MODELOWANIE PREFERENCJI A RYZYKO ’14

Studia Ekonomiczne

ZESZYTY NAUKOWE WYDZIAŁOWE

UNIWERSYTETU EKONOMICZNEGO W KATOWICACH

MODELOWANIE PREFERENCJI A RYZYKO ’14

Redaktor naukowy

Tadeusz Trzaskalik

Katowice 2014

ISSN 2083-8611

Nakład: 210 egzemplarzy

Wersją pierwotną Studiów Ekonomicznych jest wersja papierowa

WYDAWNICTWO UNIWERSYTETU EKONOMICZNEGO W KATOWICACH
ul. 1 Maja 50, 40-287 Katowice, tel.: +48 32 257-76-35, faks: +48 32 257-76-43
www.wydawnictwo.ue.katowice.pl, e-mail: wydawnictwo@ue.katowice.pl

SPIS TREŚCI

WPROWADZENIE

W niniejszym Zeszycie znajdują się prace poświęcone zagadnieniom związanym z ryzykiem w ubezpieczeniach, zarządzaniem ryzykiem projektu, analizą preferencji oraz innymi zastosowaniami ryzyka. Niżej przedstawiono ich krótkie omówienie.

RYZYKO W UBEZPIECZENIACH

W pracy **„Efektywność alternatywnych inwestycji kapitałowych na przykładzie polisy inwestycyjnej”** (M. Hadaś-Dyduch) proponowaną metodę predykcji WIG20 oparto w przeważającym stopniu na transformacie falkowej i dokonano oceny efektów inwestowania.

W opracowaniu **„Koncepcja ubezpieczenia upraw rolnych od szkód spowodowanych przez zwierzynę łowną w polskim obszarze terytorialnym”** (S. Wieteska) dostarczono elementarnej wiedzy z tego zakresu, a także przedstawiono dane statystyczne dotyczące szkód łowieckich oraz metody ich szacowania.

ZARZĄDZANIE RYZYKIEM PROJEKTU

W artykule **„Projektowanie realizacji obiektów budowlanych w warunkach ryzyka”** (S. Biruk, P. Jaśkowski) przedstawiony sposób projektowania zastosowano do opracowania harmonogramu przykładowego przedsięwzięcia, a w celu oceny jakości uzyskanego harmonogramu przeprowadzono badania symulacyjne.

W pracy **„Rozmyte liczby przedziałowe w harmonogramowaniu przedsięwzięć metodą łańcucha krytycznego”** (B. Gładysz) na podstawie podanych przez ekspertów ocen czasów zadań wyestymowano funkcję przynależności czasów zadań i na ich podstawie skonstruowano harmonogram projektu oraz wyznaczono bufory czasu zabezpieczające przed ryzykiem nieukończenia projektu w terminie.

ANALIZA PREFERENCJI

„System rekomendacji doboru wag kryteriów oparty na ich charakterystyce probabilistycznej" (J. Brzostowski, E. Roszkowska) proponuje narzędzie wsparcia negocjatora w procesie analizy preferencji własnych oraz analizy preferencji potencjalnego partnera negocjacyjnego.

W opracowaniu **„Propozycja hybrydy reguł Hurwicza i Bayesa w podejmowaniu decyzji w warunkach niepewności"** (H. Gaspars-Wieloch) zaprezentowano koncepcję procedury pozwalającej wyłonić optymalną strategię czystą za pomocą wzorów uwzględniających nie tylko współczynnik pesymizmu i optymizmu, lecz także specyfikę całego zbioru wypłat.

W artykule **„Opóźnienia w modelach społecznych dynamiki replikatorowej"** (R. Jankowski) omówiono wpływ opóźnień na stabilność wewnętrznego punktu stacjonarnego.

Praca **„Wybrane własności procedury SAW w kontekście wspomagania negocjacji"** (E. Roszkowska, J. Brzostowski) przedstawia propozycje modyfikacji klasycznego algorytmu tej procedury z punktu widzenia możliwości jego zastosowania w procesie negocjacji.

Opracowanie **„Modelowanie preferencji a zagadnienia zrównoważonego (trwałego) rozwoju"** (W. Rybicki) dokonuje selektywnego przeglądu głównych kierunków badań i podstawowych ustaleń z obszaru wyceny i porównań programów długookresowych, a także fluktuacji indywidualnych preferencji podmiotów – w miarę upływu czasu.

W pracy **„Analiza preferencji ludzi młodych w ocenie atrakcyjności miast jako produktu turystycznego"** (J. Szołtysek, G. Trzpiot) omówiono wyniki badania ankietowego związanego z problematyką kształtowania się opinii i poglądów młodzieży dotyczącego powyższego zagadnienia.

W artykule **„Dlaczego w dylemat więźnia warto grać kwantowo?"** (M. Szopa) zdefiniowano kwantowy dylemat więźnia i strategie kwantowe oraz omówiono ich własności, a także przytoczono przykłady zjawisk ekonomicznych (zmowy cenowe, strategia szachowa), które odzwierciedlają równowagi Nasha kwantowego dylematu więźnia.

INNE ZASTOSOWANIA RYZYKA

W opracowaniu **„Porównywanie efektywności prognoz ex post wielkości sprzedaży w pewnym przedsiębiorstwie wyznaczonych za pomocą rozkładu**

gamma i modeli z wahaniami sezonowymi" (M. Doszyń, K. Dmytrów) przeanalizowano szeregi czasowe tygodniowych wielkości sprzedaży i na ich podstawie dokonano porównania prognoz.

„Przykłady zastosowania macierzy migracji w zarządzaniu ryzykiem finansowym" (U. Grzybowska, M. Karwański) to artykuł, w którym przedstawiono przykłady zastosowania tych macierzy w szacowaniu ryzyka kredytowego, w ubezpieczeniach komunikacyjnych typu Bonus-Malus oraz w celu prognozowania zmian wzorców zachowań użytkowników kart kredytowych.

Praca **„Modelowanie zmian zmiennych stanu w modelu dwumianowym do celów wyceny opcji realnych"** (K. Targiel) określa (na podstawie przeszłych zmian) typ procesu stochastycznego modelującego zmiany zmiennej stanu oraz kratę najlepiej pokrywającą przyszłe trajektorie tej zmiennej.

W opracowaniu **„Analiza ryzyka starzenia demograficznego wybranych miast w Polsce"** (G. Trzpiot, A. Ojrzyńska) oceniono lukę demograficzną postrzeganą jako ryzyko funkcjonowania w zdrowiu.

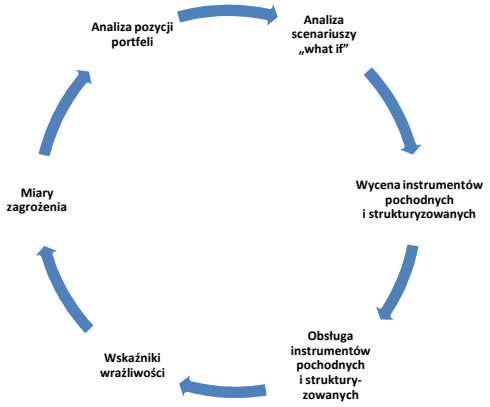
Tadeusz Trzaskalik

Monika Hadaś-Dyduch
Uniwersytet Ekonomiczny w Katowicach

EFEKTYWNOŚĆ ALTERNATYWNYCH INWESTYCJI KAPITAŁOWYCH NA PRZYKŁADZIE POLISY INWESTYCYJNEJ

1. Szczegółowa charakterystyka wycenianego instrumentu

Jednym z kluczowych elementów zarządzania ryzykiem jest wycena instrumentów pochodnych i strukturyzowanych (rys. 1). Wycena taka umożliwia wyznaczenie bieżącej wartości portfela, w której partycypuje wyceniany instrument, jak również oszacowanie rzeczywistej ceny instrumentu przed jego nabyciem na rynku lub bezpośrednio od oferenta.



Rys. 1. Kluczowe funkcjonalności zarządzania ryzykiem*

* **Analiza scenariuszy** *what if* umożliwia aktywną ocenę wpływu potencjalnych zmian czynników rynkowych (m.in. z uwzględnieniem stress testów) i modyfikacji udziałów instrumentów na wartość portfela oraz wartości wyliczanych wskaźników, warunkując tym samym efektywne zarządzanie portfelami, które ma na celu maksymalizację rentowności inwestycji przy jednoczesnej minimalizacji podejmowanego ryzyka. Konstrukcja analiz symulacyjnych daje możli-

Przyjęta do wyceny polisa inwestycyjna jest w formie grupowego ubezpieczenia na życie i dożycie z instrumentem bazowym WIG20. Jest to polisa będąca odzwierciedleniem polis oferowanych na rynku. Okres ubezpieczenia trwa 36 miesięcy (12.03.2009-12.03.2012 r.). Opłata wstępna związana z przystąpieniem do inwestycji wynosi 0,8-5%.

Premia z tytułu Polisy X jest uzależniona od wybranego przez inwestora wariantu*. Dla WARIANTU I kupon jest równy wzrostowi indeksu WIG20, jeśli maksymalna wartość indeksu WIG20 jest mniejsza od 150% wartości początkowej lub kupon jest równy 10,5-16,5%, jeśli maksymalna wartość indeksu WIG20 jest większa lub równa 150% wartości początkowej. Dla WARIANTU II kupon jest równy wzrostowi indeksu WIG20, jeśli maksymalna wartość indeksu jest mniejsza od 150% wartości początkowej oraz indeks WIG20 na koniec inwestycji jest równy lub powyżej wartości początkowej lub kupon jest równy 21-29%, jeśli maksymalna wartość indeksu WIG20 jest większa lub równa 150% wartości początkowej i jeżeli indeks WIG20 na koniec inwestycji jest powyżej wartości początkowej. Jeżeli natomiast wartość indeksu WIG20 na koniec inwestycji jest poniżej wartości początkowej, kupon jest dodawany do 90% składki zainwestowanej.

Polisa obejmuje gwarancją zainwestowaną składkę. Dla WARIANTU I – 100% po zakończeniu okresu ubezpieczenia, a dla WARIANTU II – 90% po zakończeniu okresu ubezpieczenia.

Z przytoczonych powyżej uproszczonych reguł związanych z inwestycją w polisę wynika, że ubezpieczonego w zależności od wybranego wariantu mogą spotkać różne scenariusze w związku z inwestycją w Polisę X. Możliwe scenariusze dla polisy inwestycyjnej w ramach:

1. Wariantu I:
   - **I.I.** Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji wzrośnie, ale nie osiągnie 150% wartości początkowej indeksu, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości odpowiadającej wzrostowi indeksu, np. wartość indeksu wzrośnie o 2,77%, klient otrzyma na koniec inwestycji 100% składki zainwestowanej + kupon w wysokości 2,77%.

wość sprawdzenia, jak zachowałyby się wartości bieżących pozycji czy też wyliczanych wskaźników w przypadku innej struktury portfela (niż obecna) lub w przypadku zrealizowania wybranych scenariuszy dynamiki stóp zwrotu instrumentów.

* Ostateczna wartość kuponu jest ustalana po zakończeniu okresu subskrypcji. Warunki polisy odzwierciedlają warunki jednej z polis oferowanych na rynku.

- **I.II.** Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji osiągnie lub przekroczy 150% wartości początkowej, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości 10,5-16,5%.
- **I.III.** Jeśli WIG20 na koniec okresu inwestycji będzie poniżej wartości początkowej i w trakcie trwania inwestycji osiągnie lub przekroczy 150% wartości początkowej indeksu, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości 10,5-16,5%.
- **I.IV.** Jeśli WIG20 na koniec okresu inwestycji będzie poniżej wartości początkowej i w trakcie trwania inwestycji nie osiągnie ani nie przekroczy 150% wartości początkowej indeksu, klient otrzyma na koniec 100% składki zainwestowanej.

2. Wariantu II:

- **II.I.** Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji nie osiągnie 150% wartości początkowej indeksu oraz indeks WIG20 na koniec inwestycji jest równy lub powyżej wartości początkowej, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości odpowiadającej wzrostowi indeksu, np. jeśli wartość indeksu wzrośnie o 3,94%, klient otrzyma na koniec inwestycji 100% składki zainwestowanej + kupon w wysokości 3,94%.
- **II.II.** Jeśli wartość WIG20 w trakcie trwania okresu inwestycji osiągnie lub przekroczy 150% wartości początkowej i na koniec będzie równy wartości początkowej, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości 21-29%.
- **II.III.** Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji osiągnie lub przekroczy 150% wartości początkowej i na koniec powyżej wartości początkowej klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości 21-29%.
- **II.IV.** Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji osiągnie lub przekroczy 150% wartości początkowej, ale na koniec będzie poniżej wartości początkowej, klient otrzyma na koniec 90% składki zainwestowanej + kupon w wysokości 21-29%.
- **II.V.** Jeśli wartość WIG20 na koniec okresu inwestycji będzie poniżej wartości początkowej i w trakcie trwania inwestycji nie osiągnie ani nie przekroczy 150% wartości początkowej indeksu, inwestor otrzyma na koniec 90% składki zainwestowanej.

2. Opis algorytmu predykcji instrumentu bazowego

Z opisu polisy inwestycyjnej jasno wynika, że kluczową rolę w inwestycji odgrywa zmienność cen i cena instrumentu bazowego, w tym przypadku WIG20. Zmienność cen instrumentu bazowego najprościej można określić jako miarę wskazującą, o ile dany instrument może zmienić swoją wartość w danym okresie. Zatem celem oszacowania korzyści z inwestycji w polisę inwestycyjną należy oszacować wartość instrumentów bazowych na dzień zapadalności polisy.

W artykule proponuje się oszacowanie instrumentów wchodzących w skład koszyka tworzącego w całości instrument bazowy na podstawie przedstawionego poniżej uogólnionego autorskiego modelu [Dyduch 2006], którego uproszczony ogólny algorytm przedstawiono na rys. 2.

Rys. 2. Schemat predykcji instrumentu bazowego

Proponowany algorytm opiera się na analizie falkowej [Dyduch 2008]. W literaturze analiza falkowa jest przedstawiana w dwóch odmianach: dyskret-

nej DWT (*Discrete Wavelet Transform*) i ciągłej CWT (*Continuous Wavelet Transform*).

3. Algorytm predykcji instrumentu bazowego

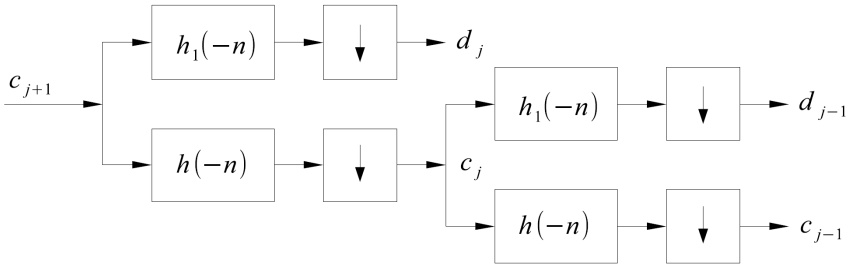
Predykcję instrumentu bazowego, jakim jest w analizowanym przypadku WIG20, oparto na indeksach giełdy japońskiej, niemieckiej, amerykańskiej oraz chińskiej. Szeregi indeksów giełdowych uwzględnionych w badaniu prezentują okres czasowy 23.04.1991-01.03.2009 r. Szeregi Dow Jones, DAX, Nikkei, Hang Seng, WIG prezentują indeksy różnych giełd światowych, zatem mimo że mieszczą się w tym samym przedziale czasowym, nie są równoliczne. Zatem konieczna jest m.in. standaryzacja czasowa szeregów.

Każdy z pięciu szeregów podzielono na podszeregi, tzw. próbki o parzystej liczbie obserwacji, będące wielokrotnością liczby 2. Możliwości podziału jest wiele, można ograniczyć każdy z szeregów do wielokrotności liczby 2 lub utworzyć kilkanaście szeregów dwuelementowych, czteroelementowych, ośmioelementowych, szesnastoelementowych itd. Następnie, po wygenerowaniu odpowiedniej falki, wyznaczono współczynniki falkowe szeregów odpowiednim algorytmem (rys. 3), który w uproszczeniu można przedstawić następująco [Dyduch 2008]:

1. Określenie współczynników filtrów: dolnoprzepustowego i górnoprzepustowego.
2. Splot sygnału wejściowego ze współczynnikami filtru dolnoprzepustowego, co prowadzi do otrzymania dolnoprzepustowej informacji o sygnale. W wyniku operacji splotu otrzymuje się:

   $$N + M - 1 \text{ próbek},$$

   gdzie:
   N – ilość próbek sygnału,
   M – długość filtru.
3. Splot sygnału wejściowego ze współczynnikami filtru górnoprzepustowego, co prowadzi do otrzymania górnoprzepustowej informacji o sygnale.
4. Przekształcenie otrzymanych wektorów, tzn. odrzucenie z każdego z otrzymanych wektorów co drugiej próbki i otrzymanie współczynników aproksymacji c i detali d.



Rys. 3. Schemat wyznaczania współczynników dyskretnej transformaty falkowej za pomocą banku filtrów – analiza wielopoziomowa

Macierze współczynników falkowych dla poszczególnych szeregów giełdowych zostały wyznaczone dla każdego rozpatrywanego indeksu przy podziale każdego szeregu na podszeregi dwuelementowe i jednym poziomie rozdzielczości falki. Są one następujące:

1. Dla szeregu prezentującego indeks Down Jones:

$$C^{(1)} = \begin{pmatrix} 4122,4820 & 4086,6034 & 31,0718 & -31,0718 \\ 4122,7048 & 4101,0461 & 18,7570 & -18,7570 \\ 4163,8195 & 4249,6764 & -74,3543 & 74,3543 \\ \vdots & \vdots & \vdots & \vdots \\ 15849,3283 & 16044,9352 & -169,4006 & 169,4006 \\ 15865,2134 & 15649,8853 & 186,4797 & -186,4797 \\ 15757,2135 & 15856,7812 & -86,2282 & 86,2282 \end{pmatrix}$$

2. Dla szeregu prezentującego indeks DAX:

$$C^{(2)} = \begin{pmatrix} 2247,4328 & 2268,5753 & -18,3099 & 18,3099 \\ 2278,3334 & 2275,2928 & 2,6332 & -2,6332 \\ 2374,6060 & 2381,1114 & -5,6338 & 5,6338 \\ \vdots & \vdots & \vdots & \vdots \\ 7568,7437 & 7623,5869 & -47,4956 & 47,4956 \\ 7500,4513 & 7274,1489 & 195,9837 & -195,9837 \\ 7322,5008 & 7519,6139 & -170,7049 & 170,7049 \end{pmatrix}$$

3. Dla szeregu prezentującego indeks Nikkei:

$$C^{(2)} = \begin{pmatrix} 37021,9897 & 36992,2913 & 25,7196 & -25,7196 \\ 36668,0828 & 36227,5553 & 381,5080 & -381,5080 \\ 36075,1738 & 36366,5018 & -252,2974 & 252,2974 \\ \vdots & \vdots & \vdots & \vdots \\ 12385,8203 & 12484,9496 & -85,8485 & 85,8485 \\ 12302,6610 & 12188,9723 & 98,4572 & -98,4572 \\ 12203,2594 & 12218,6956 & -13,3681 & 13,3681 \end{pmatrix}$$

4. Dla szeregu prezentującego indeks Hang Seng:

$$C^{(2)} = \begin{pmatrix} 5080,5622 & 5110,2607 & -25,7196 & 25,7196 \\ 5368,7082 & 5482,5524 & -98,5920 & 98,5920 \\ 5095,0579 & 5173,5468 & -67,9733 & 67,9733 \\ \vdots & \vdots & \vdots & \vdots \\ 28166,9057 & 28395,3294 & -197,8208 & 197,8208 \\ 28022,0584 & 27503,6006 & 448,9976 & -448,9976 \\ 27165,4975 & 27144,6166 & 18,0834 & -18,0834 \end{pmatrix}$$

5. Dla szeregu prezentującego indeks WIG20:

$$C^{(2)} = \begin{pmatrix} 5080,5622 & 5110,2607 & -25,7196 & 25,7196 \\ 5368,7082 & 5482,5524 & -98,5920 & 98,5920 \\ 5095,0579 & 5173,5468 & -67,9733 & 67,9733 \\ \vdots & \vdots & \vdots & \vdots \\ 28166,9057 & 28395,3294 & -197,8208 & 197,8208 \\ 28022,0584 & 27503,6006 & 448,9976 & -448,9976 \\ 27165,4975 & 27144,6166 & 18,0834 & -18,0834 \end{pmatrix}$$

Następnie każdemu podszeregowi szeregu n przypisano wygenerowane współczynniki falkowe na różnych poziomach rozdzielczości i zainicjalizowano kolejny krok algorytmu związanego z inicjalizacją sztucznej sieci neuronowej*, efektem którego jest wyjściowa macierz współczynników falkowych szeregu WIG20 w postaci:

$$WY^{(II)} = \begin{pmatrix} 9284,0645 & 9173,2609 & 95,9588 & -95,9588 \\ 9164,6696 & 9221,8038 & -49,4797 & 49,4797 \\ 9481,1352 & 9496,0552 & -12,9211 & 12,9211 \\ \vdots & \vdots & \vdots & \vdots \\ 51874,6015 & 52839,3851 & -835,5271 & 835,5271 \\ 52908,7452 & 53417,8126 & -440,8653 & 440,8653 \\ 52332,2764 & 51611,4729 & 624,2341 & -624,2341 \end{pmatrix}$$

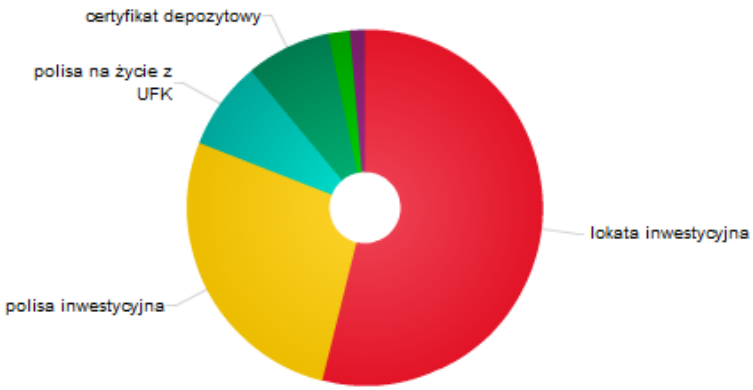
Dysponując wygenerowanymi współczynnikami transformaty falkowej dla przyszłych wartości indeksu WIG, zastosowano algorytm odwrotnej transformaty falkowej, dający w efekcie wartości przyszłe, tzn. prognozy szeregu WIG20 na dzień wskazany w opisie subskrypcyjnym produktu strukturyzowanego. Zatem otrzymujemy wartość WIG20 na dzień 09.03.2012 r. wynoszącą 2288.

4. Wyniki oszacowania polisy

Otrzymana wartość WIG20 jest obarczona błędem na poziomie 0,6%. Bezbłędne oszacowanie jest niemożliwe m.in. z uwagi na długi horyzont prognozy oraz na dużą obserwowalną zmienność notowań wynoszącą 13,37%.

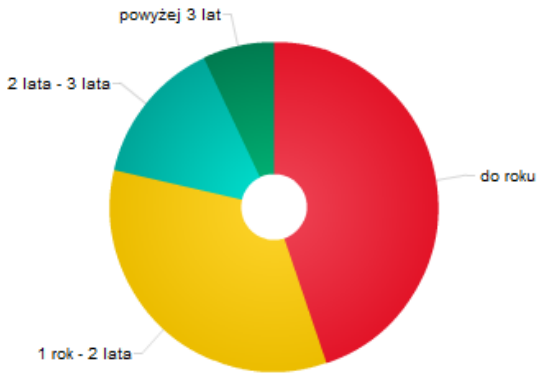
* Proces uczenia sieci neuronowej ma zgodnie z intencją prowadzić do odwzorowania zależności reprezentowanych przez zbiór uczący. Kształt powierzchni funkcji odwzorowującej w sieci neuronowej wektor wejść na wektor wyjść jest modyfikowany za pomocą zmian wartości wag w taki sposób, by jak najlepiej odpowiadał faktycznemu kształtowi przybliżanej funkcji. W rzeczywistości jednak sieć neuronowa dopasowuje wartości swoich wag do obserwacji ze zbioru uczącego, które mogą zawierać zniekształconą lub niepełną informację o zależnościach zjawiska lub funkcji, które sieć stara się odwzorować. Ponadto sieć neuronowa jest uczona z myślą o zdolności do trafnego przybliżania wyjść odpowiadających tym punktom w przestrzeni wejść, które nie należą do zbioru treningowego, czyli do możliwości generalizowania. Innymi słowy na tym etapie sieć neuronowa wykorzystuje swoją zdolność do uogólniania, czyli to, że wytrenowana jest zdolna w sposób mniej lub bardziej sensowny do generalizowania i wygenerowania rozwiązania dla dowolnego nowego, nieznanego wektora (niewchodzącego w skład zbioru treningowego) podanego jej na wejście.

Na podstawie otrzymanej prognozy inwestor może wnioskować, że kurs WIG20 na dzień zapadalności polisy inwestycyjnej prawdopodobnie będzie się mieścił w przedziale ⟨2274,27;2301,73⟩. Kurs początkowy wynoszący 1477,48 nie należy do przedziału ⟨2274,27;2301,73⟩. Zatem prognozowany kurs WIG20 na dzień zapadalności polisy inwestycyjnej jest powyżej kursu początkowego. Wyznaczona prognoza pozwala ze względu na warunek: „Kurs WIG20 na dzień zapadalności polisy inwestycyjnej jest powyżej kursu początkowego" przyjąć za możliwe do zrealizowania dla inwestora w polisę X scenariusze I.I, I.II, II.I, II.III. Każdy z tych scenariuszy jest korzystny dla inwestora, ponieważ daje premię w postaci kuponu. Podsumowując, w przypadku realizacji dla inwestora w polisę X SCENARIUSZA I.I głoszącego, że „Jeśli wartość WIG20 w trakcie trwania lub na koniec okresu inwestycji wzrośnie, ale nie osiągnie 150% wartości początkowej indeksu, klient otrzyma na koniec 100% składki zainwestowanej + kupon w wysokości odpowiadającej wzrostowi indeksu", inwestor otrzyma 100% składki zainwestowanej + kupon w wysokości 54,85%. Zatem przy inwestycji w polisę inwestycyjną kapitału w wysokości 10 000 zł inwestor otrzyma w dniu wypłaty środków z inwestycji 15 485,83 zł.



Rys. 4. Podział produktów strukturyzowanych ze względu na konstrukcję, stan na 01.09.2013 r.

Inwestor usatysfakcjonowany z inwestycji w polisę zapewnie znowu ulokuje swój kapitał lub jego część w polisę inwestycyjną [Dyduch 2011]. Jak pokazują dane, około 25% produktów strukturyzowanych sprzedawanych obecnie na polskim rynku jest w formie polisy inwestycyjnej. Około 20% inwestycji to inwestycje powyżej 2 lat.



Rys. 5. Podział produktów strukturyzowanych ze względu na okres inwestycji, stan na 01.09.2013 r.

Zakończenie

W dzisiejszych czasach każde przedsiębiorstwo oraz każdy inwestor, aby osiągnąć zaplanowane wyniki, podejmuje ryzyko. Wolny rynek stwarza zarówno szanse na osiągnięcie ponadplanowych zysków, jak i ryzyko strat w wyniku niekorzystnych zmian w otoczeniu przedsiębiorstwa oraz inwestora. Wszystkie decyzje biznesowe są obarczone ryzykiem. Ryzyko to ma szczególny wymiar, gdyż dotyczy zaangażowanych środków finansowych, dlatego niezwykle istotne jest właściwe zarządzanie ryzykiem oraz podejmowanie właściwych decyzji inwestycyjnych.

Bibliografia

Adekwatność kapitałowa i zarządzanie ryzykiem w PZU AM na 31 grudnia 2011 r. Raport, Warszawa, dnia 25 maja 2012 r.

Biegański M., Janc A. (red.), 2001: Hedging i nowoczesne usługi finansowe. Wydawnictwo Akademii Ekonomicznej, Poznań.

Dyduch M., 2006: Zastosowanie sieci falkowo-neuronowej do predykcji ekonomicznych szeregów czasowych. W: Prognozowanie w zarządzaniu firmą. P. Dittmann, J. Krupowicz (red.). Wydawnictwo UE, Wrocław.

Dyduch M., 2008: Wykorzystanie metod sztucznej inteligencji do wspomagania decyzji inwestycyjnych. Inwestowanie na rynku kapitałowym, Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania nr 10, Wydawnictwo Uniwersytetu Szczecińskiego, Szczecin.

Dyduch M., 2011: Grupowanie produktów strukturyzowanych. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu nr 185, Wrocław, s. 159-169.

Głuchowski J. (red.), 2001: Leksykon finansów. Polskie Wydawnictwo Ekonomiczne, Warszawa.

Gontarek D., Maksymiuk R., Krysiak M., Witkowski Ł., 2001: Nowoczesne metody zarządzania ryzykiem finansowym. WIG-Press, Warszawa.

Jajuga K., Jajuga T., 2011: Instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa. Wydawnictwo Naukowe PWN, Warszawa.

EFFECTIVENESS OF ALTERNATIVE INVESTMENT CAPITAL ON EXAMPLE INVESTMENT POLICY

Summary

The effectiveness of alternative investments are presented in the article as an example constructed investment policy. Adopted for the valuation of investment policy was a form of group life insurance and endowment of the underlying WIG20 index investment policy predictions based on the model copyright.

Stanisław Wieteska
Uniwersytet Jana Kochanowskiego w Kielcach
Filia w Piotrkowie Trybunalskim

KONCEPCJA UBEZPIECZENIA UPRAW ROLNYCH OD SZKÓD SPOWODOWANYCH PRZEZ ZWIERZYNĘ ŁOWNĄ W POLSKIM OBSZARZE TERYTORIALNYM

1. Postawienie problemu

W lasach w Polsce żyje kilka gatunków zwierząt łownych. Można do nich zaliczyć: dziki, sarny, jelenie, daniele, łosie, zające, lisy, muflony, kuropatwy, bażanty.

Według stanu na koniec 2011 r. w polskim obszarze terytorialnym żyje około 4862 łosi; 26 517 danieli; 194,7 tys. jeleni; 829,9 tys. saren; 267,8 tys. dzików, rozmieszczonych w sposób losowy w poszczególnych województwach [Dane Agencji Nieruchomości Rolnych…, 2011, tab. 4(133), s. 159]. Zwierzęta te oprócz żerowisk w lasach bardzo często wkraczają na tereny rolnicze, niszcząc płody rolne gospodarstw rolnych. Z tego tytułu powstają tzw. szkody łowieckie. Odszkodowania z tytułu szkód łowieckich są wypłacane rolnikom indywidualnym oraz producentom warzyw i owoców.

Powstawanie szkód łowieckich należy traktować jako zdarzenia losowe, zdefiniowane przez ustawę o działalności ubezpieczeniowej. Zachowanie się zwierząt łownych ma cechy losowości zdarzeń naturalnych. Do tych cech należy zaliczyć: wrażliwość, wyostrzone zmysły, naturalną płochliwość, cierpienie z powodu braku pokarmu, żerowanie połączone z wędrówką, tworzenie naturalnych warunków przetrwania i prokreacji.

Przez polski obszar terytorialny będzie tu rozumiana przestrzeń zawarta między granicami państwa.

W ustawie z 7 lipca 2005 r. o dopłatach do ubezpieczeń rolnych i zwierząt gospodarskich (DzU nr 150, poz. 1249 wraz z późniejszymi zmianami) wyszczególnia się wiele zagrożeń naturalnych, od których uprawy są obejmowane ochroną ubezpieczeniową. Pomija się tu zagrożenia naturalne dla upraw rolnych powodowane przez zwierzęta łowne.

Powstaje więc pytanie: czy jest możliwe ubezpieczenie upraw rolnych od strat spowodowanych przez zwierzynę łowną?

Warto przypomnieć, że w czasie prac nad ustawą o dopłatach do składek w ubezpieczeniach upraw rolnych rozważano także propozycję rozszerzającą zakres odpowiedzialności o szkody łowieckie. Niestety ta propozycja nie została przyjęta, gdyż na przeszkodzie stanęły wytyczne Wspólnoty w sprawie pomocy państwa w sektorze rolnym i leśnym na lata 2007-2013 oraz Rozporządzenie Komisji (WE) nr 1857/2006 z 15 grudnia 2006 r. w sprawie stosowania art. 87 i 88 Traktatu w odniesieniu do pomocy państwa dla MSP prowadzących działalność związaną z wytwarzaniem produktów rolnych, a także Rozporządzenie (WE) nr 70/2001, które nie przewidywało możliwości udzielenia pomocy ze środków publicznych w formie dopłat do składek ubezpieczeń producentów rolnych z tytułu szkód spowodowanych w uprawach rolnych przez zwierzynę łowną [Szymankiewicz 2008, s. 24-26].

Pomimo licznych zabiegów rolników, kół myśliwskich, szkody łowieckie powstają od wielu lat i nic nie wskazuje na to, że nie będą powstawały w przyszłości. Zwierzęta te zachowują się w sposób przypadkowy co do czasu i miejsca żerowania.

Celem głównym artykułu jest dostarczenie podstawowej wiedzy na temat szkód łowieckich powstających na terenie Polski. Celem dodatkowym jest próba konstrukcji ogólnych obowiązkowych warunków ubezpieczenia upraw rolnych, roślin okopowych i warzyw, a także składek od szkód spowodowanych przez zwierzynę łowną. Postawiono tu tezę, że szkody łowieckie powstają w wyniku naturalnego zachowania się zwierzyny łownej i powinny być włączone do zakresu obowiązkowej ochrony ubezpieczeń upraw rolnych bez dopłat do składki z tego tytułu[*].

[*] Dotychczas w literaturze przedmiotu można się było spotkać z programem grupowego ubezpieczenia odpowiedzialności cywilnej kół łowieckich. Do tego ubezpieczenia przystąpiło około 90% kół łowieckich. Grupowa umowa ubezpieczenia została podpisana między Zarządem Głównym Polskiego Związku Łowieckiego a Towarzystwem Ubezpieczeniowym Allianz. W ramach tego ubezpieczenia Towarzystwo odpowiada za szkody (roszczenia) wyrządzone przez koło łowieckie na skutek działania bądź realizacji zadań statutowych z wyłączeniem szkód powstałych z winy umyślnej. Jak dotąd najwięcej szkód powstaje w wyniku zderzenia zwierzyny łownej z pojazdami mechanicznymi (np. w czasie polowania).

Artykuł jest skierowany do zakładów ubezpieczeń majątkowo-osobowych z myślą objęcia ochroną ubezpieczeniową tego mało znanego zagrożenia naturalnego. Autor artykułu nie pretenduje do bycia znawcą problematyki szkód łowieckich, lecz z racji zajmowania się szkodami naturalnymi proponuje działom: aktuarialnym, oceny ryzyka, wdrożenie nowych produktów, zainteresowanie się tą sprawą.

2. Pojęcie szkód łowieckich

Pod pojęciem szkód spowodowanych przez zwierzynę łowną będziemy rozumieć straty spowodowane przez żerowanie w uprawach rolnych (np. zbożach, roślinach okopowych, warzywach). Do szkód łowieckich będziemy także zaliczać straty powstałe w trakcie polowań na zwierzynę.

W myśl ustawy z 16 kwietnia 2004 r. (DzU nr 92, poz. 880) o ochronie przyrody, Skarb Państwa odpowiada za szkody wyrządzone przez żubry, wilki, rysie, niedźwiedzie i bobry. Nie obejmują one utraconych korzyści. Oględzin i szacowania szkód, a także ustalenia wysokości odszkodowań i jego wypłaty dokonuje wojewoda, a na obszarze parku narodowego jego dyrektor.

Badania nad szkodami łowieckimi w Polsce wykazują, że rozmiar szkód jest uzależniony od [Szkody łowieckie…, 2005]:

- wielkości populacji, konkurencji między zwierzętami,
- rozmieszczenia przestrzennego gatunków zwierząt,
- dostępu do pokarmu, jego jakości,
- strategii dostępu do pokarmu, warunków osłonowych.

Szkody łowieckie powstają w ciągu całego roku. Duże szkody występują w okresie jesienno-zimowym, zwłaszcza w oziminach i rzepakach. Największym sprawcą szkód są dziki (około 80%), reszta (około 20%) przypada na sarny, jelenie, żubry i inne. W sensie fizycznym szkody łowieckie polegają na zjadaniu plantacji, stratowaniu, deptaniu, utworzeniu ścieżek itp. Wyłączamy z rozważań przypadki, gdy rolnik nie zebrał na czas plonów, a zwierzęta zniszczyły te uprawy.

Warto zwrócić uwagę, że ochrona upraw rolnych przed zwierzętami łownymi jest coraz trudniejsza. Stosowane środki ochrony, odstraszania, środki chemiczne, stają się mniej skuteczne, zatem problem szkód łowieckich będzie narastał.

W tabeli 1 prezentujemy potencjalne struktury upraw rolnych w skali województw narażonych na szkody łowieckie (dane za 2008 r.).

Tabela 1

Powierzchnia zasiewów w 2008 r. w tys. ha

Lp.	Województwa	Ogółem	W tym:		
			ziemniaki	buraki cukrowe	rzepak i rzepik
1	POLSKA	11631,1	548,9	187,5	771, 1
2	Dolnośląskie	741,0	28,4	19,9	113,4
3	Kujawsko-pomorskie	975,1	26,7	30,9	105,5
4	Lubelskie	1209,0	44,4	30,0	43,1
5	Lubuskie	325,2	11,8	1,4	26,0
6	Łódzkie	858,1	62,9	6,7	15,4
7	Małopolskie	421,9	46,4	1,0	4,5
8	Mazowieckie	1381,4	82,1	16,1	28,9
9	Opolskie	484,9	13,7	10,8	74;2
10	Podkarpackie	417,5	50,5	4,0	11,3
11	Podlaskie	706,0	24,0	0,0	3,6
12	Pomorskie	573,2	28,8	8,5	51,9
13	Śląskie	291,3	14,9	1,5	18,2
14	Świętokrzyskie	392,6	30,1	5,7	6,3
15	Warmińsko-mazurskie	616,4	12,3	2,8	58,2
16	Wielkopolskie	1529,0	46,7	39,6	112,4
17	Zachodniopomorskie	708,6	25,2	8,7	98,4

Źródło: Rocznik Statystyczny Województw [2009, tab. 3 (180)].

Z danych zawartych w tabeli 1 wynika, że w każdym województwie występuje podobna struktura upraw rolnych. Na szkody łowieckie są narażone praktycznie wszystkie rodzaje upraw rolnych, warzyw i traw. Najbardziej są narażone uprawy o wysokiej kaloryczności i atrakcyjności pokarmowej.

3. Skala wypłaconych szkód łowieckich w Polsce w latach 2000-2011

Szkody łowieckie powstają już od dość dawna. Rejestrację szkód łowieckich w latach 2000-2011 prowadzi m.in. Dyrekcja Generalna Lasów Państwowych (tabela 2).

Z danych zawartych w tabeli 2 wynika, że odszkodowania łowieckie są wypłacane: osobom fizycznym i prawnym, posiadaczom upraw i płodów rolnych uszkodzonych przez zwierzynę łowną. Wysokość odszkodowań łowieckich ma generalnie tendencję rosnącą.

Tabela 2

Odszkodowania łowieckie w Polsce w latach 2000-2011*

Wyszczególnienie	Ogółem	Państwowe Gospodarstwa Leśne, Lasy Państwowe	Agencja Nieruchomości Rolnych	Polski Związek Łowiecki
	w tys. zł (ceny bieżące)			
2000/2001	26 369	7244	257	18 868
2001/2002	29 803	7497	278	22 028
2002/2003	25 322	5296	151	19 876
2003/2004	26 775	5975	208	20 591
2004/2005	35 117	7309	331	27 472
2005/2006	31 179,0	6772,0	180,0	24 222,0
2006/2007	28 487,9	5632,0	171,0	22 684,9
2007/2008	41 488,1	7977,0	144,0	33 367,1
2008/2009	55 535,0	10 238,0	244,0	45 053,0
2009/2010	49 513,9	9011,0	157,9	40 345,0
2010/2011	57 376,2	9939,6	586,6	46 850,0

* Wypłacone osobom fizycznym lub prawnym – posiadaczom upraw i płodów rolnych uszkodzonych przez dziki, łosie, jelenie, daniele i sarny oraz za szkody powstałe podczas polowania. Dane dotyczą łowieckiego roku gospodarczego liczonego od 1 IV bieżącego roku do 31 III roku następnego.

Źródło: Dane Agencji Nieruchomości Rolnych... [2006; 2011, tab. 9/138, s. 162].

Najwięcej odszkodowań wypłacają Państwowe Gospodarstwa Leśne, Lasy Państwowe (około 18-19%) oraz Polski Związek Łowiecki (około 80-81%).

Prezentowane dane statystyczne dotyczą szkód łowieckich wypłacanych w skali całej Polski. Brakuje jednak danych o szkodach wypłacanych w skali regionalnej czy też kół łowieckich. Niestety statystyka publiczna nie gromadzi danych o szkodach w ujęciu wojewódzkim.

Z badań przeprowadzonych na Wyżynie Lubelskiej w okresach 1999-2000 i 2008-2009 wystąpił prawie 2,5-krotny wzrost szkód. Najwięcej szkód zostało spowodowanych przez dziki w ziemniakach w okresach kwiecień-maj oraz kukurydzy. Żerowiska na tych uprawach okazały się najbardziej atrakcyjne. Pomimo 8-krotnego wzrostu pozyskania łowieckiego i 2,5-krotnego wzrostu liczebności dzików szkody łowieckie mają tendencję rosnącą [Flis 2009, s. 179-187].

4. Metody oszacowania szkód i wysokości odszkodowań

Ważnym elementem z punktu widzenia zakładów ubezpieczeń jest likwidacja szkód. W tym przypadku likwidacja szkód łowieckich wymaga odrębnego potraktowania niż szkód spowodowanych innymi żywiołami. Przy likwidacji szkód łowieckich możemy skorzystać z ustaleń zawartych w Prawie łowieckim,

a także w Rozporządzeniu obowiązującym w latach 2002-2009 dotyczącym szacowania szkód oraz wypłat odszkodowań[*]. W Rozporządzeniu tym były zawarte:
- procedury zgłaszania szkód (terminy, formy),
- procedury oględzin i szacowania szkód,
- ustalanie ilości i wartości szkód,
- ustalanie przyczyny szkód (gatunek zwierząt, które spowodowały szkodę).

W zakresie likwidacji szkód łowieckich aktualnie obowiązuje Rozporządzenie Ministra Środowiska z 8 marca 2010 r. W Rozporządzeniu podaje się szczegółowo procedury szacowania szkód spowodowanych przez zwierzynę łowną[**].

Podstawą wypłacenia odszkodowań są ceny wolnorynkowe uszkodzonych płodów rolnych obowiązujące w czasie szacowania szkód. Wysokość odszkodowania oblicza się, mnożąc rozmiar szkód przez cenę skupu (cenę rynkową) płodu rolnego.

Wysokość odszkodowań za szkody wyrządzone przez dziki na łąkach i pastwiskach ustala się na podstawie ostatecznego szacowania, uwzględniając wartość utraconego plonu. Szczegółowy opis sposobu szacowania szkód w ziemniakach zaprezentowano w pracy M. Flisa. Autor podaje dwie metody określenia wielkości utraconego plonu [Flis 2008, s. 117-123]. Pierwsza polega na ustaleniu procentowego stopnia zniszczenia plantacji, druga (wagowa) – na ustaleniu wydajności plonów ziemniaka, a następnie ustaleniu szkód.

Pomimo znacznego postępu w sposobach i procedurze likwidacji szkód łowieckich w dalszym ciągu powinny być one doskonalone. Warto w tym miejscu zwrócić uwagę, że wypłacanie rolnikom odszkodowań za szkody łowieckie nie koresponduje z dopłatami unijnymi otrzymywanymi przez rolników. Konieczne są w tym zakresie uregulowania prawne. Należałoby także rozważyć możliwość partycypacji kół łowieckich w szkodach łowieckich w ramach tzw. pozyskania łowieckiego.

5. Częstość i intensywność szkód łowieckich

Na potrzeby kalkulacji stóp składek ubezpieczeniowych konieczne jest obliczenie częstości szkód. Przez częstość szkód będziemy rozumieć relację liczby

[*] Rozporządzenie Ministra Środowiska z 15 lipca 2002 r. w sprawie sposobu postępowania przy szacowaniu szkód oraz wypłat za szkody w uprawach i płodach rolnych (DzU nr 126 z 9 sierpnia 2002 r., poz. 1081).

[**] Rozporządzenie Ministra Środowiska z dnia 8 marca 2010 r. w sprawie sposobu postępowania przy szacowaniu szkód oraz wypłat odszkodowań za szkody w uprawach i płodach rolnych (DzU nr 45/2010, poz. 272).

przypadków szkód spowodowanych przez zwierzęta łowne do ilości ubezpieczonych ryzyk w ciągu roku.

Obliczenie rocznej częstości szkód spowodowanych przez zwierzynę łowną jest niezwykle trudne, gdyż może ona wielokrotnie nawiedzać uprawy rolne. Stąd na potrzeby ubezpieczeniowe możemy użyć pojęcia intensywności szkód rozumianego jako relacja powierzchni uszkodzonej do ogólnej powierzchni danej uprawy rolnej. Tak obliczony wskaźnik intensywności szkód może przybierać wartości (0, 1]. W skrajnym przypadku uprawy mogą być zniszczone w 100%. Nie dysponujemy danymi potrzebnymi do obliczenia tak zdefiniowanego wskaźnika intensywności szkód. Jest on możliwy do obliczenia na szczeblu zakładu ubezpieczeń, gdzie mogą być ubezpieczone dowolnej wielkości obszary płodów rolnych z konkretną lokalizacją przestrzenną.

Dla zobrazowania wysokości składki brutto możemy obliczyć wskaźnik intensywności szkód, dzieląc wypłacone odszkodowania przez powierzchnię upraw, uwzględniając przy tym koszty działalności zakładu ubezpieczeń. Zakładając równomierność szkód w skali kraju w 2008 r. szacujemy, że składka ubezpieczeniowa powinna wynieść dla tego roku około 5 zł/ha. Tak obliczona składka jest bardzo ogólna. Konieczne jest zróżnicowanie składki ze względu na zagrożenia dla różnych płodów rolnych. Potrzebne jest również zróżnicowanie składek ze względu na przestrzenny rozkład upraw. Niestety brak dostępu do danych szczegółowych uniemożliwia jej zróżnicowanie.

Ze względu na wolnorynkowe ruchy cen płodów rolnych składka ubezpieczeniowa powinna być kalkulowana w każdym roku.

6. Założenia ogólnych warunków ubezpieczeń

Dla celów praktyki ubezpieczeniowej jest potrzebna konstrukcja ogólnych warunków ubezpieczeń. Jest to bardzo ważne dla obu stron umów ubezpieczenia: poszkodowanych i zakładów ubezpieczeń. Wskażmy tutaj na najważniejsze elementy ogólnych warunków proponowanego ubezpieczenia.

Przedmiotem ubezpieczenia powinny być wyszczególnione uprawy rolne, ogrody, sady, warzywa, ich lokalizacja i powierzchnia. Nie powinny być wzięte pod ochronę ubezpieczeniową szkody w lasach spowodowane przez zwierzynę łowną. Wymaga to odrębnego ubezpieczenia.

Ubezpieczającym powinny być osoby indywidualne (gospodarstwa rolne) oraz osoby prawne (np. spółdzielnie produkcyjne).

Zakres ubezpieczenia powinien obejmować szkody spowodowane przez: dziki, sarny, jelenie, łosie, daniele, żubry.

Suma ubezpieczenia jako górna granica odpowiedzialności zakładu ubezpieczeń powinna stanowić przewidywaną wartość rynkową plonów z ubezpieczonych upraw; jest liczona w cenach skupu lub cenach rynkowych. Dla celów ubezpieczeniowych można wykorzystać maksymalne sumy ubezpieczenia publikowane corocznie przez Ministra Rolnictwa i Rozwoju Wsi*.

Składka ubezpieczeniowa powinna być obliczona jako iloczyn stopy składki przez sumę ubezpieczenia danej uprawy rolnej. Stopa składki powinna być obliczona początkowo na podstawie danych z kół łowieckich lub obwodów łowieckich, a w miarę zdobywania doświadczenia z wykorzystaniem danych zakładu ubezpieczeń.

Odszkodowania wypłacone powinny być ustalone zgodnie z aktualnie obowiązującymi przepisami prawa.

Dla celów ograniczenia wysokości szkód warto wykorzystać ustalenia art. 47 Prawa łowieckiego. Artykuł ten mówi, że właściciele lub posiadacze gruntów rolnych i leśnych powinni zgodnie z potrzebami współdziałać z dzierżawcami i zarządcami obwodów łowieckich w zabezpieczeniu gruntów przed szkodami, o których mowa w art. 46 [Radecki 2010, s. 263]. Ponieważ wspomniane współdziałanie może być trudne do osiągnięcia, koniecznością byłoby wkalkulowanie w składkę znanych instrumentów prewencyjnych, a tym samym zwiększenie jej efektywności. Tymi instrumentami zwiększającymi efektywność zapobiegania szkodom w ubezpieczeniach mogą być np. franszyza czy udziały własne.

7. Obowiązkowy czy dobrowolny charakter ubezpieczenia

W praktyce ubezpieczeniowej konieczne jest rozstrzygnięcie, czy proponowane ubezpieczenia będą miały charakter obowiązkowy czy dobrowolny. W przypadku ubezpieczeń dobrowolnych zakres odpowiedzialności, a także wyłączenia mogą być w bardzo różny sposób kształtowane przez różne zakłady ubezpieczeń. Różny zakres odpowiedzialności może wynikać z następujących przyczyn:

- podejścia do oceny ryzyka ubezpieczeniowego, a więc także skali strat, jakie mogą powodować zwierzęta łowne,
- doświadczeń w zakresie akwizycji i likwidacji szkód,

* Na przykład Rozporządzenie Ministra Rolnictwa i Rozwoju Wsi z 23 lipca 2007 r. w sprawie maksymalnych sum ubezpieczenia dla poszczególnych upraw rolnych i zwierząt gospodarskich (DzU nr 144, poz. 1010).

– dotychczasowych wyników ekonomicznych, o ile takie ubezpieczenie było prowadzone (np. wskaźniki szkodowości, suma ubezpieczenia, regionalne zróżnicowanie szkód itp.),
– konkurencyjności stosowanych taryf ubezpieczeniowych,
– zróżnicowanych form zabezpieczeń przed szkodami.

W sumie kształt ogólnych warunków ubezpieczeń może zawierać dowolne sformułowania, procedury likwidacyjne, stopy składek itp. W praktyce popyt na tego rodzaju ubezpieczenia będzie bardziej dotyczył rolników wielokrotnie poszkodowanych przez zwierzęta łowne, stąd składka na takie ubezpieczenie może być dość wysoka.

Obowiązkowy charakter ubezpieczenia to stan narzucony przez ustawodawcę. Jest to forma przymusu normatywnego (rozporządzenia, ustawa), a także przymus sytuacyjny związany z powstaniem zdarzeń niezaplanowanych, nieoczekiwanych, niepożądanych, innymi słowy losowych. Wprowadzenie obowiązku ubezpieczenia w warunkach gospodarki rynkowej ma swoje zalety i wady.

Do zalet należy zaliczyć:
– rozłożenie ciężaru szkód na wszystkich ubezpieczonych, w naszym przypadku na całość upraw w Polsce; zrozumiałe jest, że obciążenie składką powinno być zróżnicowane regionalnie (np. według województw czy powiatów),
– gwarancję wypłacenia odszkodowania ubezpieczonym,
– państwo nie będzie ponosiło ciężaru szkód, lecz jedynie zakłady ubezpieczeń,
– łatwość zawierania (akwizycji) umów ubezpieczenia,
– łatwość policzenia stóp składek,
– ubezpieczenie to pełni rolę stabilizatora produkcji rolnej,
– istnieje możliwość objęcia ochroną ubezpieczeniową jednocześnie „dobrych" i „złych" ryzyk,
– dzięki ubezpieczeniom obowiązkowym jest możliwość skierowania części środków na działalność prewencyjną.

Do wad tego ubezpieczenia można zaliczyć:
– możliwość „ukarania" rolników za niezapłacenie składki,
– przymus ubezpieczenia zawsze będzie odbierany przez rolników w sposób negatywny,
– zapłacenie składki dla słabszych ekonomicznie rolników może być zbyt uciążliwe.

Biorąc pod uwagę zalety i wady, widać, że zalet jest więcej niż wad. Z poprzednich rozważań wynika, że obszar szkód łowieckich jest dość dobrze rozpoznany, zatem istnieją przesłanki do wprowadzenia takiego produktu ubezpieczeniowego do praktyki rolniczej za pomocą właściwego rozporządzenia. Jego treść

powinna być uprzednio przedyskutowana z zakładami ubezpieczeń, służbami rolnymi, np. ośrodkami doradztwa rolniczego i kołami łowieckimi.

W ogólnych warunkach ubezpieczeń zawartych w rozporządzeniu powinny być: wytyczne w zakresie prewencji przed szkodami spowodowanymi przez zwierzynę łowną, a także określenie franszyz czy też udziałów własnych. Są to instrumenty dyskusyjne, o ich wysokości powinna zdecydować praktyka ubezpieczeniowa. Te instrumenty miałyby na celu zmniejszenie ilości, częstości, a także wartości szkód spowodowanych przez zwierzynę łowną. W ramach prewencji technicznej (ogrodzenia, systemy alarmowe) należałoby włączyć działalność kół łowieckich.

Zakończenie i wnioski

Z przeprowadzonych rozważań wynika, że szkody spowodowane przez zwierzynę łowną są dość dużych rozmiarów i mają w czasie tendencję rosnącą. Szkody łowieckie powstają w wyniku naturalnego zachowania się zwierzyny łownej. Można je traktować jako zdarzenia naturalne, za które mogą odpowiadać zakłady ubezpieczeń majątkowo-osobowych (dział II).

Z artykułu wynikają następujące wnioski:

- możliwe jest objęcie ochroną ubezpieczeń obowiązkowych upraw rolnych przed szkodami powodowanymi przez zwierzynę łowną,
- wstępne szacunki wskazują, że składka ubezpieczeniowa powinna być zróżnicowana.

Ograniczone ramy artykułu spowodowały, że problematyka nie została wyczerpana, lecz jedynie zasygnalizowana. Konieczne są dalsze badania, które powinny być ukierunkowane na uzależnienie stóp składek od rodzaju uprawy i regionalnego ich zróżnicowania.

Bibliografia

Wydawnictwa zwarte i ciągłe

Dane Agencji Nieruchomości Rolnych, Dyrekcji Lasów Państwowych i Zarządu Głównego Polskiego Związku Łowieckiego, Leśnictwo, 2006. GUS, Warszawa, tab. 9/138, s. 162.

Flis M., 2008: Procedura szacowania szkód wyrządzonych przez zwierzęta w uprawach rolniczych. „Biuletyn IHAR", 248.

Flis M., 2009: Wielkość szkód wyrządzonych przez dziki w uprawach rolniczych w obwodzie łowieckim polnym w latach 1998-2000 i 2008-2009. „Biuletyn IHAR", 254.

Gaweł J., 2006: OC kół łowieckich. „Łowiec Polski", 3.

Radecki W., 2010: Prawo łowieckie, komentarz. Wydanie III zaktualizowane. Stan prawny na 31.12.2010. Difin, Warszawa.

Rocznik Statystyczny Województw, 2009. GUS, Warszawa, tab. 3 (180).

Szkody łowieckie – uwarunkowania i możliwości zapobiegania, 2005. Instytut Ochrony Roślin, Poznań.

Szymankiewicz N., 2008: Myśliwi w gorszej pozycji. „Łowiec Polski", 9.

Trębska L., 2006: Sukces ubezpieczeń. „Łowiec Polski", 3.

Akty prawne

Rozporządzenie Ministra Rolnictwa i Rozwoju Wsi z 23 lipca 2007 r. w sprawie maksymalnych sum ubezpieczenia dla poszczególnych upraw rolnych i zwierząt gospodarskich (DzU nr 144, poz. 1010).

Rozporządzenie Ministra Środowiska z 15 lipca 2002 r. w sprawie sposobu postępowania przy szacowaniu szkód oraz wypłat za szkody w uprawach i płodach rolnych (DzU nr 126 z 9 sierpnia 2002 r., poz. 1081).

Rozporządzenie Ministra Środowiska z dnia 8 marca 2010 r. w sprawie sposobu postępowania przy szacowaniu szkód oraz wypłat odszkodowań za szkody w uprawach i płodach rolnych (DzU 2010, nr 45, poz. 272).

THE CONCEPT OF INSURANCE OF AGRICULTURAL CROPS FROM DAMAGE CAUSED BY GAME IN POLISH TERRITORIAL AREA

Summary

In Poland, there are many species of game animals. These animals are causing damage to crops. This article aims to propose a compulsory insurance of agricultural crops from damage caused by game. The article describes the size of the occurrence and the damage liquidation procedures.

ZARZĄDZANIE RYZYKIEM PROJEKTU

Sławomir Biruk
Piotr Jaśkowski
Politechnika Lubelska

PROJEKTOWANIE REALIZACJI OBIEKTÓW BUDOWLANYCH W WARUNKACH RYZYKA*

Wprowadzenie

Optymalne projektowanie pracy jednostek wykonawczych w funkcji czasu wymaga zapewnienia ich harmonizacji poprzez ich odpowiedni dobór i rozmieszczenie (zsynchronizowanie) działań w czasie [Dyżewski 1965, s. 866; Jaworski 1999, s. 17; Rowiński 1982, s. 46]. Celem harmonizacji jest wyeliminowanie nieuzasadnionych przerw w pracy zasobów (w odniesieniu do całego planu zadań przedsiębiorstwa oraz w skali poszczególnych inwestycji i budów).

Przedsięwzięcia budowlane, z uwagi na odmienny sposób modelowania ich struktury organizacyjnej, klasyfikuje się na dwa podstawowe rodzaje: przedsięwzięcia typu kompleks operacji i przedsięwzięcia, które mogą być zorganizowane zgodnie z zasadami metody pracy równomiernej, której określone wersje nazywa się coraz częściej potokowymi metodami organizacji pracy [Jaworski 1999, s. 39-40]. Przedsięwzięcia typu kompleks operacji obejmują niejednorodne pod względem technologicznym procesy, które nie charakteryzują się cyklicznością i rytmicznością realizacji. W metodach potokowych te same brygady robocze (w niezmiennym składzie) wykonują procesy budowlane na kolejnych obiektach w tej samej kolejności (cyklicznie, równomiernie i równolegle). Następny proces może być rozpoczęty po zakończeniu pracy poprzedniej brygady na danym obiekcie.

* Wyniki prac były finansowane ze środków statutowych przyznanych przez Ministerstwo Nauki i Szkolnictwa Wyższego (S/63/2013).

Zasady projektowania realizacji przedsięwzięć zgodnie z zasadami metody pracy równomiernej przedstawił Dyżewski [1965, s. 947-984], a następnie rozwinął Rowiński [1982, s. 112-152]. W literaturze angielskojęzycznej ta metoda harmonogramowania nosi nazwę LOB – *Line of Balance* lub *Linear Scheduling Method*. Harris i Ioannou [1998, s. 269-278] na podstawie analizy sprzężeń pomiędzy procesami realizowanymi na kolejnych działkach roboczych ustalili ciąg czynności (*controlling sequence*) wpływający na termin realizacji całego przedsięwzięcia, który ma takie samo znaczenie praktyczne, jak ścieżka krytyczna w metodzie CPM. Podejście to nazwano metodą RSM (*Repetitive Scheduling Method*).

Najwyższy stopień harmonizacji pracy brygad jest osiągany w warunkach pracy równomiernej i rytmicznej, gdy pracochłonność robót jednego rodzaju na poszczególnych obiektach jest taka sama (obiekty i procesy jednotypowe) lub proporcjonalna do wielkości obiektu (obiekty jednorodne). Przy zastosowaniu tej metody organizacji uzyskuje się minimalizację czasu wykonania wszystkich obiektów, ciągłość pracy brygad i ciągłość pracy na obiektach jednotypowych. W przypadku obiektów niejednorodnych (różnych pod względem wielkości, rzutu, konstrukcji i technologii procesów podstawowych) na czas i inne parametry organizacji wpływa kolejność zajmowania obiektów przez brygady.

Problem ustalenia optymalnej kolejności realizacji niejednorodnych obiektów, w celu skrócenia czasu realizacji przedsięwzięcia remontowego przy ciągłej pracy brygad, analizował Orłowski [1997]. Zastosował on w tym celu algorytm S. Johnsona, mający zastosowanie w problemie harmonogramowania pracy dwóch brygad roboczych.

Mrozowicz [1997] i Hejducki [2000] opracowali metody optymalnego planowania, projektowania i sterowania realizacją złożonych procesów niejednorodnych (zgodnie z założeniami międzynarodowej szkoły potokowych metod pracy). W metodach tych eksponuje się jakościowy charakter sprzężeń występujących między poszczególnymi robotami. Rodzaje występujących sprzężeń stanowią podstawę klasyfikacji tych metod. Przedstawiono rozwiązanie zagadnienia kombinatorycznego szeregowania dowolnej liczby zadań w poszczególnych metodach (z wykorzystaniem algorytmu podziału i ograniczeń), co umożliwia wyznaczenie optymalnej, według różnych kryteriów, kolejności prowadzenia robót na frontach roboczych.

Problem poszukiwania optymalnej kolejności realizacji obiektów i terminów realizacji procesów przez brygady specjalistyczne na obiektach realizowanych metodami potokowymi był przedmiotem prac również Marcinkowskiego [1990, 2002]. Opracował on metodę harmonogramowania przedsięwzięć reali-

zowanych sposobem potokowym dla kryterium czasowo-kosztowego w celu jak najefektywniejszego wykorzystania brygad w dostępnym programie inwestycyjnym. Jaśkowski i Biruk [2005, 2010] do rozwiązania problemu ustalania kolejności realizacji obiektów dla różnych kryteriów optymalizacji i z uwzględnieniem różnych ograniczeń zastosowali algorytmy opracowane dla problemu komiwojażera.

Większość opracowanych modeli i metod harmonogramowania ma zastosowanie w deterministycznych warunkach realizacji. Znacznie ułatwia to modelowanie złożonych przedsięwzięć, jednak tak opracowane harmonogramy są podatne na dezaktualizację w rzeczywistych warunkach realizacji. Szczególnie dotyczy to produkcji budowlanej, której specyficzne cechy (zmienne fronty robót, oddziaływanie warunków atmosferycznych, zmienność produktu, współpraca wielu wykonawców wymagająca koordynacji itd.) znacznie odróżniają ją od masowej produkcji przemysłowej, powtarzanej w stabilnych warunkach techniczno-organizacyjnych. Projekt realizacji powinien być opracowany dla dopuszczalnego poziomu ryzyka. Powszechnie stosowanym sposobem zwiększania odporności harmonogramów na zakłócenia jest alokacja buforów (rezerw) czasu. Celem ich stosowania jest zapewnienie terminowej realizacji poszczególnych procesów, etapów lub całego przedsięwzięcia. Istotnym problemem jest określenie wielkości buforów czasu i ich lokalizacji w harmonogramie [Jaśkowski, Biruk 2011].

Próbę rozwiązania problemu projektowania realizacji jednorodnych procesów w metodzie pracy równomiernej z uwzględnieniem stochastycznego charakteru przebiegu ich realizacji podjął Kapliński [1974]. Wykorzystał on w tym celu teorię masowej obsługi oraz badania symulacyjne. Przeprowadzone badania modelu pozwoliły na wyjaśnienie mechanizmu powstawania zakłóceń. Autor wskazał na potrzebę stosowania różnego rodzaju rezerw (opóźnień) we włączaniu ciągów procesów do realizacji w celu zwiększenia niezawodności dotrzymania terminu zakończenia przedsięwzięcia.

Zakłócenia losowe wpływają nie tylko na czas realizacji przedsięwzięcia, lecz również są źródłem kosztów (strat) związanych z niewykorzystaniem potencjału wykonawczego jednostek organizacyjnych na skutek przerw w ich pracy. Z tego powodu opracowano metodę analityczną określania wielkości rezerw w celu zapewnienia ciągłości w pracy brygad przy ustalonym dopuszczalnym poziomie prawdopodobieństwa wystąpienia przestojów. Ze względu na złożoność analizowanego problemu w artykule jest rozpatrywany jedynie przypadek harmonizacji dwóch ciągów produkcyjnych.

1. Metoda analityczna określania buforów czasu

Przedsięwzięcie obejmuje realizację złożonego procesu budowlanego na *m* niejednorodnych działkach roboczych.

Złożony proces budowlany podzielono na dwa procesy prostsze, powierzane do wykonania dwóm brygadom roboczym (o niezmiennym składzie kwalifikacyjnym i liczebności), zajmującym w ustalonej kolejności poszczególne działki robocze (wydzielone fronty robót na obiektach).

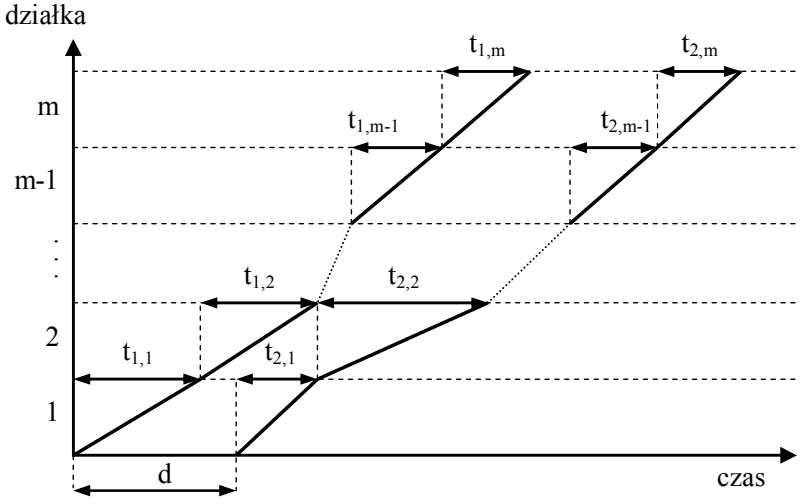
Szybka realizacja zadań na działkach roboczych wymaga maksymalnego stopnia wykorzystania dostępnych frontów robót.

Liczebności poszczególnych brygad – zgodnie z zasadami projektowania pracy równomiernej [Rowiński 1982, s. 141-152] – należy ustalać jako maksymalne tak, aby długości ich frontów pracy były równe długości frontów robót na działkach o najmniejszych pracochłonnościach.

Na każdej działce *j* ($j = 1, 2, ..., m$) musi być zrealizowany ciąg procesów *i* ($i = 1, 2$) w niezmiennej kolejności technologicznej.

Czas wykonania robót przez brygadę *i* na działce *j* wynosi t_{ij}. Termin rozpoczęcia pierwszego procesu na pierwszej działce jest równy $0 - t_{11}^{r} = 0$ (rys. 1).

Najkrótszy cykl budowy uzyskuje się przy krytycznym (maksymalnym) zbliżeniu sąsiadujących ze sobą linii łamanych (odcinków cyklogramu), które odwzorowują przebieg realizacji poszczególnych procesów technologicznych.



Rys. 1. Cyklogram dla ciągłej realizacji dwóch procesów

Proces 2 powinien być rozpoczęty z opóźnieniem *d*, które można obliczyć na podstawie następującej zależności:

$$d = \max\begin{cases} t_{11} \\ t_{11} + t_{12} - t_{21} \\ \dots \\ \sum_{j=1}^{m} t_{1j} - \sum_{j=1}^{m-1} t_{2j} \end{cases}.$$

Terminy zakończenia procesów ciągu pierwszego na kolejnych działkach $j\ (j = 1, 2, \dots, m)$ określa się na podstawie zależności:

$$t_{1j}^{z} = \sum_{k=1}^{j} t_{1k},$$

natomiast termin rozpoczynania kolejnego procesu można określić w sposób następujący:

$$t_{2j}^{r} = d + \sum_{k=1}^{j-1} t_{2k}.$$

W warunkach losowych czasy t_{ij} wykonania procesów $i\ (i = 1, 2)$ na kolejnych działkach $j\ (j = 1, 2, \dots, m)$ są zmiennymi losowymi z wartością oczekiwaną (przeciętną) $E(t_{ij})$ oraz odchyleniem standardowym $\sigma(t_{ij})$. Parametry zmiennych losowych można określić analogicznie do metody *PERT*, na podstawie oszacowań: t_{ij}^{a} – optymistycznego, t_{ij}^{m} – najbardziej prawdopodobnego i t_{ij}^{b} – pesymistycznego. Przyjmując założenie, że zmienne losowe t_{ij} są wzajemnie niezależne oraz pomijając wpływ ewentualnych przestojów brygady 2 na działkach poprzednich, wartość oczekiwana zapasu czasu na działce $j\ (j = 1, 2, \dots, m)$ będzie równa:

$$E(\psi_j) = \sum_{k=1}^{j-1} E(t_{2k}) - \sum_{k=1}^{j} E(t_{1k}) + d,$$

a jej wariancja:

$$D^2(\psi_j) = \sum_{k=1}^{j} D^2(t_{1k}) + \sum_{k=1}^{j-1} D^2(t_{2k}).$$

W celu przeciwdziałania losowym wahaniom czasu realizacji należy w szczególności zabezpieczyć ciągłość pracy brygady wykonującej proces 2 oraz wprowadzić dodatkową rezerwę czasu x zwiększającą termin włączenia ciągu technologicznego 2 (opóźniającą rozpoczęcie procesu 2 na pierwszej działce).

W artykule wielkość tej rezerwy jest ustalana tak, aby prawdopodobieństwo braku przerwy w pracy brygady 2 na skutek oczekiwania na zakończenie robót przez brygadę 1 na każdej działce było co najmniej równe wartości zdefiniowanej przez decydenta, określanej w literaturze mianem niezawodności dopuszczalnej [Jaworski 1999]. Dla każdej działki wymaganą wielkość opóźnienia $x_j \ (j = 1, 2, \ldots, m)$ w rozpoczęciu procesu 2 należy wyznaczyć tak, aby skutecznie przeciwdziałać losowym wahaniom czasu realizacji procesów. Wartości te można wyznaczyć z następującej zależności:

$$P(\psi_j \le E(\psi_j) + x_j) = R_{dop}, \ j = 1, 2, \ldots, m.$$

Zgodnie z centralnym twierdzeniem granicznym rozkład zmiennych losowych $\psi_j \ (j = 1, 2, \ldots, m)$ dąży do rozkładu normalnego.

Po dokonaniu standaryzacji rozkładu zmiennej losowej ψ_j warunek przybiera zatem następującą postać:

$$\frac{1}{2\pi} \int_{-\infty}^{u_j} \exp\left(-\frac{u^2}{2}\right) du = R_{dop}, \ j = 1, 2, \ldots, m,$$

gdzie R_{dop} to założona wartość niezawodności dopuszczalnej.

Górna granica całkowania odpowiadająca zmiennej standaryzowanej rozkładu normalnego jest wyznaczana z zależności:

$$u_j = \frac{x_j}{\sqrt{D^2(\psi_j)}}.$$

Korzystając z tablic rozkładu $N(0, 1)$, można obliczyć wartości u_j, a następnie poszukiwane wartości wymaganych opóźnień na poszczególnych działkach roboczych.

Opóźnienie d ciągu 2 należy zwiększyć o rezerwę równą:

$$x = \max_{j=1, 2, \ldots, m} x_j ,$$

tak aby w wymaganym stopniu chronić ciągłość pracy brygady 2 i zapewnić dostępność frontu robót na każdej działce.

2. Weryfikacja metody

W celu weryfikacji poprawności przyjętych założeń w proponowanej metodzie przeprowadzono badania symulacyjne realizacji dwóch przedsięwzięć (przykład 1 i 2). Dane do przykładów zamieszczono w tabelach 1 i 2.

Tabela 1

Oszacowania czasów wykonania procesów na działkach roboczych [j.cz.] (przykład 1)

Nr działki j	Czas wykonania procesu 1			Czas wykonania procesu 2		
	t_{1j}^a	t_{1j}^m	t_{1j}^b	t_{2j}^a	t_{2j}^m	t_{2j}^b
1	3,00	4,00	6,00	1,00	3,00	6,00
2	5,00	7,00	11,00	5,00	6,00	8,00
3	6,00	8,00	12,00	7,00	10,00	11,00
4	7,00	9,00	12,00	7,00	8,00	9,00
5	4,00	6,00	9,00	5,00	7,00	11,00

Tabela 2

Oszacowania czasów wykonania procesów na działkach roboczych [j.cz.] (przykład 2)

Nr działki j	Czas wykonania procesu 1			Czas wykonania procesu 2		
	t_{1j}^a	t_{1j}^m	t_{1j}^b	t_{2j}^a	t_{2j}^m	t_{2j}^b
1	15,00	24,00	30,00	23,50	28,00	31,00
2	20,00	27,00	32,00	26,67	19,00	24,00
3	18,00	22,00	26,00	22,00	16,00	22,00
4	16,00	19,00	25,00	19,50	19,00	25,00
5	30,00	35,00	42,00	35,33	23,00	26,00

Symulator został sporządzony w języku *GPSS*, a badania symulacyjne zostały przeprowadzone w programie *GPSS World* firmy *Minuteman Software*.

Rozkłady generowano, wykorzystując predefiniowane generatory w postaci: *beta*(*RN*, t_a, t_b, α, β), gdzie *RN* – numer generatora liczb pseudolosowych. Aby zapewnić niezależność zmiennych losowych czasu trwania procesów, wykorzystano generatory liczb pseudolosowych o różnych numerach.

Parametry kształtu rozkładu *beta* PERT czasów wykonania procesów określono według następujących formuł [Davis 2008]:

$$\alpha = \left(\frac{t_{ij}^m - t_{ij}^a}{t_{ij}^b - t_{ij}^a} \right) \left(\frac{\left(t_{ij}^m - t_{ij}^a\right)\left(t_{ij}^b - t_{ij}^m\right)}{\sigma_{ij}^{\,2}} - 1 \right),$$

$$\beta = \left(\frac{t_{ij}^b - t_{ij}^m}{t_{ij}^m - t_{ij}^a} \right) \cdot \alpha .$$

Wyniki obliczeń i badań symulacyjnych zestawiono w tabelach 3 oraz 4. W obu przykładach przeprowadzono taką samą liczbę przebiegów symulacyjnych (1000 przebiegów; szerokość przedziału ufności dla wartości średniej czasu realizacji przedsięwzięcia około 0,4 j.cz.). Maksymalne wymagane wielkości opóźnień x_j uzyskano w przykładzie 1 na działce nr 3, natomiast w przykładzie 2 – na działce nr 5.

W obu przykładach częstość przestojów w pracy brygady 2 na działkach, ustalona w badaniach symulacyjnych, jest najbardziej zbliżona do założeń w przypadku dużej wartości ryzyka dopuszczalnego. Mniejsze wartości częstości od założonych (lepszą ochronę ciągłości pracy brygady 2 na działce z maksymalną wartością opóźnienia x_j) uzyskano w przypadku mniejszych poziomów ryzyka dopuszczalnego. Wynika to z wpływu przerw w pracy brygady 2 na działkach poprzedzających działkę z największą wartością obliczonego opóźnienia.

Przerwy te powodują przesunięcie terminów rozpoczynania procesu 2 na kolejnych frontach, zatem zwiększają rzeczywistą wielkość rezerwy czasu. Z drugiej strony zakłócenia ciągłości realizacji procesu 2 (przy małej wartości ryzyka dopuszczalnego) są źródłem zwiększenia wartości oczekiwanej czasu realizacji przedsięwzięcia w stosunku do wartości ustalonej analitycznie, bez uwzględnienia tego wpływu.

Tabela 3

Wyniki obliczeń i badań symulacyjnych (przykład 1)

R_{dop}	x [j.cz.]	Wartość oczekiwana czasu realizacji przedsięwzięcia ustalona analitycznie [j.cz.]	Wartość oczekiwana czasu realizacji przedsięwzięcia ustalona w badaniach symulacyjnych [j.cz.]	Częstość przerw w pracy brygady 2 pomiędzy zakończeniem robót na działce nr 2 i rozpoczęciem pracy na działce nr 3
0,500	0	44,68	45,79	0,513
0,600	0,45	45,28	46,02	0,594
0,700	0,93	45,78	46,41	0,697
0,800	1,50	46,33	46,72	0,795
0,900	2,29	47,12	47,35	0,876
0,950	2,93	47,76	47,91	0,936
0,990	5,54	50,37	50,44	0,997

Tabela 4

Wyniki obliczeń i badań symulacyjnych (przykład 2)

R_{dop}	x [j.cz.]	Wartość oczekiwana czasu realizacji przedsięwzięcia ustalona analitycznie [j.cz.]	Wartość oczekiwana czasu realizacji przedsięwzięcia ustalona w badaniach symulacyjnych [j.cz.]	Częstość przerw w pracy brygady 2 pomiędzy zakończeniem robót na działce nr 4 i rozpoczęciem pracy na działce nr 5
0,500	0	153,33	156,35	0,564
0,600	1,49	153,83	156,90	0,613
0,700	3,11	156,44	157,77	0,694
0,800	5,02	158,36	159,15	0,792
0,900	7,65	160,99	161,36	0,888
0,950	9,81	163,14	163,33	0,933
0,990	17,34	170,67	170,70	0,999

Podsumowanie

Uwzględnienie wpływu warunków losowych przy harmonogramowaniu przedsięwzięć pozwala na lepsze odzwierciedlenie rzeczywistych warunków realizacyjnych i zwiększenie odporności planów na zakłócenia. Zastosowanie w praktyce modeli i metod probabilistycznych i symulacyjnych jest ograniczone m.in. ze względu na dostęp do wiarygodnych danych. Normy pracochłonności, które stanowią zwykle podstawę do ustalenia czasu realizacji procesów, dostar-

czają jedynie informacji o wartości mediany rozkładu nakładów czasu pracy. Dokładne ustalenie parametrów i typu rozkładów wymagałoby przeprowadzenia odpowiedniej (dużej) liczby badań chronometrażowych procesów dla przyjętego poziomu ufności, zapewniającej uzyskanie właściwej wiarygodności statystycznej wyników. Niezbędne jest zatem rozwijanie metod uproszczonych, o małej złożoności obliczeniowej, lecz jednocześnie mniej dokładnych ze względu na założenia upraszczające.

W proponowanej metodzie wpływ na dokładność ustalenia niezbędnej rezerwy mają: opinie ekspertów przy określaniu czasu pesymistycznego, optymistycznego i najbardziej prawdopodobnego, założenie o typie rozkładu czasu procesów i ciągu czynności, nieuwzględnienie wpływu przestojów brygady na działkach poprzedzających oraz korelacji między zmiennymi losowymi czasu. Błąd oszacowania rezerwy (wynikający z założeń upraszczających), zweryfikowany w badaniach symulacyjnych, wydaje się być dopuszczalny na poziomie zastosowań inżynierskich.

Bibliografia

Biruk S., Jaśkowski P., 2010: Harmonogramowanie przedsięwzięć wieloobiektowych z ciągłą realizacją procesów na działkach roboczych. „Zeszyty Naukowe Wyższej Szkoły Oficerskiej Wojsk Lądowych im. gen. T. Kościuszki we Wrocławiu", Rocznik XLII, 3(157), lipiec-wrzesień.

Davis R., 2008: Teaching Project Simulation in Excel Using PERT-Beta Distributions. „INFORMS Transactions on Education" 2008, Vol. 8(3).

Dyżewski A., 1965: Technologia i organizacja budowy. Arkady, Warszawa.

Harris R.B., Ioannou P.G., 1998: Scheduling Project with Repeating Activities. „Journal of Construction Engineering and Management", Vol. 124 (4).

Hejducki Z., 2000: Sprzężenia czasowe w metodach organizacji złożonych procesów budowlanych. Prace Naukowe Instytutu Budownictwa Politechniki Wrocławskiej, Monografie nr 34, Wrocław.

Jaworski K.M., 1999: Metodologia projektowania realizacji budowy. Wydawnictwo Naukowe PWN, Warszawa.

Jaśkowski P., Biruk S., 2005: Analiza algorytmów minimalizacji przestoju brygad roboczych przy ciągłej realizacji obiektów budowlanych. „Przegląd Budowlany", nr 11.

Jaśkowski P., Biruk S., 2011: The Method for Improving Stability of Construction Project Schedules through Buffer Allocation. „Technological and Economic Development of Economy", Vol. 17, No. 3.

Kapliński O., 1974: Niektóre problemy metody pracy równomiernej w warunkach stochastycznych. Materiały XX Jubileuszowej Konferencji Naukowej KILiW PAN i KN PZITB, Kraków-Krynica.

Marcinkowski R., 1990: Harmonogramowanie zadań inżynieryjno-budowlanych według wybranych kryteriów decyzyjnych. WAT, Warszawa (praca doktorska).

Marcinkowski R., 2002: Metody rozdziału zasobów realizatora w działalności inżynieryjno-budowlanej. WAT, Warszawa.

Mrozowicz J., 1997: Metody organizacji procesów budowlanych uwzględniające sprzężenia czasowe. Dolnośląskie Wydawnictwo Edukacyjne, Wrocław.

Orłowski Z., 1997: Harmonogramowanie realizacji robót remontowych metodą programowania dynamicznego. Materiały Międzynarodowej Konferencji Naukowej z cyklu: Mieszkanie XXI wieku, Białystok 8-10 V.

Rowiński L., 1982: Organizacja produkcji budowlanej. Arkady, Warszawa.

CONSIDERING RISK IN PROJECT SCHEDULING

Summary

Scheduling projects of linear or repetitive character (roads, pipelines, high-rise buildings) involves harmonizing a number of continuous construction processes to be conducted by specialized crews or machine sets executed at the same time in a number of work sections. Such projects are often modeled by time-distance diagrams that are represented graphically as an X-Y plot where one axis represents location, and the other time. Project planning involves allowing for construction-specific risks and is aimed at providing reliable schedules. These are to help the manager to assure that the project is completed by the predefined due date and, at the same time, that interruptions in work flow are avoided. In the case of repetitive processes, schedule robustness can be improved by providing time buffers between consecutive activities. The paper proposes an analytic method of sizing these buffers that assumes (as in PERT) that activity durations are stochastic variables whose distribution parameters can be defined on the basis of optimistic, pessimistic and most likely estimates. The method was used to construct a case-study schedule, and the schedule quality was tested by means of simulation.

Barbara Gładysz
Politechnika Wrocławska

ROZMYTE LICZBY PRZEDZIAŁOWE W HARMONOGRAMOWANIU PRZEDSIĘWZIĘĆ METODĄ ŁAŃCUCHA KRYTYCZNEGO

Wstęp

Pierwszą przedstawioną w literaturze techniką planowania przedsięwzięć jest metoda deterministyczna CPM (*Critical Path Method).* Dzisiaj równolegle funkcjonują probabilistyczne i rozmyte techniki harmonogramowania. W latach 1956-1957 zaproponowano metodę PERT (*Program Evaluation and Review Technique*), w której przyjmuje się, że czas realizacji zadań jest rozkładem beta. W literaturze są rozważane również metody analizy czasowej przedsięwzięć dla przypadków, gdy czasy zadań mają inne rozkłady prawdopodobieństwa, np.: rozkład wykładniczy, rozkład Weibulla, rozkład Golenko-Ginzburga, rozkład beta z grubymi ogonami. Równolegle są rozwijane metody analizy czasowej przedsięwzięć dla czasów zadań zadanych w postaci liczb rozmytych. Przegląd metod oraz technik harmonogramowania przedsięwzięć można znaleźć w pracy Kuchty [2011].

W roku 1997 Goldratt zaproponował metodę łańcucha krytycznego, w której zastosował teorię ograniczeń w konstrukcji bezpiecznego harmonogramu przedsięwzięć [Goldratt 1997]. Proponuje on uwzględnienie przy konstrukcji harmonogramu syndromu studenta oraz prawa Parkinsona, zgodnie z którymi: czas zadań jest przeszacowywany, a zadania są realizowane w ostatniej chwili. Goldratt przyjmuje planowany czas realizacji zadania na poziomie 50% kwantyla 90% (80%) przewidywanego czasu zadania. Ponadto proponuje on przyjęcie bufora czasu zabezpieczającego przed ryzykiem niedotrzymania terminu projektu na poziomie 50% planowanego czasu zadań. Takie założenia prowadzą często do bardzo dużych buforów projektu, a tym samym długiego planowanego terminu zakończenia całego projektu.

W literaturze zaproponowano wiele modyfikacji pierwotnej koncepcji łańcucha krytycznego. Ashtiani i in. [2007] oraz Fallah i in. [2010] rozważają harmonogramy przedsięwzięć dla logarytmicznie normalnego rozkładu czasu realizacji zadań. Millian [2005] analizuje własności buforów czasu łańcuchów projektu dla wybranych rozkładów prawdopodobieństwa czasów zadań. Konstrukcji harmonogramów w sytuacji, gdy czasy zadań są zadane jako liczby rozmyte, jest poświęconych wiele prac [m.in. Ma Guo-Feng, Yiang Yong-Bin 2012; Kulejewski i in. 2011; Zhao i in. 2008]. Połoński i Pruszyński [2008a, 2008b] wprowadzają dodatkowe bufory w łańcuchu krytycznym zabezpieczające przed możliwością powstania w trakcie realizacji projektu dodatkowych łańcuchów krytycznych.

W niniejszym artykule do opisu charakterystyki czasowej przedsięwzięcia wykorzystamy dwa rodzaje opisu niepewności: probabilistyczną i rozmytą.

W probabilistycznej analizie przedsięwzięć PERT przyjmuje się, że czas wykonania zadań ma rozkład beta na odcinku $[a, b]$ o funkcji gęstości:

$$f(t) = \frac{1}{B(\alpha,\beta)} \frac{(t-a)^{\alpha-1}(b-t)^{\beta-1}}{(b-a)^{\alpha+\beta-1}} \tag{1}$$

Wartość oczekiwana, wariancja i wartość najbardziej prawdopodobna w rozkładzie beta wyrażają się wzorami (dla $\alpha > 1$, $\beta > 1$):

$$E(T) = \frac{\alpha b+\beta a}{\alpha+\beta} \tag{2}$$

$$D^2(T) = \frac{\alpha\beta(b-a)^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{3}$$

$$M(T) = \frac{(\alpha-1)b+(\beta-1)a}{\alpha+\beta-2} \tag{4}$$

Parametry rozkładu czasu wykonania poszczególnych zadań przedsięwzięcia wyznacza się na podstawie wiedzy ekspertów. Ekspert podaje trzy charakterystyki czasowe: ocenę optymistyczną a, pesymistyczną b i najbardziej prawdopodobną (dominantę) m. W klasycznym podejściu PERT przyjmuje się, że wartość oczekiwana i wariancja czasu zadania są następujące:

$$E(T) = \frac{a+4m+b}{6} \tag{5}$$

$$D^2(T) = \left(\frac{b-a}{6}\right)^2 \tag{6}$$

Krytyczne uwagi do stosowania wzorów (5)-(6) w praktyce dotyczą m.in. faktu, że wszystkie zadania, dla których różnica pomiędzy pesymistycznym i optymistycznym czasem jest jednakowa, mają taką samą wariancję. Kolejnym założeniem przyjętym w metodzie PERT jest założenie, że eksperci znają naturę rozkładu beta i podane przez nich oceny czasu są zgodne z własnościami tego rozkładu. Tak nie jest zawsze. Często oszacowane przez ekspertów charakterystyki czasu zadań nie spełniają układu równań (2)-(4).

Wzory (5) i (6) określają odpowiednio wartość oczekiwaną i wariancję rozkładu beta, wówczas gdy $\alpha = 3 \pm \sqrt{2}$, $\beta = 3 \mp \sqrt{2}$ lub gdy $\alpha = \beta = 4$. Jeżeli $\alpha = \beta = 4$, to rozkład beta jest rozkładem symetrycznym [Grubbs 1962]. Gdy $\alpha = 3 - \sqrt{2}$, $\beta = 3 + \sqrt{2}$, to rozkład beta jest rozkładem prawoskośnym, a gdy $\alpha = 3 + \sqrt{2}$, $\beta = 3 - \sqrt{2}$ – rozkładem lewoskośnym. Jeżeli $\alpha = 3 \pm \sqrt{2}$, $\beta = 3 \mp \sqrt{2}$, to zgodnie ze wzorem (4) wartość najbardziej prawdopodobna jest równa $M(T) = \frac{(2\pm\sqrt{2})a+(2\mp\sqrt{2})b}{4}$. Jak wspomniano wcześniej, często jednak podana przez ekspertów najbardziej prawdopodobna wartość czasu zadania m nie odpowiada teoretycznej wartości dominanty $M(T)$ rozkładu beta.

Przedstawimy teraz elementy teorii zbiorów rozmytych. Koncepcję zbioru rozmytego zaproponował w 1965 r. Zadeh [1965]:

Przedziałową liczbą rozmytą $\tilde{X}$ nazywamy rodzinę rzeczywistych przedziałów domkniętych $[\tilde{X}]_\lambda$, gdzie $\lambda \in [0,1]$ taką, że: $\lambda_1 < \lambda_2 \Rightarrow [\tilde{X}]_{\lambda_1} \subset [\tilde{X}]_{\lambda_2}$ oraz $I \subseteq [0,1] \Rightarrow [\tilde{X}]_{\sup I} = \bigcap_{\lambda \in I} [\tilde{X}]_\lambda$. Przedział $[\tilde{X}]_\lambda$ dla ustalonego $\lambda \in [0,1]$ nazywa się λ-poziomem liczby rozmytej $\tilde{X}$. Będziemy go oznaczać jako $[\tilde{X}]_\lambda = \left[\underline{x}(\lambda), \overline{x}(\lambda)\right]$.

Dubois i Prade [1978] wprowadzili następującą użyteczną definicję klasy przedziałowych liczb rozmytych typu L-R. Przedziałową liczbę rozmytą $\tilde{X}$ nazywa się liczbą rozmytą typu L-R, jeśli jej funkcja przynależności przyjmuje postać:

$$\mu_X(x) = \begin{cases} L\left(\frac{\underline{m}-x}{\underline{m}-a}\right) & \text{dla} \quad x < \underline{m} \\ 1 & \text{dla} \quad \underline{m} \le x \le \overline{m} \\ R\left(\frac{x-\overline{m}}{b-\overline{m}}\right) & \text{dla} \quad x > \overline{m} \end{cases} \tag{7}$$

gdzie: $L(x)$, $R(x)$ – ciągłe nierosnące funkcje x.

Funkcje $L(x)$, $R(x)$ są zwane funkcjami kształtu liczby rozmytej. Najczęściej stosowanymi postaciami funkcji kształtu są: $\max\{0,1-x^p\}$ oraz

$\exp(-x^p)$, $x \in [0,+\infty)$, $p \geq 1$. Przedziałową liczbę rozmytą, dla której $L(x)$, $R(x) = \max\{0,1-x^p\}$ oraz $\underline{m} = \overline{m} = m$, nazywa się trójkątną liczbą rozmytą.

Medianą $ME(\tilde{X})$ liczby rozmytej $\tilde{X}$ typu L-R nazywamy liczbę rzeczywistą Me, dla której jest spełniona równość [Bodjanova 2005]:

$$\int_{-\infty}^{Me} \mu_X(x)dx = \int_{Me}^{+\infty} \mu_X(x)dx = 0{,}5 \cdot card(\tilde{X}) \tag{8}$$

gdzie $card(\tilde{X})=\int_{-\infty}^{+\infty} \mu_X(x)dx$ – moc zbioru rozmytego $\tilde{X}$.

Analogicznie można zdefiniować pojęcie kwantyla rzędu q liczby rozmytej $\tilde{X}$ jako liczbę rzeczywistą, dla której zachodzi:

$$\int_{-\infty}^{q} \mu_X(x)dx = q \cdot card(\tilde{X}) \tag{9}$$

Przypuśćmy teraz, że mamy dwa zbiory rozmyte $\tilde{X}$, $\tilde{Y}$. Funkcja przynależności przekroju tych zbiorów $\tilde{Z} = \tilde{X} \cap \tilde{Y}$ ma postać [Zadeh 1965]:

$$\mu_Z(x) = min(\mu_X(x), \mu_Y(x)) \tag{10}$$

Niech $\tilde{X}, \tilde{Y}$ będą dwiema liczbami rozmytymi o funkcjach przynależności odpowiednio $\mu_X(x), \mu_Y(y)$. Wówczas, zgodnie z zasadą rozszerzania Zadeha [1965], funkcja przynależności sumy $Z = X + Y$ przyjmuje postać:

$$\mu_Z(z) = \sup_{z=x+y} \left(min(\mu_X(x), \mu_Y(y))\right) \tag{11}$$

Jeżeli chcemy porównać dwie liczby rozmyte, tzn. chcemy określić możliwość, że realizacja $\tilde{X}$ (wartość przyjęta przez $\tilde{X}$) będzie większa równa (nie mniejsza) od realizacji $\tilde{Y}$, to możemy skorzystać z indeksu zaproponowanego przez Dubois i Prade [1988]:

$$\mathrm{Pos}(X \geq Y) = \sup_{x \geq y} \left(\min\left(\mu_X(x), \mu_Y(y)\right)\right) \tag{12}$$

1. Rozmyta metoda łańcucha krytycznego

Załóżmy, że ekspert podaje trzy oceny czasu zadania $(i,j) \in A$: czas minimalny (a_{ij}), czas maksymalny (b_{ij}) oraz wartość najpewniejszą m_{ij} (lub: $[\underline{m_{ij}}, \overline{m_{ij}}]$).

Oceny te wykorzystamy do konstrukcji rozkładu prawdopodobieństwa (dystrybuanty $F_{ij}(t)$) czasu zadania $T_{i,j}$ według następujących reguł. Jeżeli we-

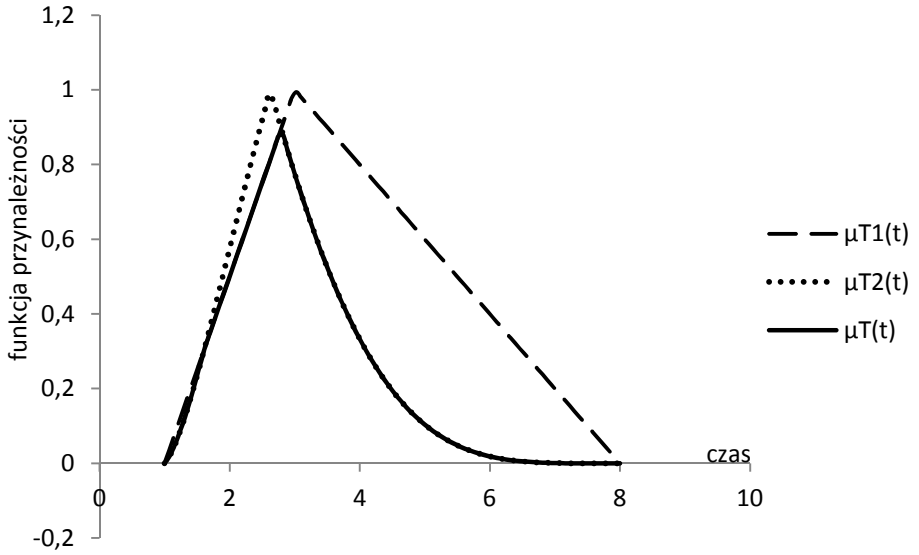
dług eksperta mamy do czynienia z rozkładem prawoskośnym, to w estymacji rozkładu prawdopodobieństwa beta przyjmiemy $\alpha = 3 - \sqrt{2}$, $\beta = 3 + \sqrt{2}$. Jeżeli czas zadania w opinii eksperta jest rozkładem lewoskośnym, to $\alpha = 3 + \sqrt{2}$, $\beta = 3 - \sqrt{2}$. Jeżeli natomiast czas zadania jest rozkładem symetrycznym, to przyjmiemy $\alpha = \beta = 4$.

Skonstruujemy teraz dwie rozmyte oceny czasu $\tilde{T}_{ij}^1$, $\tilde{T}_{ij}^2$ zadania (i,j). Pierwszą ocenę czasu $\tilde{T}_{ij}^1$ wyznaczymy ad hoc na podstawie wzoru (7), przyjmując np. liniowy kształt funkcji przynależności $\mu_{T_{ij}^1}(t)$. Drugą ocenę $\tilde{T}_{ij}^2$ skonstruujemy na podstawie rozkładu prawdopodobieństwa czasu zadania $F_{ij}(t)$ metodą zaproponowaną przez Beliakova [1996], przyjmując za funkcję przynależności $\mu_{T_{ij}^2}(t)$ unormowaną funkcję:

$$\mu_{T_{ij}^2}(t) = 2 * min\{F_{ij}(t), 1 - F_{ij}(t))\} \tag{13}$$

Dysponujemy teraz dwiema rozmytymi ocenami czasu zadania. Pierwsza jest oceną ad hoc wyznaczoną na podstawie podanych przez eksperta trzech ocen czasu zadania. Druga ocena czasu zadania została określona na postawie rozkładu prawdopodobieństwa beta z uwzględnieniem charakterystyki czasu zadania podanej przez eksperta. Za funkcję przynależności czasu zadania (i,j) przyjmiemy funkcję przynależności zbioru rozmytego $\tilde{T}_{i,j} = \tilde{T}_{ij}^1 \cap \tilde{T}_{ij}^2$.

Przypuśćmy, że ekspert podał następujące oceny czasu pewnego zadania: $a = 1$, $m = 3$, $b = 8$. Według ocen czasu podanych przez eksperta rozkład prawdopodobieństwa czasu zadania jest rozkładem prawoskośnym (rozkład beta dla $\alpha = 3 - \sqrt{2}$ i $\beta = 3 + \sqrt{2}$). Na rysunku 1 przedstawiono trzy funkcje przynależności czasu zadania (trójkątny rozkład ad hoc) $\mu_{T^1}(t)$, funkcję przynależności $\mu_{T^2}(t)$ wyznaczoną metodą Beliakova na podstawie rozkładu prawdopodobieństwa beta zgodnie ze wzorem (13) oraz funkcję przynależności przekroju $\tilde{T}_{i,j}$ zbiorów rozmytych $\tilde{T}^1$, $\tilde{T}^2$: $\mu_T(t) = min(\mu_{T^1}(t), \mu_{T^1}(t))$.



Rys. 1. Funkcje przynależności czasu zadania

Tak wyznaczoną ocenę czasu zadania wykorzystamy w konstrukcji harmonogramu przedsięwzięcia. Niżej przedstawimy odpowiedni algorytm.

Algorytm konstrukcji harmonogramu przedsięwzięcia

Krok 1. Na podstawie podanych przez ekspertów charakterystyk czasów a_{ij}, b_{ij}, m_{ij} zadań (i,j) przedsięwzięcia: dla każdego zadania wyznacz rozkład prawdopodobieństwa beta czasu (przyjmując parametry α, β zgodnie z podaną przez eksperta asymetrią) oraz wyznacz na podstawie wzorów (7) i (13) funkcje przynależności rozmytych ocen czasu $\tilde{T}_{ij}^1$, $\tilde{T}_{ij}^2$.

Krok 2. Dla każdego zadania wyznacz funkcję przynależności czasu zadania $\tilde{T}_{i,j} = \tilde{T}_{ij}^1 \cap \tilde{T}_{ij}^2$ zgodnie ze wzorem (10).

Krok 3. Przyjmij, przez analogię do propozycji Goldratta, że planowanym czasem $T(i,j)$ realizacji zadania (i,j) jest mediana $ME(\tilde{T}_{i,j})$ czasu $\tilde{T}_{i,j}$ (wzór 8).

Krok 4. Wyznacz łańcuch krytyczny CC oraz zapasy czasu zadań leżących na łańcuchach zasilających, przyjmując za czasy realizacji zadań $Me(i,j) = ME(\tilde{T}_{i,j})$.

Krok 5. Dla każdego zadania wyznacz bufor czasu $B(i,j)$ jako połowę różnicy między kwantylem 0,9 (wzór 10) a medianą czasu zadania $Me(i,j)$.

Krok 6. Wyznacz bufor projektu BP jako sumę buforów czasów zadań leżących na łańcuchu krytycznym.

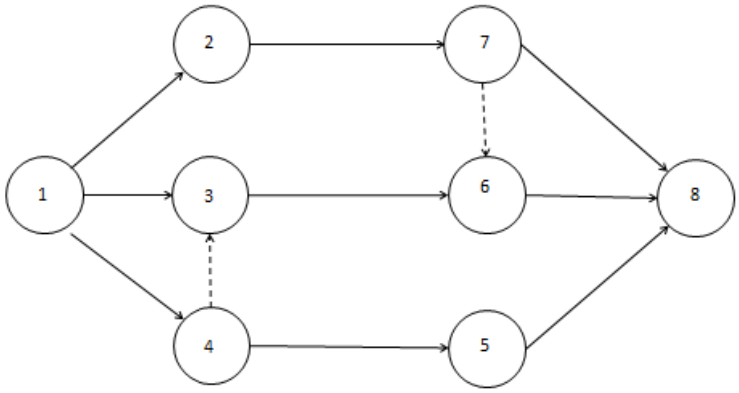
Krok 7. Dla każdego łańcucha zasilającego wyznacz BZS (suma buforów czasów zadań leżących na łańcuchu) oraz bufor łańcucha $BZ = min(BZS, FF)$, gdzie FF jest swobodnym zapasem czasu ostatniego zadania w łańcuchu zasilającym.

Krok 8. Wyznacz funkcję przynależności czasu realizacji projektu, tj. funkcję przynależności sumy czasów wykonania zadań łańcucha krytycznego *CC*.

Zaproponowany algorytm wymaga obliczeń numerycznych. Nie można bowiem wyznaczyć analitycznej postaci dystrybuanty rozkładu beta, jak również nie można wyznaczyć analitycznej postaci mediany czasów zadań czy analitycznej postaci funkcji przynależności czasu realizacji projektu. Zaletą zaproponowanego algorytmu jest uwzględnienie w analizie czasowej przedsięwzięcia wszystkich trzech ocen czasów zadań podanych przez eksperta zgodnie z jego intuicją.

2. Przykład

Rozważmy przykład z pracy Li, Chen [2007]. Przedsięwzięcie jest złożone z 8 zadań. Do realizacji zadań są wymagane zasoby A, B, C, D, E, F. Graf przedsięwzięcia z uwzględnieniem zależności technologicznych oraz ograniczeń na zasoby (łuki opisane linią przerywaną) przedstawiono na rys. 2. Zmodyfikowane dane czasowe podano w tabeli 1, przyjmując wartość $m_{ij} = 1/2(\underline{m}_{ij} + \overline{m}_{ij})$.



Rys. 2. Graf przedsięwzięcia

Tabela 1

Charakterystyka zadań projektu

Zadanie	Zasób	Charakterystyka czasowa			
		a	m	b	Skośność
1	A	2	5	8	Symetryczny
2	B	1	5	10	Prawoskośny
3	C	12	18	30	Prawoskośny
4	C	3	4,5	6	Symetryczny
5	D	8	11	16	Prawoskośny
6	E	6	9	12	Symetryczny
7	E	10	25	40	Symetryczny
8	F	2	5	8	Symetryczny

Źródło: Opracowanie własne na podstawie: Li, Chen [2007].

Charakterystyki czasowe projektu wyznaczone zgodnie z podanym algorytmem przestawiono w tabeli 2.

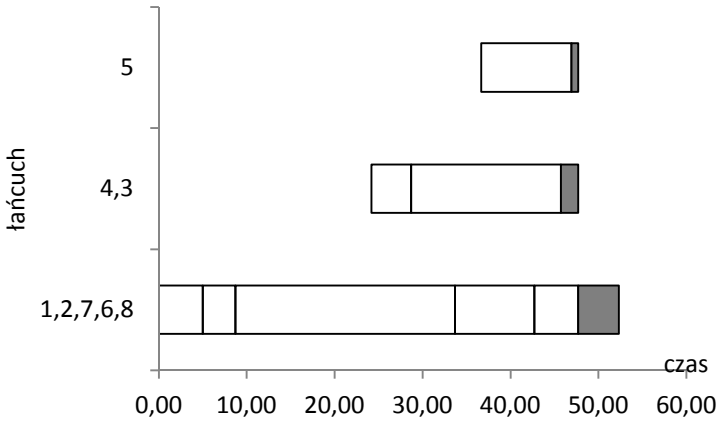
Tabela 2

Charakterystyka czasowa projektu

Zadanie	Zasób	Charakterystyka czasowa								
		a	m	b	Mediana	Kwantyl 0,9	Bufor	ES	LF	FF
1	A	2	5	8	5	5,96	0,48	0	5	0
2	B	1	5	10	3,7	5,40	0,85	5	8,7	0
3	C	12	18	30	17,03	20,5	1,74	9,5	33,7	7,17
4	C	3	4,5	6	4,5	5	0,25	5	29,2	19,7
5	D	8	11	16	10,26	11,8	0,77	9,5	42,7	22,94
6	E	6	9	12	9	9,94	0,48	33,7	42,7	0
7	E	10	25	40	25	29,7	2,35	8,7	33,7	0
8	F	2	5	8	5	5,96	0,36	42,7	47,7	0

Źródło: Opracowanie własne.

Łańcuchem krytycznym w analizowanym przedsięwzięciu jest ciąg zadań 1,2,7,6,8. Długość łańcucha krytycznego wynosi 47,7, a bufor projektu 4,63. W projekcie mamy dwa łańcuchy zasilające. Pierwszy jest złożony z zadań 4 i 3 z buforem 1,99, a drugi z jednego zadania 5 z buforem 0,77. Na rysunku 3 przedstawiono harmonogram projektu z zaznaczeniem buforów czasu.

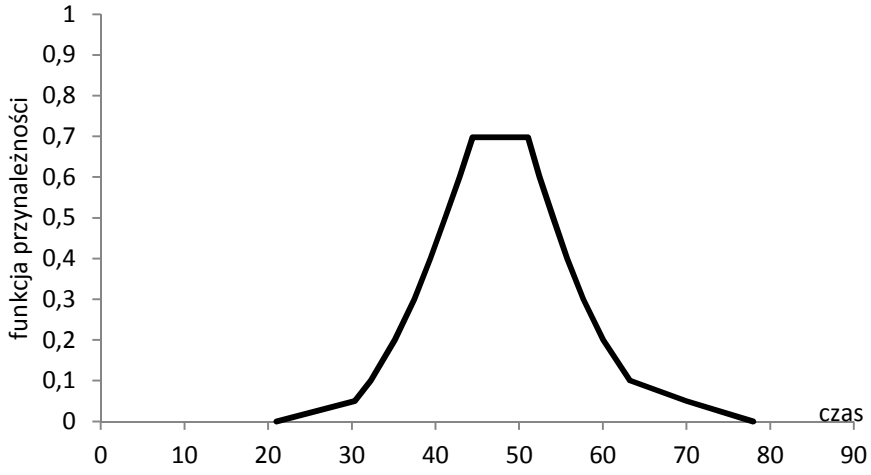


Rys. 3. Harmonogram projektu z zaznaczonymi buforami czasu

Wyznaczmy teraz funkcję przynależności czasu ukończenia projektu $\tilde{T}$, tj. czasu wykonania zadań łańcucha krytycznego *CC*:

$$\mu_T(t) = \mu_{\sum_{(i,j)\in CC} T_{ij}}(t)$$

Czasy realizacji zadań projektu nie są liczbami typu L-R. Funkcję przynależności czasu ukończenia projektu wyznaczymy numerycznie. Zastosujemy tu metodę sumowania liczb rozmytych na λ-poziomach (dla λ = 0, 0,05, 0,1, …, 1). Tak aproksymowaną funkcję przynależności czasu zakończenia projektu $\mu_T(t)$ przedstawiono na rys. 4.



Rys. 4. Funkcja przynależności czasu realizacji projektu

Możliwość, że projekt ukończymy w planowanym terminie, wynosi $Pos(\tilde{T} <= 47{,}7) = 0{,}7$. Podobnie możliwość ukończenia projektu w czasie 50,33 (planowany czas wykonania zadań łańcucha krytycznego + bufor czasu) to 0,7.

Jeżeli harmonogram dla tego projektu wyznaczylibyśmy metodą zaproponowaną przez Goldratta, planowany czas realizacji wynosiłby 73,875 [Chen, Li 2007], czyli byłby znacznie dłuższy. Goldratt przyjmuje znacznie dłuższy bufor projektu.

Podsumowanie

W artykule do opisu czasu zadań przedsięwzięcia wykorzystano aparat rachunku prawdopodobieństwa oraz aparat zbiorów rozmytych. Kształt funkcji przynależności czasów zadań estymuje się na podstawie podanych przez ekspertów trzech ocen czasów. Harmonogram przedsięwzięcia konstruuje się na podstawie wartości kwantyli liczb rozmytych. Należy podkreślić, że z uwagi na postać rozkładu prawdopodobieństwa beta wykorzystywanego w estymacji funkcji przynależności czasów zadań, zaproponowana metoda wymaga obliczeń numerycznych.

Bibliografia

Ashtiani B., Jajali G.R., Aryanezhad M.B., Makui A., 2007: A New Approach for Buffer Sizing in Critical Chain Scheduling. Proceedings of the 2007 IEEE Industrial Engineering and Engineering Management, s. 1037-1041.

Beliakov G., 1996: Fuzzy Sets and Memberships Functions Based on Probabilities. „Information Sciences", Vol. 91, Iss. 1-2, s. 95-111.

Bodjanova S., 2005: Median Value and Median Interval of a Fuzzy Number. „Information Sciences", 172, s. 73-89.

Dubois D., Prade H., 1978: Algorithmes de plus courts chemins pour traiter des donnees floues. RAIRO Recherché Operationelle/Operations Research 12, s. 213-227.

Dubois D., Prade H., 1988: Possibility Theory: An Approach to Computerized Processing of Uncertainty. Plenum Press, New York.

Fallah M., Ashiani B., Aryanezhad M.B., 2010: Critical Chain Project Scheduling: Utilizing Uncertainty for Buffer Sizing. „International Journal of Research and Reviews in Applied Sciences", Vol. 3, No. 3, s. 280-289.

Goldratt E.M., 1977: Critical Chain. The North River Press, Great Barrington, MA.

Kuchta D., 2011: Zagadnienie czasu i kosztów w zarządzaniu projektami. Wybrane metody planowania i kontroli. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.

Kulejewski J., Ibadov N., Zieliński B., 2011: Zastosowanie teorii zbiorów rozmytych w harmonogramowaniu robót budowlanych metodą łańcucha krytycznego. „Budownictwo i Inżynieria Środowiska", 2, s. 331-338.

Li K., Chen Y., 2007: Proceedings of the 2007 Applying Critical Chain in Project Scheduling and Estimating Buffer Size Based on Fuzzy Technique, Proceedings of the 2007 IEEE Industrial Engineering and Engineering Management, s. 1068-1072.

Ma G.-F., Yiang Y.-B., 2012: Improved Fuzzy Critical Chain Buffer Method. International Conference on Information Management, Innovation Management and Industrial Engineering, s. 240-243.

Millian Z., 2005: Notes of Time Buffers' Estimations in CCPM. „Archives and Civil and Mechanical Engineering", Vol. V, No. 1, s. 19-38.

Połoński M., Pruszyński K., 2008a: Lokalizacja buforów czasu w metodzie łańcucha krytycznego w harmonogramach robót budowlanych (cz. I) – podstawy teoretyczne. „Przegląd Budowlany", 2, s. 45-49.

Połoński M., Pruszyński K., 2008b: Lokalizacja buforów czasu w metodzie łańcucha krytycznego w harmonogramach robót budowlanych (cz. II) – praktyczne zastosowania. „Przegląd Budowlany" 3, s. 55-62.

Zadeh L.A., Fuzzy Sets, 1965: „Information and Control", Vol. 8, s. 338-353.

Zadeh L.A., 1973: The Concept of Linguistic Variables and Its Application to Approximate Reasoning. „Information Sciences", Vol. 8, s. 199-249.

Zhao Z.-Y., You W.-Y., Lv Q.-L., 2008: Application of Fuzzy Critical Chain Method in Project Scheduling. Proceedings of Forth International Conference on Natural Computation, s. 473-477.

FUZZY INTERVAL NUMBERS IN CRITICAL CHAIN SCHEDULING

Summary

Probabilistic critical path method PERT assumes beta distribution as probability distribution of task duration. PERT assumes that experts estimate parameters of task duration ("most likely time", "optimistic time", and "pessimistic time" for each activity) according to beta distribution nature. It's not always true. In the paper we assume that the durations of tasks are fuzzy sets. On the base of expert data we construct membership function of task duration. These fuzzy sets are used for project scheduling and for buffer size calculation. The illustrative example is presented.

ANALIZA PREFERENCJI

Jakub Brzostowski
Politechnika Śląska
Ewa Roszkowska
Uniwersytet w Białymstoku

SYSTEM REKOMENDACJI DOBORU WAG KRYTERIÓW OPARTY NA ICH CHARAKTERYSTYCE PROBABILISTYCZNEJ*

Wstęp

Ważnym etapem w procesie negocjacji jest faza prenegocjacyjna obejmująca wieloaspektowe przygotowanie do negocjacji [Kersten 1985]. W ramach zadań wstępnych realizowanych podczas tej fazy decydent dokonuje strukturyzacji problemu negocjacyjnego, wyodrębnia zagadnienia negocjacyjne wraz z określeniem stopnia ich istotności, wyznacza oferty negocjacyjne, które mogą być przedmiotem rozmów, uwzględniając przy tym preferencje własne oraz oszacowane preferencje partnera negocjacji [Roszkowska 2011]. Następnym krokiem jest budowa systemu oceny ofert (np. z wykorzystaniem metod wielokryterialnego podejmowania decyzji [Hwang i Yoon 1981; Belton i Stuart 2002]) której celem jest przedstawienie z perspektywy negocjatora wypadkowej wartości oferty negocjacyjnej za pomocą jednej zagregowanej wielkości. Analiza wartości ofert negocjacyjnych z własnego punktu widzenia oraz z punktu widzenia drugiej strony pozwala odpowiednio zaplanować strategię negocjacyjną.

W sytuacji braku możliwości dokonania bezpośredniej analizy preferencji (np. nikłe doświadczenie, niepełna lub nieprecyzyjna informacja) zachodzi potrzeba wsparcia negocjatora w fazie prenegocjacyjnej w celu rozpoznania wła-

* Praca została sfinansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/03/B/HS4/03857.

snych preferencji oraz przyjęcia założeń co do preferencji drugiej strony. Niniejszy artykuł skupia się na prezentacji autorskiego narzędzia, które umożliwi analizę własnych preferencji oraz przewidywanie preferencji potencjalnego partnera negocjacji. Proponowane narzędzie docelowo ma umożliwić wgląd w preferencje typowe dla grupy, do której należy negocjator, oraz wgląd w preferencje grupy, do której zostaje zakwalifikowany partner. Mając do dyspozycji dane opisujące preferencje szerokiej populacji negocjatorów, można skonstruować modele preferencji zbiorczych* dla różnych typów (profilów) negocjatorów. Proponowany model preferencji zbiorczych jest budowany w postaci rozkładu prawdopodobieństwa nad wektorem wag kwestii (kryteriów) negocjacyjnych. Proces tworzenia profilów negocjatora jest uzależniony od dostępności danych i może brać pod uwagę zarówno cechy demograficzne, jak i psychologiczne negocjatora.

W procesie analizy preferencji są wykorzystywane metody wielokryterialnego podejmowania decyzji (WPD) [Figueira i in. 2005; Kilgour i in. 2010; Salo i Hämäläinen 2012; Brzostowski i in. 2012], przy czym wiele z tych metod stosuje pojęcie wag przypisywanych rozpatrywanym kryteriom w problemie decyzyjnym. Stanowią one istotną część opisu preferencji decydenta, a proces szacowania wag kryteriów może mu stwarzać pewne problemy. W literaturze przedmiotu opisano wiele metod subiektywnych, obiektywnych oraz integracyjnych szacowania współczynników wagowych, których zadaniem jest ułatwienie decydentowi tej części procesu analizy preferencji [Tzeng i in. 1998]. Metody subiektywne są oparte na analizie indywidualnych preferencji decydenta, obiektywne wykorzystują dane zawarte w macierzy decyzyjnej, a integracyjne łączą oba wspomniane podejścia. Do metod subiektywnych można zaliczyć: metody oparte na rangach (*rank ordering methods*) [Stillwell i in. 1981], metody kompensacyjne (*tradeoff method, pricing-out method*) [Keeney i Raiffa 1976], metodę opartą na proporcjach (*ratio method*) [Edwards 1977], metodę AHP (*Analytic Hierarchy Process*) [Saaty 1980], metodę bezpośredniej oceny (*direct rating* – DR) [Bottomley i Doyle 2001], metodę delficką (*Delphi method*) [Hwang i Yoon 1981], metodę wektorów własnych (*eigenvector method*) [Takeda i in. 1987]. Natomiast do metod obiektywnych należą: metoda oparta na odchyleniu standardowym (*standard deviation* – SD) [Diakoulaki i in. 1995], metoda współczynnika korelacji CRITIC (*Criteria Importance Through Intercriteria Correlation*) [Diakoulaki i in. 1995], metoda punktu idealnego (*ideal point*) [Ma i in. 1999], metoda maksymalizacji odchylenia standardowego (*maximizing deviation*

* Określenie „preferencje zbiorcze" oznacza preferencje osób zakwalifikowanych do rozważanego profilu. Nie należy utożsamiać określenia preferencji zbiorczych z preferencjami grupowymi czy zagregowanymi.

method) [Wu i Chen 2007] oraz metoda oparta na koncepcji entropii Shannona (*entropy method*) [Hwang i Yoon 1981].

Proponowane w pracy podejście szacowania wag kwestii negocjacyjnych wykorzystuje dane historyczne. Informacje dotyczące przyjmowanego przez różnych decydentów w tym samym problemie decyzyjnym wektora wag kwestii negocjacyjnych są ujmowane w postaci rozkładu prawdopodobieństwa wystąpienia danego wektora wag. Dostępność danych dotyczących preferencji większej grupy decydentów rozwiązujących podobny lub ten sam problem decyzyjny umożliwia analizę potencjalnych wzorców w zachowaniu takiej właśnie grupy.

Poza wsparciem podczas analizy własnych preferencji negocjator może być zainteresowany przewidywaniem preferencji potencjalnego partnera. W takiej sytuacji dla danego problemu negocjacyjnego negocjator wybiera typ partnera, dla którego narzędzie może zilustrować opis preferencji zbiorczych. Wgląd w preferencje potencjalnego partnera pozwala na lepsze przygotowanie do procesu negocjacji pod względem doboru strategii negocjacyjnej.

W artykule przedstawiono podstawy konstrukcji modelu preferencji zbiorczych opartego na ich probabilistycznej charakterystyce, opisano fazy działania systemu wspomagania analizy preferencji zbiorczych, zaproponowano moduły, które mogą być zaimplementowane w ramach proponowanego narzędzia wspomagającego. Na przykładzie empirycznym pokazano praktyczne możliwości działania systemu rekomendacji doboru wag kryteriów w problemie negocjacyjnym opartego na ich probabilistycznej charakterystyce.

1. Konstrukcja modelu preferencji zbiorczych

W pierwszej fazie szacowania własnych wag kwestii negocjacyjnych oraz przewidywania wag kwestii przyjętych przez partnera system ma pozwalać decydentowi na zakwalifikowanie siebie do danej grupy decydentów oraz określenie typu potencjalnego partnera[*]. Typologia decydenta może być dokonana ze względu na jedną lub wiele cech, co będzie wymagać wglądu w preferencje różnych grup. Systemy wspomagania negocjacji [Kersten 2007] mogą rejestrować dane decydenta dotyczące różnych cech demograficznych. Przykładowo system wspomagania Inspire [Kersten 1999] tworzy bazę danych negocjatorów, uwzględniając: typ kwalifikacji, język ojczysty, płeć, wiek, kraj pochodzenia, kraj zamieszkiwany, znajomość języka angielskiego, poziom wykształcenia. Po-

[*] Przyjęto założenie, że każdej kwestii negocjacyjnej K_j została przyporządkowana dodatnia waga w_j $w_j = w(K_j)$ określająca jego względną ważność, gdzie $w_1 + w_2 + \ldots + w_n = 1$ ($j = 1,2,\ldots n$).

nadto decydenci wypełniający kwestionariusz Thomasa-Killmana [Killman i Thomas 1985] mogą zostać zakwalifikowani do różnych typów ze względu na stosowany styl negocjacji: współpracy, kooperacji, dostosowania się, unikania, kompromisu. Powyższe zestawy cech rozpatrywane z różnych perspektyw stwarzają możliwość analizy różnych typów negocjatora poprzez wydzielanie ze zbioru danych opisujących preferencje decydenta ze względu na przynależność do danego typu. Wyodrębnione profile negocjatora stanowią podstawę do budowy probabilistycznego opisu preferencji zbiorczych.

Załóżmy, że problem p jest reprezentowany formalnie za pomocą zbioru m cech opisujących negocjatora:

$$F_p = \{f_{p1}, f_{p2}, \ldots, f_{pm}\} \tag{1}$$

Z kolei każda zmienna f_{pj} jest zmienną wskazującą, czy cecha jest przez decydenta uwzględniana (u_{pj}) oraz jaka wartość została wybrana (v_{pj}) (wartość nominalna bądź porządkowa):

$$f_{pj} = (u_{pj}, v_{pj}) \quad u_{pj} \in \{0,1\} \quad v_{pj} \in \{1,\ldots,c_j\} \subset N \tag{2}$$

Realizacja F_p odpowiada doborowi typu decydenta, co pozwala na ekstrakcję z pełnego zbioru danych podzbioru dotyczącego tego właśnie typu. Zatem funkcja $s_F : \overline{w} \mapsto \{0,1\}$ wskazuje, czy wektor wag kryteriów (opis preferencji indywidualnego decydenta) jest brany pod uwagę w budowaniu opisu preferencji zbiorczych.

Kolejna faza analizy obejmuje konstrukcję rozkładu prawdopodobieństwa nad wektorem wag kryteriów występujących w danym problemie decyzyjnym (bądź negocjacyjnym) oraz dotyczących wybranej kategorii negocjatora. Otrzymany rozkład wag kryteriów stanowi syntetyczny opis wiedzy na temat preferencji w zakresie wektora wag przypisanego określonemu profilowi negocjatora. Analiza takiego rozkładu może wskazać na pewne cechy wspólne dla samego problemu decyzyjnego w postaci dominującej wartości wagi danego kryterium czy też wektora wag, gdy rozpatruje się zestaw kryteriów razem. Wysoka specyficzność rozkładu i jedna wartość dominująca wskazuje na tendencję statystycznego decydenta do określania wagi lub wektora wag w otoczeniu wartości występującej najczęściej. Analiza rozkładu wag lub wektora wag pozwala na określenie typu takiego rozkładu, co z kolei sprzyja możliwości opisywania preferencji zbiorczych za pomocą kilku charakterystyk liczbowych, takich jak war-

tość oczekiwana oraz odchylenie standardowe. Syntetyczny i prosty opis preferencji zbiorczych pozwala z kolei na przedstawienie potencjalnemu decydentowi sposobu postępowania innych decydentów w trakcie analizy preferencji bez ujawniania preferencji indywidualnych. Taka ilustracja pomaga decydentowi w procesie analizy jego własnych preferencji, zwłaszcza gdy mamy do czynienia z brakiem wiedzy, nikłym doświadczeniem bądź wiedzą niepewną na temat preferencji. Wgląd w preferencje zbiorcze sugeruje i pomaga w procesie analizy własnych preferencji bez narzucania rozwiązania w tej fazie rozwiązywania problemu decyzyjnego. Ponadto zbiorczy charakter wiedzy tego typu nie daje jednoznacznej sugestii na temat tego, jakie powinny być owe preferencje, co jest zgodne z subiektywną i indywidualną naturą decydenta.

Warto zaznaczyć, że z preferencjami zbiorczymi mamy do czynienia także w sytuacji podejmowania decyzji grupowych, gdzie decyzje indywidualne członków grupy mają doprowadzić do podjęcia wspólnej decyzji czy też do wyznaczenia konsensusu. W proponowanym podejściu rekomendowania preferencji indywidualnych na podstawie preferencji zbiorczych występuje pewne podobieństwo do metody podejmowania decyzji grupowych zwanej Delfi [Linstone, Turoff 1975]. Metoda Delfi opiera się na integracji wyników niezależnych decyzji indywidualnych w celu znalezienia wspólnego rozwiązania bez sugerowania członkom grupy jakiegokolwiek rozwiązania, lecz poprzez przedstawienie im wyniku ich decyzji w poprzedniej iteracji procesu decyzyjnego.

Mając do dyspozycji następujący zestaw kryteriów (kwestii):

$$D = \{K_1, K_2, ..., K_k\} \tag{3}$$

oraz próbę wektorów wag dla ustalonych kryteriów wybranych ze zbioru danych na podstawie dokonanej kwalifikacji negocjatora w fazie poprzedniej:

$$\overline{w}_j = (w_{1j}, w_{2j}, \ldots, w_{kj}) \in [0,1]^k \quad j \in \{1, \ldots, n\} \quad s_F(\overline{w}_j) = 1 \tag{4}$$

zostaje skonstruowany rozkład prawdopodobieństwa nad wektorem wag kryteriów dla wybranego problemu *p*:

$$p_{F_p} : (w_1, w_2, \ldots, w_k) \mapsto p_{12\ldots k} \tag{5}$$

gdzie $p_{12\ldots k}$ stanowi częstość względną wektora (w_1, w_2,…,w_k) w danej próbie. Rozkład p_{Fp} stanowi model preferencji zbiorczych wybranego profilu negocjatora.

Na podstawie rozkładu p_{Fp} można wyznaczyć rozkłady warunkowe dla każdej z kwestii K_i w postaci:

$$p_{F_p,\widetilde{K}_i}(w_i \mid w_1, w_2, \ldots, w_{i-1}, w_{i+1}, \ldots, w_k) = \frac{p_{F_p}(w_1, w_2, \ldots, w_k)}{\int_0^1 p_{F_p}(w_1, w_2, \ldots, w_k) dw_i} \tag{6}$$

Ustalanie przez negocjatora analizującego dane zakresu wektora wag różnych kwestii stanowi dla systemu dodatkową informację, która pozwala na skonstruowanie dla każdej kwestii rozkładu warunkowego zależnego od wartości pozostałych kwestii. Przyjmijmy, że negocjator ustalił następujące zakresy dla kwestii $\{K_1, K_2, ..., K_i, K_{i+1}, \ldots, K_k\}$:

$$w_o = w(K_o) \in [l_o, r_o] \subset [0,1] \quad o \in \{1, \ldots, m\} - \{i\} \tag{7}$$

W takiej sytuacji rozkład warunkowy wagi dla kwestii K_i zależny od zakresów pozostałych kwestii ma następującą postać:

$$p_{F,\widetilde{K}_i}(w_i) = \frac{\int_{l_1}^{r_1} \ldots \int_{l_{i-1}}^{r_{i-1}} \int_{l_{i+1}}^{r_{i+1}} \ldots \int_{l_m}^{r_m} p_F(w_1, \ldots w_{i-1}, w_{i+1}, \ldots, w_m) dw_m \ldots dw_{i+1} dw_{i-1} \ldots dw_1}{\int_{l_1}^{r_1} \ldots \int_{l_{i-1}}^{r_{i-1}} \int_{l_{i+1}}^{r_{i+1}} \ldots \int_{l_m}^{r_m} \int_0^1 p_F(w_1, \ldots w_{i-1}, w_{i+1}, \ldots, w_m) dw_i dw_m \ldots dw_{i+1} dw_{i-1} \ldots dw_1} \tag{8}$$

Zestaw rozkładów warunkowych dla danej kwestii zależnych od zakresów pozostałych kwestii może być przedstawiony negocjatorowi dokonującemu analizy własnych preferencji ze względu na istotność wag bądź przewidywania preferencji wag potencjalnego partnera. Ponadto zakresy wag dla poszczególnych kwestii mogą być iteracyjnie modyfikowane przez decydenta, co pozwala na dynamiczną zmianę rozkładów warunkowych wraz z ich ilustracją.

W sytuacji dostępności dużych zbiorów danych zachodzi możliwość zbadania normalności rozkładu wektora wag kwestii. Dodatkowa funkcjonalność systemu obejmuje zbadanie tego typu prawidłowości, co pozwala na opis modelu preferencji zbiorczych w postaci zestawu parametrów obejmującego: średni wektor wag, odchylenie standardowe oraz macierz kowariancji dla zestawu wag

kwestii. System będzie pozwalać na zbadanie, czy dla wybranego zbioru danych prawidłowość tego typu występuje poprzez zastosowanie testu Mardii [Mardia 1970]. Mając do dyspozycji próbę *n* wektorów wag dla *m* kryteriów w postaci: $\overline{w}_j = (w_{1j}, w_{2j}, \ldots, w_{mj})$, zostaje wyznaczona wartość średnia $\overline{\mu}$ oraz macierz kowariancji $\hat{\Sigma}$ w postaci:

$$\overline{\mu} = \frac{1}{n}\sum_{j=1}^{n} \overline{w}_j \tag{9}$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n} (\overline{w}_j - \overline{\mu})(\overline{w}_j - \overline{\mu})^T \tag{10}$$

Współczynnik wielowymiarowej skośności ma postać:

$$A = \frac{1}{6n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(\overline{w}_i - \overline{\mu})^T \hat{\Sigma}^{-1} (\overline{w}_j - \overline{\mu})\right]^3 \tag{11}$$

Wartość kurtozy wielowymiarowej jest określona wzorem:

$$B = \sqrt{\frac{n}{8m(m+2)}}\left\{\frac{1}{n}\sum_{i=1}^{n}\left[(\overline{w}_i - \overline{\mu})^T \hat{\Sigma}^{-1} (\overline{w}_i - \overline{\mu})\right]^2 - m(m+2)\right\} \tag{12}$$

Przy założeniu hipotezy zerowej, że rozkład wag ze względu na wyróżnione kryteria ma wielowymiarowy rozkład normalny, statystyka A ma w przybliżeniu rozkład chi-kwadrat z $\frac{1}{6}m(m+1)(m+2)$ stopniami swobody. Natomiast statystyka B ma w przybliżeniu rozkład normalny standaryzowany.

Zakładając, że w problemie decyzyjnym uwzględniono *k* kryteriów:

$$D = \{K_1, K_2, ..., K_k\} \tag{13}$$

stosujemy test normalności dla wszystkich dwuelementowych podzbiorów D_p zbioru D:

$$D_p \subset D (D_p \in 2^D) \quad \overline{D}_p = 2 \tag{14}$$

Dla danego podzbioru D_p narzędzie wspomagające wyznacza współczynnik skośności A_p oraz wartość kurtozy B_p oraz sprawdza, czy te wartości mieszczą się poza obszarami krytycznymi $T_{\chi^2}(\overline{D}_p)$ (dla A_p – przedział zależny od liczby kryteriów $\overline{D}_p$) oraz $T_{N(0,1)}$ (dla B_p). Funkcja MardiaTest realizuje test Mardii [Mardia 1970] oraz pozwala na stwierdzenie, czy jest podstawa do odrzucenia hipotezy zerowej (mówiącej o normalność rozkładu):

$$MardiaTest(D_p) = \begin{cases} 0 & A_p \in T_{\chi^2}(\overline{D}_p) \vee B_p \in T_{N(0,1)} \\ 1 & A_p \notin T_{\chi^2}(\overline{D}_p) \wedge B_p \notin T_{N(0,1)} \end{cases} \tag{15}$$

W przypadku braku podstaw do odrzucenia hipotezy zerowej stosujemy model preferencji zbiorczych składający się z zestawu rozkładów dla par kwestii:

$$p_{\{K_i,K_j\}}(w_i, w_j) = \frac{1}{2\pi\sqrt{|\Sigma_{i,j}|}} \exp\left(-\frac{1}{2}([w_i - \mu_i, w_j - \mu_j]^T \Sigma_{i,j}^{-1} [w_i - \mu_i, w_j - \mu_j])\right) \tag{16}$$

gdzie:

$$\Sigma_{i,j} = \begin{pmatrix} \sigma_i^2 & \rho_{i,j}\sigma_i\sigma_j \\ \rho_{i,j}\sigma_i\sigma_j & \sigma_j^2 \end{pmatrix} \tag{17}$$

przy czym $\rho_{i,j}$ jest poziomem korelacji ważenia kwestii K_i oraz K_j, σ_i – odchyleniem standardowym ważenia kwestii K_i, natomiast zmienne w_i oraz w_j odpowiadają wagom wybranym kwestiom.

Jeśli przyjmiemy, że wartość wagi jednej z kwestii została ustalona, rozkład warunkowy dla kwestii drugiej ma następującą postać:

$$p_{\widetilde{K}_i,a}(w_i) = \frac{1}{2\pi\sqrt{(1-\rho_{ij}^2)\sigma_i^2}} \exp\left(-\frac{\left(w_i - \left(\mu_i + \frac{\sigma_i}{\sigma_j}\rho_{ij}(a-\mu_j)\right)\right)^2}{2(1-\rho_{ij}^2)\sigma_i^2}\right) \tag{18}$$

gdzie wartość a jest ustalonym poziomem ważenia kwestii K_j. Ze względu na wyposażenie narzędzia w możliwość ustalania zakresów wag kwestii otrzymujemy zestaw rozkładów dla poszczególnych kwestii zależnych warunkowo od

przedziałów wagi dla kwestii sprzężonej po scałkowaniu funkcji gęstości rozkładu warunkowego:

$$p_{\tilde{K}_i,[c,d]}(w_i) = \frac{\int_c^d p_{K_i,K_j}(w_i, w_j)dw_j}{\int_c^d \int_0^1 p_{K_i,K_j}(w_i, w_j)dw_i dw_j} \tag{19}$$

otrzymując w konsekwencji funkcję gęstości rozkładu w postaci:

$$p_{\tilde{K}_i,[c,d]}(w_i) = \\ = \frac{\frac{1}{\sigma_i\sqrt{2\pi}}\exp(-\frac{(w_i-\mu_i)^2}{2\sigma_i^2}(\Phi_{0,1}(r(c,d,w_i))-\Phi_{0,1}(l(c,d,w_i))}{\Phi_{0,1}(d)-\Phi_{0,1}(c)} \tag{20}$$

gdzie:

$$l(c,d,w_i) = \min(g(c,w_i),(g(d,w_i)) \tag{21}$$

$$r(c,d,w_i) = \max(g(c,w_i),(g(d,w_i)) \tag{22}$$

$$g(v,w_i) = \frac{\frac{(v-\mu_j)}{\sigma_j} - \rho_{i,j}\frac{(w_i-\mu_i)}{\sigma_i}}{\sqrt{1-\rho_{i,j}^2}} \tag{23}$$

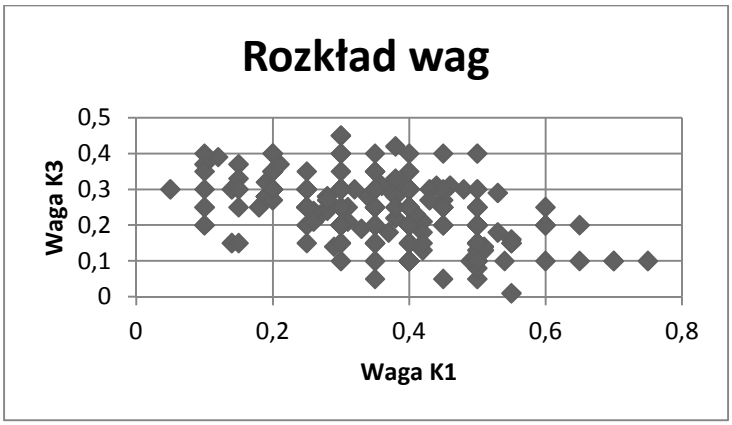
Przy czym $\Phi_{0,1}$ jest dystrybuantą rozkładu normalnego standaryzowanego.

2. Przykład empiryczny

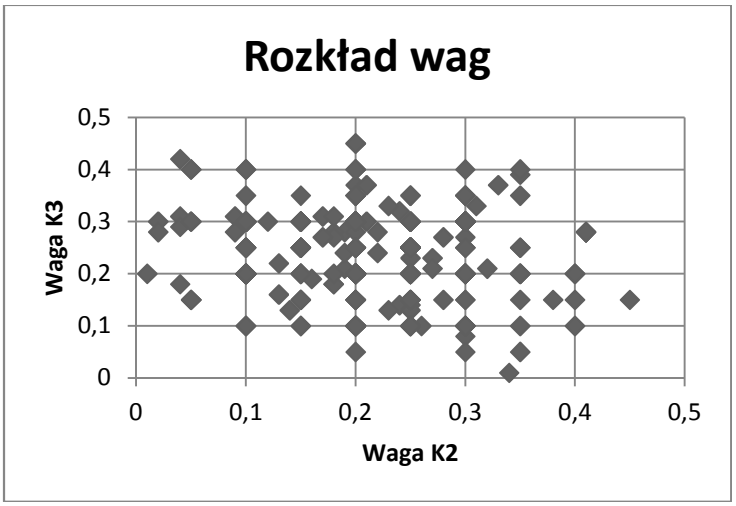
Korzystamy z danych systemu wspomagania negocjacji Inspire [Kersten 1999] dotyczących problemu symulacji negocjacji kontraktu pomiędzy reprezentantem firmy *CypresHill* – dystrybutorem części rowerowych oraz reprezentantem firmy *ItexManufacturing* zajmującej się montażem rowerów. Struktura pro-

blemu została negocjatorom narzucona – uwzględniono następujące zagadnienia negocjacyjne:

1. Cena (K1).
2. Czas zapłaty (K2).
3. Czas dostawy (K3).
4. Warunki zwrotu (dotyczące gwarancji) (K4).



Rys. 1. Ilustracja rozkładu wag kwestii: cena oraz czas dostawy



Rys. 2. Ilustracja rozkładu wag kwestii: czas dostawy oraz czas zapłaty

W przykładzie badamy możliwość wprowadzenia ciągłego modelu preferencji zbiorczych. Mając do dyspozycji zbiór danych opisujący preferencje 323 negocjatorów, przeprowadzamy testy normalności rozkładu wektora wag kwestii w różnych konfiguracjach kwestii za pomocą testu Mardii. Wyniki otrzymane dla różnych zestawów kwestii zestawiono w tabeli 1.

Tabela 1

Ilustracja wyników testów normalności rozkładów wektorów ważenia kwestii

Zestaw kryteriów	Stat. A	$\overline{\chi}^2$ – obszar krytyczny $(\alpha = 0{,}05)$	$\overline{\chi}^2$ st.swob.	Stat. B	N(0,1) – obszar krytyczny $(\alpha = 0{,}05)$	K
$\{K_1, K_2, K_3, K_4\}$	2755,94	[31,41;+∞)	20	4,9116	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	4
$\{K_1, K_2, K_3\}$	30,6978	[18,3;+∞)	10	0,0137	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	3
$\{K_1, K_2, K_4\}$	29,2555	[18,3;+∞)	10	0,0140	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	3
$\{K_2, K_3, K_4\}$	27,4475	[18,3;+∞)	10	0,0072	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	3
$\{K_1, K_3, K_4\}$	29,4075	[18,3;+∞)	10	0,0281	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	3
$\{K_1, K_2\}$	5,0001	[9,48;+∞)	4	0,1171	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2
$\{K_1, K_3\}$	2,6114	[9,48;+∞)	4	0,0916	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2
$\{K_1, K_4\}$	17,5699	[9,48;+∞)	4	-0,0098	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2
$\{K_2, K_3\}$	2,8251	[9,48;+∞)	4	-0,0840	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2
$\{K_2, K_4\}$	13,0834	[9,48;+∞)	4	-0,0506	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2
$\{K_3, K_4\}$	23,7767	[9,48;+∞)	4	-0,0503	$(-\infty,-1{,}96] \cup [1{,}96,+\infty)$	2

Źródło: Obliczenia własne.

Jak widać z tabeli 1, kwestie K_1 oraz K_2 brane parami wykazują asymptotyczny rozkład normalny. Na poziomie istotności 0,05 dla statystki A oraz na poziomie istotności 0,05 dla statystyki B nie ma podstaw do odrzucenia hipotezy zerowej, że rozkład wag kwestii K_1 oraz K_2 jest dwuwymiarowym rozkładem normalnym. Ponadto analogiczna regularność występuje dla kwestii K_2 i K_3 oraz dla kwestii K_1 i K_3. W przypadku pozostałych par kwestii nie odnotowano po-

dobnej regularności. Na podstawie powyższych obserwacji można skonstruować model preferencji zbiorczych w postaci trzech rozkładów dwuwymiarowych.

3. Zasady działania systemu rekomendacji doboru wag kryteriów opartego na ich charakterystyce probabilistycznej

Działanie systemu rekomendacji doboru wag kryteriów w kolejnych fazach można ująć następująco:

Faza 1. Ustalenie typu wsparcia:

a. Analiza preferencji własnych ze względu na wektor wag kwestii negocjacyjnych.
b. Przewidywanie preferencji potencjalnego partnera ze względu na wektor wag kwestii negocjacyjnych.

Faza 2. Tworzenie profilu negocjatora.

Typowanie grupy negocjatorów na podstawie kategorii cech lub zestawu cech (np. dane demograficzne, sposób odpowiedzi na konflikt).

Faza 3. Pobór danych dotyczących wytypowanej grupy (profilu) negocjatorów z bazy danych.

Faza 4. Ustalenie typu ogólnego modelu preferencji zbiorczych na podstawie wstępnej analizy danych:

a. Dyskretny rozkład prawdopodobieństwa nad wektorem ważenia kwestii.
b. Zbiór rozkładów ciągłych dla ustalonych zestawów kwestii.

Faza 5. Konstrukcja modelu preferencji zbiorczych w postaci rozkładu prawdopodobieństwa:

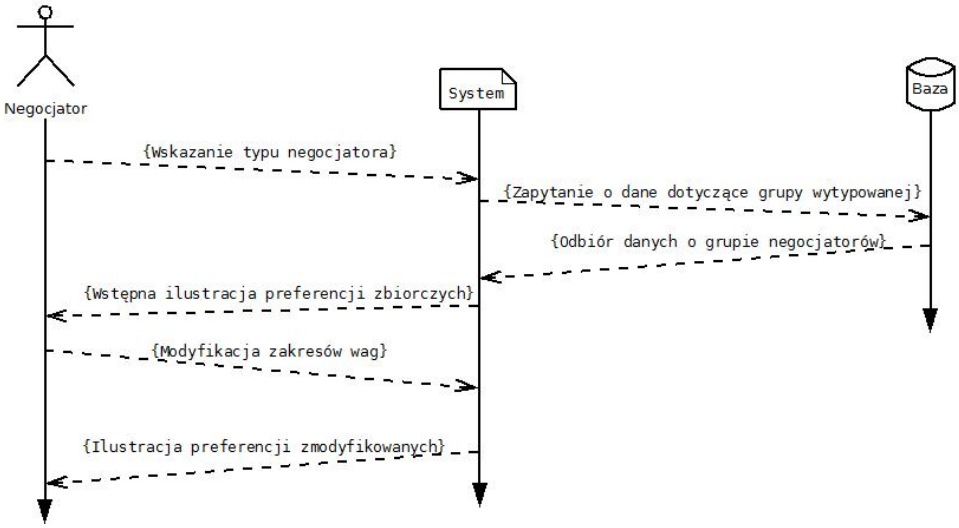
a. Konstrukcja modelu preferencji zbiorczych ogólnych.
b. Konstrukcja rozkładów warunkowych prawdopodobieństwa dla każdej z rozpatrywanych kwestii.

Faza 6. Wprowadzenie lub modyfikacja przez użytkownika zakresów ważenia kwestii.

Faza 7. Dla każdej z kwestii następuje konstrukcja rozkładu warunkowo zależnego od zakresów wag kwestii pozostałych.

Faza 8. Iterowane oddziaływanie użytkownika z systemem – powrót do fazy 6.

Sposób oddziaływania użytkownika z systemem oraz systemu z bazą przedstawiono na rys. 3.



Rys. 3. Ilustracja oddziaływania użytkownika z systemem oraz systemu z bazą

W ramach narzędzia wspomagającego zostaną zaimplementowane następujące moduły:

1. Moduł wydzielania danych dotyczących danego problemu negocjacyjnego oraz danego typu negocjatora podlegającego analizie zidentyfikowanego w pierwszym kroku działania programu.
2. Moduł konstrukcji rozkładu dyskretnego dla wydzielonego zbioru danych wraz z rozkładami warunkowymi dla poszczególnych kwestii.
3. Moduł konstrukcji rozkładu ciągłego w wypadku wystąpienia regularności w zbiorze danych wraz z rozkładami warunkowymi. Moduł zakłada testowanie normalności rozkładu dla różnych zestawów kwestii.
4. Moduł pozwalający na prezentację jednowymiarowych rozkładów ciągłych oraz dyskretnych poprzez warstwę komunikacyjną użytkownika.

Podsumowanie

W procesie negocjacyjnym preferencje własne stanowią podstawę do przygotowania oraz oceny ofert przedstawianych drugiej stronie oraz pozwalają na ocenę ofert otrzymywanych od partnera. Natomiast przewidywanie preferencji partnera pozwala na lepsze przygotowanie się zarówno w przypadku weryfikacji własnych preferencji, jak i opracowania odpowiedniej strategii negocjacyjnej. Kluczowym elementem formalnego opisu preferencji są wagi zagadnień negocjacyjnych używane zwykle do konstrukcji systemu oceny ofert negocjacyjnych.

Zapoznanie się ze sposobem ważenia zagadnień negocjacyjnych pozwala na zweryfikowanie własnych preferencji, zwłaszcza w sytuacji niepewnej wiedzy oraz nikłego doświadczenia w konstruowaniu formalnego modelu preferencji. Proponowane narzędzie wspomagania w zakresie doboru wag kryteriów pozwoli na pozyskiwanie szerszej wiedzy na temat preferencji różnych typów negocjatorów, co z kolei ma sprzyjać usprawnieniu procesu negocjacyjnego.

Zaprezentowana metoda szacowania wag jest zupełnie odmienna od metod opisywanych w literaturze przedmiotu, gdyż jako źródło wiedzy wykorzystuje preferencje zbiorcze populacji decydentów. Preferencje zbiorcze stanowią punkt wyjścia w procesie analizy, pozwalając decydentowi etapowo doprecyzowywać wektor wag kryteriów, ilustrując równocześnie zależność pomiędzy jego preferencjami indywidualnymi a preferencjami zbiorczymi w postaci rozkładów warunkowych nad wagami poszczególnych kryteriów.

Bibliografia

Belton V., Stewart T.J., 2002: Multiple Criteria Decision Analysis: An Integrated Approach. Kluwer, Dordrecht.

Bottomley P.A., Doyle J.R., 2001: A Comparison of Three Weight Elicitation Methods: Good, Better, and Best. „Omega" 29, 553-560.

Brzostowski J., Roszkowska E., Wachowicz T., 2012: Supporting Negotiation by Multi-Criteria Decision Making Methods. „Optimum – Studia Ekonomiczne", nr 5(59), 3-29.

Diakoulaki D., Mavrotas G., Papayannakis L., 1995: Determining Objective Weights in Multiple Criteria Problems: The Critic Method. „Computers & Operations Research", No. 22, 763-770.

Edwards W., 1977: How to Use Multi-attribute Utility Analysis for Social Decision Making. „IEEE Trans. Syst. Man Cybernet" SMC-7, 326-340.

Figueira J., Greco S., Ehrgott M., 2005: Multiplie Criteria Decision Analysis: State of the Art Surveys. Springer, New York.

Hwang C.L., Yoon K., 1981: Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag, New York.

Keeney R.L., Raiffa H., 1976: Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Wiley, New York.

Kersten G.E., 1985: An Interactive Procedure for Solving Group Decision Problems. W: Decision Making with Multiple Objectives. V. Chankong, Y.Y. Haimes (eds.). Springer Verlag, No. 242, 331-344.

Kersten G.E, Noronha S., 1999: WWW-based Negotiation Support: Design, Implementation, and Use. „Decision Support Systems", No. 25(2), 135-154.

Kersten G.E., Lai H., 2007: Negotiation Support and E-Negotiation Systems. „Group

Decision & Negotiation", No. 16(6), 553-586.

Kilgour D.M., Chen Y., Hipel K.W., 2010: Multiple Criteria Approaches to Group Decision and Negotiation. W: Trends in Multiple Criteria Decision Analysis. M. Ehrgott (ed.). International Series in Operations Research and Management Science, Springer, No. 142, 317-338.

Kilmann R., Thomas K.W., 1983: The Thomas-Kilmann Conflict Mode Instrument. The Organizational Development Institute, Cleveland, OH, 57-64.

Linstone H.A., Turoff M., 1975: The Delphi Method: Techniques and Applications. Reading, Mass, Addison-Wesley.

Ma J., Fan Z.P., Huang L.H., 1999: A Subjective and Objective Integrated Approach to Determine Attribute Weights. „European Journal of Operational Research" No. 112, 397-404.

Mardia K.V., 1970: Measures of Multivariate Skewness and Kurtosis with Applications. „Biometrika", No. 57, 519-530.

Roszkowska E., 2011: Wybrane modele negocjacji. Wydawnictwo UwB, Białystok.

Saaty T.L., 1980: The Analytical Hierarchy Process. McGraw-Hill, New York.

Salo A., Hämäläinen R.P., 2012: Multicriteria Decision Analysis in Group Decision Processes. W: Handbook of Group Decision and Negotiation. D.M. Kilgour, C. Eden (eds.). Springer, Dordrecht, 269-284.

Stillwell W.G., Seaver D.A., Edwards W., 1981: A Comparison of Weight Approximation Techniques in Multiattribute Utility Decision Making. „Organizational Behavior and Human Performance", No. 28, 62-77.

Takeda E., Cogger K.O., Yu P.L., 1987: Estimating Criterion Weights Using Eigenvectors: A Comparative Study. „European Journal of Operational Research" No. 29, 360-369.

Tzeng G.-H., Chen T.-Y., Wang J.C., 1998: A Weight Assessing Method with Habitual Domains. „European Journal of Operational Research" 110(2), 342-367.

Wu Z., Chen Y., 2007: The Maximizing Deviation Method for Group Multiple Attribute Decision Making under Linguistic Environment. „Fuzzy Sets and Systems", No. 158(14), 1608-1617.

SYSTEM SUPPORTING THE DECISION CRITERIA WEIGHTS SPECIFICATION BASED ON THEIR PROBABILISTIC CHARACTERISTICS

Summary

A tool for supporting the negotiator during the process of the analysis of own preferences and the analysis of the preferences of the potential partner is proposed in this work. The approach is based on the construction of the collective preferences model for a selected negotiator's profile in the form of multivariate probability distribution over the space of negotiation issue weights vectors. In the process of user interaction with the system the ranges of issue weights are modified that allows for the decomposition of the

general multi-variate distribution into series of uni-variate distributions corresponding to single issues. Such distributions conditionally depend on the issue weight ranges set by the decision-maker for all the remaining issues. Moreover, in the work we consider the possibility of constructing the collective preferences model in a continuous form in the case normally distributed weights for some sets of issues. The data from the Negotiation Support System Inspire [Kersten 1999] were used to examine the normality of the issue weights distribution for different issue sets.

Helena Gaspars-Wieloch
Uniwersytet Ekonomiczny w Poznaniu

PROPOZYCJA HYBRYDY REGUŁ HURWICZA I BAYESA W PODEJMOWANIU DECYZJI W WARUNKACH NIEPEWNOŚCI

Wstęp

Z podejmowaniem decyzji w warunkach niepewności (PDWN) mamy do czynienia wtedy, gdy sytuacja decyzyjna może być scharakteryzowana za pomocą listy decyzji dopuszczalnych (wariantów decyzyjnych, strategii) oraz stanów otaczającej nas rzeczywistości. Stany te w istotny sposób oddziałują na otrzymany wynik, ale w momencie podjęcia decyzji nie wiemy, który z nich wystąpi, i nie mamy na to żadnego wpływu [Pazek, Rozman 2009, s. 45-50; Trzaskalik 2008]. W odróżnieniu od podejmowania decyzji w warunkach ryzyka, PDWN cechuje się tym, iż decydent nie ma możliwości określenia prawdopodobieństwa wystąpienia danego stanu [Groenewald, Pretorius 2011; Render, Stair, Hanna 2006; Siddiqui, Chronopoulos, w druku; Sikora, red., 2008]. Knight [1921] jako pierwszy zaproponował wykorzystanie tak rozumianego ryzyka i niepewności w ekonomii, ale te dwie kategorie zostały formalnie wprowadzone do teorii ekonomii dopiero w 1944 r. przez Neumanna i Morgensterna [Neumann, Morgenstern 1944].

Tabela 1

Macierz wypłat (przypadek ogólny)

Scenariusze	Decyzje		
	D_1	D_j	D_n
S_1	a_{11}	a_{1j}	a_{1n}
S_i	a_{i1}	a_{ij}	a_{in}
S_m	a_{m1}	a_{mj}	a_{mn}

Nasze rozważania będą dotyczyć gry z naturą, czyli sytuacji, w której decydent ma do czynienia ze zjawiskiem (np. pogodą) mogącym przyjmować różne scenariusze (stany). Załóżymy, iż jesteśmy w stanie przewidzieć, jaka korzyść (strata) wynika z kolejnych decyzji przy zaistnieniu poszczególnych stanów natury. Zarówno liczba możliwych decyzji, jak i liczba stanów natury jest skończona (ang. decision making under „scenarios" uncertainty), a każdej kombinacji decyzja-scenariusz odpowiada dokładnie jedna wypłata, zatem owe korzyści można przedstawić w postaci macierzy wypłat (tabela 1, gdzie m – liczba scenariuszy S_1, …, S_i, …, S_m (m>1), n – liczba decyzji D_1, …, D_j , …., D_n (n>1), a_{ij} – wypłata związana z i-tym scenariuszem i j-tą decyzją). Zbiór potencjalnych wypłat dotyczący danej strategii jest też skończony i może być multizbiorem[*]. Celem decydenta jest wybór strategii maksymalizującej korzyść.

W artykule skoncentrujemy się na wyborze optymalnej strategii czystej, tj. rozwiązania zakładającego, iż decydent wybiera i realizuje tylko jeden wariant decyzyjny [Sikora, red., 2008]. Każda decyzja będzie opisana za pomocą jednego, wspólnego kryterium lub jednego miernika syntetycznego ukazującego realizację różnych celów, zatem rozważania będą dotyczyć problemów jednokryterialnych lub sprowadzalnych do tychże problemów[**].

W literaturze można znaleźć opis różnych metod mających zastosowanie przy podejmowaniu decyzji w warunkach niepewności. Jedną z nich jest reguła Hurwicza. Ta zasada prowadzi zazwyczaj do logicznych i racjonalnych odpowiedzi, ale w pewnych szczególnych przypadkach rezultaty otrzymane za pomocą tej metody mogą być zdumiewające.

Artykuł ma następującą strukturę. Pierwsza część zawiera krótki opis najbardziej znanych reguł stosowanych w przypadku PDWN. W części drugiej skupiono się na analizie istoty i mankamentów reguły Hurwicza. W części trzeciej przedstawiono nowe podejście posiadające znamiona reguł Hurwicza i Bayesa. Wnioski zebrano w zakończeniu.

1. Reguły podejmowania decyzji w warunkach niepewności

Reguły PDWN są powszechnie znane w środowisku ludzi zajmujących się teorią podejmowania decyzji i metodami ilościowymi w ekonomii, lecz dla ułatwienia

---

[*] Jeżeli natomiast scenariuszy jest nieskończenie wiele, to zamiast macierzy wypłat podaje się dla każdej strategii przedział możliwych korzyści [w_j, m_j] (ang. decision making under interval uncertainty) [Huynh, Hu, Nakamori, Kreinovich 2007].

[**] Wśród interesujących prac dotyczących wielokryterialnego podejmowania decyzji w warunkach niepewności (WPDWN) warto wymienić: Dominiak [2006, 2009].

dalszego wywodu zostaną tu one krótko scharakteryzowane, przy czym z racji podjętego tematu, reguły Hurwicza i Bayesa omówimy nieco dokładniej. Szersze omówienie niżej przedstawionych zasad można znaleźć w licznych pozycjach literaturowych [Ignasiak, red., 1996; Kaufmann, Faure 1974; Pazek, Rozman 2009; Sikora, red., 2008]. Jak będzie można zaobserwować, wybór reguły decyzyjnej powinien zależeć od preferencji decydenta i przyjętych przez niego założeń.

Reguła Walda (reguła maximin) zakłada, iż niezależnie od wybranego wariantu, decydenta spotka zawsze najniższa z wypłat odpowiadających temu wariantowi [Wald 1950a, s. 656-668, 1950b]. Jest więc ona podejściem skrajnie pesymistycznym, sprowadzającym się do wyznaczenia wskaźnika w_j (poziomu bezpieczeństwa) dla każdej alternatywy i do wyboru decyzji maksymalizującej ten wskaźnik. W przeciwieństwie do zasady Walda, reguła maximax jest bardzo optymistycznym podejściem, w którym decydent zawsze spodziewa się wystąpienia najkorzystniejszego stanu natury i wybiera decyzję maksymalizującą wskaźnik maximax m_j (poziom optymizmu).

W regule Hurwicza przyjmuje się, iż decydent powinien przeprowadzić ranking strategii na podstawie średniej ważonej poziomu bezpieczeństwa i poziomu optymizmu [Hurwicz 1952, 1951] zgodnie ze wzorem (1), gdzie h_j to wskaźnik Hurwicza, w_j i m_j to minimalna i maksymalna wypłata związana z j-tą decyzją, a $\alpha \in <0,1>$ oznacza współczynnik pesymizmu (ostrożności). Dla skrajnych optymistów parametr ten jest bliski zeru, z kolei dla skrajnych pesymistów współczynnik α dąży do jedności. Optymalną strategią czystą jest ta, której wartość wskaźnika h_j jest najwyższa* (wzór (2)).

$$h_j = \alpha \cdot w_j + (1-\alpha) \cdot m_j \tag{1}$$

$$h_j^* = \max_j \{h_j\} = \max_j \{\alpha \cdot w_j + (1-\alpha) \cdot m_j\} \tag{2}$$

Reguła Savage'a [1961, s. 173-188] jest także nazywana regułą minimalnego żalu lub regułą minimax i polega na wyborze decyzji minimalizującej maksymalną względną stratę na podstawie macierzy względnych strat.

Reguła Bayesa, zwana również regułą Laplace'a lub regułą niedostatecznej racji [Render, Stair, Hanna 2006], zakłada, że skoro decydent nie jest w stanie określić, który scenariusz ostatecznie wystąpi, to może on przyjąć, iż wszystkie stany natury są równoprawdopodobne. Wówczas wystarczy obliczyć dla każdej

* Warto podkreślić, że w literaturze parametr α opisuje czasami poziom optymizmu decydenta, a nie poziom jego pesymizmu. Wówczas wzór (1) ma oczywiście nieco inną postać [Huynh, Hu, Nakamori, Kreinovich 2007; Groenewald, Pretorius 2011].

opcji wskaźnik Bayesa jako oczekiwaną wypłatę (3) i wybrać tę decyzję, której wskaźnik przyjmuje największą wartość (4):

$$b_j = \frac{1}{m}\sum_i a_{ij} \tag{3}$$

$$b_j^* = \max_j \{b_j\} \tag{4}$$

Stosowanie poszczególnych reguł PDWN może doprowadzić do uzyskania zupełnie odmiennych rankingów i do wskazania różnych optymalnych strategii. Z wyjątkiem zasady Bayesa, opisane reguły są przeznaczone wyłącznie dla decyzji realizowanych jednokrotnie*. Dodajmy także, że reguła Bayesa mocno odbiega od pozostałych reguł klasycznych, gdyż jako jedyna odwołuje się do rachunku prawdopodobieństwa.

2. Analiza reguły Hurwicza

W tej części artykułu przeprowadzimy dokładną analizę istoty reguły Hurwicza i konsekwencji wynikających z jej stosowania. Jak już zasygnalizowano na początku, korzystanie z tej zasady w podejmowaniu decyzji w warunkach niepewności w większości analizowanych sytuacji decyzyjnych jest rozsądne. Jednak warto zaznaczyć, że ta reguła prowadzi w bardzo specyficznych przypadkach do rezultatów sprzecznych z logiką i deklarowanymi przez decydenta preferencjami.

Przyjrzyjmy się bliżej konstrukcji wzoru (1), który służy do wyznaczenia wskaźnika Hurwicza dla poszczególnych wariantów.

Wzór uwzględnia jedynie skrajne wypłaty. Pośrednie wartości, tj. $a_{ij} \in (w_j, m_j)$, nie są w ogóle brane pod uwagę. Nie ma więc znaczenia, czy wśród pośrednich wyników dotyczących danej strategii większość jest bliska parametrowi w_j bądź m_j, czy też rozkładają się one w miarę symetrycznie. Pozycja konkretnej decyzji w rankingu jest zdeterminowana wyłącznie przez w_j i m_j, przy czym $w_j \neq m_j$.

Powyższa cecha reguły Hurwicza implikuje dla pary wariantów decyzyjnych D_e i D_f następujące konsekwencje:

$$\left(w_e = w_f\right) \wedge (m_e > m_f) \wedge (\alpha \in < 0,1)) \Rightarrow (h_e > h_f) \tag{5}$$

* Zaprezentowane zasady omówiono dla problemów maksymalizowanych. Sposoby zastosowania wymienionych reguł w przypadku kryterium minimalizowanego przedstawiono m.in. w pracy Gaspars-Wieloch [2012, s. 303-324].

$$(w_e = w_f) \wedge (m_e > m_f) \wedge (\alpha = 1) \Rightarrow (h_e = h_f) \quad (6)$$
$$(w_e > w_f) \wedge (m_e = m_f) \wedge (\alpha \in (0,1>) \Rightarrow (h_e > h_f) \quad (7)$$
$$(w_e > w_f) \wedge (m_e = m_f) \wedge (\alpha = 0) \Rightarrow (h_e = h_f) \quad (8)$$
$$(w_e = w_f) \wedge (m_e = m_f) \wedge (\alpha \in <0,1>) \Rightarrow (h_e = h_f) \quad (9)$$

Odnotujmy dodatkowo, że:

$$\begin{aligned}
&(h_e > h_f) \Leftrightarrow (\alpha \cdot w_e + (1-\alpha) \cdot m_e > \alpha \cdot w_f + (1-\alpha) \cdot m_f) \\
&(h_e > h_f) \Leftrightarrow (\alpha \cdot w_e + m_e - \alpha \cdot m_e > \alpha \cdot w_f + m_f - \alpha \cdot m_f) \\
&(h_e > h_f) \Leftrightarrow (\alpha \cdot (w_e - m_e) + m_e > \alpha \cdot (w_f - m_f) + m_f) \\
&(h_e > h_f) \Leftrightarrow (\alpha \cdot (-r_e) + m_e > \alpha \cdot (-r_f) + m_f) \\
&(h_e > h_f) \Leftrightarrow (\alpha \cdot (r_f - r_e) > m_f - m_e) \qquad (10)
\end{aligned}$$

gdzie r_e i r_f oznaczają rozstępy między maksymalną a minimalną wypłatą związanymi odpowiednio z decyzjami D_e i D_f.

Zauważmy, że powyższe implikacje występują zawsze, nawet wówczas, gdy decydent jest pesymistą i zachodzi sytuacja opisana za pomocą wzoru (11) – wszystkie wypłaty pośrednie strategii D_e są niższe od wypłat pośrednich strategii D_f:

$$\left(\bigvee_{\substack{a_{i,e} \neq m_e \\ a_{i,e} \neq w_e}} \bigvee_{\substack{a_{i,f} \neq m_f \\ a_{i,f} \neq w_f}} a_{i,e} < a_{i,f} \right) \wedge \left(\exists_{a_{i,e}} a_{i,e} = m_e \right) \wedge \left(\exists_{a_{i,e}} a_{i,e} = w_e \right) \wedge$$
$$\wedge \left(\exists_{a_{i,f}} a_{i,f} = m_f \right) \wedge \left(\exists_{a_{i,f}} a_{i,f} = w_f \right) \quad (11)$$

Szczególnym przypadkiem wyżej wspomnianej zależności jest sytuacja, w której nierosnące m-elementowe ciągi wypłat rozpatrywanych wariantów spełniają warunki (12)-(13):

$$(S_e = (a_{1,e}, a_{2,e}, ..., a_{m,e})) \wedge (a_{1,e} = m_e) \wedge (a_{2,e} = ... = a_{m-1,e} = a_{m,e} = w_e) \quad (12)$$
$$(S_f = (a_{1,f}, a_{2,f}, ..., a_{m,f})) \wedge (a_{m,f} = w_f) \wedge (a_{1,f} = a_{2,f} = ... = a_{m-1,f} = m_f) \quad (13)$$

Wymienione przypadki, choć niezwykle rzadko występujące w praktyce, uzmysławiają nam zatem bardzo jasno, iż reguła Hurwicza, niezależnie od wartości współczynnika pesymizmu, pomija istotną informację, jaką jest częstotliwość występowania, w multizbiorze wszystkich wypłat związanych z danym warian-

tem decyzyjnym, wypłat bliskich parametrom w_j i m_j. Wracając do analizowanej pary decyzji D_e i D_f, można stwierdzić, że zaobserwowane następstwa wynikające z zastosowania zasady Hurwicza są dość krzywdzące dla decyzji D_f w przypadku relatywnie wyższych wypłat pośrednich, ponieważ nawet wówczas nie będzie mogła zająć wyższej pozycji w rankingu niż strategia D_e[*].

Analizując specyfikę reguły Hurwicza, warto również sprawdzić, jakie znaczenie mają wspomniane konsekwencje dla poszczególnych decydentów.

Decydent optymista (o współczynniku pesymizmu bliskim lub równym zero), a więc osoba zakładająca, iż czeka ją najwyższa bądź prawie najwyższa wypłata związana z podjętą decyzją, nie przejmuje się zbytnio wielkością pozostałych, niższych wypłat. Taki decydent nie poszukuje za wszelką cenę bezpiecznych wariantów decyzyjnych (przyjęto tu, iż decyzja jest tym bardziej bezpieczna, im jej wypłaty pośrednie są bliższe wypłacie maksymalnej). Rekomendowane przez regułę Hurwicza strategie dla decydenta optymisty odzwierciedlają więc raczej jego preferencje. Poważny problem pojawia się dopiero wtedy, gdy współczynnik pesymizmu (zwany również współczynnikiem ostrożności) decydenta rośnie, a więc gdy w procesie decyzyjnym udział bierze pesymista. Postawa pesymisty wyraża się w skłonności do dostrzegania tylko ujemnych stron życia, negatywnej oceny rzeczywistości oraz przyszłości. Stosunek pesymisty do świata jest nacechowany lękiem i poczuciem bezsilności [http://wikipedia.pl]. Pesymista spodziewa się, iż spotka go raczej scenariusz związany z niską wypłatą. Skoro tak jest, to pesymista stara się unikać decyzji, dla których większość scenariuszy oferuje niskie wypłaty. Takie warianty decyzyjne będziemy nazywać w skrócie strategiami niebezpiecznymi[**]. Tymczasem okazuje się, że im wyższy współczynnik ostrożności, tym gorzej jest uwzględniane w rankingu ustalonym na podstawie zasady Hurwicza ostrożnościowe podejście decydenta. Powróćmy do rozpatrywanej wcześniej pary decyzji D_e i D_f. Ze wzorów (5) i (7) wynika, że ta reguła nigdy nie zarekomenduje osobom o wysokim współczynniku pesymizmu decyzji D_f, nawet wtedy, gdyby wystąpił przypadek (11) lub (12)-(13), a różnice $m_e - m_f$ (wzór 5) i $w_e - w_f$ (wzór 7) były bliskie zeru, choć wiadomo, że pesymista postępuje ostrożnie i woli wybierać decyzje bezpieczne.

Kolejna kwestia warta poruszenia dotyczy strategii niezdominowanych. Jeżeli potencjalne decyzje są Pareto-optymalne, to oczekiwalibyśmy, że ranking tych decyzji będzie, wraz ze zmianą poziomu współczynnika ostrożności, również ulegał

---

[*] Przykładowo jeżeli $S_e = (5,1,1,1,1)$ i $S_f = (5,5,5,5,1)$, to $h_e = h_f = \alpha \cdot 1 + (1 - \alpha) \cdot 5 = 5 - 4\alpha$.

[**] Jeżeli przyjmiemy (zgodnie z definicją podaną na stronie http://wikipedia.pl), że ryzyko to niebezpieczeństwo wynikające z możliwych konsekwencji podjęcia danej decyzji, to dla strategii niebezpieczniej można używać zamiennie pojęcia „strategia ryzykowna".

zmianom. Tymczasem wzór (10) pokazuje, że kolejność wariantów decyzyjnych, poza współczynnikiem pesymizmu, zależy tylko od wartości maksymalnych i od rozstępów wypłat poszczególnych decyzji, które są stałe, co oznacza, iż dla niektórych sytuacji decyzyjnych ranking będzie się zmieniał bardzo rzadko.

Skoro w regule Hurwicza ważne są tylko skrajne wypłaty, nasuwa się wniosek, iż powinno się ją raczej stosować w problemach, w których dla każdej *j*-tej decyzji wszystkie (a nie tylko niektóre) wypłaty z przedziału $<w_j,m_j>$ są możliwe, a więc gdy liczba scenariuszy jest nieskończona. Natomiast w przypadku zadań ze skończoną liczbą wypłat reguła Hurwicza owszem dobrze uwzględnia preferencje decydenta, lecz jedynie wtedy, gdy dla każdego wariantu decyzyjnego rozkład wypłat jest w miarę symetryczny, tj. gdy nie występują w nadmiarze wartości zbliżone do jednej ze skrajnych wypłat [Gaspars-Wieloch 2013].

Zanim przejdziemy do przedstawienia propozycji nowego podejścia, które będzie uwzględniać nie tylko preferencje decydenta i skrajne wypłaty dotyczące poszczególnych strategii, lecz także wartości wszystkich wypłat pośrednich, co umożliwi uzyskanie sensownych rekomendacji dla szerszego spektrum problemów decyzyjnych, zilustrujemy sformułowane wcześniej wnioski fikcyjnym przykładem.

Załóżmy, że inwestorzy A i B zamierzają wybrać najlepszy projekt, mając do dyspozycji cztery różne biznesplany (P1, P2, P3, P4), przy czym eksperci przewidują, iż w przyszłości może wystąpić jeden z sześciu scenariuszy (S1, S2, S3, S4, S5, S6). Inwestor A jest umiarkowanym optymistą i określił swój współczynnik pesymizmu α_A na poziomie 0.3. Inwestor B lubi postępować ostrożnie – jego współczynnik pesymizmu α_B wynosi 0.7. W tabeli 2 podano przewidywane roczne zyski (w tys. euro) dla poszczególnych projektów w zależności od scenariusza.

Tabela 2

Macierz wypłat (studium przypadku)

Scenariusz	Projekt			
	P1	P2	P3	P4
S1	5	1	1.5	1
S2	1	4.8	3	1
S3	1	4.8	1.5	4.8
S4	1	4.8	3	1
S5	1	4.8	2	4.8
S6	1	4.8	2	1

Analizę porównawczą rozpatrywanych projektów ułatwi nam tabela 3, w której zebrano następujące informacje:

w_j – wskaźnik Walda, poziom bezpieczeństwa (wypłata minimalna),

m_j – wskaźnik maximax, poziom optymizmu (wypłata maksymalna),
r_j – rozstęp ($m_j - w_j$),
N_j – liczba scenariuszy, w których j-ty projekt uzyskuje najlepszy wynik,
n_j – liczba scenariuszy, w których j-ty projekt uzyskuje najgorszy wynik.

Tabela 3

Porównanie projektów (studium przypadku)

Charakterystyki	Projekt			
	P1	P2	P3	P4
w_j	1	1	**1.5**	1
m_j	**5**	4.8	3	4.8
r_j	4	3.8	**1.5**	3.8
N_j	1	**5**	0	2
n_j	5	1	**0**	4

Tabela 4

Wskaźniki Hurwicza (studium przypadku)

Projekty	Inwestor	
	A (umiarkowany optymista) ($\alpha_A = 0.3$)	B (umiarkowany pesymista) ($\alpha_B = 0.7$)
P1	$h_{P1}^{A} = 0.3 \cdot 1 + 0.7 \cdot 5 = 3.80$	$h_{P1}^{B} = 0.7 \cdot 1 + 0.3 \cdot 5 = 2.20$
P2	$h_{P2}^{A} = 0.3 \cdot 1 + 0.7 \cdot 4.8 = 3.66$	$h_{P2}^{B} = 0.7 \cdot 1 + 0.3 \cdot 4.8 = 2.14$
P3	$h_{P3}^{A} = 0.3 \cdot 1.5 + 0.7 \cdot 3 = 2.55$	$h_{P3}^{B} = 0.7 \cdot 1.5 + 0.3 \cdot 3 = 1.95$
P4	$h_{P4}^{A} = 0.3 \cdot 1 + 0.7 \cdot 4.8 = 3.66$	$h_{P4}^{B} = 0.7 \cdot 1 + 0.3 \cdot 4.8 = 2.14$
Ranking	I. P1 II. P2 i P4 III. P3	I. P1 II. P2 i P4 III. P3

Jak widać, największym rozstępem charakteryzuje się projekt P1 ($r_1 = 4$). To także ten właśnie projekt ma szansę wygenerować najwyższe zyski ($m_1 = 5$). Jego słabą cechą jest to, iż tylko w przypadku jednego scenariusza (S1) plasuje się on na pierwszym miejscu. W pozostałych pięciu stanach odnotowuje najgorsze wyniki. Projekty P2 i P4 mają trzy wspólne cechy (rozstęp, wartość minimalną oraz wartość maksymalną), lecz różnią się istotnie pod względem liczby najwyższych i najniższych zysków. W przypadku tych dwóch kryteriów projekt P2 jest zdecydowanie lepszy. Projekt P3, niezależnie od stanu, przynosi dość zbliżone i relatywnie małe korzyści. Jednakże w przypadku tego właśnie projektu aż trzy charakterystyki ujęte w tabeli 3 wypadają najlepiej.

Teraz przyjrzyjmy się wskaźnikom Hurwicza obliczonym dla poszczególnych projektów w zależności od preferencji inwestora oraz ostatecznym rankingom (tabela 4). Analizując otrzymane rezultaty, dochodzimy do następujących, momentami absurdalnych, wniosków, które jednak potwierdzają wymienione wcześniej konsekwencje wynikające ze stosowania reguły Hurwicza:

1) Ranking projektów jest identyczny dla umiarkowanego optymisty i pesymisty, co może być trochę zaskakujące, biorąc pod uwagę specyfikę poszczególnych biznesplanów.
2) Niezależnie od rozpatrywanego inwestora wartości wskaźników Hurwicza dla projektów P2 i P4 są parami takie same, choć doskonale wiadomo, iż projekt P2 dominuje nad P4 (patrz dwie ostatnie charakterystyki ujęte w tabeli 3). Para projektów P2 i P4 ilustruje przypadek opisany za pomocą wzoru (9).
3) Projekt P1 zdobył pierwsze miejsce w obu rankingach. Oznacza to, że zgodnie z regułą Hurwicza jest to rekomendowana optymalna strategia czysta dla obu inwestorów, nawet dla umiarkowanego pesymisty. Jest to o tyle zdumiewające, że projekt P1 jest przecież dość niebezpieczny. Owszem, można w przypadku stanu S1 liczyć na najwyższą wypłatę (aż 5 tys. euro), ale jeżeli wystąpi jakikolwiek inny scenariusz, zyski z realizacji projektu P1 będą najniższe. Wybór projektu P1 jest więc dla pesymisty raczej nieracjonalnym posunięciem. Posunięciem sprzecznym z deklarowanym przez niego współczynnikiem ostrożności. Projekt P1 ustępuje w rankingu projektowi P3 dopiero przy współczynniku pesymizmu przekraczającym 0.8.
4) Reguła Hurwicza wyżej plasuje projekt P1 niż P2, co w przypadku inwestora pesymisty także budzi kontrowersje. To przecież projekt P2 wydaje się bardziej pożądany dla inwestora B, gdyż w aż pięciu przypadkach na sześć daje możliwość realizacji całkiem wysokiego zysku. Dodajmy, iż nawet gdyby współczynnik ostrożności pesymisty wzrósł do 0.99, to i tak projekt P1 nadal uzyskiwałby wyższą wartość wskaźnika niż projekt P2. Relacja pomiędzy projektami P1 i P2 odpowiada właśnie sytuacji opisanej za pomocą wzorów (5), (11)-(13) i nawiązuje do sformułowanego wcześniej ogólnego wniosku dotyczącego rankingów gorzej uwzględniających preferencje decydentów o wysokich współczynnikach pesymizmu.

3. Reguła H+B, czyli hybryda reguły Hurwicza i reguły Bayesa – prezentacja podejścia

W dotychczasowych rozważaniach stwierdziliśmy, że rekomendacje uzyskiwane za pomocą reguły Hurwicza mogą nie do końca odzwierciedlać upodobania

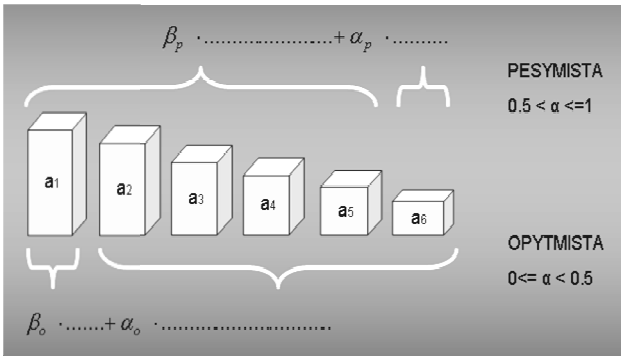
decydenta. Wniosek ten jednak dotyczył jedynie sytuacji, w której przynajmniej jeden z wariantów decyzyjnych charakteryzuje się zbiorem wypłat z przeważającą liczbą zysków albo bliskich m_j, albo bliskich w_j. W pozostałych przypadkach (tj. gdy wypłaty są rozłożone w miarę symetrycznie) zasada Hurwicza proponuje sensowne rozwiązania.

W tej części artykułu zostanie zaprezentowana pewna istotna modyfikacja analizowanej reguły, dzięki której będzie można znacząco zredukować skutki zaobserwowanego mankamentu i otrzymać logiczne odpowiedzi dla szerszego spektrum problemów decyzyjnych. Proponowana zasada, podobnie jak reguła Hurwicza, pozwala wskazać optymalną strategię czystą przy założeniu, że liczba stanów jest skończona oraz że decydent jest w stanie określić swój współczynnik pesymizmu. Metoda ta bierze pod uwagę nie tylko postawę decydenta, lecz także specyfikę zbiorów wypłat wszystkich alternatyw.

Idea proponowanego podejścia, tj. reguły H+B, jest bardzo prosta. Podobnie jak w przypadku oryginalnej wersji reguły Hurwicza, tu również trzeba będzie wyznaczyć dla każdej decyzji pewną średnią ważoną wypłat, a następnie wskazać tę decyzję, której wskaźnik jest najwyższy. Jednak zupełnie inaczej będzie obliczany ten właśnie wskaźnik, nazwijmy go wskaźnikiem H+B, czyli hb_j. Na początek trzeba będzie przedstawić zbiór wypłat w postaci nierosnącego ciągu wypłat. Gdy deklarowany współczynnik pesymizmu będzie się mieścić w przedziale <0, 0.5>, to parametr α będziemy mnożyć nie przez najniższą wypłatę, lecz przez sumę $(m - 1)$ najniższych wyrazów tego ciągu (gdzie m oznacza liczbę scenariuszy), natomiast współczynnik optymizmu (czyli β) – jedynie przez najwyższą wypłatę. Z kolei gdy deklarowany współczynnik pesymizmu będzie należeć do przedziału <0.5, 1>, to parametr α będziemy mnożyć przez najmniej korzystną wypłatę, a współczynnik optymizmu – przez sumę $(m - 1)$ najwyższych wypłat (przy $\alpha = 0.5$ nie ma znaczenia, którą metodę zastosujemy). Jak widać, przy obliczeniach weźmiemy pod uwagę wszystkie wypłaty, a takie założenie jest z kolei charakterystyczne dla zasady Bayesa. Dlatego właśnie proponowaną regułę nazwano hybrydą reguł Hurwicza i Bayesa. Aby wskaźniki hb_j mieściły się w przedziale $<w_j, m_j>$, wspomniane sumy odpowiednio zważonych wypłat zostaną na koniec podzielone przez sumę wszystkich wag.

Powyższa propozycja wymaga wyjaśnienia. Otóż dzięki takiemu podejściu mamy możliwość uwzględnienia częstotliwości występowania wypłat najwyższych i najniższych. Oznacza to, że pesymiście zostanie wskazana ta strategia, której najniższa wypłata jest względnie najwyższa lub której najwyższe wypłaty występują stosunkowo często (taki rozkład jest preferowany przez pesymistów, gdyż daje poczucie bezpieczeństwa), ponieważ jego współczynnik optymizmu będzie wagą dla każdej wypłaty oprócz jednej, tej, która znajduje się na końcu

nierosnącego ciągu wypłat. Z kolei optymiście, zgodnie z regułą H+B, zostanie zarekomendowana ta strategia, której najwyższa wypłata jest względnie największa, lecz której najkorzystniejsze wypłaty występują niekoniecznie często, ponieważ optymista takiego zabezpieczenia nie potrzebuje – liczy na to, że będzie miał szczęście i że wystąpi akurat najlepszy scenariusz dla wybranej decyzji. Współczynnik pesymizmu będzie w tym przypadku wagą dla każdej wypłaty oprócz jednej, tej, która zajmuje pierwszą pozycję w ciągu wypłat (patrz rys. 1).



Rys. 1. Ważenie wypłat w regule H+B

Stosowanie reguły H+B sprowadza się do realizacji następujących kroków:
Krok 1. Wyznaczyć dla każdej decyzji nierosnący ciąg wypłat Sq_j:

$$Sq_j = (a_{1j},...,a_{sj},...,a_{mj}) \tag{14}$$

gdzie:
– s jest numerem wyrazu tego ciągu,
– $a_{s,j} \geq a_{s+1,j}$ (s = 1, 2, …, m-1), $a_{1j} = m_j$, $a_{mj} = w_j$.

Krok 2. Obliczyć dla każdej decyzji wskaźnik hb_j (hb_j^p, hb_j^o lub $hb_j^{0.5}$ w zależności od parametru α).
1) Jeżeli $\alpha \in (0.5,1>$, obliczyć wskaźnik hb_j^p zgodnie ze wzorem (15):

$$hb_j^p = \frac{\alpha_p \cdot a_{mj} + \beta_p \cdot \sum_{s=1}^{m-1} a_{sj}}{(m-1)(1-\alpha)+\alpha} \tag{15}$$

2) Jeżeli $\alpha \in <0,0.5)$, obliczyć wskaźnik hb_j^o zgodnie ze wzorem (16):

$$hb_j^o = \frac{\alpha_o \cdot \sum_{s=2}^{m} a_{sj} + \beta_o \cdot a_{1j}}{(m-1)\cdot\alpha + 1 - \alpha} \tag{16}$$

3) Jeżeli $\alpha = 0.5$, obliczyć dla każdej decyzji wskaźnik hb_j^{05}, korzystając ze wzoru (17):

$$hb_j^{0.5} = hb_j^p = hb_j^o = b_j \tag{17}$$

Jak widać, w tym przypadku nie ma znaczenia, która formuła zostanie zastosowana (hb_j^p czy hb_j^o), ponieważ oba wzory prowadzą do uzyskania tych samych wartości, które są zresztą równe wskaźnikom Bayesa, gdyż wagi dla wszystkich wypłat są identyczne.

Krok 3. Wybrać tę strategię, która spełnia warunek (18):

$$hb_j^* = \max_j \{hb_j\} \tag{18}$$

Z konstrukcji metody wynika, że w odróżnieniu od reguły Hurwicza, suma wszystkich współczynników użytych jako wagi dla poszczególnych wypłat będzie zazwyczaj większa od jedności. Aby więc otrzymać wskaźniki spełniające warunek (19), co oczywiście nie zmienia rankingu decyzji, lecz jedynie proporcjonalnie obniża wartości wszystkich wskaźników, warto podzielić ważone sumy wypłat przez sumę wszystkich użytych wag. W przypadku pesymisty raz jest stosowany współczynnik pesymizmu (α) i (m – 1) razy jest wykorzystywany współczynnik optymizmu ($\beta = 1 - \alpha$), co w sumie daje (m – 1)(1 – α) + α (patrz mianownik we wzorze (15)). Natomiast w przypadku optymisty jedna wypłata jest mnożona przez współczynnik optymizmu ($\beta = 1 - \alpha$), a (m – 1) wypłat mnożymy przez współczynnik ostrożności (α), czyli suma wszystkich wag wynosi $(m-1)\cdot\alpha + 1 - \alpha$ (patrz mianownik we wzorze (16)).

$$w_j \le hb_j \le m_j \tag{19}$$

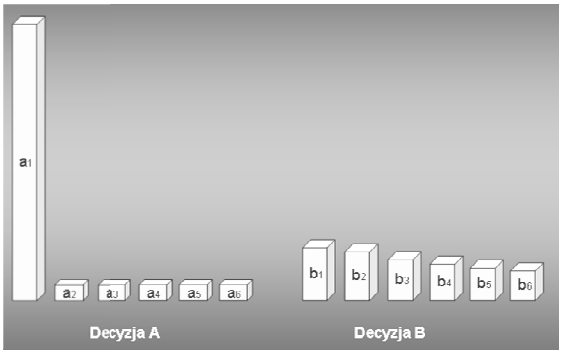
Z konstrukcji proponowanych wzorów (15) i (16) wynika, że we wskaźnikach ustalanych dla pesymisty wraz ze wzrostem liczby scenariuszy maleje znaczenie minimalnej wypłaty i rośnie znaczenie sumy (m – 1) najwyższych wypłat,

zaś we wskaźnikach obliczanych dla optymisty maleje znaczenie maksymalnej wypłaty i rośnie znaczenie sumy (m – 1) najniższych wypłat. Taka zależność nie występuje w pierwotnej wersji reguły Hurwicza – niezależnie od liczby stanów ważne są tylko dwie wypłaty: maksymalna i minimalna. Tu natomiast jest widoczny wpływ zasady Bayesa, z której wynika, że szansa wystąpienia danego scenariusza (także tego najbardziej i najmniej korzystnego) maleje, im scenariuszy jest więcej. Reguła H+B daje więc następuje korzyści:

– dla pesymisty: im więcej jest stanów, tym wyższy wskaźnik otrzyma wariant decyzyjny, dla którego większość scenariuszy oferuje wypłaty bliskie wartości maksymalnej związanej z tym wariantem (jest to pewna forma zabezpieczenia dla pesymisty),
– dla optymisty: im więcej jest scenariuszy, tym mniejszy wskaźnik otrzyma wariant decyzyjny, dla którego większość stanów oferuje wypłaty bliskie wartości minimalnej związanej z tym wariantem (dzięki temu optymiście zostanie zarekomendowana strategia, której najwyższe wypłaty są znacznie wyższe od najwyższych wypłat innych strategii).

Ostatecznie więc możemy stwierdzić, że wskaźniki H+B uwzględniają dwie informacje: 1) poziom pesymizmu (bądź optymizmu) decydenta (cecha charakterystyczna dla reguły Hurwicza), 2) szanse realizacji poszczególnych wypłat (cecha reguły Bayesa).

Ważenie wszystkich wypłat, a nie tylko skrajnych wartości, daje istotną przewagę reguły H+B nad regułą Hurwicza np. wówczas, gdy decydent ma do wyboru dwie decyzje (D_a i D_b) o zbiorach wypłat przedstawionych na rys. 2. Decyzja D_a charakteryzuje się jedną bardzo wysoką wypłatą (kilkakrotnie przewyższającą maksymalną wypłatę związaną z decyzją D_b), natomiast pozostałe jej wypłaty są kilkakrotnie niższe od minimalnej wypłaty decyzji D_b. Z kolei wypłaty decyzji D_b są dość wyrównane i, poza jednym scenariuszem, wyższe od wypłat decyzji D_a.



Rys. 2. Zbiór wypłat dla decyzji D_a i D_b

Decydent o nastawieniu mocno pesymistycznym powinien raczej być zainteresowany realizacją strategii D_b, której rozkład wypłat jest bardziej bezpieczny. Reguła Hurwicza i reguła H+B zaproponują decydentowi strategię D_a, gdy maksymalna wypłata tej decyzji będzie spełniać odpowiednio warunek (20) i (21):

$$\begin{gathered}(1-\alpha)\cdot a_1+\alpha\cdot a_6>(1-\alpha)\cdot b_1+\alpha\cdot b_6\\ a_1>\frac{(1-\alpha)\cdot b_1+\alpha\cdot(b_6-a_6)}{1-\alpha}\end{gathered} \tag{20}$$

$$\begin{gathered}\frac{(1-\alpha)\cdot(a_1+a_2+a_3+a_4+a_5)+\alpha\cdot a_6}{5\cdot(1-\alpha)+\alpha}>\frac{(1-\alpha)\cdot(b_1+b_2+b_3+b_4+b_5)+\alpha\cdot b_6}{5\cdot(1-\alpha)+\alpha}\\ a_1>\frac{(1-\alpha)\cdot(b_1+b_2+b_3+b_4+b_5-a_2-a_3-a_4-a_5)+\alpha\cdot(b_6-a_6)}{1-\alpha}\end{gathered} \tag{21}$$

Jak widać, minimalny poziom parametru a_1 dla reguły Hurwicza jest znacznie niższy od minimalnego poziomu tego parametru dla reguły H+B (wynika to m.in. z tego, że w liczniku we wzorze (21) występują ze znakiem dodatnim wszystkie wypłaty decyzji D_b, a nie tylko skrajne wartości). Zatem konkluzja jest taka, że klasyczna reguła Hurwicza sugeruje pesymistom decyzje o rzadkich wysokich wypłatach przy stosunkowo małej przewadze maksymalnej wypłaty nad maksymalnymi wypłatami innych strategii.

Reguła H+B stanowi pewne pośrednie narzędzie decyzyjne między regułą Hurwicza a regułą Bayesa. Dla współczynników α oscylujących wokół wartości 0.5 prezentowana zasada proponuje takie same rankingi, jak rankingi generowane za pomocą reguły Bayesa, co jest zrozumiałe, gdyż podobnie są ważone wszystkie wypłaty. W przypadku skrajnych pesymistów bądź optymistów rankingi H+B przypominają rankingi uzyskiwane na podstawie zasady Hurwicza, zwłaszcza wówczas, gdy liczba scenariuszy jest mała. W pozostałych sytuacjach, a więc gdy mamy do czynienia z umiarkowanymi pesymistami bądź optymistami lub gdy liczba stanów natury jest duża, ranking wariantów proponowany przez zasadę H+B nie przypomina już rankingów otrzymywanych za pomocą wspomnianych reguł klasycznych. W przypadku decyzji niezdominowanych rankingi generowane przez zasadę H+B będą bardziej wrażliwe na poziom parametru α, niż ma to miejsce przy regule Hurwicza, gdyż zmiana współczynnika ostrożności będzie wymagać zmiany wag dla wszystkich wypłat (a nie tylko skrajnych) we wskaźniku hb_j.

Tabela 5

Wskaźniki H+B (studium przypadku)

Projekty	Inwestor	
	A (umiarkowany optymista) ($\alpha_A = 0.3$)	B (umiarkowany pesymista) ($\alpha_B = 0.7$)
P1	$h^o_{P1} = \frac{0.3 \cdot 5 + 0.7 \cdot 5}{(6-1) \cdot 0.3 + (1-0.3)} = 2.27$	$h^p_{P1} = \frac{0.7 \cdot 1 + 0.3 \cdot 9}{0.7 + (6-1)(1-0.7)} = 1.55$
P2	$h^o_{P2} = \frac{0.3 \cdot 20.2 + 0.7 \cdot 4.8}{(6-1) \cdot 0.3 + (1-0.3)} = 4.28$	$h^p_{P2} = \frac{0.7 \cdot 1 + 0.3 \cdot 24}{0.7 + (6-1)(1-0.7)} = 3.59$
P3	$h^o_{P3} = \frac{0.3 \cdot 10 + 0.7 \cdot 3}{(6-1) \cdot 0.3 + (1-0.3)} = 2.32$	$h^p_{P3} = \frac{0.7 \cdot 1.5 + 0.3 \cdot 11.5}{0.7 + (6-1)(1-0.7)} = 2.05$
P4	$h^o_{P4} = \frac{0.3 \cdot 8.8 + 0.7 \cdot 4.8}{(6-1) \cdot 0.3 + (1-0.3)} = 2.73$	$h^p_{P4} = \frac{0.7 \cdot 1 + 0.3 \cdot 12.6}{0.7 + (6-1)(1-0.7)} = 2.04$
Ranking	I. P2 II. P4 III. P3 IV. P1	I. P2 II. P3 III. P4 IV. P1

Na koniec tej części powróćmy do przykładu analizowanego wcześniej. W tabeli 5 pokazano sposób wyznaczania wskaźników H+B dla umiarkowanego optymisty i umiarkowanego pesymisty:

1) Tym razem rankingi wygenerowane dla obu inwestorów nieco się różnią. Wynika to z faktu, że zmiana wartości parametru α powoduje zmianę wartości wag dotyczących wszystkich wypłat (a nie tylko skrajnych).
2) W obu przypadkach jest rekomendowany projekt P2, i jest to raczej rozsądny wybór, gdyż ani inwestor A nie jest skrajnym optymistą, by zdecydować się na projekt P1, ani inwestor B nie jest radykalnym pesymistą, by zdecydować się na projekt P3. A projekt P2 z jednej strony, o ile wystąpią korzystne scenariusze, daje wysokie wypłaty, a z drugiej jest relatywnie bezpieczny, gdyż częstotliwość pożądanych wyników jest tutaj najwyższa.
3) Projekt P4 uzyskuje niższą wartość wskaźnika H+B niż projekt P2, co jest logiczne, gdyż projekt P4 jest zdominowany przez ten projekt. Na taką dominację nie wskazywały natomiast rezultaty otrzymane za pomocą klasycznej reguły Hurwicza. Tam zawsze $h_{P2} = h_{P4}$.

Zakończenie

W pierwszych dwóch częściach artykułu pokazano, iż w niektórych przypadkach reguła Hurwicza stosowana przy PDWN może prowadzić do dość zaskakujących wyników, proponując rankingi słabo uwzględniające preferencje

decydentów (zwłaszcza pesymistów). Zasugerowano również, by tę regułę wykorzystywać tylko w przypadku problemów, w których: 1) liczba scenariuszy jest skończona, ale rozkład wypłat dla poszczególnych decyzji jest w miarę symetryczny lub gdy 2) liczba scenariuszy jest nieskończona i każda wypłata z przedziału $<w_j, m_j>$ jest możliwa.

Natomiast w trzeciej części zaproponowano regułę H+B, która stanowi hybrydę reguły Hurwicza i reguły Bayesa. Opracowano ją głównie z myślą o zadaniach ze skończoną liczbą stanów natury i ze skończonym zbiorem wypłat dla każdej decyzji, choć przy odpowiedniej modyfikacji zaproponowanych wzorów zaprezentowaną regułę można także stosować w zadaniach z ciągłym rozkładem wypłat. Reguła H+B była testowana nie tylko w ramach omówionego studium przypadku, lecz także w innych problemach decyzyjnych (z typowymi i nietypowymi rozkładami wypłat). Przy typowych rozkładach rankingi uzyskiwane za pomocą reguły Hurwicza i podejścia zmodyfikowanego są bardzo zbliżone i racjonalne. Ze względu na to, iż reguła H+B uśrednia, choć różnymi wagami, wszystkie wypłaty, może ona być przydatna zarówno przy jednokrotnej, jak i wielokrotnej realizacji wybranej decyzji, choć w tej drugiej sytuacji należy być ostrożnym, gdyż w dłuższym okresie współczynnik pesymizmu decydenta może się zmienić! Reguła H+B nie tylko pozwala deklarować swoje preferencje, ale je faktycznie uwzględnia w proponowanych rankingach. Bierze ona bowiem pod uwagę współczynnik ostrożności decydenta oraz częstotliwość występowania korzystnych i niekorzystnych wypłat. Dzięki takiej konstrukcji, nawet w przypadku problemów z dość specyficzną tabelą wypłat, rankingi otrzymane za pomocą tej metody lepiej odzwierciedlają faktyczne preferencje decydentów, a więc mogą być dla nich skutecznym narzędziem wspierającym proces decyzyjny.

Na koniec dodajmy, że zaproponowana koncepcja poszukiwania optymalnej strategii czystej nie jest całkowicie nowym pomysłem. W literaturze można znaleźć wiele opracowań zawierających opisy różnych metod stanowiących hybrydę kryterium Hurwicza i kryterium Baysesa (lub hybrydę innych klasycznych reguł PDWN). Warto jednak podkreślić, że istniejące już procedury wykorzystujące oba podejścia charakteryzują się odmienną konstrukcją i przeznaczeniem (dotyczą one raczej podejmowania decyzji w warunkach ryzyka ze względu na wprowadzenie prawdopodobieństwa) niż reguła H+B. Przykładowo w książce Ellsberga [2001] można znaleźć opis metody nazwanej „the restricted Bayes/Hurwicz criterion", w której wybór optymalnego wariantu jest uzależniony nie tylko od współczynnika pesymizmu $\alpha \in <0,1>$, lecz także od współczynnika niejednoznaczności $\rho \in <0,1>$ (ang. *degree of ambiguity*). Determinuje on prawdopodobieństwa poszczególnych wypłat. Im niższy, tym większe znaczenie przypi-

suje się wartościom skrajnym. Z kolei Basili i Zappia [2010, s. 449-474] oraz Ghiradato i inni [2004, s. 133-173] proponują kryterium „α-MEU ” (ang. α-*maxmin expected utility decision criterion*), w którym uwzględnia się zarówno prawdopodobieństwa dla poszczególnych wypłat, jak i poziom pesymizmu [Marinacci 2002, s. 755-764]. Na podstawie koncepcji CEU (ang. *Choquet expected utility*) Basili opracował także metodę przypisującą inne prawdopodobieństwa wartościom pośrednim i inne prawdopodobieństwa wartościom skrajnym [Basili 2006, s. 1721-1728; Basili, Chateauneuf, Fontini 2008, s. 485-491; Gilboa 2009]. Obok wymienionych już podejść warto wspomnieć też o pracy Nakamury [1986, s. 147-162][*], w której zaprezentowano pewne rozmyte rozszerzenie reguły Hurwicza oparte na odległości Hamminga między dwoma zbiorami (g.u.s i g.l.s, ang. *the greatest upper set, the greatest lower set*) rozmytej użyteczności.

Bibliografia

Basili M., 2006: A Rational Decision Rule with Extreme Events. „Risk Analysis”, Vol. 26, 1721-1728.

Basili M., Chateauneuf A., Fontini F., 2008: Precautionary Principle as a Rule of Choice with Optimism on Windfall Gains and Pessimism on Catastrophic Losses. „Ecological Economics”, Vol. 67, 485-491.

Basili M., Zappia C., 2010: Ambiguity and Uncertainty in Ellsberg and Shackle. „Cambridge Journal of Economics”, 34(3), 449-474.

Dominiak C., 2006: Multicriteria Decision Aid under Uncertainty. W: Multiple Criteria Decision Making’ 05. T. Trzaskalik (ed.). Publisher of the Karol Adamiecki University of Economics, Katowice.

Dominiak C., 2009: Multicriteria Decision Aiding Procedure under Risk and Uncertainty. W: Multiple Criteria Decision Making’ 08. T. Trzaskalik (ed.). Publisher of the Karol Adamiecki University of Economics, Katowice.

Ellsberg D., 2001: Risk, Ambiguity and Decision. Garland Publishing, New York, NY, USA.

Gaspars-Wieloch H., 2013: Modifications of the Hurwicz’s Decision Rule. „Central European Journal of Operations Research”, Springer, DOI 10.1007/s10100-013-0302-y, May.

Gaspars-Wieloch H., 2012: Ograniczona skuteczność metod optymalizacyjnych w rozwiązywaniu ekonomicznych problemów decyzyjnych. „Ekonomista”, 3, 303-324.

Ghirardato P., Maccheroni F., Marinacci M., 2004: Differentiating Ambiguity and Ambiguity Attitude. „Journal of Economic Theory”, Vol. 118, 133-173.

Gilboa I., 2009: Theory of Decision under Uncertainty. Cambridge University Press.

[*] Opis podejścia Nakamury można znaleźć też w pracy Piaseckiego [1990].

Groenewald M.E., Pretorius P.D., 2011: Comparison of Decision-making under Uncertainty Investment Strategies with the Money Market. „Journal of Financial Studies and Research".

Hurwicz L., 1952: A Criterion for Decision Making under Uncertainty. „Technical Report 355", Cowles Commission.

Hurwicz L., 1951: The Generalized Bayes Minimax Principle: A Criterion for Decision Making Under Uncertainty. Cowles Commission, „Discussion Paper Statistics 335".

Huynh V.N., Hu C., Nakamori Y., Kreinovich V., 2007: On Decision Making under Interval Uncertainty: A New Justification of Hurwicz Optimism-Pessimism Approach and Its Use in Group Decision Making. „Departmental Technical Reports (CS)", Paper 107.

Ignasiak E. (red.), 1996: Badania operacyjne. Polskie Wydawnictwo Ekonomiczne, Warszawa.

Kaufmann A., Faure R., 1974: Invitations a la recherche operationnelle. Paris, Dunod.

Knight F.H., 1921: Risk, Uncertainty, Profit. Hart, Schaffner & Marx, Houghton Mifflin Co Boston, MA.

Marinacci M., 2002: Probabilistic Sophistication and Multiple Priors. „Econometrica", Vol. 70, 755-764.

Nakamura K., 1986: Preference Relations on a Set of Fuzzy Utilities as a Basis for Decision Making. „Fuzzy Sets and Systems", Vol. 20, 147-162.

Neumann J., Morgenstern O., 1944: Theory of Games and Economic Behavior. Princeton University Press.

Pazek K., Rozman C., 2009: Decision Making under Conditions of Uncertainty in Agriculture: A Case Study of Oil Crops. „Poljoprivreda (Osijek)", Vol. 15(1), 45-50.

Piasecki K., 1990: Decyzje i wiarygodne prognozy. Zeszyty Naukowe nr 106, Akademia Ekonomiczna, Poznań.

Render B., Stair R.M., Hanna M.E., 2006: Quantitative Analysis for Management. Pearson Prentice Hall, Upper Saddle River, New Jersey.

Savage L.J., 1961: The Foundations of Statistics Reconsidered. W: Studies in Subjective Probability. H.E. Kyburg, H.E. Smokler (eds.). New York, Wiley, 173-188.

Siddiqui A., Chronopoulos M.: Optimal Investment and Operational Decision Making under Risk Aversion and Uncertainty. „European Journal Operations Research" (w druku).

Sikora W. (red.), 2008: Badania operacyjne. Polskie Wydawnictwo Ekonomiczne, Warszawa.

Trzaskalik T., 2008: Wprowadzenie do badań operacyjnych z komputerem. Wydanie II zmienione. Polskie Wydawnictwo Ekonomiczne, Warszawa.

Wald A., 1950a: Basic Ideas of a General Theory of Statistical Decisions Rules. W: Selected Papers in Statistics and Probability. A. Wald (ed.). McGraw-Hill, New York, 656-668.

Wald A., 1950b: Statistical Decision Functions. Wiley, New York.

A HYBRID OF THE HURWICZ AND BAYES RULES IN THE DECISION MAKING UNDER UNCERTAINTY

Summary

The Hurwicz rule and the Bayes rule are classical approaches applied in the decision making under uncertainty. This situation occurs when the decision maker may choose one of several alternatives and he or she is only able to assign to each of them an interval of potential payoffs or a set of possible profits. In both cases the answer obtained depends on the state of nature (scenario) which will happen, but in the first case the set of scenarios is infinite and in the second one – it is finite. The Hurwicz measure, with the aid of the coefficient of pessimism and the coefficient of optimism, enables to find the optimal pure strategy when the decision selected is performed only once. Meanwhile the Bayes criterion is designed to indicate the optimal pure or mixed strategy when the variant chosen is performed once or many times.

In the first part of the article the author analyzes the Hurwicz rule and illustrates cases when the use of this criterion leads to quite unexpected results which seem to be contradictory with the logic and do not reflect the decision maker's preferences. In the second part a proposal of an approach for optimal pure strategy searching (by means of formulas considering both the coefficients of pessimism and optimism, as well as the whole set of payoffs) is presented. This procedure (H+B rule) combines elements of the Hurwicz criterion and the Bayes criterion, but is deprived of disadvantages typical of the Hurwicz rule. The rule suggested takes into consideration both extreme payoffs and intermediate payoffs, which enables to receive rational recommendations for a larger spectrum of decision problems. The H+B rule may be applied in the decision making process under uncertainty when the number of potential scenarios and the set of possible payoffs are finite, however a slight modification of the equations proposed enables to use this procedure in problems with continuous payoffs.

Robert Jankowski
Uniwersytet w Białymstoku

OPÓŹNIENIA W MODELACH SPOŁECZNYCH DYNAMIKI REPLIKATOROWEJ

Wstęp

Pionierem w dziedzinie badania ewolucji populacji z wykorzystaniem analizy teorii gier był John Maynard Smith. Wprowadził on definicję strategii ewolucyjnie stabilnej. Jeżeli wszyscy grają taką strategię, to wówczas żadna inna strategia nie może rozpowszechniać się w populacji. Kolejnym krokiem było stworzenie systemu równań różnicowych (różniczkowych), zwanych równaniami replikatorowymi [Taylor, Jonker 1978; Hofbauer 1979; Zeeman 1981], opisujących ewolucję w czasie częstotliwości strategii graczy. Wiadomo przy tym, że każda równowaga Nasha będąca ewolucyjnie stabilną strategią jest lokalnie asymptotycznie stabilnym stacjonarnym punktem w takim dynamicznym modelu [Hofbauer, Sigmund 1988; Weibull 1997].

Zazwyczaj zakłada się, iż efekt interakcji pomiędzy graczami jest natychmiastowy. W rzeczywistości rezultat może się pojawić w przyszłości. Przykładem są modele dynamiki replikatorowej, w których gracze wybierają swoje strategie na podstawie wydarzeń z przeszłości [Albosza, Miękisz 2004].

W poniższym artykule zostanie omówiony wpływ opóźnień na stabilność wewnętrznego punktu stacjonarnego dynamiki replikatorowej (decyzji optymalnej w populacji) w modelach społecznych. Zaprezentowane wyniki zostaną poparte przykładami i symulacjami. Obok standardowego modelu dynamiki replikatorowej zostanie przedstawiony model zmodyfikowany zakładający, iż gracze naśladują przeciwników stosujących strategie z wyższą średnią wypłatą, z uwzględnieniem opóźnionej informacji, oraz model, w którym każda strategia będzie miała własne opóźnienie.

1. Standardowa dynamika replikatorowa

Załóżmy, że gracze w pewnej dużej, ale skończonej populacji mają do dyspozycji tylko dwie różne czyste strategie: A i B. W każdym momencie czasu gracze parują się i odbywają dwuosobową grę, której macierz wypłat jest następująca:

$$W = \begin{matrix} \\ A \\ B \end{matrix} \begin{matrix} \begin{matrix} A & B \end{matrix} \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \end{matrix},$$

gdzie W_{kl}, k, l = A, B jest wypłatą pierwszego (wierszowego) gracza grającego strategię k przy założeniu, że gracz drugi (kolumnowy) gra strategię l. Załóżmy, że obaj gracze są tacy sami, a wypłaty gracza kolumnowego są dane poprzez macierz transponowaną do W. Takie gry nazywamy symetrycznymi. Co więcej, aby istniała unikalna mieszana równowaga Nasha, należy założyć, że zawsze c > a oraz b > d lub c < a oraz b < d. $x^* = \frac{\Delta_2}{\Delta}$, gdzie $\Delta_1 = c - a$, $\Delta_2 = b - d$ i $\Delta = \Delta_1 + \Delta_2$ jest optymalną (w sensie równowagi Nasha) częstotliwością występowania strategii A w populacji.

Niech $g_i(t)$, $i = A, B$ będzie liczbą graczy grających odpowiednio strategię A oraz B w chwili t. Wówczas $g(t) = g_A(t) + g_B(t)$ jest liczbą wszystkich graczy, a $x(t) = \frac{g_A(t)}{g(t)}$ jest odsetkiem populacji graczy grających strategię A. Zakładamy, że w krótkim okresie ε tylko część populacji bierze udział w grach (mianowicie ε). Mamy więc:

$$g_i(t+\varepsilon) = (1-\varepsilon)g_i(t) + \varepsilon g_i(t)P_i(t); \quad i = A, B, \tag{1}$$

gdzie $P_A(t) = ax(t) + b(1 - x(t))$ i $P_B(t) = cx(t) + d(1 - x(t))$ są średnimi wypłatami graczy, którzy grają strategię A i B odpowiednio. Całkowita liczba graczy da się opisać za pomocą następującego równania:

$$g(t+\varepsilon) = (1-\varepsilon)g(t) + \varepsilon g(t)\overline{P}(t), \tag{2}$$

gdzie $\overline{P}(t) = x(t)P_A(t) + (1 - x(t))P_B(t)$ jest średnią wypłatą w populacji w chwili t. Otrzymujemy natychmiast równanie opisujące częstotliwość występowania strategii A:

$$x(t+\varepsilon)-x(t)=\varepsilon\frac{x(t)[P_A(t)-\overline{P}(t)]}{1-\varepsilon+\varepsilon\overline{P}(t)}. \tag{3}$$

Jeżeli podzielimy teraz obie strony równania przez ε i przejdziemy do granicy $\varepsilon \to 0$, to otrzymamy różniczkowe równanie replikatorowe:

$$\frac{dx(t)}{dt}=x(t)[P_A(t)-\overline{P}(t)]. \tag{4}$$

Równanie (4) można równoważnie zapisać jako:

$$\frac{dx(t)}{dt}=x(t)(1-x(t))[P_A(t)-P_B(t)]=(a-c+d-b)x(t)(1-x(t))(x(t)-x^*). \tag{5}$$

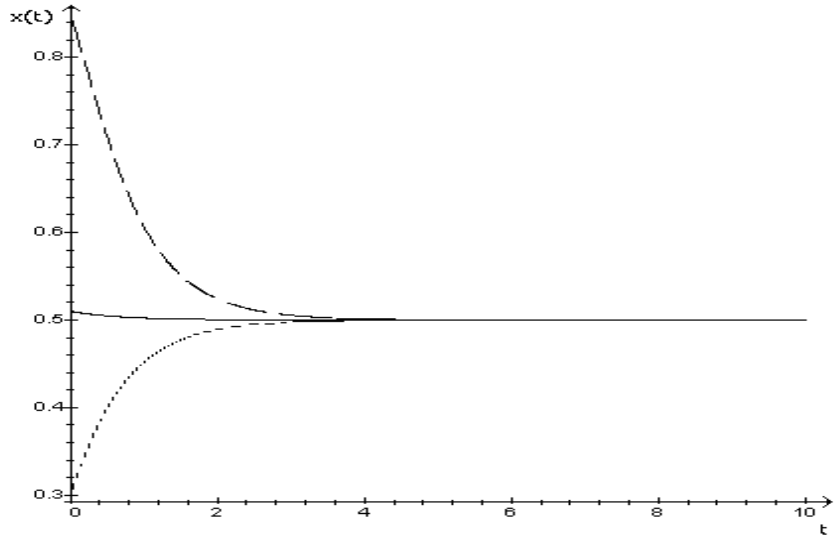
Widzimy, że punkt x* jest asymptotycznie stabilny. Oznacza to, że populacja graczy prędzej czy później będzie otrzymywała optymalne (w sensie równowagi Nasha) wypłaty.

Przykład 1

Rozważmy grę z następującą macierzą wypłat:

$$W=\begin{array}{c} \\ A \\ B \end{array}\begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \end{array}.$$

Jak łatwo zauważyć, x*=1/2. Oznacza to, że niezależnie od warunków początkowych, przy założeniu standardowej dynamiki replikatorowej, po pewnym czasie, z jednakowym prawdopodobieństwem, będzie można znaleźć w populacji graczy stosujących zarówno strategię A, jak i B. Wykorzystując program Maple, można prześledzić symulacje dynamiki częstotliwości wybierania przez graczy strategii A w nieustannie rosnącej populacji:



Rys. 1. Zmiana częstotliwości x(t) w zależności od czasu, przy różnych warunkach początkowych – 0,85; 0,51; 0,3; ε = 0,01

Na rysunku 1 widać, że niezależnie od warunków początkowych strategia x* stabilizuje się. Maksymalnie po 5 jednostkach czasu (przyjmijmy, że są to sekundy) w całej populacji tak samo często stosuje się strategię A, jak i B.

2. Dynamika replikatorowa z jednym opóźnieniem

Zakładamy, że gracze w czasie t zwiększają swoją liczebność zgodnie ze średnimi wypłatami uzyskanymi w czasie $t-\tau$ dla $\tau > 0$. Tak jak w standardowej dynamice replikatorowej, zakładamy, że w krótkim okresie ε tylko część populacji bierze udział w grach.

Niech $g_i(t)$, $i = A, B$ będzie liczbą graczy grających odpowiednio strategię A oraz B w chwili t. Wówczas $g(t) = g_A(t) + g_B(t)$ jest liczbą wszystkich graczy, a $x(t) = \dfrac{g_A(t)}{g(t)}$ jest odsetkiem populacji graczy grających strategię A. Otrzymujemy zatem następujące równanie:

$$g_i(t+\varepsilon) = (1-\varepsilon)g_i(t) + \varepsilon g_i(t)P_i(t-\tau)\,;\ i = A, B. \tag{6}$$

Wówczas całkowita liczba graczy da się opisać równaniem:

$$g(t+\varepsilon) = (1-\varepsilon)g(t) + \varepsilon g(t)\overline{P'}(t-\tau), \tag{7}$$

gdzie $\overline{P'}(t-\tau) = x(t)P_A(t-\tau) + (1-x(t))P_B(t-\tau)$. Wzory (6) i (7) pokazują, jak zmienia się częstotliwość występowania strategii A:

$$x(t+\varepsilon) - x(t) = \varepsilon \frac{x(t)[P_A(t-\tau) - \overline{P'}(t-\tau)]}{1-\varepsilon+\varepsilon\overline{P'}(t-\tau)}. \tag{8}$$

Jeżeli podzielimy teraz obie strony równania przez ε i przejdziemy do granicy $\varepsilon \to 0$, to otrzymamy opóźnione różniczkowe równanie replikatorowe:

$$\frac{dx(t)}{dt} = x(t)[P_A(t-\tau) - \overline{P'}(t-\tau)]. \tag{9}$$

Można je równoważnie zapisać jako:

$$\frac{dx(t)}{dt} = x(t)(1-x(t))[P_A(t-\tau) - P_B(t-\tau)] = -\Delta x(t)(1-x(t))(x(t-\tau) - x^*). \tag{10}$$

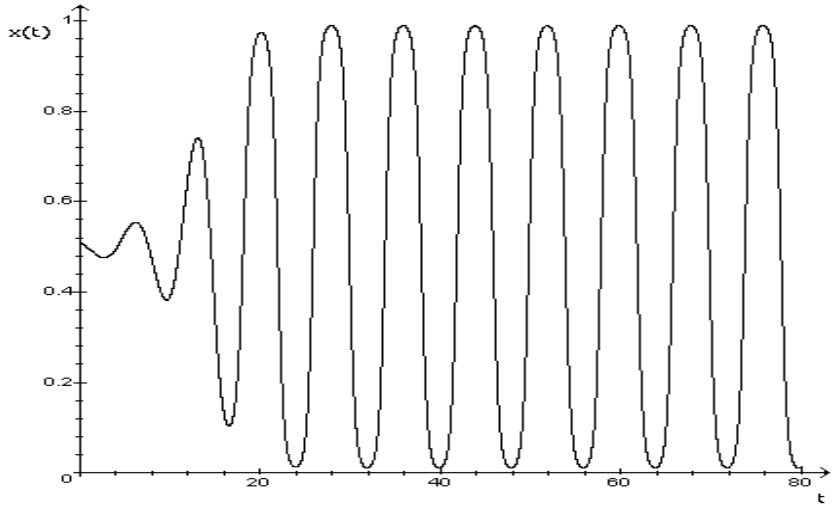
Załóżmy, że losowo wybrany gracz naśladuje losowo wybranego przeciwnika. Wówczas prawdopodobieństwo, że gracz, który gra strategię A, będzie chciał naśladować przeciwnika grającego B w czasie t, wynosi dokładnie $x(t)(1 - x(t))$. Natężenie naśladowania zależy od opóźnionej informacji o różnicy odpowiednich wypłat w chwili $t - \tau$.

Tao i Wang [1997] pokazali, że jeżeli $\tau < (c-a+b-d)\pi / 2(c-a)(b-d)$, to równowaga Nasha (w strategiach mieszanych) x^* jest punktem asymptotycznie stabilnym opóźnionego równania różniczkowego. Jeżeli natomiast opóźnienie jest większe niż $(c - a + b - d)\pi / 2(c - a)(b - d)$, to x^* staje się punktem niestabilnym.

Przykład 2

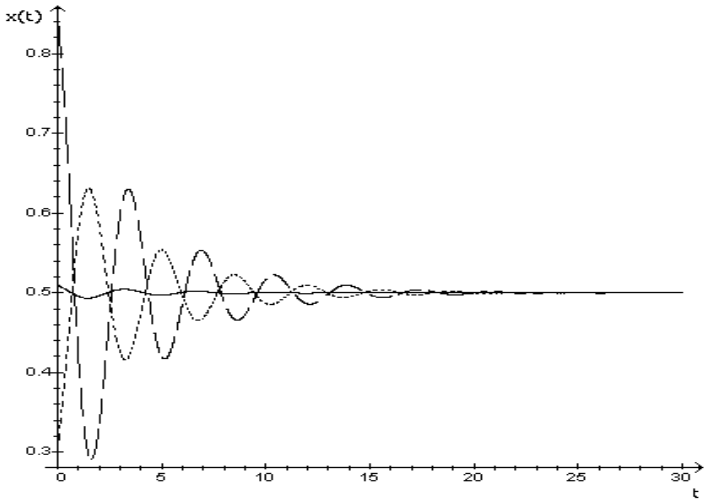
Rozważmy ponownie macierz wypłat z przykładu 1. Tak jak poprzednio x* = 1/2. Tym razem przed wyborem strategii pozwólmy graczom sprawdzać średnie wypłaty z gry z przeszłości. Aby w populacji utrwaliła się optymalna strategia x* (w sensie równowagi Nasha), horyzont czasowy nie może być większy niż $\frac{\pi}{3}$. Sytu-

ację, w której gracze podejmują swoje decyzje na podstawie wypłat uzyskanych $\tau = 1{,}3$ jednostek czasu (sekundy) temu pokazuje rys. 2.



Rys. 2. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,51; $\tau = 1{,}3$; $\varepsilon = 0{,}01$ – brak stabilności strategii x*

Nawet niewiele opóźniona informacja prowadzi bardzo szybko do mocnych wahań częstotliwości x(t). Zbyt stare informacje powodują rozchwianie populacji. W jednym momencie czasu prawie wszyscy gracze stosują strategię A, a po chwili sytuacja całkowicie się odwraca. W populacji prawie wcale nie stosuje się optymalnej (w sensie równowagi Nasha) strategii x*. Należy jednak pamiętać o tym, że opóźnienie powodujące rozchwianie częstotliwości x(t) zależy tylko i wyłącznie od macierzy wypłat. Niemniej jednak można przewidzieć, z jaką częstotliwością będzie występowała strategia A w populacji w konkretnym momencie czasu. Przeanalizujmy jeszcze sytuację, gdy $\tau = 0{,}8$ (rys. 3).



Rys. 3. Zmiana częstotliwości x(t) w zależności od czasu, przy różnych warunkach początkowych - 0,85; 0,51; 0,3; τ = 0,8; ε = 0,01

Jak pokazuje rys. 3, jeżeli zostanie spełnione kryterium $\tau < \frac{\pi}{3}$, to niezależnie od warunków początkowych strategia x* się stabilizuje. Zauważmy jeszcze, że im warunek początkowy jest bliższy x*, tym szybciej możemy się spodziewać sytuacji, w której prawdopodobieństwo znalezienia w populacji gracza grającego strategię A wynosi 1/2.

3. Dynamika replikatorowa z dwoma opóźnieniami

Rozważmy teraz sytuację, w której gracze w czasie t zwiększają swoją liczebność zgodnie ze średnimi wypłatami uzyskanymi przez ich strategie w momentach czasu $t - \tau_A$ i $t - \tau_B$ odpowiednio.

Niech $g_i(t)$, $i = A, B$ będzie liczbą graczy grających strategię A oraz B odpowiednio w chwili t. Zatem $g(t) = g_A(t) + g_B(t)$ jest liczbą wszystkich graczy, a $x(t) = \frac{g_A(t)}{g(t)}$ jest odsetkiem populacji graczy grających strategię A. Mamy wówczas następujące równanie:

$$g_i(t+\varepsilon) = (1-\varepsilon)g_i(t) + \varepsilon g_i(t)P_i(t-\tau_i); \; i = A, B. \tag{11}$$

Całkowita liczba graczy da się opisać poprzez równanie:

$$g(t+\varepsilon) = (1-\varepsilon)g(t) + \varepsilon(r_A(t)P_A(t-\tau_A) + r_B(t)P_B(t-\tau_B)). \quad (12)$$

Równania (11) i (12) pokazują, jak zmienia się częstotliwość występowania strategii A:

$$x(t+\varepsilon) - x(t) = \varepsilon \frac{x(t)P_A(t-\tau_A) - x^2(t)P_A(t-\tau_A) - x(t)(1-x(t))P_B(t-\tau_B)}{1-\varepsilon+\varepsilon(x(t)P_A(t-\tau_A) + (1-x(t))P_B(t-\tau_B))}. \quad (13)$$

Jeżeli podzielimy teraz obie strony równania przez ε i przejdziemy do granicy $\varepsilon \to 0$, to otrzymamy opóźnione różniczkowe równanie replikatorowe:

$$\frac{dx(t)}{dt} = x(t)(1-x(t))[P_A(t-\tau_A) - P_B(t-\tau_B)]. \quad (14)$$

Załóżmy, że losowo wybrany gracz naśladuje losowo wybranego przeciwnika. Wówczas prawdopodobieństwo, że gracz, który gra strategię A, będzie chciał naśladować przeciwnika grającego B w czasie t, wynosi dokładnie $x(t)(1 - x(t))$. Natężenie naśladowania zależy od opóźnionych informacji o różnicy odpowiednich wypłat we wcześniejszych momentach czasu.

Twierdzenie 1*

Jeżeli jeden z poniższych warunków będzie spełniony:

(1) $b-a+c-d>0$, $|b-a|\tau_A + |c-d|\tau_B < \dfrac{b-a+c-d}{|b-a|+|c-d|}$;

(2) $b-a>0$, $(b-a)\tau_A < \dfrac{b-a-|c-d|}{b-a+|c-d|}$;

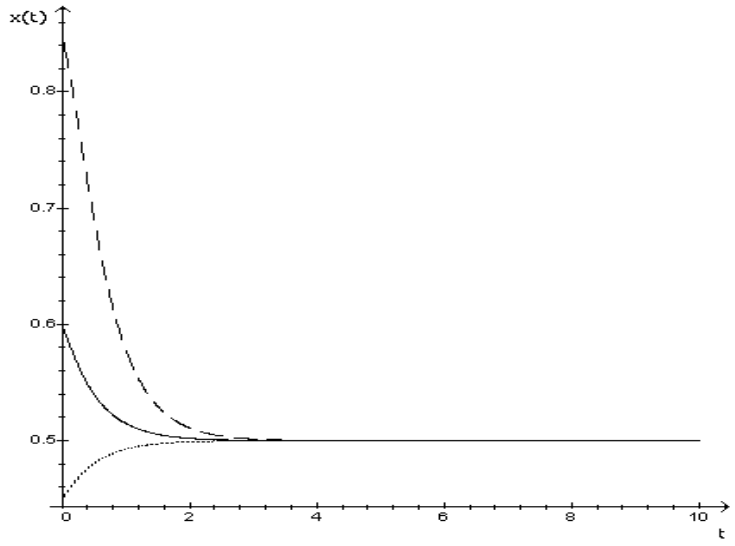
(3) $c-d>0$, $(c-d)\tau_B < \dfrac{c-d-|b-a|}{|b-a|+c-d}$,

to x* będzie punktem asymptotycznie stabilnym równania (14).

* Berezansky, Braverman [2006, s. 391-411]; Li, Ruin, Wei [1999, s. 254-280].

Przykład 3

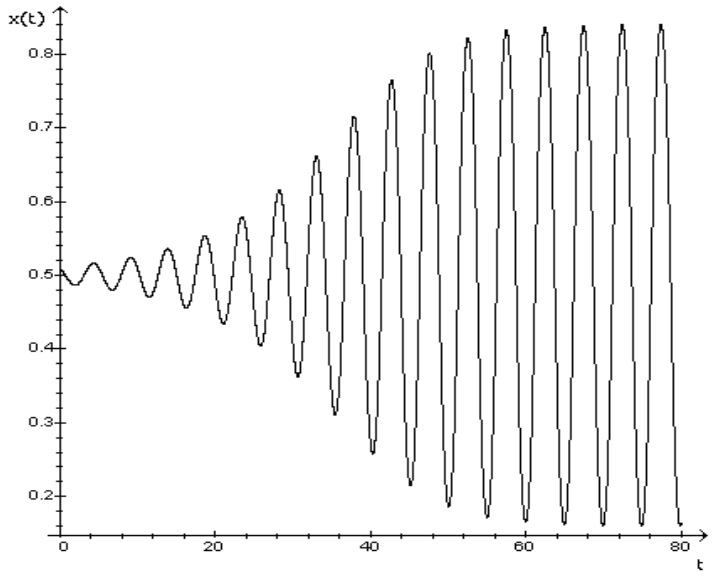
Rozważmy ponownie macierz wypłat z przykładu 1. Tak jak poprzednio x* = 1/2. Wykorzystując twierdzenie 1, zauważamy, że $b - a + c - d > 0$. Zatem jeżeli $\tau_A + \tau_B < \frac{1}{3}$, to możemy oczekiwać, że po pewnym czasie z jednakowym prawdopodobieństwem w populacji będzie występowała zarówno strategia A, jak i B. Pozostałe warunki nie mogą być spełnione, ponieważ zawsze $\tau_i > 0$ dla i = A, B. Przeprowadźmy więc symulacje w przypadku, gdy τ_A = 0,16, a τ_B = 0,14 (rys. 4).



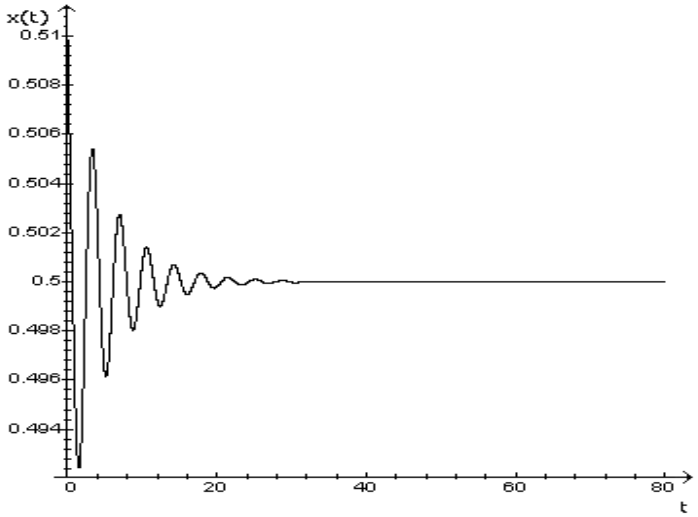
Rys. 4. Zmiana częstotliwości x(t) w zależności od czasu, przy różnych warunkach początkowych – 0,85; 0,6; 0,45; τ_A = 0,16; τ_B = 0,14; ε = 0,01

W efekcie otrzymujemy wykres bardzo podobny do rys. 1. Nie powinno to dziwić, gdyż opóźnienia są bardzo mocno zbliżone do zera.

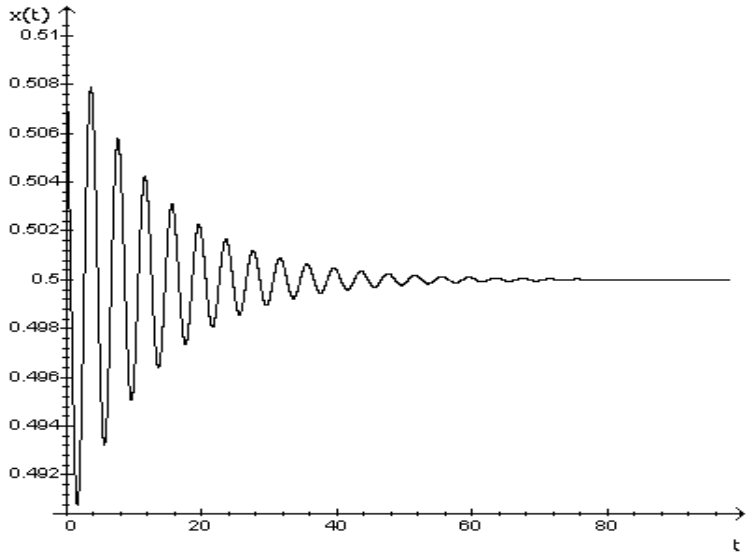
W twierdzeniu 1 nie występuje równoważność, ale sprawdźmy, co się stanie, gdy np. τ_A = 1,1, a τ_B = 1,4 (rys. 5). Nietrudno się domyślić, że x* przestanie być punktem stacjonarnym. Przeprowadzone symulacje, przedstawione na rys. 5, 6, 7, pokazują, że dzieje się tak nie dlatego, że $\tau_A + \tau_B > \frac{1}{3}$, a bardziej z powodu niespełnienia warunku $\tau_i < \frac{\pi}{3} = (c - a + b - d)\pi / 2(c - a)(b - d)$ dla i = A, B.



Rys. 5. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,51; $\tau_A = 1{,}1$; $\tau_B = 1{,}4$; $\varepsilon = 0{,}01$



Rys. 6. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,51; $\tau_A = 0{,}8$; $\tau_B = 0{,}9$; $\varepsilon = 0{,}01$



Rys. 7. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,51; $\tau_A = 1{,}1$; $\tau_B = 0{,}8$; $\varepsilon = 0{,}01$

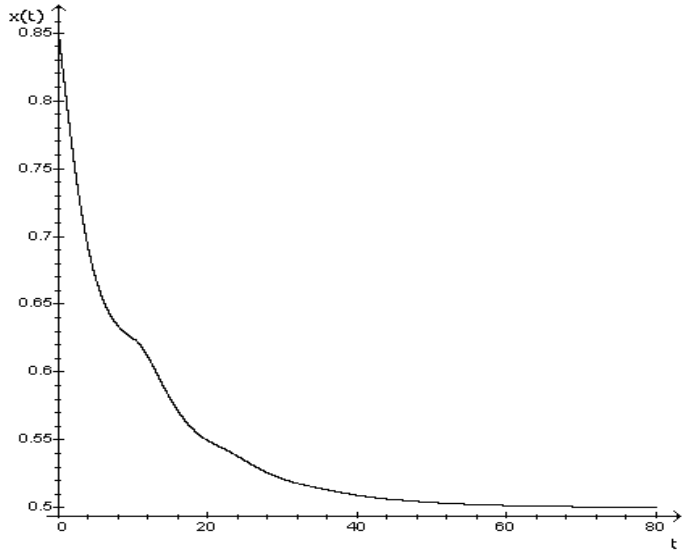
Przykład 4

Rozważmy grę znaną w literaturze pod nazwą „jastrząb-gołąb". Została ona szczegółowo opisana przez Smitha oraz Price'a [1973]. Gracze walczą o zasoby. W biologicznym sensie im więcej ich posiadają (może to być np. powierzchnia potrzebna do rozrodu), tym liczniejsze mają potomstwo. Każdy z graczy ma do dyspozycji tylko dwie strategie: jastrzębia – zawsze walczyć o zasoby, i gołębia – nie wdawać się w konflikty. Jastrzębie mają 50% szansy na wygraną z innym jastrzębiem. Gołębie zawsze uciekają, jeżeli tylko ktoś je zaatakuje. Jastrzębie walcząc odnoszą rany, które kosztują je dwa razy więcej niż zysk zasobu. Natomiast gołębie zawsze się wszystkim dzielą ze sobą. Przyjmiemy, że zasób ma wartość 1. Wówczas macierz wypłat takiej gry jest następująca:

$$W = \begin{matrix} \\ J \\ G \end{matrix} \begin{matrix} \begin{matrix} J & G \end{matrix} \\ \begin{bmatrix} -½ & 1 \\ 0 & ½ \end{bmatrix} \end{matrix}.$$

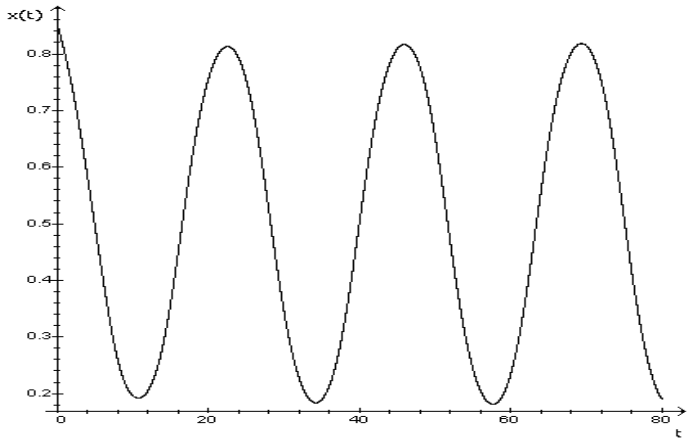
Załóżmy, że każda strategia ma swoje własne opóźnienie. Zatem jak daleko mogą sięgać gracze pamięcią wstecz, aby stan x* był asymptotycznie stabilny?

Rozważmy warunek (2) twierdzenia 1: $b - a = \frac{3}{2} > 0$, więc jeżeli tylko $\tau_J < \frac{1}{3}$, to w populacji utrzyma się równowaga w stosowaniu strategii A i B na poziomie x* = 1/2. Zauważmy ponadto, że przy stosowaniu strategii gołębia możemy sprawdzać dowolnie odległe (stare) średnie wypłaty z gry. Przeprowadźmy przykładową symulację, w której przyjmiemy $\tau_J = 0{,}2$, $\tau_G = 10$. Otrzymany w jej wyniku wykres przedstawiony na rys. 8 jest bardzo zbliżony do rys. 1.



Rys. 8. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,85; $\tau_J = 0{,}2$; $\tau_G = 10$; $\varepsilon = 0{,}01$

Analizując rys. 8, warto zauważyć, że znacząco wydłużył się czas potrzebny do osiągnięcia równowagi. Jest to spowodowane bardzo dużym opóźnieniem strategii gołębia. Kiedy zatem strategia x* nie będzie stanem stabilnym? Zgodnie z przypuszczeniami z przykładu 3 niech wypłaty obu strategii będą opóźnione o więcej niż $2\pi = (c - a + b - d)\pi/2(c - a)(b - d)$. Przeprowadźmy symulację, w której $\tau_J = 7$, $\tau_G = 10$ (rys. 9).



Rys. 9. Zmiana częstotliwości x(t) w zależności od czasu, przy warunku początkowym 0,85; $\tau_J = 7$; $\tau_G = 10$; $\varepsilon = 0{,}01$

Na rysunku 9 widać wyraźnie brak stabilności x*. Wydłużył się również znacząco czas potrzebny na przejście populacji od stanu „prawie wszyscy stosują strategię J" do „prawie wszyscy stosują strategię G".

Podsumowanie

Zaprezentowane modele mają na celu pokazanie sposobu podejmowania decyzji w populacji. Zanim cokolwiek kupimy, sprzedamy, zainwestujemy, sprawdzamy, czy nasze działanie przyniesie korzyści. Jak to zrobić najprościej? Wystarczy skorzystać z doświadczenia swojego lub bliskich. Wystarczy przypomnieć sobie, czy dawniej dana strategia dała nam realne zyski w porównaniu z inną opcją. Niekiedy napotykamy dodatkowy problem – decyzje nie mogą być porównywane w tym samym punkcie przeszłości.

W zaprezentowanych modelach istnieją ostre warunki zapewniające stabilność równowagi Nasha. Niezależnie od tego, czy mamy jedno opóźnienie, czy dwa, potrafimy wyznaczyć maksymalnie odległe punkty w przeszłości, w których zawarte informacje zapewnią optymalne wypłaty graczom. Jest to bardzo naturalne – również w prawdziwym życiu nieaktualne informacje są nieprzydatne i mogą prowadzić do przekłamań o rzeczywistości.

Oczywiście przedstawione modele dotyczą sytuacji uproszczonych, zazwyczaj mamy bowiem więcej możliwości do wyboru niż tylko dwie strategie. Niemniej jednak jest to dobry punkt wyjścia do rozważań na temat podświadomego dążenia do równowagi Nasha w społeczeństwie.

Bibliografia

Albosza J., Miękisz J., 2004: Stability of Evolutionarily Stable Strategies in Discrete Replicator Dynamics. „Journal of Theoretical Biology", Vol. 231, No. 2, 175-179.

Berezansky L., Braverman E., 2006: On Stability of Some Linear and Nonlinear Delay Differential Equations. „Journal of Mathematical Analysis and Applications", Vol. 314, No. 2, 391-411.

Friedman D., 1998: On Economic Applications of Evolutionary Game Theory. „Journal of Evolutionary Economics", No. 8, 15-43.

Hofbauer J., Shuster P., Sigmund K., 1979: A Note on Evolutionarily Stable Strategies and Game Dynamics. „Journal of Theoretical Biology", Vol. 81, No. 3, 609-612.

Hofbauer J., Sigmund K., 1988: The Theory of Evolution and Dynamical Systems. Cambridge University Press, Cambridge.

Iijima R., 2012: On Delayed Discrete Evolutionary Dynamics. „Journal of Theoretical Biology", Vol. 300, 1-6.

Li X., Ruin S., Wei J., 1999: Stability and Bifurcation in Delay-differential Equations with Two Delays. „Journal of Mathematical Analysis and Applications", Vol. 236, No. 2, 254-280.

Maynard Smith J., 1974: The Theory of Games and the Evolution of Animal Conflicts. „Journal of Theoretical Biology", Vol. 47, No. 1, 209-221.

Maynard Smith J., Price G.R., 1973: The Logic of Animal Conflict. „Nature", Vol. 246, No. 5427, 15-18.

Maynard Smith J., 1982: Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Moreira J.A., Pinheiro F.L., Nunes A., Pacheco J.M., 2012: Evolutionary Dynamics of Collective Action when Individual Fitness Derives from Group Decisions Taken in the Past. „Journal of Theoretical Biology", Vol. 298, 8-15.

Tao Y., Wang Z., 1997: Effect of Time Delay and Evolutionarily Stable Strategy. „Journal of Theoretical Biology", Vol. 187, No. 1, 111-116.

Taylor P.D., Jonker L.B., 1978: Evolutionarily Stable Strategy and Game Dynamics. „Mathematical Biosciences", Vol. 40, No. 1-2, 145-156.

Weibull J., 1995: Evolutionary Game Theory. MIT Press, Cambridge MA.

Zeeman E., 1981: Dynamics of the Evolution of Animal Conflicts. „Journal of Theoretical Biology", Vol. 89, No. 2, 249-270.

DELAYS IN SOCIAL MODELS OF REPLICATOR DYNAMICS

Summary

This paper discusses the impact of delayed information on the social models of replicator dynamic. Internal fixed point is here the optimal decision in the population. Examples and simulations confirm presented results. Here is described both the standard replicator dynamic and the modified model, which assumes that players imitate opponents taking strategies with a higher average payoff in the past. Moreover the model, in which each strategy has its own delay is presented.

Ewa Roszkowska
Uniwersytet w Białymstoku
Jakub Brzostowski
Politechnika Śląska

WYBRANE WŁASNOŚCI I ODMIANY PROCEDURY SAW W KONTEKŚCIE WSPOMAGANIA NEGOCJACJI*

Wstęp

Metody wielokryterialnej analizy problemu decyzyjnego dostarczają zestawu narzędzi wykorzystywanych do rozwiązywania problemów negocjacyjnych, w których występuje jednocześnie wiele konfliktowych kwestii stanowiących przedmiot rozmów [Hwang i Yoon 1981; Figueira i in. 2005; Salo i Hämäläinen 2012]. Dobór narzędzia jest uzależniony od struktury problemu negocjacyjnego, stopnia jego złożoności, zakresu i rodzaju dostępnej informacji, znajomości i prostoty obliczeniowej algorytmu oraz systemu preferencji negocjatora [Guitouni i Martel 1998]. Procedury wielokryterialnego wspomagania decyzji można podzielić na metody oparte na funkcji wartości/użyteczności (tzw. szkoła amerykańska) oraz modelu relacyjnym (tzw. szkoła europejska) [Greco i in 2001; Figueira i in. 2005]. W systemach wspomagania decyzji są wykorzystywane trzy główne rodzaje modeli agregacji preferencji: oparte na funkcji użyteczności (wartości), systemie relacyjnym (relacja przewyższania) oraz na zbiorze reguł decyzyjnych [Słowiński 2007].

W artykule dokonano przeglądu własności metody SAW (*Simple Additive Weighting*), jednej z najprostszych oraz najczęściej stosowanych technik wielokryterialnych, w kontekście wspomagania negocjacji dwustronnych. Metoda SAW oparta na funkcji wartości/użyteczności wykorzystującej agregację preferencji opartą na wartościach wariantów decyzyjnych jest wykorzystywana

* Praca została sfinansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/03/B/HS4/03857.

w wielu systemach wspomagania negocjacji, w tym w Inspire [Kersten i Noronha 1998], SmartSettle [Thiessen i Soberg 2003], NegoCalc [Wachowicz 2008]. W artykule omówiono zalety i ograniczenia proponowanego podejścia, a także przedstawiono propozycje modyfikacji klasycznego algorytmu z punktu widzenia możliwości jego zastosowania w procesie negocjacji. Funkcja przypisująca pakietom negocjacyjnych ocenę punktową (tzw. funkcja scoringowa) wyznaczona za pomocą zmodyfikowanej procedury SAW jest użytecznym narzędziem umożliwiającym m.in. liniowe uporządkowanie pakietów negocjacyjnych, szacowanie wartości potencjalnych ustępstw lub korzyści, programowanie strategii postępowania oraz analizę porozumień negocjacyjnych.

1. Algorytm klasycznej procedury SAW

Wyróżnia się dwa główne podzbiory metod rozwiązywania wielokryterialnych problemów decyzyjnych (MCDM), tzw. metody szkoły amerykańskiej oraz metody szkoły francuskiej [Greco i in 2001; Figueira i in. 2005]. Grupy tych metod różnią się sposobem traktowania kryteriów, których nie można bezpośrednio porównać, oraz w konsekwencji sposobem analizy oraz porównywania wariantów decyzyjnych. Metody szkoły amerykańskiej (np. SAW, TOPSIS, AHP) oparte na tzw. kryterium syntetycznym zakładają tworzenie funkcji agregującej wartości wariantów decyzyjnych ze względu na poszczególne kryteria. Metody szkoły francuskiej (np. rodzina metod ELECTRE) pozwalają na klasyfikację wariantów decyzyjnych z wykorzystaniem mechanizmu ustalania odpowiednich progów wzajemnych zależności między wartościami kryteriów.

Wielokryterialne metody porządkujące charakteryzują się tym, że zadany jest zbiór wariantów decyzyjnych, które należy uporządkować lub wybrać najlepszy z nich; zbiór kryteriów decyzyjnych; wektor współczynników wagowych przypisanych kryteriom oraz macierz decyzyjna zawierająca wartości badanych wariantów decyzyjnych. Na podstawie tych danych wyznacza się funkcję agregującą wartości przypisywane każdemu wariantowi decyzyjnemu. Następnie na podstawie kryterium syntetycznego jest tworzony ranking wariantów decyzyjnych. W zależności od systemu preferencji decydenta, jego oczekiwań oraz dostępnych danych funkcja agregująca przybiera różną postać. Dowolny wielokryterialny problem dyskretny można przedstawić w postaci tabeli 1.

Tabela 1

Reprezentacja dyskretnego problemu decyzyjnego

Warianty decyzyjne	Kryteria			
	C_1	C_2	…	C_n
A_1	x_{11}	x_{12}	…	x_{1n}
A_2	x_{21}	x_{22}	…	x_{2n}
…	…	…	…	…
A_m	x_{m1}	x_{m2}	…	x_{mn}
Współczynniki wagowe	w_1	w_2	…	w_n

gdzie:

- $\boldsymbol{A} = \{A_1, A_2, ..., A_m\}$ – zbiór wariantów decyzyjnych,
- $\boldsymbol{C} = \{C_1, C_2, ..., C_n\}$ – zbiór kryteriów, przy czym $\boldsymbol{C} = I \cup J$, gdzie I – zbiór kryteriów typu „zysk" (im więcej, tym lepiej), J – zbiór kryteriów typu „koszt" (im mniej, tym lepiej),
- $x_{ij} \in \Re$ – wartość *i*-tego wariantu decyzyjnego ze względu na *j*-te kryterium,
- $w_j \in \Re^+$ – współczynnik wagowy określający względną istotność *j*-tego kryterium.

Wartości wariantów decyzyjnych tworzą tzw. macierz decyzyjną $X = [x_{ij}]_{mxn}$, współczynniki wagowe wektor $w = [w_1, ..., w_n]$ spełniający warunek $w_1 + ... + w_n = 1$.

Do najczęściej stosowanych wielokryterialnych metod rankingowych zalicza się technikę *Simple Additive Weighting* (SAW), *Analytical Hierachy Process* (AHP) oraz *Technique for Order Preference by Similarity to Ideal Solution* (TOPSIS).

W celu ujednolicenia charakteru wartości poszczególnych kryteriów oraz umożliwienia porównań tych wartości dokonuje się normalizacji. Najczęściej wykorzystywane formuły to [Hwang, Yoon 1981; Wysocki 2010]:

$$z_{ij} = \begin{cases} \dfrac{x_{ij}}{\sqrt{\sum_{i=1}^{m} (x_{ij})^2}} & \text{dla } i \in I \\ 1 - \dfrac{x_{ij}}{\sqrt{\sum_{i=1}^{m} (x_{ij})^2}} & \text{dla } i \in J \end{cases} \qquad \text{(normalizacja wektorowa)} \qquad (1)$$

$$z_{ij} = \begin{cases} \dfrac{x_{ij} - \min\limits_{i} x_{ij}}{\max\limits_{i} x_{ij} - \min\limits_{i} x_{ij}} & \text{dla } i \in I \\ 1 - \dfrac{x_{ij} - \min\limits_{i} x_{ij}}{\max\limits_{i} x_{ij} - \min\limits_{i} x_{ij}} & \text{dla } i \in J \end{cases} \qquad \text{(normalizacja liniowa I typu)} \qquad (2)$$

$$z_{ij} = \begin{cases} \dfrac{x_{ij}}{\max\limits_{i} x_{ij}} & \text{dla } i \in I \\ 1 - \dfrac{x_{ij}}{\max\limits_{i} x_{ij}} & \text{dla } i \in J \end{cases} \qquad \text{(normalizacja liniowa II typu)} \qquad (3)$$

$$z_{ij} = \begin{cases} \dfrac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} & \text{dla } i \in I \\ 1 - \dfrac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} & \text{dla } i \in J \end{cases} \qquad \text{(normalizacja liniowa III typu)} \qquad (4)$$

dla $i = 1,2,\ldots,m$.

W klasycznej procedurze SAW funkcja agregująca S przypisuje wariantowi decyzyjnemu A_i kombinację liniową wektora wagowego oraz znormalizowanych wartości wariantu decyzyjnego zgodnie ze wzorem [np. Hwang i Yoon 1981]:

$$S(A_i) = \sum_{j=1}^{n} w_j \cdot z_{ij} \qquad (i = 1,2,\ldots,m) \qquad (5)$$

gdzie: z_{ij} – znormalizowana wartość *i*-tego wariantu decyzyjnego ze względu na *j*-te kryterium, $w_j \in \Re^+$ współczynnik wagowy *j*-tego kryterium ($j = 1,2,\ldots,n$).

Warianty decyzyjne ze zbioru *A* porządkuje się liniowo ze względu na wartość funkcji *S*, przy czym wyższe wartości $S(A_i)$ świadczą o wyższej pozycji w rankingu *i*-tego wariantu. Warunkiem poprawnej stosowalności procedury SAW jest założenie wzajemnej niezależności kryteriów (ze względu na relację preferencji między nimi). Zaletą algorytmu SAW jest jego prostota obliczeniowa, łatwość interpretacji uzyskanego wyniku. Wadą zaś zależność końcowego rankingu wariantów od przyjętej metody normalizacji, możliwość zmiany uporządkowania wariantów decyzyjnych w sytuacji usunięcia lub dołączenia nowego wariantu do rozważanego zbioru alternatyw, tzw. *rank reversal* [García-Cascales i Lamata 2012] (zob. też przykład obliczeniowy).

W klasycznym algorytmie SAW zakłada się, że warianty decyzyjne są reprezentowane przez liczby rzeczywiste. Przy werbalnym opisie kryteriów można zastosować odpowiednie ekwiwalenty numeryczne [np. Jadidi i in. 2008]. Przykładowe skale ocen opcji jakościowych zawiera tabela 2.

Tabela 2

Skalowanie opcji jakościowych

Ocena	Wartość
Odpowiednia (OD)	1
Dostateczna (DST)	3
Dobra (DB)	5
Bardzo dobra (BDB)	7
Wyróżniająca (W)	9
Wartości pośrednie między ocenami	2,4,6,8

Źródło: Na podstawie: Jadidi i in. [2008].

2. Problem negocjacyjny jako problem wielokryterialnego podejmowania decyzji

Założenia modelu negocjacyjnego. Przyjęto założenie, że wariantem decyzyjnym jest pakiet negocjacyjny, który negocjator może przedstawić jako ofertę lub otrzymać od oponenta, kryterium wariantu decyzyjnego – zagadnienie negocjacyjne, wartością kryterium – opcja zagadnienia negocjacyjnego [por. Roszkowska 2012; Roszkowska i Wachowicz 2012].

Niech $\boldsymbol{Z} = \{Z_1,Z_2,\ldots,Z_n\}$ oznacza zbiór zagadnień negocjacyjnych, $Z = I \cup J$, gdzie *I* jest zbiorem zagadnień typ „zysk" (im więcej, tym lepiej), *J* zbiorem zagadnień typu „koszt" (im mniej, tym lepiej). Każdemu zagadnieniu negocjacyj-

nemu Z_j zostaje przyporządkowana dodatnia waga określająca jego względną istotność, przy czym $w_1 + \ldots + w_n = 1$. Przegląd metod wyznaczania współczynników wagowych zawierają m.in. prace: Barron i Barrett [1996b], Tzeng i in. [1998], Belton i Stewart [2002]. Do wyznaczenia współczynników wagowych zagadnień negocjacyjnych można wykorzystać np. metody punktowe, metodę AHP [Saaty 1980], procedurę Simosa [Simos 1990], funkcje rangujące [Stillwell i in. 1981; Barron i Barrett 1996a, 1996b; Solymosi i Dompi 1985].

Niech dalej:

- x_j^+ poziom aspiracji, czyli najlepsza z możliwych do zaakceptowania opcja ze względu na *j*-te zagadnienie negocjacyjne,
- x_j^- poziom rezerwacji, czyli najgorsza z możliwych do zaakceptowania opcja ze względu na *j*-te zagadnienie negocjacyjne.

Wartości x_j^+, x_j^- reprezentują maksymalną granicę żądań oraz minimalną granicę ustępstw ze względu na *j*-te zagadnienie. Przez D_j oznaczmy dziedzinę, czyli zakres możliwych opcji *j*-tego zagadnienia negocjacyjnego ($j = 1,2,\ldots,n$). Zachodzi przy tym:

$$D_j \subseteq < x_j^-; x_j^+ > \subset \Re^+ \quad \text{gdy} \quad j \in I$$

oraz:

$$D_j \subseteq < x_j^+; x_j^- > \subset \Re^+ \quad \text{gdy} \quad j \in J$$

Dowolny pakiet negocjacyjny jest reprezentowany przez wektor w postaci $P = [x_1,\ldots, x_n]$, gdzie $x_j \in D_j$ ($j = 1,2,\ldots,n$).

Negocjacje to złożony proces interakcji między co najmniej dwiema stronami, który polega na wymianie ofert, czynieniu ustępstw, a jest zorientowany na podjęcie wspólnej decyzji umożliwiającej realizację interesów wszystkim stronom [Roszkowska 2011]. Negocjacje kończą się, gdy strony osiągną kompromis lub jedna ze stron zerwie rozmowy.

System punktowy. W celu określenia strategii postępowania negocjator tworzy tzw. system punktowy (scoringowy) przypisujący pakietom negocjacyjnym oceny punktowe, które są podstawą uporządkowania ofert od najlepszej do najgorszej. Negocjator może początkowo wyselekcjonować ograniczoną liczbę pakietów $\boldsymbol{P} = \{P_1, P_2,\ldots, P_m\}$ do oceny w ten sposób, aby zorientować się w zbiorze wszystkich rozwiązań. W trakcie negocjacji negocjator powinien mieć sposobność uogólnienia oceny punktowej na wszystkie pozostałe pakiety, które nie zostały wybrane wstępnie do oceny. Pożądaną własnością systemu ocen punkto-

wych jest stabilność polegająca na tym, że modyfikacja wyjściowego zbioru pakietów ***P*** (tzn. usunięcie pakietu ze zbioru ***P*** lub dodanie nowego pakietu do tego zbioru) nie zmienia ocen punktowych ani uporządkowania pozostałych pakietów.

W dalszej części artykułu dokonano analizy możliwości zastosowania procedury SAW do budowy systemu ocen pakietów negocjacyjnych. Warunkiem koniecznym przy konstrukcji dowolnego kryterium syntetycznego jest normalizacja zmiennych*. Przy wyborze formuły normalizacyjnej należy uwzględnić charakter sytuacji negocjacyjnej, rodzaj skali pomiaru zmiennych opisujących zagadnienia negocjacyjne, własności zmiennych znormalizowanych. Na proces normalizacji można nałożyć warunki związane np. z ujednoliceniem charakteru rozważanych zagadnień, stałość poziomu zmienności, zachowanie poziomu asymetrii dla zmiennej wyjściowej oraz przekształconej.

Zauważmy, że normalizacja określona wzorem (1) oraz (4) prowadzi do niestabilnego systemu ocen pakietów negocjacyjnych, gdyż usunięcie lub dołączenie nowego pakietu wiąże się z koniecznością ponownego wyznaczenia ocen wszystkich pakietów (por. przykład obliczeniowy). W celu otrzymania stabilnego systemu scoringowego proponuje się modyfikację formuły (2) lub (3). Modyfikacja ta polega na włączeniu do systemu scoringowego dwóch dodatkowych pakietów, tzw: idealnego $P_I = [x_1^+, x_2^+, ..., x_n^+]$ oraz antyidealnego $P_{AI} = [x_1^-, x_2^-, ..., x_n^-]$ wyznaczonych na podstawie poziomów aspiracji i rezerwacji zagadnień negocjacyjnych. Pakiety P_I oraz P_{AI} pełnią rolę stabilnych punktów referencyjnych, czyli wzorca i antywzorca.

Niech $P = [x_1, x_2, \ldots, x_n]$ będzie dowolnym pakietem negocjacyjnym, gdzie $x_j \in D_j$. Zmodyfikowane formuły normalizacyjne są określone następująco:

$$z_j^1 = \left(\frac{x_j - x_j^-}{x_j^+ - x_j^-} \right)^r \qquad \text{(zmodyfikowana normalizacja I typu)} \qquad (6)$$

$$z_j^2 = \begin{cases} \left(\dfrac{x_j}{x_j^+} \right)^r & \text{dla } i \in I \\ 1 - \left(\dfrac{x_j}{x_j^-} \right)^r & \text{dla } i \in J \end{cases} \qquad \text{(zmodyfikowana normalizacja II typu)} \qquad (7)$$

* Szerzej o własnościach formuł normalizacyjnych, zagadnieniu doboru metody normalizacyjnej w zależności od skal pomiaru zmiennych traktują np. prace [Kukuła 2000; Walesiak 2002].

$r \in \Re^+$ – parametr. Parametr r pozwala dopasować postać formuły normalizacyjnej do systemu preferencji decydenta, dla $r = 1$ mamy normalizację liniową.

Funkcje scoringowe. Funkcje przypisujące pakietowi negocjacyjnemu P ocenę punktową bazującą na procedurze SAW oraz formułach (6) i (7) mają postać:

$$S_1^{\text{mod}}(P) = S_1^{\text{mod}}(x_1,...,x_n) = \sum_{j=1}^{n} w_j \left(\frac{x_j - x_j^-}{x_j^+ - x_j^-} \right)^{r_j} \tag{8}$$

$$S_2^{\text{mod}}(P) = S_2^{\text{mod}}(x_1,...,x_n) = \sum_{j \in I} w_j \left(\frac{x_j}{x_j^+} \right)^{r_j} + \sum_{j \in J} w_j \left(1 - \frac{x_j}{x_j^-} \right)^{r_j} \tag{9}$$

gdzie w_j – waga j-tego zagadnienia negocjacyjnego, x_j – wartość opcji pakietu P ze względu na j-te zagadnienie negocjacyjne, x_j^+ – wartość opcji pakietu P_I ze względu na j-te zagadnienie negocjacyjne, x_j^- – wartość opcji pakietu P_{AI} ze względu na j-te zagadnienie negocjacyjne, $r_j \in \Re^+$ – parametr ($j = 1,2,\ldots,n$).

Normalizacja oparta na pakietach P_I oraz P_{AI} gwarantuje stabilność systemu scoringowego. W sytuacji gdy nowy pakiet $P = [x_1, x_2\ldots, x_n]$ nie zmienia granicy żądań i ustępstw ze względu na zagadnienia negocjacyjne, tzn. $x_j \in D_j$, wystarczy tylko wyznaczyć ocenę punktową tego pakietu zgodnie ze wzorem (8) lub (9) oraz dołączyć do zbioru ocen pozostałych pakietów. Dla dowolnego pakietu P mamy:

$$S_k^{\text{mod}}(P_{AI}) \le S_k^{\text{mod}}(P) \le S_k^{\text{mod}}(P_I) \ (k = 1,2) \tag{10}$$

Ponadto można pokazać, że:

$$S_1^{\text{mod}}(P_I) = 1 \tag{11}$$

oraz:

$$S_1^{\text{mod}}(P_{AI}) = 0 \tag{12}$$

$$S_2^{\text{mod}}(P_I) = \sum_{j \in I} w_j + \sum_{j \in J} w_j \left(1 - \frac{x_j^+}{x_j^-} \right)^{r_j} \le 1 \tag{13}$$

$$S_2^{\mathrm{mod}}(P_{AI}) = \sum_{j \in I} w_j \left(\frac{x_j^-}{x_j^+} \right)^{r_j} \geq 0 \tag{14}$$

W przypadku normalizacji (6) rozstęp między oceną punktową pakietu idealnego oraz antyidealnego jest zawsze stały i wynosi 1, stąd ocena punktowa pakietu P otrzymana na podstawie wzoru (8) pozwala bezpośrednio ocenić jego wartość w relacji do pakietów P_I oraz P_{AI}. Przy normalizacji typu (7) rozstęp między ocenami punktowymi pakietów P_1 oraz P_{AI} jest zmienny i zależny od problemu decyzyjnego, o czym należy pamiętać przy interpretacji oceny punktowej pakietów.

Szacowanie ustępstw/korzyści. Funkcje (8) oraz (9) zakładają możliwość kompensacji między kryteriami, tzn. poczynione ustępstwo w ramach jednego zagadnienia może być zrekompensowane korzyścią uzyskaną na innym zagadnieniu. Niech $P_0 = [x_1,\ldots,x_n]$, $P_1 = [x_1 + \Delta x_1,\ldots,x_n + \Delta x_n]$ to pakiety negocjacyjne, gdzie $x_j, x_j + \Delta x_j \in D_j$. Różnica $\Delta S_k^{\mathrm{mod}}{}_{1/0} = S_k^{\mathrm{mod}}(P_1) - S_k^{\mathrm{mod}}(P_0)$ ($k = 1,2$) ocen punktowych pakietów P_1 oraz P_0 stanowi miarę ustępstwa/korzyści lub miarę kompensacji między kryteriami w przypadku zmiany oferty z P_0 na P_1. Można pokazać, że dla $r = 1$ zachodzi:

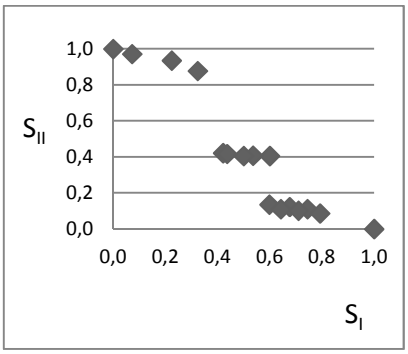
$$\Delta S_1^{\mathrm{mod}}{}_{1/0} = \sum_{j=1}^{n} w_j \frac{\Delta x_j}{x_j^+ - x_j^-} \tag{15}$$

$$\Delta S_2^{\mathrm{mod}}{}_{1/0} = \sum_{j \in I}^{n} w_j \frac{\Delta x_j}{x_j^+} - \sum_{j \in I}^{n} w_j \frac{\Delta x_j}{x_j^-} \tag{16}$$

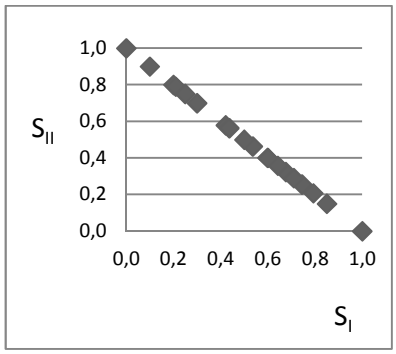
Wzory (15)-(16) pozwalają w prosty sposób wyznaczyć różnicę ocen punktowych między ofertami, czyli oszacować wartość ustępstwa $\left(\Delta S_k^{\mathrm{mod}}{}_{1/0} < 0\right)$, korzyści $\left(\Delta S_k^{\mathrm{mod}}{}_{1/0} > 0\right)$ lub wyznaczyć pakiet alternatywny $\left(\Delta S_k^{\mathrm{mod}}{}_{1/0} = 0\right)$.

Analiza porozumienia. Przyjmijmy, że dowolne porozumienie negocjacyjne jest reprezentowane przez punkt $(S_{\mathrm{I}}(\mathrm{P_k}), S_{\mathrm{II}}(\mathrm{P}_k)) \subseteq \Re^2$, gdzie $S_i(\mathrm{P}_k)$ oznacza wartość miernika oceny pakietu dla i-tej strony negocjacji (I = I,II) uzyskanego np. zmodyfikowaną metodą SAW. Strony negocjacji mogą dokonać analizy potencjalnych rozwiązań z uwzględnieniem zarówno własnych interesów, jak i interesów drugiej strony, wyznaczyć rozwiązania kompromisowe, stosując procedury przetargowe (np. rozwiązanie arbitrażowe Nasha), czy poszukiwać rozwiązań Pareto-optymalnych.

Ze względu na nastawienie do sytuacji negocjacyjnej oraz charakter zależności pomiędzy interesami stron wyróżnia się dwa rodzaje negocjacji: pozycyjne oraz integracyjne [Roszkowska 2011]. Negocjacje pozycyjne są związane z sytuacją, gdy zysk jednej ze stron odpowiada stracie drugiej, a każda ze stron pragnie maksymalizować wynik negocjacji. Negocjacje integracyjne odpowiadają sytuacji, gdy interesy stron są częściowo sprzeczne, częściowo zgodne i polegają na poszukiwaniu rozwiązania, które zaspokajałoby interesy obu stron. Negocjacje pozycyjne spełniają więc warunek $S_{\mathrm{I}}(\mathrm{P_k}) + \mathrm{S_{II}}(\mathrm{P}_k) = a$, gdzie a – stała dla dowolnego pakietu P_k, w przeciwnym wypadku są to negocjacje integracyjne (rys. 1 oraz 2). W sytuacji negocjacji integracyjnych strony mogą poszukiwać rozwiązania Pareto-optymalnego.



Rys. 1. Analiza negocjacyjnego porozumienia – negocjacje integracyjne



Rys. 2. Analiza negocjacyjnego porozumienia – negocjacje pozycyjne

3. Uogólnienia procedury SAW

Do budowy systemu ocen pakietów negocjacyjnych można wykorzystać metodę bazującą na koncepcji pomiaru odległości między pakietami negocjacyjnymi [Hwang i Yoon 1981; Figueira i in. 2005]. W pierwszym wariancie tej metody decydent dokonuje porównania ofert z pakietem idealnym, wyznaczając odległość każdego ocenianego pakietu P od pakietu idealnego P_{I} zgodnie ze wzorem[*]:

[*] Wzór na odległość Minkowskiego, gdzie parametr p jest przyjmowany przez decydenta. Dla $p = 1$ mamy odległość liniową (metryka miejska), dla $p = 2$ odległość euklidesową. Wraz ze wzrostem p zagadnienia przyjmujące wyższe wartości i wariancje nabierają większego znaczenia i wykazują tendencje do dominacji nad innymi kryteriami. W przypadku metryki miejskiej (p = 1) wartości odstające przyjmowane przez kryteria mają mniejszy wpływ na wartość odległości niż w przypadku odległości euklidesowej (p = 2) [por. Wysocki 2010, s. 64; Walesiak 2002].

$$D_k^p(P,P_I)=\left(\sum_{j=1}^{n}\left(w_j z_{jI}^k - w_j z_j^k\right)^p\right)^{\frac{1}{p}} \tag{17}$$

gdzie: w_j – waga j-tego zagadnienia negocjacyjnego, z_j^k – znormalizowana wartość opcji pakietu negocjacyjnego P ze względu na j-te kryterium, z_{jI}^k – znormalizowana wartość opcji pakietu negocjacyjnego P_I ze względu na j-te kryterium, $p \geq 0$ – parametr (j = 1,2,…,n). Dla k = 1 przyjmujemy normalizację określoną wzorem (6), dla k = 2 wzorem (7). W szczególności mamy:

$$D_1^p(P,P_I)=\left(\sum_{j=1}^{n}\left(w_j \frac{x_j^+ - x_j}{x_j^+ - x_j^-}\right)^p\right)^{\frac{1}{p}} \tag{18}$$

$$D_2^p(P,P_I)=\left(\sum_{j\in I}\left(w_j \frac{x_j^+ - x_j}{x_j^+}\right)^p + \sum_{j\in J}\left(w_j \frac{x_j - x_j^+}{x_j^-}\right)^p\right)^{\frac{1}{p}} \tag{19}$$

Następnie dokonuje się uporządkowania pakietów negocjacyjnych według rosnącej odległości do pakietu idealnego. Im bliższą odległość ma pakiet P od pakietu idealnego, tym jest lepszy.

W drugim wariancie negocjator dokonuje porównania oferty P z pakietem antyidealnym, wyznaczając odległość ocenianego pakietu P od pakietu P_{AI} zgodnie ze wzorem:

$$d_k^p(P,P_{AI})=\left(\sum_{j=1}^{n}\left(w_j z_j^k - w_j z_{jAI}^k\right)^p\right)^{\frac{1}{p}} \tag{20}$$

gdzie: w_j – waga j-tego zagadnienia negocjacyjnego, z_j^k – znormalizowana wartość pakietu negocjacyjnego P ze względu na j-te kryterium, z_{jAI}^k – znormalizowana wartość pakietu negocjacyjnego P_I ze względu na j-te kryterium, $p \geq 0$ – parametr (j = 1,2,…,n). Dla k = 1 przyjmujemy normalizację określoną wzorem (6), dla k = 2 wzorem (7). W szczególności zachodzi:

$$d_1^p(P,P_{AI}) = \left(\sum_{j=1}^{n} \left(w_j \frac{x_j - x_j^-}{x_j^+ - x_j^-} \right)^p \right)^{\frac{1}{p}} \tag{21}$$

$$d_2^p(P,P_{AI}) = \left(\sum_{j\in I} \left(w_j \frac{x_j - x_j^-}{x_j^+} \right)^p + \sum_{j\in J} \left(w_j \frac{x_j^- - x_j}{x_j^-} \right)^p \right)^{\frac{1}{p}} \tag{22}$$

Następnie dokonuje się uporządkowania liniowego pakietów według malejącej odległości od pakietu antyidealnego. Im bliższą odległość ma pakiet *P* od pakietu antyidealnego, tym jest gorszy.

Trzecią propozycją tworzenia systemu ocen pakietów jest metoda TOPSIS, gdzie negocjator wyznacza odległość każdego ocenianego pakietu *P* od pakietu idealnego oraz antyidealnego, następnie wyznacza ocenę punktową pakietu negocjacyjnego *P* zgodnie ze wzorem [Hwang i Yong 1981]:

$$T_k^p(P) = \frac{d_k^p(P,P_{AI})}{d_k^p(P,P_{AI}) + D_k^p(P,P_I)} \tag{23}$$

Zachodzi przy tym $0 \le T_k^p(P) \le 1$. Wyższe wartości miernika $T_k^p(P)$ świadczą o wyższej pozycji w rankingu tego pakietu negocjacyjnego. Według procedury TOPSIS najlepsza oferta negocjacyjna to taka, która posiada najkrótszą odległość od pakietu idealnego (aspiracji), a zarazem najdalszą od pakietu antyidealnego (rezerwacji).

W polskiej literaturze ekonomicznej Hellwig [1968] wprowadził kryterium syntetyczne, tzw. taksonomiczną miarę rozwoju opartą na koncepcji pomiaru odległości od wzorca. Zmodyfikowana funkcja scoringowa oparta na kryterium syntetycznym Hellwiga przyjmuje postać:

$$Hm_k^p(P) = 1 - \frac{D_k^p(P,P_I)}{D_k^p(P_{AI},P_I)} = 1 - \frac{D_k^p(P,P_I)}{d_k^p(P_{AI},P_I)} \tag{24}$$

Można łatwo pokazać, że metody bazujące na koncepcji odległości między pakietami są rozszerzeniem procedury SAW. Istotnie, dla $k = 1$ oraz $p = 1$ zachodzi:

$$d_1^1(P,P_I) = S_{\text{mod}}^1(P) \tag{25}$$

$$D_1^1(P, P_I) = 1 - S_{\text{mod}}^1(P) \tag{26}$$

$$T_1^1(P) = S_{\text{mod}}^1(P) \tag{27}$$

$$Hm_1^1(P) = S_{\text{mod}}^1(P) \tag{28}$$

W przypadku metod opartych na odległości między pakietami zarówno metoda normalizacji, jak i sposób mierzenia odległości ma wpływ na końcowy ranking (por. przykład obliczeniowy).

Warto również zaznaczyć, że metryka Minkowskiego jest najbardziej znaną i najczęściej stosowaną metodą pomiaru odległości. W sytuacji gdy zmienne opisujące kryteria wykazują silne związki korelacyjne, do obliczenia odległości można zastosować np. wzór Mahalanobisa, a gdy zmienne opisujące kryteria są mierzone na skalach różnych rodzajów, użyteczna jest uogólniona miara odległości (GDM) Walesiaka [2002].

4. Przykład obliczeniowy

Prezentowany przykład, oparty na danych umownych, pokazuje możliwości zastosowania procedury SAW do oceny pakietów negocjacyjnych. Zakładamy, że negocjator dokonuje oceny pakietów negocjacyjnych ze względu na trzy zagadnienia typu zysk: Z_1, Z_2, Z_3. Obszary negocjacji dla zagadnień są określone następująco: <5,20> dla Z_1; <1,5> dla Z_2 oraz <20,35 > dla Z_3. Wektor wag ma postać w = [0,4; 0,3;0,3]. Na etapie wstępnym wybrano do oceny 4 pakiety. Zestawienie pakietów wraz z poziomem aspiracji, rezerwacji oraz wektorem wag przedstawia tabela 3.

Tabela 3

Problem decyzyjny negocjatora

Pakiety negocjacyjne	Zagadnienia		
	Z_1	Z_2	Z_3
P_1	15	2	21
P_2	12	3	20
P_3	8	4	25
P_4	6	4	35
Poziom rezerwacji	5	1	20
Poziom aspiracji	20	5	35
Waga	**0,4**	**0,3**	**0,3**

Źródło: Opracowanie własne.

Zauważmy, że wszystkie opcje kwestii negocjacyjnych są zawarte w obszarze negocjacji. Pakiet P_1 przyjmuje ze względu na zagadnienie Z_1 wartość najwyższą, Z_2 – najniższą, Z_3 – bliską poziomowi rezerwacji (chociaż nie najniższą). Pakiet P_2 przyjmuje wartości pośrednie ze względu na wszystkie zagadnienia negocjacyjne. Pakiet P_4 jest najlepszy ze względu na zagadnienie Z_2, opcje pozostałych zagadnień przyjmują wartości średnie, przy czym w porównaniu z pakietem P_2 wyższą wartość dla Z_3, a niższą dla Z_1. Pakiet P_4 przyjmuje najwyższą wartość ze względu na zagadnienia Z_1 oraz Z_2, natomiast najniższą na najważniejsze dla niego zagadnienie Z_1.

W dalszej części dokonano zestawienia oraz zaprezentowano skróconą analizę ocen punktowych pakietów oraz ich ranking otrzymanych metodą SAW (klasyczną i zmodyfikowaną) oraz metodami opartymi na pomiarze odległości między pakietami z wykorzystaniem różnych formuł normalizacyjnych oraz miar odległości.

Wartości klasycznej funkcji $S(P_i)$ otrzymane za pomocą różnych formuł normalizacyjnych oraz ranking pakietów negocjacyjnych zawiera tabela 4.

Tabela 4

System ocen pakietów negocjacyjnych uzyskanych klasyczną metodą SAW dla różnych metod normalizacji oraz zbiorów pakietów

Pakiet	Normalizacja wektorowa		Normalizacja liniowa I typu		Normalizacja liniowa II typu		Normalizacja liniowa III typu	
P_1	0,488**(2)**	-	0,420**(3)**	-	0,730**(2)**	-	0,2549**(1)**	-
P_2	0,472**(3)**	0,574**(1)**	0,417**(4)**	0,400**(3)**	0,716**(4)**	0,796**(2)**	0,2457**(3)**	0,341**(1)**
P_3	0,471**(4)**	0,550**(3)**	0,489**(2)**	0,533**(2)**	0,728**(3)**	0,781**(3)**	0,2446**(4)**	0,326**(3)**
P_4	0,492**(1)**	0,562**(2)**	0,600**(1)**	0,600**(1)**	0,760**(1)**	0,800**(1)**	0,2548**(2)**	0,333**(2)**

Źródło: Opracowanie własne na podstawie danych tabeli 3.

Formuła normalizacyjna przyjęta w procedurze SAW ma wpływ na końcowy ranking pakietów negocjacyjnych. Przykładowo pakiet P_1 został oceniony jako najlepszy, prawie tak samo, jak pakiet P_4 (wzór (4)), drugi (wzór (1) oraz (3)), trzeci (wzór (2)). Odrzucenie pakietu P_1 spowodowało nie tylko zmianę wartości punktowej ocen pozostałych pakietów, ale także w przypadku normalizacji wektorowej oraz liniowej II typu zmianę ich uporządkowania (*rank reversal*).

System ocen pakietów negocjacyjnych uzyskanych różnymi metodami z wykorzystaniem zmodyfikowanej formuły normalizacyjnej I typu przedstawia tabela 5.

Tabela 5

System ocen pakietów negocjacyjnych uzyskanych różnymi metodami z zastosowaniem zmodyfikowanej liniowej formuły normalizacyjnej I typu

Pakiet	$S_1^{\mathrm{mod}}(P_i)$	$D_1^1(P_i)$	$D_1^2(P_i)$	$d_1^1(P_i)$	$d_1^2(P_i)$	$T_1^1(P_i)$	$T_1^2(P_i)$
P_1	0,362**(3)**	0,638**(3)**	0,383**(2)**	0,362**(3)**	0,278**(2)**	0,362**(3)**	0,420**(2)**
P_2	0,337**(4)**	0,663**(4)**	0,398**(4)**	0,337**(4)**	0,239**(4)**	0,337**(4)**	0,376**(4)**
P_3	0,405**(2)**	0,595**(2)**	0,385**(3)**	0,405**(2)**	0,259**(3)**	0,405**(2)**	0,402**(3)**
P_4	0,552**(1)**	0,448**(1)**	0,381**(1)**	0,552**(1)**	0,376**(1)**	0,552**(1)**	0,497**(1)**
$\boldsymbol{P_I}$	**1,000**	**0,000**	**0,000**	**1,000**	**0,583**	**1,000**	**1,000**
$\boldsymbol{P_{AI}}$	**0,000**	**1,000**	**0,583**	**0,000**	**0,000**	**0,000**	**0,000**

Źródło: Opracowanie własne na podstawie danych tabeli 2.

Przyjęcie zmodyfikowanej formuły normalizacyjnej zapewniło stabilność systemu oceny ofert, tzn. odrzucenie pakietu P_1 nie zmienia ani wartości punktowej, ani uporządkowania pozostałych pakietów. Niemniej jednak można zauważyć, że system oceny ofert jest wrażliwy na przyjętą miarę odległości. Dla $p = 1$ pakiet P_4 oceniono jako najlepszy, a pakiet P_2 – najgorszy we wszystkich rankingach. Pakiety P_1 oraz P_3 zajmują pozycję drugą lub trzecią w zależności od przyjętej metody pomiaru odległości między pakietami. Ponadto $S_1^{\mathrm{mod}}(P_i) = d_1^1(P_i) = 1 - D_1^1(P_i) = T_1^1(P_i)$. Różnica ocen punktowych dwóch pakietów stanowi miarę ustępstwa lub korzyści negocjatora. Można zauważyć, że przy stosowaniu systemów scoringowych opartych na odległości euklidesowej różnice punktowe między pakietami negocjacyjnymi są dużo mniejsze.

System ocen pakietów negocjacyjnych uzyskanych różnymi metodami z wykorzystaniem zmodyfikowanej formuły normalizacyjnej typu II przedstawia tabela 6.

Tabela 6

System ocen pakietów negocjacyjnych uzyskanych różnymi metodami z zastosowaniem zmodyfikowanej liniowej formuły normalizacyjnej II typu

Pakiet	$S_2^{\mathrm{mod}}(P_i)$	$D_2^1(P_i)$	$D_2^2(P_i)$	$d_2^1(P_i)$	$d_2^2(P_i)$	$T_2^1(P_i)$	$T_2^2(P_i)$
P_1	0,600**(3)**	0,400**(3)**	0,2383**(2)**	0,269**(3)**	0,209**(2)**	0,402**(3)**	0,46720**(1)**
P_2	0,591**(4)**	0,409**(4)**	0,2378**(1)**	0,260**(4)**	0,184**(4)**	0,389**(4)**	0,43679**(3)**
P_3	0,614**(2)**	0,386**(2)**	0,2618**(3)**	0,283**(2)**	0,195**(3)**	0,423**(2)**	0,42626**(4)**
P_4	0,660**(1)**	0,340**(1)**	0,2864**(4)**	0,329**(1)**	0,222**(1)**	0,491**(1)**	0,43682**(2)**
$\boldsymbol{P_I}$	**1,000**	**0,000**	**0,0000**	**0,669**	**0,405**	**1,000**	**1,0000**
$\boldsymbol{P_{AI}}$	**0,331**	**0,669**	**0,4051**	**0,000**	**0,000**	**0,000**	**0,0000**

Źródło: Opracowanie własne na podstawie danych tabeli 2.

Zauważmy, że w analizowanym przykładzie ranking pakietów negocjacyjnych uzyskanych zmodyfikowaną metodą SAW nie zależy od przyjętej formuły normalizacyjnej, chociaż pakiety przyjmują różne wartości ocen punktowych. Na uporządkowanie pakietów ma wpływ sposób pomiaru odległości. Dla $p = 1$ uporządkowanie jest zgodne z rankingiem uzyskanym zmodyfikowaną metodą SAW, dla $p = 2$ zaobserwowano brak zgodności. Przykładowo pakiet P_2 oceniono najwyżej, pakiet P_4 najniżej z zastosowaniem systemu scoringowego opartego na funkcji $D_2^2(P_i)$. W przypadku pozostałych metod pakiet P_2 oceniono najniżej lub na trzeciej pozycji, pakiet P_4 oceniono najwyżej lub na drugiej pozycji.

Przykład pokazuje, że uporządkowanie pakietów negocjacyjnych jest zależne od przyjętej formuły normalizacyjnej, metody mierzenia odległości czy funkcji agregującej. Stąd w praktyce wybór ten powinien być przemyślany i dokonany z uwzględnieniem zarówno własności poszczególnych formuł, jak i preferencji decydenta [Wysocki 2010; Walesiak 2006]. Można także dodatkowo dokonać analizy wrażliwości uporządkowania pakietów ze względu na modyfikację wag zagadnień negocjacyjnych. Wyznaczenie krytycznych kryteriów oraz minimalnych modyfikacji wag, które skutkują zmianą uporządkowania pakietów negocjacyjnych, pozwala ocenić stabilność systemu ocen punktowych [Triantaphyllou i in. 2000].

W przypadku zmodyfikowanej procedury SAW ocena punktowa dowolnego pakietu negocjacyjnego P jest sumą ocen punktowych ze względu na poszczególne kryteria:

$$S_k^{\text{mod}}(P) = \sum_{j=1}^{n} S_{kj}^{\text{mod}}(x_j) \qquad (k = 1,2) \tag{29}$$

gdzie $S_{kj}^{\text{mod}}(x_j)$ ocena punktowa pakietu ze względu na j-te kryterium ($j = 1,2,\ldots,n$).

Taki rozkład pozwala ocenić wkład poszczególnych opcji w ocenę końcową pakietu oraz zaplanować strategię negocjacyjną, uwzględniając nie tylko ocenę punktową całego pakietu, ale także wartości progowe ocen ze względu na poszczególne zagadnienia. Rozkład oceny punktowej pakietu ze względu na zagadnienia dla zmodyfikowanych funkcji SAW przedstawia tabela 7.

Tabela 7

Rozkład oceny punktowej pakietów negocjacyjnych ze względu na poszczególne zagadnienia dla zmodyfikowanych funkcji SAW

Pakiet	$S_{11}^{\mathrm{mod}}(x_1)$	$S_{12}^{\mathrm{mod}}(x_2)$	$S_{13}^{\mathrm{mod}}(x_3$	$S_1^{\mathrm{mod}}(P_i)$	$S_{21}^{\mathrm{mod}}(x_1)$	$S_{22}^{\mathrm{mod}}(x_2)$	$S_{23}^{\mathrm{mod}}(x_3)$	$S_2^{\mathrm{mod}}(P_i)$
P_1	0,267	0,075	0,020	**0,362**	0,300	0,120	0,180	**0,600**
P_2	0,187	0,150	0,000	**0,337**	0,240	0,180	0,171	**0,591**
P_3	0,080	0,225	0,100	**0,405**	0,160	0,240	0,214	**0,614**
P_4	0,027	0,225	0,300	**0,552**	0,120	0,240	0,300	**0,660**
$\boldsymbol{P_I}$	**0,400**	**0,300**	**0,300**	**1,000**	**0,400**	**0,300**	**0,300**	**1,000**
$\boldsymbol{P_{AI}}$	**0,000**	**0,000**	**0,000**	**0,000**	**0,100**	**0,060**	**0,171**	**0,331**

Źródło: Opracowanie własne na podstawie danych tabeli 2.

Podsumowanie

W artykule zaprezentowano wiele prostych sposobów tworzenia systemów scoringowych wykorzystujących procedurę SAW oraz metody oparte na pomiarze odległości między pakietami. Metody te znajdują zastosowanie w sytuacjach, gdy zagadnienia negocjacyjne są opisane wartościami precyzyjnymi czy też w postaci słów, a wybór odpowiedniej metody zależy od potrzeb i preferencji negocjatora.

Modyfikacja klasycznej procedury SAW oraz metod bazujących na pomiarze odległości poprzez wykorzystanie formuł normalizacyjnych opartych na pakiecie idealnym oraz antyidealnym ułatwia ocenę nowych pakietów negocjacyjnych, zachowując uporządkowanie oraz ocenę punktową pakietów wybranych wstępnie do oceny. Należy jednak pamiętać, że zarówno metoda normalizacji, jak i miara odległości może mieć istoty wpływ na końcowe uporządkowanie pakietów negocjacyjnych. Stąd wybór ten powinien być dokonany z uwzględnieniem własności poszczególnych formuł oraz preferencji decydenta.

Bibliografia

Barron F.H, Barrett B.E., 1996a: The Efficacy of SMARTER: Simple Multi-attribute Rating Technique Extended to Ranking. „Acta Psychologica", 93(1-3), s. 23-36.

Barron F.H., Barrett B.E., 1996b: Decision Quality Using Ranked Attribute Weights. „Management Science", 42, 1515-1523.

Belton V., Stewart T.J., 2002: Multiple Criteria Decision Analysis: An Integrated Approch. Kluwer, Dordrecht.

Figueira J., Greco S., Ehrgott M., 2005: Multiplie Criteria Decision Analysis: State of the Art Surveys. Springer, New York.

García-Cascales S.M., Lamata M.T., 2012: On Rank Reversal and TOPSIS Method. „Mathematical and Computer Modelling", 56, s. 123-132.

Greco S., Matarazzo B., Słowiński R., 2001: Rough Sets Theory for Multicriteria Decision Analysis. „European Journal of Operational Research", s. 129.

Guitouni A., Martel J.M., 1998: Tentative Guidelines to Help Choosing an Appropriate MCDA Method. „European Journal of Operational Research", 109, s. 501-521.

Hwang C.L., Yoon K., 1981: Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag, Berlin.

Jadidi O., Hong T.S., Firouzi F., Yusuff R.M., 2008: An Optimal Grey Based Approach Based on TOPSIS Concept for Supplier Selection Problem. „International Journal of Management Science and Engineering Management", Vol. 4, No. 2, s. 104-117.

Kersten G.E., Noronha S.J., 1999: WWW-based Negotiation Support: Design, Implementation and Use. „Decision Support Systems", 25(2), s. 135-154.

Kukuła K., 2000: Metoda unitaryzacji zerowanej. Wydawnictwo Naukowe PWN, Warszawa.

Roszkowska E., 2011: Wybrane modele negocjacji. Wydawnictwo UwB, Białystok.

Roszkowska E., 2012: Zastosowanie metody TOPSIS do wspomagania procesu negocjacji. W: Taksonomia 19. Klasyfikacja i analiza danych – teoria i zastosowania. K. Jajuga, M. Walesiak (red.). Wydawnictwo UE, Wrocław, s. 68-75.

Roszkowska E., Wachowicz T., 2012: Negotiation Support with Fuzzy TOPSIS. W: Group Decision and Negotiation, Proceedings. A.T. Almeida, D.C. Morais, S.F. Daher (eds.). Recife, Brasil, s. 161-175.

Salo A., Hämäläinen R.P., 2012: Multicriteria Decision Analysis in Group Decision Processes. W: Handbook of Group Decision and Negotiation. D.M. Kilgour, C. Eden (eds.). Springer, Dordrecht, s. 269-284.

Saaty T.L., 1980: The Analytical Hierarchy Process. McGraw-Hill, New York.

Simos J., 1990: Evaluer l'impact sur l'environnement: Une approche originale par l'analyse multicrite` re et la ne´ gociation. Presses Polytechniques et Universitaires Romandes, Lausanne.

Słowiński R., 2007: Podejście regresji porządkowej do wielokryterialnego porządkowania wariantów decyzyjnych. W: Techniki informacyjne w badaniach systemowych. P. Kulczycki, O. Hryniewicz, J. Kacprzyk (red.). Wydawnictwo Naukowo-Techniczne, Warszawa.

Solymosi T., Dompi J., 1985: A Method for Determining the Weights of Criteria: The Centralized Weights. „European Journal of Operational Research", 26, s. 35-41.

Stillwell W.G., Seaver D.A., Edwards W., 1981: A Comparison of Weight Approximation Techniques in Multiattribute Utility Decision Making. „Organizational Behavior and Human Performance", 28, s. 62-77.

Thiessen E.M., Soberg A., 2003: Smartsettle Described with the Montreal Taxonomy. „Group Decision and Negotiation", 12, s. 165-170.

Triantaphyllou E.B., Shu S., Nieto S., Ray T., 2000: Multi-Criteria Decision Making: An Operations Research Approach. John Wiley & Sons, New York.

Tzeng G.H., Chen T.Y., Wang J.C., 1998: A Weight Assessing Method with Habitual Domains. „European Journal of Operational Research", 110(2), s. 342-367.

Wachowicz T., 2008: NegoCalc: Spreadsheet Based Negotiation Support Tool with Even-Swap Analysis. W: J. Climaco, G. Kersten, J.P. Costa (eds.). Group Decision and Negotiation 2008: Proceedings – Full Papers, INESC Coimbra, s. 323-329.

Walesiak M., 2006: Uogólniona miara odległości w statystycznej analizie wielowymiarowej. Wydawnictwo Akademii Ekonomicznej, Wrocław.

Wysocki F., 2010: Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich. Wydawnictwo Uniwersytetu Przyrodniczego, Poznań.

SOME PROPERTIES AND TYPES OF THE SAW PROCEDURE FROM THE PERSPECTIVE OF SUPPORTING NEGOTIATION

Summary

This work provides a survey of the properties of SAW method (*Simple Additive Weighting*) which is one of the simplest and mostly used multiple criteria techniques. The work is presented by focusing mostly on the application of SAW in the support of bilateral negotiations. The strengths and limitations of the proposed approach are discussed and the suggestions of modifications of the classical algorithm are presented from the viewpoint of applications in the negotiation process. The function assigning a score to the negotiation packages, determined by the use of modified SAW procedure is a useful tool facilitating linear ordering of negotiation packages, the estimation of potential concessions, the implementation of a negotiation strategy and the analysis of negotiation compromises.

Wojciech Rybicki
Wyższa Szkoła Oficerska Wojsk Lądowych
im. gen. Tadeusza Kościuszki we Wrocławiu

MODELOWANIE PREFERENCJI A ZAGADNIENIA ZRÓWNOWAŻONEGO (TRWAŁEGO) ROZWOJU*

The time of which a man exists cannot affect the value of his happiness from universal point… the interests of posterity must concern a utilitarian as much as those of his contemporaries.

H. Sidgwick (1874)

[...] mocny, etyczny wymóg jednakowego traktowania wszystkich pokoleń, sam w sobie racjonalny, stoi w sprzeczności z równie mocnym, intuicyjnym przekonaniem, że niemoralne jest oczekiwanie nadmiernie wysokiej stopy oszczędności wobec jednego pokolenia czy wobec wszystkich kolejnych pokoleń.

K.J. Arrow (1995)

Wstęp

Poszukiwanie „właściwych" metod porównań i wyceny strumieni konsumpcji, oszczędności i obciążeń (w czasie) stanowi jeden z odwiecznych tematów zainteresowań ekonomistów. Ogólniej: chodzi o porównywanie użyteczności sekwencji stanów w projektach wieloetapowych oraz o sposoby egzekwowania (różnie rozumianej) sprawiedliwości międzypokoleniowej. O fundamentalnych wątpliwościach w tej subtelnej kwestii mogą świadczyć m.in. zacytowane jako motta niniejszego artykułu dwie opinie wybitnych uczonych, filozofa i ekonomisty, uwypuklające różne aspekty problematyki. O braku jednoznacznych rozstrzygnięć (na gruncie formalnym) będzie mowa w dalszym ciągu opracowania.

* Prezentowane wyniki badań, zrealizowane w ramach tematu nr WZ/009/DzS, zostały sfinansowane ze środków własnych Wyższej Szkoły Oficerskiej Wojsk Lądowych im. gen. Tadeusza Kościuszki.

Metodyka ewaluacji korzyści i kosztów, wynikających z przedsięwzięć podejmowanych obecnie, materializowanych zaś w odległej przyszłości wiąże się wyborem formuł dyskontowania takich dalekookresowych programów. W artykule przytacza się niektóre próby (oraz ujawniane, przy tej okazji, sprzeczności) aksjomatyzacji preferencji w zbiorach ciągów, mających na celu uwzględnienie postulatów typu efektywności, sprawiedliwości i trwałości. Widoczny jest konflikt postulatów typu sprawiedliwości (bezstronności, anonimowości etc.) z postulatami z rodziny efektywności (różne wersje monotoniczności itp.). Pojawiają się kwestie zupełności preporządków oraz ich konstruktywnego definiowania. Większość twierdzeń ma charakter negatywny („brak", „niemożność"), a podstawowe wyniki pozytywne są formułowane w postaci twierdzeń o istnieniu. Niektóre z przytoczonych ustaleń są zaskakujące.

Kluczową rolę pełni kwestia istnienia i kształtu funkcyjnego reprezentacji (numerycznej) preferencji – funkcjonałów dyskontujących, pytań o ich zgodność dynamiczną, „odwracanie" preferencji (z upływem czasu) oraz tzw. malejącą niecierpliwość. Zjawiska te dobrze opisuje się tzw. hiperbolicznymi funkcjami dyskontującymi. Nie stanowią one jednak „panaceum" wyjaśniającego wszelkie odstępstwa realnych procesów od kanonicznych norm racjonalności. W zakończeniu przytoczono niehiperboliczne modele opisu tych zjawisk, które pojawiły się niedawno w literaturze przedmiotu. Rozważa się też możliwość użycia nowego wzorca wyceny w czasie – dyskontowania log-normalnego (propozycja ta jednak nie jest poparta obserwacjami empirycznymi); dokładniejsze studium tego przypadku przygotowuje się w artykule Rybickiego [2014].

Znamienny jest brak klarownych powiązań „efektów pikoekonomicznych" (kwantyfikacja wzorców behawioralno-psychologicznych) z „rekomendacjami makroekonomicznymi" dotyczącymi kształtowania społecznych (międzypokoleniowych) stóp dyskontowych – w obydwu klasach zagadnień pojawiają się, pozornie jednobrzmiące, sugestie formalnego modelowania, podczas gdy problemy te dotykają absolutnie odmiennych „światów". W związku z tym szczególną rangę zyskują zadania agregacji funkcjonałów dyskontujących. Propozycje takich agregacji sformułowano w ostatnim dziesięcioleciu [Reinschmidt 2002; Gollier 2007; Nocetti i in. 2008]. Zagregowane społeczne stopy dyskontowe okazują się niższe niż indywidualne (a czynniki dyskontujące również wolniej maleją). Pokazano też, że agregaty wykładniczych funkcji dyskontujących (stacjonarnych – w sensie parametru „niecierpliwości") są funkcjami hiperbolicznymi (których istotą jest brak stacjonarności, czyli „malejąca niecierpliwość" [Weitzman 1998; Weitzman 2001; Reinschmidt 2002]). Podstawowym celem pracy jest zasygnalizowanie (nieodosobnionego) poglądu autora o „ukonstytuowaniu się" w interdyscyplinar-

nym obszarze badawczym teorii ekonomii, psychologii, biologii, socjologii i matematyki nowego, samodzielnego bytu naukowego, w którym splatają się rozliczne nurty tych dyscyplin. Bardzo selektywny wybór „faktów naukowych" (ustaleń formalnych i empirycznych, głównie z ostatniego półwiecza) ma służyć poparciu sformułowanej tezy. Artykuł ma więc charakter przeglądowy – „z przesłaniem". Niezwykle ważną – konstrukcyjnie i merytorycznie – rolę wyznaczono tu dosyć wszechstronnej bibliografii. Nie wszystkie występujące w spisie literatury pozycje „aktywnie wspierają" prowadzone poniżej wywody. Artykuł nie jest notą bibliograficzną sensu stricto, załączona literatura stanowi jednak „tło dokumentacyjne", które można polecić zainteresowanemu uzupełnieniem informacji Czytelnikowi. Nie pretenduje ona – w żadnej mierze – do kompletności. Dość obszerną bibliografię przedmiotu zawierają prace [Rybicki 2009, 2010b, 2011, 2012]. Artykuł został przygotowany w „konwencji mieszanej": przeważają wywody o charakterze eseistycznym, są też fragmenty sformalizowane oraz komentarze do twierdzeń.

1. Dylematy modelowania sprawiedliwości międzypokoleniowej, realizacji postulatów typu „sustainability" i dyskontowania

Refleksja nad „samopodtrzymywaniem" jako pryncypialnym wymogiem stawianym procesom rozwoju uświadamia konieczność „reformalizacji" niektórych pojęć (rewizji istoty relacji „efektywność – sprawiedliwość") i ich „wkomponowania" w całą mechanikę rozwoju.

Nie do przyjęcia jest „dyktat" współczesnych, przyznających im (poprzez „podparcie się" odpowiednią filozofią i techniką wyceny) przywileje w sferze konsumpcji oraz krótkookresowej ekspansywnej polityki eksploatacji zasobów, tolerując lekceważenie skutków takiej strategii dla jakości życia w przyszłości oraz zagrożenia bezpieczeństwa egzystencji w ogóle. Nie do zaakceptowania jest także odwrócenie proporcji – poświęcenia jakości życia i złożenia na barki współczesnych całokształtu kosztów inwestycji „adresowanych dla potomnych" (i przez nich głównie konsumowanych) [Czerwiński 1992]. Obydwie skrajności łatwo dyskwalifikuje wielookresowy rachunek korzyści i kosztów, w połączeniu z szacowaniem ryzyka zagrożeń oraz „pozostałego czasu życia ekonomii".

Powyższe dylematy są przedmiotem intensywnych badań od około 60 lat – zajmują się nimi teoretycy ekonomii, matematycy oraz przedstawiciele szeroko rozumianych badań operacyjnych w dwóch komplementarnych płaszczyznach:

a) Modelowania preferencji w przestrzeniach nieskończonych ciągów wielkości ekonomicznych o rozmaitej naturze i wykładni, które uwzględniałyby – formali-

zując i rozstrzygając kwestie egzystencjalne – aprioryczne postulaty etyczno-pragmatyczne. Można by było określić ją jako „warstwę aksjomatyczno-formalną". W tym kontekście termin „modelowanie preferencji" brzmi szczególnie trafnie, zawiera bowiem esencję dialektyki procesu: „odzwierciedlanie-abstrahowanie".

b) W drugiej („operacyjnej") warstwie próbuje się owe problemy sprowadzić do usankcjonowania matematyczno-ekonomicznych technik dyskontowania strumieni kosztów i korzyści. Formalne ustalenia z „pierwszej warstwy" przekładają się na metodykę wyceny: konstruowane funkcjonały pełnią rolę kryteriów dla porównywania programów, indukując preporządki w odpowiednich przestrzeniach (ciągów, funkcji).

Pewne wskazówki płyną nie tylko z empirii i obserwacji zjawisk oraz procesów ekonomicznych, ale z obserwacji natury psychologicznej, i to w odniesieniu zarówno do „racjonalnie zachowujących się" podmiotów z gatunku *homo oeconomicus*, jak i do instynktownych zachowań przedstawicieli innych gatunków świata fauny. W obydwu kategoriach odnotowano permanentne „pogwałcanie normatywnych kanonów racjonalności" – szczególnie wyraziście manifestowało się zjawisko niekonsekwencji (odwracania!) preferencji wraz z upływem czasu. Spostrzeżenia te zaowocowały zmianami w narzędziach kwantyfikacji procesów – dostosowania ich do realiów. Rozliczne, nieklasyczne metody dyskontowania sprowadzają się przede wszystkim do zarzucenia formuł wykładniczych (tradycyjnych, „logicznych", „kupieckich" – w najlepszym tego słowa znaczeniu – reguł kalkulacyjnych) – całkowicie nieadekwatnych do ideologii „*sustainability*" oraz sfalsyfikowanych empirycznie (jak wspomniano wyżej). Najbardziej spektakularnym efektem jest „rozsadzanie paradygmatu stacjonarnej niecierpliwości" związanej ze stałością stopy dyskontowej. Od kilku dziesięcioleci są podejmowane próby modyfikacji samuelsonowskiego dyskontowania wykładniczego – w kierunku spowolnienia zbieżności do zera czynnika dyskontującego.

Warte podkreślenia są interdyscyplinarne konotacje problematyki poszukiwania „właściwej" (kompromisowej) metodyki dyskontowania i modelowania preferencji w przestrzeniach ścieżek rozwoju: oprócz poszukiwania (formalnych, cząstkowych) „kompromisowych" rozstrzygnięć fundamentalnej kwestii „*equity vs efficiency*" mieści się w nich wspomniane zagadnienie optymalizacji społecznej stopy dyskontowej, problemy agregacji „stóp indywidualnych", pojawiają się związki (naturalne!) z teorią wzrostu, teoriami oszczędzania (cyklu życia oraz permanentnego dochodu), kontrowersjami wokół teorematu Ricardo-Barro, „krótkowzrocznością" preferencji – powszechnego fenomenu psychologii ekonomicznej (ekonomii behawioralnej), konfliktów i gier międzypokoleniowych, subtelnych kwestii dotykających „handlu czasem".

Za cenę utraty (formalnej) dynamicznej koherencji preferencji uzyskuje się większą zgodność z obserwowanymi zachowaniami podmiotów w skali „mikro", zaś „grube ogony" funkcji dyskontujących (np. hiperbolicznych) przyczyniają się do zmniejszenia stopnia „dyskryminacji" wskaźników dobrobytu odległych w czasie przyszłych pokoleń. Zgoda na niezgodność dynamiczną w modelowaniu procesów decyzyjnych jest ceną, którą warto zapłacić za „ochronę" przed „pułapkami dyskontowania" pojawiającymi się przy rachunku korzyści i strat w tzw. bardzo długim okresie, o porażającej wręcz skali (nieodwracalnych) konsekwencji. Jest to niewielka cena aspiracji do strategicznego zarządzania w sferze „*sustainability*". Czy można jednak mówić o niezgodności (czasowej) formuł dyskontowych różnych od eksponencjalnej? Defekty natury formalnej, np. strata tzw. efektu stopera (jeden z przejawów odwracania preferencji [Bleichrodt i in. 2008]), stają się paradoksalnie atutami w zakresie adekwatności samego odzwierciedlania rzeczywistych przebiegów owych (mentalnych) procesów, w których rzeczonej zgodności po prostu nie ma. Trudno więc nawet mówić o „cenie za niezgodność".

Na zakończenie tego fragmentu zostaną przytoczone niektóre spostrzeżenia związane z konfrontacją wartości dyskontowania „dalekiej przyszłości" według formuł wykładniczych oraz hiperbolicznych, a także przykłady odwracania preferencji.

Harvey [1995] uzasadnia (na gruncie aksjomatycznym) stosowanie hiperbolicznych funkcji dyskontujących w postaci $h(t) = \frac{b}{b+t}$, które przypisują znacznie większe znaczenie przyszłym wielkościom niż geometryczne $g(t) = d^t$. Spektakularny przykład numeryczny, który podał Harvey, dotyczy sytuacji, kiedy stopa dyskontowa wynosi 10% (w skali roku). Zdyskontowana w sposób klasyczny wartość (dzisiejsza) wyniku, równego nominalnie dzisiejszemu, po okresie 100 lat jest 14 000 razy mniejsza od nominału. Dyskontowanie hiperboliczne o identycznej stopie początkowej (następnie – „programowo" malejącej) prowadzi do wyceny jedynie 11 razy „mniej ważnej" za 100 lat niż obecna.

O'Donoghue i Rabin [1999] podają następujący przykład. Decydenci stojący na początku lutego przed wyborem wykonania nieprzyjemnego zadania, które ma im zająć 7 godzin (np. wypełnienie PIT-u) w dniu 15 kwietnia albo takiego samego typu zadania przez 8 godzin w dniu 30 kwietnia, będą w większości preferowali pierwszy wariant. Postawieni wobec tej alternatywy tuż przed 15 kwietnia na ogół zmienią swoje preferencje. Zakładając, milcząco, ciągłość (w czasie) owego „mentalnego" eksperymentu, dochodzi się do wniosku o istnieniu takiej chwili, w której podmiot decyzyjny jest indyferentny w stosunku do przedstawionych wariantów – hiperbole ich dyskonta się przecinają.

Dasgupta i Maskin [2005] podają przykład „dyskontowania instynktowego". Kos decyduje się „zawisnąć" nad krzakami malin w oczekiwaniu na moment, kiedy dojrzeją. W międzyczasie maliny mogą jednak paść łupem stada wron, które „nie wybrzydzając" zjedzą wszystkie owoce. Kos musi przeprowadzić na bieżąco rachunek kosztów i korzyści, dyskontując wartość (kaloryczną) owoców V w stosunku do zagrożenia przylotem wron: jego „naturalną" stopą dyskontową może być tzw. stopa hazardowa dla „procesu nalotów wron", czyli (w chwili t) prawdopodobieństwo warunkowe przybycia wron w przedziale Δt czasowym $(t, t + \Delta t)$ podzielone przez długość przedziału Δt. Tylko dla poissonowskiego procesu „odwiedzin przez wrony" ta stopa jest stała, co „upoważniałoby kosa" do dyskontowania wykładniczego wartości V, wyceniającego wartość obecną jako $e^{-rt}V$. Bardziej realistyczne wydaje się jednak rozważanie zmiennej (zależnej od czasu) stopy $r(t)$. Jej malenie może, po części, wiązać się z kosztami alternatywnymi (zmęczenie organizmu). Tak więc instynkt może w pewnej chwili zmusić (niewykształconego w zakresie teorii dyskontowania) kosa do odlotu.

Na jeszcze inny aspekt wyceny strumieni wielkości w czasie zwraca uwagę Asheim [1996]. Załóżmy, że kolejne pokolenia „biorą sobie do serca" ogólne wytyczne trwałości rozwoju, zarządzają kapitałem oraz zasobami naturalnymi zgodnie z subiektywnymi preferencjami, które z kolei „rekomendują" pewien stopień altruizmu w stosunku do potomnych. Okazuje się, że – wbrew intencjom decydentów – ów altruizm może nie wystarczyć do zagwarantowania przyszłym pokoleniom bezpieczeństwa ekonomicznego. Załóżmy, że subiektywny dobrobyt generacji t (czyli w_t) zależy od poziomu jego własnej „użyteczności statusu" (u_t) oraz – w pewnym stopniu – od subiektywnego (odczucia) dobrobytu przez jej bezpośrednich potomków:

$$w_t = u_t + bw_{t+1} \qquad 0 < b < 1 \tag{1}$$

łatwo zauważyć, że powyższa rekurencja prowadzi do określenia *explicite* poziomu dobrobytu:

$$w_t = \sum_{s=t}^{\infty} b^{s-t} u_t. \tag{2}$$

Jest to formuła zdyskontowanej utylitarystycznej wyceny (o której będzie mowa w następnym punkcie), która ewidentnie zaniża znaczenie użyteczności (statusu) przyszłych pokoleń w stosunku do obecnego.

2. P. Samuelson, T.C. Koopmans i R. Strotz

Celowe wydaje się krótkie odniesienie się do wydarzeń naukowych o kapitalnym dla rozwoju badań w omawianym obszarze znaczeniu i ich wybitnych animatorów. Bez wątpienia rolę takich „kamieni milowych" odegrały w XX w. prace: Samuelsona [1937], Koopmansa [1960] oraz Strotza [1955]. Pierwsze dwie z nich wyznaczyły paradygmat filozofii wyceny i rangowania projektów wielookresowych, a trzecia podważyła ten paradygmat. Dały początek wielowątkowej teorii – w obiegowym słownictwie przedmiotu używa się terminu „dyskontowy paradygmat Samuelsona", a badania w zakresie ewaluacji i dyskontowania strumieni rozdziela cezura lat 1955-60: na „erę przed Koopmansem i Strotzem oraz po nich".

Rozważmy przestrzeń $\boldsymbol{X}$ nieskończonych ciągów elementów z pewnego wypukłego podzbioru $\boldsymbol{C}$ przestrzeni $\boldsymbol{R^n}$ ($\boldsymbol{C}$ może być dowolną, ośrodkową, spójną przestrzenią topologiczną, co umożliwia wielość interpretacji natury badanych strumieni $\boldsymbol{X}$ [Bleichrodt i in. 2008]):

$$\boldsymbol{X} = \{\boldsymbol{x}: x = (x_t;\, t \in \boldsymbol{T},\, x_t \in \boldsymbol{C})\}\,. \tag{3}$$

Niech dana będzie funkcja („użyteczności chwilowej") $\mathbf{u:}\ \boldsymbol{C} \to \boldsymbol{R}$ rosnąca i wklęsła. Symbolem $\boldsymbol{U}$ oznaczymy zbiór wszystkich rzeczywistych ciągów $u = (u_t)$, gdzie $u_t = u(x_t)$ (dla pewnego $t \in \boldsymbol{T}$):

$$\boldsymbol{U} = \{u:\ u\ =\ (u_t;\, t \in \boldsymbol{T})\}\,. \tag{4}$$

Na oznaczenie „czasu" przyjęto symbol $\boldsymbol{T}$. W omawianym przypadku $\boldsymbol{T} = \boldsymbol{N}$ lub $\boldsymbol{T} = \boldsymbol{N}_+$, (w wersji ciągłej, do której będziemy się incydentalnie odwoływać $\boldsymbol{T} = \langle 0, \infty)$ lub $\boldsymbol{T} = \langle 0, T \rangle$).

Ogólna postać (separowalnego addytywnie) funkcjonału wyceny (dyskontującego) jest następująca:

$$U(\boldsymbol{x}) = \sum_{t \in \boldsymbol{T}} d(t)\, u(x_t)\,, \tag{5}$$

gdzie d jest dowolną funkcją rzeczywistą („czasu" t), a szereg po prawej stronie równości jest bezwzględnie zbieżny. Od pojawienia się pracy Samuelsona [1937] rangę kanonu dyskontowania (strumieni z $\boldsymbol{X}$) zyskała formuła tzw. zdyskontowanego utylitarystycznego (DU) funkcjonału:

$$U(x) = \sum_{t \in T} d^t \cdot u_t\,, \text{ gdzie } 0 < d < 1. \tag{6}$$

Jest on dobrze określony co najmniej dla ograniczonych funkcji użyteczności. W tak zwanej wersji ciągłej sumę zastępuje całka odpowiadająca ciągłemu oprocentowaniu (złożonemu):

$$U(x) = \int_T e^{-\delta t} u(x_t) dt, \quad 0 < \delta < 1. \tag{6a}$$

W obydwu przypadkach z definicji stosunek $\frac{d(t+1)}{d(t)}$ (lub $\frac{d'(t)}{d(t)}$) jest stały, stała jest stopa dyskontowa, stała (stacjonarna) tzw. stopa niecierpliwości. Zgodność („logiczna" i „analityczna") powyższej procedury ewaluacyjnej wyraża się niezależnością „siły" dyskontowania od początku przedziału dyskontowania. Nie zależy od czasu oczekiwania na rozpoczęcie procesu dyskontowania. Zależy jedynie „addytywno-multyplikatywnie" od długości przedziału dyskontowania:

$$d^t \cdot d^s = d^{t+s}; \; e^{-\delta t} \cdot e^{-\delta s} = e^{-\delta(t+s)}. \tag{7}$$

Przypomnijmy, że funkcje wykładnicze są jedynymi funkcjami o tej własności (w rodzinach ciągów oraz funkcji ciągłych).

Niezależnie od regularności analitycznej i „zdroworozsądkowej" techniki kalkulacyjnej dyskontowanie DU nie miało „pierwotnego umocowania" w filozofii rangowania strumieni użyteczności. Lukę tę wypełniła znamienna praca Koopmansa [1960], w której podano aksjomatyzację preferencji w przestrzeni ciągów „użyteczności chwilowych", implikującą własności „stacjonarno-wykładnicze" dyskonta.

Aksjomaty Koopmansa dotyczyły własności relacji preporządku w przestrzeni liczbowych ciągów (nieskończonych) i odzwierciedlały intuicyjne postulaty: tzw. separowalności wszystkich okresów oraz stacjonarności (pojęcia te będą uściślone poniżej). Pierwszy z nich jest utrzymany w duchu aksjomatu niezależności [von Neuman-Morgenstern 1944], stanowiącego fundament ich aksjomatyki oczekiwanej użyteczności oraz jego odpowiednika w „subiektywnym" wariancie Savage'a [1954] – „*sure things principle*". Drugi oznacza niezależność relacji między dwoma dowolnymi ciągami od ich początkowych odcinków, jeśli są one – na tych segmentach – identyczne.

Bleichrodt i in. [2008] uściślili i uprościli aksjomatykę Koopmansa oraz uogólnili jego twierdzenie o istnieniu (takich) preferencji i ich reprezentacji DU (uściślenie dotyczyło doprecyzowania dziedziny relacji oraz kwestii ograniczoności funkcji użyteczności „brzegowych"). Poniżej zrelacjonujemy w syntetycznej formie tę (zmodyfikowaną) konstrukcję. Nie ma oczywiście sensu w każdym wzmiankowanym przypadku referować detali koncepcji autorów. Wyjątek od

reguły czynimy w tym artykule trzykrotnie – dla pomysłów (i wyników) o szczególnym znaczeniu dla omawianej tematyki.

Rozważa się przestrzeń $\boldsymbol{X}$ programów (konsumpcji) wprowadzoną już w poprzednim punkcie, we wzorze: $x \in \boldsymbol{X} \Leftrightarrow x = (x_1, x_2, ...),\ x_K \in \boldsymbol{C}$. Dla zadanej w $\boldsymbol{X}$ relacji $\succcurlyeq$, preporządku liniowego (relacja przechodnia i zupełna), jej część asymetryczną oznacza się tradycyjnie symbolem $\succ$, symetryczną – symbolem $\sim$, dziedzinę – symbolem $\boldsymbol{F}$.

Mówi się, że w $\boldsymbol{F}$ jest określona zdyskontowana użyteczność, jeśli istnieją: funkcja (użyteczności) $u : \boldsymbol{C} \to \boldsymbol{R}$ oraz czynnik dyskontujący $\rho \in (0,1)$ takie, że każdy program $\boldsymbol{x} \in \boldsymbol{F}$ jest wyceniany według wzoru:

$$DU(\boldsymbol{x}) = \sum_{t=1}^{\infty} \rho^{t-1} u(\boldsymbol{x}_t), \tag{6b}$$

szeregi powyższe są sumowane, a wartości ich sum (dla ustalonych $\boldsymbol{x}$) nazywa się zdyskontowanymi użytecznościami (DU) strumieni $\boldsymbol{x}$; strumienie o większej wartości DU są – z definicji – przedkładane nad strumienie o mniejszej DU.

Przyjmuje się następujące umowy co do oznaczeń. Dla strumienia $x = (x_1, x_2, ...)$ oraz (jednookresowej) konsumpcji $\alpha \in \boldsymbol{C}$ $\alpha x = (\alpha, x_1, x_2, ...)$ (konsekwentnie: $\alpha\beta x = (\alpha, \beta, x_1, x_2, ...)$). Preferencje (indukowane przez relację „globalną") w zbiorze konsumpcji $\boldsymbol{C}$ są – z definicji – zgodne z preferencjami w całej dziedzinie: $\alpha \succcurlyeq \beta \Leftrightarrow (\alpha, \alpha, ...) \succcurlyeq (\beta, \beta, ...)$. Preferencje nazywają się monotoniczne, jeśli z zachodzenia relacji $x_t \succcurlyeq y_t\ (\forall t \in 1, 2, ...)$ wynika zachodzenie relacji $x \succcurlyeq y$. Jeśli ponadto dla pewnego t jest $x_t \succ y_t$, to także $x \succ y$. Preporządek $\succcurlyeq$ jest „wrażliwy na preferencje teraźniejsze" (w pierwszym okresie), jeżeli istnieją także $\alpha, \beta \in \boldsymbol{C}$ oraz $x \in \boldsymbol{X}$, że $\alpha x \succ \beta x$. Dzięki temu eliminuje się trywialny przypadek – indyferencji (wzajemnej) wszystkich programów z $\boldsymbol{F}$ oraz całkowite pominięcie wagi pierwszego okresu. Uściślenie wspomnianego wcześniej aksjomatu niezależności (separowalności) wyraża się równoważnością:

$$\alpha\beta x \succcurlyeq \gamma\delta x \Leftrightarrow \alpha\beta y \succcurlyeq \gamma\delta y, \tag{8}$$

dla wszystkich programów x, y oraz „konsumpcji brzegowych" $\alpha, \beta, \gamma, \delta$. Warunek ten wyraża niezależność skutków modyfikacji ciągów na dwóch pierwszych miejscach od całej przyszłej konsumpcji. Tę separowalność można przenieść na dowolne segmenty strumieni.

Jeśli wszystkie okresy są separowalne, to preferencje spełniają warunek łącznej niezależności. Z kolei warunek stacjonarności preferencji formalizuje poniższa równoważność:

$$\alpha x \succ \alpha y \Leftrightarrow x \succ y \tag{9}$$

dla wszystkich programów $x, y \in \boldsymbol{F}$ oraz wszystkich $\alpha \in \boldsymbol{C}$.

W konsekwencji dla określenia preferencji między programami wystarczy rozważać podprogramy zaczynające się od pierwszych wyrazów, które są różne – w wyjściowych ciągach. Program nazywa się T-prawie stały, jeśli istnieje indeks $T \in \boldsymbol{N}_+$ taki, że $x = (x_1, x_2, ..., x_T, \alpha, \alpha, ...)$. Dla ustalonego T zbiór takich programów oznacza się symbolem $\boldsymbol{X_T}$: łatwo zauważyć zbiór $\boldsymbol{X_T}$ wyznaczony jest T+1-wymiarowy produkt zbiorów **C** (utworzony z wektorów w postaci $(x_1, ..., x_T, \alpha)$. „Skończoną – prawie ciągłość" relacji μ definiuje się jako koniunkcję warunków ciągłości każdego jej zawężenia do zbioru $\boldsymbol{X_T}$ (godne podkreślenia jest sprowadzenie rozważań natury topologicznej do „przypadków skończenie wymiarowych").

A oto pozostałe definicje pomocne przy sformułowaniu aksjomatyki Koopmansa oraz twierdzenia o reprezentacji DU. Program $\boldsymbol{x} \in \boldsymbol{F}$ spełnia warunek „równoważności ze stałą", jeśli istnieje równoważny z nim program stały $x \sim (\alpha, \alpha, \alpha, ...)$. Program $\boldsymbol{x} \in \boldsymbol{F}$ jest „odporny na zachowanie w odległej przyszłości" (tail-robust), jeśli dla wszystkich wyników $\beta \in \boldsymbol{C}$ zachodzi implikacja:

$$x \succ \beta \ (x \prec \beta) \Leftrightarrow \exists T \in N_+ \ x_T\beta \succ \beta \ (x_T\beta \prec \beta) \ \forall t > T. \tag{10}$$

Twierdzenie ([Bleichrodt i in. 2008] „Preference Axiomatization for Discounted Utility).

Niech relacja μ będzie określona na dziedzinie $\boldsymbol{F}$ zawierającej wszystkie prawie stałe programy. Następujące stwierdzenia są równoważne:

i. na $\boldsymbol{F}$ jest określona zdyskontowana użyteczność, przy czym funkcja użyteczności u jest ciągła i nie jest stała,

ii. relacja ⩾ spełnia poniższe warunki:

a) ⩾ jest preporządkiem wrażliwym na preferencje teraźniejsze, skończenie prawie ciągłym,

b) ⩾ spełnia warunki równoważności ze stałą oraz odporności na odległą przeszłość,

c) ⩾ ma własność separowalności dwóch początkowych okresów i stacjonarności.

Twierdzenie Koopmansa usankcjonowało paradygmat DU od strony formalnej. Niemniej jednak w tej samej pracy Koopmans [1960] sformułował dwa fundamentalne pytania dotyczące istnienia tzw. etycznych preferencji i ich reprezentacji. Zacznijmy od dwóch definicji.

A. Relacja preferencji $\succcurlyeq$ określona w dodatnim stożku:

$$\boldsymbol{L}_t^{\infty} = \left\{ x = (x_0, x_1, ...) : x_t \geq 0,\ \forall t \in \boldsymbol{N},\ \sup x_t < \infty \right\} \tag{11}$$

nazywa się (skończenie) międzypokoleniowo bezstronną (symetryczną, niezmienniczą, anonimową), jeśli dla każdej permutacji $\pi_f : \boldsymbol{L}_+^{\infty} \to \boldsymbol{L}_+^{\infty}$, niezmienniczej dla prawie wszystkich współrzędnych („pokoleń") i dla wszystkich $x, y \in \boldsymbol{L}_+^{\infty}$:

$$x \succcurlyeq y \Leftrightarrow \pi_f(x) \succcurlyeq \pi_f(y). \tag{12}$$

B. Relacja preferencji μ w przestrzeni dodatnich, ograniczonych ciągów ($\boldsymbol{L}_+^{\infty}$) nazywa się słabo monotoniczna (Pareto-efektywna, PA), jeśli zachodzi implikacja:

$$(\forall t \in \boldsymbol{N}\ \ x_t > y_t) \Leftrightarrow x \succ y. \tag{13}$$

Relacja spełniająca powyższe dwa postulaty (lub ich modyfikacje w kierunku „zwariantowania" Pareto-efektywności) nazywa się relacją międzypokoleniowego, społecznego dobrobytu (używa się także syntetycznej nazwy: „preferencje etyczne" – terminologię tę utrwalił G. Asheim [1996]).

Dwa pytania Koopmansa brzmiały: czy istnieją (w $\boldsymbol{L}_+^{\infty}$) preferencje etyczne, a jeśli tak, to czy mają rzeczywistą reprezentację, czyli („agregujący historię") funkcjonał wyceny, nazywany też międzypokoleniowym funkcjonałem dobrobytu społecznego. Postawienie tych kwestii przyczyniło się do intensyfikacji badań teoretycznych, wpisujących się w ogólniejszy nurt tematyki sprawiedliwości międzypokoleniowej oraz trwałości rozwoju (niektórzy autorzy utożsamiają te subdyscypliny [Fiedor 2007]). Bardzo szerokie spektrum tej tematyki obejmowało ogólne, filozoficzno-etyczne polemiki dotykające samej istoty, interpretacji i „operacjonalizacji" kategorii sprawiedliwości – w kontekście społecznej sprawiedliwości dystrybutywnej [Rybicki 2009, 2010a], ale także wolności wyboru [Sen 1988] oraz „imperatywu zasłony niewiedzy" Kanta-Rawlsa [1970], mającej gwarantować „sterylne" warunki realizacji procesów wyboru, decyzji i podziału, nieskażone świadomością podmiotu o jego położeniu (pozycji w społeczeństwie, w czasie). Pomimo iż filozoficzna koncepcja „totalnej anonimowości dla sprawiedliwości" nie była wolna od pewnych wad natury lo-

giczno-formalnej, jej konsekwencja kwantytatywna (czyli tzw. kryterium „lexy-minu", też krytykowane – paradoksalnie – za „asymetrię" w traktowaniu jednostek czy pokoleń) zyskała wielu zwolenników, a w zmodyfikowanych wersjach weszła do kanonu metod ewaluacji strumieni użyteczności [Sakai 2003; Basu, Mitra 2003; Alcantud, Garcia-Sanz 2010]. Propozycję implementacji tej zasady do zagadnień teorii wzrostu gospodarczego podał m.in. Solow [1974] w pracy przedstawionej w trakcie słynnego sympozjum poświęconego teorii wzrostu w warunkach ograniczonych zasobów, co było niezwykle radykalne. W cytowanej wyżej pracy z 1996 r. G. Asheim stwierdził: „with resource constraints it seems right for Solow (1974) to be >plus Rawlsian que le Rawls<". „Jego" preferencje etyczne w zbiorze użyteczności ścieżek wzrostu porządkowały je pod kątem porównań wartości użyteczności najsłabszych generacji $(\inf_{t\geq 1} u_t)$. W klasycznym modelu wzrostu (z ograniczonymi zasobami) Dasgupty-Heala-Solowa („DHS": synteza prac [Dasgupta, Heal 1974] oraz [Solow 1974] ze wspomnianego sympozjum z 1974 r.) kryterium to prowadzi do całkowicie egalitarnych, stagnacyjnych (constans) ścieżek wzrostu [Asheim 1996]. Całe ostatnie półwiecze obfitowało w badania nad możliwościami formalnego pogodzenia – przeciwstawnych w gruncie rzeczy – postulatów Koopmansa. W następnym punkcie naszkicowano dwie propozycje modelowe w tym zakresie.

Ogólnie rzecz biorąc, ustalenia teoretyczne były ambiwalentne. Od pracy Diamonda [1965], w której autor podał negatywną odpowiedź na pytania Koopmansa, w przypadku jeśli żąda się ciągłości preferencji w metryce jednostajnej, we wszystkich pracach „balansowano" pomiędzy różnymi odcieniami założeń bezstronności (m.in. rozważano niezmienniczość preporządków względem różnych podklas przestrzeni wszystkich permutacji indeksów [Lauwers , Liedekerke 1997; Basu, Mitra 2002; Banerjee 2008; Alcantud, Garcia-Sanz 2010] oraz efektywności (a'la Pareto), rezygnowano też z postulatów o charakterze topologicznym [Basu, Mitra 2003]), co prowadziło do rozmaitych wyników. Wszystkie miały charakter cząstkowy i warunkowy, większość – negatywny lub niekonstruktywny. Kwestie te zostały ostatnio wyjaśnione na gruncie pogłębionej analizy formalnej [Lauwers 2010; Żylicz 2001], o czym będzie mowa dalej.

Czasem drobne modyfikacje założeń prowadziły do przeciwstawnych konkluzji (np. [Alcantud, Garcia-Sanz 2010]), czasem „sprytny" ich dobór zapewniał istnienie międzypokoleniowej relacji społecznego dobrobytu, wykluczając jednocześnie możliwość rzeczywistej („jednowymiarowej") reprezentacji tych etycznych preferencji (por. np. [Basu, Mitra 2003]). Dość obszerną bibliografię kierunków rozwoju i postępów badań w tym obszarze zawierają prace [Rybicki

2009, 2010b, 2011]. G. Chichilnisky zamieściła w pracy [Chichilnisky 1997] krytyczny przegląd „jednowymiarowych" kryteriów (czasem przybierają one postać sekwencyjnych procedur porównywania określonych wskaźników), generujących preporządki w przestrzeniach ciągów (por. także [Rybicki 2011; Asheim 2010]). Konstrukcja takich kryteriów również odzwierciedla aspekty sprawiedliwości, niecierpliwości i uprzywilejowania określonych generacji. Tu podstawową osią kontrowersji – zarówno w planie filozoficznym, jak i formalno-matematycznym – jest kwestia interpretacji obiektywizmu. Utylitarystyczne, pozornie „obiektywne" kryterium Ramseya – sumowania szeregów użyteczności bez wag (zerowa stopa dyskontowa) oraz jego mutacja – skumulowanych odległości od błogostanu („bliss") [Asheim 1996], zdecydowanie krzywdzą pierwsze pokolenie:

$$\sum_{t=1}^{\infty} u_t; \quad \sum_{t=1}^{\infty} (b - u_t). \tag{14}$$

W podobnej „ideologii" jest skonstruowane kryterium zaspokojenia podstawowych potrzeb, autorstwa Chichilnisky [1977], które rekomenduje porównywanie liczebności sum częściowych szeregów użyteczności, przekraczających po raz pierwszy ustalony poziom. Wszelkie kryteria są obarczone pewnymi defektami: czy to w sferze racjonalności, czy też formalnej (sumowalność szeregów użyteczności), czy też z powodu dyskryminacji (w wycenie) znaczenia pokoleń współczesnych albo przyszłych. W świetle uwag poczynionych wcześniej, odnoszących się do niekonstruktywnego charakteru ustaleń „o istnieniu" etycznych preferencji, warto przypomnieć koncepcję von Weizseckera [1965] tzw. kryterium doganiania (*over-taking, catch-up-criterion*): wyżej są wyceniane strumienie, których prawie wszystkie (skończone) sumy częściowe są większe. Kryterium to generuje, oczywiście, preporządek częściowy i wydaje się być „wymyślone ad hoc". W 1970 r. Brock podał aksjomatyczne umocowanie tej techniki. Idea rozszerzania subrelacji na wstępujących rodzinach skończonych podciągów przestrzeni wszystkich ciągów (w najprostszym przypadku: rzeczywistych, dodatnich i ograniczonych) była kilkakrotnie podejmowana (por. np. [Lauwers 1995; Asheim i in. 2010b; Sakai 2010]). Jest to chyba jedyna metodyka oferująca konstruktywne rozwiązania tej klasy zadań (por. [Sakai 2010]).

Z tak zwanych wyników pozytywnych na uwagę zasługują ustalenia zawarte w pracy Banerjee, Mitra [2007] (por. także [Asheim i in. 2010a]). Autorzy tej pracy rozpatrują „swój" warunek niecierpliwości dla programu $\boldsymbol{x}$ (a'la Koopmans). Wyraża on („z grubsza") wyższą wycenę strumieni, konsumpcji, w których para wyrazów o własności „większe x_i występuje wcześniej niż mniejsze x_j" przesądza

o przewadze nad strumieniem, w którym ta para pojawia się w odwrotnej kolejności. Ponadto (również wzorując się na wspominanej już idei Von Neumanna-Morgensterna, Savage'a oraz Koopmansa) zakładają niezależność preferencji indukowanych „na fragmencie strumienia" od dopełniczych przebiegów (o ile są one identyczne). Przywołani autorzy otrzymują mocne, ciekawe wnioski:

1) jeśli relacja $\succcurlyeq$ jest paretowska, to zbiór strumieni spełniających warunek niecierpliwości jest mocy continuum,

2) jeśli relacja $\succcurlyeq$ jest paretowska na zbiorze $\boldsymbol{X} = \langle 0;1\rangle^N$ z metryką Czebyszewa (d), to zbiór ten jest gęstym podzbiorem przestrzeni metrycznej ($\boldsymbol{X}$, d).

Słowami: w przestrzeni ciągów ograniczonych jest „bardzo dużo" programów, w których przedkłada się wcześniejszą realizację większej wartości nad późniejszą. Jeśli rozpatrzymy tylko ciągi o wyrazach z odcinka ⟨0; 1⟩, to zbiór ten „wypełnia" prawie całą przestrzeń. Są to mocne argumenty (dla racjonalnych „paretowskich") podmiotów, między innymi dla agitacji za paradygmatem DU. Jednak dopełnienia zbiorów, o których mowa w twierdzeniu, też mogą być mocy continuum oraz gęste.

Spory „kłopot" sprawiło odkrycie Strotza [1955], którego istotą jest zanegowanie – na gruncie empirii – „ideologii" DU, wraz z wykazaniem niezgodności dynamicznej rzeczywistych preferencji, oraz pionierskie ustalenia w zakresie dyskontowania hiperbolicznego. W tym miejscu ograniczymy się jedynie do przytoczenia jednego z oryginalnych kontrprzykładów R. Strotza.

Na początku roku kalendarzowego wielu ludzi planuje zachować określone środki finansowe na okolice Bożego Narodzenia. Jednak w miarę upływu czasu (w lecie) pojawia się okazja wydania części z nich na wakacje i wielu z nich decyduje o zmianie przeznaczenia tych pieniędzy. Następnie rozpoczyna się rok szkolny i znów ludzie zmieniają (co najmniej modyfikują) pierwotne preferencje i wydają część pieniędzy na wyposażenie szkolne („wyprawka" oraz odzież). Tak więc stają się bardziej niecierpliwi niż w styczniu. Stopa niecierpliwości (stopa dyskonta) „kurczy się", gdy czas do realizacji jest większy (w rozpatrywanym przypadku musimy „podążać za czasem wstecz"), jej zmienność skutkuje odwracaniem zwrotu preferencji, co z kolei znajduje wyraz w hiperbolicznej (odwrotnie proporcjonalnej do czasu) funkcji dyskontującej [Dasgupta , Maskin 2005; Strotz 1955].

W jeszcze większą „konfuzję" może wprawić argumentacja prowadzona na gruncie teorii wzrostu. Asheim [1996] przedstawił fundamentalną – zarazem klarowną – krytykę metodyki dyskonta geometrycznego w kontekście modelu DHS. Poniżej przytoczono cytat z pracy Asheima i Mitry [2010], w którym autorzy „streszczają" owo uzasadnienie: „ When applied to some models of economic growth, DU leads to seemingly unappealing consequences. In particular, in model of capital accumulation and resource depletion first analyzed by

Dasgupta and Heal [1974] and Solow [1974] (…) the application of DU forces consumption to approach zero as time goes to infinity, even though sustainable streams with constant or increasing consumption are feasible. Moreover, this result holds for any discount factor δ smaller than one; even when δ is close to one so the discounting is small (…) when applied to DHS model, the use of DU undermines the livelihood in the far future also when each generation is given almost the same weight as its predecessor".

3. Etyczne preferencje w przestrzeniach ciągów i ich reprezentacje – wybrane modele i dyskusja

Nie tylko modyfikacje, stopniowanie „siły" warunków sprawiedliwości oraz efektywności, ale także ustalanie zbiorów wartości chwilowych użyteczności doprowadziły w ostatnim dwudziestoleciu do sformułowania licznych ustaleń (często ambiwalentnych, jak już wspomniano), odnoszących się do kwestii istnienia i (ewentualnej) reprezentacji liczbowej etycznych preferencji. Można mówić o „naukowej batalii" o osiągnięcie formalno-metodologicznego kompromisu w poszukiwaniu „słusznej, dalekowzrocznej" wyceny rzeczy i procesów – na kształt „iustum pretium" św. Tomasza z Akwinu, we współczesnym, międzypokoleniowym „kostiumie ekonomiczno-matematycznym". Poniżej zasygnalizowano niektóre sprzeczności w godzeniu podstawowych postulatów, próby „ominięcia" tych sprzeczności, a także stwierdzenia (odwołujące się do podstaw matematyki) wyjaśniające istotę kłopotów z „trade-off" między „efficiency" a „equity".

Svensson [1980] wykazał istnienie w zbiorze ciągów rzeczywistych z odpowiednio zdefiniowaną metryką (por. także [Rybicki 2010a; Żylicz 2001] relacji preferencji etycznych, ciągłych w „jego" metryce, a Bossert i in. [2007] wzmocnili jego wynik (dodając postulat spełniania zasady transferów Pigou-Daltona lub „sprawiedliwości Hammonda" [Hammond 1976]). Przywołane twierdzenia miały tezy stricte egzystencjalne oraz niekonstruktywne dowody: w kluczowych punktach korzysta się z twierdzenia E. Szpilrajna [1930] o istnieniu rozszerzenia (pre)porządków częściowych do liniowych (zupełnych), które z kolei bezpośrednio zależy od aksjomatu wyboru (ewentualnie lematu Zorna). Analogiczna – w przypadkach „pozytywnych rozstrzygnięć" – jest idea rozumowań prowadzonych w bardzo ważnych pracach [Basu, Mitra 2003; Banerjee 2006], w których nie rozważa się topologii w $\boldsymbol{X}$. Bardzo ciekawe są w tym kontekście mocno sceptyczne ustalenia z artykułu Zame [2007]. Wykazano tam, że:

a) każda relacja preferencji etycznych (spełniających warunek PA) jest „bardzo niezupełna": prawie wszystkie (w sensie miary zewnętrznej) pary strumieni użyteczności są indyferentne,

b) jeśli nawet zrezygnujemy z wymogu zupełności i warunku PA, to okazuje się, że dla każdej przeciwzwrotnej relacji (w zbiorze strumieni użyteczności) prawie wszystkie pary są nieporównywalne,

c) żadna relacja preferencji etycznych nie jest mierzalna (jako podzbiór kwadratu kartezjańskiego $\boldsymbol{Y}^N \times \boldsymbol{Y}^N$); w konsekwencji istnienie takich (powszechnie akceptowalnych – z perspektywy budowy modeli teorioekonomicznych) relacji jest niezależne od tzw. aksjomatu zależnego wyboru (fundamentalnego dla aksjomatyki Zermelo-Fraenkla); wreszcie jeśli taka relacja istnieje, to nie może być *explicite* zdefiniowana.

Postawiono też problem związku istnienia preferencji etycznych z tzw. zbiorami Ramseya [Lauwers 2010; Fleurbaey, Michel 2003]. W pracy z 2010 r. L. Lauwers scharakteryzował te związki. Podstawowa teza tej pracy (w nieco „zbeletryzowanej" wersji) brzmi: skończenie permutacyjne bezstronne i paretowska relacja preporządku w zbiorze nieskończonych ciągów zero-jedynkowych albo jest niezupełna, albo jest „obiektem niekonstruktywnym" (co pociąga za sobą niemożność podania jej – *explicite* – opisu). Autor wykazuje też związki między istnieniem etycznych preferencji a ultrafiltrami [Luwers 2010; Kulpa 2009] oraz zbiorami nieramseyowskimi [Ramsey 1928; Mathias 1977].

W cytowanej wyżej pracy Lauwers trawestuje opinię H. Weila o niekonstruktywnych tezach w matematyce: „the world is informed about a treasure (a finite anonymous and Paretian ordering) and about the fact, that – with certainty – no one will ever disclose its location"!

W licznych pracach widać wyraźnie współzależności wspomnianych na wstępie odcieni założeń. Tak zwany słaby aksjomat Pareta (Weak Pareto Axiom – WPA) w wersji [Basu, Mitra 2003] ma następującą postać:

a) dla dowolnych strumieni $x, y \in \boldsymbol{X}$, jeśli istnieje taki wskaźnik $j \in N$, że $x_j > y_j$ oraz dla wszystkich $k \neq j$, $x_k \geq y_k$, wówczas $W(x) > W(y)$,

b) dla dowolnych $x, y \in \boldsymbol{X}$, jeśli wszystkie współrzędne pierwszego strumienia są ostro większe od odpowiednich wyrazów w drugim strumieniu, ($x \gg y$), to $W(x) > W(y)$ (W oznacza międzypokoleniowy funkcjonał społecznego dobrobytu, czyli rzeczywistą reprezentację badanych preferencji, Social Welfare Functional – SWF). Alcantud i Garcia-Sanz [2010] pokazują, że jeśli zbiór wartości użyteczności chwilowych jest podzbiorem zbioru liczb naturalnych, to istnieje zarówno relacja spełniająca WPA oraz bardzo słabą formą anonimowości (niezmienniczość ewaluacji W względem permutacji dwuelementowych zbiorów wskaźników, przy pozostałych – odpowiednio – równych), jak i jej reprezentacja. Zastąpienie WAP jego mocniejszą wersją – „zwyczajnej" ścisłej monotoniczności W względem porządku Pareto (Pareto Axiom, PA), a przeciwdziedzi-

ny u dowolnym podzbiorem zbioru liczb rzeczywistych $\boldsymbol{R}$ prowadzi do wyniku negatywnego – nieistnienia reprezentacji (funkcjonału SWF). Z kolei nawet ograniczenie zbioru wartości wyrazów ciągu do zbioru dwupunktowego {0, 1} nie „ratuje" tezy o istnieniu, podczas gdy zastąpienie PA przez WPA pozwala ją utrzymać przy dowolnym zbiorze wartości wyrazów ciągu (np. [0, 1]. Nieco wcześniej Lauwers [1997] rozważał możliwość „godzenia" PA (w jego terminologii – strong *sensitivity*) z tzw. ścisłą neutralnością (bardzo mocna wersja bezstronności: niezmienniczości preferencji względem zbioru wszystkich permutacji indeksów, czyli wszystkich wzajemnie jednoznacznych odwzorowań całego zbioru liczb naturalnych N na siebie). Otrzymał – oczywiście – wynik negatywny, nawet dla najmniejszego, nietrywialnego zbioru wartości użyteczności {0; 1}.

Ważną rolę w relacjonowanym nurcie badań odgrywają definicje sprawiedliwości w ujęciu P. Hammonda. Tak zwany oryginalny postulat HE (Hammond Equity) sformułowany w pracy Hammonda [1976] został następnie – w pracach Lauwersa z lat dziewięćdziesiątych oraz Asheima i in. [2001] – zmodyfikowany pod kątem zastosowania w kontekście międzypokoleniowym (a także: większej czytelności).

A. Warunek (Lauwersa): „słabej wersji niesubstytucyjności" – Weak Non Substitubility (WNS, por. także [Asheim i in. 2010b; Lauwers 1998]).

Dla każdych czterech wartości (wyrazów strumienia użyteczności) x, y, z, v zachodzi implikacja: jeśli $v > z$, to strumień (program) w postaci (x, z, z, z, …) nie jest preferowany nad (żaden) program (y, v, v, v, …).

B. Warunek (Hammonda) – „sprawiedliwości wobec przyszłości" (Hammond Equity for Future, HEF [Asheim i in. 2010b; Alcantud, Garcia-Sanz 2010; Asheim, Mitra 2010]).

Dla dowolnych czterech wyrazów x, y, z, $v \in Y$ (zbiór wartości funkcji użyteczności): jeśli $x > y > v > z$, to strumień (program) w postaci (x, z, z, …) nie jest preferowany nad (żaden) program (y, v, v, v, …).

Należy podkreślić przyjmowane, implicite, założenie o tym, że „indykatory dobrobytu" (użyteczności) poszczególnych pokoleń można mierzyć przynajmniej w skali porządkowej raz, że ich poziomy można (sensownie!) porównywać w planie horyzontalnym – między generacjami. Warunek HEF wiąże się z porównywaniem „poświęcenia" jednego pokolenia z jednolitym zyskiem dla każdego z nieskończonego zbioru pokoleń, które są w gorszej sytuacji: jeśli poświęcenie przez obecne pokolenie (będące w lepszej sytuacji niż przyszłe) skutkuje zwiększeniem użyteczności wszystkich potomnych, to taki transfer jest (co najmniej słabo) preferowany (w stosunku do status quo). Asheim i in. [2010] podkreślają, w związku z tym, że „filozofia" HEF znajduje oparcie zarówno z pozycji egalitaryzmu, jak i utylitaryzmu. Pierwszy znaczący wynik na ścieżce

poszukiwań reprezentacji numerycznej (SWF) dla takiej wersji (formalnej) sprawiedliwości był jednak negatywny. Banerjee [2006] wykazał niemożliwość takiej reprezentacji dla ciągów o wyrazach z przedziału $\boldsymbol{Y} = \langle 0; 1\rangle$ (nawet po zastąpieniu silnego wymogu efektywności PA przez słabą dominację (WPA)).

Z kolei Alcantud i Garcia-Sanz [2010], zastępując zbiór wartości $\boldsymbol{Y}$ zbiorem wszystkich liczb naturalnych, otrzymali wynik pozytywny – dla nieco „zaostrzonej" wersji HEF i silnego porządku Pareto. Godne podkreślenia jest także to, że ich twierdzenie ma charakter konstruktywny: podany jest algorytm konstrukcji SWF dla HEF^+ („ich zaostrzenia" HEF).

Jedną z najciekawszych propozycji ostatniego dwudziestolecia była koncepcja (i konstrukcja) [Chichilnisky 1996] tzw. trwałych (*sustainable*) preferencji. G. Chichilnisky sformułowała dwa postulaty natury etycznej: preferencje (i kryteria) nie powinny dyskryminować ani przyszłości, ani teraźniejszości. W jej terminologii oznacza to wykluczenie (w procedurach ewaluacji i porównań strumieni) tzw. dyktatury teraźniejszości oraz dyktatury przyszłości: „No Dictatoral Role of the Present" (NDP), „No Dictatorial Role of the Future" (NDF). Oto krótki zarys tej koncepcji – z pominięciem szczegółów technicznych. Punktem wyjścia jest przestrzeń wszystkich „osiągalnych" (*feasible*) „ścieżek konsumpcji" $\boldsymbol{F}$, gdzie:

$$\boldsymbol{F} \subset \{x : x = (x_g),\ g = 1, 2, \ldots\ x_g \in \boldsymbol{R^n}\}\ . \tag{15}$$

Konsumpcja dóbr jest tu rozumiana bardzo ogólnie, m.in. jako eksploatacja odnawialnych (nieodnawialnych) zasobów, wykorzystywanie kapitału (poddanego stałym procesom akumulacji oraz deprecjacji). Zakłada się istnienie rzeczywistej reprezentacji (użyteczności) kwantyfikującej status quo każdego pokolenia – jest to funkcja u: $R^n \to R$, taka, że zbiory jej wartości na elementach z $\boldsymbol{F}$ są ograniczone, a użyteczności poszczególnych pokoleń są porównywalne:

$$(\alpha)\ \sup(u_g(x_g))_{x_g \in \boldsymbol{R^n}} \leq 1\ .$$

Przestrzeń strumieni użyteczności jest oznaczona symbolem $\boldsymbol{\Omega}$.

$$(\beta)\ \boldsymbol{\Omega} = \{\alpha : \alpha = (\alpha_g),\ \alpha_g = u_g(x_g)\ g = 1, 2, \ldots;\ \ x \in \boldsymbol{F}\} \subset \boldsymbol{L^\infty}\ .$$

„Poszukiwane" kryterium (funkcjonał SWF) powinien być rosnącą (w sensie WPA) funkcją:

$$(\gamma)\ W : \boldsymbol{L^\infty} \to \boldsymbol{R^+}.$$

Przyjmijmy następującą konwencję. Jest to para definicji:

(δ) Teraźniejszą przestrzeni ***F*** nazywa się zbiór wszystkich „*K*-obcięć" strumieni z **Ω**, czyli ciągów powstałych z elementów **Ω** przez zastąpienie (w każdym z nich) wyrazów o indeksach wyższych od *K* zerami (pierwsze *K* wyrazów się nie zmienia); $K \in N_+$.

(ε) Przyszłością (przestrzeni ***F***) nazywa się zbiór wszystkich „*K*-ogonów" strumieni z **Ω**, czyli ciągów, w których początkowe *K* wyrazów zastąpiono zerami, pozostawiając pozostałe bez zmian.

Teraźniejszość **Ω** będzie oznaczana symbolem ***P*(Ω)**, a przyszłość – symbolem ***F*(Ω)**. Elementy ***P*(Ω)** będą oznaczane symbolami α^K, β^K; elementy ***F*(Ω)** – symbolami γ^K, δ^K.

DEFINICJA 1 [Chichilnisky 1996]

(φ) Kryterium *W* ustanawia dyktaturę teraźniejszości (DP), jeśli jest „wrażliwe" tylko na ***P*(Ω).** Formalnie: dla dowolnych dwóch strumieni użyteczności α i β.

(DP) $W(\alpha) > W(\beta) \Leftrightarrow \exists N = N(\alpha, \beta)$ takie, że jeśli $K > N$, to $W(\alpha^K, \gamma_K) > W(\beta^K, \delta_K)$ „dla każdej pary γ, δ – strumieni z **Ω**", gdzie ciąg (α^K, γ^K) powstaje przez „zestawienie *K*-obcięcia elementu α z *K*-ogonem elementu β, co „wypełnia ich segmenty złożone z zer".

(χ) Kryterium *W* ustanawia dyktaturę przyszłości (DF), jeśli nie zależy od elementów z ***P*(Ω)**, jest „wrażliwe" tylko na ***F*(Ω)**. Formalnie: dla dowolnych dwóch strumieni $\alpha, \beta \in$ **Ω**.

(DF) $W(\alpha) > W(\beta) \Leftrightarrow \exists N = N(\alpha, \beta)$, takie, że jeśli $K > N$, to $W(\gamma^K, \alpha_K) > W(\delta^K, \beta_K)$ dla każdej pary γ, δ – strumieni z **Ω**.

AKSJOMATY CHICHILNISKY (ACH)

Preferencje w **Ω** są indukowane przez SWF (kryterium *W*), która nie może mieć własności DP ani DF.

DEFINICJA 2 [Chichilnisky 1996]

Preferencje:

(κ) które spełniają ACH,

(λ) których reprezentacja rzeczywista $W : \mathbf{\Omega} \to R$ jest funkcją rosnącą – „słabo paretowską" (spełnia WAP), nazywa się preferencjami trwałego (zrównoważonego) rozwoju (*sustainable preferences*, SP).

TWIERDZENIE [Chichilnisky 1996]

Istnieje relacja SP na przestrzeni F i jest ona reprezentowana przez kryterium W_{SP}^{CH} określone wzorem:

$$W_{SP}^{CH}(u) = \sum_{g=1}^{\infty} \lambda_g u_g + \Phi(u), \tag{16}$$

gdzie $\lambda_g > 0, \ \forall_g \in N_+, \ \sum_{g=1}^{\infty} \lambda_g < \infty,$ a uogólniona granica $\Phi(u) = \lim_g(u_g)$ jest rozszerzeniem do l_∞ (via twierdzenie Hahna-Banacha) funkcjonału „zwykłej granicy".

KRYTERIUM CHICHILNISKY (ważonej sumy DU oraz GL użyteczności przyszłych pokoleń)

Rodzina kryteriów – funkcjonałów dyskontujących, zdefiniowanych wzorami:

$$W_{SP}^{CH}(\alpha;u) = \alpha \sum_{g=0}^{\infty} \frac{u_g}{(1+r)^g} + (1-\alpha)\lim_g(u_g); \ \ \alpha \in \langle 0;1 \rangle \tag{17}$$

indukuje w $\boldsymbol{\Omega}$ preferencje trwałego rozwoju.

Komentarze:

1) „Filozofię dyskontowania" zawartą w propozycji G. Chichilnisky można (umownie) określić jako „zdyskontowany samuelsonowski utylitaryzm z szumem przyszłości".

2) Arbitralność doboru wag (α) ma swoje zalety („elastyczność") oraz wady tkwiące w każdym przypadku dopuszczającym arbitralność.

3) Alternatywne podejście, uwzględniające sprawiedliwość typu HEF, zaprezentowali w latach 2001, 2009, 2010, 2012 Asheim i inni. Syntezę ich koncepcji stanowią idee rekurencyjnych funkcjonałów dobrobytu społecznego [Asheim 2010b] oraz „zrównoważonego dyskontowania utylitarystycznego", skwantyfikowane w formułach tzw. SDU SWF podanych w pracy [Asheim, Banerjee 2010], w której podano także (konstruktywny) dowód istnienia takich relacji i ich reprezentacji – spełniają one m.in. obydwa postulaty Chichilnisky (NDP, NDF) oraz postulat HEF.

W nurt ten wpisuje się także wzmiankowana wyżej praca [Alcantud, Garcia-Sanz 2010], która również zawiera algorytm wyceny. Konstrukcję kryterium dobrobytu (w kontekście sprawiedliwości międzypokoleniowej) podał też w pracy z 2010 r. T. Sakai. Jedną z najnowszych prac z omawianego nurtu jest artykuł

z 2012 r. autorstwa Asheima i Zuber [2012], w którym zastosowano bardzo interesującą, ,,egalitaryzującą" modyfikację koncepcji RDEU (*rank depend expected utility, anticipated utility*) J. Quiggina i M. Yaari'ego z lat osiemdziesiątych ubiegłego wieku [Quiggin 1982; Yaari 1987] – „ważenia" użyteczności kardynalnej (stanowiącej jądro całkowego funkcjonału oczekiwanej użyteczności) czynnikiem funkcyjnym, „deformującym" jej wartości w zależności od względnej ich wartości (pozycji, rang). „Rangowanie" Asheima-Zuber wyraża się w przyporządkowaniu „klasycznych" (geometrycznych) wag dyskontowych wyrazom uporządkowanego w rosnącej kolejności wycenionego ciągu użyteczności, a nie w jego „oryginalnym" porządku (w czasie). Zasługą autorów jest m.in. uogólnienie, na przestrzeń nieskończonych ciągów, metodyki konstrukcji funkcji dobrobytu społecznego J. Weymarka oraz U. Eberta z lat 1981 oraz 1988, odpowiednio [Weymark 1981; Ebert 1988], a'la Gini. Indukowane preporządki są częściowe, bo nie każdy ciąg nieskończony (użyteczności) daje się efektywnie przestawić do wymaganej przez ich kryterium postaci.

4. Dyskontowanie: paradygmat P. Samuelsona v. rzeczywistość psychologii i ekonomia przetrwania

„Pęknięcie" w spójnej i eleganckiej metodyce DU, dostrzeżone przez R. Strotza (także przez psychologów i biologów [Herrnstein 1961; Ainslie 1974]) spowodowało nie tylko intensyfikację badań w sferze cyzelowania własności preporządków w przestrzeniach ciągów użyteczności, ale przede wszystkim – w obrębie weryfikacji poczynionych spostrzeżeń niezgodności dynamicznej zachowań podmiotów i identyfikacji jej źródeł – poszerzenia spektrum. Głównym „konkurentem filozofii geometryczno-wykładniczej" (funkcji dyskontującej) okazało się tzw. dyskontowanie hiperboliczne. Idea tej formalizacji (popartej dokumentacją empiryczną) wyraża się w odwrotnej proporcjonalności wartości dyskonta i długości opóźnienia (czyli dystansu czasowego dzielącego „teraźniejszość" – aktualny moment wyceny od momentu wystąpienia ewaluowanego zdarzenia).

Poniżej przytoczono wybrane koncepcje formalne o „rodowodzie" ekonomicznym oraz psychologicznym (uwzględniające element niezgodności czasowej). W nurcie ekonomicznym za pionierskie (a zarazem mające znamiona uniwersalizmu) można uznać prace E. Phelpsa i R. Pollaka z 1968 r. oraz F. Kydlanda i J. Prescota z 1977 r. Z kolei w badaniach „zakotwiczonych" w psychologii i biologii eksponowano (i próbowano wyjaśniać) tzw. anomalie (niekonsekwencje, sprzeczności rozmaitej natury) wyboru międzyokresowego [Rachlin, Green 1972; Ainslie

1974; Thaler 1981; Mazur 1987; Prelec 2004]. Phelps i Pollak [1968] sięgają do modelowania teoriogrowego (gier powtarzalnych) i proponują stosować tzw. quasi-geometryczną (quasi-hiperboliczną) regułę dyskontowania. Czynnik dyskontujący $d(t)$ występujący w ogólnej addytywnej reprezentacji funkcjonału dyskontującego ma postać:

$$d(t) = \begin{cases} 1 & \text{dla} \quad t = 0 \\ \beta\delta^t & \text{dla} \quad t \in N_t; \quad \delta > 0; \; \beta \in \langle 0;1 \rangle. \end{cases} \tag{18}$$

Warto zwrócić uwagę na dwa aspekty podanej reguły. Po pierwsze: jednookresowy czynnik dyskontujący (miara „niecierpliwości") dla pierwszego okresu, który zaczyna się w chwili $t = 0$, a kończy w $t = 1$, jest równy $\beta\delta$, a dla pozostałych – δ. Od drugiego okresu (włącznie) ta stopa jest stała, większa niż w pierwszym. Drugim jest niezgodność czasowa – występuje brak tzw. efektu stopera charakterystycznego dla stacjonarnego dyskontowania: dyskonto okresu $\langle t; t + 1\rangle$ w chwili zerowej jest „u Samuelsona" takie samo, jak wówczas, gdy w chwili t „cofnie się czas do zera" i dyskontuje ten okres (jako pierwszy). Teraz „ponowne włączenie stopera" w chwili t zmienia (wyjściową) ewaluację niecierpliwości z δ na $\beta\delta$.

Niektóre „paradoksy behawioralne" znacznie wcześniej zostały zaobserwowane w kontekście teorii wyboru w warunkach ryzyka oraz niepewności (prace z lat pięćdziesiątych i sześćdziesiątych XX w., spostrzeżenia Alaisa, Ellsberga i in.). Odnotowane odstępstwa do normatywów teorii oczekiwanej użyteczności oraz konstruktywne próby jej sanacji od końca lat siedemdziesiątych [Kahneman, Tversky 1979; Quiggin 1982; Machina 1982; Yaari 1987] inspirowały ideę poszukiwania niezgodności (w mniejszym stopniu metodę ich kwantyfikacji). Propagatorami „nowego" dyskontowania byli Loewenstein, Prelec [1992] oraz Laibson [1997].

Uogólniona hiperboliczna funkcja dyskontująca jest funkcją zależną od trzech parametrów dodatnich γ, δ, μ:

$$d(t) = (1 + \mu t)^{-\frac{\gamma}{\delta}}, \tag{19}$$

a jej uproszczony wariant (autorstwa Laibsona z 1997 r.):

$$d(t) = (1 + \mu t)^{-\gamma} \tag{20}$$

już tylko od dwóch.

Ważną rolę odegrały prace Ch. Harveya z 1986 i 1995 r. oraz artykuły Fishburna i Rubinsteina z 1981 r., a także Fishburna z Ewardsem z 1997 r. W pracach tych podjęto (skuteczną) próbę wyprowadzenia ogólnych hiperbolicznych reguł kalkulacyjnych z „postulatów pierwotnych" (aksjomatów) sformułowanych „pod adresem" preferencji w zbiorach tzw. konsekwencji, czyli par („wynik", „czas"):

$$(\boldsymbol{x}, \boldsymbol{t}) = \{(x_1, t_1),\ (x_2, t_2),\ ....\}\ . \tag{21}$$

Przytoczymy kilka przykładowych postulatów z pracy Harveya [1995] i ich analityczne konsekwencje dla funkcji $d(t)$. Punktem wyjścia jest wyrażenie „postawy podmiotu względem czasu" przez ogólne własności funkcji $d(t)$ – z równoczesną charakteryzacją tych postaw w języku preferencji w zbiorze dwuwymiarowych strumieni $(\boldsymbol{x}, \boldsymbol{t})$. I tak: preferencje podmiotu wyrażają jego „niechęć względem upływu czasu" (są „timing averse"), gdy funkcja dyskontująca jest malejąca; „skłonność" (jest „timing pron"), gdy $d(t)$ rośnie oraz jest neutralny, gdy jest stała (w innej terminologii neutralność jest nazywana cierpliwością, a „skłonność" wiąże się z międzypokoleniowym altruizmem [Rybicki 2009, 2010b]. Awersja (podmiotów) względem opóźnień jest definiowana w języku relacji zachodzących między wyrazami strumieni warunkiem: $(x,s) \succ (x,t)$ dla dowolnego wyniku x oraz chwil $s < t$. Własność stacjonarności preferencji określa się implikacją:

$$(\boldsymbol{x},\boldsymbol{s}) \sim (\boldsymbol{y},\boldsymbol{t}) \Leftrightarrow \forall\, \mathrm{h} > 0\ (\boldsymbol{x},\boldsymbol{t}+h) \sim (\boldsymbol{y},\boldsymbol{t}+h)\ , \tag{22}$$

gdzie symbol $\boldsymbol{t}' + h$ oznacza ciąg równości $t'_n = t_n + h\ \ \forall n \in N$.

A oto tzw. bezwzględna malejąca awersja względem czasu:

$$(x,s) \sim (y,t) \Leftrightarrow (x, s+h) \prec (y, t+h)\ \ \forall h > 0, \tag{23}$$

jeśli $s < t$ oraz $0 < x < y$.

Logarytm (malejącej) funkcji dyskontującej $g(t) = \ln d(t)$ musi być wówczas funkcją ściśle wypukłą. Harvey [1995] udowodnił, że jest to warunek konieczny i wystarczający, skąd wynika, że dyskontowanie wykładnicze tego postulatu nie spełnia, czyli dyskonta stacjonarnego nie znamionuje malejąca awersja względem czasu.

Ważną rolę w konstrukcjach Ch. Harveya spełniają postulaty tzw. względnej czasowej neutralności, rozciągliwości („stretching") oraz „proporcjonalności – w stosunku do czasów opóźnień" (timing proportional preferences, TPP [Harvey 1986, 1995]. Stanowią one istotny segment aksjomatycznego umocowania formalnego stworzonego przez niego dla uzasadnienia „filozofii dyskon-

towania" hiperbolicznego. Oto pierwszy z nich, nazwany przez Ch. Harveya własnością „relative timing constant" (RTC preferences):

$$\begin{gathered}\exists b>0:(x,s)\sim(y,l)\Rightarrow\\ \forall m>0\ (x,s+m(b+s))\sim(y,t+m(b+t)).\end{gathered}\qquad \text{(RTC) (24)}$$

Występujące powyżej relacje indyferencji odnoszą się do „chwilowych" (liczbowych) par, tzw. konsekwencji (wynik, czas realizacji). We wcześniejszej pracy (z 1986 r.) cytowanego autora jest wprowadzone pojęcie rozciągliwości (nieco inaczej formalizujące ideę (RTC) – przygotowujące grunt pod koncepcję TPP: podmiot przedkładający wypłatę x w chwili s nad wypłatę y w chwili t dokona takiego samego wyboru dla okresów, odpowiednio, $D(s + 1) - 1$ oraz $D(t + 1) - 1$, gdzie D oznacza opóźnienie:

$$(x,s) \succ (x,t) \Rightarrow (x, D(s + 1) - 1) \succ (y, D(t + 1) - 1). \qquad (25)$$

Istotę owych preferencji wyjaśnia w swoim artykule sam Harvey, a przybliża ją, na sugestywnym przykładzie, praca Cieślak [2003].

„Odmiennie niż w przypadku aksjomatu stacjonarności, rozciągliwość jest wrażliwa na przesunięcie w czasie. Wyobraźmy sobie osobę, która woli 100 j.p. (jednostek pieniężnych) dziś od 110 j.p. jutro. Rozciągliwość oznacza, że preferencje te nie zmieniają się w sytuacji wyboru między 100 j.p. za 10 dni, a 110 j.p. za 21 dni (parametr $d = 11$). Jednak ta sama osoba, postawiona przed alternatywą 100 j.p. za 10 dni lub 100 j.p. za 11 dni, może zmienić swe upodobania na korzyść wypłaty wyższej i późniejszej". Upraszczając wywody pracy Harveya, można przytoczyć jedną z jej konkluzji: w addytywnej reprezentacji funkcjonału wyceny, wagi mają postać funkcyjną $d(t) = \left(\frac{b}{b+t}\right)^r$; $t > 0$ dla pewnych wartości parametrów $b > 0$ oraz $-\infty < t < \infty$, wtedy i tylko wtedy, gdy preferencje mają własność RTC. Dalsze specyfikacje $d(t)$ otrzymuje się jako konsekwencję postulatu TPP. Symbolicznie:

$$\begin{gathered}\forall x,y>0 \quad \forall s_0,t_0\\ (x,s_0)\sim(y,t_0);\ \Delta s>0,\ \Delta t>0,\quad \frac{\Delta t}{\Delta s}=\frac{y}{x}\\ \Rightarrow (x,\ s_0+\Delta s)\sim(y,\ t_0+\Delta t).\end{gathered}\qquad \text{(TPP) (26)}$$

Harvey [1995] wykazał, że tak właśnie sformułowanemu postulatowi czynią zadość reprezentacje preferencji neutralnych czasowo oraz preferencje o „powolnie malejącej awersji do czasu", w których funkcjonał dyskontujący określają wagi hiperboliczne:

$$d(t) = \frac{b}{b+t} \tag{27}$$

(jest to jednoparametryczna rodzina funkcji charakteryzująca się malejącą stopą niecierpliwości – czasowej).

Podsumowanie

Koncepcja dyskontowania hiperbolicznego „zadomowiła się" w ostatnim dwudziestoleciu w obszarze badawczym wyceny programów dalekookresowych, zyskując status najnowszej klasyki przedmiotu. Różnorodność potencjalnych kontekstów rzeczywistych przerasta jednak możliwości deskryptywne tej kategorii funkcji. Poniżej zasygnalizowano niektóre alternatywne propozycje ujęcia zagadnień dyskontowania. Ważnym „tropem" jest podążanie za analogiami z dziedziny teorii podejmowania decyzji w warunkach ryzyka (i niepewności). Najogólniej rzecz sprowadza się do „wyznaczenia osi czasu roli osi stanów" [Prelec, Loewenstein 1989; Quiggin, Horowitz 1995; Fishburn, Edwards 1997]. Przykładem tego zabiegu jest reinterpretacja paradoksu Petersburskiego przedstawiona w pracy Fishburna i Edwardsa z 1997 r. Prawdopodobieństwa realizacji wypłat (poprzedzonych seriami porażek) można traktować jako wagi dyskontowe. Klasyczny paradygmat oczekiwanej użyteczności (expected utility – EU) ma naturalny odpowiednik w „wersji dynamicznej" (deterministycznej) – „równie klasyczny" paradygmat DU. Można to symbolicznie zapisać w postaci podwójnej równości:

$$EU = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \cdot u(2^k) = DU. \tag{28}$$

Niektóre koncepcje eksploatowania tej analogii zreferowano w cytowanych pracach. W tym miejscu ograniczymy się do przedstawienia kilku niesformalizowanych spostrzeżeń oddających główne idee tego ujęcia. Rodzaje i siła dyskontowania odpowiadają rodzajom postaw (podmiotów) w stosunku do ryzyka. W nomenklaturze teorii wyboru międzyokresowego i relacji międzypokoleniowych kategoria awersji do ryzyka została zastąpiona pojęciem niecierpliwości i rozważa

się kwestie zmian tego czynnika (w czasie: stałość, malenie, wzrost, tempo malenia [Bleichrodt i in. 2008; Prelec 2004; Price 2004]). Te (lokalne, w czasie) własności preferencji w przestrzeni strumieni można wyrażać miarami zależącymi od własności funkcji dyskontujących. W pierwszej kolejności wchodzi tu porównywanie stopnia wypukłości i wskaźniki wzorowane na „logarytmicznej wypukłości (wklęsłości)" miar awersji do ryzyka Arrowa-Prata. Już „zwykła" wypukłość" malejącej funkcji dyskontującej $f(t)$ informuje o silniejszym dyskontowaniu w okresach bliższych „teraźniejszości" niż w dalszej przyszłości. Inną możliwością jest badanie zachowania chwilowych stóp dyskontowych w postaci $-\frac{f'(t)}{f(t)}$ (czas ciągły) i porównywanie tempa malenia tych funkcji (pojawiają się tu klasy równoważności funkcji o tym samym stopniu niestacjonarności); podobnie można wykorzystać porównywanie stopnia wypukłości logarytmów funkcji dyskonta. Prelec [2004] wprowadza tzw. funkcję niecierpliwości mającą postać (dla czasu ciągłego):

$$g(t)=\begin{cases}\dfrac{f'(t)}{f'(o)} & t>0\\ 1 & t=0\end{cases} \tag{29}$$

i zauważa zależność $g(t) \le f(t)$, przy czym równość tych funkcji ma miejsce tylko w przypadku stacjonarnej (wykładniczej) formuły dyskontowej f.

Bleichrodt i in. [2008] definiują dwa wskaźniki niecierpliwości (o charakterze absolutnym oraz relatywnym, odpowiednio). Są to „miary różniczkowe":

$$\gamma(t)=-\frac{[\ln\varphi(t)]''}{[\ln\varphi(t)]'} \tag{30}$$

oraz:

$$\delta(t)=t\cdot\gamma(t)\ . \tag{31}$$

Typologia dyskontowania zaproponowana przez tych autorów opiera się na własnościach tych miar i prowadzi do rozważenia, innych niż hiperboliczne, funkcji dyskonta (por. [Bleichrodt i in. 2008].

Pogłębioną analizę uwarunkowań decydentów zależnych od: wielkości wypłaty, miary probabilistycznej (prawdopodobieństwa subiektywnego podmiotów) oraz parametru czasowego przeprowadza Read [2001]. Podstawową konkluzją jego artykułu jest supozycja o tzw. subaddytywnym mechanizmie dyskontowania: dyskontowanie wielkości „rozłożony na raty" (wielokrotnie – w danym okresie) prowadzi do większej wartości dyskonta niż w przypadku aktów jednorazowych

(w skali całego tego okresu). Sama idea (ilustrowana przykładami) jest pokrewna koncepcji nieaddytywnych miar probabilistycznych, sięgającej G. Choqueta, rozwiniętej w kontekście „statycznej" teorii podejmowania decyzji w warunkach niepewności, przede wszystkim przez D. Schmeidlera [1987]. Istotne dla tej filozofii dyskontowania okazuje się subtelne rozróżnienie zaakcentowane przez Reada [2001] – między opóźnieniem („bezwzględnym" – delay) a długością okresu dyskontowania (interval). Okazuje się, że obydwie „konkurencyjne" formuły (wykładnicza oraz hiperboliczna) mają wspólną cechę – „dyskontują addytywnie".

W przygotowywanej pracy [Rybicki 2014] proponuje się rozważenie „nowego benchmarku" dla funkcji dyskontujących: proporcjonalnych do gęstości rozkładów logarytmiczno-normalnych. Charakteryzują się one wolniejszym tempem malenia niż wykładnicze, czyli w mniejszym stopniu „dyskryminują przyszłość". Należą do szerszej rodziny funkcji podwykładniczych, które są funkcjami o tzw. grubych (ciężkich) „ogonach" (heavy-tailed). Z kolei maleją one szybciej niż funkcje hiperboliczne. Mają własność malejącej niecierpliwości, można efektywnie obliczyć odpowiednie miary (a także ich dyskretne odpowiedniki) i badać asymptotykę. Ich „pośrednie" usytuowanie (pomiędzy „skrajnościami") nasuwa podejrzenie, że mogą one charakteryzować preferencje, w jakimś sensie „podstacjonarne", w zbiorach programów (klasę dotychczas niebadaną i niezidentyfikowaną empirycznie).

Wydaje się, że powyższy – pobieżny – przegląd podstawowych aspektów problematyki porównań i wyceny strumieni wielkości ekonomicznych w czasie może dać pewien pogląd na całokształt tych zagadnień, uwypuklając ich „wielobarwność" i znaczenie zarówno teoretyczno-modelowe, jak i w sferze *Praxis*.

Bibliografia

Ainslie G., 1974: Impulse Control in Pigeons. „Journal of Experimental Analysis of Behavior", 21, 485-489.

Allais M., 1952: Fondaments d'une théorie des choix compartant un risqué et critique des postulates et axioms de l'ecole Américaine. International Conference on Risk. Centre National de La Recherche Scientifique, May, Colleges International XL, Paris.

Alcantud J., Garcia-Sanz M., 2010: Paretian Evaluation of Infinite Utility Streams: An Egalitarian Criterion. „Economic Letters", No. 106, 209-211.

Arrow K.J., 1999: Discounting, Morality, and Gaming. W: Discounting and Intergenerational Equity. P.R., Portney, J.P. Weyant (eds.). Resources for the Future. Washington, DC.

Asheim G., 1996: Ethical Preferences in the Presence of Resource Constraints. „Nordic Journal of Political Economy", No. 23, 55-67.

Asheim i in., 2001: Justifying Sustainability. „Journal of Environ. Econom. Management", No. 41, 252-268.

Asheim G. i in., 2010: Generalized Time-invariant Overtaking. „Journal of Mathem. Economics", No. 46, 519-533.

Asheim G. i in., 2010: Sustainable Recursive Social Welfare Functions. „Econ. Theory", No. 16, 183-219.

Asheim G., Banerjee K., 2010: Fixed-step Anonymous Overtaking and Catching Up. „International Journal of Economic Theory", No. 6, 149-165.

Asheim G., Mitra T., 2010: Sustainability and Discounted Utilitarianism in Models of Economic Growth. „Math. Soc. Sciences", No. 59, 148-169.

Asheim G., Zuber S., 2012: Justifying Social Discounting: The Rank-discounted Utilitarian Approach. „Journal of Economic Theory", No. 147, 1572-1601.

Banerjee K., 2006: On the Equity-efficiency Trade-off in Aggregatiing Utility Streams. „Economic Letters", No. 93, 63-67.

Banerjee K., Mitra T., 2007: On the Impatience Implications of Paretian Social Welfare Functions. „Journal of Math. Economics", Vol. 43, 236-248.

Basu K., Mitra T., 2002: Aggregating Infinite Utility Streams with Inter-Generational Equity. MIT Department of Economics Working Paper, No. 02-19.

Basu K., Mitra T., 2003: Aggregating Infinite Streams with Intergenerational Equity: The Impossibility Being Paretian. ,,Econometrica", No. 32, 1557-1563.

Berbeka K., 2008: Problemy stosowania międzypokoleniowej stopy dyskontowej. ,,Ekonomista", 233-242.

Bleichrod H., Rohde K., Wakker P., 2009: Non-hyperbolic Time Inconsistency. ,,Game and Economic Behavior", No. 66, 27-38.

Bleichrodt H. i in., 2008: Koopmans' Constant Discounting for Inter-temporal Choice: A Simplification and a Generalization. „Journal of Mathematical Psychology", No. 52, 341-347.

Bossert W. i in., 2007: Utilitarianism for Infinite Utility Streams: A New Welfare Criterion and Its Axiomatic Characterization. „Journal of Economic Theory", No. 135(1), 579-589.

Brock W., 1970: An Axiomatics Basis for the Ramsey-Weizsecker Overtaking Criterion. „Econometrica", No. 38, 927-929.

Chichilnisky G., 1977: Economic Development and Efficiency Criteria in the Satisfaction of Basic Needs. „Applied Math. Modelling", No. 1(6), 290-297.

Chichilnisky G., 1997: What Is Sustainable Development. „Land Economics", No. 74(4), 467-491.

Chichilnisky G., 1996: An Axiomatic Approach to Sustainable Development. ,,Social Choice and Welfare", No. 13, 231-257.

Cieślak A., 2003: Behawioralna ekonomia finansowa. Modyfikacja paradygmatów funkcjonujących w nowoczesnej teorii finansów. Materiały i Studia NBP, z. 165, Warszawa, październik.

Cowen T., 1997: Discounting and Restitution. Philosophy and Public Affairs, Spring, No. 26, 2, 168-185.

Czerwiński Z., 1992: Jak porównać jutrzejszy dobrobyt z dzisiejszą biedą czyli o optymalnej intensywności inwestowania. W: Dylematy ekonomiczne. PWE, Warszawa.

Dasgupta P., Heal G.M., 1974: The Optimal Depletion of Exhaustible Resources. „Review of Economic Studies", No. 38 (Symposium), 331-339.

Dasgupta P., Maskin E., 2005: Uncertainly and Hyperbolic Discounting. „The Americal Economic Review", No. 95(4), 1290-1299.

Diamond P., 1965: The Evaluation of Infinite Utility Streams. „Econometrica", No. 33, 170-177.

Ebert U., 1988: Measurement of Inequality: An Attempt at Unification and Generalization. „Soc. Choice and Welfare", No. 5, 147-169.

Ellsberg D., 1961: Risk. Ambiquity and Savage Axiomic. „Quarterly Journal of Economics", Vol. 75, 643-669.

Fiedor B., 2007: Wzrost zrównoważony w ekonomii głównego nurtu i w ujęciu ekonomii środowiska. W: Zrównoważony rozwój w teorii ekonomii i w praktyce. A. Graczyk (red.). Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu nr 1190, 40-64.

Fishburn P., Rubinstein A., 1982: Time Preferences. „Internat. Economics Review", Vol. 23, 677-694.

Fishburn P., Edwards W., 1997: Discount Neutral Utility Models for Denumerable Time Streams. „Theory and Decision", No. 43, 136-167.

Fleurbaey M., Michel P., 2003: Intertemporal Equity and Extensions of the Ramsey Criterion. „Journal of Math. Economics", Vol. 39, 777-802.

Gollier R., Zeckhauser R., 2005: Aggregation of Heterogeneous Preferences. „Journal of Political Economy", No. 113(4), 878-896.

Gollier Ch., 2008: Discounting with Fat-tailed Economic Growth. „Journal of Risk and Uncertainty", No. 37, 171-186.

Gollier C., 2010: Ecological Discounting. „Journal of Economic Theory", No. 145, 812-829.

Hammond P.J., 1976: Equity, Arrow's Conditions and Rawl's Difference Principle. „Econometrica", No. 44, 793-804.

Harvey Ch., 1995: Proportional Discounting of Future Costs and Benefits. „Mathemtics of Operation Research", No. 20, 381-399.

Harvey Ch., 1986: Value Functions for Infinite Period Planning. „Management Science", No. 32, 1123-1139.

Herrnstein R., 1961: Relative and Absolute Strength of Response as a Function of Frequency of Reinforcements. „Journal of Exper. and. Behav.", No. 4, 267-272.

Jech T., 1978: Set Theory. Academic Press, New York.

Kahneman D., Tversky A., 1979: Prospect Theory: An Analysis of Decision under Risk. „Econometrica", Vol. 47, No. 2, 363-391.

Kirby K.N., Herrnstein R.J., 1995: Preference Reversals Due to Myopic Discounting of Delayed Reward. „Psychological Science", No. 6(2), 83-89.

Koopmans T.C., 1960: Stationary Ordinal Utility and Impatience. „Econometrica", No. 28, 287-309.

Krusell P., Smith A., 2003: Consumption – Savings Decisions with Quasi-geometric Discounting. „Econometrica", No. 71, 365-375.

Kulpa W., 2009: Topologia a ekonomia. Wydawnictwo Kardynała Stefana Wyszyńskiego, Warszawa.

Kydland F., Prescot E., 1977: Rules Rather than Discretion: The Inconsistency of Optimal Plans. „Journal of Political Economy", No. 85(3), 473-491.

Laibson D., 1997: Golden Eggs and Hyperbolic Discounting. „Quarterly Journal of Economics", No. 62, 443-479.

Lauwers L., 1995: Time – Neutrality and Linearity. „Journal of Mathematical Economics", No. 24, 342-351.

Lauwers L., Liedekerke L.V., 1997: Sacrificing the Patrol: Utilitarianism, Future Generations and Infinity. „Economics and Philosophy", No. 13, 159-174.

Lauwers L., 1998: Intertemporal Objective Functions. Strong Pareto versus Anonymity. „Math. Social Sciences", Vol. 35, 37-55.

Lauwers L., 2010: Ordering Infinite Utility Streams Comes of the Cost of a non-Ramsey Set. „Journal of Mathematical Economics", No. 46, 32-37.

Loewenstein G., Prelec D., 1992: Anomalies in Intertemporal Choice: Evidence and an Interpretation. „Quarterly Journal of Economics", No. 107, 573-597.

Machina M., 1983: Expected Utility Analysis without Independence Axiom. „Econometrica", No. 50, 227-323.

Mathematical Social Sciences 2010, No. 59 (Special Issue on Sustainability).

Mathias A., 1977: Happy Families. „Annals of Pure and Applied Logic", No. 12, 59-111.

Mazur J., 1987: An Adjusting Procedure for Studying Delayed Reinforcement. W: M. Commins, J. Mazur, J. Nevin, H. Rachlin (eds.). „Quantitative Analyses of Behaviour", No. 5, 55-73.

Neumann J. von, Morgenstern O., 1944: Theory of Games and Economic Behavior. „Princeton Univ. Press", Princeton, NJ.

Nocetti D. i in., 2008: Properties of the Social Discount Rate in a Benthamite Framework with Heterogeneous Degrees of Impatience. „Management Science", No. 54(10), 1882-1886.

O'Donog'hue T., Rabin M., 1999: Doing It Now or Later. „The Americal Economic Review", No. 89(1), 103-124.

Phelps E., Pollak R.A., 1968: On Second-Best National Saving and Game-Equilibrium Growth. „Review of Economic Studies", No. 35(2).

Pitman E., 1980: Sub-exponential Distribution Functions. „Journal of Australian Math. Society" (Series A), No. 29, 337-347.

Portney P.R., Weyant J.P. (eds.), 1999: Discounting and Intergenerational Equity. „Resources for the Future" Washingon, DC.

Prelec D., 2004: Decreasing Impatience: A Criterion for Non-stationary Time Preference and „Hyperbolic" Discounting. „Scand. Journal of Economics", No. 106, 511-532.

Prelec D., Loewenstein G., 1991: Decision-making over Time and under Uncertainty: A Common Approach. „Management Science", No. 37, 770-786.

Price C., 2004: Hyperbole, Hypocrisy and Discounting that Slowly Fades Away. „Scandinavian Forest Economics", No. 40, 343-59.

Price C., 2005: How Sustainable Is Discounting? W: S. Kant, R.A. Berry (eds): Economics, Natural Resources, and Sustainability: Economics of Sustainable Forest Management. Kluwer, Dordrecht, 106-35.

Quiggin J., 1982: A Theory of Anticipated Utility. „Journal of Econ. Behavior and Organ.", No. 3, 323-344.

Quiggin J., Horowitz J., 1995: Time and Risk. „Journal of Risk and Uncertainty", No. 10, 37-55.

Rachlin H., Green L., 1972: Commitment, Choice and Self-Control. „Journal of Experimental Analysis and Behavior", No. 17, 15-22.

Ramsey F., 1928b: On a Problem of Formal Logic. „Proceedings of the London Mathematical Society", Series 2, 30.4, 339-384.

Ramsey F.P., 1928: A Mathematical Theory of Saving. „Economics Journal", No. 38, 543-559.

Rawls J., 1970: The Theory of Justice. Oxford University Press, Oxford.

Read D., 2001: Is Time-Discounting Hyperbolic or Subadditive? „Journal of Risk and Uncertainty", No. 23, 5-32.

Reinschmidt K., 2002: Aggregate Social Discount Rate Derived from Individual Discount Rates. „Management Science", No. 48(2), 307-312.

Robson A., 2001: The Biological Basis of Economic Behavior. „Journal of Econ. Literature", No. 39, 11-23.

Roemer J., Suzumura K., 2007: Intergenerational Equity and Sustainability. Palgrave Macmillan, Houndmils-Basingstoke-Hampshire.

Rybicki W., 2009: Sprawiedliwość międzypokoleniowa i sekwencyjne modelowanie preferencji. W: B. Fiedor, Z. Hockuba (red.): Nauki ekonomiczne wobec wyzwań współczesności. PTE, Warszawa, 175-187.

Rybicki W., 2010a: O realokacji dóbr i sprawiedliwości międzypokoleniowej. W: M. Kulesza, W. Ostasiewicz (red.): Pragmata Tes Oikonomias. AJD, Częstochowa, Vol. IV, 110-132.

Rybicki W., 2010b: O sprawiedliwości międzypokoleniowej (nota bibliograficzna). W: J. Pociecha (red.): Aktualne zagadnienia modelowania i prognozowania zjawisk społeczno-gospodarczych. Studia i Prace Uniwersytetu Ekonomicznego w Krakowie nr 10, 141-155.

Rybicki W., 2011: Sprawiedliwość międzypokoleniowa i „kłopoty z dyskontowaniem przyszłości". W: Modelowanie i prognozowanie gospodarki narodowej. Prace i materiały Wydziału Zarządzania Uniwersytetu Gdańskiego, z. 4/8, 151-165.

Rybicki W., 2012: Discounting and Ideas of Intergenerational Equity and Sustainability. „Operations Research and Decisions", No. 22(1), 63-84.

Rybicki W., 2014: O preferencjach etycznych i wolno malejących funkcjach dyskontujących. Artykuł przygotowywany do druku w kwartalniku „Ekonomista".

Sakai T., 2010: Intergenerational Equity and Explicit Construction of Welfare Criteria. „Social Choice and Welfare", No. 35, 393-414.

Samuelson P., 1937: A Note on Measurement of Utility. „Review of Economic Studies", IV, 2, 155-161.

Savage L., 1954: The Foundations of Statistics. J. Wiley&Sons, New York.

Sen A.K., 1988: Freedom of Choice. Concept and Contents. „Eur. Econ. Review", Vol. 32, 269-294.

Schmeidler D., 1987: Subjective Probability and Expected Utility without Additivity. „Econometrica", No. 57, 571-587.

Sidgwick H., 1907: The Methods of Ethics. Macmillan, London (7th edition).

Sikora R.J., Barry B., 1978: Obligations to Future Generations. Temple, Philadelphia.

Solow R.M., 1974: Intergenerational Equity and Exhaustible Resources. „The Review of Economic Studies", Symposium, 29-45.

Special Issue on Discounting Dilemmas. „Journal of Risk and Uncertainty", Vol. 37, No. 2/3.

Strotz R.M., 1955: Myopia and Inconsistency in Dynamic Utility Maximization. „Rev. Econ. Studies", No. 23, 165-180.

Svensson L., 1980: Equity among Generations. „Econometrica", No. 48, 1251-1256.

Szpilrajn E. (Marczewski E.), 1930: Sur l'extension de l'odre partiel. „Fundamenta Mathematicae", No. 16, 386-389.

Thaler R., 1981: Some Empirical Evidence on Dynamic Inconsistency. „Econ. Letters", No. 8, 201-207.

Weitzman M., 2007: A Review of the Stern Review of the Economics of Climate Change. „Journal of Econ. Literature", No. 45, 703-724.

Weitzman M., 2001: Gamma Discounting. „American Economic Review", No. 91, 260-271.

Weitzman M., 1998: Why the Far – Distant Future Should Be Discounted at Its Lowest Possible Rate. „Journal of Environmental Economic and Management", No. 36, 201-208.

Weizsäcker von C.C., 1965: Existence of Optimal Programs of Accumulation for an Infinite Time Horizon. „Review of Economic Studies", No. 32, 85-104.

Weymark J., 1981: Generalized Gini Inequality Indices. „Mathemat. Social Sciences", No. 1, 409-430.

Yaari M., 1987: The Dual Theory of Choice under Risk. „Econometrica", No. 55, 95-115.

Zame W.R., 2007: Can Intergenerational Equity Be Operationalized. „Theoretical Economics", No. 2, 187-202.

Żylicz T., 2001: Wiara i nauka a sprawiedliwość międzypokoleniowa. W: U progu trzeciego tysiąclecia. A. Białecka, J.J. Jadacki (red.). Wyd. Naukowe Semper, Warszawa.

PREFERENCE'S MODELING AND PROBLEMS OF SUSTAINABLE DEVELOPMENT

Summary

In the paper the selected review of main research directions and basic statements from the area of evaluations and comparisons of long-term project is given as well as – questions concerning an adjustment of discounting functions to the "true" fluctuations of individual preferences (in the passage of time) are asked. The above research is carried out in two (complementary) planes. At first the analyses of preorders in sets of infinite sequences of (instantaneous) utilities are shortly reported. The subject of the second part of the article is an identification of the analytical shapes of the "exact" and "just-intergenerational" discounting functions.

Jacek Szołtysek
Grażyna Trzpiot
Uniwersytet Ekonomiczny w Katowicach

ANALIZA PREFERENCJI LUDZI MŁODYCH W OCENIE ATRAKCYJNOŚCI MIAST JAKO PRODUKTU TURYSTYCZNEGO

Przedmiotem artykułu jest omówienie wyników badania ankietowego związanego z problematyką kształtowania się opinii i poglądów młodzieży na temat miasta jako produktu turystycznego. W badaniach związanych ze zmianami zachodzącymi w otoczeniu miejskim, w jakim funkcjonujemy, oraz w projektowaniu miast pojawiają się nowe kategorie postrzegania miast – nowe funkcje miast, znane z literatury jako bunt miast. Badamy postrzeganie miast w kategoriach atrakcyjności turystycznej oraz w kategoriach produktu turystycznego. Badamy również kreatywne postrzeganie miast w odniesieniu do nowych nurtów kształtowania przestrzeni miejskich, takich jak gry miejskie (*city games*) czy legendy miejskie (*urban legend*). Na potrzeby analizy wyników badań wykorzystujemy metody wielowymiarowe.

1. Miasto jako nierozszyfrowany fenomen

„Z miastami jest jak ze snami: wszystko, co wyobrażalne, może się przyśnić, ale nawet najbardziej zaskakujący sen jest rebusem, który kryje w sobie pragnienie lub jego odwrotną stronę – lęk" [Calvino 1975, s. 34]. Tymi słowami Italo Calvino opisuje miasto nie konkretne, lecz niewidoczne. Nie ma powodu, by w celu zrozumienia fenomenu miasta sięgać po opisy bytów konkretnych bądź próbować je badać samemu. Miasto jest „(…) najspójniejszą i, ogólnie rzecz biorąc, najbardziej udaną próbą, jaką kiedykolwiek podjął człowiek w przekształcaniu świata, w którym żyje, zgodnie z głosem swojego serca. Jednakże jeżeli miasto jest światem

stworzonym przez człowieka, to jest to też świat, w którym jest on odtąd zmuszony żyć. W ten sposób, pośrednio i bez pełnej świadomości istoty swojego zadania, stwarzając miasto, człowiek też przekształcił samego siebie" [Park 1967, s. 3, cyt. za: Harvey 2012, s. 21]. Ten pogląd, z którego wynika wiele konsekwencji zarówno dla miasta, jak i ludzi w nim zamieszkujących bądź wiążących z nim swoje plany, jest początkiem działań praktycznych, co do których możemy podejrzewać, że pomysły ludzi, ich wspólne działanie oraz „kolektywna wizja przyszłości", wsparta narzędziami ułatwiającymi wdrożenie pomysłu w życie, może być motorem napędzającym i kreującym nową miejską rzeczywistość. „Żyjemy w miejscach i poprzez miejsca, tak zresztą, jak i miejsca istnieją dzięki nam. Ta koegzystencja miejsca i człowieka jest zbudowana z wielu różnorodnych i złożonych przeżyć. Ma ona także swą wewnętrzną osobną budowę, związaną z bogactwem świata ukonstytuowanego" [Buczyńska-Grajewicz 2006, s. 5]. Miasto jest więc nie tylko przestrzenią i konsekwencją jej pokonywania, lecz również wynikiem swoistej interakcji człowieka z nim, dokonywanej w każdym codziennym działaniu. „Bez elementów miasta nie ma miasta, lecz elementy nie tworzą miasta. Linią łuku dla miasta jest jego nastrój i tajemne więzi w nim ukryte" [Buczyńska-Grajewicz 2006, s. 301]. Ów nastrój znajduje swoje ujście w wielu elementach życia codziennego miasta. Ich kreowaniem zajmuje się często elita miasta – artyści, dziennikarze, wyspecjalizowane agencje, ale też szeregowi mieszkańcy, którzy coraz częściej są twórcami zdarzeń alternatywnych. Nastrój i tajemne więzi znajdują swoje odbicie również w legendach miejskich. Legendy te, „(…) zwane również współczesnymi, miejskimi mitami lub makroplotkami, to opowieści współczesnego folkloru, które mogą być rozpowszechniane metodami tradycyjnymi, takimi jak przekaz ustny, albo nowoczesnymi, czyli za pośrednictwem e-maili, faksów i Internetu. Jedne są prawdziwe, inne zawierają jedynie ziarnko prawdy. Niekiedy granica między fikcją a faktem staje się na tyle niewyraźna, że nie daje się ich rozgraniczyć" [Barber 2007, s. 11 i 13]. Miasto, zgodnie z poglądami większości specjalistów, profesjonalnie zajmujących się kwestiami funkcjonowania miast i problemami tzw. miejskości, należy do osób je zamieszkujących. Powinno tym osobom gwarantować szereg praw (np. do samorealizacji, poszanowania odrębności czy bezpieczeństwa), w zamian oczekując solidarności. Dzieje się tak dlatego, że to właśnie ludzie, poprzez swoje wybory życiowe, są jego twórcami. Dyskutowane powszechnie prawo do miasta oznacza „(…) żądanie pewnego rodzaju władzy nad kształtowaniem procesów urbanizacyjnych nad sposobami, w jakie nasze miasta są tworzone i przekształcane, i zrobienie tego w fundamentalny i radykalny sposób" [Harvey 2012, s. 23]. Aby budować przyszłość miasta, trzeba rozpoznać opinie użytkowników miasta

w zakresie tego, jak postrzegają pożądany stan przyszły miasta, w jaki sposób do tego stanu należy zmierzać i z jakimi wyrzeczeniami przyjdzie się zmierzyć. Takie badania znajdują swoje istotne uzasadnienie w dwóch okolicznościach, wcześniej sygnalizowanych:

1. Po pierwsze – o przyszłości miasta decydują jego właściciele – mieszkańcy.
2. Po drugie – na decyzje o przyszłości ma wpływ wiele okoliczności – w tym zarówno materialny wymiar przestrzeni miasta, jak i wymiary duchowe.

Badania zaprezentowane w dalszej części artykułu zostały przeprowadzone na grupie studentów i nie pretendują do miana badań reprezentatywnych. Intencją autorów jest wskazanie potencjału takich badań oraz przykładu doboru narzędzi ułatwiających wnioskowanie i planowanie konkretnych działań rozwojowych [Szołtysek, Trzpiot 2011, s. 28-33; Szołtysek, Trzpiot 2011, s. 21-32; Szołtysek, Trzpiot 2012, s. 75-85].

2. Miasto jako produkt turystyczny

W publikacjach dotyczących aspektów ekonomicznych miasta szeroko jest stosowany termin „produkt turystyczny" (ang. *tourist product*). W wąskim znaczeniu oznacza on wszystko, co jest dla turystów przedmiotem zakupu (np. usługi hotelarskie, transportowe). W szerszym rozumieniu obejmuje zasoby (walory) turystyczne obszaru, zagospodarowanie turystyczne (wraz z usługami) oraz całość doświadczeń turysty od chwili opuszczenia miejsca zamieszkania do czasu powrotu. Zdaniem J. Altkorna produktem turystycznym może być „miejsce w hotelu, wycieczka krajoznawcza, pobyt w uzdrowisku, zwiedzanie miasta i rejs dookoła świata". Według tego autora produkt turystyczny składa się z tzw. produktu rzeczywistego (*tangible product*), czyli niezbędnych dóbr i usług pozwalających zaspokoić potrzeby turystów, oraz tzw. produktu poszerzonego (*augmented product*), na który składają się dobra i usługi pożądane, ale nie niezbędne. Przedmiotem polityki produktu w aspekcie terytorialnym jest tzw. złożony produkt turystyczny oferowany przez dany obszar (*community tourism product*). Generalnie na złożony produkt turystyczny składają się naturalne i sztuczne dobra turystyczne, określone usługi i towary, a także udogodnienia umożliwiające korzystanie z dóbr turystycznych oraz nabywanie towarów i usług. Usługi świadczone przez wytwórców (np. przejazdy, noclegi, wyżywienie) oraz udogodnienia o charakterze nierynkowym (sieć drogowa, ochrona środowiska itp.) są podporządkowane realizacji celów podstawowych [Strzelecki 2010, s. 5-8].

Celem artykułu jest analiza preferencji, do której wykorzystamy wielowymiarową analizę statystyczną miasta jako produktu turystycznego – atrakcyjność miasta dla młodych ludzi z województwa śląskiego. Analizy dokonano na podstawie badania ankietowego przeprowadzonego wśród młodych ludzi zamieszkujących i studiujących w województwie śląskim. Ankieta była złożona z kilku bloków pytań. Odpowiedzi respondenci udzielali na skali porządkowej od -5 do 5. Ocena -5 oznacza najniższą wartość badanego zjawiska, 0 – stan obojętności, a ocena 5 – najwyższą wartość badanego zjawiska.

W pierwszej kolejności respondenci zostali poproszeni o ocenę elementów (atrybutów) miasta decydujących o jego atrakcyjności turystycznej. Ocenie podlegały: architektura; zabytki; parki, miejsca spacerowe, ogrody; galerie; ogólny wygląd miasta; infrastruktura; klimat, atmosfera miasta; hotele, restauracje, kluby, puby; centra handlowe; imprezy kulturalne; przeloty nad miastem.

Drugie zagadnienie dotyczyło oceny najbardziej atrakcyjnych określeń dla miasta jako atrakcji turystycznej: miasto turystyczne; miasto spotkań; miasto otwarte na ludzi; miasto rozrywek i bogatego życia towarzyskiego; miasto edukacji i nauki; miasto zabytków; miasto wydarzeń kulturalnych; miasto kultu religijnego; miasto biznesu i inwestycji; miasto ciągłego rozwoju i postępu.

Trzeci blok pytań ankietowych dotyczył oceny najbardziej atrakcyjnych obszarów dla miasta jako produktu turystycznego: turystyka piesza; turystyka kulturowa (festiwale, obchody, koncerty); seksturystyka; turystyka religijna (zabytki sakralne, pielgrzymki, imprezy wyznaniowe); udział w imprezach sportowych (zawody, mecze, turnieje); turystyka rowerowa; sporty zimowe; tenis; turystyka wodna; turystyka uzdrowiskowa; Wellness, Spa; zabiegi lekarskie; zakupy; konferencje, kongresy, targi branżowe; podróże służbowe; turystyka przygraniczna; turystyka poznawcza (zwiedzanie).

Respondenci byli proszeni o podanie źródła utrzymania (samodzielnie się utrzymuję, na utrzymaniu innej osoby/osób, częściowo samodzielnie, a częściowo na utrzymaniu innych osób); oceny statusu materialnego (jestem zadowolony/zadowolona, nie jestem zadowolony/zadowolona, jestem częściowo zadowolony/zadowolona); budżetu rozporządzalnego na rozrywkę i turystykę miesięcznie (zł); ustosunkowanie się do stwierdzenia, iż w mieście można poprawić swój status materialny (tak, nie, nie wiem); odpowiedzi na pytanie, czy lubi miasto (tak, nie, nie wiem); oraz czy w przyszłości chciałby mieszkać mieście (tak, nie, nie wiem).

3. Próba badawcza

Badaniem zostało objętych 87 respondentów. Statystyki opisowe dotyczące zmiennych ilościowych charakteryzujących respondentów, tj. wiek, liczba mieszkańców miejscowości, którą obecnie zamieszkują, budżet rozporządzalny na rozrywkę i turystykę miesięcznie (zł) zamieszczono w tabeli 1.

Tabela 1

Statystyki opisowe zmiennych ilościowych charakteryzujących respondentów

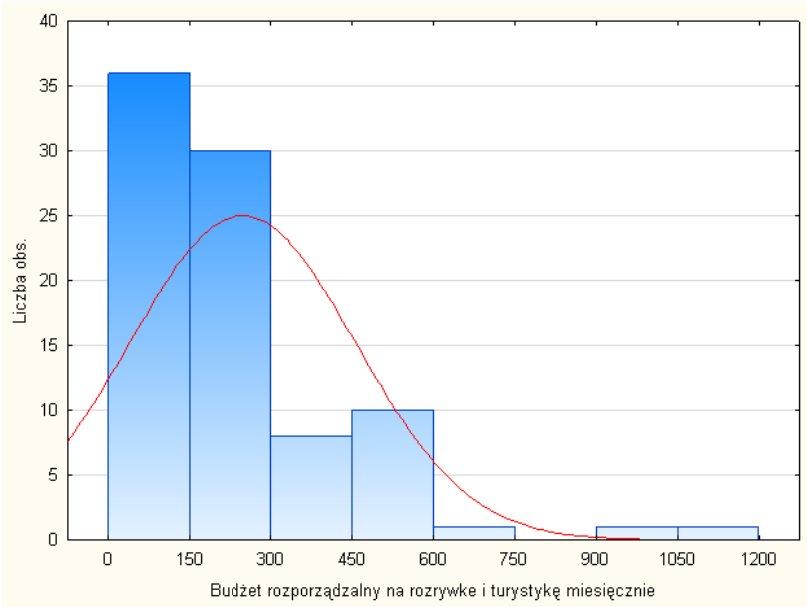
Wyszczególnienie	Średnia	Mediana	Moda	Min	Max	Rozstęp	Odchylenie standardowe	Współczynnik zmienności	Skośność
Wiek	21,7	21	21	20	26	6	1,1	5,21	1,157
Liczba mieszkańców	157 892	150 000	300 000	600	400 000	399 400	111 106,1	70,36	0,073
Budżet na rozrywkę i turystykę miesięcznie	246,1	200	100,	0	1200	1200	208,2	84,60	1,89

Średnia wieku respondentów wynosiła 21,7 lat (tabela 1). Mediana wynosząca 21 lat informuje o tym, że przynajmniej połowa badanych respondentów miała nie więcej 21 lat. Rozstęp wieku wśród badanych wyniósł 6 lat, najmłodszy uczestnik badania miał 20 lat, najstarszy 26 lat. Niski współczynnik zmienności (5,2%) świadczy o niewielkiej sile dyspersji, wiek respondentów charakteryzuje jednorodność.

Respondenci posiadający wykształcenie średnie (rys. 1) zamieszkują miejscowości, w których średnia liczba mieszkańców wynosi 158 000, największa liczba ankietowanych zamieszkuje miejscowość o liczbie 300 000 mieszkańców – Katowice. Wysoki współczynnik zmienności (70,4%) świadczy o dużej sile dyspersji, liczba mieszkańców charakteryzuje się dużą niejednorodnością. Współczynnik skośności wynoszący 0,07 informuje o występowaniu asymetrii prawostronnej rozkładu liczby mieszkańców, co oznacza, że większa część respondentów mieszka w miejscowościach o liczbie mieszkańców poniżej 300 000.

Rozważając budżet, którym dysponują respondenci na rozrywkę i turystykę miesięcznie (zł), średnio była to kwota 246,1 zł. Przynajmniej połowa badanych respondentów miesięcznie rozporządzała na rozrywkę i turystykę kwotą 200 zł. Najczęściej respondenci na ten cel przeznaczali 100 zł miesięcznie. Rozstęp dla miesięcznych wydatków na turystykę i rekreację wyniósł 1200 zł, wśród badanych znajdowali się tacy, którzy nie dysponowali środkami pieniężnymi na ten cel, oraz tacy, którzy przeznaczali na niego miesięcznie 1200 zł. Zmienna ta cha-

rakteryzuje się dużą niejednorodnością, o czym świadczy współczynnik zmienności równy 85%, oraz bardzo silną asymetrią prawostronną (rys. 1).



Rys. 1. Struktura wydatków na rozrywkę i turystykę

Cechy jakościowe charakteryzujące respondentów: płeć, wykształcenie, zamieszkanie, źródło utrzymania, ocena statusu materialnego przedstawiono za pomocą tabeli liczności (tabele 2-5).

Tabela 2

Tabela liczności cechy płeć i miejsce zamieszkania

Płeć	Liczba	Procent	Miejsce zamieszkania	Liczba	Procent
Mężczyzna	63	72,4%	**miasto**	75	86,2%
Kobieta	24	27,6%	**wieś**	12	13,8%

Tabela 3

Tabela liczności cechy wykształcenie

Wykształcenie	Liczba	Procent
Średnie	81	93,1%
Licencjackie	5	5,7%
Magisterskie	1	1,2%

Większość ankietowanych stanowili mężczyźni (tabela 2), większość – 93,1% respondentów – posiadała wykształcenie średnie (tabela 2).

Tabela 4

Tabela liczności cechy źródło utrzymania

Źródło utrzymania	Liczba	Procent
Częściowo samodzielnie, a częściowo na utrzymaniu innych osób	47	54%
Na utrzymaniu innej osoby/osób	38	43,7%
Samodzielnie się utrzymuję	2	2,3%

Większość ankietowanych zamieszkiwała miasto (tabela 4), aż 54,1% respondentów częściowo utrzymywało się samodzielnie, a częściowo pozostawało na utrzymaniu innych osób. Na utrzymaniu innej osoby było 43,7% ankietowanych (tabela 5). Ponad 47% respondentów jest częściowo zadowolonych ze swojego statusu materialnego, natomiast ponad 41% respondentów jest z niego zadowolonych (tabela 5).

Tabela 5

Tabela liczności zmiennej ocena statusu materialnego

Ocena statusu materialnego	Liczba	Procent
Nie jestem zadowolony/zadowolona	10	11,5%
Jestem zadowolony/zadowolona	36	41,3%
Jestem częściowo zadowolony/zadowolona	41	47,2%

Podsumowując, można stwierdzić, iż badaniu zostali w większości poddani mężczyźni w wieku 21 lat, z wykształceniem średnim, zamieszkujący miasto, utrzymujący się częściowo samodzielnie, będący częściowo zadowoleni ze swojego statusu materialnego, przeznaczający miesięcznie na rozrywkę i turystykę średnio 246,1 zł.

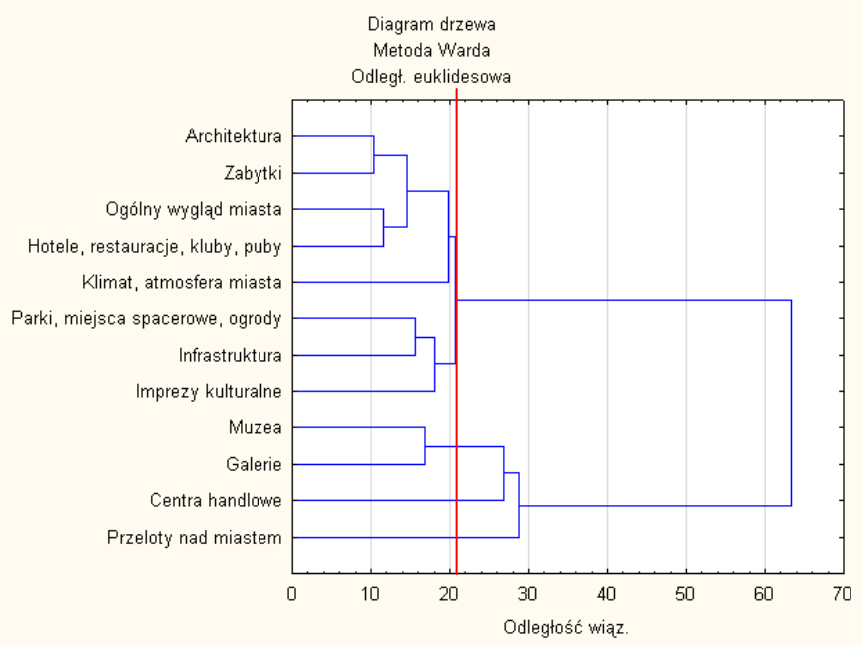
4. Analiza preferencji z wykorzystaniem statystycznej analizy danych

Analiza wielowymiarowa została wykorzystana do wskazania powiązań preferencji badanych respondentów w grupy jednorodne. Analiza skupień jest narzędziem analizy danych służącym do grupowania obiektów opisanych za pomocą cech w możliwie „jednorodne" grupy – skupienia. Zastosowaną w tym artykule metodą do analizy skupień jest metoda hierarchiczna [Everitt 2001; Gordon 1999; Grabiński 1992]. Wspólną cechą krokowych algorytmów tej metody jest wyznaczenie skupień przez łączenie (aglomerację) powstałych, w poprzednich krokach algorytmu, mniejszych skupień. Graficzną ilustracją przebie-

gu aglomeracji jest wykres zwany dendrogramem. Jest to binarne drzewo, którego węzły reprezentują skupienia, a liście pojedyncze obiekty. Liście są umieszczone na poziomie zerowym, pozostałe węzły drzewa są umieszczone na wysokości odpowiadającej mierze niepodobieństwa między skupieniami reprezentowanymi przez węzły potomki. Algorytm ten wykorzystuje nie tylko miary niepodobieństwa między obiektami, ale także metody wiązania skupień [Krzyśko et al. 2008, s. 345-348].

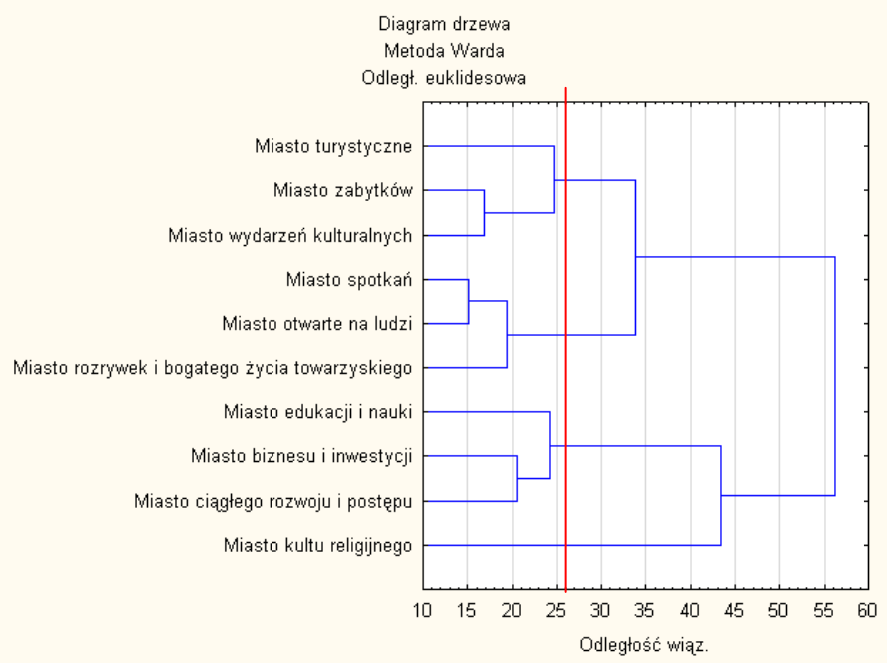
Rozpoczynamy analizę od pierwszego bloku pytań. Analiza skupień metodą aglomeracji, z wykorzystaniem metody Warda do określenia metody łączenia skupień oraz odległości euklidesowej [Ostasiewicz (red.) 1998] przy formowaniu skupień w odniesieniu do oceny atrybutów miasta jako produktu turystycznego, doprowadziła do powstania hierarchicznego drzewa (rys. 2).

Można wyodrębnić kilka skupień w ocenie atrybutów miasta decydujących o jego atrakcyjności turystycznej. Jak wynika z analizy (w odległości 16), można wyodrębnić skupienie muzea i galerie (rys. 2). Kolejne skupienie (w odległości 18) to: parki, miejsca spacerowe, ogrody; infrastruktura; imprezy kulturowe. Następnie (w odległości 19,8) można wyodrębnić skupienie następujących atrybutów miasta: architektura; zabytki; ogólny wygląd miasta; hotele, restauracje, kluby, puby oraz klimat, atmosfera miasta.



Rys. 2. Diagram drzewa, atrybuty miasta jako produktu turystycznego

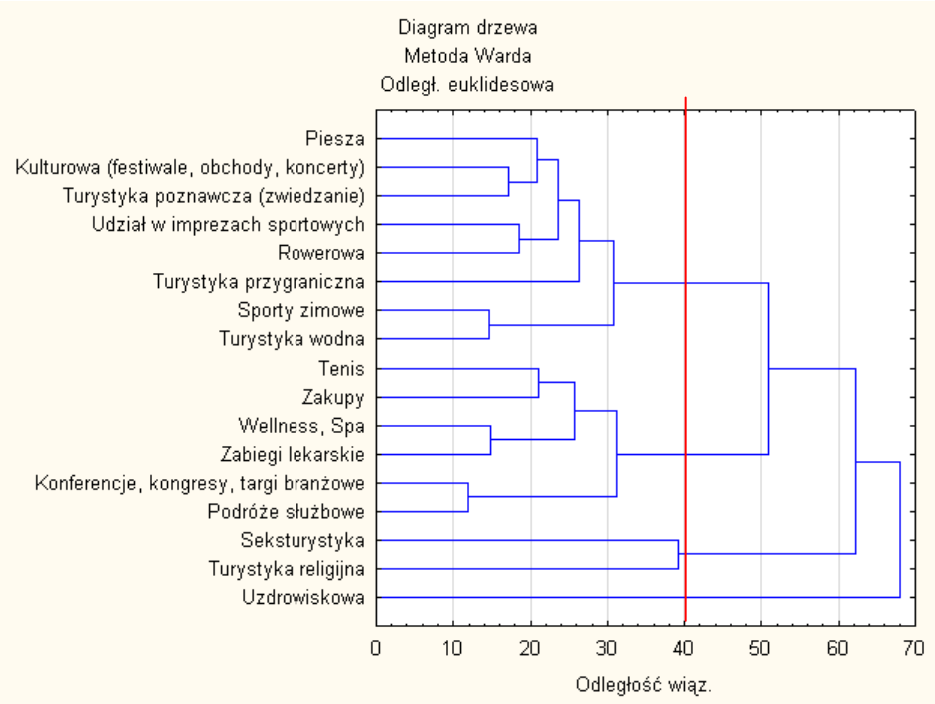
Taką samą metodologię, czyli analizę skupień metodą aglomeracji z wykorzystaniem metody Warda do określenia metody łączenia skupień oraz odległości euklidesowej przy formowaniu skupień, wykorzystano do oceny określeń miasta jako produktu turystycznego. Preferencje respondentów tym razem również grupują się, tworząc skupienia względem określeń opisujących postrzeganie miasta. Oceny młodych ludzi są powiązane z potrzebami funkcjonowania w społeczeństwie (rys. 3). Oczekuje się, że będą charakterystyczne dla badanej grupy demograficznej.



Rys. 3. Diagram drzewa, określenia miasta jako produktu turystycznego

Analiza wskazuje (rys. 3), że w ocenach respondentów dotyczących określeń miasta jako produktu turystycznego występowały trzy skupienia, pierwsze w odległości 19,5 złożone z określeń: miasto spotkań; miasto otwarte na ludzi; miasto rozrywek i bogatego życia towarzyskiego, drugie w odległości 24,3 skupiające miasto edukacji i nauki; miasto biznesu i inwestycji; miasto ciągłego rozwoju i postępu, trzecie skupienie w odległości 23,6 to miasto turystyczne; miasto zabytków; miasto wydarzeń kulturowych. Skupienia te wyraźnie zarysowały podobieństwa w ocenach respondentów określeń dla miasta jako produktu turystycznego związanych z kulturą, życiem towarzyskim oraz edukacją i biznesem.

Analogiczna analiza została przeprowadzona przy formowaniu skupień w odniesieniu do oceny obszarów miasta jako produktu turystycznego (rys. 4).



Rys. 4. Diagram drzewa, obszary miasta jako produktu turystycznego

Oceny respondentów w stosunku do obszarów miasta jako produktu turystycznego nie były do siebie na tyle podobne pomiędzy poszczególnymi obszarami miasta (rys. 4.), aby wyodrębnić tak wyraźne skupienia, jak w pytaniu dotyczącym określeń miasta jako produktu turystycznego. Na podstawie powyższego dendrogramu można stwierdzić, iż respondenci najbardziej jednorodnie ocenili obszary: konferencje, kongresy, targi branżowe oraz podróże służbowe (odległość 11,9). Obserwujemy znacznie większe odległości, a najodleglejszymi od siebie obszarami miasta jako produktu turystycznego jest seksturystyka oraz turystyka religijna.

5. Analiza rzetelności

W naukach społecznych nierzetelne pomiary ludzkich przekonań lub intencji będą przeszkodą w przewidywaniu ich zachowań. Analiza rzetelności może

być wykorzystana do zbudowania rzetelnych skal pomiarowych, do poprawy istniejących skal oraz oceny rzetelności używanych skal. Pomiar jest rzetelny, jeśli w stosunku do błędu odzwierciedla głównie wynik prawdziwy. W szczególności możemy zdefiniować współczynnik rzetelności w kategoriach proporcji zmienności wyniku prawdziwego, która jest ujęta dla wszystkich respondentów w stosunku do całkowitej obserwowanej zmienności. Jeśli na nasze pytania odpowiedziało kilka osób, to możemy obliczyć wariancję dla każdej pozycji oraz wariancję dla skali sumarycznej. Wariancja skali sumarycznej będzie mniejsza niż suma wariancji pozycji, jeśli pozycje mierzą tę samą zmienność między respondentami, tzn. jeśli mierzą pewien wynik prawdziwy. W teorii oznacza to, iż wariancja sumy dwóch pozycji jest równa sumie dwóch wariancji minus (dwa razy) kowariancja, tzn. wielkość wariancji wyniku prawdziwego wspólnej tym dwóm pozycjom.

Możemy oszacować proporcję wariancji wyniku prawdziwego, która jest udziałem danych pozycji, przez porównanie sumy wariancji pozycji i wariancji skali sumarycznej; najbardziej popularny współczynnik rzetelności to współczynnik alfa Cronbacha. Jeśli pozycje w ogóle nie dają wyniku prawdziwego, ale jedynie błąd (który jest nieznany, specyficzny i w konsekwencji nieskorelowany pomiędzy osobnikami), to wariancja sumy będzie taka sama, jak suma wariancji poszczególnych pozycji. Dlatego współczynnik alfa będzie równy zero. Jeśli wszystkie pozycje są idealnie rzetelne i mierzą tę samą rzecz (wynik prawdziwy), to współczynnik alfa jest równy 1 [Trzpiot (red.) 2013]. W celu poddania analizie rzetelności skal w pytaniach zawartych w ankiecie wyznaczono dla poszczególnych skal wartość współczynnika alfa Cronbacha.

Tabela 6

Wartości współczynnika alfa Cronbacha dla pytań ankietowych

Pytanie ankietowe	Współczynnik alfa Cronbacha
Ocena atrakcyjności elementów (atrybutów) miasta jako produktu turystycznego	0,73
Ocena atrakcyjności określeń miasta jako produktu turystycznego	0,81
Ocena atrakcyjności obszarów miasta jako produktu turystycznego	0,79

Dla wszystkich skal w pytaniach zamieszczonych w ankiecie współczynnik alfa Cronbacha wynosi powyżej 0,7, można zatem stwierdzić, iż ich rzetelność nie jest możliwie najwyższa, lecz jest przeciętna (tabela 6). Wartości współczynników alfa Cronbacha są akceptowalne.

Podsumowanie

W celu poznania opinii młodych ludzi (20-26 lat) zamieszkujących województwo śląskie na temat atrakcyjności turystycznej miasta przeprowadzono badanie ankietowe. Analiza statystyczna uzyskanych danych ankietowych była prowadzona z uwzględnieniem następujących zmiennych grupujących: płeć, wykształcenie, zamieszkanie.

Z przeprowadzonej analizy wnioskuje się iż:

Najatrakcyjniejszym atrybutem miasta, bez względu na miejsce zamieszkania i wykształcenie respondentów, jest architektura. W badaniach wyjaśniono, że pod tym pojęciem należy rozumieć unikatową bądź wyróżniającą architekturę miasta, składającą się na wygląd miasta. Wprowadzając kategoryzację względem cechy grupującej – płeć, możemy dostrzec różnice w postrzeganiu atrakcyjności miasta, ponieważ kobiety za najatrakcyjniejszy atrybut miasta uważają jego ogólny klimat, mężczyźni – zabytki. Należy dodać, iż architektura oraz zabytki tworzą jedno skupienie (wynik analizy skupień). Bezsprzecznie te informacje mogą być źródłem wielu decyzji zarządczych i realizacyjnych w zakresie dostosowania architektonicznego wyrazu miasta – począwszy od wyboru zespołu projektowego, poprzez dobór argumentów oraz programów promujących nowe idee, kończąc na doborze prezenterów koncepcji na spotkania z mieszkańcami.

Bezsprzecznie najatrakcyjniejszym określeniem dla miasta jako produktu turystycznego dla wszystkich grup respondentów jest określenie – miasto rozrywek i bogatego życia towarzyskiego. Najmniej atrakcyjnym – miasto kultu religijnego. Te stwierdzenia pozwalają na kreowanie nowego wizerunku miasta z punktu widzenia tworzenia produktu turystycznego w skali mezo.

Obszary miasta jako produktu turystycznego były tematem najbardziej spornym wśród respondentów. Kobiety oraz mieszkańcy miasta i wsi oraz osoby z wykształceniem średnim i licencjackim uznali, że obszarem takim jest turystyka uzdrowiskowa. W ramach tej kategorii mężczyźni ocenili, że sporty zimowe stanowią o atrakcyjności miasta. Respondenci z wykształceniem wyższym stwierdzili, iż turystyka kulturowa może być tym magnesem, który przyciągnie do miasta odwiedzających.

Opinie młodych ludzi na temat atrybutów, określeń oraz obszarów miasta jako produktu turystycznego w największym stopniu były zależne od ich budżetu rozporządzalnego na rozrywkę i turystykę miesięcznie.

Zaprezentowane częściowe wyniki badań mają za zadanie pokazać, w jaki sposób władze miasta mogą pozyskiwać informacje na temat pożądanych kierunków rozwoju, w tym preferencji mogących być podstawą do budowania spo-

łecznie akceptowanych programów rozwojowych. Przeprowadzenie badań w rozmaitych przekrojach społecznych pozwoli na budowanie koncepcji rozwojowych uwzględniających wyjawione preferencje badanej grupy. Jeżeli nastąpi konieczność zbliżenia rozbieżnych poglądów, wówczas władze miasta mogą sięgać do rozmaitych programów propagujących nowe kierunki, zachęcających do poparcia nowych pomysłów, bądź (lub jednocześnie) przystąpić do kreowania mitów miejskich.

Bibliografia

Barber M., 2007: Legendy miejskie. Wydawnictwo RM, Warszawa.

Buczyńska-Grajewicz H., 2006: Miejsca, strony, okolice. Przyczynek do fenomenologii przestrzeni. Universitas, Kraków.

Calvino I., 1975: Niewidzialne miasta. Tłum. Alina Kreisberg. Czytelnik.

Everitt B., 2001: Cluster Analysis. HEB, London.

Gordon A.D., 1999: Classification. Chapman & Hall, London, New York, Washington.

Grabiński T., 1992: Metody aksonometrii. AE, Kraków.

Harvey D., 2012: Bunt miast. Prawo do miasta i miejska rewolucja. Fundacja Nowej Kultury Bęc Zmiana, Warszawa.

Krzyśko M., Wołyński W., Górecki T., Skorzybut M., 2008: Systemy uczące się, rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości. WNT, Warszawa.

Ostasiewicz W. (red)., 1998: Statystyczne metody analizy danych. AE, Wrocław.

Park R., 1967: On Social Control and Collective Behavior. Chicago.

Strzelecki K., 2010: Miasto i gmina Miastko jako produkt turystyczny. Akademia Pomorska, Słupsk.

Szołtysek J., Trzpiot G., 2011: Klasyfikacja oczekiwań i preferencji komunikacyjnych studentów. „Śląski Przegląd Statystyczny", 9 (15), s. 21-32.

Szołtysek J., Trzpiot G., 2011: Preferencje komunikacyjne studentów jako przesłanki kształtowania programów mobilnościowych. „Transport Miejski i Regionalny", nr 4, s. 28-33.

Szołtysek J., Trzpiot G., 2012: Cluster Analysis in Description Some Communication Behavior. Zeszyty Naukowe Politechniki Częstochowskiej, Zarządzanie 5, s. 75-85.

Tryon R.C., 1939: Cluster Analysis. Edwards Brothers, Ann Arbor, Ml.

Trzpiot G. (red.), 2013: RExcel w analizie danych. UE, Katowice.

Trzpiot G., Szołtysek J., 2013: Atrakcyjność miasta w ocenie młodzieży – na przykładzie miasta Wałbrzych – wstępne rozpoznanie. Wydawnictwo PWSZ, Wałbrzych.

ANALYSIS OF PREFERENCE IN ASSESSMENT OF YOUNG PEOPLE CITY'S ATTRACTIVENESS AS TOURISM PRODUCT

Summary

The article is a discussion of the results of the survey related to the issues shaping the views and opinions of young people on the city as a tourist product. In studies related to changes taking place in an urban setting in which we operate, and in the design of cities, there are new categories of perception cities – new features towns, known in the literature as a revolt cities. We study the perception of cities in terms of tourist appeal and in terms of the tourism product. We also examine the creative perception of cities with regard to new trends shaping urban spaces, such as urban games or urban legends. For the analysis of results of research we use multivariate methods.

Marek Szopa
Uniwersytet Śląski w Katowicach

DLACZEGO W DYLEMAT WIĘŹNIA WARTO GRAĆ KWANTOWO?*

1. Klasyczny dylemat więźnia

Dylemat Więźnia [DW] jest bardzo znanym przykładem zastosowania teorii gier do zagadnień związanych z podejmowaniem decyzji. Po raz pierwszy opisany przez Flooda i Dreshera [Flood i in. 1952] został spopularyzowany przez Alberta Tuckera, którego historii o dwu więźniach dylemat zawdzięcza obecną nazwę. Swoja popularność DW zawdzięcza uniwersalności schematu gry, która opisuje dylemat decyzyjny powszechnie występujący w wielu sytuacjach życia codziennego. Typowy scenariusz zakłada, że dwaj gracze, Alicja i Bartek, niezależnie od siebie wybierają jedną z dwu strategii – „współpraca" (*W*) lub „odmowa" (*O*). Wybory obu graczy są podstawą do wypłacenia im wygranych, które są opisane w tabeli 1.

Tabela 1

Macierz wypłat Dylematu Więźnia

		Bartek: *W*	Bartek: *O*
Alicja	*W*	(w, w)	(f, z)
	O	(z, f)	(k, k)

Wybór „*W*" oznacza współpracę, a wybór „*O*" jej odmowę. Pierwsza liczba każdej pary to wypłata Alicji, druga to wypłata Bartka. Wypłata w to nagroda za współpracę, z to pokusa do zdrady, f to wypłata frajera, a k to kara za (obustronną) odmowę współpracy. Liczby te spełniają nierówności: $z > w > k > f$ oraz $w > \frac{z+f}{2}$ [Straffin 2001].

* Praca była dofinansowana ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji nr DEC-2011/01/B/ST6/07197 i DEC-2011/03/B/HS4/03857.

Na podstawie macierzy wypłat łatwo zauważyć, że niezależnie od wyboru przeciwnika dominującą strategią każdego gracza jest „odmowa". Para strategii (*O,O*) jest równowagą Nasha (RN) tej gry. Paradoksalnie równowaga ta odpowiada wygranej (k, k), która jest daleka od rozwiązania optymalnego w sensie Pareto, którym jest wynik (w, w). Sposobem uzyskania rozwiązania optymalnego jest obustronna współpraca (*W,W*). Jednak taka gra wymaga wzajemnego zaufania graczy, gdyż każde odstąpienie od tej strategii daje odmawiającemu nagrodę w postaci pokusy do zdrady z, a drugiemu karę w postaci wypłaty frajera f. Jeśli gra w DW toczy się pomiędzy osobami, które nie mają do siebie zaufania – a taka sytuacja społeczna jest powszechna – to najczęstszym wynikiem gry jest obustronna kara za brak współpracy. Gra ma naturalne rozszerzenie do wieloosobowego DW, w którym zasadniczy dylemat pozostaje taki sam, jak w przypadku dwuosobowym. Z punktu widzenia praktycznych zastosowań istotnym rozszerzeniem gry jest jej iterowana wersja [Hamilton i in. 1981], dla której strategie graczy zależą od historii wcześniejszych rozgrywek.

2. Definicja gry kwantowej

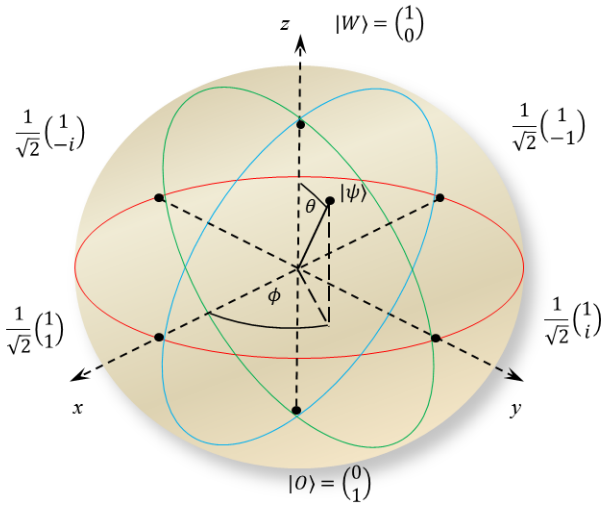
Psychologia eksperymentalna pokazuje, że realne decyzje ludzkie w sytuacji typu DW są często niezgodne z klasyczną RN. Niektórzy badacze dowodzą, że lepiej od klasycznych do wyjaśnienia decyzji ludzkich nadają się kwantowe metody opisu podejmowania decyzji [Busemeyer i in. 2006; Pothos i in. 2009]. Wraz z rozwojem badań na temat kwantowego przetwarzania informacji DW doczekał się swej kwantowej wersji [Eisert i in. 1999]. W tym ujęciu strategie graczy są operatorami w pewnej przestrzeni wektorowej zwanej sferą Blocha. Przestrzeń ta to zbiór kubitów – unormowanych wektorów o współczynnikach zespolonych rozpiętych na dwuelementowej bazie $\{|W\rangle, |O\rangle\}$, które, z dokładnością do fazy, można przedstawić w postaci:

$$|\psi\rangle = \cos\frac{\theta}{2}|W\rangle + \mathrm{e}^{\mathrm{i}\phi}\sin\frac{\theta}{2}|O\rangle, \tag{1}$$

gdzie $\theta \in [0, \pi]$ oraz $\phi \in [-\pi, \pi]$ (rys. 1).

Kubity $|W\rangle$ oraz $|O\rangle$ to kwantowe stany czyste reprezentujące „współpracę" i „odmowę". Zgodnie z zasadami mechaniki kwantowej kubit (1) jest superpozycją dwu stanów kwantowych. Oznacza to, że dopóki nie dokonamy pomiaru, nie możemy stwierdzić, w którym z dwu stanów kubit tak naprawdę się

znajduje. Jedyne, co możemy powiedzieć, to że pomiar da wynik $|W\rangle$ z prawdopodobieństwem $\cos^2\frac{\theta}{2}$ oraz wynik $|O\rangle$ z prawdopodobieństwem $\sin^2\frac{\theta}{2}$.



Rys. 1. Sfera Blocha z oznaczonym położeniem kubitów $|W\rangle = \binom{1}{0}$ – „współpraca”, $|O\rangle = \binom{0}{1}$ – „odmowa” oraz kubitów leżących w płaszczyźnie x-y i przecinających osie

W wyniku pomiaru następuje tzw. kolaps funkcji falowej kubitu, która z postaci (1) przechodzi w stan $|W\rangle$ lub $|O\rangle$. Przykładowo wszystkie kubity na równiku sfery Blocha (czerwona linia na rys. 1) reprezentują stan kwantowy, który w wyniku pomiaru kolapsuje do stanu $|W\rangle$ lub $|O\rangle$ z prawdopodobieństwem $\frac{1}{2}$. Analogiczna sytuacja ma miejsce w znanym przykładzie z „kotem Schrödingera” – w tym przypadku stany $|W\rangle$ i $|O\rangle$ odnoszą się do „stanu witalności” kota.

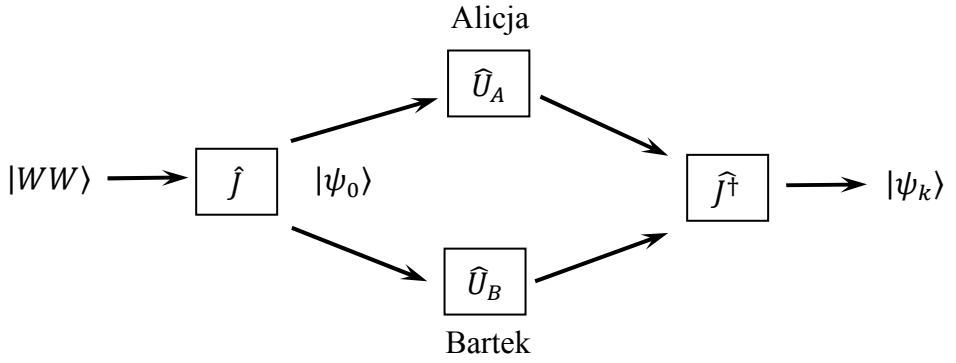
W kwantowej teorii gier to jednak nie kubity $|W\rangle$ ani $|O\rangle$ odpowiadają strategiom graczy. Strategiami są transformacje unitarne $\hat{U}_A$ – dla Alicji oraz $\hat{U}_B$ – dla Bartka działające na pewnym, specjalnie przygotowanym i znanym obu graczom, splątanym stanie kwantowym $|\psi_0\rangle = \hat{J}|WW\rangle$. Transformacje te są w ogólnej postaci obrotami na sferze Blocha $\hat{U}_X \in SU(2)$, określonymi przez macierze unitarne:

$$\hat{U}(\theta,\phi,\alpha) = \begin{pmatrix} e^{-i\phi}\cos\frac{\theta}{2} & e^{i\alpha}\sin\frac{\theta}{2} \\ -e^{-i\alpha}\sin\frac{\theta}{2} & e^{i\phi}\cos\frac{\theta}{2} \end{pmatrix}, \tag{2}$$

gdzie $\hat{U}_X = \hat{U}(\theta_X,\phi_X,\alpha_X)$, $\theta_X \in [0,\pi]$ oraz $\alpha_X,\phi_X \in [-\pi,\pi]$, $X = A,B$.

W szczególnym przypadku, gdy obrót jest określony tylko przez kąt θ, tj. $\alpha = \phi = 0$, można go wyrazić jako $\widetilde{U}(\theta) = \widehat{U}(\theta, 0,0) = \cos\frac{\theta}{2}\widehat{W} + \sin\frac{\theta}{2}\widehat{O}$. Macierz jednostkowa $\widehat{W} \equiv \widetilde{U}(0)$ odpowiada strategii „współpraca", a macierz $\widehat{O} \equiv \widetilde{U}(\pi) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (zamieniająca kubity $|W\rangle$ i $|O\rangle$) odpowiada strategii „odmowa". Strategia $\widetilde{U}(\theta)$ jest równoważna klasycznej strategii mieszanej, dla której prawdopodobieństwa obu strategii czystych W oraz O wynoszą odpowiednio $\cos^2\frac{\theta}{2}$ i $\sin^2\frac{\theta}{2}$.

Gra kwantowa, jako że dotyczy dwojga graczy, toczy się w przestrzeni par kubitów, po jednym dla każdego gracza, które są ze sobą skorelowane poprzez mechanizm splątania kwantowego (rys. 2).



Rys. 2. Schemat kwantowego DW

Gra taka może być fizycznie zrealizowana przez komputer kwantowy realizujący powyższy algorytm zależny od strategii graczy. Algorytm taki został zrealizowany eksperymentalnie [Du i in. 2002] na dwukubitowym komputerze kwantowym opartym na jądrowym rezonansie magnetycznym. Szczegóły jego działania, tj. fizyczna implementacja algorytmu kwantowego, nie są istotne dla zrozumienia gry kwantowej i w niniejszym artykule zostaną pominięte.

Stan początkowy gry jest parą kubitów $|WW\rangle$, z których pierwszy jest stanem Alicji, a drugi stanem Bartka. Wektory bazowe przestrzeni par kubitów: $|WW\rangle$, $|WO\rangle$, $|OW\rangle$, $|OO\rangle$ są stanami czystymi. Dla wygody rachunkowej oznaczymy je jako wektory: $(1,0,0,0)$ $(0,1,0,0)$, $(0,0,1,0)$, $(0,0,0,1)$ czterowymiarowej przestrzeni stanów. Na stan początkowy $|WW\rangle$ działa operator splatający, zdefiniowany w tej przestrzeni jako $\hat{J} = \frac{1}{\sqrt{2}}(\hat{I} + i\sigma_x \otimes \sigma_x)$, gdzie $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ jest jedną z macierzy Pauliego. W wyniku tego działania otrzymu-

jemy $\hat{J}\,|WW\rangle = \frac{1}{\sqrt{2}}|WW\rangle + \frac{i}{\sqrt{2}}|OO\rangle$, czyli stan splątany. Kolejno na uzyskany wynik działa iloczyn prosty operatorów $\hat{U}_A \, i \, \hat{U}_B$ reprezentujących kwantowe strategie Alicji i Bartka. Przed pomiarem stanu końcowego działamy jeszcze operatorem rozplatającym $\hat{J}^{\dagger}$. Stan końcowy $|\psi_k\rangle$ jest więc ostatecznie dany przez:

$$|\psi_k\rangle = \hat{J}^{\dagger}(\hat{U}_A \otimes \hat{U}_B)\,\hat{J}\,|WW\rangle \tag{3}$$

i jest na ogół stanem splątanym:

$$|\psi_k\rangle = p_{WW}\,|WW\rangle + p_{WO}\,|WO\rangle + p_{OW}\,|OW\rangle + p_{OO}\,|OO\rangle, \tag{4}$$

gdzie $|p_{WW}|^2$,…, $|p_{OO}|^2$ są prawdopodobieństwami, że pomiar dokonany w końcowym stanie mieszanym (4) da jeden z czterech możliwych wyników.

Wartość oczekiwana wypłaty Alicji $\$_A$ w grze kwantowej jest średnią ważoną czterech klasycznych wartości w, f, z i k z macierzy wypłat (tabela 1):

$$\$_A = w\,|p_{WW}|^2 + f\,|p_{WO}|^2 + z\,|p_{OW}|^2 + k\,|p_{OO}|^2, \tag{5}$$

gdzie wagami są kwantowe prawdopodobieństwa odpowiednich stanów czystych, z dokładnością do fazy równe [Chen i in. 2006]:

$$p_{WW} = \cos\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\cos(\phi_A + \phi_B) - \sin\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\sin(\alpha_A + \alpha_B),$$

$$p_{WO} = \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\cos(\alpha_A - \phi_B) - \cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\sin(\phi_A - \alpha_B),$$

$$p_{OW} = \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\sin(\alpha_A - \phi_B) + \cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\cos(\phi_A - \alpha_B),$$

$$p_{OO} = \cos\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\sin(\phi_A + \phi_B) + \sin\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\cos(\alpha_A + \alpha_B).$$

Wartość oczekiwaną wypłaty Bartka otrzymamy ze wzoru (5) przez zamianę rolami $f \leftrightarrow z$.

Zwróćmy uwagę, że sytuację kwantową można jednak symulować poprzez klasyczne obliczanie odpowiednich prawdopodobieństw i podstawienie ich do wzoru (5). Uzyskany wynik będzie taki sam, jak średni wynik gry kwantowej rozgrywanej wielokrotnie. W przypadku urządzenia kwantowego, które splata kubity oraz dokonuje odpowiednich przekształceń unitarnych (3), wynik gry poznamy przez pomiar stanu końcowego (4), który w wyniku kolapsu funkcji falowej da jeden z czterech możliwych stanów z właściwym prawdopodobień-

stwem. W tym przypadku również można użyć urządzeń klasycznych do losowania jednego z czterech możliwych stanów czystych $|WW\rangle$, $|WO\rangle$, $|OW\rangle$, $|OO\rangle$. Dzięki temu, że gracze wykorzystują strategie kwantowe, splątanie daje możliwości wzajemnego oddziaływania na siebie graczy, które nie ma odpowiednika w grach klasycznych.

3. Kwantowy DW w granicy klasycznej

Gra kwantowa przechodzi w grę klasyczną, jeżeli strategie (2) nie zawierają zespolonych współczynników fazowych $\alpha = \phi = 0$. Rzeczywiście, operator splątujący $\hat{J}$ jest tak dobrany, aby komutował z iloczynem prostym $\widetilde{U}_A \otimes \widetilde{U}_B$ każdej pary operatorów klasycznych, ale $\hat{J}^{\dagger}\hat{J} = \hat{I}$, więc $|\psi_k\rangle = (\widetilde{U}_A \otimes \widetilde{U}_B)\,|WW\rangle$ i wartość oczekiwana wypłaty Alicji (5) wynosi:

$$\$_A = w\,\cos^2\frac{\theta_A}{2}\cos^2\frac{\theta_B}{2} + f\cos^2\frac{\theta_A}{2}\sin^2\frac{\theta_B}{2} + \\ + z\sin^2\frac{\theta_A}{2}\cos^2\frac{\theta_B}{2} + k\,\sin^2\frac{\theta_A}{2}\sin^2\frac{\theta_B}{2}, \tag{6}$$

co daje wynik jak w klasycznej grze, kiedy obaj gracze wybierają strategie mieszane (tabela 2).

Tabela 2

Macierz wypłat kwantowego Dylematu Więźnia w granicy klasycznej $\alpha_X = \phi_X = 0$ dla $X = A, B$, gdy gracze grają strategiami mieszanymi wyznaczonymi przez θ_A i θ_B

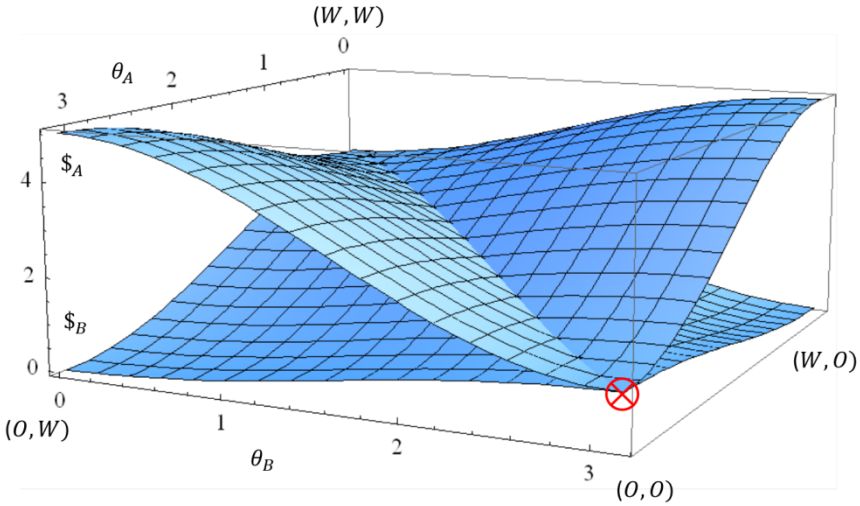
		Bartek	
		W $\left(\cos^2\frac{\theta_B}{2}\right)$	O $\left(\sin^2\frac{\theta_B}{2}\right)$
Alicja	W $\left(\cos^2\frac{\theta_A}{2}\right)$	(w, w)	(f, z)
	O $\left(\sin^2\frac{\theta_A}{2}\right)$	(z, f)	(k, k)

Wypłaty graczy należy pomnożyć przez prawdopodobieństwa wyboru poszczególnych strategii (liczby w nawiasach).

Jeśli na przykład Alicja wybierze współpracę $\theta_A = 0$, a Bartek jej odmówi $\theta_B = \pi$, to wynik gry będzie (f, z), czyli wygrana Bartka. Jedyna różnica pomiędzy klasycznym DW a granicą klasyczną kwantowego DW polega na tym, że w tym pierwszym przypadku gracz wykorzystujący strategię mieszaną sam

musi dokonać losowania, a następnie wybrać wylosowaną opcję W lub O. Pomiędzy losowaniem a wyborem opcji jest chwila, w czasie której gracz może zmienić zdanie (i zadany rozkład prawdopodobieństwa) lub przeciwnik może przechwycić informację o planowanym ruchu (i adekwatnie zareagować). W przypadku kwantowym gracz jedynie decyduje się na wybór swojej strategii danej przez kąt θ, a całą resztę wykonuje komputer kwantowy. Nie ma możliwości przechwycenia informacji kwantowej przez drugą stronę – każda taka próba zakończyłaby się kolapsem funkcji falowej i przedwczesnym zakończeniem gry.

W przypadku klasycznego DW równowagą Nasha jest para strategii (O, O). Łatwo to zobaczyć po narysowaniu funkcji wypłat obu graczy w zależności od kątów $\theta_A, \theta_B \in [0, \pi]$ (por. rys. 3, gdzie założyliśmy $z = 5, w = 3, k = 1$ i $f = 0$).



Kąt, gdzie θ = 0 odpowiada „współpracy", a θ = π „odmowie". Wypłaty $\$_A$ i $\$_B$ są rosnącymi funkcjami θ_A i θ_B, a ich wspólne maksimum odpowiada równowadze Nasha (czerwony znacznik) w punkcie (O, O).

Rys. 3. Wypłaty graczy klasycznego DW parametryzowane kątami $\theta_A, \theta_B \in [0, \pi]$

Punkt $\theta_A = \theta_B = \pi$ odpowiadający obustronnej odmowie (O, O) jest jedyną równowagą Nasha klasycznego DW, gdyż obie funkcje wypłat $\$_A(\theta_A)$ i $\$_B(\theta_B)$ rosną wraz ze swoimi argumentami, osiągając wspólne maksimum tylko w tym jedynym punkcie.

Jeśli jeden z graczy gra klasycznie, np. $\widetilde{U}(\theta)$, to drugi, który wykorzystuje kwantowe strategie, może zagrać $\widehat{U}\left(\theta + \pi, 0, -\frac{\pi}{2}\right)$. Jak łatwo sprawdzić (5), wynik gry będzie w tym przypadku równy (f, z) na korzyść gracza kwantowego. Pokazuje to przewagę strategii kwantowej nad klasyczną – niezależnie od użytej

strategii klasycznej gracz kwantowy zawsze znajdzie najlepszą odpowiedź w postaci strategii, która daje mu maksymalną wygraną f, a gracza klasycznego pozostawi z minimalną wypłatą z. Taka sytuację jednak trudno uważać za rozwiązanie gry, gdyż gracz pierwszy, dysponując większą ilością strategii, jest uprzywilejowany.

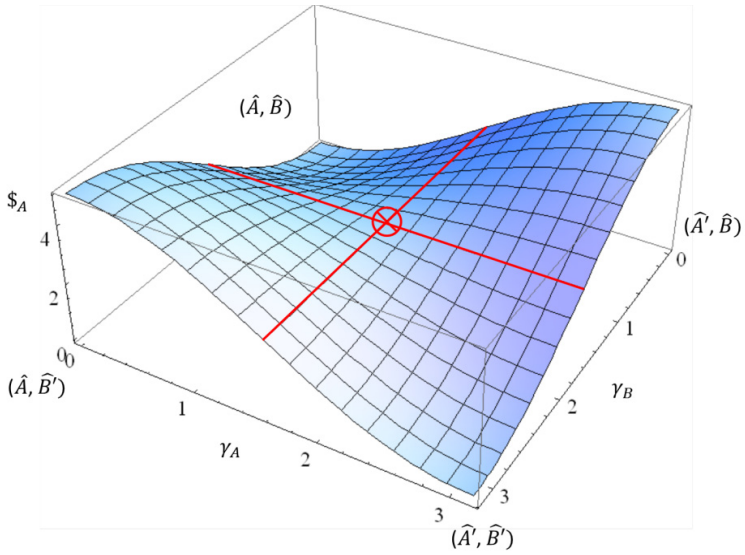
4. Równowagi Nasha kwantowego DW

Strategie kwantowe dają jednak dużo większe bogactwo możliwych wyników, niemożliwych do osiągnięcia za pomocą strategii klasycznych. Załóżmy, że Alicja wybiera dowolną strategię kwantową $\hat{A} = \hat{U}(\theta_A, \phi_A, \alpha_A)$. Bartek może wybrać strategię $\hat{B} = \hat{U}(\theta_B, \phi_B, \alpha_B) = \hat{U}\left(\theta_A + \pi, \alpha_A, \phi_A - \frac{\pi}{2}\right)$. Zauważmy, że niezależnie od wyboru Alicji, ruch Bartka daje współczynniki (4) $p_{WW} = p_{OW} = p_{OO} = 0$ oraz $|p_{WO}| = 1$. Transformacja użyta przez Bartka „unieważnia" zatem dowolny ruch Alicji i doprowadza do sytuacji, że końcową strategią Alicji zapisaną w $|\psi_k\rangle$ jest „współpraca", podczas gdy Bartek gra „odmowę". Wynikiem tej gry jest maksymalna wygrana Bartka $(\$_{A,}\$_B) = (f, z)$. Jednak gra kwantowa jest symetryczna i Alicja może odpowiedzieć na strategię Bartka $\hat{B}$ strategią $\widehat{A'} = \hat{U}(\theta'_A, \phi'_A, \alpha'_A) = \hat{U}\left(\theta_A, \phi_A - \frac{\pi}{2}, \alpha_A - \frac{\pi}{2}\right)$. Tym razem jedynym niezerowym współczynnikiem $|\psi_k\rangle$ jest $|p_{OW}| = 1$, czyli teraz Alicja gra „odmowę", podczas gdy Bartek gra „współpracę" i wypłata odwraca się $(\$_{A,}\$_B) = (z, f)$. Dla Bartka najlepszą odpowiedzią na strategię Alicji $\widehat{A'}$ jest $\widehat{B'} = \hat{U}(\theta'_B, \phi'_B, \alpha'_B) = = \hat{U}\left(\theta_A + \pi, \alpha_A - \frac{\pi}{2}, \phi_A - \pi\right)$, gdyż wynik gry jest znowu $(\$_{A,}\$_B) = (f, z)$. W końcu najlepszą odpowiedzią na $\widehat{B'}$ jest pierwotna strategia Alicji, która przywraca jej wygraną $(\$_{A,}\$_B) = (z, f)$. Załóżmy teraz, że Alicja wybierze strategię mieszającą obie swoje strategie $\cos^2\frac{\gamma_A}{2}\hat{A} + \sin^2\frac{\gamma_A}{2}\widehat{A'}$, $\gamma_A \in [0, \pi]$, a Bartek postąpi podobnie, grając $\cos^2\frac{\gamma_B}{2}\hat{B} + \sin^2\frac{\gamma_B}{2}\widehat{B'}$, $\gamma_B \in [0, \pi]$. Wartość oczekiwana wygranej Alicji (5) jest równa:

$$\$_A = f\left(\cos^2\frac{\gamma_A}{2}\cos^2\frac{\gamma_B}{2} + \sin^2\frac{\gamma_A}{2}\sin^2\frac{\gamma_B}{2}\right) + \\ + z\left(\cos^2\frac{\gamma_A}{2}\sin^2\frac{\gamma_B}{2} + \sin^2\frac{\gamma_A}{2}\cos^2\frac{\gamma_B}{2}\right). \quad (7)$$

Zauważmy, że suma wygranych Alicji i Bartka w tej grze jest stała i wynosi $\$_A + \$_B = z + f$. Na rysunku 4 przedstawiono wypłatę Alicji przy założeniu, że: $z = 5$, $w = 3$, $k = 1$ i $f = 0$, wypłata Bartka jest równa $5 - \$_A$. Gra ma

jedną równowagę Nasha dla $\gamma_A = \gamma_B = \frac{\pi}{2}$, w tym punkcie wypłaty graczy są $\$_A = \$_B = \frac{f+z}{2}$ i żaden z graczy nie powiększy swojej wypłaty przez jednostronną zmianę swojej strategii [Flitney i in. 2002]. Wystarczy, że jeden z graczy zastosuje strategię $\gamma_X = \frac{\pi}{2}$, aby wygrana obu była równa $\frac{f+z}{2}$ i niezależna od strategii drugiego (czerwone linie na rys. 4).



Rys. 4. Wypłata Alicji dla kwantowego DW, w którym przeciwnicy grają strategiami mieszanymi $\cos^2\frac{\gamma_A}{2}\hat{A} + + \sin^2\frac{\gamma_A}{2}\hat{A}'$ i $\cos^2\frac{\gamma_B}{2}\hat{B} + \sin^2\frac{\gamma_B}{2}\hat{B}'$, gdzie $\gamma_A, \gamma_B \in [0, \pi]$. Gra ma jedyną RN w punkcie siodłowym $\gamma_A = \gamma_B = \frac{\pi}{2}$ (czerwony znacznik)

Gra w tej postaci sprowadza się więc do gry o sumie stałej, a jej rozwiązaniem jest punkt siodłowy. Kwantowy DW ma nieskończenie wiele równowag Nasha, każdą wyznaczoną przez trójkę pierwotnych strategii Alicji $(\theta_A, \phi_A, \alpha_A)$. Pokazaliśmy zatem, że odpowiedni wybór mieszanych strategii kwantowych może zapewnić obu graczom wynik tylko nieco gorszy od wzajemnej współpracy (pamiętajmy, że $w > \frac{z+f}{2}$ z definicji DW). Biorąc jednak pod uwagę, że w grze klasycznej jedyną równowagą Nasha jest wynik (k, k), gra kwantowa daje graczom dużo korzystniejszą równowagę – nieosiągalną w grze klasycznej.

Zauważmy, że jeśli w szczególnym przypadku ustalimy $(\theta_A, \phi_A, \alpha_A) = = (0, 0, 0)$:

$$\hat{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \hat{I}, \hat{B} = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = -i\widehat{\sigma_x} \tag{8}$$

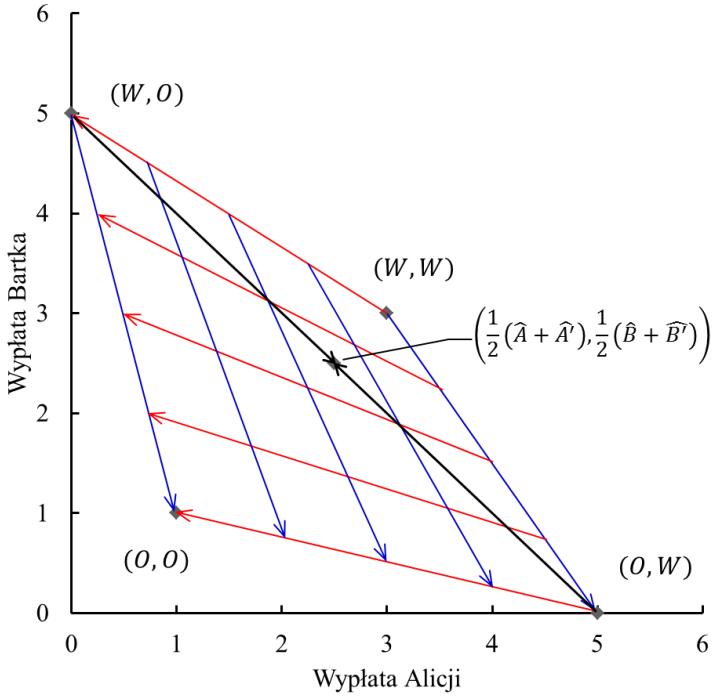
oraz:

$$\widehat{A'} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = i\widehat{\sigma_z}, \widehat{B'} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\widehat{\sigma_y} \quad (9)$$

to strategie $\hat{A}, \widehat{A'}, \hat{B}, \widehat{B'}$, z dokładnością do stałej, sprowadzają się do czterech macierzy (macierz jednostkowa plus trzy macierze Pauliego) z grupy macierzy unitarnych $SU(2)$. Macierze Pauliego są generatorami obrotów o kąt π wokół odpowiednich osi (rys. 1).

5. Czy można wykorzystać kwantowe równowagi DW?

Na rysunku 5 przedstawiono diagram użyteczności DW (dla $z = 5$, $w = 3$, $k = 1$ i $f = 0$) z zaznaczeniem czterech klasycznych strategii $(W, O), (W, W), (O, W)$ i (O, O). Seria linii tworzących widzianą pod kątem szachownicę odpowiada parom klasycznych strategii graczy $(\theta_{A,}\theta_B)$ o stałym θ_A (linie czerwone) i stałym θ_B (linie niebieskie). Strzałki wskazują preferencje graczy i jedyną klasyczną równowagę Nasha w punkcie (O, O). Czarna linia łącząca punkty (W, O) i (O, W) odpowiada grze o stałej sumie $\$_A + \$_B = z + + f = 5$ i zawiera wszystkie wypłaty kwantowych strategii $\cos^2\frac{\gamma_A}{2}\hat{A} + \sin^2\frac{\gamma_A}{2}\widehat{A'}$ Alicji oraz $\cos^2\frac{\gamma_B}{2}\hat{B} + \sin^2\frac{\gamma_B}{2}\widehat{B'}$ Bartka.



Niezależnie od strategii mieszanej Alicji (linie czerwone) dla Bartka najkorzystniejszą strategią (strzałki) jest O, podobnie w przypadku strategii mieszanych Bartka (linie niebieskie) dla Alicji najkorzystniejszą strategią jest O. Kwantowy DW ma równowagę Nasha dla pary strategii $\left(\frac{1}{2}(\hat{A}+\hat{A}'),\frac{1}{2}(\hat{B}+\hat{B}')\right)$ korzystniejszą od klasycznego (*0*, *0*).

Rys. 5. Diagram użyteczności DW dla $z = 5$, $w = 3$, $k = 1$ i $f = 0$

Jak wykazaliśmy w poprzednim rozdziale, jedyna równowaga Nasha odpowiada w tym przypadku parze strategii $\left(\frac{1}{2}(\hat{A}+\hat{A}'),\frac{1}{2}(\hat{B}+\hat{B}')\right)$. Równowaga ta daje obu graczom wygraną $\frac{f+z}{2} = 2{,}5$, korzystniejszą od klasycznej równowagi dzięki charakterystycznemu splątaniu klasycznych strategii (W, O) i (O, W) niemożliwemu do uzyskania za pomocą strategii klasycznych.

Sprowadzenie DW do gry o stałej sumie jest możliwe tylko kwantowo. Analiza tego rozwiązania pokazuje, że jego istotą jest specyficzna korelacja rozwiązań typu „współpraca” – „odmowa” w taki sposób, aby gracze nie potrafili przewidzieć, kiedy ich strategia doprowadzi do jednej bądź do drugiej opcji. Jeśli Alicja wybiera strategię $\hat{A}$, to nie wie, czy Bartek odpowie jej $\hat{B}$ czy $\hat{B}'$, tym samym nie wie, czy jej ruch jest „współpracą” czy „odmową”. To samo dotyczy strategii $\hat{A}'$ oraz sytuacji Bartka. Przy takim wyborze strategii kwantowy DW

sprowadza się więc do gry o sumie zerowej w wybór strony monety: dwaj gracze niezależnie od siebie wybierają orła lub reszkę; Alicja wygrywa, jeśli wybrane strony są różne, a Bartek – jeśli są takie same (z ang. „matching pennies"). Gra ta jest bardziej znana (w wersji z potrójnym wyborem) jako „papier, nożyce, kamień". Jak wiadomo, nie ma ona rozwiązania w strategiach czystych, a jej jedyną równowagą Nasha jest strategia losowa – wybór dowolnej opcji z jednakowym prawdopodobieństwem. Jeśli jeden z graczy zastosuje tę strategię, to drugi, niezależnie od swojej, nie może podwyższyć wyniku, którego wartość oczekiwana wynosi w tym przypadku zero.

W naturalny sposób pojawia się pytanie, czy kwantowa wersja DW może się przyczynić do rozwiązywania typowych sytuacji z życia codziennego, które mają charakter dylematu więźnia. Przykłady z innych dziedzin wskazują, że strategie kwantowe mogą lepiej niż klasyczne rozwiązywać problemy gier rynkowych i giełdy [Piotrowski i in. 2002], aukcji i konkursów [Piotrowski i in. 2008] czy hazardu [Goldenberg i in. 1999]. Jak wynika z niniejszego artykułu, rozwiązanie dylematu za pomocą kwantowych strategii może dać lepsze rezultaty niż klasyczne rozwiązania. Klasyczny DW w nieuchronny sposób doprowadza graczy do jedynego racjonalnego rozwiązania obustronnej odmowy współpracy i w konsekwencji kary za brak współpracy, podczas gdy jego kwantowy odpowiednik ma równowagi Nasha na dużo korzystniejszym poziomie średniej z „pokusy do zdrady" i „nagrody frajera". Czy zatem strategie kwantowe mogą być praktycznie wykorzystane?

Jak wiemy, jednym z pozytywnych sposobów, w jaki DW reguluje procesy rynkowe, jest równowaga cen, tzw. piękna równowaga Nasha. Dzięki niej firmy, które chciałyby sprzedawać swoje towary drożej, de facto obniżają ceny, aby optymalizować swoje zyski [Dixit i in. 2009]. Dylemat więźnia polega tu na tym, że obopólna (lub wielostronna) współpraca, polegająca na utrzymywaniu wysokich cen, jest w sytuacji rynkowej niemal niemożliwa, bo zawsze znajdzie się firma, która zechce sprzedawać taniej („odmowa"), rekompensując sobie niższe ceny zwiększoną ilością klientów i pozostawiając „na lodzie" (tzn. bez klientów) droższego producenta („współpraca"). Jednak w niektórych przypadkach działanie mechanizmu DW jest w tej dziedzinie zaburzone, dochodzi do niego w przypadku tzw. zmowy cenowej. Znany jest przykład z lat pięćdziesiątych ubiegłego wieku na rynku turbin w USA [Dixit i in. 2009]. Trzy firmy umówiły się, że będą stosować zawyżone ceny, lecz takie, aby, w zależności od terminu ogłoszenia przetargu, wygrała jedna z nich. Wygrany w danym przetargu brał wszystko (pokusa do zdrady), pozostali zostawali z niczym (nagroda frajera). Losowość terminów ogłaszania przetargów zapewniała, że wszyscy partnerzy zarabiali, każdy w stosownym terminie. Najważniejsze w całej sprawie było odpowiednie skorelowa-

nie firmy, która miała dany przetarg wygrać, z terminem jego ogłoszenia w taki sposób, aby zainteresowani nie mieli wątpliwości, a nikt poza nimi nie był w stanie przewidzieć algorytmu (bo zmowy cenowe są nielegalne). Gdyby partnerzy zmowy wykorzystali algorytm oparty na splątaniu kwantowym, żaden zewnętrzny obserwator nie byłby w stanie udowodnić im zmowy cenowej. Jednak dyrektorzy firm wylądowali w więzieniu, gdyż wykorzystali mniej wyszukany, a możliwy do odkrycia system korelujący – zwycięzcę przetargu wyłaniano na podstawie kalendarza księżycowego – ta czy inna firma wygrywała w zależności od ilości dni, które upłynęły od nowiu.

Pozytywnym przykładem działania, które znajduje analogię do rozwiązania kwantowego DW w zagadnieniach ekonomicznych, jest tzw. strategia szachowa w negocjacjach [Perrotin i in. 1994]. W strategii tej zakłada się, że niezbędnym elementem każdych negocjacji są obustronne ustępstwa, a sztuka negocjowania polega na ich umiejętnym doborze. Ustępstwa dotyczące poszczególnych kwestii (zmiennych) negocjacyjnych są skategoryzowane ze względu na ich ważność dla obu stron i naniesione na macierz strategii szachowej, której kolejne wiersze oznaczają kwestie: najważniejsze, średnio ważne i najmniej ważne dla Alicji. Kolumny w analogiczny sposób porządkują ważność kwestii dla Bartka. Zmienne, których ważność dla Alicji jest mniejsza niż dla Bartka, leżą poniżej przekątnej macierzy strategii, a te, których ważność dla Alicji jest większa niż dla Bartka, powyżej przekątnej [Szopa 2010]. Zastosowanie strategii szachowej polega m.in. na wzajemnej wymianie ustępstw tak, aby Alicja ustępowała w kwestiach poniżej, a Bartek w kwestiach powyżej przekątnej. Dzięki takiej dystrybucji ustępstw suma ich „ważności" jest minimalizowana, a wynik negocjacji zamiast klasycznego „spotkania w połowie drogi" zbliża się do rozwiązania optymalnego typu „wygrany-wygrany". Analogia do równowagi kwantowego DW jest następująca: ustępstwa Alicji to jej „współpraca", a Bartka „odmowa", w przypadku ustępstw Bartka role się odwracają. Gracze, ustępując sobie kolejno, grają kwantową sekwencją $\hat{A}$ - $\hat{B}$ - $\hat{A}'$- $\hat{B}'$ -…, gdzie każda kolejna para strategii oznacza ustępstwo jednej i wygraną drugiej strony. Sekwencyjna wymiana ustępstw nie jest niczym nowym w relacjach społecznych, wynika z głęboko zakorzenionej w naturze ludzkiej korelacji wzajemnych życzliwości, znanej w psychologii społecznej jako reguła wzajemności [Cialdini 1995].

Podsumowanie

Przeprowadzona w niniejszym artykule analiza pokazuje, że do zasymulowania strategii kwantowych w grach wystarczą klasyczne rachunki. Dzięki znajomości mechaniki kwantowej jesteśmy w stanie symulować zachowanie się cząstek kwantowych i, co za tym idzie, przewidywać wynik stosowania strategii kwantowych. Znając wynik działania tych algorytmów, nawet jeśli ich fizycznie nie implementujemy, możemy się starać je naśladować, aby wykorzystać szerszą klasę możliwych kwantowo rozwiązań gier. W artykule opisaliśmy dwa przykłady klasycznych procesów – zmowy cenowe oraz strategię szachową, które w dużym stopniu wykorzystują równowagi Nasha kwantowego DW.

Można zadać pytanie: jaki jest związek gier klasycznych ze zjawiskami kwantowymi? Jako teoria matematyczna, gry klasyczne okazują się być szczególnym przypadkiem gier kwantowych. Czy realne gry klasyczne, rozgrywane codziennie przez ludzi, mają jakiś związek z fizycznymi procesami kwantowymi? Odpowiedź na to pytanie wydaje się być twierdząca – takim procesem może być kolaps funkcji falowej. Według oszacowań [Albrecht i in. 2012] to fluktuacje kwantowe są przyczyną zjawisk makroskopowych, które uznajemy za losowe, takich jak np. rzut monetą lub kością. Według cytowanych autorów każde praktyczne użycie prawdopodobieństwa ma swoje źródło w zjawiskach kwantowych. Gdyby przyjąć taki punkt widzenia, to każde wykorzystanie w grze strategii mieszanej byłoby w istocie zjawiskiem kwantowym.

W grach kwantowych istotnym elementem mechanizmu gry jest splątanie. Czy to zjawisko również ma swój odpowiednik w realnej grze klasycznej? Czy obiekty makroskopowe, które są kontrolowane bądź tylko obserwowane przez nasze zmysły, mogą być splątane? Tego nie potrafimy udowodnić. Problemy z dekoherencją funkcji falowej powodują, że nawet na poziomie ściśle kontrolowanych eksperymentów, odbywających się w skrajnych rygorach odizolowania od otoczenia, trudno jest utrzymać dwa splątane kubity. Budowa komputera kwantowego, opartego na rejestrze wielu splątanych kubitów, poddanych unitarnym operacjom bramek kwantowych, i zdolnego do rozwiązywania za pomocą algorytmów kwantowych praktycznych problemów lub symulowania gier kwantowych jest prawdziwym wyzwaniem, które jednak współczesna fizyka nie bez sukcesów podejmuje [Vandersypen i in. 2001].

Bibliografia

Albrecht A., Phillips D., 2012: Origin of Probabilities and Their Application to the Multiverse [Online]. Cornell University Library, http://arxiv.org/pdf/1212.0953 v1.pdf [dostęp: 12.2012].

Busemeyer J.R., Wang Z., Townsend J.T., 2006: Quantum Dynamics of Human Decision Making. „Journal of Mathematical Psychology", Vol. 50, 220-241.

Chen K., Hogg T., 2006: How Well Do People Play a Quantum Prisoner's Dilemma. „Quantum Information Processing", Vol. 5(1), 43-67.

Cialdini R., 1995: Wywieranie wpływu na ludzi. Gdańskie Wydawnictwo Psychologiczne, Gdańsk.

Dixit A.K., Nalebuff B.J., 2009: Sztuka strategii. MT Biznes, Warszawa.

Du J. et al., 2002: Experimental Realization of Quantum Games on Quantum Computer. „Physical Review Letters", Vol. 88, 137902.

Eisert J., Wilkens M., Lewenstein M., 1999: Quantum Games and Quantum Strategies. „Physical Review Letters", Vol. 83, 3077, s. 3077.

Flitney A.P., Abbott D., 2002: An Introduction to Quantum Game Theory. „Fluct. Noise Lett", Vol. 2, R175-87.

Flood M.M., Dresher M., 1952: Research Memorandum. RM-789-1-PR. RAND Corporation, Santa-Monica, Ca.

Goldenberg L., Vaidman L., Wiesner S., 1999: Quantum Gambling. „Physical Review Letters", Vol. 82, 3356.

Hamilton W.D., Axelrod R., 1981: The Evolution of Cooperation. „Science", Vol. 211.27, 1390-1396.

Perrotin R., Heusschen P., 1994: Kupić z zyskiem. Poltext, Warszawa.

Piotrowski E., Sładkowski J., 2008: Quantum Auctions: Facts and Myths. „Physica A", 15, Vol. 387, 3949-3953.

–, 2002: Quantum Market Games. „Physica A", 1-2, Vol. 312, 208-216.

Pothos E.M., Busemeyer J.R., 2009: A Quantum Probability Explanation for Violations of 'Rational' Decision Theory. Proceedings of the Royal Society B., Vol. 276, 2171-2178.

Straffin P.D., 2001: Teoria gier. WN Scholar, Warszawa.

Szopa M., 2010: Teoria gier w negocjacjach. Skrypty dla studentów Ekonofizyki na Uniwersytecie Śląskim [Online]. http://el.us.edu.pl/ekonofizyka/index.php/ Teoria_gier/strategie_taktyki_negocjacji.

Vandersypen L.M.K. et al., 2001: Experimental Realization of Shor's Quantum Factoring Algorithm Using Nuclear Magnetic Resonance. „Nature", 6866, Vol. 414, 883-887.

WHY IS IT WORTH PLAYING QUANTUM PRISONER'S DILEMMA?

Summary

The Prisoner's Dilemma [PD] is the best known example of a two-person, simultaneous game, for which the Nash equilibrium is far from Pareto-optimal solutions. In this paper we define a quantum PD, for which player's strategies are defined as rotations of the SU(2) group, parameterized by three angles. Quantum strategies are correlated through the mechanism of quantum entanglement and the result of the game is obtained by the collapse of the wave function. Classic PD is a particular case of the quantum game for which the set of rotations is limited to one dimension. Each quantum strategy can be, by appropriate choice of counter-strategy, interpreted as a "cooperation" or "defection". Quantum PD has Nash equilibria that are more favorable than the classic PD and close to the Pareto optimal solutions. With proper selection of strategies, quantum PD can be reduced to the classic, zero-sum, "matching pennies" game. In this paper we show examples of economic phenomena (price collusion, the chess strategy) that mimics the Nash equilibria of quantum PD.

INNE ZASTOSOWANIA RYZYKA

Mariusz Doszyń
Krzysztof Dmytrów
Uniwersytet Szczeciński

PORÓWNYWANIE EFEKTYWNOŚCI PROGNOZ EX POST WIELKOŚCI SPRZEDAŻY W PEWNYM PRZEDSIĘBIORSTWIE WYZNACZONYCH ZA POMOCĄ ROZKŁADU GAMMA I MODELI Z WAHANIAMI SEZONOWYMI

Wstęp

Gospodarowanie zapasami i zakupami produktów jest istotnym z punktu widzenia kosztów elementem zarządzania przedsiębiorstwem. Ma ono istotne znaczenie zarówno w przypadku przedsiębiorstwa produkcyjnego, w którym zapasy materiałów, surowców czy półproduktów zapewniają ciągłość procesu produkcyjnego, jak i dla przedsiębiorstwa zajmującego się dystrybucją produktów końcowych. Tutaj zapasy zapewniają pokrycie zapotrzebowania zgłaszanego przez klientów. W pierwszym przypadku popyt na zapasy produkcyjne jest popytem zależnym (zależy on od popytu na produkty końcowe, na podstawie którego ustala się popyt na materiały, surowce czy półprodukty). W drugim przypadku mamy do czynienia z zapasami dystrybucyjnymi. Popyt na nie jest popytem niezależnym. Prognozuje się go za pomocą znanych metod, np. metod analizy szeregów czasowych czy prognozowania na podstawie rozkładów (empirycznych lub teoretycznych).

W klasycznym modelu zapasów należy odpowiedzieć na pytanie: jak często i ile należy zamawiać, żeby zaspokoić w danym stopniu zapotrzebowanie zgła-

szane na produkty [zob. Sarjusz-Wolski 2000]. Zakładamy, że znamy wartości jednostkowych kosztów magazynowania oraz niedoborów produktów, znamy także koszty złożenia i realizacji pojedynczego zamówienia. Znajomość tych kosztów ma bardzo istotne znaczenie, ponieważ uzupełniając zapasy, możemy albo zamawiać częściej mniejsze partie, albo zamawiać rzadziej większe partie. W pierwszym przypadku ponosimy niższe koszty magazynowania (ponieważ mamy mniejsze zapasy), za to ponosimy wyższe koszty zamawiania (bo zamawiamy częściej). W drugim zaś przypadku mamy sytuację odwrotną – ponosimy niższe koszty zamawiania (ponieważ realizujemy je rzadziej), za to godzimy się na wyższe koszty magazynowania. Znajomość kosztów pozwala nam wyznaczyć optymalną wielkość zamawiania, przy której łączne koszty będą minimalne.

W badanym przedsiębiorstwie jednak nie można zastosować klasycznego podejścia w celu wyznaczenia optymalnej wielkości zamówienia za pomocą któregoś ze znanych modeli gospodarowania zapasami. Jest to spowodowane specyfiką działania niniejszej firmy. Jest ona jednym z oddziałów dużego przedsiębiorstwa i zajmuje się dystrybucją produktów sprowadzanych z centrali. Centrala, chcąc ujednolicić system uzupełniania zapasów, zdecydowała, że oddziały mają składać zamówienia uzupełniające co pięć tygodni i w każdym zamówieniu zamawiać wszystkie produkty, których zapasy wymagają uzupełnienia.

Z powyższych powodów kluczowym elementem efektywnego gospodarowania zapasami w badanym przedsiębiorstwie jest prawidłowe wyznaczanie przyszłego, pięciotygodniowego poziomu sprzedaży. Brane są tutaj pod uwagę różne kryteria. Przede wszystkim należy dążyć do w miarę pełnego zaspokojenia zapotrzebowania zgłaszanego na produkty przez odbiorców, przy jednoczesnym kontrolowaniu poziomu zapasów. W analizowanym przedsiębiorstwie, z ofertą składającą się z wielu produktów, pojawia się konieczność tworzenia systemu prognoz pozwalających na przewidywanie przyszłej sprzedaży. Większość ilościowych metod prognostycznych bazuje na, znanych z literatury, ekonometrycznych metodach analizy szeregów czasowych, jednak metody te nie zawsze dają zadowalające rezultaty, dlatego w analizowanym przypadku postanowiono porównać wyniki prognozowania za pomocą jednej z metod analizy szeregów czasowych z wynikami prognozowania na podstawie rozkładów.

1. Opis zastosowanych metod i procedur

Przeprowadzone badanie dotyczy efektywności prognoz ex post wyznaczanych dla okresów 5-tygodniowych. Przyjęcie takiego horyzontu prognozy jest

związane z długością cyklu składania zamówień. Wykorzystane informacje statystyczne dotyczą przedsiębiorstwa, w którego asortymencie znajduje się ponad 10 000 produktów. Obserwacje odnoszą się do tygodniowych wielkości sprzedaży poszczególnych produktów. Okres objęty analizą obejmuje 210 tygodni (ostatni tydzień, z którego pochodzą dane, to 3-10 II 2013 r.). Długość poszczególnych szeregów czasowych jest często różna. Jest to związane z ciągłymi zmianami asortymentu. Są wprowadzane nowe produkty, z kolei produkty, na które nie jest zgłaszany popyt, są wycofywane. Wyznaczane prognozy dotyczą produktów aktualnie dostępnych w asortymencie. Dane nie są zagregowane, ukazują sprzedaż każdego z produktów, w związku z czym często bardzo trudno doszukać się jednoznacznych prawidłowości w zakresie kształtowania się wielkości sprzedaży w czasie.

W pierwszym etapie badania zostały wyselekcjonowane te produkty, w przypadku których odpowiednie testy statystyczne wskazywały na występowanie istotnych statystycznie wahań sezonowych oraz na zgodność empirycznego rozkładu wielkości sprzedaży z rozkładem gamma[*]. Powodem wyboru liniowego modelu trendu ze stałą sezonowością była po pierwsze prostota obliczeń i interpretacji modelu, a po drugie dla danych tygodniowych branie pod uwagę dodatkowo zmiennej sezonowości spowodowałoby prawie dwukrotne zwiększenie liczby szacowanych parametrów, a co za tym idzie, zmniejszenie liczby stopni swobody, co nawet przy dużej liczbie obserwacji (210) znacznie pogorszyłoby własności analityczne i prognostyczne modeli. Z kolei powodem wyboru rozkładu gamma był fakt, iż jest to najczęściej występujący w praktyce rozkład zapotrzebowania [zob. Całczyński 2000]. Nawet gdyby rozkład zapotrzebowania nie był znany, to ze względu na swoją uniwersalność zastosowanie rozkładu gamma byłoby uzasadnione [zob. Tadikamalla 1984]. Jest to spowodowane kilkoma czynnikami. W bardzo wielu przypadkach występuje duża liczba okresów, w których nie występowało zapotrzebowanie na produkty oraz relatywnie rzadko zdarzały się okresy, w których zapotrzebowanie było bardzo wysokie (w odniesieniu np. do mediany). Rozkłady charakteryzowały się więc w większości przypadków skrajną asymetrią prawostronną, a co za tym idzie, bardzo dużą zmiennością (ze współczynnikami zmienności wynoszącymi nawet powyżej 1000%). Po przyjęciu poziomu istotności $\alpha = 0{,}1$ w przypadku 72 produktów wielkość sprzedaży wykazywała istotne wahania sezonowe, a rozkłady wielkości sprzedaży były zgodne z rozkładem gamma.

[*] Wszystkie obliczenia zostały wykonane z wykorzystaniem skryptów napisanych w języku *hansl*, który jest dostępny w pakiecie do obliczeń ekonometrycznych *Gretl*.

Wśród analizowanych produktów bardzo zróżnicowana była częstość zamówień, rozumiana jako udział tygodni z niezerowymi zamówieniami w liczbie tygodni, w których produkt był oferowany do sprzedaży. Znaczna cześć produktów była zamawiana bardzo rzadko, w wielu przypadkach dominowały pojedyncze zamówienia. Test na występowanie wahań sezonowych oraz na zgodność rozkładu zamówień z rozkładem gamma był stosowany tylko do tych produktów, w przypadku których częstość zamówień była większa od 0,5, czyli zamówienia były składane średnio częściej niż co dwa tygodnie. Jeżeli zamówienia były składane rzadziej, wówczas stosowano inne metody prognozowania sprzedaży, w której wykorzystywano zarówno poziom zamówień niezerowych, jak i średnie odstępy pomiędzy nimi.

Model ze zmiennymi sztucznymi, w którym uwzględnia się wahania tygodniowe, można zapisać następująco:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{51} d_i Q_{it} + u_t \tag{1}$$

gdzie:
y_t – wielkość sprzedaży w tygodniu *t*,
α_0, α_1, d_i $(i = 1,..,51)$ – parametry modelu,
t – zmienna czasowa,
Q_{it} – tygodniowe zmienne zero-jedynkowe,
u_t – składnik losowy.

Model (1) jest szacowany przy założeniu, że $\sum_{i=1}^{52} d_i = 0$. Zmienna Q_{it}=1 w *i*-tym tygodniu oraz zero w pozostałych tygodniach (poza 52). Zmienna Q_{it} dla ostatniego podokresu (52 tydzień) jest pomijana, a wartość pozostałych zmiennych sztucznych w tym podokresie jest równa –1. Odchylenie dla ostatniego podokresu można wyznaczyć ze wzoru: $d_{52} = -\sum_{i=1}^{51} d_i$. W modelu (1) założono liniową funkcję trendu, aby uwzględnić ewentualne systematyczne zmiany poziomu zamówień, choć analiza wyselekcjonowanych szeregów czasowych wskazuje na to, iż wielkości zamówień oscylowały zazwyczaj wokół pewnego stałego poziomu.

Test na istotność wahań sezonowych bazuje na rozkładzie *F*-Fishera-Snedecora, w którym bada się łączną istotność wpływu sztucznych (tygodniowych) zmiennych zero-jedynkowych:

$$F = \frac{(SSR_r - SSR_u)/j}{SSR_u/df_u} \tag{2}$$

gdzie:

SSR_r – suma kwadratów reszt w modelu z restrykcjami (w modelu bez zmiennych zero-jedynkowych, w którym nie uwzględnia się wahań sezonowych),

SSR_u – suma kwadratów reszt w modelu bez restrykcji (w modelu ze zmiennymi zero-jedynkowymi, w którym uwzględnia się wahania sezonowe),

j – liczba restrykcji,

df_u – liczba stopni swobody w modelu bez restrykcji.

Prognozy były wyznaczane również na poziomie mediany obliczanej na podstawie funkcji gęstości rozkładu gamma:

$$f(y;\lambda,\eta) = \frac{\lambda^{\eta}}{\Gamma(\eta)} y^{\eta-1} e^{-\lambda y}, y \geq 0, \lambda > 0, \eta > 0 \tag{3}$$

gdzie $y \geq 0, \lambda > 0, \eta > 0$.

Parametry kształtu (λ) oraz skali (η) są równe odpowiednio:

$$\lambda = \frac{\bar{y}^2}{S^2(y)} \tag{4}$$

$$\eta = \frac{S^2(y)}{\bar{y}} \tag{5}$$

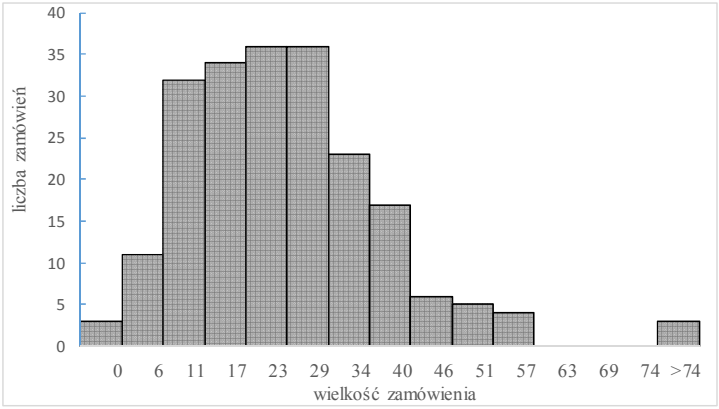
Hipoteza o zgodności rozkładu wielkości zamówień z rozkładem gamma była weryfikowana na podstawie nieparametrycznego testu Locke'a, w którym przyjmuje się, iż parametry rozkładu nie są znane [Locke 1976; Shapiro, Chen 2001]. Test ten odwołuje się do następującej zasady: jeżeli zmienne *X* i *Y* są niezależne, to zmienne *X*/*Y* i *X*+*Y* są niezależne wtedy, gdy mają rozkład gamma [Locke 1976]. Test Locke'a może być stosowany wtedy, gdy parametry rozkładu nie są znane, ponieważ powyższa własność cechuje wszelkiego typu rozkłady gamma.

W teście tym próba jest dzielona losowo na dwie podpróby o takiej samej liczebności *n*. Jeśli liczba obserwacji jest nieparzysta, jedna z nich jest losowo pomijana, co może się przyczyniać do tego, że kolejne zastosowania tego testu mogą prowadzić do nieznacznie różniących się wyników. W kolejnym etapie są tworzone pary następujących zmiennych: $U_i = X_{2i-1} + X_{2i}$; $V_i = max\left\{\frac{X_{2i-1}}{X_{2i}}, \frac{X_{2i}}{X_{2i-1}}\right\}$. Niezależność zmiennych U_i i V_i bada się na podstawie korelacji rang Kendalla. Szczegółowy opis całej procedury zawiera artykuł Locke'a [1976].

Test Locke'a był stosowany wtedy, gdy liczba obserwacji była większa od 30, a częstość tygodniowa sprzedaży (rozumiana jako udział tygodni z dodatnią wielkością sprzedaży) przekraczała 0,5. W przypadku wyników uzyskanych po

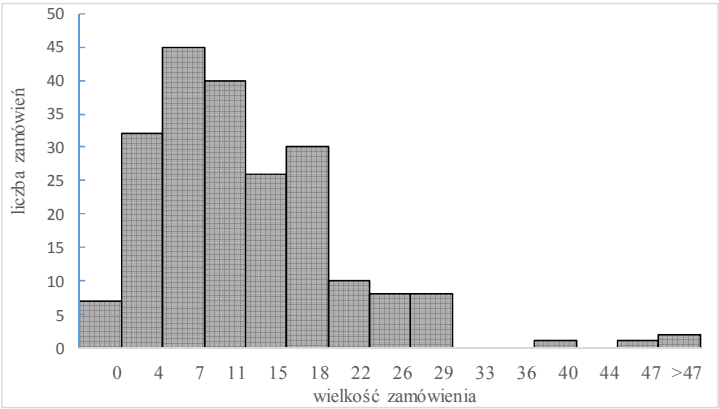
zastosowaniu testu Locke'a można stwierdzić, iż hipotezy o zgodności rozkładu wielkości zamówień z rozkładem gamma nie można było odrzucić w przypadku 343 produktów (poziom istotności 0,1)*.

Rozkłady wielkości zamówień przykładowych produktów przedstawiono na rys. 1-4.



Rys. 1. Rozkład wielkości zamówień produktu *a*

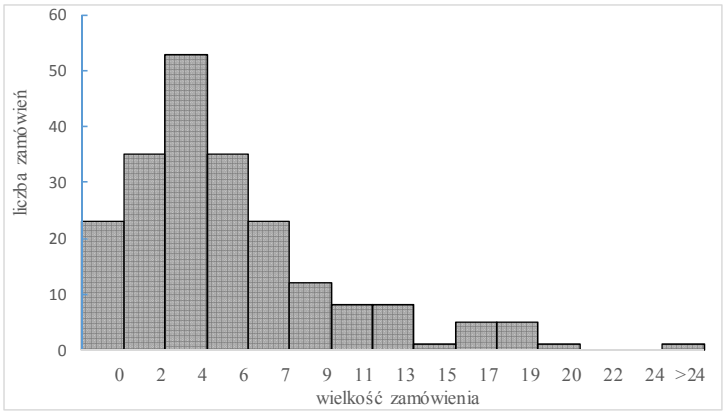
Źródło: Opracowanie własne.



Rys. 2. Rozkład wielkości zamówień produktu *b*

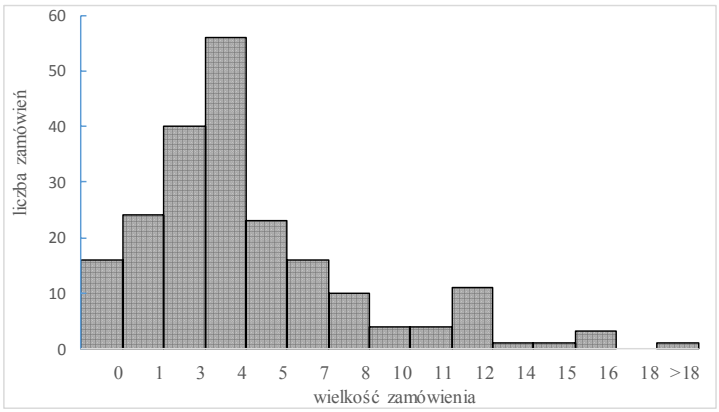
Źródło: Opracowanie własne.

* Szczegółowe wyniki nie zostały tu zamieszczone ze względu na ich dużą objętość.



Rys. 3. Rozkład wielkości zamówień produktu *c*

Źródło: Opracowanie własne.



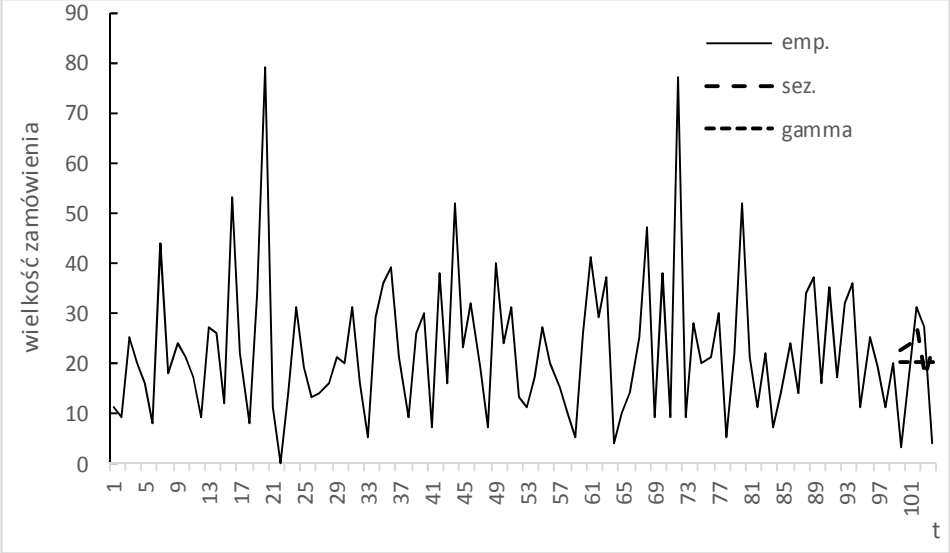
Rys. 4. Rozkład wielkości zamówień produktu *d*

Źródło: Opracowanie własne.

2. Charakterystyka otrzymanych prognoz

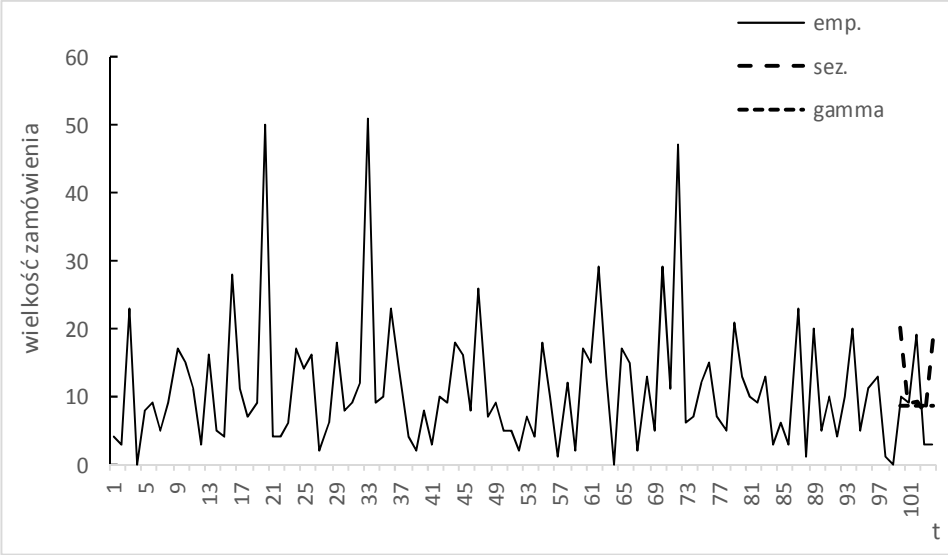
Prognozy wielkości sprzedaży na kolejnych 5 tygodni zostały wyznaczone dla 72 produktów zarówno z wykorzystaniem modelu z wahaniami sezonowymi, jak i na podstawie rozkładu gamma. Tylko w przypadku 72 produktów test *F* wskazywał na istotność wahań sezonowych, a test Locke'a na zgodność rozkładu empirycznego z rozkładem gamma. Szeregi czasowe skrócono o 5 ostatnich tygodni. Na

podstawie wcześniejszych obserwacji dokonano prognoz na 5 tygodni, po czym porównano prognozy z rzeczywistymi wartościami sprzedaży w tych tygodniach. W przypadku rozkładu gamma za prognozę była przyjmowana mediana wyznaczona dla otrzymanej funkcji gęstości. Można zadać pytanie, dlaczego przyjęto właśnie ten poziom kwantyla. Powodów było kilka. Po pierwsze – w toku prowadzonych badań i dokonywanych symulacji stwierdzono, że właśnie ten poziom kwantyla daje średnio najmniejsze wielkości błędów prognoz. Po drugie – mimo że w wielu przypadkach poziom prognoz na poziomie mediany może być zaniżony, to jednak w wielu przypadkach zamówienia nie występują w każdym tygodniu (a prognozę na podstawie rozkładu gamma uzyskujemy na każdy tydzień). Przykładowe prognozy wraz z wielkościami zamówień w ostatnich 104 tygodniach wyznaczone za pomocą każdej z metod przedstawiają rys. 1-4.



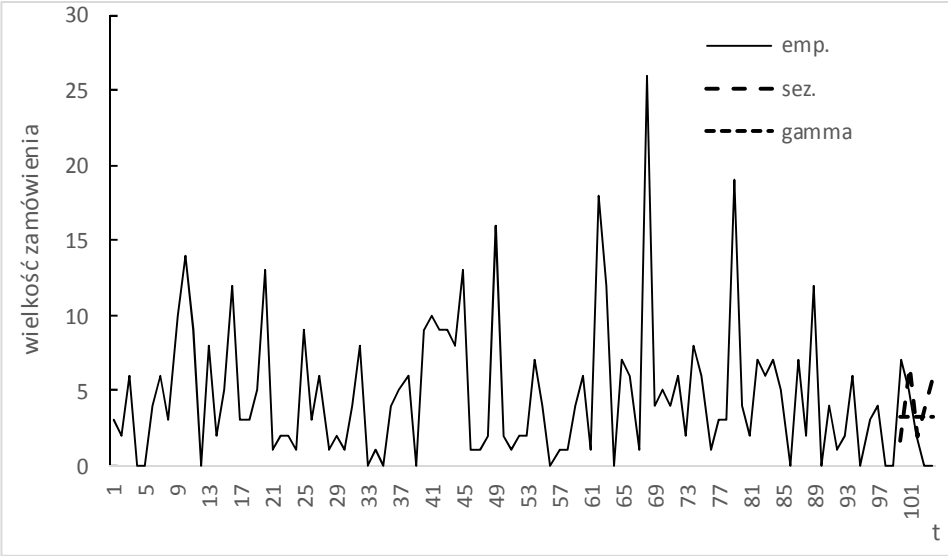
Rys. 5. Dynamika wielkości zamówień produktu *a* wraz z prognozami wyznaczonymi na podstawie modelu z wahaniami sezonowymi oraz rozkładu gamma

Źródło: Opracowanie własne.



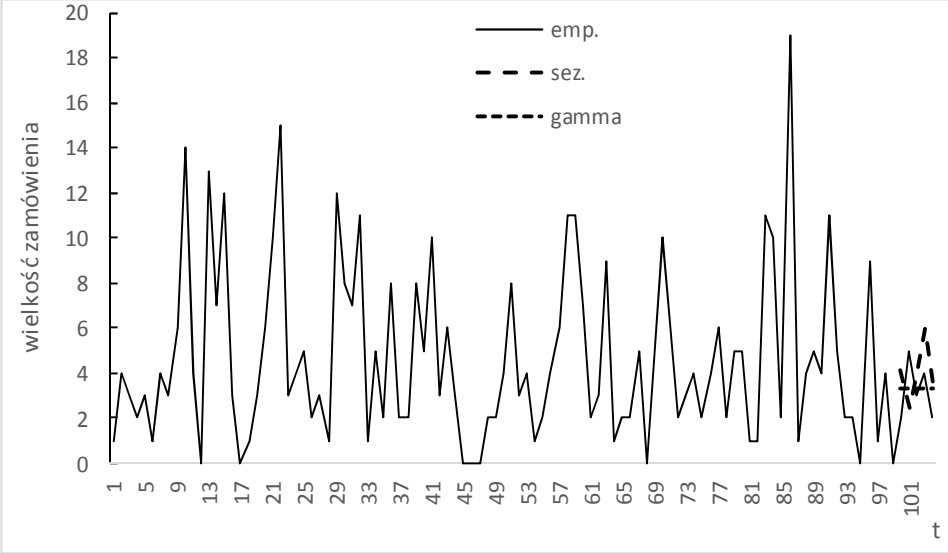
Rys. 6. Dynamika wielkości zamówień produktu *b* wraz z prognozami wyznaczonymi na podstawie modelu z wahaniami sezonowymi oraz rozkładu gamma

Źródło: Opracowanie własne.



Rys. 7. Dynamika wielkości zamówień produktu *c* wraz z prognozami wyznaczonymi na podstawie modelu z wahaniami sezonowymi oraz rozkładu gamma

Źródło: Opracowanie własne.



Rys. 8. Dynamika wielkości zamówień produktu *d* wraz z prognozami wyznaczonymi na podstawie modelu z wahaniami sezonowymi oraz rozkładu gamma

Źródło: Opracowanie własne.

Analizowane szeregi czasowe charakteryzowały się dość znaczną zmiennością. Współczynnik zmienności oscylował w przedziale 55%-281%. Wartości wybranych kryteriów charakteryzujących oszacowane modele z wahaniami sezonowymi przedstawiono w tabeli 1.

Tabela 1

Wybrane charakterystyki modeli z wahaniami sezonowymi

Kryterium	Vs_e	R^2	*AIC*	*BIC*	*HQC*
Min	0,21	0,3164	43,84	149,26	84,50
Max	2,28	0,9998	3166,70	3342,82	3237,94

Źródło: Opracowanie własne.

Dopasowanie modeli do wartości empirycznych było zróżnicowane, co potwierdzają wartości współczynnika determinacji zawierające się w przedziale 31,64%-99,98%. Dosyć zróżnicowane wartości przyjmował również współczynnik zmienności losowej (21%-228%). Kryteria informacyjne *AIC*, *BIC* i *HQC* przyjmowały wartości dodatnie w zróżnicowanym zakresie.

Dokładność prognoz zbadano za pomocą dwóch wielkości. Pierwszą z nich był względny błąd prognoz wyznaczony jako pierwiastek współczynnika zbieżności Theila:

$$I = \sqrt{\frac{\sum_{T=1}^{r}\left(y_T - y_{pT}\right)^2}{\sum_{T=1}^{r} y_T^2}} \tag{6}$$

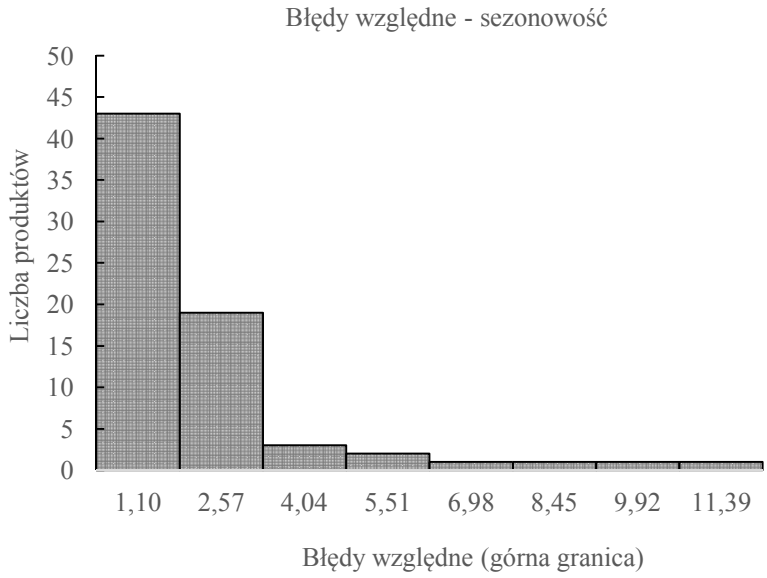
gdzie:
y_T – rzeczywista wielkość sprzedaży,
y_{pT} – prognoza wielkości sprzedaży,
T – okres prognozowany,
r – liczba okresów, na które jest dokonywana prognoza.

Drugą miarą była relacja sumy prognoz do sumy wielkości rzeczywistych. Współczynnik ten liczy się na podstawie wzoru:

$$R = \frac{\sum_{T=1}^{r} y_{pT}}{\sum_{T=1}^{r} y_T} \tag{7}$$

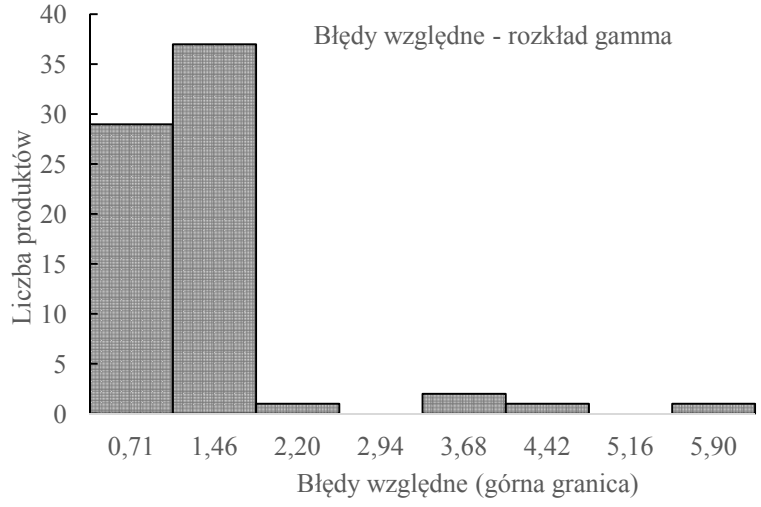
gdzie oznaczenia są takie same, jak we wzorze (6).

W przypadku badanej firmy była to ważniejsza miara dokładności prognoz. Przyczyną tego jest fakt, iż firmie nie zależało specjalnie, aby dokładnie przewidzieć wielkość sprzedaży w każdym tygodniu, tylko żeby przewidzieć zapotrzebowanie na 5 tygodni, gdyż na taki okres było składane zamówienie uzupełniające. Jeżeli chodzi o względny błąd prognoz, to oczywiście jest pożądana jak najmniejsza wartość tego błędu. Jeżeli z kolei chodzi o relacje sumy prognoz do sumy wielkości rzeczywistych, to najbardziej pożądaną wielkością tego błędu jest jedność. Oznaczałoby to, że idealnie przewidzieliśmy wielkość sprzedaży w prognozowanych pięciu tygodniach. Dla wszystkich 72 produktów obliczono oba rodzaje błędów zarówno dla prognoz uzyskanych za pomocą modeli z sezonowością, jak i dla prognoz uzyskanych za pomocą rozkładu gamma. Następnie zbadano rozkłady tych błędów. Przedstawiają je rys. 9-12.



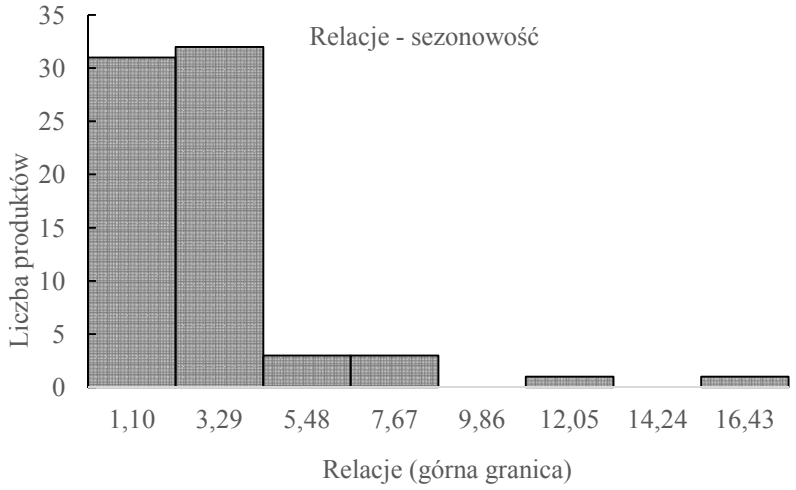
Rys. 9. Rozkłady błędów względnych prognoz uzyskanych dla modeli z sezonowością

Źródło: Opracowanie własne.



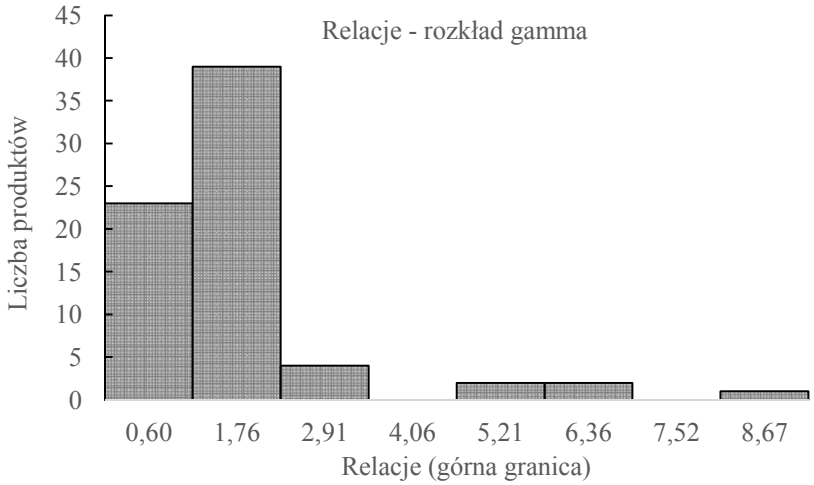
Rys. 10. Rozkłady błędów względnych prognoz uzyskanych dla rozkładu gamma

Źródło: Opracowanie własne.



Rys. 11. Rozkłady błędów prognoz uzyskanych dla modeli z sezonowością

Źródło: Opracowanie własne.



Rys. 12. Rozkłady błędów prognoz uzyskanych dla rozkładu gamma

Źródło: Opracowanie własne.

Jeżeli zbadamy rozkłady błędów prognoz, to widać, że w każdym przypadku posiadały one silną asymetrię prawostronną (dla błędów względnych była to skrajna asymetria prawostronna). Świadczy to o tym, że dominowały mniejsze wielkości błędów. Podstawowe pozycyjne statystyki opisowe struktury błędów prognoz przedstawia tabela 2.

Tabela 2

Podstawowe (pozycyjne) statystyki opisowe struktury błędów prognoz

Miary	Błędy względne – sezonowość	Błędy względne – rozkład gamma	Relacje – sezonowość	Relacje – rozkład gamma
x_{min}	0,367	0,344	0,000	0,028
x_{max}	10,654	5,531	15,333	8,093
Mediana	1,000	0,777	1,352	0,827
Kwartyl pierwszy	0,746	0,585	0,732	0,313
Kwartyl trzeci	1,433	0,940	2,179	1,265
Odchylenie ćwiartkowe	0,344	0,178	0,723	0,476
Współczynnik zmienności	34,35%	22,85%	53,51%	57,58%
Współczynnik asymetrii	0,261	-0,080	0,143	-0,079
Współczynnik spłaszczenia	0,121	0,253	0,178	0,227

Źródło: Opracowanie własne.

Jeżeli zbadamy strukturę względnych błędów prognoz, to widać, że nie jest ona zbyt korzystna. W przypadku modeli z sezonowością minimalny poziom tego błędu wyniósł ponad 36%, czyli prognozując, pomyliliśmy się średnio o ponad 36%. Jeżeli chodzi o maksymalny jego poziom, to wyniósł on aż ponad 1000%. W połowie przypadków prognozując na podstawie modeli z sezonowością pomylono się o nie więcej niż 100%. Rozkład błędów w zawężonym obszarze zmienności posiadał znaczną zmienność oraz wyraźną asymetrię prawostronną, był też wysmukły. Nieco lepiej przedstawiały się błędy względne prognoz uzyskanych za pomocą rozkładu gamma. Minimalny ich poziom wyniósł ponad 34%, a maksymalny – ponad 500%. W połowie przypadków nie przekroczył on 78%. Zmienność rozkładu względnego błędu prognozy była znaczna, natomiast asymetria w zawężonym obszarze zmienności praktycznie nie występowała.

Analizując relacje sumy prognoz do sumy wartości rzeczywistych, można zauważyć, że prognozy były znacznie lepsze, niż wynikałoby to ze względnych błędów prognoz. W przypadku sezonowości w połowie przypadków suma prognoz w stosunku do sumy wielkości rzeczywistych nie była większa niż 35%. Analizując pozostałe kwartale, widać, że w przypadku pierwszych 25% produktów stosunek sumy prognoz do sumy wielkości rzeczywistych nie przekroczył 73%, a ostatnie 25% produktów posiadało tę relację na poziomie wyższym niż 218%. Rozkład relacji posiadał bardzo dużą zmienność (na poziomie przekraczającym 53%), niewielką asymetrię prawostronną oraz był wysmukły. Lepiej prezentowały się relacje sumy prognoz do sumy wielkości rzeczywistych dla prognoz uzyskanych na podstawie rozkładu gamma. Dla pierwszej połowy pro-

duktów relacja ta nie przekroczyła 83%, czyli prognozy były zaniżone. Pierwsze 25% produktów posiadało tę wartość nie większą niż 31%, a ostatnie 25% – powyżej 126,5%. Podobnie jak w przypadku sezonowości, zmienność rozkładu błędów była bardzo duża, asymetria w zawężonym obszarze zmienności praktycznie nie występowała oraz rozkład był wysmukły.

Jak widać z powyższych wyników, nie jest łatwo prognozować wielkość sprzedaży produktów przy takim kształtowaniu się jej w czasie – występowanie bardzo dużych wahań (czyli dużej zmienności), spora liczba okresów, w których nie występowały zamówienia. Widać także, że w takim przypadku prognozowanie na podstawie rozkładu gamma daje lepsze wyniki, niż stosowanie modeli trendu z wahaniami sezonowymi. Zarówno względne błędy prognoz, jak i relacje sumy prognoz do sumy wielkości rzeczywistych były bardziej korzystne dla rozkładu gamma.

Wnioski

W artykule porównano wyniki prognozowania wielkości sprzedaży na podstawie modeli z sezonowością oraz na podstawie rozkładu gamma. W badaniu wykorzystano rzeczywiste dane dotyczące ponad 10 tys. produktów, z których wyselekcjonowano 72. Ich sprzedaż można było prognozować za pomocą obu zaprezentowanych metod. Otrzymane wyniki wskazują, że lepsze prognozy uzyskano za pomocą rozkładu gamma. Oba rodzaje błędów potwierdziły mniejszą rozbieżność otrzymanych prognoz w porównaniu z rzeczywistymi wielkościami sprzedaży. Bardzo duża zmienność zamówień na poszczególne produkty oraz brak jednoznacznych prawidłowości powodują, że trudno jest zbudować modele tendencji rozwojowych z wahaniami sezonowymi o pożądanych własnościach prognostycznych.

Należy również zauważyć, że w badanym przedsiębiorstwie najwięcej produktów było zamawianych rzadziej niż co drugi tydzień, w związku z czym zastosowanie rozkładu gamma mogłoby nie dawać zbyt dobrych wyników (suma prognoz mogłaby być zawyżona, ponieważ otrzymano by je na każdy tydzień, a rzeczywista sprzedaż wystąpiłaby znacznie rzadziej). Dlatego dalszym etapem badań będzie badanie prognoz sprzedaży produktów, których sprzedaż występuje rzadko, maksymalnie co dwa tygodnie.

Bibliografia

Całczyński A. (red.), 2000: Elementy badań operacyjnych w zarządzaniu. Tom 1. Radom.

Czerwiński Z., Guzik B., 1980: Prognozowanie ekonometryczne. PWE, Warszawa.

Dittman P., 2004: Prognozowanie w przedsiębiorstwie. Oficyna Ekonomiczna, Kraków.

Locke C., 1976: A Test for the Composite Hypothesis that Population Has a Gamma Distribution. „Communications in Statistics", 4, s. 351-364.

Sarjusz-Wolski Z., 2000: Sterowanie zapasami w przedsiębiorstwie. PWE, Warszawa.

Shapiro S.S., Chen L., 2001: Composite Test for the Gamma Distribution. „Journal of Quality Technology", Vol. 33, No. 1, s. 47-59.

Tadikamalla P.R., 1984: A Comparison of Several Approximations to the Lead Time Demand Distribution. „International Journal of Management Science", Vol. 12, No. 6, s. 575-581.

Zeliaś A., 1997: Teoria prognozy. PWE, Warszawa.

COMPARING THE EFFECTIVENESS OF EX POST FORECASTS OF SALES IN A COMPANY OBTAINED BY MEANS OF GAMMA DISTRIBUTION AND MODELS WITH SEASONAL ADJUSTMENTS

Summary

Main aim of the article was comparison of sales forecasts effectiveness in a company. In the first stage time series representing sales were selected with respect to presence of seasonality and consistency with gamma distribution. Two types of statistical tests were used: *F* test (seasonality) and nonparametric Locke tests (compatibility with gamma distribution). In the next step *ex post* forecast were computed by means of linear trend model with seasonality and gamma distribution (on the level of median of theoretical distribution). Effectiveness of forecasts was compared on the basis of chosen coefficients of forecasts errors. Forecasts obtained be means of gamma distribution turned out to be better in all considered aspects.

Urszula Grzybowska
Marek Karwański
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

PRZYKŁADY ZASTOSOWANIA MACIERZY MIGRACJI W ZARZĄDZANIU RYZYKIEM FINANSOWYM

Wstęp

W ostatnich latach coraz częściej do opisu kondycji finansowej podmiotów gospodarczych stosuje się podejście oparte na ratingach. Podmioty gospodarcze, np. banki, państwa, kredytobiorcy, są grupowane w skończoną liczbę klas obiektów o podobnej kondycji finansowej opisywanych za pomocą skal ratingowych. Ratingi wyznaczają dyskretną przestrzeń stanów. Prawdopodobieństwa przejścia elementów między klasami są opisywane przez tzw. macierze migracji. Macierze migracji stosowane w ryzyku kredytowym są macierzami przejścia dyskretnych łańcuchów Markowa z jednym stanem pochłaniającym, stanem niewypłacalności („default"). Są to macierze diagonalnie dominujące, a ich największa wartość własna wynosi 1. Macierze migracji stosowane w ubezpieczeniach komunikacyjnych lub badaniu profilu użytkowników kart kredytowych są macierzami regularnymi [Iosifescu 1987, s. 96].

W niniejszym artykule przedstawiono przykłady zastosowania macierzy migracji w szacowaniu ryzyka kredytowego, w ubezpieczeniach Bonus-Malus oraz w badaniu zmian zachowań użytkowników kart kredytowych.

1. Zastosowanie macierzy migracji w szacowaniu ryzyka kredytowego

Nowe Umowy Kapitałowe (Basel II/III) zaostrzyły zasady zarządzania ryzykiem i wprowadziły możliwość wykorzystywania przez banki systemów IRB,

czyli wewnętrznych systemów ratingowych, do celów wyznaczania zabezpieczenia kapitałowego oraz do wyznaczania wewnętrznych miar ryzyka. Podejście IRB wymaga szacowania indywidualnych PD (prawdopodobieństw migracji do stanu niewypłacalności) dla każdego kredytobiorcy. W ramach tego podejścia są wykorzystywane macierze migracji szacowane na podstawie danych historycznych, którymi dysponuje bank.

W artykule oparto się na ratingu obejmującym 8 stanów [Saunders 2001, s. 21], przy czym AAA oznacza najwyższą klasę ratingową, a D stan niewypłacalności (default). Prawdopodobieństwo p_{ij} przejścia z dowolnego stanu *i* do stanu *j* w okresie od *t-1* do *t* jest liczone jako stosunek n_{ij} liczby elementów, które przeszły ze stanu *i* do stanu *j*, do ilości wszystkich elementów, które wyszły ze stanu *i*, tj. $\sum_j n_{ij} = N_i$:

$$p_{ij} = \frac{n_{ij}}{N_i} \tag{1}$$

Prawdopodobieństwa przejścia wyznaczane według wzoru (1) są obliczane na podstawie ratingów dla dwóch wybranych dat. Nie są brane pod uwagę zmiany stanów wewnątrz okresu. T. Schuermann wprowadził następujący estymator prawdopodobieństw przejścia z uwzględnieniem T okresów:

$$p_{ij} = \frac{N_{ij}}{N_i} = \frac{\sum_{t=1}^{T} n_{ij}(t)}{\sum_{t=1}^{T} N_i(t)} \tag{2}$$

gdzie $N_i(t) = \sum_j n_{ij}(t)$ oznacza sumaryczną liczbę migracji ze stanu *i* w okresie *t-1* do *t*, zaś $n_{ij}(t)$ oznacza liczbę elementów, które przeszły ze stanu *i* do stanu *j* w okresie *t-1* do *t,* dla $t \le T, \quad t \in N$.

Macierze migracji wykorzystywane w ryzyku kredytowym są macierzami łańcuchów pochłaniających. Stąd z czasem prawdopodobieństwo migracji do stanu D nawet z najwyższej klasy wynosi 1. Porównanie macierzy migracji wyznaczonych na podstawie tych samych danych kilkoma metodami [Grzybowska, Karwański 2011] pokazało, że chociaż macierze te nie różniły się bardzo pod względem różnych miar matematycznych, różniły się znacznie pod względem prędkości zbieżności do stanu D (od 5 do 12 okresów dla migracji ze stanu CCC). Uzyskane wyniki świadczą o dużej niestabilności metod szacowania ma-

cierzy. Istotnym problemem dotyczącym macierzy migracji jest zależność od stanu gospodarki. W niniejszym artykule, opierając się na rzeczywistych macierzach migracji oszacowanych dla różnych stanów gospodarki USA (recesja i ekspansja), pokazujemy, jak z czasem zmienia się prędkość zbieżności do stanu niewypłacalności dla kolejnych klas ratingowych oraz jak macierze migracji wpływają na szacowanie wartości zagrożonej.

1.1. Opis danych

Do analizy wykorzystano rzeczywiste kwartalne macierze migracji [Bangia, Diebold, Schuermann 2000, s. 28], które przeliczono na macierze prawdopodobieństw przejścia w jednym roku. Przez E oznaczamy jednoroczną macierz migracji dla stanu ekspansji gospodarczej (tabela 1), zaś przez R macierz dla stanu recesji (tabela 2).

Tabela 1

Roczna macierz migracji dla okresu ekspansji (macierz E)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	0,9297	0,0628	0,0057	0,001	0,0008	0,0001	0	0
AA	0,0057	0,9257	0,0609	0,0056	0,0006	0,0012	0,0003	0,0001
A	0,0008	0,0201	0,9265	0,0451	0,0049	0,0025	0,0001	0,0001
BBB	0,0004	0,0031	0,0548	0,8855	0,0448	0,0096	0,0009	0,0011
BB	0,0004	0,0013	0,0086	0,0689	0,8287	0,0793	0,0064	0,0064
B	0	0,0008	0,003	0,0056	0,0599	0,8511	0,0406	0,039
CCC	0,0016	0,0002	0,0065	0,0081	0,0177	0,1108	0,5837	0,2716
D	0	0	0	0	0	0	0	1

Tabela 2

Roczna macierz migracji dla recesji gospodarczej (macierz R)

	AAA	AA	A	BBB	BB	B	CCC	D
AAA	0,9222	0,0653	0,012	0,0005	0,0001	0	0	0
AA	0,0067	0,8828	0,1009	0,0059	0,0034	0,0001	0	0
A	0,0009	0,0319	0,868	0,093	0,0058	0,0002	0	0,0002
BBB	0,0015	0,0021	0,04	0,8658	0,0818	0,0035	0,0006	0,0047
BB	0	0,0023	0,0031	0,049	0,817	0,0936	0,0155	0,0194
B	0	0,0022	0,0023	0,0045	0,0249	0,8173	0,0672	0,0816
CCC	0	0	0	0,0001	0,0004	0,0355	0,5382	0,4258
D	0	0	0	0	0	0	0	1

Porównując wyrazy macierzy E (tabela 1) i R (tabela 2), zauważamy, że w okresie recesji znacznie wyższe są prawdopodobieństwa migracji do stanu niewypłacalności, z wyjątkiem dwóch najwyższych klas ratingowych. Wyrazy na przekątnej macierzy migracji dla okresu recesji (macierz R), czyli prawdopodobieństwa, że kredytobiorcy nie zmienią klasy w ciągu kolejnego roku, są niższe niż odpowiednie wyrazy macierzy dla okresu ekspansji (macierz E). Dla okresu recesji nie zaobserwowano zajścia skrajnych wydarzeń, czyli migracji z wysokich klas ratingowych do niskich i odwrotnie, na co wskazuje wyraźna koncentracja dodatnich wartości prawdopodobieństw wokół przekątnej. Oznacza ona, że prawdopodobna jest migracja tylko do najbliższych stanów.

1.2. Zastosowanie macierzy migracji do wyznaczania PD

Na podstawie prawdopodobieństw PD wylicza się ryzyko kredytowe. Miary ryzyka odnoszą się do określonego horyzontu czasowego. W praktycznych zastosowaniach różnego rodzaju wskaźniki opierające się na prawdopodobieństwie zmiany ratingu klienta są liczone dla okresu jednego roku. Tym samym wartość ryzyka jest wykorzystywana do liczenia miar biznesowych, takich jak np. dochodowość, gdzie trzeba uwzględnić prognozy prawdopodobieństwa zmian ratingu w ciągu całego okresu „życia" klienta. Stąd niezbędne są modele pozwalające szacować dynamikę zmian macierzy migracji. W ramach teorii łańcuchów Markowa macierze przejścia stanowią grupę; zmiany macierzy przejścia w danym okresie wylicza się mnożąc macierze przejść dla okresów pośrednich. W szczególności wyrazy n-tej potęgi jednorocznej macierzy migracji oznaczają prawdopodobieństwa przejścia w n-tym roku.

Do celów prognostycznych wyznaczono kolejne potęgi macierzy jednorocznych. Obliczono 5 kolejnych potęg macierzy R i E. Istotne z punktu widzenia szacowania ryzyka wartości zostały przedstawione w tabelach 3 i 4. Otrzymane dla obu stanów gospodarki macierze przejścia w n krokach, $n = 1, 2, 3, 4, 5$, mają malejące wyrazy na przekątnych, przy czym kolejne potęgi macierzy R mają wartości niższe niż odpowiednie potęgi macierzy E. Jednocześnie stosunkowo szybko rosną wartości w ostatnich kolumnach potęg macierzy, które są prawdopodobieństwami migracji do stanu niewypłacalności w kolejnych latach. Wzrost wartości jest szczególnie szybki dla niższych klas ratingowych w okresie recesji (tabela 4). W szczególności prawdopodobieństwo migracji do stanu D w ciągu 5 lat dla kredytobiorcy przypisanego do klasy CCC wynosi aż 90%. Ze względu na wysokie prawdopodobieństwa

migracji do stanu D w stosunkowo krótkim okresie niższe klasy ratingowe określa się mianem spekulacyjnych lub nieinwestycyjnych.

Wykorzystując macierz fundamentalną pochłaniającego łańcucha Markowa [Iosifescu 1987, s. 96] dla macierzy przejścia E i R, możemy wyznaczyć średnią liczbę lat, która upłynie do momentu znalezienia się kredytobiorcy w stanie niewypłacalności. Wartości z tabeli 5 świadczą o tym, że w obu przypadkach okres ten jest bardzo długi dla wyższych klas ratingowych. Jednocześnie w czasie recesji gospodarczej średnia liczba lat przed przejściem do stanu niewypłacalności wynosi zaledwie 3 lata dla kredytobiorcy przypisanego do klasy CCC i 12 dla kredytobiorcy przypisanego do klasy B.

Tabela 3

Wyrazy kolejnych potęg jednorocznej macierzy migracji wyznaczonej dla stanu ekspansji gospodarczej

	Wyrazy na przekątnej					Wyrazy z ostatniej kolumny				
klasa	n = 1	n = 2	n = 3	n = 4	n = 5	n = 1	n = 2	n = 3	n = 4	n = 5
AAA	0,9297	0,8647	0,8046	0,749	0,6975	0	0	0,0001	0,0002	0,0003
AA	0,9257	0,8585	0,7977	0,7426	0,6927	0,0001	0,0003	0,0007	0,0012	0,0018
A	0,9265	0,8622	0,8057	0,756	0,7122	0,0001	0,0004	0,0009	0,0018	0,0029
BBB	0,8855	0,7897	0,7093	0,6416	0,5842	0,0011	0,003	0,0057	0,0093	0,0136
BB	0,8287	0,6947	0,5894	0,5062	0,4399	0,0064	0,0166	0,0298	0,0451	0,0619
B	0,8511	0,7337	0,6391	0,5617	0,4974	0,039	0,0836	0,1288	0,1722	0,2126
CCC	0,5837	0,3453	0,2083	0,1291	0,0829	0,2716	0,4346	0,5348	0,5986	0,641
D	1	1	1	1	1	1	1	1	1	1

Tabela 4

Wyrazy kolejnych potęg jednorocznej macierzy migracji wyznaczonej dla stanu recesji gospodarczej

	Wyrazy na przekątnej					Wyrazy z ostatniej kolumny				
Klasa	n = 1	n = 2	n = 3	n = 4	n = 5	n = 1	n = 2	n = 3	n = 4	n = 5
AAA	0,9222	0,8509	0,7855	0,7255	0,6705	0	0	0	0,0001	0,0003
AA	0,8828	0,783	0,6977	0,6246	0,5617	0	0,0001	0,0005	0,0011	0,0021
A	0,868	0,7604	0,6722	0,5996	0,5395	0,0002	0,0009	0,0024	0,0046	0,0079
BBB	0,8658	0,7574	0,6691	0,5966	0,5366	0,0047	0,0109	0,0193	0,03	0,0433
BB	0,817	0,6739	0,5612	0,4721	0,401	0,0194	0,0497	0,0874	0,1294	0,1735
B	0,8173	0,6727	0,5569	0,4634	0,3873	0,0816	0,1774	0,2721	0,3591	0,4361
CCC	0,5382	0,2921	0,1604	0,0897	0,0515	0,4258	0,6579	0,7862	0,8586	0,9007
D	1	1	1	1	1	1	1	1	1	1

Tabela 5

Średnia liczba lat przed migracją do klasy D z kolejnych klas ratingowych

	AAA	AA	A	BBB	BB	B	CCC
Recesja	71	60	51	40	24	12	3
Ekspansja	162	150	138	120	91	59	27

Uzyskane wyniki wydają się wskazywać na horyzonty czasowe przekraczające granice stosowane w praktyce. Jednakże bank powinien prognozować ryzyko na okres czasowy, który pozwoli na ewentualną zmianę struktury portfela. Dla produktów tzw. portfela bankowego (np. tradycyjnych produktów, takich jak kredyty) oznacza to okres kilku, a nawet kilkunastoletni. Niektóre produkty wymagają modeli szacujących ryzyko na okres kilkudziesięciu lat.

1.3. Zastosowanie macierzy migracji do obliczania wartości zagrożonej

Pomiar ryzyka kredytowego jest oparty na rozkładzie zysków i strat, zwanym dalej dla uproszczenia rozkładem strat, związanych z handlem produktami finansowymi.

Z biznesowego punktu widzenia ryzyko kredytowe oznacza koszt finansowania kredytów i składa się z dwóch składowych. Pierwsza to koszt związany bezpośrednio ze stratami wynikłymi z niewywiązywania się klientów ze swoich zobowiązań, tj. przejścia do stanu niewypłacalności, i jest mierzona wartością oczekiwaną strat. Wartość oczekiwana oznacza rzeczywistą stratę w długim okresie czasowym i musi być traktowana jako bezpośredni koszt. Druga składowa jest związana z koniecznością utrzymywania rezerw gwarantujących „bezpieczne prowadzenie działalności", czyli tzw. kapitału ekonomicznego. Kapitał ten można traktować jako źródło finansowania w przypadku wystąpienia fluktuacji np. strat katastroficznych, które w długim okresie wchodzą do średniej, ale w krótkim okresie muszą być dodatkowo pokryte z odpowiednich rezerw. Składnik ten w ramach metodologii zwanej „wartością zagrożoną" jest mierzony na podstawie rozkładu strat.

Wartość oczekiwaną strat najczęściej wyznacza się za pomocą modeli scoringowych, czyli modeli regresyjnych pozwalających wyznaczyć średnią warunkową, gdzie czynnikami kontrolnymi są odpowiednie atrybuty klienta. Pomiar części drugiej, w ramach metodologii wartości zagrożonej, polega na wyznaczeniu odpowiedniego percentyla straty. Z uwagi na krytyczne znaczenie strat nieoczekiwanych (katastroficznych) nadzór bankowy ustanowił ścisłą wartość dla tego percentyla, a mianowicie 99%.

Istnieje kilka modeli wyliczania percentyla rozkładu strat. W niniejszym artykule skupimy uwagę na modelach wyznaczania wartości zagrożonej opartych na macierzy migracji [Gupton, Finger, Bhatia 1997]. Jednym z popularnych modeli należących do tej klasy jest tzw. CreditMetrics Model.

Teoretyczną podstawę tego modelu stanowi teoria wyceny wartości firmy zaproponowana przez Mertona [Hull, Nelken, White 2004]. Według Mertona zdolność kredytową przedsiębiorstwa można wyznaczyć porównując wartość jego aktywów i pasywów. Pasywa są w modelu Mertona wartością stałą, natomiast aktywa można przedstawić w postaci procesu stochastycznego zależnego od czasu t. Na podstawie prostych analogii Merton pokazał, że proces ten może mieć postać geometrycznego ruchu Browna. Można więc zapisać wartość aktywów A_t w postaci równania:

$$A_t = A_0 exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma\sqrt{t}Z_t\right\} \tag{3}$$

gdzie A_0 jest stałą skalującą, zaś $Z_t \sim N(\mu,\sigma^2)$ zmienną losową (o rozkładzie normalnym).

W modelu Mertona stopa zwrotu na aktywach $\frac{dA_t}{A_t}$ ma rozkład normalny. Niewypłacalność następuje wówczas, gdy wartość aktywów A_t spada poniżej pewnej granicy związanej z wartością pasywów.

Model CreditMetrics rozszerza podejście Mertona, wprowadzając zmienność jakości aktywów w sposób zgodny z systemem ratingowym. Polega to na wprowadzeniu zakresu wartości (tzw. klas) do rozkładu aktywów w taki sposób, aby prawdopodobieństwa przejść między klasami replikowały prawdopodobieństwa migracji między klasami ratingowymi. Innymi słowami sprowadza się to do wyznaczenia parametrów rozkładu aktywów A_t: μ i σ^2 na bazie macierzy migracji [Crouhy, Galai, Mark 2001, s. 315-355]. W nomenklaturze CreditMetrics aktywa są nazywane portfelem.

Produkty bankowe, które mogą wchodzić w skład portfela, to przede wszystkim tzw. transakcje bazowe, w których przepływy finansowe odbywają się w określonych momentach czasowych. Inne typy transakcji są mapowane na transakcje bazowe z użyciem metody tzw. replikacji. Wycena wartości produktu zależy w modelu zarówno od klasy ratingowej, jak i od struktury terminowej płatności poprzez wprowadzenie współczynników dyskonta pozwalających przeliczać wartości przepływów pieniężnych w różnych punktach czasowych. Jest to szczególnie istotne w sytuacji, gdy chcemy pokazać różnice ryzyka portfela w różnych stanach gospodarki.

W poniższym przykładzie założono, że bank ma portfel X składający się z obligacji opisanych w tabeli 6.

Tabela 6

Atrybuty obligacji wchodzących w skład portfela X

ID	Klasa ratingowa	Kupon (wartość przepływu pieniężnego)	Zapadalność (w latach)	Wartość nominalna PLN	Wartość na koniec roku PLN
Obligacja MMM	BBB	3.00%	4	4 000 000	3 696 142.1
Obligacja XTX	A	1.00%	2	2 000 000	1 943 190.7
Obligacja YYY	CCC	10.00%	3	1 000 000	1 010 581.7

Wartość portfela X jest sumą wartości poszczególnych składowych (obligacji). Musimy uwzględnić zarówno prawdopodobieństwo przepływów finansowych związanych z „kuponami" odzwierciedlone w macierzy migracji pomiędzy klasami ratingowymi, jak i fakt, że przepływy finansowe następują w różnych punktach czasowych. Za krzywą dyskontową przyjmijmy funkcję stałą $Dyskonto_t = 0{,}06$ – taka struktura stóp procentowych pozwala na wyrażenie wyniku w kategorii testów warunków skrajnych, czyli zmian ryzyka przy zmianie wartości tej funkcji.

Aby obliczyć ryzyko, czyli średnią stratę oraz 99% percentyl strat (wartość zagrożoną) w horyzoncie czasowym 1 roku, trzeba wyznaczyć rozkład wartości poszczególnych obligacji w tym okresie. Przy założeniu, że roczne stopy zwrotu obligacji mają rozkład normalny (co wynika z modelu Mertona), można oszacować parametry rozkładu portfela na podstawie wartości średnich i macierzy migracji. Wartość średnia strat jest sumą wartości średnich strat poszczególnych obligacji, natomiast wartość zagrożoną portfela liczy się na podstawie macierzy migracji: metodą symulacyjną Monte Carlo lub metodą uproszczoną za pomocą aproksymacji rozkładem normalnym.

W poniższych obliczeniach porównano wartości ryzyka dla dwóch stanów ekonomii: cyklu recesji (R) i cyklu ekspansji (E), wyznaczone z zastosowaniem obu metod.

W tabeli 7 przedstawiono obliczenia dla portfela X opisanego w tabeli 6. Wartość zagrożona dla recesji jest wyraźnie wyższa niż dla stanu ekspansji gospodarczej.

W przypadku rzeczywistych portfeli okazuje się, że stany gospodarki wpływają nie tylko na prawdopodobieństwa migracji pomiędzy różnymi klasami, ale również na współczynniki dyskontujące. Wartość tych współczynników spada w trakcie recesji. Ma na to wpływ zarówno polityka banku centralnego (decyzje polityki pieniężnej), jak i ogólne zapotrzebowanie na środki finansowe.

Tabela 7

Porównanie ryzyka (wartości zagrożonych) dla różnych stanów gospodarki – cykl ekspansji i recesji

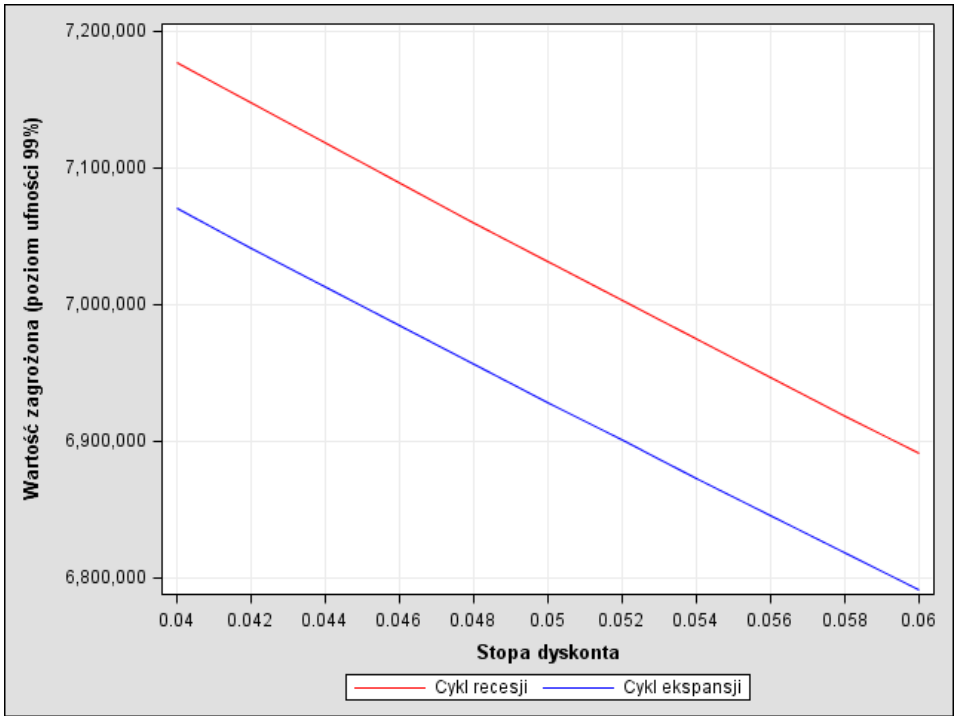
Parametry ryzyka ***Cykl ekspansji***	**Wartość**
Wartość średnia	6 504 607,13
Odchylenie stand.	349 112,91
Value at Risk (99% perc.)	6 790 857,02
Value at Risk (Normal approximation)	7 316 765,20
Parametry ryzyka ***Cykl recesji***	**Wartość**
Wartość średnia	6 146 635,78
Odchylenie stand.	630 656,28
Value at Risk (99% perc.)	6 871 216,25
Value at Risk (Normal approximation)	7 613 761,67

Zmiany makroekonomiczne wpływają na zmiany prawdopodobieństw migracji pomiędzy klasami ratingowymi, dlatego do modelowania ryzyka w okresach „ekspansji" i „recesji" trzeba użyć innych macierzy. Z drugiej strony zmiany gospodarcze wpływają na stopy dyskontowe. Oba efekty wzmacniają się, wpływając na wartość portfela.

Na rysunku 1 przedstawiono wykresy zmian wartości portfela X spowodowane zmianami stóp dyskontowych w cyklu recesji i ekspansji [Bluhm, Overbeck, Wagner 2003]. Linia czerwona i niebieska odpowiadają sytuacji opisywanej przez różne macierze migracji z uwzględnieniem zmian stóp dyskontowych. Zmiany stóp użyte w poniższych symulacjach mają postać tzw. przesunięć równoległych, tzn.:

$$Dyskonto_t^{nowe} = Dyskonto_t^{stare} + \Delta \tag{4}$$

Warto zwrócić uwagę, że w naszym modelu różnica wartości ryzyka w różnych cyklach gospodarczych wynikająca ze zmian macierzy migracji jest porównywalna z przesunięciem stóp dyskontowych o 1%. Jednocześnie daje to odpowiedź na pytanie, w jaki sposób różne macierze migracji mogą wpływać na wartość portfela w sytuacji zmian w otoczeniu ekonomicznym banku.



Rys. 1. Zmiany poziomu ryzyka (wartości zagrożonej) w funkcji zmian stóp dyskonta w różnych stanach cyklu gospodarczego

2. Zastosowanie macierzy migracji w systemach komunikacyjnych typu Bonus-Malus

System Bonus-Malus jest systemem ratingowym stosowanym w ubezpieczeniach komunikacyjnych do korekty ceny produktów komunikacyjnych. W tym systemie przejścia między klasami zależą od liczby szkód, jakie poniósł ubezpieczony w danym okresie składkowym (roku) [Denuit, Maréchal, Pitrebois 2007, s. 165-178; Lemaire 1998]. Systemy Bonus-Malus były omawiane w publikacjach polskojęzycznych [Miszczyńska, Miszczyński 2006, s. 387-398; Podgórska 2008, s. 206-231]. Macierze przejścia systemów Bonus-Malus są macierzami regularnymi. Z czasem ich potęgi stabilizują się wokół pewnego poziomu równowagi, odpowiadającego oczekiwanej liczbie szkód w ciągu roku.

Do analizy wykorzystano dane jednego z dużych amerykańskich towarzystw ubezpieczeniowych dotyczące 1500 ubezpieczonych w 4 kolejnych latach obejmujących okres 2000-2003. Dane pochodziły z systemu B-M z 13 kla-

sami typu +1/-3, tzn. każdy rok bezszkodowej jazdy powodował przesunięcie ubezpieczonego o jedną klasę w górę, zaś zgłoszenie co najmniej jednej szkody przesuwało ubezpieczonego o 3 klasy w dół. Ze względu na brak informacji o szacowaniu prawdopodobieństwa wystąpienia szkody wyznaczono macierze migracji, wykorzystując wzór (2) (tabela 8). Następnie wyznaczono wektor stacjonarny dla oszacowanej macierzy przejścia jako lewy unormowany wektor własny odpowiadający jej największej wartości własnej (tabela 9).

Tabela 8

Macierz migracji sytemu B-M

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0,2	0,8	0	0	0	0	0	0	0	0	0	0	0
2	0,163	0	0,837	0	0	0	0	0	0	0	0	0	0
3	0,075	0	0	0,925	0	0	0	0	0	0	0	0	0
4	0,124	0	0	0	0,876	0	0	0	0	0	0	0	0
5	0	0,084	0	0	0	0,916	0	0	0	0	0	0	0
6	0	0	0,062	0	0	0	0,938	0	0	0	0	0	0
7	0	0	0	0,051	0	0	0	0,949	0	0	0	0	0
8	0	0	0	0	0,06	0	0	0	0,94	0	0	0	0
9	0	0	0	0	0	0,055	0	0	0	0,945	0	0	0
10	0	0	0	0	0	0	0,063	0	0	0	0,937	0	0
11	0	0	0	0	0	0	0	0,068	0	0	0	0,932	0
12	0	0	0	0	0	0	0	0	0,059	0	0	0	0,941
13	0	0	0	0	0	0	0	0	0	0,053	0	0	0,947

Tabela 9

Wektor stacjonarny

0,01	0,02	0,02	0,04	0,09	0,13	0,46	0,78	1,00	5,32	4,99	4,65	82,49

Wyrazy wektora stacjonarnego opisują długookresowe zachowanie ubezpieczonych. Otrzymane wartości sugerują, że z czasem, niezależnie od klasy początkowej, około 82,5% ubezpieczonych znajdzie się w ostatniej, najwyższej klasie zniżkowej. Tendencję do koncentracji ubezpieczonych w najwyższych klasach zniżkowych wykazuje wiele systemów B-M. Wektor stacjonarny jest wykorzystywany do wyznaczania miar efektywności systemu B-M, porównywania różnych systemów B-M oraz obliczania tzw. średniego stacjonarnego poziomu składki.

3. Zastosowanie macierzy migracji w badaniu profili użytkowników kart kredytowych

Posiadacze kart kredytowych, którzy przynoszą zyski bankowi, np. spłacając z opóźnieniem swoje zobowiązania, są dzieleni na trzy grupy, w zależności od sposobu wywiązywania się ze swoich zobowiązań:

1. Transaktorzy (T).
2. Rewolwery (R).
3. Klienci ekonomiczni (convenience user, swingerzy (S) [Berry, Linoff 2011, s. 580].

Do grupy określanej mianem „Transaktorzy" należą osoby, które mają duże saldo na karcie, ale regularnie spłacają swoje zobowiązania. Mianem „Rewolwerów" określa się posiadaczy kart kredytowych, którzy mają duże saldo na karcie, ale nie spłacają zadłużenia regularnie. „Klienci ekonomiczni" charakteryzują się rzadkimi, ale dużymi transakcjami, które spłacają przez dłuższy czas, ponosząc koszty odsetek.

Tabela 10

Macierz A migracji profili użytkowników kart kredytowych

	R	S	T
R	0,8972	0,1028	0
S	0,3289	0,3896	0,2815
T	0	0,0720	0,9280

Dane dotyczyły 2100 posiadaczy kart kredytowych, których przynależność do klas była weryfikowana co 3 miesiące (okresy 3-miesięczne). Do analizy wykorzystano 12 okresów, przy czym pierwszy okres był dla wszystkich początkowym okresem używania karty. Na podstawie danych wyznaczono macierz migracji A według wzoru (2) (tabela 10). Ponieważ macierz przejścia użytkowników kart kredytowych jest macierzą regularną, zachowanie długookresowe można badać wykorzystując rozkład stacjonarny.

Tabela 11

Unormowany lewy wektor własny odpowiadający największej wartości własnej macierzy przejścia A

R	0,3945
S	0,1233
T	0,4821

Otrzymany wektor własny macierzy A (tabela 11) sugeruje, że z czasem, niezależnie od początkowej klasyfikacji, około 40% użytkowników kart będzie „Rewolwe-

rami", 48% „Transaktorami" i tylko 12% „Klientami ekonomicznymi". Zbieżność kolejnych potęg macierzy migracji do rozkładu stacjonarnego jest bardzo wolna.

Podsumowanie

Macierze migracji są powszechnie stosowane w zarządzaniu ryzykiem kredytowym. W artykule przedstawiono ich wykorzystanie do wyznaczania prawdopodobieństw przejścia do stanu niewypłacalności w zadanym horyzoncie czasowym dla różnych stanów gospodarki. Pokazano także, jak macierze migracji wyznaczane dla różnych stanów gospodarki wpływają na szacowanie wartości zagrożonej (ryzyka). Wartość zagrożona jest większa w okresie recesji. Na zwiększenie ryzyka ma wpływ zarówno postać macierzy migracji, jak i zmiana stóp dyskontowych.

Macierze migracji są także stosowane w systemach ubezpieczeń komunikacyjnych typu Bonus-Malus oraz w ubezpieczeniach zdrowotnych. Przyjmuje się, że migracje podmiotów są opisywane przez macierze przejścia jednorodnych łańcuchów Markowa pierwszego rzędu. Nie wszystkie dane rzeczywiste, z którymi mieliśmy okazję się zetknąć, potwierdzają tę zasadę. W szczególności, w kontekście macierzy migracji stosowanych w ryzyku kredytowym, są spotykane różne odstępstwa od przyjętego standardu. Spora część kredytobiorców, którzy stali się niewypłacalni, po spełnieniu pewnych zobowiązań powraca do bazy danych z wysokim ratingiem. Oznacza to, że stan D nie jest pochłaniający. Wyznaczone przez nas macierze na podstawie danych rzeczywistych nie były diagonalnie dominujące. W praktyce jest widoczna wyraźna tendencja opisywana w literaturze jako „rating drift" [Bangia, Diebold, Schuermann 2000, s. 21-25]. Oznacza to, że istnieje potrzeba poszukiwania innych modeli macierzy migracji, które nie są oparte na teorii jednorodnych łańcuchów Markowa pierwszego rzędu.

Bibliografia

Basel II: Revised International Capital Framework. http://www.bis.org/publ/bcbsca.htm.

Basel III: http://www.bis.org/ bcbs/basel3.htm.

Bangia A., Diebold F., Schuermann T., 2000: Rating Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing. The Wharton Financial Institution, http://www.ssc.upenn.edu/~diebold/index.html.

Berry M.J.A., Linoff G.S., 2011: Data Mining Techniques for Marketing, Sales and Consumer Relationship Management. John Wiley & Sons.

Bluhm C., Overbeck L., Wagner C., 2003: An Introduction to Credit Risk Modeling. Chapman and Hall, London.

Crouhy M., Galai D., Mark R., 2001: Risk Management. McGraw-Hill, New York.

Denuit M., Maréchal X., Pitrebois S., 2007: Actuarial Modelling of Claim Counts Risk Classification. Credibility and Bonus-Malus. John Wiley & Sons.

Grzybowska U., Karwański M., 2011: Wykorzystanie miar matematycznych i biznesowych do porównywania modeli macierzy migracji stosowanych w analizie ryzyka kredytowego. „Metody ilościowe w badaniach ekonomicznych", t. XII, nr 2, Wydawnictwo SGGW, Warszawa.

Gupton G., Finger C., Bhatia M., 1997: CreditMetrics – Technical Document. J.P. Morgan.

Hull J., Nelken I., White A., 2004: Merton's Model, Credit Risk, and Volatility Skews. University of Toronto.

Iosifescu M., 1987: Skończone łańcuchy Markowa. WNT, Warszawa.

Lemaire J., 1998: Bonus-Malus Systems: The European and Asian Approach to Merit-Rating. „North American Actuarial Journal", Vol. 2, No. 1.

Miszczyńska D., Miszczyński M., 2006: Systemy ubezpieczeń Bonus-malus. Podejście symulacyjne. W: Modelowanie preferencji a ryzyko. T. Trzaskalik (red.). AE, Katowice, s. 387-398.

Podgórska M. (red.), 2008: System Bonus-Malus sprawiedliwy w sensie przejść między klasami. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, nr 1 (1201), s. 206-231.

Saunders A., 2001: Metody pomiaru ryzyka kredytowego. Oficyna Ekonomiczna, Kraków.

Schuermann T.: Credit Migration Matrices. W: Encyclopedia Quantitative Risk Analysis & Assessment. http://www.wiley.com//legacy/wileychi/risk/.

MIGRATION MATRICES AND THEIR APPLICATION IN RISK MANAGEMENT

Summary

Migration matrices are widely used in various areas of risk management. In credit risk management an obligor is assigned to one of several rating classes and his future rating is determined by a transition matrix of a Markov chain. The probability that an obligor will change his credibility can be read off a migration matrix. The most important aspect of credit risk management is estimation of the probability that the obligor will not be able to meet his financial commitments, i.e., estimation of the probability of his default. The other field were migration matrices are used are insurance Bonus-Malus systems. The aim of the paper is to present applications of migration matrices in credit risk management, automobile insurance and in investigation of credit card users profiles.

Krzysztof Targiel
Uniwersytet Ekonomiczny w Katowicach

MODELOWANIE ZMIAN ZMIENNYCH STANU W MODELU DWUMIANOWYM DO CELÓW WYCENY OPCJI REALNYCH*

Wprowadzenie

Najpowszechniej stosowanym podejściem w wycenie opcji realnych jest wykorzystanie drzewa dwumianowego metodą zaproponowaną przez Coxa, Rossa i Rubinsteina [1979]. Przyjmuje się, że wartość opcji zależy od pewnej wielkości ekonomicznej nazywanej zmienną stanu. Dodatkowa wartość sytuacji opcyjnej wynika z faktu, iż zmienna stanu porusza się w pewnym procesie stochastycznym, mogąc dzięki temu osiągnąć korzystne wartości. Metoda drzew dwumianowych polega na pokryciu tego procesu drzewem, w którym za każdym razem wartość zmiennej stanu może wzrosnąć lub spaść. Utworzony graf wartości jest drzewem, w węzłach którego znajdują się rozważane wartości zmiennej stanu, natomiast łuki pokazują możliwe przejścia pomiędzy tymi wartościami. Bardzo istotne jest właściwe dopasowanie drzewa dwumianowego. Jest to jeden pierwszych etapów wyceny opcji realnych. Jest dokonywany na podstawie historycznych danych procesu stochastycznego. Niniejszy artykuł jest poświęcony problemowi doboru parametrów drzewa dwumianowego.

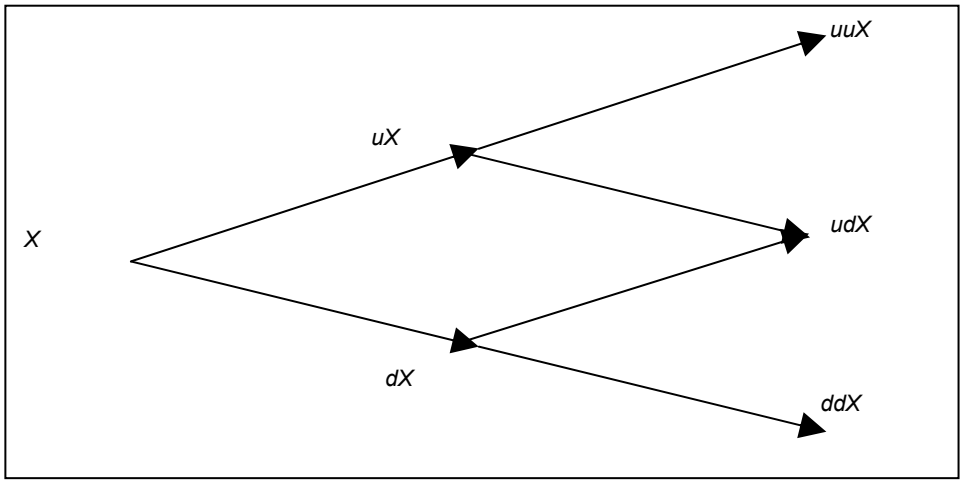
Rozdział pierwszy przedstawia istotę problemu. Na podstawie kilku przykładów pokazujących niewłaściwie dobrane drzewa jest zobrazowana waga problemu. Kolejny rozdział przedstawia rozważane w literaturze przedmiotu modele procesów stochastycznych. Wybór odpowiedniego procesu stochastycznego ma decydujące znaczenie dla sposobu obliczania węzłów drzewa dwumianowego, co jest przedmiotem kolejnego rozdziału. Całość kończy przykład oblicze-

* Projekt został sfinansowany ze środków Narodowego Centrum Nauki jako projekt badawczy nr N N 111 477740.

niowy doboru drzewa dwumianowego do rzeczywistego szeregu czasowego przedstawiającego kurs wymiany euro do PLN. Przykład oparto na danych historycznych, dzięki czemu była możliwa weryfikacja procedury.

1. Drzewa dwumianowe

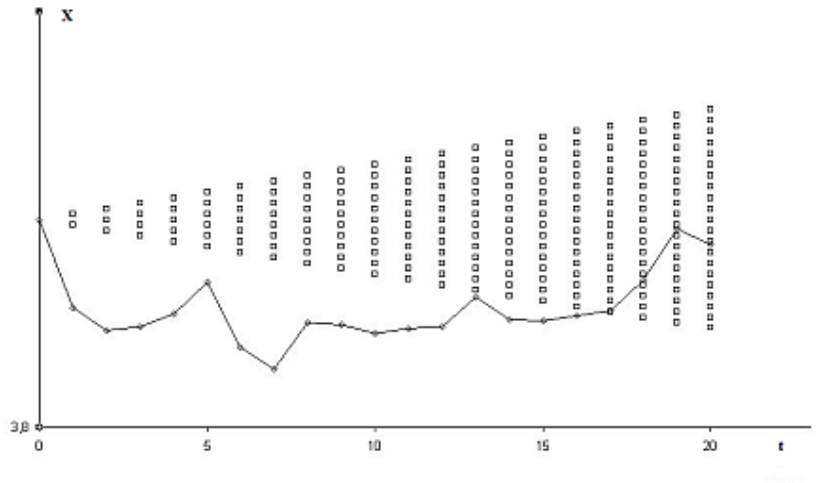
Drzewo dwumianowe, którego przykład pokazano na rys. 1, pełni rolę swoistego scenariusza możliwych zmian zmiennej stanu. Rozpatruje się tylko wzrost w stopniu u oraz spadek wartości w stopniu d. Rozważania rozpoczyna się od znanej obecnej wartości zmiennej stanu (X). Na podstawie historii zmian tej wielkości należy ustalić odpowiednie wartości u oraz d. Po ustaleniu horyzontu czasowego i podzieleniu go na okresy, reprezentowane przez węzły grafu, zostaje utworzone drzewo możliwych zmian zmiennej stanu, możliwych scenariuszy rozwoju sytuacji.



Rys. 1. Struktura drzewa dwumianowego

Źródło: Opracowanie własne.

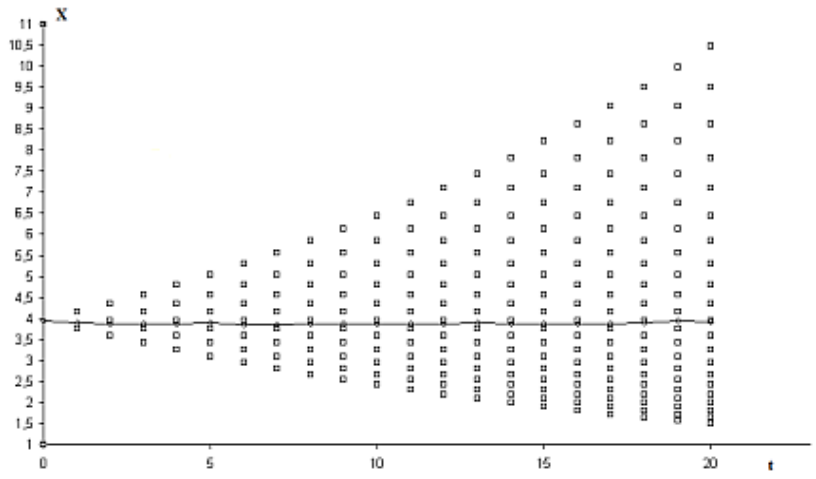
Dobranie zbyt małych wartości parametrów wzrostu i spadku może spowodować, iż drzewo dwumianowe nie pokryje przyszłych zmian procesu stochastycznego wartości X, jak pokazano na rys. 2.



Rys. 2. Drzewo dwumianowe dla u = 1,001 i d = 0,9990

Źródło: Opracowanie własne.

Dobranie zbyt dużych wartości parametrów wzrostu i spadku może spowodować, iż drzewo dwumianowe obejmie zbyt duże spektrum wartości, nie pokrywając drobnych zmian zmiennej stanu, co jest widoczne na rys. 3 przedstawiającym ten sam proces zmiennej stanu X.

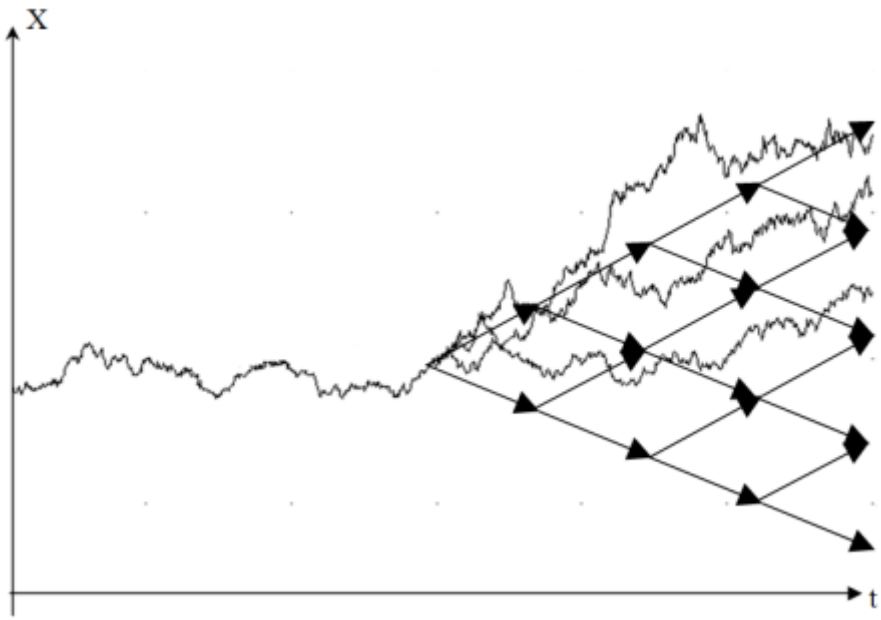


Rys. 3. Drzewo dwumianowe dla u = 1,05 i d = 0,9523

Źródło: Opracowanie własne.

Estymacja parametrów musi się rozpocząć od dobrania odpowiedniego procesu stochastycznego modelującego zmienność zmiennej stanu. Konsekwencją wybranego modelu jest sposób obliczania parametrów wzrostu i spadku oraz prawdopodobieństwa tych zmian. Na rysunku 4 pokazano drzewo dwumianowe dobrane na podstawie innego modelu, niż wygenerowane trzy ścieżki procesu stochastycznego. Niektóre ścieżki wychodzą poza obszar, który obejmuje drzewo. Dodatkowo wygenerowane procesy kształtują się w górnej części drzewa dwumianowego.

Wybór typu modelu jest nazywany modelowaniem, natomiast dobór parametrów modelu – kalibracją [Seydel 2009, s. 53].



Rys. 4. Błędnie dobrany model drzewa dwumianowego (drzewo dla procesu BM, ścieżki dla procesu GBM)

Źródło: Opracowanie własne.

2. Modele procesów zmian zmiennej stanu

W literaturze przedmiotu [Dixit, Pindyck 1994] są rozważane trzy zasadnicze grupy procesów stochastycznych wykorzystywanych do modelowania zmian zmiennej stanu. Są to procesy dyfuzyjne, procesy Poissona oraz procesy mieszane. W niniejszym artykule skupiono się na pierwszej grupie procesów. Omówiono tu procesy Arytmetycznego Ruchu Browna, Geometrycznego Ruchu

Browna oraz proces Orsteina-Uhlenbecka jako szczególny przypadek procesu z powrotem do średniej.

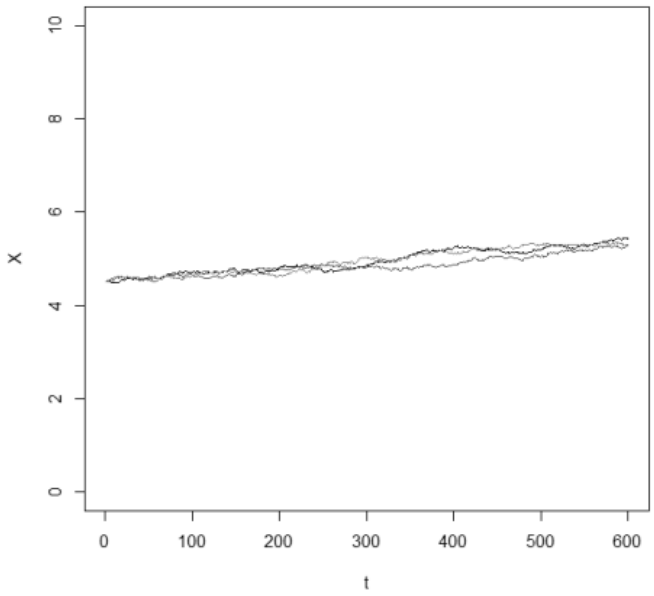
2.1. Arytmetyczny Ruch Browna

Proces stochastyczny nazywany Arytmetycznym Ruchem Browna (*Brownian Motion* – BM) jest dany równaniem różniczkowym [Weron, Weron 1998, s. 166]:

$$dX_t = \mu\, dt + \sigma\, dW_t \tag{1}$$

gdzie:
W_t – proces Wienera,
X_t – zmienna stanu,
μ – dryf procesu,
σ – zmienność procesu.

Wybór tego typu modelu może mieć uzasadnienie zwłaszcza w przypadku wskaźników technicznych. Przykładowe trajektorie takiego procesu przedstawiono na rys. 5.



Rys. 5. Arytmetyczny Ruch Browna dla $X_0 = 4{,}5$, $\mu = 0{,}08$ i $\sigma = 0{,}1$

Źródło: Opracowanie własne w programie R.

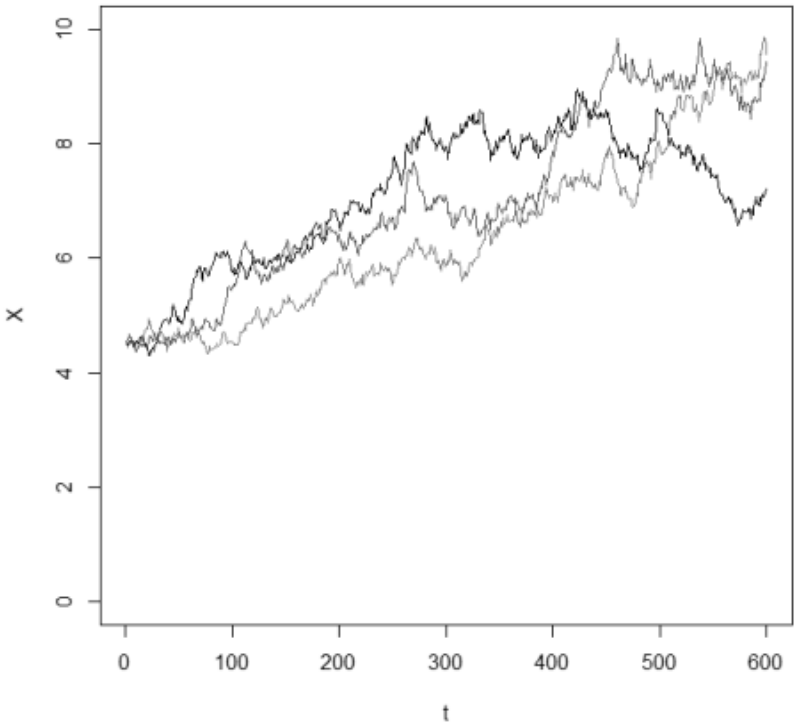
2.2. Geometryczny Ruch Browna

Proces stochastyczny nazywany Geometrycznym Ruchem Browna (*Geometric Brownian Motion* – GBM) jest dany równaniem różniczkowym [Weron, Weron 1998, s. 167]:

$$dX_t = \mu X_t\, dt + \sigma X_t\, dW_t \tag{2}$$

gdzie:
W_t – proces Wienera,
X_t – zmienna stanu,
μ – dryf procesu,
σ – zmienność procesu.

Przykładowe trajektorie takiego procesu przedstawiono na rys. 6.



Rys. 6. Geometryczny Ruch Browna $X_0 = 4{,}5$, $\mu = 0{,}08$ i $\sigma = 0{,}1$

Źródło: Opracowanie własne w programie R.

Wykorzystanie tego procesu ma szczególne uzasadnienie w przypadku, gdy zmienną stanu jest instrument finansowy notowany na giełdzie, taki jak akcje, dla których jest obserwowany ekspotencjalny wzrost wartości.

2.3. Procesy z powrotem do średniej

Grupa procesów, w których obserwuje się powrót do średniej (*Mean Reverting Model* – MRM), jest dana równaniem różniczkowym [Seydel 2009, s. 39]:

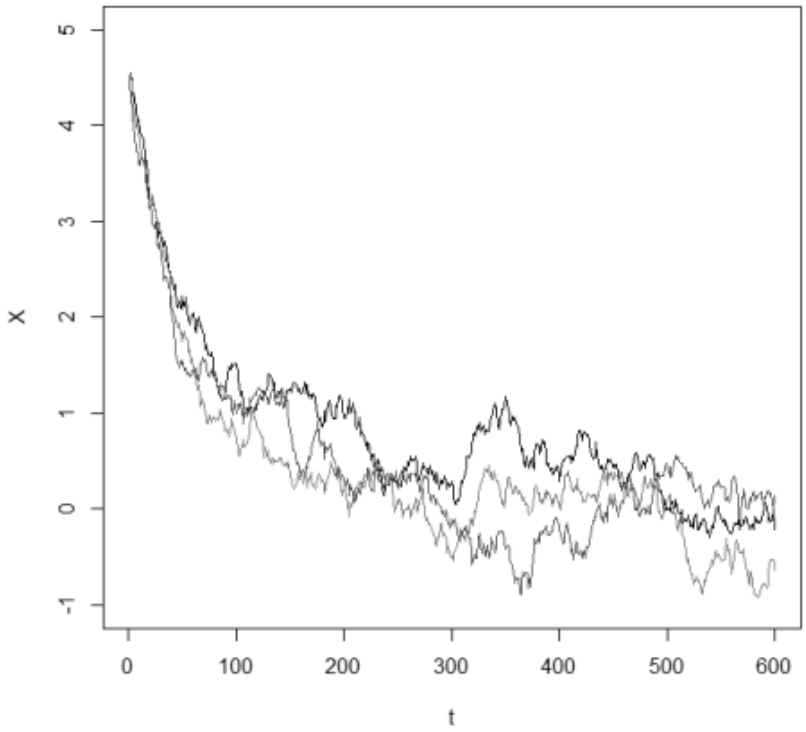
$$dX_t = \lambda(R - X_t)dt + \sigma X_t^{\beta} dW_t \qquad \text{dla } \lambda > 0 \tag{3}$$

gdzie:
W_t – proces Wienera,
X_t – zmienna stanu,
λ – parametr określający szybkość powrotu do średniej,
β – parametr,
σ – zmienność procesu,
R – średnia.

Dla szczególnych wartości parametrów uzyskujemy znane z literatury procesy [Gątarek, Maksymiuk 1998]:
– $\beta = 0$, $R = 0$ – proces Orsteina-Uhlenbecka,
– $\beta = 0$, $R > 0$ – model Vasička,
– $\beta = 1/2$, $R > 0$ – proces Coxa-Ingersolla-Rossa (CIR).

Wykorzystanie tej grupy procesów ma szczególne uzasadnienie ekonomiczne. Dla wielu wielkości ekonomicznych jest obserwowany efekt powrotu do pewnego ustalonego poziomu cen.

Przykładowe trzy trajektorie procesu Orsteina-Uhlenbecka przedstawiono na rys. 7.



Rys. 7. Proces Orsteina-Uhlenbecka dla $X_0 = 4{,}5$, $\lambda = 1{,}0$ i $\sigma = 0{,}5$

Źródło: Opracowanie własne w programie R.

3. Estymacja parametrów drzewa dwumianowego

Oceny właściwych wartości drzewa dwumianowego dokonujemy poprzez porównanie własności probabilistycznych wynikających z teoretycznego modelu oraz jego dyskretnego przybliżenia w postaci drzewa dwumianowego. Porównywane będą dwa pierwsze momenty, tj. wartość oczekiwana oraz wariancja procesu.

Dla ogólnej postaci procesu dyfuzyjnego:

$$dX_t = \alpha\,(X_t,\, t)\,dt + \sigma\,(X_t,\, t)\,dW_t \tag{4}$$

wartość oczekiwana jest dana wzorem:

$$\mathbf{E}[dX_t\,] = \alpha(X_t,\, t)\,dt \tag{5}$$

natomiast wariancja procesu jest dana wzorem:

$$\mathbf{Var}[\mathrm{dX_t}] = \sigma^2(X_t, t)\, dt \tag{6}$$

Poprzez porównanie powyższych zależności z wartościami dla modelu dyskretnego w postaci drzewa dwumianowego uzyskujemy wzory na parametry drzewa. Są one zależne od obserwowanej realizacji procesu stochastycznego, które Guthrie [2009, s. 267] zaleca obliczać jako średnią arytmetyczną ($\hat{v}$) z danych historycznych dla dryfu ($\hat{\mu}$), procesów GBM oraz BM:

$$\hat{\mu} = \frac{\hat{\nu}}{\Delta t_d} \tag{7}$$

Natomiast zmienność procesu ($\hat{\sigma}$) ocenia się na podstawie odchylenia standardowego ($\hat{\phi}$) z danych historycznych:

$$\hat{\sigma} = \frac{\hat{\phi}}{\sqrt{\Delta t_d}} \tag{8}$$

W obydwu przypadkach Δt_d oznacza część roku, dla której są podawane dane historyczne. Dla danych dziennych jest to 1/250 część roku, przy założeniu, iż mamy 250 dni roboczych w roku.

3.1. Dla Arytmetycznego Ruchu Browna

Na podstawie znajomości estymatora zmienności Arytmetycznego Procesu Browna ($\hat{\sigma}$) można obliczyć stopień wzrostu oraz spadku w drzewie dwumianowym obejmującym w najlepszym stopniu taki proces [Guthrie 2009, s. 327]:

$$u = \hat{\sigma}\sqrt{\Delta t_m} \tag{9}$$

$$d = -\hat{\sigma}\sqrt{\Delta t_m} \tag{10}$$

W tym przypadku Δt_m oznacza część roku odpowiadającej etapowi siatki dwumianowej.

Węzły siatki drzewa dwumianowego są obliczane ze wzoru [Guthrie 2009, s. 327]:

$$X(i,n) = X_0 + (n-2i)\cdot\hat{\sigma}\sqrt{\Delta t_m} \tag{11}$$

gdzie:
i – liczba spadków,
n – numer etapu.

3.2. Dla Geometrycznego Ruchu Browna

Na podstawie znajomości estymatora $\hat{\sigma}$ można obliczyć stopień wzrostu oraz spadku w drzewie dwumianowym obejmującym w najlepszym stopniu proces GBM [Guthrie 2009, s. 268]:

$$u = e^{\hat{\sigma}\sqrt{\Delta t_m}} \tag{12}$$

$$d = e^{-\hat{\sigma}\sqrt{\Delta t_m}} \tag{13}$$

Guthrie [2009, s. 268] podaje w przypadku wyboru Geometrycznego Ruchu Browna jako procesu zmian zmiennej stanu wzór na obliczanie węzłów siatki drzewa dwumianowego:

$$X(i,n) = X_0 e^{(n-2i)\cdot\hat{\sigma}\sqrt{\Delta t_m}} \tag{14}$$

gdzie:
i – liczba spadków,
n – numer etapu.

3.3. Dla procesów z powrotem do średniej

Do oceny estymatora zmienności ($\hat{\sigma}$) Guthrie [2009, s. 272] dla procesu Orsteina-Uhlenbecka wykorzystuje model autoregresyjny stopnia pierwszego (AR(1)) w postaci:

$$x_{j+1} - x_j = \alpha_0 + \alpha_1 x_j + u_{j+1},\ u_{j+1} \sim N(0,\ \phi^2) \tag{15}$$

gdzie:

$$x_j = ln(X_i)$$

Stąd estymatory modelu procesu stochastycznego mogą byś obliczone ze wzorów [Guthrie 2009, s. 274]:

$$\hat{R} = \frac{-\hat{\alpha}_0}{\hat{\alpha}_1} \tag{16}$$

$$\hat{\lambda} = \frac{-ln(1+\hat{\alpha}_1)}{\Delta t_d} \tag{17}$$

$$\hat{\sigma} = \hat{\phi}\sqrt{\frac{2\,ln(1+\hat{\alpha}_1)}{\hat{\alpha}_1(2+\hat{\alpha}_1)\Delta t_d}} \tag{18}$$

Dla procesów z powrotem do średniej sposób obliczania parametrów *u* i *d* jest identyczny jak dla procesu GBM (wzory 12 i 13). W ten sam sposób oblicza się także węzły siatki drzewa dwumianowego (wzór 14).

3.4. Weryfikacja modelu dyskretnego

Jakość pokrycia przez drzewo dwumianowe przyszłych zmian procesu stochastycznego ocenimy poprzez miernik średniokwadratowej odległości realizacji procesu od najbliższego węzła drzewa dwumianowego:

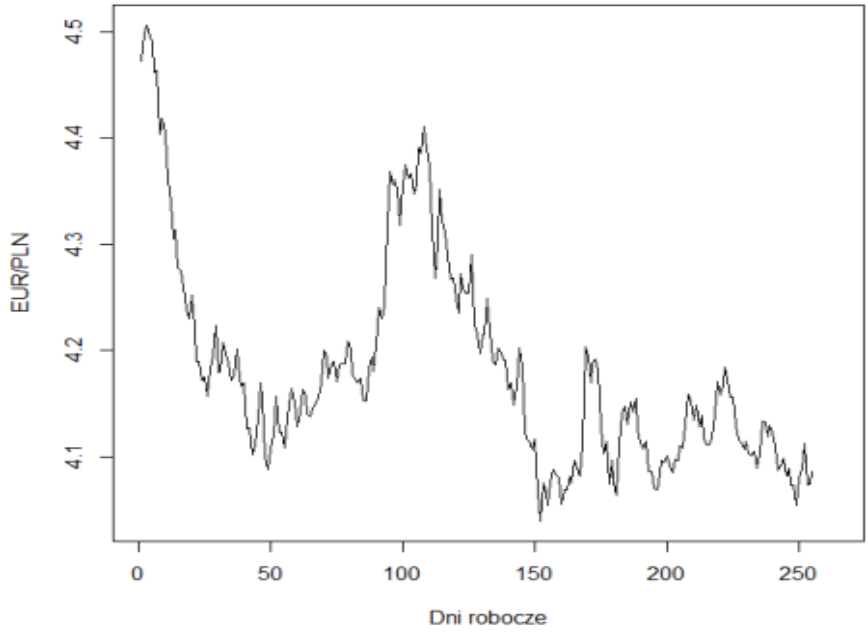
$$MAE_{min} = \frac{1}{M}\sum_{n=1}^{M}\left(\min_i |X(i,n) - x_n|\right) \tag{19}$$

gdzie:
$X(i,n)$ – dla $n = 1,..,M$, $i = 1,..,M$ – i-ty węzeł drzewa dwumianowego na n-tym etapie,
x_n – realizacja procesu losowego w momencie rozpoczęcia n-tego etapu,
M – liczba etapów w horyzoncie czasowym.

4. Przykład obliczeniowy

Jako przykład ilustrujący przedstawiane metody obliczono parametry siatki drzewa dwumianowego dla kursu wymiany euro do złotego (EUR/PLN). Na podstawie dziennych danych z 2012 r. (kursu wymiany pokazanego na rys. 8) wyestymowano parametry trzech rozważanych procesów stochastycznych, a na-

stępnie wykorzystano te parametry do utworzenia drzewa dwumianowego pokrywającego zmiany rozważanego kursu w styczniu 2013 r. Do weryfikacji uzyskanego rezultatu wykorzystano miarę $MAE_{\min}$.



Rys. 8. Proces zmian kursu wymiany EUR/PLN w 2012 r.

Źródło: Opracowanie własne.

Na podstawie notowań kursu wymiany EUR/PLN z 2012 r. estymowano średnią oraz odchylenie standardowe z dziennych wartości kursu zamknięcia. Dane te posłużyły do oszacowania parametrów procesu stochastycznego. Na ich podstawie obliczono właściwy stopień wzrostu (u) i spadku (d) wartości zmiennej stanu. Parametry te posłużyły do utworzenia drzewa dwumianowego z etapami dziennymi (M = 20). Na tak utworzone drzewo nałożono rzeczywisty przebieg kursów wymiany w styczniu 2013 r. Za pomocą wskaźnika $MAE_{\min}$ oceniono, jak blisko proces ten przechodził obok węzłów utworzonego drzewa. Uzyskane wyniki dla procesów BM i GBM przedstawiono w tabelach 1-2.

Tabela 1

Obliczenia dla BM

Δt_d	$\hat{\nu}$	$\hat{\mu}$	$\hat{\phi}$	$\hat{\sigma}$	Δt_m	u	d	$MAE_{\min}$
1/250	-0,00035	-0,0875	0,0054	0,0851	1/250	0,0054	-0,0054	0,0060

Źródło: Opracowanie własne.

Tabela 2

Obliczenia dla GBM

Δt_d	$\hat{\nu}$	$\hat{\mu}$	$\hat{\phi}$	$\hat{\sigma}$	Δt_m	u	d	$MAE_{\min}$
1/250	-0,00035	-0,0875	0,0054	0,0851	1/250	1,0054	0,9946	0,0112

Źródło: Opracowanie własne.

Dla modelu MRM-OU (procesu Orsteina-Uhlenbecka) wyestymowano parametry modelu AR(1), a następnie wykorzystano je do obliczenia estymatora $\hat{\sigma}$ zgodnie ze wzorem (18). Wyniki przedstawiono w tabeli 3.

Tabela 3

Obliczenia dla MRM-OU

Δt_d	$\hat{\nu}$	$\hat{\mu}$	$\hat{\phi}$	$\hat{\sigma}$	Δt_m	u	d	$MAE_{\min}$
1/250	-0,00035	-0,0875	0,0068	0,1375	1/250	1,0352	0,9660	0,0718

Źródło: Opracowanie własne.

Sumaryczne wyniki przedstawiono w tabeli 4.

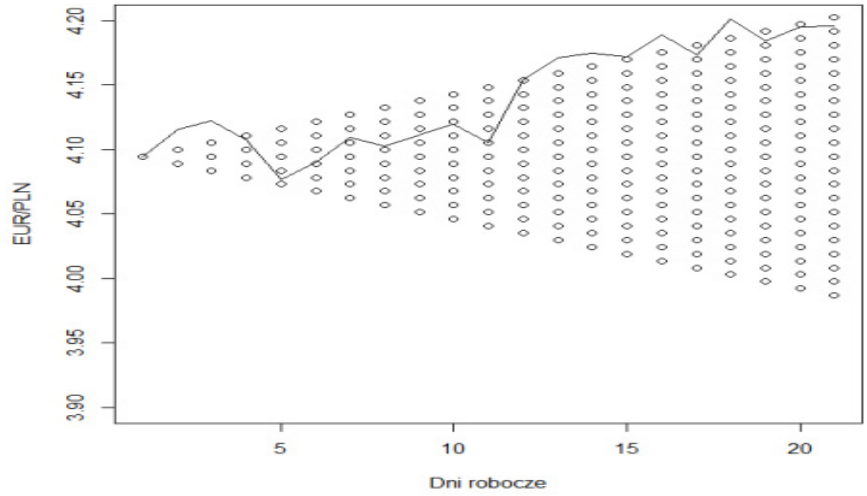
Tabela 4

Porównanie wyników dla modeli

Model	u	d	$MAE_{\min}$
BM	0,0054	– 0,0054	0,0060
GBM	1,0054	0,9946	0,0112
MRM-OU	1,0352	0,9660	0,0718

Źródło: Opracowanie własne.

Na rysunku 9 przedstawiono najlepsze pokrycie uzyskane dla modelu dla węzłów drzewa generowanych dla każdego dnia stycznia 2013 r., które uzyskano dla modelu Arytmetycznego Ruchu Browna (BM).



Rys. 9. Drzewo dwumianowe najlepiej dopasowane do zmian kursu wymiany (M = 20, u = 0,0054, model BM)

Źródło: Opracowanie własne.

Na przedstawionym rysunku widać, że pomimo najmniejszych odległości od węzłów uzyskanych dla procesu BM, siatka nie obejmuje całego spektrum zmian procesu kursu wymiany EUR/PLN.

Wnioski

W artykule podjęto tematykę dopasowywania drzew decyzyjnych do przyszłych przebiegów pewnych wielkości ekonomicznych. Problem jest istotny w wycenie opcji realnych za pomocą metody Coxa-Rossa-Rubinsteina. Dopasowanie odbywa się najpierw poprzez wybór modelu procesu stochastycznego obserwowanego szeregu czasowego, a następnie obliczenie parametrów drzewa dwumianowego, którymi są stopień wzrostu (u) i spadku (d) na każdym etapie. Na przykładach pokazano, jak niewłaściwy dobór tych parametrów może wpłynąć na fakt, iż drzewo dwumianowe nie pokryje przyszłych ścieżek procesu stochastycznego opisującego zmiany rozważanej wielkości ekonomicznej nazywanej w wycenie opcyjnej zmienną stanu lub może pokryć zbyt mało dokładnie. Do ilościowej oceny dopasowania wykorzystano wskaźnik bezwzględnej odległości realizacji procesu od najbliższego węzła drzewa dwumianowego. Przedstawione przykłady pokazały jednak niedoskonałość wykorzystanego wskaźnika. Rodzi to konieczność dal-

szych badań nad metodami weryfikacji dopasowania drzewa dwumianowego do rozważanego szeregu czasowego.

Bibliografia

Cox J.C., Ross S.A., Rubinstein M., 1979: Option Pricing: A Simplified Approach. „Journal of Financial Economics", No. 7, 229-263.

Dixit A.K., Pindyck R.S., 1994: Investment under Uncertainty. Princeton University Press, Princeton.

Gątarek D., Maksymiuk R., 1998: Wycena i zabezpieczenie pochodnych instrumentów finansowych. Wydawnictwo Liber, Warszawa.

Guthrie G., 2009: Real Options in Theory and Practice. Oxford University Press, Oxford.

Seydel R.U., 2009: Tools for Computational Finance. IV Ed. Springer-Verlag, Berlin.

Weron A., Weron R., 1998: Inżynieria finansowa. WNT, Warszawa.

MODELING CHANGES IN STATE VARIABLE FOR PURPOSE OF REAL OPTIONS VALUATION

Summary

The concept of real options mean the actual (real) opportunities arising in business processes. We are not obliged to use them. Noticing these capabilities creates added value of the project. Its use depends on quantitative measurement. It is assumed that this value is dependent on some economic size called state variable. Additional value is derived from the fact, that the state variable moves in a stochastic process, thus being able to achieve a advantageous level.

Widely used method for the valuation of real options is binomial tree method (CRR – Cox, Ross and Rubinstein). The idea is to cover the future trajectory of the state variable with lattice. The size of the lattice depends on the nature of the stochastic process, which we can model the state variable changes.

The presented work is devoted to determining, on the basis of past changes, the type of stochastic process which is best for modeling the state variable changes, and determine on this basis of the best lattice covering the future trajectory of this variable.

Grażyna Trzpiot
Anna Ojrzyńska
Uniwersytet Ekonomiczny w Katowicach

ANALIZA RYZYKA STARZENIA DEMOGRAFICZNEGO WYBRANYCH MIAST W POLSCE

Wstęp

Starzenie się populacji to pochodna przede wszystkim dwóch czynników: wzrostu przeciętnej długości życia oraz spadku płodności kobiet. Tempo tych przemian jest zróżnicowane regionalnie w zależności od szeregu czynników dodatkowych: migracji wewnętrznych i zagranicznych, polityki socjalnej wpływającej na podejmowanie decyzji o zakładaniu rodzin, stylu życia, modelu rodziny czy wzorców kariery zawodowej. W 2011 r. odsetek ludności w wieku 60/65 lat i więcej wynosił około 17,3%, natomiast odsetek ludności młodej (0-14 lat) – 15,1%. Systematyczny wzrost udziału osób starszych w populacji Polski wiąże się z wieloma konsekwencjami dla gospodarki kraju, w szczególności z sukcesywnym zmniejszaniem się liczby osób aktywnych zawodowo oraz wzrostem nakładów z budżetu państwa na system rent i emerytur oraz inne formy wsparcia i opieki dla osób starszych oferowane w zakresie polityki socjalnej państwa.

Celem artykułu jest analiza przestrzennego zróżnicowania dynamiki i zaawansowania starzenia się populacji wybranych miast Polski. Ujęcie przestrzenne pozwala dostrzec niejednorodność w strukturze ludności poszczególnych miast oraz w przebiegu przemian demograficznych. Wyznaczone wskaźniki starzenia się demograficznego oraz wskaźnik selektywności napływu do miast ludzi powyżej 50 roku życia pozwolą ocenić lukę demograficzną postrzeganą jako ryzyko funkcjonowania w zdrowiu w wybranych miastach w Polsce.

1. Metodologia badania starzenia się społeczeństwa

Wśród prostych miar zaawansowania starości w danym momencie czasu kalendarzowego można wymienić współczynnik starości demograficznej. Jest on wskaźnikiem struktury pokazującym udział frakcji traktowanej jako starszej w całej populacji [Cieślak 1992, s. 106]:

$$W_s = \frac{L_s}{L_{og}} \tag{1}$$

gdzie:

L_s – liczebność starszej części zbiorowości, np. kobiet w wieku 60 lat i mężczyzn w wieku 65 lat i więcej,

L_{og} – liczebność całej populacji.

Natomiast wśród najczęściej stosowanych miar, bazujących na ilościowych relacjach pomiędzy grupami wieku, można wymienić indeks starości demograficznej i współczynnik obciążenia demograficznego osobami starszymi. Indeks starości demograficznej I_{st} jest ilorazem liczby osób w starszym wieku L_s przez liczbę dzieci (0-14 lat) L_m w tej samej populacji [Długosz 1998, s. 19]:

$$I_{st} = \frac{L_s}{L_m} \tag{2}$$

Jak wiadomo, im wartość tego wskaźnika jest wyższa, tym społeczeństwo jest starsze, gdyż więcej ludności starszej przypada na określoną populację ludzi młodych. Z kolei współczynnik obciążenia demograficznego W_{od} osobami starszymi jest relacją pomiędzy liczebnością subpopulacji osób starszych L_s i liczebnością subpopulacji pozostałych osób dorosłych L_d żyjących w badanej populacji (np. w wieku 15-59/64 lata) [Cieślak 1992, s. 106]:

$$W_{od} = \frac{L_s}{L_d} \tag{3}$$

Natomiast do oceny zmian w procesie starzenia się społeczeństwa w określonym przedziale czasu można wykorzystać zaproponowany przez Z. Długosza wskaźnik starzenia się demograficznego (W_{sd}) [Długosz 1998, s. 19]:

$$W_{sd} = \left[U_{(0-14)t} - U_{(0-14)t+n}\right] + \left[U_{(>65)t+n} - U_{(>65)t}\right] \tag{4}$$

gdzie:

$U_{(0-14)t}$ – udział ludności w wieku 0-14 lat na początku badanego okresu,

$U_{(0-14)t+n}$ – udział ludności w wieku 0-14 lat na koniec badanego okresu,

$U_{(>65)t}$ – udział ludności w wieku 65 lat i więcej na początku badanego okresu,

$U_{(>65)t+n}$ – udział ludności w wieku 65 lat i więcej na koniec badanego okresu.

Im wartości wskaźnika W_{sd} będą niższe od 0, tym w większym stopniu będziemy mieli do czynienia z odmładzaniem się społeczeństwa, natomiast im wskaźnik ten będzie wyższy od 0, tym starzenie się ludności będzie dynamiczniejsze.

2. Metodologia badania migracji wewnętrznych ludności

Charakterystyczną cechą zjawiska migracji jest selektywność. W szerokim rozumieniu przejawem selektywności jest zróżnicowanie wpływu określonych cech populacji na skłonność do zaistnienia pewnego zjawiska. W badaniu jako miarę skłonności wykorzystano Współczynnik Selektywności Migracji (*WSM*) [Cieślak 1992, s. 248] zaadaptowany do opisu przepływów ludności wybranych miast Polski w wieku 50 lat i więcej.

Definicja miernika jest następująca:

$$WSM_{V=i} = \frac{\frac{M_{V=i}}{M} - \frac{P_{V=i}}{P}}{\frac{P_{V=i}}{P}} \tag{5}$$

gdzie:

V – zmienna, ze względu na którą badamy selektywność zjawiska (np. płeć, obecne i poprzednie miejsce zamieszkania),

i – kategoria zmiennej V, dla której jest liczona wartość współczynnika (np. kobiety, miasto),

$WSM_{V=i}$ – współczynnik selektywności ze względu na zmienną V dla kategorii i,

$M_{V=i}$ – liczebność podpopulacji badanej należącej do kategorii i oraz zmiennej V,

M – liczebność podpopulacji badanej ogółem,

$P_{V=i}$ – liczebność populacji badanej należąca do kategorii i oraz zmiennej V,

P – całkowita liczebność populacji badanej.

Współczynnik ten może przyjmować wartości z zakresu $[-1;+\infty]$. Wartości dodatnie świadczą o występowaniu dodatniej selektywności, tym wyższej, im wyższa wartość współczynnika. Oznacza to, że w danym zjawisku uczestniczy więcej jednostek danej kategorii niż wynikałoby to z ich proporcji w całej populacji [Mioduszewska 2008, s. 16]. Mówimy, że zjawisko selektywności nie występuje, gdy wartości omawianego współczynnika wynoszą zero lub są bliskie zeru. WSM może również posłużyć do określenia, które z badanych cech silniej selekcjonują podpopulację z danego obszaru.

3. Metodologia badania trwania życia w zdrowiu

Koncepcja oczekiwanej długości życia w zdrowiu została opracowana początkowo w celu oceny, czy dłuższy czas trwania życia jest związany z wydłużeniem się czasu życia w dobrym zdrowiu, czy też w złym. Zatem konieczne stało się określenie relacji między trwaniem życia a trwaniem życia w zdrowiu, nazywanej „luką zdrowotną". Miarami pozwalającymi na określenie lat życia przeżytych w pełnym zdrowiu oraz lat przeżytych z pewnymi dysfunkcjami i niesprawnością jako równoważnik tych pierwszych są m.in.: oczekiwana długość życia bez niesprawności oraz oczekiwana długość życia korygowana niesprawnością. Miary te należą do grupy wskaźników nazywanych sumarycznymi miarami stanu zdrowia.

Oczekiwana długość życia bez niesprawności (*Disability Free Life Expectancy* – **DFLE**) została zaproponowana jako jeden ze wskaźników monitorujących stan zdrowia w krajach europejskich. Od 2004 r. wskaźniki te są publikowane pod nazwą lata życia w zdrowiu (*Healthy Life Years* – **HLY**) i są opracowywane przez Eurostat dla krajów Unii Europejskiej [Wróblewska 2008, s. 153-154]. Metoda ta zakłada, że:

$$DFLE_x = \frac{\sum_{i=x}^{\omega} YWD_i}{l_x} \tag{6}$$

$$YWD_x = L_x \cdot (1 - prev_x) \tag{7}$$

gdzie:
YWD – liczba lat życia bez niesprawności w wieku *x*,
l_x – liczba osób dożywających wieku *x* ukończonych lat,
L_x – liczba lat życia przeżytych w wieku *x* na podstawie tablic trwania życia,
$prev_x$ – częstość występowania niesprawności w stanie zdrowia.

Tak więc metoda ta opiera się na dwóch miarach: częstości występowania niepełnosprawności w populacji w określonym wieku oraz umieralności. HLY oblicza się na podstawie tablic umieralności i indywidualnie postrzeganej niepełnosprawności określanej z użyciem standardowych kwestionariuszy wywiadu.

Natomiast miarą używaną przez Światową Organizację Zdrowia jest oczekiwana długość życia korygowana niesprawnością *(Disability-adjusted Life Expectancy* – **DALE**). W celu obliczenia DALE, od oczekiwanej długości życia odejmuje się liczbę lat życia utraconych w wyniku jakiejkolwiek niesprawności (sumę wszystkich utraconych lat pełnego zdrowia):

$$DALE_x = \frac{\sum_{i=x}^{\omega} YLH_i}{l_x} \tag{8}$$

$$YLH_x = L_x \cdot \sum_{i}^{s} w_s \cdot D_{si} = L_x \cdot \sum_{i}^{s} (1 - dw_s \cdot D_{si}) \tag{9}$$

gdzie:

YLH_x – liczba lat przeżytych w zdrowiu w wieku *x*,

L_x – liczba lat życia przeżytych w wieku *x* na podstawie tablic trwania życia,

D_s – częstość występowania niesprawności *s*, gdzie $\sum_s D_s = 1$,

s – poziom niesprawności, gdzie $s = 0$ oznacza brak niesprawności,

w_s – waga stanu zdrowia przypisana niesprawności s ($w_s = 1 - dw_s$), gdzie stan dobrego zdrowia określa $w_s = 1$ lub $dw_s = 0$.

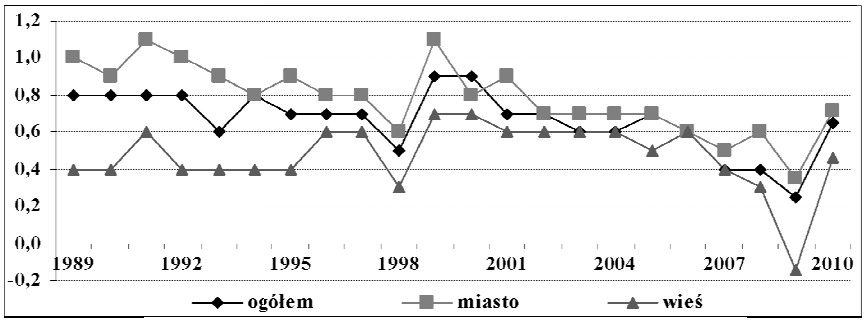
W 2001 r. wskaźnik DALE zastąpiono wskaźnikiem oczekiwanej długości życia skorygowanej ze względu na stan zdrowia (*Healthy Life Expectancy* – **HALE**). Wskaźnik HALE mówi o przeżywalności w różnych stanach zdrowia, a więc nie tylko w pełnym zdrowiu. W związku z tym charakteryzuje go duża zmienność w czasie oraz różnice pomiędzy krajami w zakresie danych dotyczących stanu zdrowia [Gromulska, Wysocki, Goryński 2008]. Jest szacowany przez WHO na podstawie danych o umieralności (oczekiwanego trwania życia według wieku i statystyki zgonów według przyczyn) oraz danych epidemiologicznych dotyczących zapadalności i chorobowości [Wróblewska 2008]. Na podstawie danych dotyczących światowego obciążenia zdrowotnego (*Global Burden Disease*) szacuje się nasilenie zachorowań ze względu na płeć i przedziały wiekowe dla każdego kraju. Metoda obliczania DALE/HALE jest bar-

dziej zaawansowana i złożona niż metoda obliczania HLY i w związku z tym trudniej jest zebrać kompletne dane potrzebne do obliczeń.

4. Analiza empiryczna

Badanie starzenia się społeczeństwa przeprowadzono dla wybranych miast Polski, tj.: Gdańska, Katowic, Krakowa, Poznania, Warszawy i Wrocławia. Analizę poziomu zaawansowania tego procesu oraz zachodzących w czasie zmian przeprowadzono na podstawie danych Głównego Urzędu Statystycznego na temat stanu i struktury ludności w latach 1989-2011. Z kolei informacje o rozmiarach i strukturze migracji wewnętrznych dotyczą 2011 r. Podstawą analizy empirycznej trwania życia w zdrowiu były dane zaczerpnięte z Europejskiego System Statystycznego, a także baza danych oraz publikacje przygotowywane przez Światową Organizację Zdrowia (WHO).

Celem dokładniejszego uchwycenia trendów w ostatnim dwudziestopięcioleciu, w którym zachodziły dość znaczące zmiany w procesie starzenia się ludności Polski, podjęto próbę wykorzystania wskaźnika starzenia się demograficznego (W_{sd}). Posługując się rocznymi zmianami tego wskaźnika, w przekroju lat 1989-2011, obliczono dla miast i wsi wskaźnik starzenia (odmładzania) się ludności Polski ogółem (rys. 1).



Rys. 1. Roczne zmiany wskaźnika starzenia się demograficznego ludności Polski w punktach procentowych

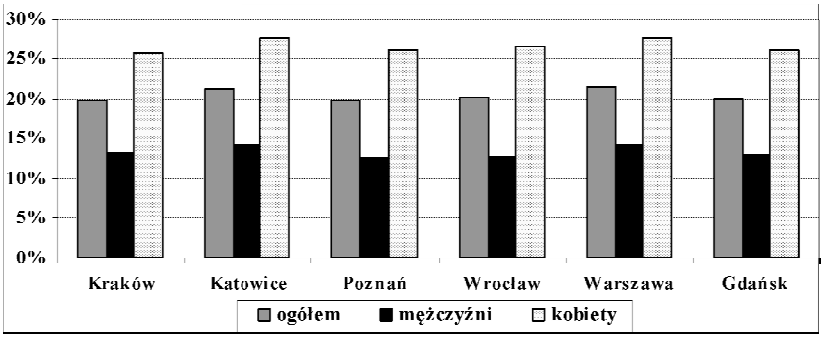
Cykliczność wielkości wskaźników W_{sd} w miastach i na terenach wiejskich pokrywa się z tym, że wielkości wskaźnika w całym okresie badania dla miast była wyższa niż na terenach wiejskich. Dodatkowo przyjmując 1989 r. za początek okresu badania, a 2011 r. za koniec, można zauważyć, że proces starzenia się ludności był znacznie dynamiczniejszy w miastach niż na wsi (tabela 1).

Tabela 1

Wskaźnik starzenia się demograficznego ludności Polski w latach 1989-2011 w punktach procentowych

Wskaźnik starzenia się demograficznego		
ogółem	**miasto**	**wieś**
10,00	17,16	10,42

Dalsza część badania charakteryzuje proces starzenia się ludności w wybranych miastach Polski. W najprostszy sposób poziom starości określono obliczając udział najstarszej grupy wiekowej w ogóle populacji w 2011 r. (rys. 2). Najwyższym odsetkiem ludności starszej charakteryzowali się mieszkańcy Katowic i Warszawy (27%). Potwierdzają to również obliczone wartości tego współczynnika oddzielnie dla kobiet i mężczyzn. Stan zaawansowania starości demograficznej był wyższy dla kobiet niż dla mężczyzn, głównie ze względu na różnice wynikające z przeciętnej długości trwania życia według płci.

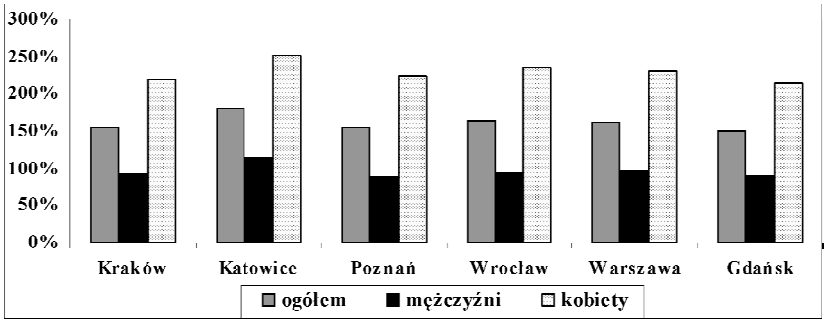


Rys. 2. Współczynnik starości demograficznej wybranych miast w 2011 r.

Stosowanie jednej wartości, w tym przypadku udziału najstarszej grupy wieku, do oceny stanu procesu starzenia się nie zawsze jest adekwatne do rzeczywistości, gdyż w pełni nie oddaje sytuacji demograficznej na badanym obszarze. Dlatego wykorzystano również mierniki bazujące na dwóch kryteriach liczbowych, tj. indeks starości demograficznej i współczynnik obciążenia demograficznego osobami starszymi.

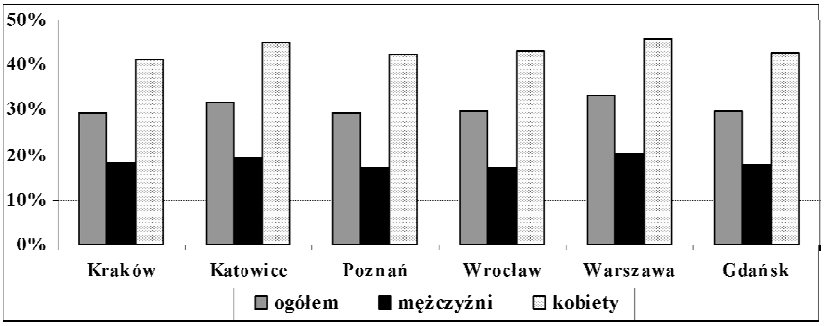
Najwyższy poziom zaawansowania procesu starzenia się ludności, mierzony obciążeniem grupy najmłodszej grupą najstarszą, ma miejsce w Katowicach (rys. 3), a udział kobiet w wieku 60 lat i więcej stanowi 252% liczby kobiet w wieku 0-14 lat. Wartość tego indeksu dla kobiet mieszkających w Gdańsku wynosi 214%. Znacznie mniej rozwinięty proces starzenia dotyczy subpopulacji mężczyzn. Spośród wybranych miast wartość tego indeksu jest najwyższa również w Katowicach

(112%), a najniższa w Gdańsku (90%). Zastosowany indeks jednocześnie uwzględnił sytuację w najmłodszej grupie wieku, determinowaną poziomem urodzeń.



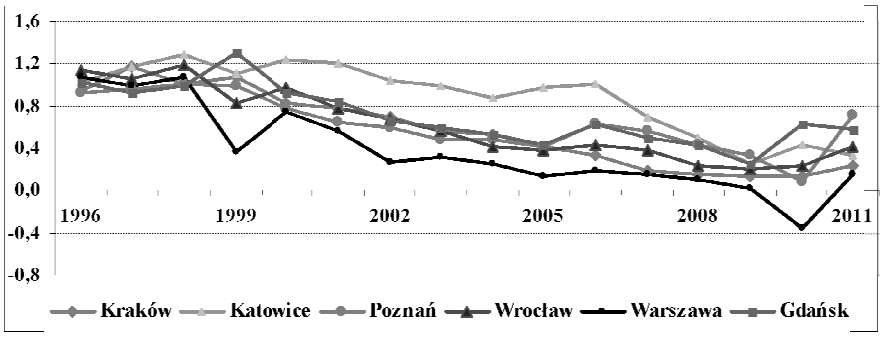
Rys. 3. Indeks starości demograficznej wybranych miast w 2011 r.

Dodatkowo dla wybranych miast Polski wyznaczono relację pomiędzy liczebnością ludności starszej (w wieku 60/65 lat i więcej) a liczebnością pozostałych osób dorosłych żyjących w badanej populacji, tj. w wieku 15-59/64 lat. Wyniki przedstawiono na rys. 4. Najwyższe obciążenie ludnością starszą ludności w wieku produkcyjnym dotyczyło w 2011 r. Warszawy i Katowic. Natomiast najmniej zaawansowany proces starzenia się dotyczył mieszkańców Krakowa.



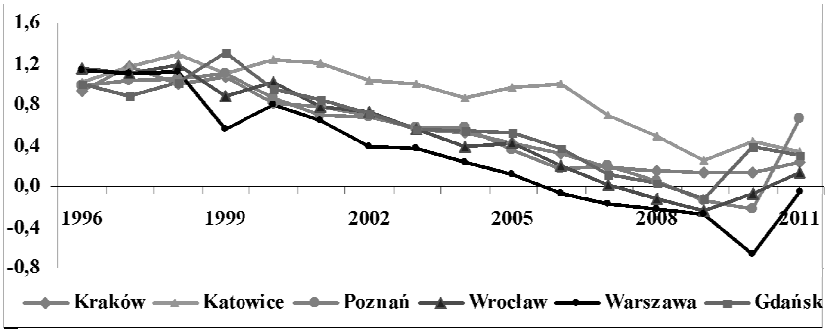
Rys. 4. Współczynnik obciążenia demograficznego osobami starszymi wybranych miast w 2011 r.

Dotychczas obliczone mierniki były stosowane do określenia zaawansowania starości w danym momencie czasu kalendarzowego, czyli w 2011 r. Do określenia zmian w procesie starzenia się wybranych miast Polski w latach 1995-2011 wykorzystano wskaźnik starzenia się demograficznego, a wyniki przedstawiono na rys. 5-7.



Rys. 5. Roczne zmiany wskaźnika starzenia się demograficznego ogółem wybranych miast w punktach procentowych

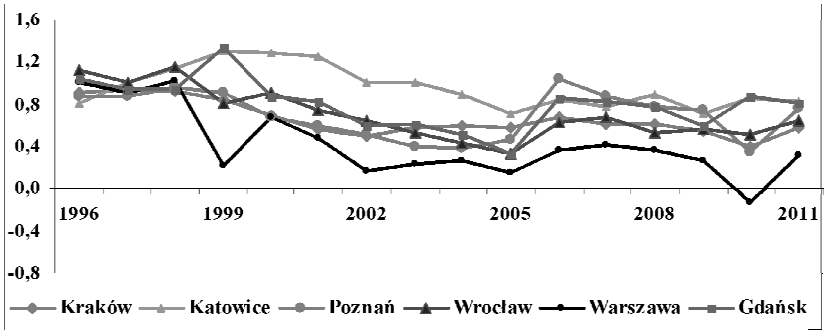
Roczne zmiany tego wskaźnika, w przekroju lat 1995-2011, obliczone dla wybranych miast wskazują, iż nastąpiła zmiana w szybkości zaawansowania procesu starzenia się ludności. Dodatnie wartości tego wskaźnika w badanym okresie (z wyjątkiem Warszawy w 2009 r.) obliczone dla ludności ogółem wskazują, iż rozwój procesu starzenia się jest obecnie mniej dynamiczny niż na początku okresu badania.



Rys. 6. Roczne zmiany wskaźnika starzenia się demograficznego mężczyzn wybranych miast w punktach procentowych

Z kolei na podstawie obliczonych wartości tego wskaźnika dla mężczyzn można zaobserwować, iż w latach 2008-2010 w takich miastach, jak Warszawa, Wrocław, Poznań i Gdańsk, wystąpiło zjawisko odmładzania się społeczeństwa. Odmiennie kształtowało się to zjawisko w subpopulacji kobiet. Z wyjątkiem Warszawy oraz Katowic tempo zmian tego procesu kształtowało się na stabilnym poziomie 0,6 p.p. rocznego wzrostu. Natomiast w Warszawie dynamika zmian procesu starzenia się ludności była znacznie wyższa na początku badane-

go okresu niż obecnie. Dodatkowo można zaobserwować krótkotrwałe zjawisko odmładzania się społeczeństwa tego miasta w 2010 r. Z kolei w Katowicach najbardziej dynamiczny rozwój tego procesu dotyczy lat 1998-2002, po których nastąpiła stabilizacja na poziomie wzrostu o 0,8 p.p. z roku na rok.



Rys. 7. Roczne zmiany wskaźnika starzenia się demograficznego kobiet wybranych miast w punktach procentowych

Dodatkowo porównując zmianę w strukturze ludności według wieku pomiędzy 1995 a 2011 r. można zauważyć, że proces starzenia się ludności rozwijał się najszybciej w Katowicach (tabela 2).

Tabela 2

Wskaźnik starzenia się demograficznego wybranych miast w latach 1995-2011 w punktach procentowych

Wyszczególnienie	Kraków	Katowice	Poznań	Wrocław	Warszawa	Gdańsk
Ogółem	10,0	14,8	10,1	9,9	6,1	11,3
Mężczyźni	9,2	14,1	8,7	8,2	5,1	9,5
Kobiety	10,6	15,3	11,2	11,3	6,8	12,8

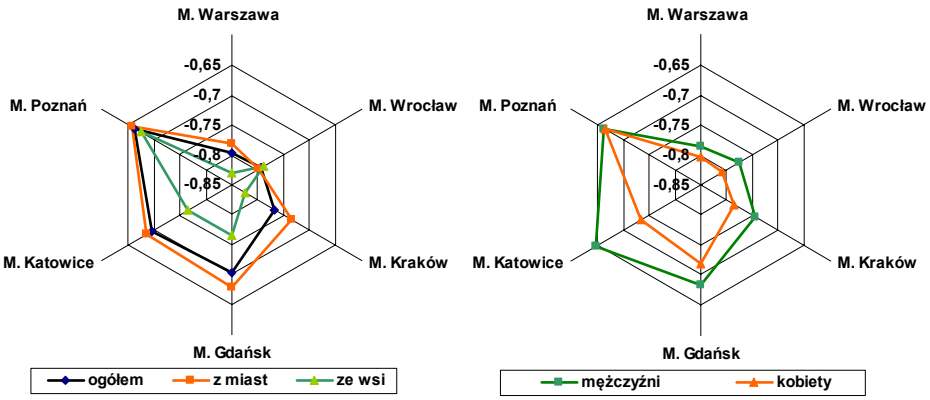
Jednym z czynników dodatkowych wpływających na tempo zmian procesu starzenia się ludności jest zjawisko migracji. Dlatego też w dalszej części artykułu dokonano oceny selektywności migracji do i z miast osób powyżej 50 roku życia. Na podstawie przedstawionych w tabeli 3 udziałów ludności w wieku 50 lat i więcej w napływie i odpływie z wybranych miast Polski w 2011 r. można stwierdzić, iż zaawansowanie procesu starzenia się strumieni napływu i odpływu jest niewielkie.

Tabela 3

Natężenie migracji ludności w wieku 50 lat i więcej w 2011 r.

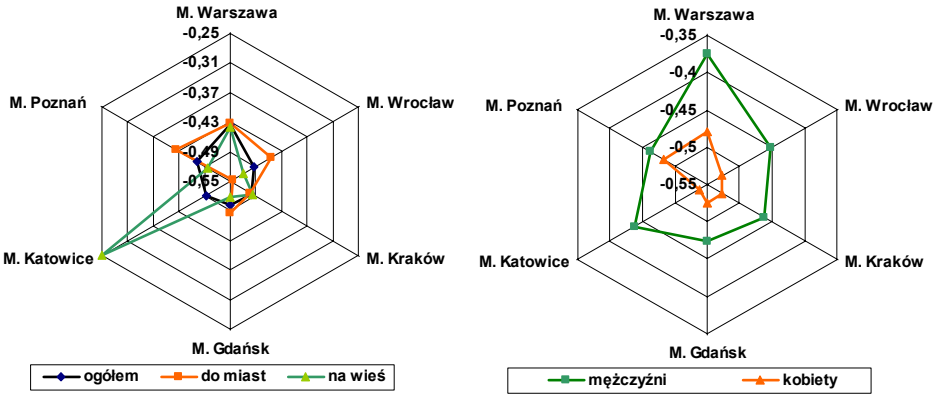
Udział ludności 50+	Kraków	Katowice	Poznań	Wrocław	Warszawa	Gdańsk
W napływie	8,6%	12,2%	10,7%	7,9%	8,0%	11,3%
W odpływie	18,5%	20,5%	16,7%	19,5%	22,2%	18,3%

Dodatkowo aby ocenić skłonność do migracji osób w wieku 50 lat i więcej, zostały wyznaczone wskaźniki selektywności migracji przedstawione na rys. 8 i 9. Ujemne wartości tego wskaźnika wskazują na brak skłonności migracji poza granice swojego miasta osób w badanym wieku. Większą awersję do napływu do miast miały osoby zamieszkujące wcześniej obszary wiejskie, w szczególności dotyczy to napływu do takich miast, jak Warszawa, Wrocław i Kraków. Można również zauważyć, iż kobiety rzadziej niż mężczyźni decydowały się na imigrację do miast.



Rys. 8. Wartości WSM napływu ludności w wieku 50 lat i więcej do wybranych miast w 2011 r.

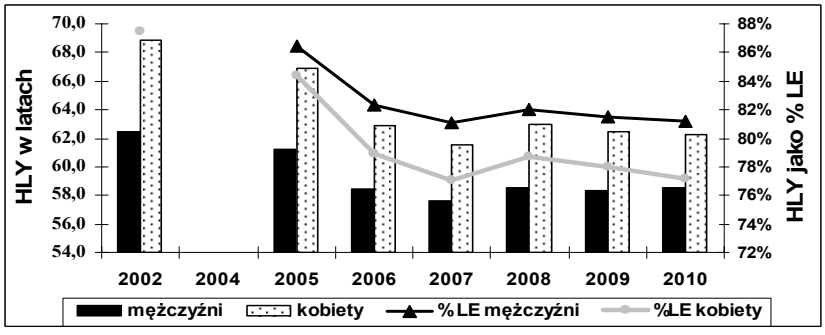
Ponadto osoby w wieku 50 lat i więcej nie wykazywały skłonności do emigracji z miast, a wręcz charakteryzowała je silna awersja do opuszczenia swojego miasta. Wyjątek stanowili mieszkańcy Katowic, dla których awersja do emigracji nie była tak silna, w szczególności jeżeli dotyczyła emigracji na wieś.



Rys. 9. Wartości WSM odpływu ludności w wieku 50 lat i więcej z wybranych miast w 2011 r.

Podobnie jak w przypadku napływu do miast, tak i w przypadku odpływu z miast kobiety zdecydowanie mocniej nie wykazywały skłonności do migracji niż mężczyźni.

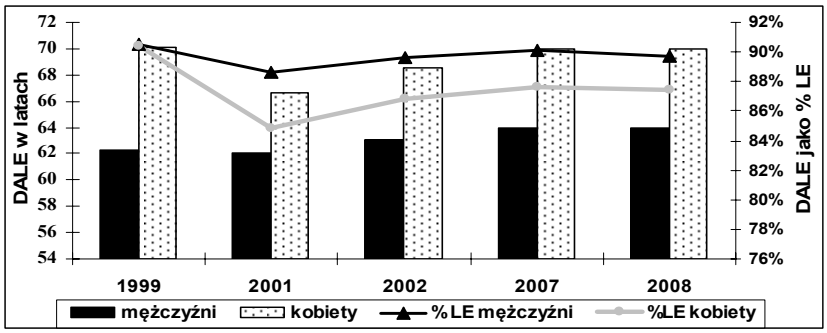
Wydłużony okres dalszego trwania życia, obok spadku płodności, stanowi powszechną przyczynę starzenia się ludności. Oczekiwana długość życia w Polsce w 2010 r. wyniosła 80,7 lat dla kobiet i 72,1 lat dla mężczyzn. Ważne zdaje się być pytanie, czy dłuższy czas trwania życia jest związany z wydłużeniem się czasu życia w dobrym zdrowiu. Odpowiedź na nie mogą dać obliczone sumaryczne miary stanu zdrowia. W 2010 r. długość trwania życia w zdrowiu (HLY) w Polsce wyniosła 62,3 i 58,5 lat odpowiednio dla kobiet i mężczyzn. Porównując oczekiwane lata życia w zdrowiu z przeciętnym dalszym trwaniem życia, możemy stwierdzić, iż w badanym roku 77,2% lat życia kobiet i 81,14% lat życia mężczyzn w Polsce mija bez negatywnych ocen stanu zdrowia.



Rys. 10. Oczekiwane lata życia w zdrowiu HLY oraz udział oczekiwanych lat życia w zdrowiu w oczekiwanej długości trwania życia

Porównując 2002 r. z 2010 możemy stwierdzić, iż oczekiwana liczba lat życia w zdrowiu Polek zmniejszyła się o 6,6 lat. Tym samym udział lat życia w zdrowiu w stosunku do przeciętnego dalszego trwania życia Polek obniżył się z poziomu 88,4% w 2002 r. do 77,2% w 2010 r. Oznacza to, iż wzrostowi oczekiwanego trwania życia kobiet w Polsce nie towarzyszy wzrost (a niestety spadek) lat przeżytych w zdrowiu. Należy zwrócić uwagę na fakt, iż procentowe udziały HLY w oczekiwanej długości życia kobiet są niższe niż mężczyzn. Przyczyną takiej relacji może być m.in. dłuższe oczekiwane trwanie życia kobiet. Odnotowano również spadek lat życia w zdrowiu mężczyzn, porównując 2002 r. do 2010.

Na podstawie danych o umieralności oraz danych epidemiologicznych dotyczących zapadalności i chorobowości została oszacowana przez Światową Organizację Zdrowia oczekiwana długość życia skorygowana o stan zdrowia DALE/HALE (rys. 11).



Rys. 11. Oczekiwana długość życia skorygowana ze względu na zdrowie DALE/HALE oraz jej udział w oczekiwanej długości trwania życia

W przypadku kobiet najniższą wartość DALE/HALE odnotowano w 2001 r. W kolejnych latach możemy zauważyć poprawę w poziomie tego wskaźnika. Natomiast w przypadku mężczyzn z roku na rok obserwujemy poprawę oczekiwanej długości lat życia skorygowanych ze względu na stan zdrowia. Analizując udział procentowy DALE/HALE w oczekiwanej długości życia kobiet od 2001 r., możemy zauważyć delikatny trend wzrostowy, co oznacza szybszy wzrost wartości DALE/HALE w porównaniu ze wzrostem oczekiwanego trwania życia.

Zakończenie

Cechą charakterystyczną struktury wieku populacji Polski jest falowanie wyżów i niżów demograficznych, co przekłada się na pewną zmienność w czasie przedstawionych wskaźników. Jednakże niezależnie od perturbacji spowodowanych pofałdowaną strukturą wieku można uznać, że przebieg procesu demograficznego starzenia sie populacji Polski ma charakter systematyczny. Zjawisko to jest silniej odczuwalne na obszarach miejskich niż wiejskich. Z roku na rok jest obserwowany rozwój tego procesu w badanych miastach Polski. Najbardziej zaawansowany proces starzenia się ludności jest widoczny w Katowicach. Z kolei zdecydowanie mniejszą dynamiką charakteryzuje się to zjawisko w Warszawie, gdzie można było obserwować krótkookresowe odmładzanie się społeczeństwa tego miasta. Ponadto starzenie się ludności silniej dotyczy subpopulacji kobiet niż mężczyzn. Jednakże na przebieg starzenia się ludności badanych miast nieznaczny wpływ miały migracje wewnętrzne ludności w wieku 50 lat i więcej. Co więcej, subpopulacja ta wykazywała w 2011 r. brak skłonności do ruchu wędrówkowego, a szczególnie wyraźnie jest to zaznaczone w przypadku kobiet. Na podstawie lat życia w zdrowiu obliczonych za pomocą samooceny stanu zdrowia możemy wnioskować, iż ryzyko „luki zdrowotnej" w ostatnich latach się zwiększa. Do odwrotnych wniosków można dojść, analizując oczekiwaną długość życia skorygowaną o stan zdrowia. Na podstawie tej miary można twierdzić, iż w ostatnich latach zaobserwowano, iż relacja trwania życia oraz trwania życia w zdrowiu utrzymuje się na niezmienionym poziomie, a nawet w ostatnim roku udział lat przeżytych w zdrowiu do ogółu lat życia się zwiększył. Tak więc zmniejszyło się ryzyko „luki zdrowotnej". Problem z porównywalnością sumarycznych miar stanu zdrowia wynika z faktu, że są one kombinacją oczekiwanego trwania życia oraz życia w określonych stanach zdrowia, które mogą być różnie definiowane. Nie bez znaczenia jest to, że oszacowania tych miar pochodzą z różnych baz danych i są obliczane dla różnych okresów.

Bibliografia

Cieślak M., 1992: Demografia. Metody analizy i prognozowania. Wydawnictwo Naukowe PWN, Warszawa.

Cieślak M., 2004: Pomiar procesu starzenia się. „Studia Demograficzne", nr 2/146.

Długosz Z., 1998: Próba określania zmian starości demograficznej Polski w ujęciu przestrzennym. „Wiadomości Statystyczne", nr 3.

Gromulska L., Wysocki M.J, Goryński P., 2008: Lata przeżyte w zdrowiu (Healthy Life Years, HLY) – zalecany przez Unię Europejską syntetyczny wskaźnik sytuacji zdrowotnej ludności. „Przegląd Epidemiologiczny", 62(4).

Holzer J., 2003: Demografia. Polskie Wydawnictwo Ekonomiczne, Warszawa.

Mioduszewska M., 2008: Najnowsze migracje z Polski w świetle danych Badania Aktywności Ekonomicznej Ludności. OBM WNE UW, Warszawa.

Wróblewska W., 2008: Sumaryczne miary stanu zdrowia populacji. „Studia Demograficzne", 153-154 (1-2).

RISK ANALYSIS OF DEMOGRAPHIC AGING IN SELECTED POLAND CITIES

Summary

In the paper work we will take study of the occurrence of demographic aging risk. We will use population study carried out in 2011, and the methods of analysis of demographic and regional indicators. We will begin with setting the rate of demographic aging and the selectivity index influx to urban people over the age of 50, and then move on to assess the demographic gap perceived as functioning in health risk in selected cities in Poland.