcjonowania ubezpieczonych od odpowiedzialności cywilnej wskazuje, że mamy do czynienia z nielicznymi przypadkami popełnienia błędów. Stąd prawdopodobieństwo ich wystąpienia jest rzadkie. W takich przypadkach oszacowanie prawdopodobieństwa subiektywnego jest bardziej skomplikowanie. Mimo wszystko praktyka oczekuje także na oszacowanie takiego prawdopodobieństwa. Zdaniem J. Peno, w tych przypadkach można wykorzystać metody drzew wadliwości, a także scenariuszy. Obie te metody wymagają dostarczenia szczegółowych danych o możliwych prawdopodobieństwach i przyczynach strat.

Różnym zjawiskom przyporządkowuje się wagi decyzyjne odwzorowujące postawy psychologiczne, które ludność przekształca w prawdopodobieństwa. Wagi decyzyjne niekoniecznie spełniają zasady rachunku prawdopodobieństwa. Wielu analityków badających postawy i oceny dokonywane przez różne grupy ludności zauważyło, że ludność „zawyża prawdopodobieństwo wyniku pozytywnego, a zaniża prawdopodobieństwo wyniku negatywnego" [11, s. 66]. Próby wyjaśnienia tego zjawiska doprowadziły do idei wag konfiguralnych, które dążyły do

porównywania tych ocen. Oznaczało to w rezultacie, że ocena prawdopodobieństwa zależy od pozycji danego wyniku wśród innych wyników, tzn. czy to jest np. wynik pozytywny czy też negatywny. Zauważono także – co potwierdza praktyka – że im dalej od punktu odniesienia (tzn. od daty powstania negatywnego zjawiska), tym mniejsza wrażliwość, czyli następuje działanie „prawa zmniejszającej się wrażliwości".

3. Zasady stosowania prawdopodobieństwa subiektywnego do kalkulacji stóp składek

Przy kalkulacji stóp składek w obowiązkowych ubezpieczeniach odpowiedzialności cywilnej są z góry ustalone ogólne warunki ubezpieczenia. Warunków tych nie można modyfikować, gdyż są to ustalenia rządowe lub parlamentarne, mające na celu ochronę określonych instytucji i grup zawodowych. Jedyną wielkością obowiązkowo podawaną przez ustawodawcę jest minimalna suma gwarancyjna określająca górną granicę odpowiedzialności zakładu ubezpieczeń. Są to ubezpieczenia o jednorocznym okresie ochrony ubezpieczeniowej.

Powstaje pytanie, w jaki sposób użyć prawdopodobieństwa subiektywnego, aby zakład ubezpieczeń wywiązał się z narzuconego na niego obowiązku i nie poniósł strat? Odpowiedź na to pytanie nie należy do łatwych, gdyż szkody z tytułu zawartych umów ubezpieczenia mogą się pojawiać parę lat po zawarciu umowy ubezpieczenia.

Zasada pierwsza to wykorzystanie wiedzy na temat ubezpieczanego produktu. Literatura polska i zagraniczna, zwłaszcza opracowania specjalistyczne, dostarczają wielu cennych informacji, które mogą być przydatne przy ocenie ryzyka ubezpieczanego. Stawiamy tutaj tezę, że im więcej wiemy na temat nieznanego nam ryzyka ubezpieczanego, tym mniejszy możemy popełnić błąd przy kalkulacji stóp składek. Ciągłe pogłębianie wiedzy o przedmiocie ubezpieczenia może się przyczynić do zmniejszenia marginesu niepewności, a tym samym wzrostu precyzyjności obliczeń aktuarialnych.

Druga zasada to oszacowanie prawdopodobieństwa subiektywnego. Stawiamy tezę, że w warunkach braku informacji, zwłaszcza statystycznych, żadna wartość prawdopodobieństwa subiektywnego nie może zostać uznana za w pełni wiarygodną. A więc z góry należy zakładać popełnienie błędu przy kalkulacji stopy składki. Błąd może polegać m.in. na:

– pominięciu istotnych danych,
– złej interpretacji danych,

– niewzięciu pod uwagę opinii eksperta,
– błędnej metodologii pomiaru,
– nieanalizowaniu wyroków sądowych, rozstrzygających sytuacje konfliktowe na linii ubezpieczony–zakład ubezpieczeń.

Trzecia zasada to postępowanie ostrożnościowe. Możemy tutaj postępować na dwa sposoby:

1) zawieranie kontraktów ubezpieczeniowych na okresy krótsze niż jeden rok; stawiamy tutaj tezę, że ubezpieczenia krótkoterminowe są bardziej bezpieczne i opłacalne,
2) porównywanie między ustaloną stopą składki a rzeczywistą osiągniętą, tzn. która została osiągnięta w drodze realizacji ubezpieczeń; w tym przypadku powinniśmy dążyć, aby różnica między tymi wielkościami była jak najmniejsza.

Zarówno w pierwszym, jak i drugim przypadku musimy dokładnie (na podstawie akt szkodowych) analizować zbierane składki, jak i każdą powstałą szkodę (przyczyny, proces likwidacji, procedury wyceny itp.). Nie bez znaczenia jest tutaj obserwacja wyniku rachunku technicznego, który zawiera m.in. parametr rezerw techniczno-ubezpieczeniowych. Rachunek ten powinien być sporządzany na koniec każdego miesiąca po to, aby można było na bieżąco obserwować ekonomiczną stronę ubezpieczonego ryzyka.

Przez pryzmat wyniku na rachunku technicznym i znanych wskaźnikach techniczno-ekonomicznych powinniśmy obserwować adekwatność stóp składek ustalonych w przeszłości. Ważne jest, aby przy konstrukcji taryf ubezpieczeniowych zastosować odpowiednio skonstruowane franszyzy, udziały własne, a także zwyżki i zniżki. Parametry te mogą skutecznie chronić zakład ubezpieczeń przed stratami.

4. Możliwości wykorzystania wiedzy ekspertów i firm konsultingowych do szacowania prawdopodobieństwa subiektywnego

W celu oszacowania PS możemy wykorzystać wiedzę ekspertów. Pod pojęciem „ekspertów" będziemy rozumieć specjalistów-praktyków z wieloletnim doświadczeniem, posiadających znajomość w ocenie danego zjawiska, którzy mają dostatecznie dużą wiedzę w przedmiocie objętym ochroną ubezpieczeniową. Nie negując wiedzy eksperta należy podchodzić do oszacowania przez niego PS z ostrożnością. Wynika to z faktu, że są to osoby mogące się mylić w swoich sądach (ocenach i opiniach). Do najczęściej popełnianych błędów przez ekspertów zalicza się:

– odchyłki, gdy np. ekspert wykazuje duże rozbieżności w ocenie prawdopodobieństwa zdarzenia,
– brak koherencji, tj. odstępstwa od formalizmu teorii prawdopodobieństwa,
– brak kalibracji, tj. zgodności szacunków z rzeczywistością [1].

W naszym przypadku będziemy zakładać, że eksperci postępują w kategoriach zachowania racjonalnego.

Warto tutaj zwrócić uwagę, że w przypadku zadania ekspertowi pytania o częstość występowania zjawiska może on mieć trudności w oszacowaniu. Stąd przy wyborze eksperta trzeba zwrócić uwagę na stopień znajomości eksperta z zakresu rachunku prawdopodobieństwa i statystyki matematycznej. Wiedza ta z pewnością mogłaby zmniejszyć skalę błędów popełnianych przez niego przy szacowaniu PS.

W celu ograniczenia niepewności przy stosowaniu prawdopodobieństwa subiektywnego możliwe jest wykorzystanie wiedzy firmy konsultingowej. Konsulting za Międzynarodowym Biurem Pracy rozumiany jest jako „[…] zestaw usług świadczonych przez niezależną i wykwalifikowaną osobę lub osoby w zakresie rozpoznania i badania problemów dotyczących polityki, organizacji procedur i metod, zlecania podjęcia stosownego działania usprawniającego oraz udzielania pomocy we wdrażaniu rekomendowanych rozwiązań" [2, s. 23]. Konsulting może także dotyczyć doradztwa w szeroko rozumianych finansach, w tym ubezpieczeniach. Zadaniem firm konsultingowych jest pomoc klientowi w rozwiązaniu problemu. Jak dotychczas, rzadko można spotkać w firmach konsultingowych osoby zajmujące się doradztwem w zakresie kalkulacji składek zarówno w ubezpieczeniach na życie, jak i majątkowych. Stąd przy korzystaniu z usług firmy konsultingowej trzeba dokonać rozpoznania odnośnie do kompetencji*, zaufania i wiarygodności świadczonych usług. Innymi słowy, trzeba podchodzić z ostrożnością do uzyskanych wyników.

5. Możliwości wykorzystania intuicji przy szacowaniu stóp składek

Człowiek najczęściej szacuje prawdopodobieństwa bardzo niedokładnie, intuicyjnie. Jak wynika z definicji określenia prawdopodobieństwa subiektywnego, w wyrażaniu opinii towarzyszy nam intuicja. Każdy człowiek ma intuicję wynikającą z doświadczenia, treningu, obserwacji. Pod pojęciem „intuicji" będziemy rozumieć „twórczość w pomysłach i kreatywności [...], momentu olśnienia [...], nagłym zwrocie myśli w kierunku rozwiązania analizowanego problemu [8].

* Pojęcie „kompetencji" zostało szczegółowo omówione w pracy [9].

Zdaniem M. Laszczuka, intuicja to „forma bezpośredniego poznania, którego wiarygodność jest zawarta w nim samym i nie wymaga dowodów" [7, s. 63]. Intuicja m.in.:

– jest naturalną częścią myślenia człowieka,
– występuje w warunkach braku wiedzy przy podejmowaniu decyzji, traktowana jest jako uzupełnienie wiedzy,
– należy do kategorii subiektywnych,
– subiektywnie postrzega otoczenie przedsiębiorstw.

Wyniki badań przeprowadzonych przez tego autora dowodzą m.in. że:

– kobiety mają większą intuicję niż mężczyźni,
– im krótszy staż pracy, tym częściej sięgamy po intuicję,
– intuicja skraca czas podejmowania decyzji,
– wobec otaczającej nas lawiny informacyjnej częściej sięgamy do intuicji,
– im bardziej przewidywalne otoczenie, tym chęć wykorzystania intuicji jest mniejsza.

W warunkach braku informacji o ubezpieczonym przedmiocie możemy sięgać do intuicji. Postępowanie to możliwe jest do wykorzystania przez kierownictwo zakładu ubezpieczeń, działy oceny ryzyka i działy aktuarialne. Stosowanie intuicji przy szacowaniu stopy składki może być trafne albo chybione, zawodne. Jeśli zastosowanie stopy składki będzie chybione może doprowadzić do strat, które będzie trudno pokryć dochodami z innych ubezpieczeń, dlatego posługiwanie się intuicją musi być bardzo ostrożne i rozważne.

Wyżej wymienione spostrzeżenia i czynniki oddziaływania można odnieść do decyzji ostatecznych, jakie muszą podjąć zarządy zakładów ubezpieczeń przy ustalaniu taryf ubezpieczeniowych w omawianych rodzajach ubezpieczeń. Skorzystanie z intuicji – jak się wydaje – jest konieczne, ale trzeba mieć na uwadze dodatkowe utrudnienia. W przypadku ubezpieczeń odpowiedzialności cywilnej zawodowej możemy mieć do czynienia z tzw. długim ogonem ryzyka. W praktyce np. błędy popełnione przez projektanta z zakresu budownictwa mogą się ujawnić dopiero po wielu latach eksploatacji, np. katastrofa hali w Katowicach. Stąd kalkulacja stóp składek w oparciu o prawdopodobieństwo subiektywne może być obarczona dużym błędem. Należy także zauważyć, że w przypadku ubezpieczeń OC zawodowej bardzo trudno jest oszacować maksymalną możliwą szkodę [15, s. 16-17]. Wprawdzie mamy minimalną sumę gwarancyjną, jednak znajomość maksymalnej możliwej szkody jest istotna dla celów kalkulacji stopy składki. Wykorzystanie i w tym przypadku określenia prawdopodobieństwa subiektywnego powstania maksymalnej szkody może być obarczone dużym błędem.

Ważna dla celów kalkulacji stóp składek jest znajomość tzw. triggerów zdarzenia [6]. Triggery zdarzenia odnoszą się do zagadnień procesu szkodowego i odpowiadają na pytania: kiedy szkoda powstała?, kiedy została ujawniona?, kiedy została zgłoszona?, kiedy została zlikwidowana? Każde z tych pytań zawiera trudności w określeniu i dochodzeniu w czasie likwidacji szkód. Dla potrzeb kalkulacji stóp składek ważne jest, że ze zdarzeniami określonymi jako triggery spotykamy się w praktyce ubezpieczeniowej.

Warto w tym miejscu zwrócić uwagę na regulacje Kodeksu cywilnego. W myśl art. 442, roszczenie o naprawienie szkody wyrządzonej czynem niedozwolonym ulega przedawnieniu z upływem 3 lat od dnia, w którym poszkodowany dowiedział się o szkodzie i o osobie obowiązanej do jej naprawienia. Termin ten nie może być dłuższy niż 10 lat od dnia, w którym nastąpiło zdarzenie wywołujące szkodę. W świetle takich ustaleń kalkulacja stóp składek w przypadku pokrycia roszczeń spornych w OC zawodowej jest bardzo utrudniona. Stosowanie więc prawdopodobieństwa subiektywnego przy kalkulacji stóp składek na tak odległe okresy realizacji zobowiązań jest bardzo problematyczne.

Podsumowanie

Przedstawiony na wstępie problem prawdopodobieństwa subiektywnego pokazuje złożoność problematyki. W warunkach skrajnego braku danych o ubezpieczonym ryzyku użycie prawdopodobieństwa subiektywnego budzi wątpliwości. Zarówno działy aktuarialne, jak i zarządy zakładów ubezpieczeń powinny dążyć do ograniczenia jego stosowania. W naszej ocenie lepiej jest wydać środki na badania empiryczne niż ponosić straty w przyszłości wynikające z błędu kalkulacji stopy składki. Rozważnie także trzeba się sugerować opiniami rzeczoznawców, ekspertów i firm konsultingowych. Bardzo ostrożnie należy stosować intuicję w podejmowaniu ostatecznych decyzji o wysokości taryf ubezpieczeniowych. Wynika to z bardzo złożonej problematyki odpowiedzialności cywilnej zawodów zaufania publicznego.

Aby uniknąć trudności w kalkulacji stóp składek konieczna jest nowelizacja rozporządzeń, która powinna pójść w kierunku powiększenia dostępu do danych o dotychczasowych szkodach różnych ubezpieczonych grup zawodowych. Leży to w interesie zakładów ubezpieczeń i samych ubezpieczonych.

Podjęta problematyka nie została wyczerpana, lecz jedynie zasygnalizowana, konieczne są dalsze uwagi i polemiki.

Literatura

1. Brandawski A., Estymacja prawdopodobieństwa subiektywnego w modelowaniu ryzyka, „Problemy Eksploatacji" 2005, nr 3.
2. Chruścicki Z., Konsulting w zarządzaniu, Polska Fundacja Promocji Kadr, Warszawa 1997.
3. Cyrul T., Analiza czynników racjonalizujących transfer ryzyka w podziemnym zakładzie górniczym, Prace Instytutu Mechaniki Górotworu, PAN 1-4/2006.
4. Dziurosz-Serafinowicz P., Subiektywne prawdopodobieństwo i problem addytywnej przeliczalności, „Filozofia Nauki" 2009, nr 1.
5. Kościelniak M., Analiza porównawcza sum gwarancyjnych zakresu odpowiedzialności oraz wyłączeń odpowiedzialności w obowiązkowych umowach ubezpieczenia odpowiedzialności cywilnej, Monitor Ubezpieczeniowy, czerwiec 2011.
6. Królikowska J., Triggery w OC zawodowej, „Miesięcznik Ubezpieczeniowy" 2008, nr 7-8.
7. Laszczuk M., Intuicja w podejmowaniu decyzji strategicznych, ATH, Bielsko Biała 2010.
8. Mielus M., Wykorzystanie intuicji w działalności konsultingowej, w: Konsulting. Uwarunkowania i perspektywy rozwoju w Polsce, red. M. Ćwiklicki, M. Jabłoński, Poligrafia ITS, Kraków 2009.
9. Niewiadomski P., Stern K., Model kluczowych kompetencji współczesnego konsultanta-doradcy rolniczego, w: Konsulting. Uwarunkowania i perspektywy rozwoju w Polsce, red. M. Ćwiklicki, M. Jabłoński, Poligrafia ITS, Kraków 2009.
10. Peno J., Analiza porównawcza wybranych metod estymacji prawdopodobieństwa subiektywnego, „Przedsiębiorczość i Zarządzanie", T. XII, zesz. 10.
11. Sokołowska J., Optymizm a prawidłowości percepcyjne przy wyrażaniu prawdopodobieństwa, „Prakseologia" 2006, nr 146.
12. Sukiennik P., Brokerzy i trigger „zdarzenia", „Miesięcznik Ubezpieczeniowy" 2008, nr 7-8.
13. Szreder M., Od klasycznej do częstościowej i personalistycznej interpretacji prawdopodobieństwa, „Wiadomości Statystyczne" 2004, nr 8.
14. Zeeuw G. De, Wagenaar W.A., Czy prawdopodobieństwa subiektywne są prawdopodobieństwami, „Prakseologia" 1974, nr 3.
15. Żyra Z., Kowalczyk K., Pan od ryzyka, „Miesięcznik Ubezpieczeniowy" 2009, styczeń.

USEFULNESS OF SUBJECTIVE PROBABILITY TO CALCULATE PREMIUMS IN THE OBLIGATORY CIVIL PROFESSIONAL INSURANCE

Summary

In the years 2003-2009 the several acts of law concerning obligatory insurance of various professional groups were created. Insurances protect these groups especially from errors that occur during performing work. The article presents the problem of difficulties in calculating the contribution rates as a result of lack of data. But in order to calculate the rate of contribution the article attempts to use of subjective probability.

Michał Stachura
Barbara Wodecka
Uniwersytet Jana Kochanowskiego w Kielcach

OPTYMALIZACJA PORTFELA Z ZASTOSOWANIEM KOPULI NIESYMETRYCZNYCH

Wprowadzenie

Analizy dotyczące portfeli inwestycyjnych w sposób szczególny wymagają uwzględnienia współzależności występujących pomiędzy składowymi portfela. Fakt ten ma podstawowe znaczenie z punktu widzenia podejmowania właściwych decyzji ze względu na dwa zasadnicze kryteria, jakie stanowią maksymalizacja zysku i minimalizacja ryzyka. Niewłaściwe wychwycenie zależności pomiędzy składowymi portfela może sprawić, że ryzyko będzie znacząco niedoszacowane.

Jednym z podejść, pozwalającym na ujęcie współzależności między składowymi portfela jest modelowanie za pomocą kopuli. Taka metodologia jest efektywna i dość popularna. Daje się ona stosować zarówno w statycznym podejściu do danych, w którym analizowane są bezwarunkowe stopy zwrotu, jak i w podejściu dynamicznym, uwzględniającym m.in. autokorelacje, czy też warunkową zmienność. Należy jednak zwrócić uwagę, że stosowane w praktyce i opisywane od strony teoretycznej modele bazują jedynie na kopulach symetrycznych, a do tego w zdecydowanej większości przypadków na archimedesowych. Mimo że modele te dają dość satysfakcjonujące efekty, to jednak ich istota stoi w sprzeczności z faktem, że składowe portfela (na ogół) nie mogą być traktowane jako wzajemnie zastępowalne elementy. Wobec tego symetria wektora losowego jest tu rozumiana jako symetria kopuli (w sensie wymienności jej argumentów) łączącej współrzędne tego wektora. Oznacza to, że wektor $(Y_1, \ldots, Y_k)$ jest symetryczny, jeżeli wektory $(F_1(Y_1),\ldots. F_k(Y_k))$ i $(F_{\tau(1)}(Y_{\tau(1)}),\ldots. F_{\tau(k)}(Y_{\tau(k)}))$ mają ten sam rozkład dla dowolnej permutacji τ, gdzie F_i są dystrybuantami brzegowymi.

W związku z powyższym w opracowaniu zaproponowano odejście od modelowania zależności między rozkładami brzegowymi przy użyciu kopuli symetrycznych na rzecz modelowania kopulami niesymetrycznymi, które służą opisowi zależności między brzegowymi rozkładami stóp zwrotu pojedynczych walorów tworzących portfel. W tym kontekście celem opracowania jest prezentacja możliwości efektywnego zastosowania kopuli niesymetrycznych do optymalizacji portfela. Prezentacja ta jest prowadzona na podstawie przykładowego, arbitralnie wybranego portfela dwuskładnikowego.

Od kilku lat w literaturze zaczyna się odnotowywać fakt, że modele bazujące na kopulach symetrycznych są nieadekwatne. Prezentowane przez autorów ujęcia wywodzą się z grubsza z dwóch nurtów. Pierwszy uwypukla konieczność stosowania kopuli o różnych dolnym i górnym współczynnikach zależności ekstremalnych. Drugi dotyczący tzw. modelu kaskady kopuli dwuwymiarowych, stanowi próbę oddania specyficznych zależności wielowymiarowych [1]. Ujęcia te jednak nie w pełni i nie wprost realizują istotę nadmienionej asymetrii.

1. Optymalizacja portfela

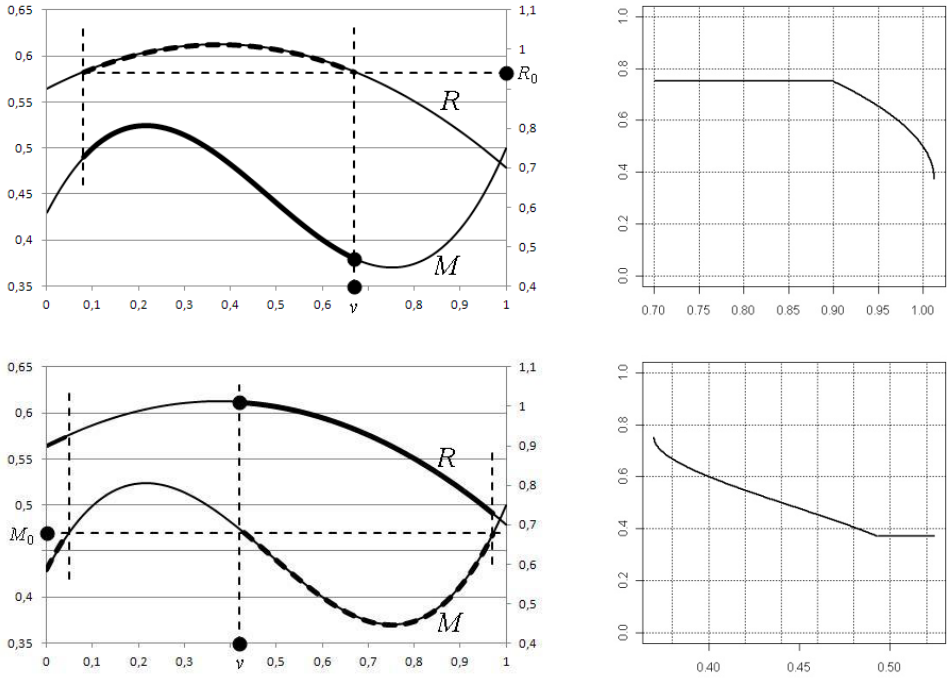
Opis stosowanej w opracowaniu metodologii zaczniemy od prezentacji istoty optymalizacji portfela złożonego z walorów $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(k)}$, którą można ująć w następujący sposób*:

Załóżmy, że znane są oczekiwana stopa zwrotu z portfela $R(\boldsymbol{v})$ i pewna przyjęta miara ryzyka $M(\boldsymbol{v})$ dla wszystkich możliwych udziałów $\boldsymbol{v} = (v_1,\ldots, v_k)$ ($\sum_{i=1}^{k} v_i = 1$, $v_i \geq 0$) walorów pod względem ilości w portfelu. Wówczas albo przy zadanej najmniejszej dopuszczalnej oczekiwanej stopie zwrotu R_0 wyznacza się skład portfela $\boldsymbol{v}$ minimalizujący miarę ryzyka (ujęcie optymalizacyjne ozn. (O1)), albo przy zadanej największej dopuszczalnej wartości miary ryzyka M_0 wyznacza się skład portfela $\boldsymbol{v}$ maksymalizujący oczekiwaną stopę zwrotu (ujęcie optymalizacyjne ozn. (O2)).

Graficzną ilustrację istoty obu ujęć optymalizacji w przypadku portfela dwuskładnikowego ($k = 2$, $\boldsymbol{v} = (v, 1 - v)$) prezentują wykresy zamieszczone w lewej kolumnie na rys 1. Natomiast w prawej kolumnie tego rysunku umieszczone są dodatkowo wykresy odzwierciedlające zależności udziału v pierwszej składo-

* Przedstawione tu ujęcie optymalizacji jest bardzo ogólne, a szczegółowe podejścia można znaleźć np. w pracach [4; 7; 9].

wej portfela odpowiednio od R_0 ($v = v(R_0)$, w ujęciu (O1)) bądź od M_0 ($v = v(M_0)$, w ujęciu (O2)).



Rys. 1. Ilustracja graficzna sposobu optymalizacji portfela dla umownych przebiegów funkcji R i M*

W dalszym ciągu rozważane są opisane pokrótce miary ryzyka oraz metoda szacowania oczekiwanej stopy zwrotu z portfela.

Po pierwsze, jako miary M przyjęte są wartość zagrożona (VaR) i warunkowa wartość zagrożona (CVaR). Dla prostoty miary te są rozważane jako stosowne parametry rozkładu stóp zwrotu z portfela, w szczególności VaR jest rozumiane jako kwantyl tego rozkładu, a nie odpowiadająca temu kwantylowi strata na wartości portfela. Ponadto VaR i CVaR są wyznaczane z perspektywy krótkiej pozycji – odnoszą się zatem do prawych ogonów rozkładu stóp zwrotu. Za metodę estymacji tych miar – w prezentowanym dalej przykładzie empirycznym – przyjęto metodę symulacji Monte Carlo, a uzyskane tak oszacowania miar przedstawione zostaną przez pryzmat wykresów będących analogonami wykresów z prawej kolumny z rys. 1.

* Wykresy z pierwszej kolumny zostały wykonane w programie MS Excel, natomiast wykresy z drugiej kolumny, jak również wszystkie pozostałe wykresy zamieszczone w opracowaniu – w środowisku *R*.

Po drugie, stosowaną procedurę szacowania oczekiwanej stopy zwrotu z portfela można określić następująco: dla dowolnego składu portfela odtwarzane są dzienne realizacje jego stóp zwrotu w określonym horyzoncie czasowym, których średnia przyjmowana jest jako oczekiwana stopa zwrotu na kolejny dzień.

W dalszym ciągu, bez straty ogólności rozważań, analizowany jest jedynie przypadek portfela dwuskładnikowego*.

2. Kopule niesymetryczne

Spośród wielu możliwości opisanych w literaturze, w opracowaniu wykorzystany jest szczególny przypadek ogólniejszej konstrukcji dwuwymiarowych kopuli niesymetrycznych zaproponowanej w pracy [2]. Zgodnie z wynikami tej pracy (wnioski 3 i 4, s. 387-388), kopulę niesymetryczną można utworzyć z dwóch dowolnych kopuli A i B za pomocą m.in. funkcji: $f(x) = x^{\alpha}$ $g(x) = x^{\beta}$ o parametrach $\alpha, \beta \in (0,1)$. Jeśli bowiem $\alpha \neq 1/2$ lub $\beta \neq 1/2$, to kopula dana wzorem:

$$C_{\alpha,\beta}(x,y) = A(x^{1-\alpha}, y^{1-\beta}) \cdot B(x^{\alpha}, y^{\beta}) \tag{1}$$

jest niesymetryczna.

W opracowaniu za kopulę A przyjęta jest archimedesowa kopula Claytona $C_{\gamma}(x,y) = (x^{-\gamma} + y^{-\gamma} - 1)^{-1/\gamma}$, z parametrem $\gamma > 0$. Natomiast za kopulę B przyjęto tzw. kopulę przetrwania Claytona, czyli: $\hat{C}_{\gamma}(x,y) = x + y - 1 + C_{\gamma}(1-x, 1-y)$, która nie jest archimedesowa. Dobór typu tych kopuli jest arbitralny. Użycie kopuli przetrwania podyktowane jest faktem, że kopula Claytona ma niezerowy dolny współczynnik zależności ekstremalnych (równy $2^{-1/\gamma}$) i zerowy górny, zaś w przypadku kopuli przetrwania wartości tych współczynników są zamienione. Dzięki temu w kopuli utworzonej na podstawie wzoru (1), a wyrażonej jako:

$$\begin{aligned} C_{\alpha,\beta,\gamma,\delta}(x,y) &= C_{\gamma}(x^{1-\alpha}, y^{1-\beta}) \cdot \hat{C}_{\delta}(x^{\alpha}, y^{\beta}) = \\ &= (x^{(\alpha-1)\gamma} + y^{(\beta-1)\gamma} - 1)^{-1/\gamma} \cdot \\ &\cdot (x^{\alpha} + y^{\beta} - 1 + ((1-x^{\alpha})^{-\delta} + (1-y^{\beta})^{-\delta} - 1)^{-1/\delta}) \end{aligned} \tag{2}$$

oba współczynniki zależności ekstremalnych są niezerowe, choć nie dają się one w prosty sposób wyrazić wzorami w zależności od parametrów $\alpha, \beta \in (0,1)$ i $\gamma, \delta > 0$.

* Przypadek więcejwymiarowy wymaga jedynie użycia innych metod konstrukcji kopuli niesymetrycznych.

W prezentowanym dalej przykładzie empirycznym do estymacji parametrów kopuli danej wzorem (2), a następnie do symulacji prób według tak uzyskanego rozkładu stosowano metody ogólne [3; 8]. Ponadto z uwagi na brak wzorów analitycznych stosownych funkcji pomocniczych nieuniknione okazały się metody numeryczne*.

3. Metoda Monte Carlo

Niech dane będą notowania $X_{1,t}$, $X_{2,t}$, dla $t \in \{1, \ldots, T\}$, walorów $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ oraz zrealizowane stopy zwrotu $R_{1,t}$, $R_{2,t}$ tych walorów. Przy tych oznaczeniach implementacja metody Monte Carlo (MC) do szacowania miar ryzyka (VaR i CVaR) portfela dwuskładnikowego przeprowadzona jest w sposób następujący:

1. Na podstawie zrealizowanych stóp zwrotu $R_{1,t}$, $R_{2,t}$ estymowane są (metodą największej wiarogodności) parametry rozkładów brzegowych, za które przyjęto dwustronne uogólnione rozkłady Pareto.
2. Następnie (przy użyciu numerycznej metody największej wiarogodności) estymowane są parametry kopuli łączącej powyższe rozkłady brzegowe, danej wzorem (2).
3. W kolejnym kroku zgodnie z rozkładem łącznym, zadanym przez kopulę i rozkłady brzegowe, symulowana jest próba liczebności T odtwarzająca realizacje $\widetilde{R}_{1,t}$, $\widetilde{R}_{2,t}$ stóp zwrotu składowych portfela.
4. Na podstawie tak otrzymanej próby odtwarzane są realizacje notowań obu walorów $\widetilde{X}_{1,t}$, $\widetilde{X}_{2,t}$. Z uwagi jednak na możliwość swobodnego zadania wartości początkowych ciągów notowań przyjęto zasadę równości median, tzn. $\mathrm{med}(\widetilde{X}_{1,t}) = \mathrm{med}(X_{1,t})$ i $\mathrm{med}(\widetilde{X}_{2,t}) = \mathrm{med}(X_{2,t})$.
5. Próba $\widetilde{X}_{1,t}$, $\widetilde{X}_{2,t}$, przy zadanym udziale v pierwszego waloru, służy do wyznaczenia ciągów: $\widetilde{Z}_t(v)$, $\widetilde{R}_t(v)$ odtwarzających realizacje wartości oraz realizacje stóp zwrotu stosownego portfela.
6. Z tak otrzymanych ciągów $\widetilde{R}_t(v)$ wyznacza się oszacowania $\mathrm{VaR}(v)$ i $\mathrm{CVaR}(v)$.

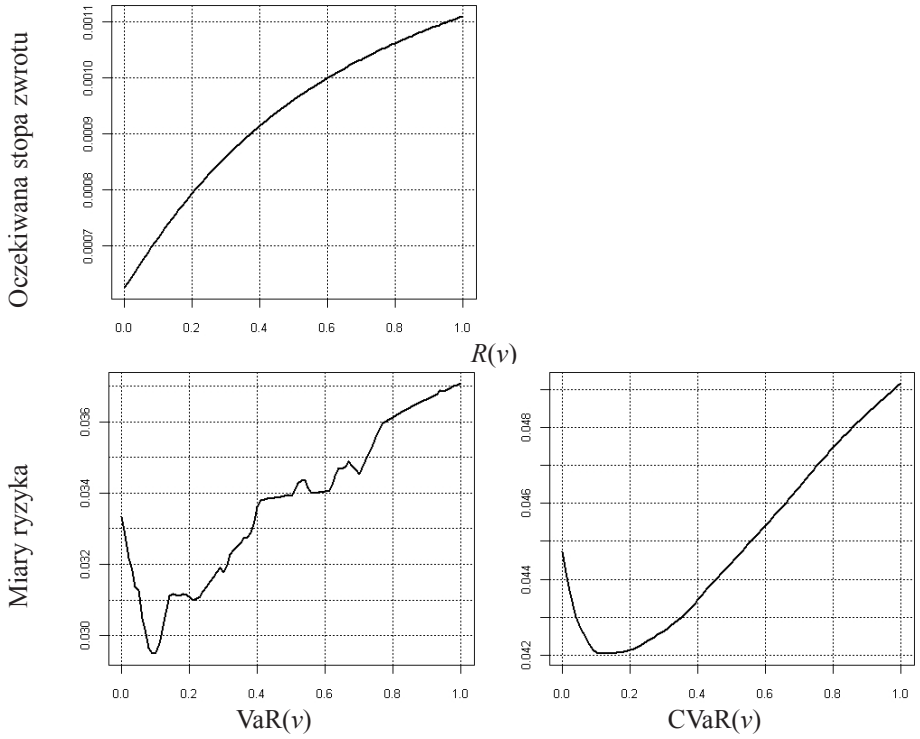
Kroki 3-6 powtarzane są 1000 razy, a jako ostateczne oszacowania $\mathrm{VaR}(v)$ i $\mathrm{CVaR}(v)$ przyjmowane są mediany z wartości uzyskanych w poszczególnych repetycjach.

* Wszystkie obliczenia zostały wykonane w środowisku *R*.

4. Przykład empiryczny

Przykładowym analizom poddany jest portfel tworzony z akcji KGHM i PKN Orlen na 30.09.2011 r. z horyzontem czasowym wynoszącym 1 dzień roboczy. Szacowanie oczekiwanej stopy zwrotu bazuje na $T = 500$ ostatnich realizacjach* stopy zwrotu dla wszystkich możliwych proporcji udziałów walorów w portfelu zadanych przez v – stanowiące udział KGHM – zmieniające się od 0 do 1 z krokiem 0,001.

Dla każdego v z podanego zakresu, zgodnie z wcześniej opisaną metodologią, wyznaczone zostały oczekiwane stopy zwrotu ($R(v)$) oraz miary ryzyka (VaR(v) i CVaR(v)), które przedstawiają odpowiednie wykresy na rys. 2.

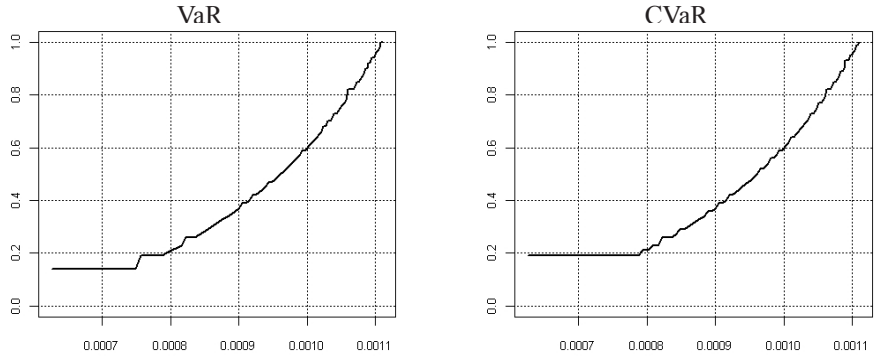


Rys. 2. Stopy zwrotu i miary ryzyka z portfela uzyskane metodą MC (w zależności od v – udziału KGHM)

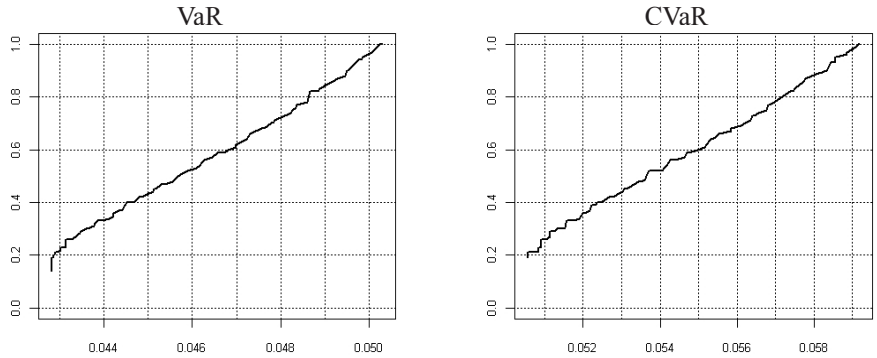
Na tej podstawie wyznaczone są udziały v optymalizujące analizowany portfel w rozumieniu metod (O1) i (O2). Zależność udziałów v od najmniejszej do-

* Wobec tego podstawą wszelkich obliczeń są dzienne notowania na zamknięcie kursów akcji obu spółek z okresu 8.10.2010-30.09.2011 (po 500 obserwacji) z GPW w Warszawie.

puszczalnej oczekiwanej stopy zwrotu R_0 (w sensie (O1)) prezentują wykresy z rys. 3, natomiast zależność udziałów v od największej dopuszczalnej wartości miary ryzyka M_0 (w sensie (O2)) – wykresy z rys. 4.



Rys. 3. Skład portfela $v(R_0)$ uzyskany metodą (O1)



Rys. 4. Skład portfela $v(M_0)$ uzyskany metodą (O2)

Zaprezentowany przykład empiryczny pozwala wnioskować, że:

– Szacowanie miar ryzyka (VaR, CVaR) poprzez odtwarzanie wektora stóp zwrotu metodą MC choć jest złożone metodologicznie (przez co wymaga znacznych „nakładów obliczeniowych"), to jednak daje się efektywnie stosować.
– Uzyskane oszacowania stóp zwrotu oraz miar ryzyka dla różnych składów portfela pozwalają w prosty sposób zoptymalizować portfel.

Podsumowanie

Oprócz wniosków wynikających bezpośrednio z zaprezentowanego przykładu zastosowania kopuli niesymetrycznej należy uwypuklić fakt ogólniejszy – – dotyczący istoty asymetrii (w sensie zaprezentowanym w opracowaniu) i jej uwzględnienia w modelu. Mianowicie, stosowanie kopuli niesymetrycznych jest bardziej zasadne merytorycznie ze względu na specyfikę zależności między składowymi portfela. Niniejsze opracowanie jest swego rodzaju przyczynkiem do dyskusji nad potrzebą, a zarazem możliwością zastosowań kopuli niesymetrycznych, które poszerzają perspektywę spojrzenia na modelowanie zależności między rozkładami brzegowymi. Obszar ten w zakresie aplikacyjnym jest wciąż zbyt mało eksplorowany.

Wobec powyższego, w kontekście przedstawionego przykładu empirycznego perspektywa dalszych badań powinna objąć chociażby takie aspekty, jak: rozszerzenie analiz na portfele wielowymiarowe; dynamiczna weryfikacja uzyskiwanych wyników względem pojawiających się kolejnych realizacji; porównanie z innymi metodami.

Literatura

1. Doman R., Zastosowania kopuli w modelowaniu dynamiki zależności na rynkach finansowych, Uniwersytet Ekonomiczny, Poznań 2011.
2. Durante F., Construction of non-exchangeable bivariate distribution functions, „Stat Papers" 2009, No. 50.
3. Heilpern S., Funkcje łączące, WAE, Wrocław 2007.
4. Jajuga K., Jajuga T., Inwestycje. Instrumenty finansowe, aktywa finansowe, ryzyko finansowe, inżynieria finansowa, Wydawnictwo Naukowe PWN, Warszawa 2006.
5. Kuziak K., Koncepcja wartości zagrożonej VaR (Value at Risk), StatSoft, 2003.
6. Liebscher E., Construction of asymmetric multivariate copulas, „J. Multivariate Anal." 2008, No. 99.
7. Markowitz H., Portfolio selection. Efficient diversification of investments, John Wiley & Sons, Nowy Jork 1959.
8. Nelsen R.B., An Introduction to Copulas, Springer, New York 1999.

9. Tarczyński W., Rynki kapitałowe. Metody ilościowe, t. 2, Placet, Warszawa 1997.
10. Tsay R. S., Analysis of Financial Time Series, Wiley, Hoboken 2005.
11. R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

APPLICATION OF ASYMMETRIC COPULAS IN PORTFOLIO OPTIMIZATION

Summary

This paper is some kind of a discussion about both the necessity and possibility of asymmetric copula applications. Presented deliberation is settled in the context of financial portfolio analysis that, in a specific way, requires taking the correlations of the component assets into consideration, which creates an opportunity for asymmetric copula implementation. Mentioned issues are exemplified by real two-asset portfolio optimization.

Henryk Gurgul
Tomasz Wójtowicz
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

ZALEŻNOŚCI DŁUGOOKRESOWE ZMIENNOŚCI STÓP ZWROTU I WIELKOŚCI OBROTÓW NA GPW W WARSZAWIE

Wprowadzenie

Jednym z głównych zagadnień dotyczących rynku akcji jest badanie procesu kształtowania się cen. Rozważania teoretyczne mające na celu wyjaśnienie obserwowanych własności stóp zwrotu, np. występowanie grubych ogonów, pojawiają się już od połowy ubiegłego wieku. Jedną z ważniejszych hipotez dotyczącą procesu kształtowania się cen jest hipoteza o mieszance rozkładów (Mixture Distribution Hypothesis), zgodnie z którą całkowita zmiana cen akcji w ciągu dnia jest sumą zmian cen akcji dokonujących się pod wpływem pojawiających się na rynku nowych informacji. Ponieważ przybliżoną miarą tempa napływu informacji jest wolumen, oznacza to, że dzienna wariancja stóp zwrotu (mierząca ich zmienność) oraz wolumen wspólnie zależą od nieobserwowalnego procesu napływu informacji oraz mają zbliżone do niego własności. Ponieważ na rynek stale napływają informacje o zróżnicowanym znaczeniu i czasie oddziaływania ich akumulacja powoduje m.in. występowanie zależności długookresowych (tzw. długiej pamięci) w szeregach czasowych zmienności i wolumenu, jak również ich długookresową współzależność (m.in. wspólną długą pamięć).

Dotychczasowe wyniki badania długookresowych współzależności zmienności i wolumenu nie są jednoznaczne. Z jednej strony, Bollerslev i Jubinski [1] na podstawie danych z lat 1962-1995 stu spółek z indeksu S&P100 wykazali, że zmienność i wolumen charakteryzują się zbliżonymi wartościami parametrów długiej pamięci (tzw. wspólna długa pamięć). Z drugiej zaś Lobato i Velasco [7] oraz Gurgul i Wójtowicz [3; 4] wykazali na przykładzie spółek z indeksów DJIA

i DAX, że pomimo zbliżonych własności długiej pamięci, pomiędzy zmiennością a wolumenem nie występuje zależność długookresowa (ułamkowa kointegracja). Ponadto, w przypadku spółek z GPW Gurgul i Wójtowicz [5] stwierdzili brak wspólnej długiej pamięci i odmienną długookresową dynamikę zmienności i wolumenu. We wspomnianych wyżej pracach miarą zmienności stóp zwrotu były ich wartości bezwzględne lub kwadraty, jednak, jak pokazują badania ostatnich lat, lepszą miarą zmienności jest zrealizowana wariancja. Miarę tę wykorzystali Fleming i Kirby [2]. Ich wyniki sugerują jeszcze słabszą zależność długookresową niż wynika to z wcześniejszych badań.

W opracowaniu zaprezentowano badanie długookresowych zależności pomiędzy zmiennością stóp zwrotu i wolumenem 20 najbardziej płynnych spółek notowanych na GPW w Warszawie w okresie od stycznia 2005 r. do września 2011 r. W dalszej części zdefiniowano i omówiono własność długiej pamięci, następnie scharakteryzowano dane oraz uzyskane wyniki badań empirycznych.

1. Długa pamięć

Szereg czasowy charakteryzuje się długą pamięcią, jeżeli jego funkcja gęstości spektralnej $f_x(\lambda)$ jest nieograniczona w otoczeniu zera i spełnia warunek:

$$f_x(\lambda) \sim G\lambda^{-2d} \tag{1}$$

gdy $\lambda \to 0^+$. Jeśli parametr długiej pamięci (ułamkowej integracji) d jest dodatni, to powyższy warunek jest równoważny hiperbolicznemu zanikaniu funkcji autokorelacji ρ_k, tzn.:

$$\rho_k \sim c_\rho k^{2d-1} \tag{2}$$

gdy $k \to +\infty$. Jeśli $d < 0{,}5$, to proces jest stacjonarny, natomiast dla $d > 0{,}5$ proces jest niestacjonarny. Własność (2) oznacza, że nawet pomiędzy oddalonymi od siebie obserwacjami występują istotne zależności korelacyjne. Przykładem procesów z długą pamięcią są m.in. procesy ARFIMA.

Jedną z podstawowych metod estymacji parametru d długiej pamięci jest semiparametryczna metoda zaproponowana przez Robinsona [9], tzw. lokalna metoda Whittle'a. Wartość estymatora parametru d jest obliczana na podstawie periodogramu dla m najniższych częstości opisujących długookresowe zachowanie procesu. Właściwy dobór parametru m uwalnia estymator d od wpływu krótkookresowych zaburzeń badanego procesu. Estymator Whittle'a jest zgodny o asymptotycznym rozkładzie normalnym również w przypadku niektórych niestacjonarnych szeregów czasowych.

2. Dane

W przeprowadzonych obliczeniach jako miarę zmienności zastosowano zrealizowaną wariancję obliczoną analogicznie, jak w pracy Fleminga i Kirby [2] jako sumę kwadratów pięciominutowych stóp zwrotu. Ponieważ wykorzystując dane wysokiej częstości należy uwzględnić zjawiska wynikające z mikrostruktury rynku, wprowadzono poprawkę uwzględniającą autokorelację pięciominutowych stóp zwrotu aż do rzędu 3, czyli do opóźnienia 15 minut. Przy konstrukcji zrealizowanej zmienności wykorzystano również stopy zwrotu pomiędzy zamknięciem a otwarciem notowań, tzw. overnight returns [6]. Jako miara wielkości obrotów został zastosowany dzienny wolumen, czyli liczba sprzedanych akcji danej spółki. W celu eliminacji wpływu stałego rozwoju giełdy z logarytmów obu rozważanych zmiennych usunięto składnik deterministyczny w postaci wielomianu trzeciego stopnia.

3. Wyniki empiryczne

W pierwszym etapie zbadana została długa pamięć pojedynczych szeregów czasowych zmienności i wolumenu rozważanych spółek. Uzyskane wartości estymatorów parametru długiej pamięci w całym okresie 2005-2011 oraz w dwóch równolicznych podokresach (1.01.2005 r.-6.05.2008 r. oraz 7.05.2008 r.-1.09.2011 r.) zestawiono w tab. 1. Estymację parametru d przeprowadzono na podstawie periodogramu obliczonego dla $m = 70$ najniższych częstości. Wartość ta odpowiada okresowi o długości co najmniej 24 dni. W ten sposób uzyskano ocenę parametru d opisującego dynamikę średnio- i długookresową. Wszystkie wartości estymatorów w tab. 1 są istotnie większe od zera i istotnie mniejsze od 1. Natomiast w części przypadków są większe od 0,5, co wskazuje na ewentualną niestacjonarność danych.

Aby możliwe było poprawne wnioskowanie na podstawie uzyskanych wyników przeprowadzony został również test pozornej długiej pamięci (spurious long memory), tzn. występowania istotnej wartości parametru długiej pamięci spowodowanej jednak zmianami strukturalnymi, np. zmianą w poziomie średniej, a nie zależnościami autokorelacyjnymi. Test pozornej długiej pamięci zaproponowany przez Shimotsu polega na podziale całej próby na k podprób, a następnie na weryfikacji równości parametrów długiej pamięci zarówno w podpróbach, jak i w całym rozważanym okresie.

Tabela 1

Długa pamięć zmienności i wolumenu

	Zmienność			Wolumen		
	01.05-09.11	01.05-05.08	05.08-09.11	01.05-09.11	01.05-05.08	05.08-09.11
ACP	0,339	0,133	0,454	0,226	0,153	0,418
BHW	0,446	0,296	0,533	0,487	0,611	0,427
BRE	0,387	0,077	0,440	0,309	0,260	0,462
BRS	0,416	0,395	0,427	0,458	0,578	0,361
BSK	0,257	0,292	0,240	0,373	0,461	0,373
BZW	0,497	0,459	0,504	0,478	0,356	0,548
GTC	0,375	0,114	0,520	0,522	0,527	0,337
GTN	0,607	0,576	0,591	0,322	0,220	0,454
IPX	0,350	0,291	0,398	0,513	0,550	0,503
KGH	0,486	0,478	0,449	0,360	0,433	0,323
MCI	0,487	0,416	0,446	0,513	0,599	0,485
MIL	0,449	0,424	0,407	0,286	0,255	0,346
NET	0,481	0,624	0,417	0,481	0,567	0,475
PBG	0,394	0,354	0,444	0,453	0,328	0,283
PEO	0,568	0,491	0,536	0,289	0,439	0,258
PKN	0,512	0,473	0,451	0,297	0,132	0,385
PKO	0,409	0,297	0,417	0,385	0,390	0,380
PXM	0,505	0,380	0,516	0,614	0,644	0,614
TPS	0,472	0,552	0,427	0,144	0,049	0,293
TVN	0,443	0,314	0,416	0,485	0,542	0,461

Ponieważ rozważane szeregi czasowe są stosunkowo krótkie, to test pozornej długiej pamięci został przeprowadzony z podziałem na dwa równoliczne podokresy wyszczególnione wcześniej. Uzyskane wyniki wskazują, że dla wszystkich rozważanych spółek wolumen charakteryzuje się istotną długą pamięcią, natomiast w przypadku zmienności trzech spółek (ACP, BRE, GTC) na poziomie istotności 0,05 można odrzucić hipotezę o równości parametrów długiej pamięci w podokresach i całym okresie, co wskazuje, że wykryta długa pamięć może być pozorna.

Warto w tym miejscu podkreślić istotną różnicę w zachowaniu się rynków w rozważanych podokresach. Pierwszy z nich obejmuje czas hossy poprzedzającej kryzys na rynkach finansowych. Drugi obejmuje już czas samego kryzysu i wychodzenia z niego. Jeżeli zatem w rozważanych szeregach mogłaby wystąpić jakaś zmiana strukturalna skutkująca pozorną długą pamięcią, to przy takim podziale na podokresy powinna być ona najbardziej widoczna.

Na podstawie wyników przeprowadzonego testu można przyjąć, że w badanych szeregach czasowych występuje długa pamięć. W związku z tym ma sens badanie czy zmienność i wolumen mają podobną dynamikę długookresową. Pierwszym etapem jest zbadanie czy ich parametry ułamkowej długiej pamięci są równe, czy też różnią się istotnie. Jest to warunek konieczny występowania ułamkowej integracji i istnienia wspólnej składowej długookresowej.

Do testowania hipotezy zerowej o równości parametrów długiej pamięci zmienności i wolumenu wykorzystana zostanie statystyka [8]:

$$T = \frac{\sqrt{m}(\hat{d}_1 - \hat{d}_2)}{\sqrt{\frac{1}{2}\left(1 - \frac{\hat{G}_{12}^2}{\hat{G}_{11}\hat{G}_{22}}\right) + \frac{1}{\ln(n)}}} \tag{3}$$

gdzie:

$\hat{d}_1$ i $\hat{d}_2$ – odpowiednie estymatory parametrów długiej pamięci,

$\hat{G}_{ij}$ – elementy estymatora dwuwymiarowej macierzy gęstości spektralnej w 0.

Przy założeniu braku kointegracji statystyka T ma asymptotyczny rozkład $N(0,1)$. Wyniki przeprowadzonych testów wraz z wyestymowaną wartością wspólnego parametru długiej pamięci zestawiono w tab. 2. W większości przypadków nie ma podstaw do odrzucenia hipotezy zerowej o wspólnej wartości parametru długiej pamięci zmienności i wolumenu. Hipoteza ta została odrzucona jedynie w przypadku sześciu spółek: GTN, IPX, MIL, PEO, PKN i TPS. Wyniki te wskazują, że w porównaniu do wcześniejszych badań [5] nastąpiła zmiana zależności długookresowych pomiędzy wolumenem i zmiennością stóp zwrotu akcji spółek notowanych na GPW w Warszawie.

Występowanie wspólnego parametru długiej pamięci jest warunkiem koniecznym występowania ułamkowej kointegracji. Pojęcie to jest analogiczne do tradycyjnej definicji kointegracji, np. w przypadku procesów $I(1)$. Dwa procesy z długą pamięcią (tzn. $X_t \sim I(d_1)$ i $Y_t \sim I(d_2)$) są ułamkowo zintegrowane, jeżeli:

– $d_1 = d_2 = d$, tzn. mają ten sam stopień ułamkowej integracji,
– istnieje $a \in \mathbf{R}$, takie, że $X_t + aY_t \sim I(b)$ i $b < d$, tzn. istnieje kombinacja liniowa procesów X_t i Y_t, której parametr długiej pamięci jest mniejszy niż parametr długiej pamięci procesów X_t i Y_t.

Tabela 2

Wspólna długa pamięć i ułamkowa kointegracja

	Wspólna długa pamięć		Ułamkowa kointegracja		
	d_{wsp}	T	wartości własne		liczba relacji kointegrujących
			w_1	w_2	
ACP	0,279	-1,23	1,36	0,64	0
BHW	0,466	0,41	1,12	0,88	0
BRE	0,348	-0,79	1,28	0,72	0
BRS	0,437	0,46	1,47	0,53	0
BSK	0,314	1,21	1,29	0,71	0
BZW	0,488	-0,20	1,21	0,79	0
GTC	0,448	1,46	1,21	0,79	0
GTN	0,464	-3,23***	1,52	0,48	0
IPX	0,434	1,78*	1,37	0,63	0
KGH	0,423	-1,42	1,60	0,40	0
MCI	0,500	0,32	1,67	0,33	1
MIL	0,367	-1,74*	1,42	0,58	0
NET	0,481	0,00	1,36	0,64	0
PBG	0,418	0,48	1,33	0,67	0
PEO	0,428	-3,05***	1,48	0,52	0
PKN	0,404	-2,40**	1,48	0,52	0
PKO	0,395	-0,30	1,48	0,52	0
PXM	0,556	1,05	1,29	0,71	0
TPS	0,308	-3,30***	1,30	0,70	0
TVN	0,464	0,42	1,21	0,79	0

*, **, *** – istotność na poziomie 0,1, 0,05 i 0,01.

Badanie ułamkowej kointegracji, podobnie jak w przypadku klasycznym, polega na analizie wartości własnych i rzędu macierzy gęstości spektralnej G w 0 (lub alternatywnie macierzy korelacji P). Procedura ustalania liczby relacji kointegrujących jest następująca [10]:

Niech $\hat{\delta}_i$ ($i = 1,2$) będą wartościami własnymi estymatora $\hat{P}$ macierzy korelacji ustawionymi w kolejności malejącej. Wówczas liczbę relacji kointegrujących γ można wyznaczyć jako:

$$r = \operatorname{argmin}\{L(0), L(1)\} \tag{4}$$

gdzie:

$$L(u) = m^{-0,15}(2-u) - \sum_{i=1}^{2-u} \hat{\delta}_i \tag{5}$$

W tab. 2 zestawiono wyniki odpowiednich obliczeń również dla spółek, dla których hipoteza o wspólnej długiej pamięci została odrzucona. Ponieważ w większości przypadków żadna z wartości własnych macierzy korelacji nie jest bliska 0 oznacza to, że macierz gęstości spektralnej dwuwymiarowego procesu zmienności i wolumenu jest pełnego rzędu, czyli nie występuje relacja kointegrująca te dwie zmienne. Jedynym wyjątkiem są akcje MCI, dla których zastosowana procedura pozwoliła na wykrycie ułamkowej kointegracji pomiędzy zmiennością i wolumenem. Oznacza to, że tylko w przypadku akcji tej spółki możemy mówić o wspólnej długookresowej dynamice wolumenu i zmienności.

Podsumowanie

W opracowaniu zaprezentowano wyniki badania wspólnej długiej pamięci wolumenu i zmienności stóp zwrotu najbardziej płynnych spółek notowanych na GPW w Warszawie w latach 2005-2011. Ich celem była praktyczna weryfikacja hipotezy o mieszance rozkładów, z której wynika podobna długookresowa dynamika wolumenu i zmienności. Uzyskane wyniki świadczą o istnieniu wspólnej długiej pamięci zmienności i wielkości obrotów. Jednocześnie wskazują na różną długookresową dynamikę tych zmiennych i brak wspólnej składowej długookresowej. Świadczą o tym wyniki badania ułamkowej kointegracji. Uzyskane w trakcie badania rezultaty są analogiczne do występujących w literaturze wyników badań rozwiniętych rynków akcji [np. 7].

Literatura

1. Bollerslev T., Jubinski D., Equity trading volume and volatility: latent information arrivals and common long-run dependencies. "Journal of Business and Economic Statistics" 1999, T. 17.
2. Fleming J., Kirby B., Long memory in volatility and trading volume. "Journal of Banking and Finance" 2011, No. 35(7).
3. Gurgul H., Wójtowicz T., Long Memory on German Stock Exchange, "Finance a úvěr – Czech Journal of Economics and Finance" 2006, No. 56 (9-10).

4. Gurgul H., Wójtowicz T., Długookresowe własności wolumenu obrotów i zmienności cen akcji na przykładzie spółek z indeksu DJIA, „Badania Operacyjne i Decyzje" 2006, nr 3-4.
5. Gurgul H., Wójtowicz T., Długa pamięć wolumenu. Porównanie giełd w Warszawie i Frankfurcie, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 462 (6), Szczecin 2007.
6. Hansen P., Lunde A., A realized variance for the whole day based on intermittent high-frequency data, "Journal of Financial Econometrics" 2005, No. 3.
7. Lobato I.N., Velasco C., Long memory in stock-market trading volume, "Journal of Business and Economic Statistics" 2000, No. 18(4).
8. Nielsen M.O., Shimotsu K., Determining the cointegration rank in nonstationary fractional systems by the exact local Whittle approach, "Journal of Econometrics" 2007, No. 141.
9. Robinson P.M., Gaussian semiparametric estimation of long range dependence, "Annals of Statistics" 1995, No. 23.
10. Robinson P.M., Yajima Y., Determination of cointegrating rank in fractional systems, "Journal of Econometrics" 2002, No. 106 (2).

LONG-RUN DEPENDENCIES IN RETURN VOLATILITY AND TRADING VOLUME ON WSE

Summary

In the paper common long-term dynamics of return volatility and trading volume of the largest companies listed on Warsaw Stock Exchange in 2005-2011 is examined. The existence of contemporaneous relationship between volatility and volume is implied by the Mixture Distribution Hypothesis which states that volatility and trading volume are jointly generated by information flow process. In the study realized volatility computed on the basis of high frequency data is used as a measure of return volatility. It is more efficient measure of daily return volatility than commonly used absolute or squared returns.

Tomasz Zdanowicz
Uniwersytet Mikołaja Kopernika w Toruniu

ZMIENNOŚĆ MOMENTÓW WYŻSZYCH RZĘDÓW NA RYNKACH FINANSOWYCH*

Wprowadzenie

Dynamiczny rozwój chińskiej gospodarki obserwowany od lat 80. XX w. coraz częściej zwraca uwagę na ten kraj jako kandydata na lidera handlu światowego. Wzrost gospodarczy w ciągu ostatnich lat utrzymujący się na ponad 9% poziomie, ogromny rynek pracy i potężny strumień inwestycji zagranicznych, pomimo rosnącej inflacji sprawiają, iż gospodarka Chin stała się drugą gospodarką świata zaraz za gospodarką USA. Ten gwałtowny rozwój nie byłby możliwy bez wsparcia dwóch potężnych giełd w Szanghaju i Hongkongu oraz dużych zmian w prawie. Od 2001 r. stopniowo otwierano rynek finansowy dla instytucji zagranicznych. Chiński fenomen stanowi obecnie przedmiot zainteresowania szerokiego grona potencjalnych inwestorów oraz badaczy (Osińska, Zdanowicz [5] czy Fałdziński, Osińska, Zdanowicz [2]), szczególnie w latach kryzysu finansowego. Ciekawe rezultaty uzyskano w ostatniej ze wspomnianych prac, gdzie prezentowane były wyniki testów przyczynowości w sensie Grangera dla średniej, wariancji i ryzyka między szeregami stóp zwrotu głównych indeksów i walut światowych oraz analogicznymi szeregami dla Chin. Zarówno testy Pierce'a i Haugha, Cheunga i Ng, testy oparte na wielowymiarowych modelach GARCH wskazały na występowanie zależności między analizowanymi szeregami. Trochę gorzej na tym tle prezentowały się wyniki testu Honga, Liu i Wanga przyczynowości w ryzyku. Uzyskane wyniki stanowiły inspirację do dalszych badań nad powiązaniami między kursami walut i wartościami indeksów, które zostaną przeprowadzone przy zastosowaniu modeli pozwalających na opis zmienności całego rozkładu warunkowego. W tym celu można zastosować modele zmiennej warunkowej funkcji gęstości (ARCD – Autoregressive Conditional Density). Modele te pozwalają na modelowanie zmienności

* Publikacja finansowana ze środków na naukę w ramach grantu MNiSW N N111 328839.

całego rozkładu warunkowego przy stosunkowo prostej parametryzacji. Szeroki jest też wybór możliwych do zastosowania rozkładów warunkowych. Modelowanie rozkładu warunkowego szybko znalazło zastosowanie przy opisie zmienności asymetrii i kurtozy warunkowej, a także przy wyznaczaniu wartości narażonej na ryzyko (VaR) i oczekiwanego niedoboru (Expected Shortfall – ES).

Celem niniejszego opracowania jest prezentacja wyników wstępnych badań dotyczących modelowania momentów warunkowych wyższych rzędów dla wybranych szeregów. Analiza stanowi część większego badania mającego dać odpowiedź na wpływ gospodarki chińskiej na gospodarki światowe.

1. Autoregresyjne Modele Warunkowej Funkcji Gęstości

Klasa modeli ARCD została zaproponowana przez Hansena w 1994 r. jako pewnego rodzaju rozszerzenie modeli GARCH na parametry kształtu i asymetrii rozkładu warunkowego. Hansen nie precyzował konkretnego równania dla każdego z parametrów, przyjął, że a priori nie ma żadnej informacji o kształcie oraz zmienności rozkładu warunkowego w czasie, stąd należy wybrać odpowiednio elastyczny rozkład – skośny t-Studenta, a także zdefiniować bardzo ogólne równania dla parametrów tego rozkładu. W przypadku parametru asymetrii i kształtu objaśniającej w modelu dla bonów skarbowych zastosowano równanie uwzględniające opóźnienia reszt i zmiennej objaśniającej, ich iloczyn, a także kwadraty. W przypadku modelu dla kursu wymiany dolara do franka szwajcarskiego zastosowano tylko opóźnione reszty i ich kwadraty. Uzyskane przez Hansena wyniki potwierdzały tezę o zmienności parametrów kształtu i asymetrii rozkładów warunkowych. Należy jednak mieć na uwadze, że taki rezultat dotyczył tylko dwóch szeregów. Krótki przegląd literatury związanej z modelowaniem rozkładu warunkowego oraz asymetrii i kurtozy warunkowej można znaleźć w pracy [1].

Ze względu na ograniczenie w postaci posiadanej bazy o stosunkowo małej liczbie obserwacji zdecydowano o wykorzystaniu w pracy dwóch najczęściej spotykanych w modelowaniu finansowych szeregów czasowych rozkładów – skośnego t-Studenta, według specyfikacji Hansena (STD) oraz skośnego uogólnionego rozkładu błędu (SGED) według parametryzacji Theodossiou [6]. Oba rozkłady mają parametr asymetrii zawierający się w przedziale (-1,1) oraz parametr kształtu. W pracy wprowadzono ograniczenia dla oryginalnych parametrów λ i κ, tak aby możliwa była estymacja modelu ARCD z wybranymi równaniami. I tak dla parametru λ przyjęto transformację logistyczną daną równaniem 1 oraz następujące równania opisujące zmienność parametru – równania 2 do 6:

$$\lambda_t = -0{,}95 + 1{,}9(1 + \exp(-\lambda_t^*))^{-1} \tag{1}$$

$$\lambda_t^* = \gamma_0 + \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-1}^2 + \gamma_3 \lambda_{t-1}^* \tag{2}$$

$$\lambda_t^* = \gamma_0 + \gamma_1 \varepsilon_{t-1}^3 + \gamma_2 \lambda^{-1} \tag{3}$$

$$\lambda_t^* = \gamma_0 + \gamma_1 I^-(\varepsilon_{t-1})\varepsilon_{t-1} + \gamma_2 I^+(\varepsilon_{t-1})\varepsilon_{t-1} + \gamma_3 \lambda_{t-1}^* \tag{4}$$

$$\lambda_t^* = \gamma_0 + \gamma_1 \varepsilon_{t-1} + \gamma_2 \lambda_{t-1}^* \tag{5}$$

$$\lambda_t^* = \gamma_0 + \gamma_1 \varepsilon_{t-1}^2 + \gamma_2 \lambda_{t-1}^* \tag{6}$$

gdzie $I^-(x) = 1$, gdy $x < 0$ oraz $I^-(x) = 0$, gdy $x >= 0$, $I^+(x) = 1 - I^-(x)$.

W przypadku paramertu kształtu i rozkładu STD wprowadzono tranformację daną wzorem 7 dla zapewnienia określoności rozkładu:

$$\kappa_t = 0{,}05 + 2(1 + \exp(-\kappa_t^*)^{-1} \tag{7}$$

natomiast dla rozkładu SGED odpowiednią transformację zapisano wzorem:

$$\kappa_t = 4{,}1 + 30(1 + \exp(-\kappa_t^*)^{-1} \tag{8}$$

gdzie dodatkowo parametr κ_t^* jest zapisany jednym z równań 9-12:

$$\kappa_t^* = \vartheta_0 + \vartheta_1 \varepsilon_{t-1}^2 + \vartheta_2 \kappa_{t-1}^* \tag{9}$$

$$\kappa_t^* = \vartheta_0 + \vartheta_1 \varepsilon_{t-1}^4 + \vartheta_2 \kappa_{t-1}^* \tag{10}$$

$$\kappa_t^* = \vartheta_0 + \vartheta_1 I^+(\varepsilon_{t-1})\varepsilon_{t-1} + \vartheta_2 I^-(\varepsilon_{t-1})\varepsilon_{t-1} + \vartheta_3 \kappa_{t-1}^* \tag{11}$$

$$\kappa_t^* = \vartheta_0 + \vartheta_1 \varepsilon_{t-1}^4 + \vartheta_2 I^-(\varepsilon_{t-1})\varepsilon_{t-1}^4 + \vartheta_3 \kappa_{t-1}^* \tag{12}$$

Specyfika wspomnianych modeli wymaga również niestandardowych testów do oceny ich jakości. W literaturze najczęściej spotykane są dwa testy: Nybloma oraz ortogonalności odpowiednio zdefiniowanych warunków.

2. Zastosowanie modeli ARCD do opisu zmienności rozkładów warunkowych szeregów czasowych

Do analizy zostały wybrane szeregi stóp zwrotu z indeksów światowych, dla których przeprowadzano badanie przyczynowości w średniej, wariancji i ryzyku. Uzyskane wyniki świadczą o występowaniu różnokierunkowych zależności w średniej, wariancji i ryzyku między indeksami giełd i walut światowych oraz Chin. Dane pochodziły z okresu 01.02.2006-18.02.2011, co łącznie dawało 1325

obserwacji dla każdego z indeksów. W przypadku równania dla średniej i wariancji przyjęto standardowy model ARMA-GARCH. Wyboru opóźnień w modelu GARCH dokonano a priori ustawiając opóźnienia części ARCH i GARCH równe 1. Założenie to wynika z analizy literatury w zakresie modelowania wariancji warunkowej z użyciem modeli GARCH, gdzie w przeważającej części model GARCH (1,1) jest preferowany ponad inne z większymi stopniami opóźnień. W przypadku równania ARMA przyjęto graniczną wartość opóźnienia równą 5. Wszystkie modele były estymowane metodą maksymalizacji funkcji wiarygodności z użyciem algorytmu BHHH. Następnie wyznaczano statystyki testów Nybloma. Modele ARCD zdefinowane w tej pracy nie pozwalają na bezpośrednią ocenę wpływu zmienności szeregów z giełd chińskich na szeregi z innych giełd światowych. Możliwa jest jedynie wstępna ocena zmienności rozkładów warunkowych badanych szeregów czasowych. Dopiero wprowadzenie do wspomnianych równań dla poszczególych parametrów dodatkowej zmiennej objaśniajacej – stóp zwrotu indeksów z giełd w Szanghaju i Hongkongu, umożliwi weryfikację wyników testów przyczynowości. W pierwszej kolejności należy jednak zweryfikować hipotezę o zmienności rozkładów warunkowych.

Ze względu na stosunkowo dużą liczbę estymowanych modeli w tab. 1 zaprezentowano przyjęte w pracy oznaczenia.

Tabela 1

Oznaczenia estymowanych modeli

Model	λ^*	κ^*	Model	λ^*	κ^*	Model	λ^*	κ^*
M_1	const	const	M_11	2	9	M_21	4	11
M_2	2	const	M_12	2	10	M_22	4	12
M_3	3	const	M_13	2	11	M_23	5	9
M_4	4	const	M_14	2	12	M_24	5	10
M_5	5	const	M_15	3	9	M_25	5	11
M_6	6	const	M_16	3	10	M_26	5	12
M_7	const	9	M_17	3	1	M_27	6	9
M_8	const	10	M_18	3	12	M_28	6	10
M_9	const	11	M_19	4	9	M_29	6	11
M_10	const	12	M_20	4	10	M_30	6	12

* const. oznacza, że w danym modelu estymowany parametr nie zmieniał się w czasie, natomiast kolejne numery odwołują się do odpowiednich równań prezentowanych w pracy.

Estymowane były modele z dwoma rozkładami warunkowymi STD oraz SGED. W przypadku tego drugiego napotkano duże problemy estymacyjne, co było spowodowane problemami przy wyznaczaniu pochodnych oraz trudnościa-

mi z osiągnięciem zbieżności w algorytmach optymalizacyjnych. Z tej przyczyny wyniki estymacji modeli z rozkładem SGED nie będą prezentowane.

Tabela 2 zawiera podsumowanie wyników modeli z rozkładem STD. Prezentowane są dwa modele – klasyczny AR-GARCH (M_1) oraz najlepszy z modeli konkurencyjnych. Przedstawione zostały wartości logarytmu wiarygodności (LogL), wartości kryteriów informacyjnych Akaike'a i Shwartza oraz statystyka testu stabilności Nybloma dla całego wektora parametrów. Dla wszystkich szeregów lepsze pod względem logarytmu wiarygodności okazały się modele ze zmienną funkcją gęstości warunkowej. Należy jednak podkreślić, że poprawa wartości LogL po dołożeniu dodatkowych parametrów w większości szeregów nie jest istotna. Podobne wnioski można wyciągnąć analizując wartości kryteriów informacyjnych. Jest to wynik, który wskazuje na brak zmienności parametrów asymetrii i kształtu w rozkładach warunkowych dla badanych szeregów. Należy jednak pamiętać o wielkości próby, która została wybrana do badania, w której znajdowały się szeregi liczące po 1325 obserwacji, podczas gdy Dark [1] wskazuje minimalną próbę do estymacji modeli ARCD – powyżej 3000. Ciekawy jest przypadek indeksu HSI, dla którego lepszy okazał się model oznaczony jako M_26 – stosunkowo proste równanie dla asymetrii i bardziej złożone dla parametru kształtu. Analiza wartości parametrów tego modelu nakazuje odrzucić go jako faworyta dla HSI. Model ten miał bardzo duże co do modułu wartości parametrów, dodatkowo na jego niekorzyść przemawia wartość statystyki testu Nybloma przekraczająca wartość krytyczną przy 1% poziomie istotności. Mając na uwadze wspomniane ograniczenia należy uznać uzyskane wyniki za poprawne, jednak niepreferujące modeli ARCD nad prostsze w estymacji i interpretacji modele z klasy ARMA-GARCH.

Tabela 2

Wyniki estymacji modeli AR-GARCH oraz ARCD

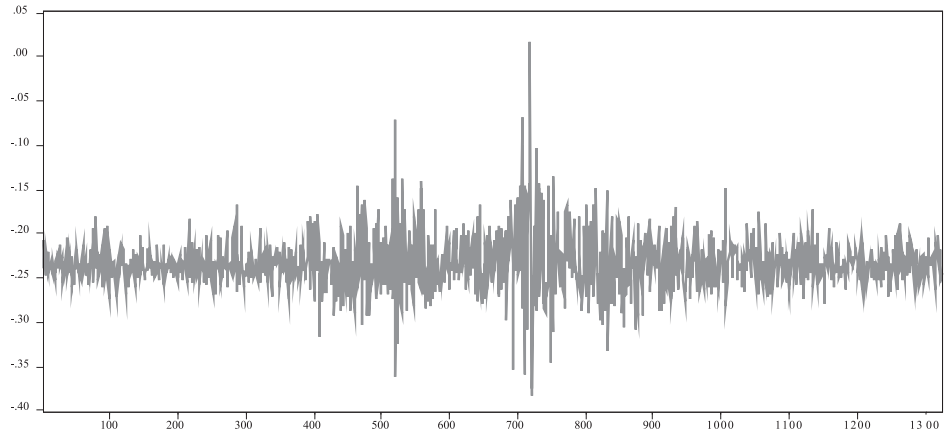
	AEX		AMEX		ATG		ATX	
Model:	M_1	M_22	M_1	M_14	M_1	M_14	M_1	M_14
Logl	3958,11	3965,27	4079,45	4082,48	3618,51	3625,32	3650,7	3654,13
Aic/Bic	-5,96/-5,94	-5,97/-5,93	-6,15/-6,12	-6,15/-6,10	-5,45/-5,42	-5,46/-5,41	-5,50/-5,47	-5,50/-5,46
Nyblom	1,6943	2,2749	1,3057	1,5169	2,6104	3,4743	1,1412	1,6575
	BEL20		BOVESPA		BSESN		BUX	
Model:	M_1	M_14	M_1	M_14	M_1	M_27	M_1	M_22
Logl	4000,05	4007,51	3542,68	3549,77	3645,75	3647,36	3589,07	3591,22
Aic/Bic	-6,03/-6,00	-6,03/-5,99	-5,34/-5,31	-5,34/-5,30	-5,50| -5,47	-5,50/-5,46	-5,41/-5,38	-5,40/-5,36
Nyblom	1,3319	2,0235	0,9582	1,5366	1,8232	2,4675	0,8129	1,5931
	CAC40		DAX		DJCA		DJIA	
Model:	M_1	M_26	M_1	M_18	M_1	M_22	M_1	M_9
Logl	3875,92	3892,95	3955,79	3969,42	4112,35	4118,29	4247,01	4250,46

cd. tabeli 2

Aic/Bic	-5,84| -5,81	-5,86/-5,82	-5,96/-5,93	-5,98/-5,94	-6,20/-6,17	-6,20/-6,16	-6,40/-6,37	-6,40/-6,37
Nyblom	1,5215	1,7805	1,4363	3,6922	1,4265	2,4354	0,9853	2,2553
	FTSE100		HEX		HSI		IPCMEX	
Model:	M_1	M_18	M_1	M_22	M_1	M_26	M_1	M_14
Logl	4059,37	4062,62	3813,88	3817,72	3704,27	3715,68	3894,65	3898,41
Aic/Bic	-6,12/-6,09	-6,12/-6,08	-5,75/-5,72	-5,75/-5,70	-5,58/-5,55	-5,60/-5,55	-5,87/-5,84	-5,87/-5,83
Nyblom	2,04	4,7353	1,0112	1,8478	1,7953	5,4186	1,4628	2,4202
	ISE100		JKSE		KLSE		KOSPI	
Model:	M_1	M_22	M_1	M_22	M_1	M_20	M_1	M_30
Logl	3489,79	3495,96	3822,46	3833,66	4543,74	4544,51	3928,03	3939,57
Aic/Bic	-5,26/-5,23	-5,26/-5,22	-5,76/-5,73	-5,77/-5,73	-6,85/-6,82	-6,85/-6,81	-5,92/-5,89	-5,93/-5,89
Nyblom	1,2351	2,5055	0,9147	2,6673	2,5615	4,9085	0,4993	4,0274
	MERVAL		NIKK225		NSDQCOMP		NZ50	
Model:	M_1	M_14	M_1	M_26	M_1	M_19	M_1	M_22
Logl	3647,08	3650,73	3795,73	3795,84	3970,4	3970,52	4678,34	4683,54
Aic/Bic	-5,50/-5,47	-5,50/-5,45	-5,72/-5,70	-5,72/-5,68	-5,98/-5,96	-5,98/-5,94	-7,05/-7,02	-7,05/-7,01
Nyblom	1,8657	3,1808	1,5434	2,2748	1,9486	3,1647	1,4544	2,1879
	OMXSPI		PX50		RTS		SMSI	
Model:	M_1	M_13	M_1	M_22	M_1	M_5	M_1	M_14
Logl	3879,21	3886,74	3801,18	3805,27	3390,67	3393,53	3872,57	3879,81
Aic/Bic	-5,85/-5,82	-5,85/-5,81	-5,73/-5,70	-5,73/-5,68	-5,11/-5,08	-5,11/-5,08	-5,84/-5,81	-5,84/-5,80
Nyblom	2,1099	2,0837	1,5532	2,4626	1,7385	2,041	2,2231	4,5065
	SP500		SSE		SSMI		STI	
Model:	M_1	M_20	M_1	M_18	M_1	M_18	M_1	M_18
Logl	4188,08	4188,32	3479,14	3484,68	4187,69	4195,93	4027,4	4034,97
Aic/Bic	-6,31/-6,28	-6,31/-6,27	-5,24/-5,21	-5,25/-5,21	-6,31/-6,28	-6,32/-6,28	-6,07/-6,04	-6,08/-6,04
Nyblom	1,9017	1,8557	2,0108	2,9147	0,8503	1,7504	1,5037	2,2791

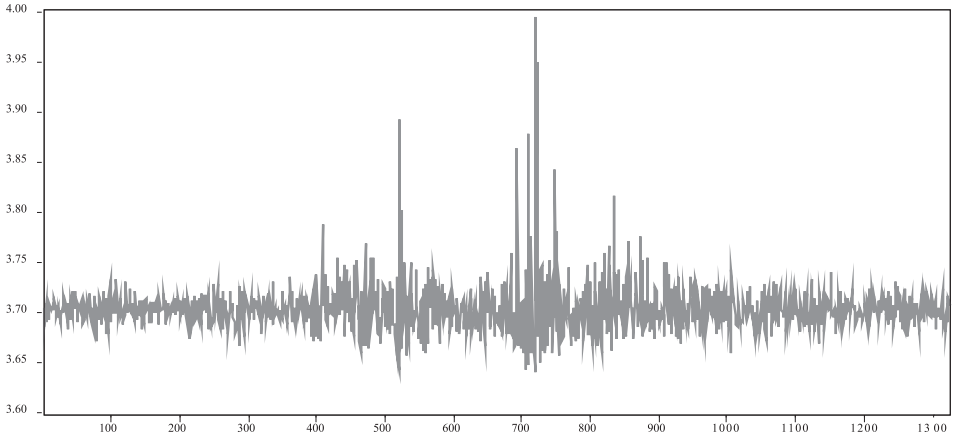
Przeprowadzone badanie miało na celu wstępne stwierdzenie występowania zmienności rozkładu warunkowego przed dalszym pogłębionym badaniem z użyciem wybranych indeksów w roli zmiennych objaśniających – tak jak sugerowałyby to wyniki testów przyczynowości. Dalsze badania będą musiały być przeprowadzone w oparciu o dłuższą próbę, tak aby maksymalnie pominąć efekt zbyt krótkiej próby.

Dla zobrazowania zmienności trzeciego i czwartego momentu warunkowego przedstawione zostały wykresy szeregów asymetrii i kurtozy warunkowej uzyskanych z modelu dla dwóch szeregów HSI i SSE. Wyraźnie odznacza się moment zwiększonej zmienności (około 400 obserwacji – połowa 2007 r.) oznaczający początek kryzysu finansowego. W pierwszej fazie widoczne są dość duże zmiany stóp zwrotu, których zmienność po pewnym czasie stabilizuje się.



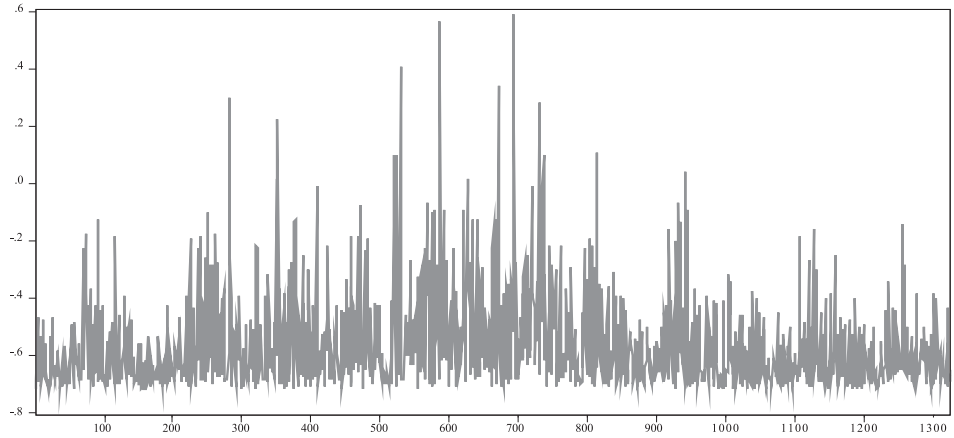
Rys. 1. Asymetria warunkowa z modelu M_26 – szereg HSI

W okresie kryzysu cały rozkład warunkowy HSI ulegał dość dużym zmianom pod wpływem napływających informacji. Po okresie podwyższonej zmienności oba momenty warunkowe uległy stabilizacji.



Rys. 2. Kurtoza warunkowa z modelu M_26 – szereg HSI

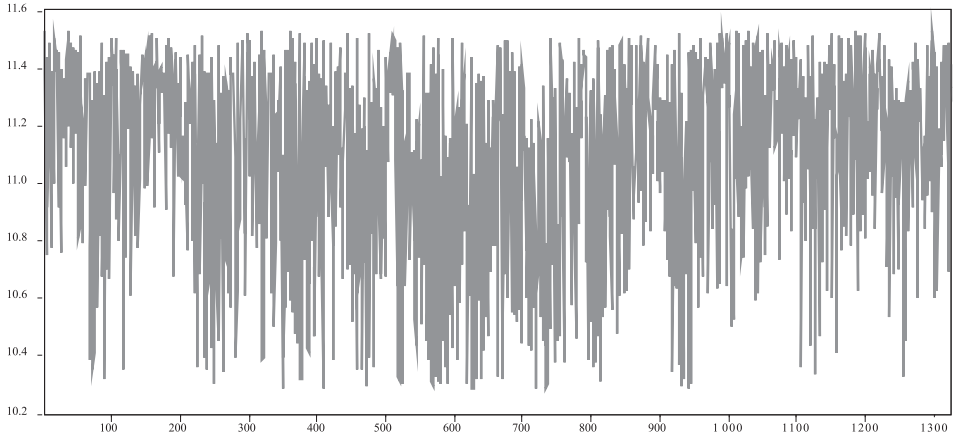
Jednocześnie kurtoza warunkowa utrzymuje się na zaskakująco niskim poziomie – około 3,7, co oznacza, że rozkład warunkowy jest bardziej zbliżony do normalnego niż w rozkładzie stóp zwrotu, w którym kurtoza jest równa 11.



Rys. 3. Asymetria warunkowa z modelu M_18 – szereg SSE

Podobne wnioski można wyciągnąć analizując otrzymane dane dla szeregu stóp zwrotu indeksu giełdy w Szanghaju. Okres dużej zmienności rozpoczyna się od 300 obserwacji (marzec 2007 r.), jednak w tym przypadku przebieg kryzysu jest dużo spokojniejszy, co może mieć związek z regulacjami obowiązującymi w Chinach (w skład SSE wchodzą tylko przedsiębiorstwa chińskie, nie do końca uwolniony kurs juana). Nie ma bardzo dużych gwałtownych zmian i trudno jest wskazać poszczególne okresy analizując tylko przebieg stóp zwrotu.

Rozkład SSE charakteryzuje się za to dość dużą lewostronną asymetrią równą -0,4, kurtoza w rozkładzie jest równa 5,548. Czwarty moment warunkowy jest, podobnie jak w poprzednim przypadku, na zupełnie innym poziomie niż ma to miejsce w rozkładzie stóp zwrotu.



Rys. 4. Kurtoza warunkowa z modelu M_18 – szereg SSE

Uzyskany wynik może być trochę zaskakujący, jednak należy pamiętać, że w przypadku rozkładu STD asymetria zależy tylko od parametru λ, podczas gdy czwarty moment centralny zależy zarówno od λ, jak i parametru kształtu – κ.

Podsumowanie

W pracy zaprezentowano wyniki estymacji autoregresyjnych modeli warunkowej funkcji gęstości dla szeregów stóp zwrotu z 35 indeksów giełd światowych. Badanie będące kontynuacją wcześniejszych prac autora należy traktować jako wstęp do głębszej analizy wpływu gospodarki Chin na główne gospodarki świata (reprezentowane przez indeksy giełdowe oraz kursy walutowe). Dla analizowanych szeregów należy wskazać modele AR-GARCH jako lepsze niż modele ARCD, co sugerowałoby brak zmienności rozkładu warunkowego. W pewnym sensie potwierdzają się wyniki testów przyczynowości w ryzyku, które wskazały na zależności między badanymi szeregami w kwantylach rozkładów warunkowych – sugerują jako przyczynę zmienności rozkładu wpływ innych zmiennych niż uwzględnione w pracy. Przeprowadzone badanie nie jest do końca wiarygodne ze względu na stosunkowo krótką próbę. Wskazuje to jednoznacznie na potrzebę przeprowadzenia kolejnych analiz na odpowiednio dłuższych szeregach, zarówno testów przyczynowości, jak i estymacji modeli ARCD.

Literatura

1. Dark J.G., Estimation of Time Varying Skewness and Kurtosis with an Application to Value at Risk, "Studies in Nonlinear Dynamics and Econometrics" 2010, Vol. 14.
2. Fałdziński M., Osińska M., Zdanowicz T., Econometric Analysis of the Risk Transfer on Capital Markets. A Case of China, Referat wygłoszony na 71. konferencji International Atlantic Economic Society w Atenach, 2011.
3. Hansen B.E., Autoregressive Conditional Density Estimation, „International Economic Review" 1994, Vol. 35.
4. Nyblom J., Testing the Constancy of Parameter over Time, „Journal of the American Statistical Association" 1989, Vol. 84.
5. Osińska M., Zdanowicz T., What Drives Chinese Financial Markets?, w: Financial Markets. Principles of Modelling, Forecasting and Decision-Making,

FindEcon Monograph Series: Advances in Financial Market Analysis, 9, eds. W. Milo, G. Szafrański, P. Wdowiński, Uniwersytet Łódzki, Łódź 2011.
6. Theodossiou P., Skewed Generalized Error Distribution of Financial Assets and Option Pricing, 2000, http://ssrn.com/abstract=219679.

HIGHER ORDER MOMENTS VARIABILITY IN THE FINANCIAL MARKETS

Summary

The dynamic development of Chinese economy observed since the 80s of the twentieth century, increasingly draws attention to this country as a candidate for the leadership of world trade. This rapid growth is supported by two powerful stock exchanges in Shanghai and Shenzhen. Chinese phenomenon is now of interest to a wide audience of potential investors and researchers [2]. Modeling time varying conditional asymmetry or kurtosis becomes more often the subject of analysis, especially during the growing world financial crisis. Autoregressive Conditional Density Models (ARCD), which were first presented in 1994 by Hansen [3], allow for modeling the conditional volatility of the entire distribution with a relatively simple parameterization. There are also extension of the classical model of Hansen on the other conditional distributions, more complex equations of shape and asymmetry parameters of the distribution. ARCD models was used to model the financial data of world stock exchanges in relation to the ranks of China's stock exchange.