

Uniwersytet
Ekonomiczny
w Krakowie

Zeszyty Naukowe

Cracow Review
of Economics
and Management

916

Metody
analizy danych

Kraków 2013

Rada Naukowa

Andrzej Antoszewski (Polska), Slavko Arsovski (Serbia), Josef Arlt (Czechy), Daniel Baier (Niemcy), Hans-Hermann Bock (Niemcy), Ryszard Borowiecki (Polska), Giovanni Lagioia (Włochy), Tadeusz Markowski (Polska), Martin Mizla (Słowacja), David Ost (USA), Józef Pociecha (Polska)

Komitet Redakcyjny

Paweł Lula, Barbara Pawełek (sekretarz), Adam Sagan, Tadeusz Sikora, Edward Smaga, Andrzej Sokołowski (przewodniczący)

Redaktor statystyczny

Andrzej Sokołowski

Redaktorzy Wydawnictwa

Hanna Wojciechowska, Seth Stevens (streszczenia w j. angielskim)

Projekt okładki i układ graficzny tekstu

Marcin Sokołowski

Streszczenia artykułów są dostępne w międzynarodowej bazie danych The Central European Journal of Social Sciences and Humanities <http://cejsh.icm.edu.pl> oraz w Central and Eastern European Online Library www.ceeol.com, a także w adnotowanej bibliografii zagadnień ekonomicznych i pokrewnych BazEkon http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

© Copyright by Uniwersytet Ekonomiczny w Krakowie, Kraków 2013

ISSN 1898-6447

Wersja pierwotna: publikacja drukowana
Publikacja jest dostępna w bazie CEEOL (www.ceeol.com)
oraz w czytelni on-line ibuk.pl (www.ibuk.pl)

Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie
31-510 Kraków, ul. Rakowicka 27, tel. 12 293 57 42, e-mail: wydaw@uek.krakow.pl
www.zeszyty-naukowe.uek.krakow.pl

Zakład Poligraficzny Uniwersytetu Ekonomicznego w Krakowie
31-510 Kraków, ul. Rakowicka 27

Objętość 4,8 ark. wyd.
Zam. 102/14

Spis treści

| | |
|--|-----------|
| Małgorzata Rószkiewicz | |
| Problemy analityczne metaanalizy – efekt procesu badawczego | 5 |
| Aneta Rybicka | |
| Od <i>conjoint analysis</i> do metod wyborów opartych na menu | 13 |
| Paweł Chlipała | |
| Orientacja hybrydowa a nowe podejścia w marketingu | 25 |
| Alicja Grześkowiak, Agnieszka Stanimir | |
| Wielowymiarowa analiza społeczno-demograficznych aspektów wykorzystania Internetu | 37 |
| Piotr Tarka | |
| Model czynnikowy ze zmiennymi dychotomicznymi w analizie ukrytej struktury zjawisk rynkowych | 47 |
| Mirosława Kaczmarek | |
| Problemy stosowania skali Likerta w pomiarze jakości użytkowej serwisów internetowych | 59 |
| Małgorzata Sej-Kolasa, Mirosława Sztemberg-Lewandowska | |
| Funkcjonalna analiza głównych składowych w badaniu zmian liczby studentów w czasie w krajach europejskich | 71 |
| Iwona Olejnik, Robert Skikiewicz | |
| Metoda <i>k</i>-średnich w segmentacji emerytów na podstawie priorytetów życiowych | 83 |

Małgorzata Rószkiewicz

Instytut Statystyki i Demografii

Szkoła Główna Handlowa w Warszawie

Problemy analityczne metaanalizy – efekt procesu badawczego*

Streszczenie

Wraz z poszerzaniem się problematyki badań społecznych poszerzają się również obszary realizowanych badań empirycznych. W związku z licznymi badaniami często odnoszącymi się do tych samych zagadnień pojawia się pytanie, jak prowadzić analizę sumaryczną, pozwalającą ustalić, co w zasadzie wiadomo na temat przebadanych obszarów. Poszczególne badania są realizowane na próbach losowych, co sprawia, że uzyskiwane wyniki nie są jednakowe. Ich zakres zmienności opisuje wariancja losowa estymatora. W artykule rozważono problemy związane z analizą fluktuacji wyników badań, których rozstrzygnięcie wyznacza paradygmat statystycznej metaanalizy.

Słowa kluczowe: metaanaliza, efekt badania, testy wiązane, bezpieczna liczba badań.

1. Wprowadzenie

Wśród najczęściej przytaczanych powodów sięgania do metaanalizy wskazuje się wzrost liczby badań i ich wyników w ramach poszczególnych domen oraz oczekiwanie, że procedura ta spowoduje wzrost mocy statystycznej [Becker i Cohen 2003]. Obok opinii bardzo pozytywnych wyrażano jednak również poglądy sceptyczne, podważające walory poznawcze tego podejścia badawczego [Hedges i Olkin 1985]. W literaturze tematu brak jest jednej, powszechnie uznanej definicji metaanalizy. Na potrzeby niniejszego opracowania przyjęto, że metaanaliza w ujęciu

* Opracowanie przygotowano w ramach grantu NCN nr UMO-2011/01/B/HS4/00970.

statystycznym polega na wypracowaniu ogólnej konkluzji z wyników pochodzących z wielu badań. Kluczową jej częścią jest ocena spójności analizowanej serii wyników. Jeśli bowiem nie wiadomo, na ile są one spójne, to nie wiadomo, w jaki sposób dokonywać ich uogólnienia. Z tego względu ważne jest rozstrzygnięcie, czy odmienności obserwowane między poszczególnymi badaniami są powodowane efektami systematycznymi, czy też stanowią jedynie artefakty wynikające ze zmienności losowej poszczególnych badań. Celem artykułu jest omówienie ścieżek postępowania w diagnozowaniu efektu procesu badawczego oraz ocena ich skuteczności.

2. Metody oceny efektu procesu badawczego

Diagnozowanie homogeniczności wyników badań jest na ogół dokonywane za pomocą testu wykorzystującego statystykę zaproponowaną przez W.G. Cochran, opartą na rozkładzie chi-kwadrat [Higgins i Thompson 2002, Rószkiewicz 2009]. Test ten zwiększa swą moc, wykazując nadmierną skłonność do wskazywania na heterogeniczność wyników, gdy liczba badań jest znaczna, a szczególnie gdy badania są realizowane na dużych próbach. Moc ta jest zależna również od wariancji wewnątrzgrupowych wyników, czyli poziomów błędu średniokwadratowego każdego wyniku. Może się zatem zdarzyć, że mimo iż po przeprowadzeniu testu o homogeniczności serii wyników nie będzie podstaw do podważania ich spójności, wystąpi widoczna (znaczna) wartość ich wariancji międzygrupowej. Problem diagnozowania homogeniczności wyników z rozważanej serii badań jest zatem konsekwencją niedoskonałości procedur weryfikujących założenie o ich spójności.

Najstarszymi podejściami rekomendowanymi w ocenie istotności różnic między wynikami niezależnych badań są procedury ważne, takie jak t -test i z -test. Procedury te określa się mianem testów wiązanych (*combined tests*). Poniżej wybrano testy prezentowane przez R.A. Fishera [1932, 1948], S.A. Stouffera *et al.* [1949] oraz B.J. Winera [1971].

R.A. Fisher [1948] zaproponował procedurę wykorzystującą istotność krytyczną dla wyników poszczególnych badań w następującej postaci:

$$\chi^2 = -2 \sum_{i=1}^k \ln p_i,$$

gdzie:

p_i – istotność krytyczna dla wyniku i -tego badania,
 $i = 1, 2, \dots, k,$

w której statystyka testująca posiada asymptotyczny rozkład chi-kwadrat o liczbie stopni swobody $\nu = 2k$. Procedura ta jest przedstawiana jako asymptotycznie optymalna w porównaniu do innych metod wiązanych [Kozil i Perlman 1978, Littell i Folks 1973]. Jest z nią związanych jednak wiele ograniczeń [Rosenthal 1984]. Może ona dostarczyć wyników niezgodnych z prostą procedurą testu istotności w sytuacji, gdy większość wyników badań będzie wskazywać na rezultaty zbieżne co do kierunku, ale nieistotne. Z kolei S.A. Stouffer zaproponował, by dla łączenia niezależnych testów realizowanych z uwzględnieniem statystyki t -Studenta sumować wartości statystyki testującej t i przekształcać do statystyki Z według wzoru:

$$Z = \frac{\sum t_i}{\sqrt{\sum n_i}},$$

gdzie:

t_i – wynik testu istotności w badaniu i ,

n_i – rozmiar próby w badaniu i .

Procedura ta opiera się na własności addytywności rozkładu normalnego z wariancją równą liczbie obserwowanych wyników prób. Z racji zbieżności rozkładu t -Studenta z rozkładem normalnym procedura ta nie jest odpowiednia dla testów pochodzących z małych próbek. Niemniej procedura Stouffera jest łatwiejsza od procedury Fishera, w której konieczne są przekształcenia logarytmiczne. Procedura zaproponowana przez B.J. Winera jest identyczna z procedurą Stouffera i przyjmuje postać:

$$Z = \frac{\sum t_i}{\sqrt{\sum \frac{df_i}{df_i - 2}}},$$

gdzie df_i oznacza liczbę stopni swobody statystyki testującej wyniki badania i .

Oba podejścia są zbieżne wraz ze wzrostem liczby próbek.

Charakterystyczną cechą najstarszych rozwiązań rekomendowanych do oceny efektu procesu badawczego jest traktowanie równorzędnie wyników wszystkich prób niezależnie od ich rozmiarów. Próby mniej liczne z oczywistych względów mogą nie spełniać wymogu reprezentatywności lub spełniać go w ograniczonym zakresie w porównaniu do prób o znacznych rozmiarach. W takich sytuacjach nadanie wszystkim wynikom jednakowych wag nie wydaje się poprawne, gdyż może prowadzić do nadawania nadmiernego znaczenia wynikom o niskiej reprezentatywności. By przełamać ograniczenia wynikające z różnej reprezentatywności prób o różnych rozmiarach F.M. Mosteller i B.R. Bush [1954] zaproponowali

nadanie wag odpowiadających odchyleniom standardowym wynikom agregowanym według procedury Stouffera, czyli:

$$Z = \frac{\sum df_i S_i}{\sqrt{\sum df_i^2}},$$

gdzie S_i jest odchyleniem standardowym wyników w badaniu i .

Zupełnie odmienne podejście do oceny homogeniczności serii wyników badań zaproponowali L.V. Hedges i I. Olkin [1985]. Wykazali oni, że testująca homogeniczność statystyka Q -Cochrana może być podzielona na dwa składniki, analogicznie do dekompozycji odchyłeń od średniej, w przypadku danych klasyfikowanych ze względu na wybraną cechę, tj. na składnik między- i wewnątrzgrupowy, gdzie grupy są definiowane przez wartości lub kategorie cech potencjalnie wpływających na wyniki badań. W myśl tego rozwiązania jeśli:

$$Q = \sum_{i=1}^k \frac{(T_i - \bar{T})^2}{v_i},$$

to:

$$(T_{ij} - \bar{T}) = (T_{ij} - \bar{T}_j) + (\bar{T}_j - \bar{T}),$$

gdzie:

T_{ij} – wynik badania i ze względu na j -tą kategorię lub wartość cechy charakteryzującej proces badawczy,

\bar{T} – wartość uogólniona wszystkich wyników badań,

\bar{T}_j – wartość uogólniona z tych wyników badań, które odnoszą się do j -tej wartości lub kategorii cechy charakteryzującej proces badawczy.

Zaproponowana dekompozycja wartości statystyki testującej pozwala zidentyfikować, które cechy badań potencjalnie wpływają na ich wyniki. Każdy z wyodrębnionych w wyniku dekompozycji składników może być testowany oddzielnie. Jeśli cecha badań, ze względu na którą dokonano dekompozycji, byłaby istotnym czynnikiem zakłócającym homogeniczność wyników, to zarówno wynik testu Q -Cochrana powinien wskazywać na niejednorodność analizowanej serii wyników badań, jak i istotny powinien okazać się składnik międzygrupowy statystyki Q . Składnik wewnątrzgrupowy statystyki Q dotyczący zmienności losowej poszczególnych wyników prób powinien być nieistotny. Gdyby składnik wewnątrzgrupowy okazał się istotny, przeprowadzoną procedurę dla wyróżnionego czynnika zakłócającego należałoby uznać za niewystarczającą. Oznaczać to może, że występują jeszcze inne, nieuwzględnione czynniki zakłócające porów-

nywalność wyników. W takiej sytuacji należałoby rozważyć włączenie do analizy dodatkowych zmiennych opisujących potencjalne różnice między badaniami.

Przedstawiona dekompozycja zakłada, że czynniki zakłócające mogą być opisywane przez zmienne nominalne lub dyskretne. Jeśli czynnikami zakłócającymi wyniki badań i odnoszącymi się do procesu badawczego są zmienne o charakterze ciągłym, to ich oddziaływanie może być opisane modelem regresji w ujęciu modelu mieszanego [Konstantopoulos 2006].

Modele mieszane wykorzystywane w metaanalizie należą do grupy ogólnych liniowych modeli mieszanych. Jeśli założyć, że efekty procesu badawczego są opisane za pomocą p predyktorów X_1, X_2, \dots, X_p , które są powiązane liniowo z wynikiem badania, to przykładowy model drugiego poziomu dla wyniku pochodzącego z i -tego badania przyjmuje postać:

$$T_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \eta_i + \varepsilon_i,$$

gdzie:

x_{ij} – poziom j -tego predyktora opisującego działanie zmiennej zakłócającej w i -tym badaniu,

ε_i – składnik losowy próby w badaniu i ,

η_i – efekt losowy specyficzny dla i -tego badania, o rozkładzie $N(0, \tau^2)$.

Jeśli wariancja międzygrupowa wyników badań (τ^2) byłaby znana, to współczynniki regresji mogłyby być szacowane ważoną metodą najmniejszych kwadratów, a ich istotność rozstrzygałaby o efektach systematycznych związanych ze specyfiką poszczególnych badań. Na ogół jednak τ^2 nie jest znane – wówczas możliwe są cztery drogi postępowania:

1) oszacowanie τ^2 na podstawie dostępnych wyników badań i wykorzystanie oceny tej wariancji do estymacji współczynników regresji ważoną metodą najmniejszych kwadratów, co umożliwi weryfikację istotności efektów systematycznych poszczególnych badań,

2) łączne oszacowanie wektora współczynników i wariancji τ^2 za pomocą nierestykcyjnej metody największej wiarygodności (REML),

3) przyjęcie, że wariancja międzygrupowa wyników badań wynosi zero (tak jak to ma miejsce w analizie efektów ustalonych), co – jak już wspomniano – może powodować obciążenie wyników, ale ponieważ jest ono tym mniejsze, im liczniejszy jest zbiór dostępnych badań, rozwiązanie takie można rekomendować w przypadku, gdy zbiór wyników badań jest dostatecznie liczny,

4) oszacowanie wektora współczynników regresji dla każdej wartości τ^2 z przedziału prawdopodobnych wartości i użycie średniej ważonej dla tych wyników, przyjmując za wagi wartości odpowiednie do prawdopodobieństw, tak jak ma to miejsce w podejściu bayesowskim.

Wprowadzając metodologię ogólnych modeli mieszanych do procedur metaanalizy, należy odnotować dwie podstawowe różnice między ogólnymi modelami mieszanymi a modelami wykorzystywanymi w metaanalizie. Po pierwsze, w metaanalizie wariancje błędu próby ε_i , co do którego zakłada się, że ma rozkład normalny z wartością oczekiwaną 0 i wariancją v_i , nie są identyczne w poszczególnych badaniach. W związku z tym założenie, że są sobie równe, jest nierealne. Jak wiadomo, wariancja błędu próby zależy od wielu czynników wynikających z konstrukcji próby, tj. rozmiaru próby i metody losowania. Po drugie, wariancje błędu poszczególnych badań w metaanalizie nie są identyczne, ale są znane. Z tego względu model wykorzystywany w metaanalizie rozważa specjalny przypadek ogólnego liniowego modelu hierarchicznego, w którym wariancje pierwszego poziomu są różne, lecz znane. Szacowanie wariancji międzygrupowej przy ustalonych cechach badań opisanych zmiennymi X_i przebiega wówczas według formuły:

$$\tau^2 = \frac{\chi^2 - k + p}{c},$$

$$c = \text{tr}(\mathbf{V}^{-1}) - \text{tr}\left[\left(\mathbf{X}^T \mathbf{V}^{-1}\right)^{-1} \mathbf{X}^T \mathbf{V}^{-2} \mathbf{X}\right],$$

$$V = \text{diag}(v_1, v_2, \dots, v_k).$$

Wnioskowanie o istotności wariancji międzygrupowej przy ustalonych cechach badań opisanych zmiennymi X_i opiera się na wartości statystyki χ^2 o rozkładzie chi-kwadrat z liczbą stopni swobody $(k - p)$, którą definiuje formuła:

$$\chi^2 = \mathbf{T}^T \left[\mathbf{V}^{-1} - \mathbf{V}^{-1} \mathbf{X} \left(\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^T \mathbf{V}^{-1} \right] \mathbf{T}.$$

Dokonując metaanalizy, należy pamiętać, że wiele sytuacji badawczych może prowadzić do obciążonych lub błędnych wyników. Należą do nich:

- wykorzystanie tylko pozytywnych wyników, a pominięcie wyników negatywnych,
- traktowanie równoważnie wyników wszystkich badań, które odnoszą się do tych samych pytań badawczych, nawet wówczas, gdy występują między nimi różnice jakościowe,
- wielokrotne wykorzystanie wyników tego samego badania,
- brak gwarancji wysokiego stopnia zgodności lub rzetelności między rangami przy kodowaniu charakterystyk badań, kiedy są podstawowe lub metodologiczne.

Można wskazać kilka sposobów przełamania tych trudności. Problem pomiaru w publikacjach wyników, które okazały się nieistotne, i występowanie nadreprezentatywnych wyników pozytywnych w literaturze z danej domeny badań R. Rosenthal [1984] nazwał problemem doboru badań, proponując określenie,

jaka liczba badań jest niezbędna, by potwierdzić hipotezę zerową o nieistotności wyniku uogólnienia. H. Cooper [1979] określił tę liczebność mianem bezpiecznej liczby badań z liczby N (N_{fs} – *fail-safe N*). Jest to dodatkowa liczba badań, niezbędna by przekształcić całkowitą istotność statystyki testu połączonych prób w wartość wyższą niż krytyczna istotność – na ogół 0,05 lub 0,01. Dla $p = 0,05$ jest to:

$$N_{fs,0,05} = \left(\frac{\sum Z}{1,645} \right)^2 - N,$$

zaś dla $p = 0,01$ jest to:

$$N_{fs,0,01} = \left(\frac{\sum Z}{2,33} \right)^2 - N,$$

gdzie:

$\sum Z$ – suma wartości statystyki Z dla prób łączonych według procedury istotności Stouffera,

N – liczba prób łączonych.

Tak wyznaczone liczebności określają zatem, ile dodatkowych badań – każde o wyniku nieistotnym lub w sumie dające $\sum Z = 0$ – powinno być włączonych, by nie było podstaw do odrzucenia hipotezy zerowej o nieistotności wyniku.

3. Podsumowanie

Warto zwrócić uwagę na korzyści analityczne, jakich dostarcza metaanaliza, a które można uznać za szczególne, gdyż są nieosiągalne w incydentalnych badaniach przekrojowych. Po pierwsze, metaanaliza stwarza możliwość wykorzystania pełnego zakresu informacji wynikowych ze zrealizowanych badań w ramach danej dziedziny, niezależnie od tego, czy wyniki te okazały się istotne statystycznie, czy też nie. Po drugie, umożliwia prowadzenia bardziej szczegółowych podziałów i klasyfikacji poprzez integrację cząstkowych zbiorów danych w jeden zbiór o znacznym rozmiarze. Ponadto wskazane ścieżki postępowania umożliwiające identyfikację czynników zakłócających procesy badawcze i powodujące ich niejednorodność dają szansę modyfikacji postępowania w przyszłych badaniach. Należy podkreślić, że ocena kontekstu badawczego jest możliwa jedynie w metaanalizie, a jej znaczenie dla właściwej interpretacji wyników poszczególnych badań jest nie do przecenienia.

Literatura

- Becker B.J., Cohen L.D. [2003], *How Meta-analysis Increases Statistical Power*, „Psychological Methods”, vol. 8, nr 3.
- Cooper H. [1979], *Statistically Combining Independent Studies: A Meta-analysis of Sex Differences in Conformity Research*, „Journal of Personality and Social Psychology”, vol. 37, nr 1.
- Fisher R.A. [1932], *Statistical Method for Research Workers*, wyd. 4, Oliver and Boyd, London.
- Fisher R.A. [1948], *Combining Independent Tests of Significance*, „American Statistician”, vol. 2, nr 5.
- Hedges L.V., Olkin I. [1985], *Statistical Methods for Meta-analysis*, Academic Press, New York.
- Higgins J.P.T., Thompson S.G. [2002], *Quantifying Heterogeneity in a Meta-analysis*, „Statistics in Medicine”, vol. 21, nr 11.
- Konstantopoulos S. [2006], *Fixed and Mixed Effects Models in Meta-analysis*, Discussion Paper Series, IZA DP, nr 2198, IZA, Bonn.
- Kozil J.A., Perlman M.D. [1978], *Combining Independent Chi-squared Tests*, „Journal of American Statistical Association”, nr 73.
- Littell R.C., Folks J.L. [1973], *Asymptotic Optimality of Fisher's Method of Combining Independent Tests II*, „Journal of American Statistical Association”, nr 68.
- Mosteller F.M., Bush R.R. [1954], *Selected Quantitative Techniques [w:] Handbook of Social Psychology*, red. G. Lindzey, vol. 1, Addison-Wesley, Cambridge.
- Rosenthal R. [1984], *Meta-analysis Procedure for Social Research*, Sage, Beverly Hills.
- Rószkiewicz M. [2009], *Możliwości i ograniczenia metaanalizy w obszarze ilościowych badań marketingowych*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 96, Wrocław.
- Stouffer S.A. et al. [1949], *The American Soldier: Adjustment during Army Life*, vol. 1, Princeton University Press, Princeton.
- Winer B.J. [1971], *Statistical Principles in Experimental Design*, wyd. 2, McGraw-Hill, New York.

Analytical Problems of Meta-analysis – the Effect of Heterogeneous Studies

The growing number of issues in social science research has led to a correspondingly large body of research studies. The sheer volume of research related to the same topics poses a question as to how to organise and summarise these findings to identify and use what is known and focus research on promising areas. Effect size estimates from the studies are not identical. This is to be expected because the estimates are based on data from samples, and random variations due to sampling should introduce fluctuations into estimates. This paper introduces methods for exploring these fluctuations as a paradigm for explorations in meta-analyses generally.

Keywords: meta-analysis, survey effect, combined tests, fail-safe N .

Aneta Rybicka

Katedra Ekonometrii i Informatyki
Uniwersytet Ekonomiczny we Wrocławiu

Od conjoint analysis do metod wyborów opartych na menu

Streszczenie

W artykule zaprezentowano metody wykorzystywane w badaniach preferencji ujawnionych – począwszy od *conjoint analysis* do metody wyborów opartej na menu.

Przeprowadzając badania z wykorzystaniem *conjoint analysis*, respondentom przedstawia się do oceny zbiór profilów. Respondenci są proszeni o porządkowanie bądź rangowanie zaprezentowanych profilów. Wykorzystując metody wyborów dyskretnych, badacz przedstawia respondentowi kilka zbiorów profilów, a ten wybiera jeden profil z każdego zbioru lub rezygnuje z wyboru. Pomiar preferencji z wykorzystaniem adaptacyjnej *conjoint analysis* odbywa się etapami: respondent określa preferencje odnośnie do poziomów każdego atrybutu, a następnie ocenia wyselekcjonowane pary i wybiera z każdej jeden profil (wyrażając jednocześnie swoje preferencje). By uzyskać dane bardziej odpowiadające zachowaniom rynkowym, zaproponowano adaptacyjną metodę wyborów dyskretnych, która równie dobrze odzwierciedla zachowania kompensacyjne, jak i niekompensacyjne oraz angażuje respondenta w badanie. Najnowszą metodą jest metoda wyborów oparta na menu, zaproponowana z myślą o rynkach produktów lub usług, na których wybór oparty jest na menu.

Słowa kluczowe: badanie preferencji ujawnionych, *conjoint analysis*, metoda wyborów dyskretnych, adaptacyjna *conjoint analysis*, adaptacyjna metoda wyborów dyskretnych, metoda wyborów opartych na menu.

1. Wprowadzenie

Badania preferencji konsumentów można podzielić na dwie grupy: badania preferencji ujawnionych (oparte na danych historycznych) oraz badania preferencji wyrażonych (wykorzystujące deklaracje, intencje respondentów).

W badaniach preferencji wyrażonych zastosowanie znajdują metody reprezentujące podejścia kompozycyjne, dekompozycyjne oraz mieszane. Ze względu na lepsze odzwierciedlenie rzeczywistych zachowań konsumentów na rynku najczęściej wykorzystuje się metody dekompozycyjne (*conjoint analysis*, metody wyborów dyskretnych, metody wyborów opartych na menu) oraz metody mieszane (adaptacyjna *conjoint analysis*, adaptacyjna metoda wyborów dyskretnych).

Celem niniejszego artykułu jest zaprezentowanie rozwoju dekompozycyjnych i mieszanych metod badania preferencji, począwszy od *conjoint analysis*, do metod wyborów opartych na menu.

2. Historia *conjoint analysis**

Na początku lat 70. XX w. profesor marketingu P.E. Green zauważył, że artykuł R.D. Luce'a i J.W. Turkeya [1964] opublikowany w czasopiśmie niemarketingowym może być wykorzystany w dziedzinie marketingu. Może pomóc zwłaszcza w zrozumieniu, jak konsumenci dokonują kompleksowych zakupów, a także w oszacowaniu preferencji konsumentów oraz ważności poszczególnych atrybutów produktów lub usług. Także artykuł P.E. Greena i V.R. Rao [1971] miał duży wpływ na *conjoint analysis* z wykorzystaniem metody pełnych profilów. Za pomocą tej metody respondenci oceniali zaprezentowane profile i porządkowali je od najlepszego do najgorszego. Metoda ta sprawdzała się w badaniach, w których liczba profilów w cząstkowym układzie czynnikowym nie była zbyt duża. Zauważono, że lepsze dane można otrzymać, zwracając się do respondenta z prośbą, by ocenił zaprezentowane profile np. w skali 10-punktowej, wykorzystując do estymacji metodę najmniejszych kwadratów. W 1975 r. P.E. Green i Y. Wind opublikowali artykuł na temat badania preferencji z wykorzystaniem *conjoint analysis*, w którym respondenci rangowali profile [Green i Wind 1975]. Metoda ta została szybko zauważona i doceniona zarówno przez badaczy, jak i dyrektorów firm.

Również na początku lat 70. XX w. praktyk R.M. Johnson szukał niezależnie rozwiązania problemu wyboru poszczególnych produktów przez konsumentów. Planował badanie preferencji konsumentów w zakresie dóbr trwałych opisanych

* Punkt ten opracowano na podstawie [Orme 2009].

28 atrybutami i na ogół 5 poziomami każdego atrybutu. Problem był bardziej skomplikowany niż ten przedstawiony i rozwiązany przez P.E. Greena i współautorów, w związku z czym Johnson zaproponował inne rozwiązanie: metodę porównywania atrybutów parami [Johnson 1974]. Zamiast zwracać się do respondenta z prośbą o ocenę wszystkich atrybutów w tym samym czasie (jak w metodzie pełnych profilów), poproszono respondentów o ocenę profilów opisanych tylko dwoma atrybutami. Tę metodę prezentacji profilów respondentowi nazwano macierzą kompromisów (metodą porównywania atrybutów parami).

Pod koniec lat 70. XX w. P.E. Green i V. Srinivasan opublikowali artykuł podsumowujący wykorzystanie metody *conjoint analysis* w przemyśle, przedstawiający perspektywę jej rozwoju oraz zawierający wskazówki na temat lepszych jej zastosowań [Green i Srinivasan 1978].

Pierwsze komercyjne zastosowanie (przez agencję badań marketingowych) *conjoint analysis* w ramach projektu badawczego miało miejsce w 1971 r. [Strojny 2003].

Na początku lat 80. XX w. *conjoint analysis* zyskała popularność zwłaszcza wśród badaczy i naukowców posiadających wiedzę statystyczną oraz umiejętności programowania komputerowego. Gdy w 1985 r. komercyjne oprogramowanie stało się dostępne, możliwości zastosowań tej metody w badaniach preferencji znacznie się zwiększyły. Na podstawie pracy P.E. Greena na temat *conjoint analysis* wykorzystującej metodę pełnych profilów S. Herman wraz z firmą Bretton-Clark Software opracował system dla komputerów osobistych.

Również w 1985 r. R.M. Johnson i jego nowa firma Sawtooth Software (założona w 1983 r.) opracowali system dla komputerów osobistych – Adaptive Conjoint Analysis (ACA).

Po wielu latach pracy i badań z wykorzystaniem macierzy kompromisów (macierz porównywania atrybutów parami) R.M. Johnson zauważył, że respondenci mają problem z liczbą macierzy do oceny. Z tego względu opracowano program komputerowy tak, by administrował gromadzone dane oraz by w czasie rzeczywistym adaptował indywidualne odpowiedzi respondentów. Chodziło o to, by w kolejnym etapie pytać respondentów już tylko o najważniejsze atrybuty.

Pomiar preferencji z wykorzystaniem adaptacyjnej *conjoint analysis* odbywa się w trzech etapach:

- 1) respondent określa preferencje odnośnie do poszczególnych poziomów każdego atrybutu;

- 2) na podstawie uzyskanych danych o preferencjach program komputerowy przedstawia respondentom ekstremalne oceny poziomów każdego atrybutu w celu oszacowania istotności różnic pomiędzy tymi skrajnymi wartościami [Walesiak i Bąk 2000, s. 90]. Szacowane są również wagi każdego atrybutu. Dzięki dwóm pierwszym etapom (odpowiadającym etapom w metodzie danych samowjaśnia-

jących) możliwe jest pierwsze oszacowanie użyteczności cząstkowych dla każdego z respondentów;

3) na podstawie wcześniej oszacowanych użyteczności cząstkowych program komputerowy wybiera pary profilów. Następnie respondent ocenia te pary i wybiera z każdej jeden profil, wyrażając jednocześnie swoje preferencje. Oceny te są na bieżąco uwzględniane w selekcji następnej pary profilów przez aktualizację wyjściowych użyteczności cząstkowych [Bąk 2000, s. 220]. Dzięki temu kolejne prezentowane pary profilów dobierane są już na podstawie pełniejszych informacji.

Istotą tej metody jest dynamiczny charakter selekcji par profilów.

Na początku lat 70. XX w. D.L. McFadden pracował nad modelowaniem wyboru spośród wielu opcji (profilów). Pracą dotyczącą idei analizy wyboru spośród dostępnych opcji (profilów) oraz wielomianowego modelu logitowego zajmował się również J. Louviere wraz ze współpracownikami. Koncepcja analizy wyboru wydawała się bardzo interesująca, ponieważ zgodnie z nią respondenci nie porządkują ani nie rangują profilów produktów lub usług, tylko oceniają zbiór takich profilów i dokonują wyboru. Stanowiło to więc odzwierciedlenie rzeczywistych zachowań konsumentów na rynku. Ze statystycznego oraz teoretycznego punktu widzenia analiza wyboru wydawała się więc lepszym rozwiązaniem w badaniach preferencji konsumentów, lecz w praktyce stanowiła pewne wyzwanie.

W latach 90. XX w. ACA była najczęściej stosowaną techniką *conjoint*, lecz już pod koniec dekady ustąpiła miejsca analizie wyborów dyskretnych (metodom wyborów dyskretnych). Było to możliwe dzięki powstałemu w 1993 r. oprogramowaniu pozwalającemu na przeprowadzenie badania z wykorzystaniem metod wyborów dyskretnych firmy Sawtooth Software oraz dzięki zastosowaniu modeli hierarchicznych Bayesa w metodach wyborów dyskretnych (w celu oszacowania parametrów na poziomie indywidualnym).

Równie istotnymi czynnikami w tym czasie były: możliwość wykorzystania modeli klas ukrytych do oszacowania parametrów na poziomie segmentowym (zarówno w *conjoint analysis*, jak i w metodach wyborów dyskretnych), wzrost szybkości komputerów (co miało znaczący wpływ na czas, w jakim przeprowadzano analizę preferencji z wykorzystaniem modeli hierarchicznych Bayesa) oraz opracowanie przez W.F. Kuhfelda, R.D. Tobiasa i M. Garratta procedur pozwalających na zaprojektowanie eksperymentów wyboru i ich analizę (*Statistical Analysis System – SAS*). Lata 90. XX w. były w związku z tym dekadą wzmożonego rozwoju *conjoint analysis* i technik pokrewnych oraz okresem wzrostu zastosowań tychże metod w wielu obszarach badawczych.

Conjoint analysis jest techniką badawczą, która ma już ponad 40 lat, nie można jednak powiedzieć, że jest techniką starą, która przechodzi stagnację.

Dzięki rozwojowi technologii komputerowych, metodologii oraz intelektualnym wysiłkom poszczególnych badaczy rozwija się bowiem nadal.

3. Historia metod wyborów dyskretnych*

Metoda wyborów dyskretnych¹ została wprowadzona do badań marketingowych w 1983 r. przez J.J. Louviere'a i G. Woodwortha [Krapp i Sattler 2001, s. 2].

Pomimo niewątpliwych zalet jak każda inna metoda także i ta ma swoje wady. W badaniach preferencji z wykorzystaniem metody wyborów dyskretnych:

- użyteczności częściowe szacowane są na poziomie zagregowanym,
- respondenci zbyt szybko dokonują wyborów (12–15 sekund przeznaczają na wybór profilu w jednym zbiorze),
- dokonywanie kolejnych wyborów może wydawać się respondentowi nudne i dlatego nie angażuje się on w wywiad – wybór w każdym kolejnym zbiorze jest dokonywany coraz szybciej, a więc wynik jest mniej wiarygodny,
- dokonywanie wyborów nie odzwierciedla niekompensacyjnego charakteru zachowań respondentów (85% respondentów skupia się na obecności lub nieobecności co najwyżej 4 poziomów atrybutów),
- istnieje ograniczona liczba atrybutów i ich poziomów.

By szacowanie parametrów było możliwe na poziomie segmentowym, zaproponowano wykorzystanie modeli klas ukrytych [DeSarbo i Ramaswamy 1995]. Z pomocą W.S. DeSarbo R.M. Johnson zaprojektował system służący do szacowania parametrów na poziomie segmentowym – *Latent Class*. Następnie, by móc oszacować parametry na poziomie indywidualnym, opracowano system *Individual Choice Estimation* (ICE). Dzięki obu programom otrzymywano dobre, wiarygodne wyniki, jednakże pod koniec lat 90. XX w. zaczęto interesować się powiązaniem *conjoint analysis* i metod wyborów dyskretnych z modelami hierarchicznymi Bayesa. Dzięki pomocy G.M. Allenby'ego R.M. Johnson opracował system CBC/HB pozwalający oszacować parametry na poziomie indywidualnym. Tym samym udoskonalono metodę wyborów dyskretnych. Można było teraz prowadzić badania preferencji konsumentów zarówno na poziomie zagregowanym, jak i indywidualnym. Dzięki temu około 2000 r. liczba zastosowań metod wyborów dyskretnych przewyższyła liczbę zastosowań adaptacyjnej metody *conjoint analysis*.

* Punkt ten opracowano m.in. na podstawie [Johnson 2011].

¹ W 2000 r. J. Heckman i D.L. McFadden otrzymali nagrodę im. Alfreda Nobla z dziedziny ekonomii – J. Heckman za rozwój teorii i metod analizy prób selektywnych, D.L. McFadden za rozwój teorii i metod analizy wyboru dyskretnego.

Po 2000 r. analitycy zauważyli, że decyzje respondentów stanowią realizację zarówno modeli kompensacyjnych, jak i niekompensacyjnych, natomiast po 2003 r. rozpoczęły się badania nad poprawą efektywności eksperymentu czynnikowego. Oba te aspekty miały poprawić jakość badania oraz otrzymywanych danych. Na podstawie obserwacji stwierdzono, że respondenci bardzo szybko dokonują wyboru z każdego prezentowanego zbioru. Wynioskowano zatem, że przy podejmowaniu decyzji respondenci biorą pod uwagę obecność lub nieobecność co najwyżej 3 poziomów atrybutów [Gilbride i Allenby 2004, s. 391–406; Hauser *et al.* 2006, s. 169–181].

W 2009 r. opracowano system *Adaptive Choice-based Conjoint Analysis* (ACBC). Możliwe stało się przeprowadzanie badań z wykorzystaniem adaptacyjnej metody wyborów dyskretnych, pozwalającej skupić się na profilach (poziomach i atrybutach) rzeczywiście istotnych dla respondenta.

Metoda ta:

- wykorzystuje metodę pełnych profili,
- angażuje respondenta w badanie,
- równie dobrze odzwierciedla zachowania kompensacyjne jak niekompensacyjne.

Pierwsze badania porównawcze z tradycyjną metodą wyborów dyskretnych pozwalają stwierdzić, że wykorzystując w badaniach adaptacyjną jej formę, otrzymuje się dane dokładniej prognozujące wybory konsumentów [ACBC *Technical Paper...* 2009, s. 2]. Metoda ta dostarcza również więcej informacji na poziomie indywidualnym niż metoda wyborów dyskretnych, a badanie może być przeprowadzone na małej próbie z wykorzystaniem 5 lub więcej atrybutów.

W trakcie wywiadu z wykorzystaniem adaptacyjnej metody wyborów dyskretnych respondent nie jest zmuszany do wykonywania szeregu monotonicznych i trudnych zadań wyboru profilu – zamiast tego jest włączany w interaktywny proces tworzenia profili.

Ankieta składa się z następujących sekcji [ACBC *Technical Paper...* 2009, s. 3–8]: konfigurator, rozważane oferty, zadania wyboru, kalibracja (opcjonalnie) (por. [Adaptive Choice-based... 2003, *Optymalizacja taryf...* 2012]).

Konfigurator (BYO – *build your own configurator*) pozwala na zdefiniowanie przez respondenta optymalnego profilu. Stworzony w ten sposób profil stanowi punkt wyjścia do określenia kształtu prezentowanych w kolejnym kroku profili.

Rozważane oferty (*screening section*) to ocena zbiorów profili nieznacznie różniących się od początkowo zdefiniowanego profilu idealnego. Etap ten pozwala na określenie głównych reguł, którymi kieruje się respondent (odpowiedzi respondenta są na bieżąco analizowane, co pozwala na wychwycenie obowiązkowych poziomów atrybutów oraz poziomów nieakceptowanych).

W trakcie badania stale zadawane są pytania o to, czy któryś z prezentowanych poziomów atrybutów jest absolutnie nie do przyjęcia lub czy któryś z prezentowanych poziomów atrybutów jest absolutnie wymagany. Zadawane są pytania nie o wszystkie poziomy atrybutów, lecz o najczęściej unikane i najczęściej wybierane.

Zadania wyboru polegają na tym, że spośród rozważanych profilów (z poprzedniej sekcji) w ramach kilku kroków dokonywany jest wybór najlepszego profilu. W tym etapie respondentowi prezentowane są zbiory profilów (zazwyczaj od 7 do 9 zbiorów zawierających od 3 do 5 profilów generowanych na podstawie odpowiedzi uzyskanych w poprzednich sekcjach). Każdy wybrany profil ze zbioru „powraca”, aż najlepszy profil zostanie wskazany.

Kalibracja pozwala na urealnienie analiz rynkowych prowadzonych w ramach symulatora – możliwe jest wskazanie takiego punktu na skali intencji zakupowych, który jest najbliższy zachowaniom rynkowym w obrębie danej kategorii produktowej (stanowi najlepsze odzwierciedlenie zachowań respondentów).

W ramach części kalibracyjnej badana jest skłonność do zakupu testowanych profilów. Oceniane są:

- profil samodzielnie skonfigurowany na początku wywiadu,
- profil wyłoniony jako najlepszy w ramach zadań wyboru,
- 4 profile obejmujące zarówno profile bliskie, jak i dalekie od oczekiwań respondenta.

Podsumowując, można stwierdzić, że kwestionariusz ACBC jest ciekawszy od wcześniej omawianych, proces podejmowania decyzji nie jest nużący, a respondent angażuje się w badanie i dokonuje bardziej przemyślanych wyborów. System ACBC zawiera wbudowaną estymację hierarchiczną Bayesa (szacowanie parametrów na poziomie indywidualnym), pozwala także na eksportowanie danych do systemu klas ukrytych (szacowanie parametrów na poziomie segmentowym). Wszystkie te zalety sprawiają, że z roku na rok adaptacyjna metoda wyborów dyskretnych jest coraz popularniejsza, a co za tym idzie – liczba jej zastosowań w badaniach preferencji rośnie.

„Najmłodszą” metodą *conjoint* jest metoda wyborów oparta na menu (*Menu Choice-based* – MCB) [Orme 2010a i 2010b]. Nazwa nawiązuje od menu w restauracji, gdzie oprócz gotowych zestawów na śniadanie, obiad czy kolację można samemu dowolnie skomponować menu, korzystając z dostępnych potraw.

Wiele produktów i usług jest pomyślanych tak, by kupujący miał wybór spośród różnych opcji przy ich konfiguracji (dopasowywaniu) i samodzielnie mógł stworzyć preferowany produkt. Przykłady tego typu działań obejmują m.in. menu w restauracji, usługi finansowe, opcje przy ubezpieczeniach, zakup samochodu, wybór telefonu czy dostawy Internetu/kabłówki.

Metoda ta jest zatem wskazana w analizowaniu mieszanych koszyków, gdzie respondenci mogą wybrać gotową ofertę (produkt czy usługę wraz ze zdefinio-

wanymi dodatkami) lub dokonać samodzielnego wyboru wszystkich (lub prawie wszystkich) elementów.

W metodzie wyborów opartej na menu na dole ekranu znajduje się cena bazowa profilu (np. konkretnego modelu samochodu). Respondent wybiera poszczególne opcje zgodnie ze swoimi potrzebami (preferencjami) – z każdą wybraną opcją cena zwiększa się. Gdy respondent zakończy konfigurację wybiera przycisk „Dalej” i przechodzi do kolejnego pytania.

Ceny, które przedstawiane są w pytaniach respondentom, mogą być takie same dla wszystkich pytań i wszystkich respondentów, ale mogą także różnić się pomiędzy respondentami, a nawet pomiędzy poszczególnymi pytaniami do tego samego respondenta. Zmiany cen w ramach pytań do tego samego respondenta pozwalają zaobserwować ewentualny wpływ tego czynnika na dokonywane wybory. Różne ceny poszczególnych opcji pozwalają ponadto na obliczenie elastyczności cenowej popytu. Metoda ta umożliwi także szacowanie efektów substytucyjnych. Można obserwować, czy respondenci wybierają gotowy, skonfigurowany produkt, czy wolą sami wybierać poszczególne jego elementy.

4. Podsumowanie

Początkowo dekompozycyjne metody badania preferencji (zarówno *conjoint analysis*, jak i metody wyborów dyskretnych) pozwalały jedynie na analizowanie preferencji, ocenianie ważności atrybutów czy prognozowanie i symulowanie udziałów w rynku dla nowych profilów. Za największą z wad w początkowym okresie rozwoju tych metod można uznać fakt, że zakładały one, iż większość konsumentów ma podobne preferencje i zachowania na rynku.

Wraz ze wzrostem świadomości konsumentkiej pojawiła się potrzeba personalizacji („dopasowania”) produktów oraz usług do potrzeb indywidualnego klienta. W związku z tym pojawiły się metody adaptacyjne ACA oraz ACBC. Pozwalają one nie tylko na przezwyciężenie ograniczeń metod *conjoint analysis* i metod wyborów dyskretnych, ale także na podstawie odpowiedzi osoby biorącej udział w badaniu umożliwiają zaobserwowanie i zbadanie jej zachowań i preferencji. Dzięki temu możliwe jest dokładniejsze przeanalizowanie potrzeb tego konsumenta – lepsza estymacja jego preferencji i przewidywanie przyszłych zachowań na rynku. Wadą podejść adaptacyjnych jest natomiast konieczność wykorzystania oprogramowania komputerowego. Często towarzyszą temu bardzo skomplikowane procedury badawcze.

Ostatnim etapem rozwoju dekompozycyjnych metod badania preferencji, który pozwala na jeszcze lepsze analizowanie zachowania konsumenta i personalizację (dopasowanie) produktów do jego indywidualnych potrzeb, są metody wyborów

oparte na menu. Tu – niczym w restauracji – respondent dokonuje swoich wyborów, konstruuje idealny produkt (wybierając spośród dostępnych opcji), a kolejne pytania pozwalają zidentyfikować cechy konieczne, jak również niepożądane w opinii respondenta. Takie podejście pozwala dopasować produkt do indywidualnych potrzeb konsumenta – zaoferować mu dokładnie to, czego oczekuje. Należy jednak zaznaczyć, że metody oparte na menu podobnie jak metody adaptacyjne wymagają użycia oprogramowania komputerowego i charakteryzują się złożoną procedurą badania.

Analizując 40-letni rozwój dekompozycyjnych i mieszanych metod badania preferencji – od *conjoint analysis* do metod wyboru opartych na menu – nie można powiedzieć, że metody starsze (tj. *conjoint analysis* oraz metody wyborów dyskretnych) nie mają już zastosowania w badaniach marketingowych. Wprost przeciwnie, metody te są dobrym narzędziem badawczym w wielu sytuacjach, np. gdy nie ma potrzeby aż tak mocnej personalizacji produktu lub usługi, ponieważ są one w miarę homogeniczne (np. dostawy energii elektrycznej czy gazu). Dodatkowo dzięki relatywnie prostej procedurze badawczej, dobrze poznanym metodom szacowania i długiej historii stosowania omawianych metod znane są ich wady, zalety i ograniczenia. Wiadomo, która i kiedy daje lepsze rezultaty.

Nowsze metody badawcze – tj. metody adaptacyjne oraz metoda wyborów oparta na menu – z jednej strony powstały jako odpowiedź na ograniczenia metod *conjoint analysis* i metod wyborów dyskretnych, a z drugiej pozwalają na personalizację produktów i usług. Można powiedzieć, że w przeciwieństwie do metod wcześniej omówionych są silnie ukierunkowane na potrzeby i zachowania jednostek. Podejście to jest szczególnie istotne w kontekście produktów i usług, które można czy wręcz trzeba dopasowywać do potrzeb klienta (np. oferta ubezpieczeniowa, bankowa; komputery osobiste).

Rozwój technologiczny oraz metodologiczny pozwala obecnie na wybór odpowiedniej metody *conjoint* do badania. *Conjoint analysis* ma bogatą i interesującą historię, i można przypuszczać, że jej przyszłość będzie równie interesująca.

Literatura

- ACBC *Technical Paper* [2009], <http://www.sawtoothsoftware.com/download/techpap/acbctech.pdf> (dostęp: 9.10.2012).
- Adaptive Choice-based Conjoint Analysis* [2003], <http://pentor-arch.tnsglobal.pl/58346.xml> (dostęp: 12.11.2012).
- Bąk A. [2000], *Możliwości wykorzystania alternatywnych algorytmów conjoint analysis w badaniach marketingowych*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, nr 874, Wrocław.
- DeSarbo W.S., Ramaswamy V. [1995], *Marketing Segmentation with Choice-based Conjoint Analysis*, „Marketing Letters”, vol. 6, nr 2.

- Gilbride T., Allenby G.M. [2004], *A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules*, „Marketing Science”, vol. 23, nr 3.
- Green P.E., Rao V.R. [1971], *Conjoint Measurement for Quantifying Judgmental Data*, „Journal of Marketing Research”, vol. 8, nr 3.
- Green P.E., Srinivasan V. [1978], *Conjoint Analysis in Consumer Research: Issues and Outlook*, „Journal of Consumer Research”, vol. 5, nr 2.
- Green P.E., Wind Y. [1975], *New Way to Measure Consumers' Judgments*, „Harvard Business Review”, July–August.
- Hauser J.R. et al. [2006], „Must Have” Aspects vs. Tradeoff Aspects in Models of Customer Decisions [w:] *Proceedings of the Sawtooth Software Conference*, Sequim, Washington.
- Johnson R.M. [1974], *Trade-off Analysis of Consumer Values*, „Journal of Marketing Research”, vol. 11, nr 2.
- Johnson R.M. [2011], *History of Sawtooth Software's CBC Program*, <http://www.sawtooth-software.com/download/techpap/historyCBC.pdf> (dostęp: 6.12.2012).
- Krapp A., Sattler H. [2001], *Rethinking Preference Measurement*, http://www.henriksattler.de/publikationen/HS_046.pdf (dostęp: 10.05.2006).
- Luce R.D., Turkey J.W. [1964], *Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement*, „Journal of Mathematical Psychology”, vol. 1, nr 1.
- Optymalizacja taryf telekomunikacyjnych* [2012], <http://pentor.com.pl/web/tel1/tel1logn.htm> (dostęp: 12.11.2012).
- Orme B. [2009], *A Short History of Conjoint Analysis*, <http://www.sawtoothsoftware.com/download/techpap/cahistory.pdf> (dostęp: 10.12.2012).
- Orme B. [2010a], *Menu-based Choice Modeling Using Traditional Tools*, <http://www.sawtoothsoftware.com/download/techpap/mbconf2010.pdf> (dostęp: 10.10.2012).
- Orme B. [2010b], *Task Order Effects in Menu-Based Choice*, <http://www.sawtoothsoftware.com/download/techpap/mbcorder2010.pdf> (dostęp: 10.10.2012).
- Strojny J. [2003], *Metoda conjoint*, http://www.ar.krakow.pl/~jstrojny/Matrix/Analiza/Metody/conjoint/etapy_rozwoju.html (dostęp: 10.01.2013).
- Walesiak M., Bąk A. [2000], *Conjoint analysis w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

From Conjoint Analysis to Menu-based Choice

One of most popular methods applied in stated preference evaluation is conjoint analysis. While doing preference evaluations with the application of conjoint analysis, a respondent is presented a set of profiles (products or services) to evaluate. Respondents are asked to rank or order profiles. While using discrete choice studies we don't ask respondents to rank or order profiles, but we present a set of profiles and ask him or her to choose either one or none of them.

Preference measurement and evaluation with the application of adaptive conjoint analysis is done in two phases. The respondent evaluates preferences towards levels of all attributes, then evaluates selected pairs and chooses one profile. This method allows for the use of a large number of attributes and levels. But because of the need to choose between pairs of profiles it is different from the consumer behaviour market. To get more real-life data, an adaptive discrete choice method was proposed to describe both compensatory and non-compensatory

behaviours. Menu-based choice, designed for markets where a consumer can make choices from the menu, is the newest method of preference analysis. The main aim of the paper is to describe all of the methods, starting from conjoint analysis on through to menu-based choice.

Keywords: revealed preference analysis, conjoint analysis, discrete choice method, adaptive conjoint analysis, adaptive discrete choice method, menu-based choice method.

Paweł Chlipała

Katedra Marketingu

Uniwersytet Ekonomiczny w Krakowie

Orientacja hybrydowa a nowe podejścia w marketingu

Streszczenie

W artykule omówiono możliwości zastosowania orientacji hybrydowej, czyli tzw. trzeciego ruchu metodologicznego, w badaniu współczesnych koncepcji marketingu, takich jak marketing wartości, marketing relacji, marketing społeczny czy marketing międzynarodowy. W tekście zaprezentowano podstawy metodologiczne podejścia zintegrowanego – jego pojęcie, istotę oraz typologie. Następnie przedstawiono możliwości i ograniczenia wykorzystania podejścia zintegrowanego w badaniach koncepcji marketingu.

Słowa kluczowe: integracja metod, integracja podejść badawczych, badania pozytywistyczne, badania interpretacyjne, badania marketingu.

1. Wprowadzenie

Celem artykułu jest zaprezentowanie możliwości zastosowania orientacji hybrydowej, czyli tzw. trzeciego ruchu metodologicznego, w badaniu współczesnych koncepcji marketingu, takich jak marketing wartości, marketing relacji, marketing społeczny czy też orientacja marketingowa. To podejście metodologiczne bazuje na dwóch szkołach badań: ilościowych i jakościowych. Jest ono próbą oparcia badań na zgoła odmiennych nurtach metodologicznych, polegającą na ich integracji w procesie badawczym. Nie jest to, jak zostało zaprezentowane w dalszej części artykułu, zadanie łatwe, jednak jest ono możliwe do realizacji. Co więcej, wydaje się szczególnie przydatne w eksploracji nowych koncepcji marketingowych, które stanowią problemowe, złożone, wielotematyczne wątki

badawcze. Ich badanie jedną ścieżką metodyczną stwarza ograniczenia w budowaniu kompleksowej wiedzy. Zaletą orientacji hybrydowej jest możliwość tworzenia szerokiej wiedzy o zróżnicowanym charakterze.

2. Podstawy metodologiczne orientacji hybrydowej

Orientacja określona w artykule mianem hybrydowej w literaturze przedmiotu nazywana jest trzecią ścieżką metodologiczną (*the third path*), trzecim paradygmatem badawczym oraz trzecim ruchem metodologicznym (*the third methodological movement*) [Teddle i Tashakkori 2009, s. 4]. Stosowane jest także pojęcie metodologii zintegrowanej, zmierzającej – co podkreśla sens orientacji – do redukcji dychotomii pomiędzy badaniami ilościowymi i jakościowymi [Ploveright 2011, s. 2–3]. W wielu pozycjach poświęconych prezentowanej problematyce używa się określenia „metody zintegrowane” lub „metody mieszane” (*mixed methods*) (np. [The Mixed Methods Reader... 2008, *Handbook of Mixed Methods*... 2010]). Lektura pozycji zawierających w tytułach nawiązania do zintegrowanych metod pozwala wnioskować, że autorzy mają na uwadze raczej integrację podejść metodologicznych niż samych metod badawczych, a więc używanie pojęcia łączonych metod ma zbyt wąski zakres przedmiotowy, wszak istnieją jeszcze inne płaszczyzny integracji. Podejście zintegrowane może dotyczyć także kompilowania teorii, danych oraz integracji badaczy [Mazurek-Łopacińska i Sobocińska 2011, s. 19; *Theory and Methods*... 2011, s. 160, 215]. Orientacja hybrydowa obejmuje łączenie procedur badawczych – tak na etapie projektowania, jak i przetwarzania wyników. Integracja dotyczy prowadzących badania, ale także przedmiotów badań. W procedurze badawczej podlegają jej aspekty teoretyczne oraz metodyczne.

Pomimo zróżnicowanego nazewnictwa prezentowane podejście oznacza łączenie (na różnych poziomach) metodologii badań jakościowych z badaniami ilościowymi. Tak rozumiana orientacja hybrydowa jako alternatywa dla badań prowadzonych wyłącznie w jednej ścieżce metodycznej spełnia dwie podstawowe funkcje, które można określić mianem kontrolnej oraz synergicznej.

Funkcja kontrolna pozwala weryfikować wiedzę gromadzoną w ramach jednej procedury badawczej danymi zebranymi w ramach odrębnej procedury badawczej. W literaturze przedmiotu często stosuje się określenie „triangulacja”, oznaczające procedurę gromadzenia danych za pomocą więcej niż jednej metody badawczej, a następnie integrowanie wyników. Ma to zminimalizować ryzyko błędnego pomiaru wynikające z ograniczeń danej metody. Funkcji kontrolnej nie należy – jak zauważył M. Hammersley [*Advances in Mixed Methods Research*... 2008, s. 23] – zawęzać wyłącznie do weryfikacji wiedzy gromadzonej z uwzględnieniem

różnych podejść metodycznych, gdyż można dokonywać triangulacji w obrębie metod stosowanych w jednym podejściu, np. dane z ankiety pocztowej mogą być konfrontowane z wynikami ustrukturalizowanego wywiadu bezpośredniego. To stanowisko potwierdza nakreślone powyżej spektrum możliwości w zakresie integrowania podejść badawczych.

Funkcja synergiczna podejścia hybrydowego pozwala przezwyciężyć ograniczenia poszczególnych podejść metodycznych. Wiedza budowana na podstawie integracji podejść metodologicznych wykracza poza wiedzę zdobytą w ramach wyłącznie jednej procedury badawczej [Mazurek-Łopacińska i Sobocińska 2011, s. 19]. Pomimo licznych trudności i ograniczeń w zakresie stosowania podejścia hybrydowego, które zostały scharakteryzowane w dalszej części artykułu, badania te eliminują ryzyko błędnego pomiaru oraz pozwalają przyjąć szeroką perspektywę, co jest bardzo ważne w badaniach w warunkach dynamicznie zmieniającego się otoczenia.

U. Flick [2011, s. 95–97, 185–186] przedstawił dwie grupy wyborów badaczy w zakresie metodologii:

- decyzje odnośnie do ścieżki metodycznej, sprowadzające się do wyboru pomiędzy badaniami ilościowymi i jakościowymi,
- decyzje o kombinacji ścieżek metodycznych, gdzie należy określić zakres i formy integracji różnych podejść.

W tabeli 1 zaprezentowano podstawowe cechy wyróżniające badania ilościowe, prowadzone na podstawie paradygmatu pozytywistycznego, oraz badania jakościowe, czerpiące z paradygmatu konstruktywistycznego.

Tabela 1. Podstawowe cechy różnicujące paradygmat konstruktywistyczny i pozytywistyczny

| Płaszczyzna porównania | Paradygmat konstruktywistyczny | Paradygmat pozytywistyczny |
|--|--|--|
| Epistemologia: relacja pomiędzy badaczem a przedmiotem badań, charakter wiedzy i jej weryfikacja | Badacz wchodzi w interakcję z obszarem badanym, jest nierozzerwalnie złączony z procesem zdobywania wiedzy | Badacz i przedmiot badań są niezależni od siebie – dualizm |
| Aksjologia: rola wartości w procesie poznania | Empiryczne poznanie jest powiązane z wartościowaniem | Empiryczne poznanie zmierza do obiektywizmu i braku wartościowania |
| Ontologia: natura rzeczywistości, bytu i prawdy | Rzeczywistość jest złożona, ukryta (trudna w identyfikacji) i holistyczna | Rzeczywistość jest prosta, jawna (łatwa w poznaniu) i poddająca się fragmentyzacji |

cd. tabeli 1

| Płaszczyzna porównania | Paradygmat konstruktywistyczny | Paradygmat pozytywistyczny |
|---|--|--|
| Możliwość określenia związków przyczynowych | Niemożliwe jest rozróżnienie przyczyn od skutków, gdyż jednostki badane podlegają ciągłym i wielorakim zmianom | Można identyfikować przyczyny zdarzeń, które w dłuższych lub krótszym czasie wywołują określone efekty |
| Możliwość generalizacji | Wyłącznie wiedza zależna od czasu i kontekstu (idiograficzna) jest możliwa | Wiedza niezależna od czasu i kontekstu, poddająca się generalizacji (nomotetyczna), jest możliwa |

Źródło: [Teddle i Tashakkori 2009, s. 86].

Paradygmaty rozwijają się w czasie, także ich główne cechy charakterystyczne są modyfikowane – jak zauważyli w wyniku krytycznego spojrzenia na własne opracowania E. Guba i Y. Lincoln [*Metody badań jakościowych...* 2010, s. 291–292]. W typologii podejść badawczych metodologię pozytywistyczną przeciwstawia się także metodologii interpretacyjnej [Sagan 2004, s. 16–19]. Niezależnie jednak od ewolucji nauki i przyjmowanej terminologii, warto mieć na uwadze dwie skrajnie odmienne metodologie badawcze, charakteryzujące się różnymi procedurami badawczymi, innym podejściem do badań i rolą badacza, gromadzonymi informacjami i wiedzą o zupełnie odrębnym charakterze. W badaniach pozytywistycznych badacz często zostaje oddzielony od procesu gromadzenia informacji, koncentruje się na zarządzaniu projektem badawczym i organizowaniu poszczególnych zadań. Na etapie projektowania badań określa się problem badawczy, dokonuje szczegółowej kategoryzacji przedmiotów badań, formułuje się hipotezy badawcze. W tego typu badaniach dąży się do obiektywnego poznania rzeczywistości – badacz stara się oddzielić materiał badawczy od swoich przekonań i wyznawanych wartości. Paradygmat pozytywistyczny zakłada, że rzeczywistość można uchwycić, poddać fragmentyzacji i zbadać. W badaniach pozytywistycznych określa się relacje przyczynowo-skutkowe oraz prognozuje – żeby tego dokonać, dąży się do osiągnięcia reprezentatywnych wyników.

Paradygmat konstruktywistyczny lub interpretacyjny obliguje badacza do bezpośredniego i osobistego związku z badanym obszarem – wchodzi on w interakcje z podmiotem badań. W badaniach interpretacyjnych na wstępie określa się problem i tworzy ogólny koncept badań. Podejście to pozwala stawiać otwarte, poznawcze pytania oraz pozostawia możliwość dokonywania późniejszych zmian kierunku badań, a nawet głównego problemu badawczego. Efektem procesu badań interpretacyjnych jest wiedza idiograficzna, budowana na drodze przyczynowego rozumowania, która nie daje podstaw do tworzenia uogólnień, praw i reguł, jednak

umożliwia zrozumienie badanych zjawisk. Cały proces poznania nie jest oddzielony od wartościowania, gdyż w konstruktywizmie zakłada się, że rzeczywistość jest ukryta i złożona. W tak skomplikowanym świecie niemożliwe jest oddzielenie przyczyn zjawisk od ich efektów. Zdobytej w procedurze badań interpretacyjnych wiedzy nie można generalizować, gdyż zawsze jest ona zależna od czasu przeprowadzenia badania i jego kontekstu.

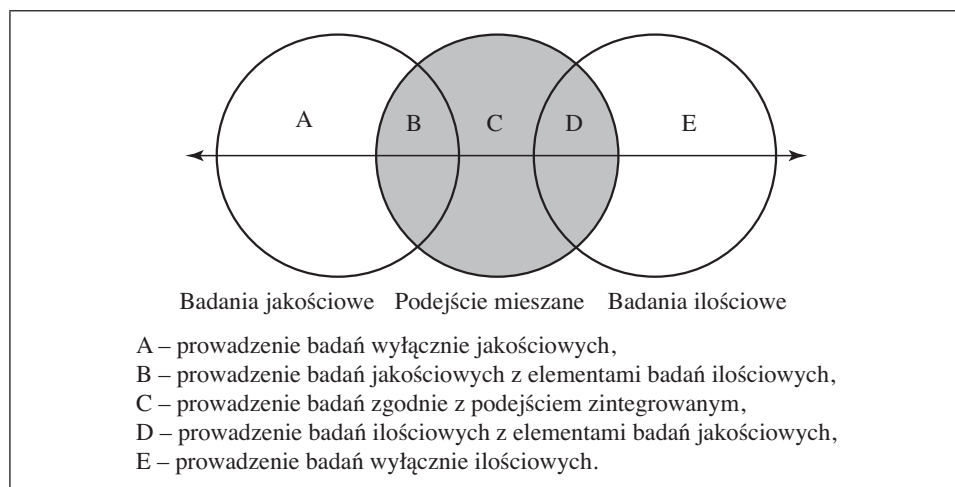
Podejścia metodyczne oparte na paradygmacie konstruktywistycznym i pozytywistycznym są tak dalece odmienne, że wśród dużej części badaczy można znaleźć zwolenników jednej lub drugiej opcji. Coraz częściej jednak w literaturze przedmiotu zauważa się, że badacz niekoniecznie musi wybierać pomiędzy dwiema metodologiami, gdyż może zastosować podejście mieszane, nazywane – jak już wspomniano – „trzecią ścieżką metodyczną”. Wówczas podejmuje on decyzję o kombinacji ścieżek metodycznych, precyzując zakres i formy integracji podejść.

J.W. Creswell [2003] wyróżnia następujące podejścia (metody badawcze):

– fazowe – prowadzone są zarówno badania ilościowe, jak i jakościowe, jednak realizowane są one odrębnie, jedno poprzedzają drugie (przy czym kolejność nie ma znaczenia). Taka procedura badawcza może zawierać więcej niż jedną, naprzemienną sekwencję badań ilościowych i jakościowych;

– zintegrowane – z podejściem dominującym. Badania prowadzone są wówczas na podstawie jednej, głównej metodologii, druga stanowi jej uzupełnienie, jednak jej udział w całych badaniach jest marginalny;

– mieszane, które łączą oba podejścia badawcze w jedną wspólną procedurę badawczą.



Rys. 1. Kontinuum podejść badawczych

Źródło: [Teddle i Tashakkori 2009, s. 28].

Podójście mieszane reprezentuje pełną integrację, dzięki której badacz czerpie z dorobku obu szkół badawczych. Formy i zakres integracji zostały scharakteryzowane na rys. 1.

Idea podejścia hybrydowego jest prosta i zakłada dążenie do powiązania badań jakościowych z badaniami ilościowymi. Dużo trudniejsze jest zastosowanie w praktyce takiego podejścia z kilku ważnych powodów. Do najważniejszych kwestii, które powinny być uwzględnione, aby można było skutecznie stosować podejście mieszane, należą:

- postawy i orientacja badaczy,
- wiedza i umiejętności badaczy,
- odrębność proceduralna badań.

Badacze są najczęściej „wychowywani” w jednej tradycji metodologicznej, z którą się utożsamiają. D. Silverman [2007, s. 56–62] opisał liczne antagonizmy pomiędzy badaczami zorientowanymi ilościowo i jakościowo. Wielu badaczy ilościowych nie uznaje jako równoważnej procedury gromadzenia informacji w trybie jakościowym – i *vice versa*. W ten sposób w badaniach dochodzi często do integracji na poziomie B i D z rys. 1, zdecydowanie rzadsza jest pełna integracja.

Stosowanie metodologii zintegrowanej wymaga podejścia opartego na pluralizmie, synkretyzmie, a więc ciągłym kompromisie i balansowaniu pomiędzy faktami i wartościami, wiedzą i mądrością, racjonalnością i podejściem emocjonalnym, idealizmem i materializmem [*Handbook of Mixed Methods...* 2010, s. 90]. Wykorzystywanie trzeciej ścieżki metodologicznej wymaga wiedzy, umiejętności i doświadczenia badawczego w różnych nurtach metodycznych.

Problemem wydaje się też odrębność proceduralna badań o charakterze pozytywistycznym i konstruktywistycznym. Jak już wcześniej zauważono, oba podejścia odróżnia specyfika procedury badawczej. Zrozumiałe jest integrowanie badań na etapie projekcyjnym oraz budowania wiedzy, natomiast trudności może sprawiać łączenie trybu gromadzenia informacji, gdyż nie można przyjmować równocześnie antagonicznych założeń każdego z podejść metodologicznych. Ten problem stanowi największe wyzwanie dla zwolenników orientacji hybrydowej w badaniach.

3. Eksploracja nowych podejść w marketingu z zastosowaniem orientacji hybrydowej

Problematyka integracji podejść badawczych jest ważna we współczesnym marketingu, który może być rozumiany jako proces społeczny i zarządczy obejmujący identyfikację, kształtowanie i zaspokajanie potrzeb klientów, aby

w długim okresie zapewnić przedsiębiorstwu przetrwanie i rozwój oraz przyrost wartości rynkowej. Współczesny marketing jako przedmiot badań jest dziedziną złożoną, więc wykraczanie poza wiedzę gromadzoną za pomocą jednej ścieżki metodycznej jest niezbędne. Badania w zakresie współczesnego marketingu mogą dotyczyć sfery ogólnej – marketingu jako dyscypliny (dziedziny naukowej – chociaż nie ma zgodności osób zajmujących się marketingiem co do jego naukowego statusu). W tym kontekście badania marketingu odnoszą się do wskazywanego przez wielu autorów (np. J. Dietla, L. Zabińskiego, H. Mruka i J. Szumilaka) problemu tożsamości marketingu, definicji jego zakresu przedmiotowego, a także kwestii spójności terminologicznej i koncepcyjnej.

Zintegrowane podejście badawcze może znacznie wzbogacić efekty w zakresie gromadzenia wiedzy o współczesnych koncepcjach marketingowych, takich jak marketing wartości, marketing relacyjny, marketing społeczny czy marketing międzynarodowy. Badania współczesnego marketingu mogą dotyczyć wielu koncepcji marketingowych i związanych z nimi problemów, których przykłady zaprezentowano w tabeli 2.

Tabela 2. Koncepcje współczesnego marketingu jako przedmiot badań

| Koncepcja | Wybrane problemy |
|--------------------------|---|
| Marketing wartości | <ul style="list-style-type: none"> – kształtowanie wartości dla klienta – wartość klienta dla przedsiębiorstwa – podnoszenie wartości przedsiębiorstwa dla akcjonariuszy |
| Marketing relacji | <ul style="list-style-type: none"> – skuteczność programów budowania lojalności – wpływ dobrych relacji na poczucie satysfakcji klientów i powtarzalność zakupów – nielojalność jako element postaw współczesnych konsumentów |
| Marketing społeczny | <ul style="list-style-type: none"> – koncepcja społecznej odpowiedzialności w działalności marketingowej przedsiębiorstw – odbiór koncepcji marketingu społecznego przez pracowników i klientów organizacji – orientacja marketingowa w realizacji idei społecznej odpowiedzialności |
| Marketing międzynarodowy | <ul style="list-style-type: none"> – problem standaryzacji i adaptacji narzędzi marketingowych na rynkach międzynarodowych – postawy kierownictwa i ich wpływ na kształtowanie strategii marketingu międzynarodowego |

Źródło: opracowanie własne.

Przegląd problemów współczesnego marketingu, w tym także w zakresie nowych koncepcji i wyzwań stawianych przed nim przez burzliwe otoczenie, prowadzi do wniosku, że badanie tej sfery z wykorzystaniem wyłącznie jednego podejścia metodologicznego nie jest wystarczające. Wiedza o marketingu lub

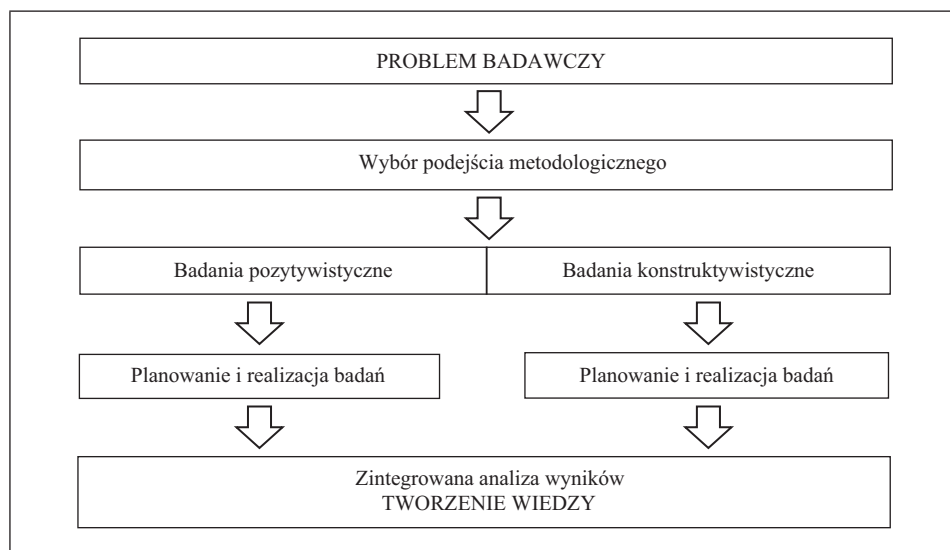
Tabela 3. Badania pozytywistyczne i konstruktywistyczne jako płaszczyzna integracji wiedzy o koncepcjach marketingowych

| Koncepcja marketingowa | Przykład zakresu przedmiotowego badań | Zakres problemowy – charakter wiedzy | |
|--------------------------|---|--|--|
| | | badania pozytywistyczne | badania konstruktywistyczne |
| Marketing wartości | wartość dla klienta | <ul style="list-style-type: none"> – strategie podnoszenia wartości dla klienta w przedsiębiorstwach – skala zjawiska – źródła satysfakcji nabywców z oferty | <ul style="list-style-type: none"> – uwarunkowania wdrażania w przedsiębiorstwie koncepcji kształtowania wartości na podstawie założeń aksjologii fenomenologicznej |
| Marketing relacji | kształtowanie lojalności klientów | <ul style="list-style-type: none"> – koszty i efekty programów lojalnościowych wdrażanych przez przedsiębiorstwa – prognoza przychodów generowanych z powtarzalnych transakcji w różnych branżach | <ul style="list-style-type: none"> – problem lojalności i nieojalności wobec sie- bie sprzedających i kupujących w kontekście podejścia do konsumpcji i stylu życia |
| Marketing społeczny | społeczna odpowiedzialność w tworzeniu oferty produktowej | <ul style="list-style-type: none"> – uwzględnianie zasad społecznej odpowiedzialności w strategii przedsiębiorstwa – stosowanie certyfikatów i oznaczeń na opakowaniach produktów | <ul style="list-style-type: none"> – przyzwyczajenie w życiu osobistym i zawodowym pracowników w kontekście projektowania oferty produktowej |
| Marketing międzynarodowy | motywy i przesłanki internacjonalizacji przedsiębiorstw | <ul style="list-style-type: none"> – hierarchia motywów rynkowych, ekonomicznych i prawnych w podejmowaniu działalności eksportowej – oceny kadry zarządzającej – motywy internacjonalizacji przedsiębiorstw a ich wielkość i skala eksportu | <ul style="list-style-type: none"> – okoliczności i uwarunkowaniach wchodzenia na rynki zagraniczne – rola charakteru, systemu wartości i utęgniwości osób podejmujących decyzje – problem przypadku w procesie internacjonalizacyjnym – jak wiele nie zależy od planów strategicznych i ich realizacji? |

Źródło: opracowanie własne.

jego wycinku może być szeroka, dynamiczna, ale płytka, jeżeli zdecydujemy się na badania w nurcie pozytywistycznym, lub będzie to wiedza pogłębiona, ale „lokalna” i niepoddająca się uogólnieniom, kiedy badania będą miały charakter interpretacyjny. Orientacja hybrydowa pozwala skorzystać z dwóch opisanych powyżej efektów łączenia podejść badawczych. Po pierwsze, możliwy jest efekt synergii ułatwiający kształtowanie szerszej i pełniejszej wiedzy na temat przedmiotu badań. Po drugie, korzystny jest efekt kontrolny, a więc triangulacja problemów badawczych, metod oraz zgromadzonych informacji. Wiedza uzyskana w ramach obu podejść badawczych wzajemnie się uzupełnia, co ilustruje tabela 3.

Zastosowanie integracji metodologii badawczych w eksploracji i eksplanacji współczesnych koncepcji marketingowych może mieć formę zrównoważoną lub formę dominacji jednego z podejść. Należy docenić zalety obu wariantów, jednak wydaje się, że najlepsze efekty przyniesie dążenie do zachowania równych proporcji w stosowaniu obu podejść metodologicznych.



Rys. 2. Integracja podejść badawczych w procesie tworzenia wiedzy o współczesnych koncepcjach marketingowych – ujęcie modelowe

Źródło: opracowanie własne.

Integracja metodologii pozytywistycznej i konstruktywistycznej nie musi polegać na wiązaniu obu procedur badawczych – już sama integracja na etapie projektu badań i budowania wiedzy (jak zaprezentowano na rys. 2) jest skuteczna. Za taką właśnie integracją opowiada się autor nieniejszego opracowania, pozwala ona bowiem korzystać z efektów synergii oraz triangulacji i jednocześnie

zachować założenia ontologiczne, aksjologiczne i epistemologiczne każdego z podejść metodycznych, a także odrębność proceduralną.

4. Zakończenie

Poznanie współczesnego marketingu oraz jego nowych koncepcji stanowi duże wyzwanie dla naukowców – także metodyczne. Omówione w artykule podejście hybrydowe to jedna z propozycji badawczych starająca się przewyciężyć ograniczenia podejścia pozytywistycznego oraz podejścia konstruktywistycznego. Zaletą takiej metodologii jest możliwość zdobywania pełniejszej i wiarygodniejszej wiedzy, niż ma to miejsce w badaniach prowadzonych w ramach jednej ścieżki metodologicznej. Należy przy tym zauważyć, że oprócz wskazanych w tekście licznych ograniczeń związanych z preferencjami badaczy, koniecznością posiadania wiedzy, umiejętności i doświadczeń w zakresie prowadzenia badań w odmiennych metodologiach oraz trudności w integrowaniu niezależnych z założenia procedur badawczych integracja badawcza wydłuża czas i zwiększa koszty badań. Wydaje się jednak, że są to koszty niezbędne, które należy ponieść dla rozwoju marketingu i pogłębiania znajomości jego koncepcji.

Literatura

- Advances in Mixed Methods Research* [2008], red. M.M. Bergman, Sage, Los Angeles.
- Creswell J.W. [2003], *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, Sage, Thousand Oaks.
- Flick U. [2011], *Introducing Research Methodology: A Beginner's Guide to Doing a Research Project*, Sage, London.
- Handbook of Mixed Methods in Social & Behavioral Research* [2010], red. A. Tashakkori, Ch. Teddlie, wyd. 2, Sage, Thousand Oaks.
- Mazurek-Łopacińska K., Sobocińska M. [2011], *Rozwój badań marketingowych – w kierunku nowych podejść i kontekstów badawczych związanych z funkcjonowaniem przedsiębiorstwa*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 236, Wrocław.
- Metody badań jakościowych* [2010], red. N.K. Denzin, Y.S. Lincoln, tłum. M. Bobako, t. 1, Wydawnictwo Naukowe PWN, Warszawa.
- Plowright D. [2011], *Using Mixed Methods: Framework for an Integrated Methodology*, Sage, London.
- Sagan A. [2004], *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Silverman D. [2007], *Interpretacja danych jakościowych*, tłum. M. Głowacka-Grajper, J. Ostrowska, Wydawnictwo Naukowe PWN, Warszawa.

Teddlie Ch., Tashakkori A. [2009], *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, Sage, Los Angeles.

The Mixed Methods Reader [2008], red. V.L. Plano Clark, J.W. Creswell, Sage, Los Angeles.

Theory and Methods in Social Research [2011], red. B. Somekh, C. Lewin, wyd. 2, Sage, London.

Hybrid Orientation and New Approaches in Marketing

The paper presents the possibility of using the hybrid orientation called “the third methodological movement” in the study of modern marketing concepts, such as value-based marketing, relationship marketing, social marketing or international marketing. This article presents the methodological basis of an integrated approach, presenting its concept, nature and typologies. It then presents the opportunities and limitations of using an integrated, hybrid approach to research marketing concepts.

Keywords: integrated methods, mixed methods, integration of research approaches, positivist research, interpretive research, research on marketing concepts.

Alicja Grześkowiak
Agnieszka Stanimir

Katedra Ekonometrii
Uniwersytet Ekonomiczny we Wrocławiu

Wielowymiarowa analiza społeczno-demograficznych aspektów wykorzystania Internetu

Streszczenie

Celem artykułu było sprawdzenie prawdziwości hipotezy o zróżnicowaniu wykorzystania komputerów do posługiwania się Internetem w celach komunikacyjnych, zakupowych, rozrywkowych i informacyjnych. Analizowano różnice dotyczące osób należących do dwóch grup wiekowych: rozpoczynających i kończących karierę zawodową w latach 2003 i 2011. Ponadto w badaniu uwzględniono dodatkowe cechy społeczno-demograficzne: płeć, wykształcenie oraz miejsce zamieszkania. Dane zaczerpnięto z Diagnozy społecznej. Analizę przeprowadzono, korzystając z regresji logistycznej oraz analizy korespondencji. W wyniku przeprowadzonych badań potwierdzono zróżnicowanie w zakresie eksploatacji zasobów Internetu przez osoby starsze i młode, z uwzględnieniem pozostałych zmiennych.

Słowa kluczowe: regresja logistyczna, analiza korespondencji, kompetencje ICT, różnice pokoleniowe.

1. Wprowadzenie

W *Strategii rozwoju społeczeństwa informacyjnego w Polsce do roku 2013* jako jeden z celów w obszarze dotyczącym kapitału ludzkiego wskazano „podniesienie poziomu motywacji, świadomości, wiedzy oraz umiejętności w zakresie wykorzystywania technologii informacyjnych i komunikacyjnych” [*Strategia*

rozwoju... 2008, s. 2]. Kompetencje w zakresie wykorzystania obszaru ICT stają się niezbędne w życiu codziennym oraz w pracy zawodowej, a ich brak pozostaje w ścisłym związku z wykluczeniem społecznym [Wykluczenie cyfrowe... 2010, s. 11]. Jak podkreślono w *Europe's Digital Competitiveness Report* [2009, s. 178–179], rozwój społeczeństwa informacyjnego w Polsce następuje bardzo powoli, a wartości wskaźników charakteryzujących rozwój sektora ICT należą do najniższych w Europie. Powstaje pytanie o determinanty niskiego wykorzystania możliwości, jakie niesie rozwój technologii informacyjnych i komunikacyjnych. Problematyka ta jest bardzo szeroka, a w niniejszej pracy podjęto próbę analizy jednego z jej aspektów związanych z uwarunkowaniami korzystania z Internetu.

Głównym celem przeprowadzonych badań było zweryfikowanie hipotezy, że na zakres wykorzystania Internetu ma wpływ wiele czynników społeczno-demograficznych, takich jak: wykształcenie, miejsce zamieszkania, płeć oraz przynależność do starszej lub młodszej grupy wiekowej.

W celu zbadania postawionej hipotezy przeprowadzono analizy wielowymiarowe: analizę korespondencji – jako metodę badania współwystępowania oraz regresję logistyczną – jako technikę badania współzależności. Analizy przeprowadzono, bazując na danych z Diagnozy społecznej z 2003 i 2011 r. dotyczących różnych obszarów wykorzystania Internetu [Rada Monitoringu Społecznego 2003, 2011]. W celu porównania sytuacji w czasie wybrano te zastosowania, o które respondenci byli pytani podczas obu edycji badania. Dotyczyły one różnych przejawów eksploatacji zasobów sieci internetowej:

- komunikowania się za pomocą narzędzi internetowych: „sprawdzanie i wysyłanie poczty elektronicznej (e-mail)”, „korzystanie z komunikatorów umożliwiających rozmowy ze znajomymi (takich jak ICQ, gadu-gadu)”, „uczestniczenie w czatach”, „uczestniczenie w grupach lub forach dyskusyjnych”,

- wyszukiwania informacji za pomocą narzędzi ICT: „przeglądanie stron WWW” – w 2003 r., „korzystanie z wyszukiwarek internetowych (np. Google)” – w 2011 r., „zbieranie materiałów potrzebnych do nauki lub pracy”,

- wykonywania zakupów przez Internet: „kupowanie produktów przez Internet”, „uczestniczenie w aukcjach internetowych”,

- rozrywek: „granie w gry sieciowe przez Internet”, „ściągnięcie darmowego oprogramowania, muzyki lub filmów”,

- zarządzania własną stroną internetową: „tworzenie lub modyfikowanie własnej strony WWW”¹.

¹ Wszystkie podane sformułowania pochodzą z Kwestionariuszy Diagnozy Społecznej [Rada Monitoringu Społecznego 2003, 2011].

W pracy rozpatrzono czynniki społeczno-demograficzne, które mogą wpływać na użytkowanie Internetu, takie jak: płeć, wykształcenie (według czterech kategorii: W1 – podstawowe i niższe, W2 – zasadnicze zawodowe/gimnazjum, W3 – średnie, W4 – policealne i wyższe), miejsce zamieszkania (według sześciu kategorii: wieś, miasta poniżej 20 tys. mieszkańców, miasta o liczbie mieszkańców 20–100 tys., miasta o liczbie mieszkańców 100–200 tys., miasta o liczbie mieszkańców 200–500 tys., miasta o liczbie mieszkańców 500 tys. i więcej) oraz przynależność do jednej z dwóch grup wiekowych: osób młodych rozpoczynających życie zawodowe lub osób starszych będących u końca kariery zawodowej w latach 2003 i 2011.

2. Identyfikacja czynników społeczno-demograficznych wpływających na korzystanie z poszczególnych obszarów ICT za pomocą regresji logistycznej

Odpowiedzi na pytania stawiane w ramach Diagnozy społecznej dotyczące rozpoznania celów, do osiągnięcia których wykorzystywane są nowoczesne technologie informacyjne, w 2011 r. miały charakter dychotomiczny, natomiast w 2003 r. zostały zdychotomizowane (respondenci zostali przyporządkowani do jednej z grup: używających bądź nieużywających określonego narzędzia). Do modelowania prawdopodobieństwa wystąpienia zmiennej binarnej w zależności od kształtowania się zmiennych niezależnych można zastosować regresję logistyczną określającą relację pomiędzy logarytmem ilorazu szans (logitem) a liniową kombinacją zmiennych niezależnych (zob. [Gruszczyński 2002, s. 58]). Estymacja regresji logistycznej jest zazwyczaj wykonywana metodą największej wiarygodności (zob. [Anderson 1997, s. 159–163]). Z punktu widzenia użyteczności interpretacyjnej można posiłkować się bądź otrzymanymi ocenami parametrów modelu, bądź ilorazem szans wskazującym na względną możliwość wystąpienia określonego zdarzenia (zob. [Rószkiewicz 2002, s. 176–178]). W niniejszej pracy do porównania siły wpływu czynników przyjęto podejście bazujące na wartościach ilorazu szans. Ocenę istotności poszczególnych parametrów przeprowadzono za pomocą testu Walda (zob. np. [Rószkiewicz 2011, s. 194–195]), przyjmując poziom istotności 0,05. Otrzymane rezultaty zaprezentowano według wyróżnionych zakresów zastosowań Internetu. W tabeli 1 przedstawiono czynniki zidentyfikowane jako istotnie wpływające na wykorzystanie poszczególnych obszarów, przy czym znakami (+) i (–) oznaczono odpowiednio przyrost i spadek szansy zastosowania danej funkcjonalności na skutek wzrostu danej zmiennej niezależnej o jedną kategorię.

Tabela 1. Czynniki istotnie wpływające na korzystanie z możliwości komunikowania się za pomocą Internetu

| Grupa aktywności | Czynność | Kierunek wpływu istotnych czynników ^a | | |
|--|---|--|---|--|
| | | 2003 | 2011 | |
| Komunikowanie się za pomocą Internetu | Wykorzystanie poczty elektronicznej | Wykształcenie (+) Wielkość miejscowości (+) | Wiek (+) Wielkość miejscowości (-) Wykształcenie (-) | |
| | Stosowanie komunikatorów | Wielkość miejscowości (+) Wiek (-) | Wykształcenie (-) | |
| | Uczestnictwo w czatach | Wielkość miejscowości (+) Wiek (-) | Wykształcenie (+) Wielkość miejscowości (+) Wiek (-) | |
| | Uczestnictwo w grupach lub forach dyskusyjnych | Wielkość miejscowości (+) Wiek (-) | Wielkość miejscowości (+) Wiek (-) | |
| Wyszukiwanie informacji za pomocą narzędzi ICT | Przeglądanie stron WWW | Wiek (-) | - | |
| | Korzystanie z wyszukiwarek internetowych | - | Wykształcenie (+) Płeć (mężczyzna +) Wiek (-) | |
| | Zbieranie materiałów potrzebnych do nauki lub pracy | Wiek (-) | Wielkość miejscowości (-) Wiek (-) | |
| Dokonywanie zakupów, rozrywka i tworzenie własnych stron internetowych | Zakupy | Kupowanie produktów i usług przez Internet | Wykształcenie (+) Płeć (mężczyzna +) Wiek (-) | |
| | | Uczestnictwo w aukcjach internetowych | Płeć (mężczyzna +) Wiek (-) Wielkość miejscowości (+) Wiek (-) Płeć (mężczyzna +) | |
| | Rozrywka | Granie w gry sieciowe przez Internet | Wielkość miejscowości (+) Płeć (mężczyzna +) Wiek (-) | Wielkość miejscowości (+) Wiek (-) |
| | | Ściąganie darmowej muzyki, filmów | Wielkość miejscowości (+) Płeć (mężczyzna +) Wiek (-) | Wielkość miejscowości (+) Wiek (-) |
| | Własne strony WWW | Tworzenie lub modyfikowanie strony WWW | Wiek (-) | Wykształcenie (+) Wielkość miejscowości (+) Wiek (-) |

^a „+” oznacza wzrost szansy, „-” oznacza spadek szansy.

Źródło: opracowanie własne na podstawie [Rada Monitoringu Społecznego 2003, 2011].

W 2003 r. zmiennymi wywierającymi duży wpływ na stosowanie różnych form komunikowania się za pomocą Internetu były przede wszystkim „wiek” i „wielkość miejscowości”, przy czym mieszkanie w większej miejscowości oraz niższy wiek zwiększały szanse na używanie internetowych narzędzi porozumiewania się. W 2011 r. zaobserwowano istotny wpływ zmiennej „wykształcenie”, której wyższe wartości predysponowały do uczestnictwa w czatach oraz, co zaskakujące, zmniejszały prawdopodobieństwo wykorzystania poczty elektronicznej i komunikatorów. Warto zauważyć, że w żadnej z kategorii zmienna „płeć” nie odgrywała znaczącej roli. Jedynym obszarem, w którym powtórzył się zestaw kluczowych czynników, było korzystanie z grup i forów dyskusyjnych.

Pytanie dotyczące przeglądania stron WWW, które było zadawane w 2003 r., nie zostało powtórzone w 2011 r. – wówczas przeformułowano je na pytanie dotyczące korzystania z wyszukiwarek internetowych. Obie kategorie są dosyć podobne, ale odmienne sformułowanie pytań nie upoważnia do przeprowadzenia bezpośrednich porównań. Należy podkreślić, że w 2003 r. jedyną zmienną istotnie wpływającą na korzystanie z zasobów Internetu (w sensie przeglądania stron WWW oraz zbierania materiałów niezbędnych w pracy lub w procesie edukacyjnym) był „wiek” (grupa wiekowa, do której przynależał respondent). Ten sam czynnik został zidentyfikowany jako istotny w 2011 r., ale ważne były także inne determinanty, jak „wykształcenie”, „płeć” oraz „wielkość miejscowości”.

Najczęściej występującym istotnym czynnikiem determinującym różnorakie wykorzystanie zasobów internetowych był „wiek”, przy czym zawsze iloraz szans dla tej zmiennej był mniejszy od jedności, czyli wśród starszych respondentów malało prawdopodobieństwo dokonywania zakupów, zabawy i relaksu z wykorzystaniem ogólnosiwiatowej sieci oraz tworzenia lub modyfikowania stron WWW. Warto zauważyć, że robienie zakupów przez Internet to jedyny obszar, w którym we wszystkich rozpatrywanych kategoriach i latach ważnym czynnikiem warunkującym podejmowanie działań była „płeć”. Co więcej, rezultaty wskazują, że zależność ta zachodziła zawsze w jednym kierunku – bycie mężczyzną zwiększało szanse na dokonywanie zakupów w sieci. Z kolei w obszarze „rozrywka” ważną determinantą była „wielkość miejscowości” – mieszkańcy większych miast byli bardziej skłonni do grania w gry sieciowe i pobierania darmowej muzyki i filmów.

3. Identyfikacja czynników społeczno-demograficznych wpływających na korzystanie z poszczególnych obszarów ICT za pomocą analizy korespondencji

Z uwagi na skalę pomiaru zmiennych właściwym postępowaniem badawczym jest poszukiwanie wniosków o interakcjach kategorii wyróżnionych zmiennych na podstawie analizy korespondencji. W prowadzonym badaniu bardzo istotne

znaczenie ze względu na zakres badania i ukierunkowanie zainteresowania interakcjami zmiennych miał wybór techniki prowadzenia analizy korespondencji. Wybrano podejście zaproponowane przez A. Stanimir [2011], łączące technikę prowadzenia analizy korespondencji dla wielowymiarowych i łączonych tablic kontyngencji².

W łączonych tablicach kontyngencji³ występują dwa typy zmiennych, które pełnią odmienne funkcje. Pierwszy z nich to zmienna łącząca tabele, nazwana w pracy A. Stanimir [2005, s. 50] cechą wspólną. W pracach angielskojęzycznych, np. M. Greenacre'a [1994, s. 143], można znaleźć określenie *dependent variable*. Z uwagi na brak definiowania funkcyjnej zależności zmiennych nie jest to właściwe określenie. Ten sam autor pozostałe zmienne nazywa *describing variables* lub *independent variables*. Za trafne określenie można uznać „zmienne opisowe”. Budując łączoną tablicę kontyngencji, tworzy się w pierwszym etapie kilka tablic kontyngencji między parami zmiennych (jak podaje M. Greenacre [1994, s. 145], można wykonać analizę dla każdej pary zmiennych oddzielnie). Następnie tablice ulegają zblokowaniu względem cechy wspólnej. Dzięki takiemu postępowaniu można dokonać opisu najistotniejszych dla badacza powiązań między zmiennymi oraz wprowadzić do badania zmienne nominalne o charakterze dychotomicznym⁴.

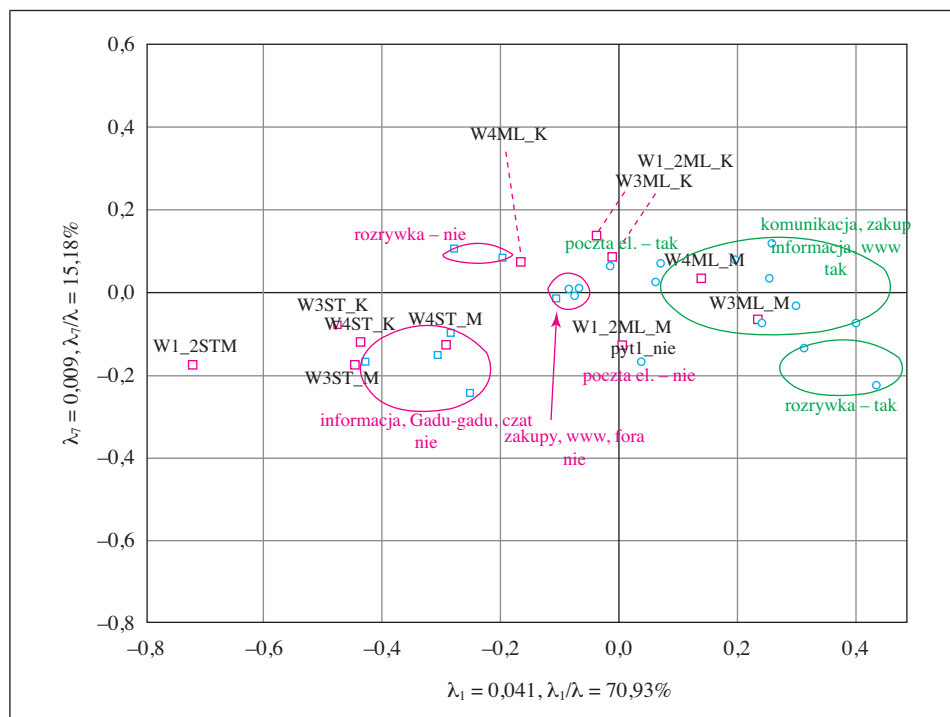
Ograniczenie analizy do łączonej tablicy kontyngencji wymusiłoby konieczność przeprowadzenia analizy kilkakrotnie – dla każdej zmiennej społeczno-demograficznej oddzielnie. Jako że celem prowadzonego badania jest rozpoznanie różnic w wykorzystaniu komputera do eksploatacji zasobów internetowych przez osoby młode i starsze, słuszne wydaje się powiązanie zmiennych w tzw. cechy kombinowane [Stanimir 2005, s. 27], tworząc warstwy w kolumnach lub wierszach. Do analizy korespondencji wybrano trzy zmienne społeczno-demograficzne: „płeć”, „wiek” (respondentów zgrupowano w dwie kategorie: starsi i młodzi) oraz „wykształcenie” (ponieważ po wykonaniu graficznej prezentacji wyników punkty obrazujące wykształcenie podstawowe i niższe oraz zasadnicze zawodowe/gimnazjum pokrywały się, to w ostatecznej analizie połączono te dwie kategorie). Z wyróżnionych zmiennych stworzono cechę kombinowaną o kategoriach według schematu: W1_2ML_M (młodzi – ML, mężczyźni – M, z wykształceniem niższym niż średnie – W1_2). Zmienna ta była w dalszym postępowaniu traktowana jako

² Szczegółowy opis algorytmu analizy korespondencji można znaleźć w pracy A. Stanimir [2005].

³ Łączona tablica kontyngencji wykorzystuje regułę addytywności statystyki χ^2 , która jest szczegółowo opisana w pracy [Yule i Kendall 1966, s. 474].

⁴ Przeprowadzenie klasycznej analizy korespondencji dla zmiennych zero-jedynkowych prowadzi do prezentacji wyników w przestrzeni jednowymiarowej, zatem „inercja całkowita jest reprezentowana przez pojedynczą linię” [Greenacre 1994, s. 145].

zmienna wspólna w łączonej tablicy kontyngencji. Działanie to miało na celu wskazanie najistotniejszych powiązań między zmiennymi z zachowaniem jak najwyższej jakości odzwierciedlenia tych relacji w przestrzeni o niskim wymiarze. Graficzną prezentacją uzyskanych wyników analizy korespondencji dla lat 2003 i 2011 są rys. 1 i 2.



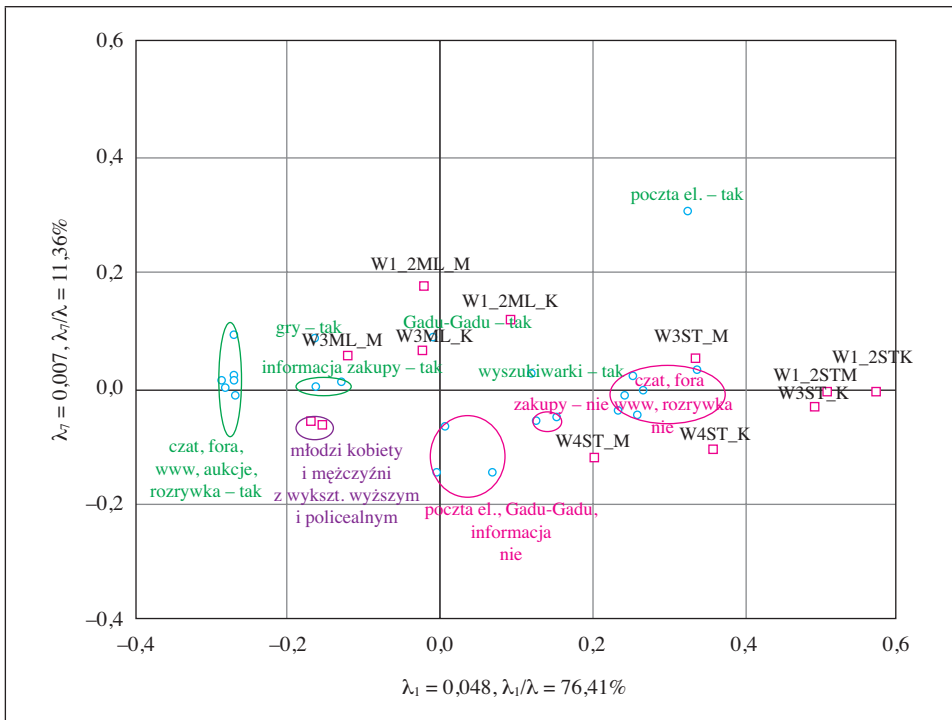
Rys. 1. Interakcje zmiennych „wykształcenie”, „wiek”, „płeć” z eksploatacją różnych obszarów zasobów internetowych w 2003 r.

Źródło: opracowanie własne z wykorzystaniem programu Statistica 10 na podstawie [Rada Monitoringu Społecznego 2003].

Dzięki jednoczesnemu zastosowaniu dwóch technik analizy korespondencji otrzymano dwuwymiarowe prezentacje powiązań kategorii analizowanych zmiennych, które odzwierciedlają 86% (2003 r. – rys. 1) oraz 88% (2011 r. – rys. 2) rzeczywistych powiązań między wyróżnionymi kategoriami zmiennych.

Analizując położenie punktów na rys. 1, można stwierdzić, że w 2003 r. badane osoby bez względu na wiek, płeć czy wykształcenie najmniej różniły się w kwestii niekorzystania z komputera w celu dokonywania zakupów, tworzenia stron WWW oraz uczestniczenia w forach. Młodzi mężczyźni z wykształceniem co najmniej

średnim korzystali z komputera w celach komunikacyjnych, dokonywali zakupów przez Internet, wyszukiwali informacje oraz tworzyli własne strony WWW. Młodzi mężczyźni z wykształceniem średnim dodatkowo korzystali z komputera w celach rozrywkowych. Młode kobiety z wykształceniem co najwyżej średnim wykorzystywały komputer, by odbierać i wysyłać pocztę elektroniczną. Dla młodych kobiet z wykształceniem co najmniej policealnym trudno wskazać najważniejsze aspekty korzystania z Internetu. Najbardziej charakterystyczne dla tych kobiet było niekorzystanie z Internetu w celu ściągania darmowego oprogramowania i grania w gry sieciowe. Młodzi mężczyźni z wykształceniem co najwyżej gimnazjalnym nie korzystali z komputera, by odbierać i wysyłać pocztę elektroniczną. Punkty obrazujące osoby starsze, niezależnie od płci i wykształcenia, są umiejscowione na rys. 1 z lewej strony w sąsiedztwie punktów wskazujących na niekorzystanie z komputera w takich celach, jak wyszukiwanie informacji, korzystanie z komunikatorów czy uczestnictwo w czatach.



Rys. 2. Interakcje zmiennych „wykształcenie”, „wiek”, „płeć” z eksploatacją różnych obszarów zasobów internetowych w 2011 r.

Źródło: opracowanie własne z wykorzystaniem programu Statistica 10 na podstawie [Rada Monitoringu Społecznego 2011].

Na rys. 2. widać większe zróżnicowanie w ocenie wykorzystania komputerów zarówno w grupie osób starszych, jak i młodych. Punkty obrazujące młode kobiety i młodych mężczyzn z wykształceniem co najmniej policealnym są położone blisko siebie, zatem postrzeganie wykorzystania komputera przez te dwie grupy było w 2011 r. zgodne. Najbardziej dla nich charakterystyczne było korzystanie z komputera w celu dokonywania zakupów, gromadzenia informacji, w celu rozrywkowym, a także w celu uczestnictwa w czatach i forach oraz tworzenia własnych stron WWW.

Na rys. 2 na uwagę zasługuje położenie punktów obrazujących niekorzystanie z poczty elektronicznej, komunikatorów i niegromadzenie informacji z Internetu. Te kategorie w 2011 r. nie odnosiły się do żadnej grupy osób. Młodzi mężczyźni z wykształceniem średnim bardzo często korzystali z komputera, by grać w gry sieciowe. Dla młodych kobiet o takim samym poziomie wykształcenia istotne było korzystanie z komunikatorów internetowych. Korzystanie z komunikatorów było również kategorią, którą można uznać za charakterystyczną dla młodych kobiet i mężczyzn z wykształceniem niższym niż średnie. Dla starszych kobiet z wyższym i policealnym wykształceniem oraz mężczyzn z wykształceniem średnim charakterystyczne było niekorzystanie z komputera w celach rozrywkowych, do tworzenia stron WWW czy uczestnictwa w forach i czatach. Punkt obrazujący korzystanie z poczty elektronicznej na rys. 2 jest położony w dużym oddaleniu od pozostałych punktów, z czego wynika, że jest to kategoria, która nie była w 2011 r. istotna dla żadnej z analizowanych grup społecznych. Prawdopodobnie poczta elektroniczna ustępuje miejsca innym komunikatorom internetowym lub usługom SMS. Punkt obrazujący niekorzystanie z wyszukiwarek na rys. 2 jest położony w jeszcze większym oddaleniu w prawą stronę od pozostałych punktów. Ta kategoria również nie charakteryzuje żadnej z analizowanych grup osób. Należy ponadto zwrócić uwagę na położenie punktu obrazującego korzystanie z wyszukiwarek internetowych. Punkt ten przy uwzględnieniu danych zarówno z 2003 r., jak i z 2011 r. jest położony blisko centrum rzutowania, co wskazuje na małe zróżnicowanie pod tym względem analizowanych grup osób. Na wykresie zawierającym dane z 2011 r. punkt ten w przeciwieństwie do wykresu za 2003 r. jest umiejscowiony po stronie obrazującej zachowania osób starszych. Wynika z tego, że coraz więcej osób należących do tej grupy wiekowej korzysta z wyszukiwarek internetowych.

4. Podsumowanie

Wyniki analizy korespondencji potwierdzają i jednocześnie uszczegóławiają wyniki uzyskane z analiz przeprowadzonych za pomocą regresji logistycznej.

Wybrane czynniki społeczno-demograficzne różnicują analizowaną populację ze względu na wykorzystanie komputera do określonych czynności. W szczególności czynnikami różnicującymi są wiek i wykształcenie. Kobiety i mężczyźni

należący do tej samej grupy wiekowej najczęściej realizują podobne modele wykorzystania Internetu.

Literatura

- Anderson E.B. [1997], *Introduction to the Statistical Analysis of Categorical Data*, Springer, Berlin.
- Europe's Digital Competitiveness Report. Main Achievements of the i2010 Strategy 2005–2009 [2009], http://ec.europa.eu/information_society/eeurope/i2010/docs/annual_report/2009/digital_competitiveness.pdf (dostęp: 15.12.2012).
- Greenacre M. [1994], *Multiple and Joint Correspondence Analysis [w:] Correspondence Analysis in the Social Sciences*, red. M. Greenacre, J. Blasius, Academic Press, Harcourt Brace & Company, San Diego.
- Gruszczyński M. [2002], *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, Oficyna Wydawnicza Szkoły Głównej Handlowej, Warszawa.
- Rada Monitoringu Społecznego [2003, 2011], *Diagnoza Społeczna: zintegrowana baza danych*, www.diagnoza.com (dostęp: 23.10.2012).
- Rószkiewicz M. [2002], *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa.
- Rószkiewicz M. [2011], *Analiza klienta*, SPSS Polska, Kraków.
- Stanimir A. [2005], *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
- Stanimir A. [2011], *Zastosowanie różnych technik analizy korespondencji w analizie danych personalnych*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 161, Wrocław.
- Strategia rozwoju społeczeństwa informacyjnego w Polsce do roku 2013. Streszczenie* [2008], <http://www.msw.gov.pl/download/56/6107/Streszczenie> (dostęp 15.12.2012).
- Wykluczenie cyfrowe na rynku pracy* [2010], red. E. Kryńska, Ł. Arendt, IPiSS, Warszawa.
- Yule G.U., Kendall M.G. [1966], *Wstęp do teorii statystyki*, tłum. B. Górecki, E. Vielrose, PWN, Warszawa.

Multivariate Analysis of Socio-demographic Determinants of Internet Use

The objective of this paper was to test a hypothesis on the diversity in the use of computers for online communication, shopping activities, entertainment and information acquisition. The differences between people belonging to two age groups – those who started and ended their professional careers either in 2003 or 2011 – were analysed.

The study included the additional socio-demographic characteristics: gender, level of education and place of residence. The data were retrieved from the “Social Diagnosis”, and logistic regression and correspondence analysis were used. The research findings confirm the differences in Internet use by the groups of younger and older people, taking into account other variables.

Keywords: logistic regression, correspondence analysis, ICT competences, intergenerational differences.

Piotr Tarka

Katedra Badań Rynku i Usług
Uniwersytet Ekonomiczny w Poznaniu

Model czynnikowy ze zmiennymi dychotomicznymi w analizie ukrytej struktury zjawisk rynkowych

Streszczenie

W artykule przedstawiono teoretyczne założenia konstrukcji wraz z praktycznymi implikacjami modelu czynnikowego w kontekście badań marketingowych. W pierwszej części zwrócono uwagę na zależność badań marketingowych od rozwoju teorii pomiaru w ramach analizy zjawisk rynkowych o charakterze ukrytym. Następnie omówiono proces konstrukcji modelu czynnikowego do analizy zmiennych dychotomicznych. Na podstawie przeprowadzonych badań empirycznych zaprezentowano funkcjonalność jednoczynnikowego modelu w sferze pomiaru postaw młodzieży (młodych konsumentów) wobec hedonistyczno-konsumpcyjnego stylu życia.

Słowa kluczowe: model czynnikowy, zmienne ukryte, dane dychotomiczne, skala nominalna.

1. Wprowadzenie

Gdyby projekty badawcze były zdominowane przez antyrealistów (kierujących się zasadą „brzytwy Ockhama”)¹, wiele pomysłów w badaniach, w tym także

¹ Zasada ta polega m.in. na tym, że wszystko to, czego nie można zmierzyć, nie istnieje.

wiele interesujących odkryć badawczych w marketingu, nigdy nie doczekałoby się realizacji. Na szczęście w metodologii badań marketingowych (w sposób szczególnie w sferze analizy zjawisk rynkowych ukrytych) zaistniały także prace realistów, według których zmienne obserwowalne podlegają statystycznym procesom łączenia i redukcji, co może przyczynić się do odkrywania zjawisk o charakterze ukrytym. Zdaniem realistów to, co jest widoczne i mierzone za pomocą zmiennych obserwowalnych, dotyczy często ukrytej rzeczywistości, którą należy rozpoznawać jedynie drogą pośrednią [Sagan 2000]. Przykładowo w badaniach marketingowych w zakresie analizy konsumentów wyrażających określone postawy i opinie wobec produktów czy usług pomiar bezpośredni byłby mało skuteczny, a otrzymana na jego podstawie informacja – mało wiarygodna. Oznacza to, że prawdziwa informacja zawarta jest w zmiennej ukrytej, która odzwierciedla pośredni wkład szeregu zmiennych obserwowalnych [Sikorska 2012].

W niniejszym artykule podjęto próbę przybliżenia modelu czynnikowego ze zmiennymi dychotomicznymi w celu rozpowszechnienia tego podejścia analitycznego wśród badaczy akademickich i praktyków. W części pierwszej zaprezentowano proces konstrukcji modelu czynnikowego do analizy danych dychotomicznych. W części drugiej zawarto zaś prostą ilustrację modelu, w ramach której ukazano i zinterpretowano wybrane składniki modelu omówionego w części pierwszej. W ramach przykładu posłużono się modelem jednoczynnikowym w sferze pomiaru postaw młodzieży (młodych konsumentów) wobec hedonistyczno-konsumpcyjnego stylu życia.

2. Znaczenie modeli czynnikowych w badaniu marketingowych zjawisk ukrytych

Analiza ukrytych postaw konsumentów, nieobserwowalnych bezpośrednio i mających znaczenie denotacyjne o głębokim „ładunku informacyjnym”, nie byłaby możliwa bez wpływu statystyki i rozwoju teorii psychometrii. Należy w tym miejscu wyraźnie podkreślić fakt, że pomiar zmiennych ukrytych dokonuje się poprzez wyznaczenie zestawu zmiennych obserwowalnych (wskaźników lub indeksów). Oczywiście charakter wskaźników generujących dany czynnik może się wyrażać dwuzakresowo – poprzez wskaźniki refleksyjne bądź formatywne, za pomocą których oceniany jest charakter wzajemnych relacji przyczynowych. W modelu czynnikowym opartym na klasycznej teorii pomiaru wykorzystuje się przede wszystkim wskaźniki refleksyjne, zaś domeną wskaźników formatywnych jest analiza głównych składowych. Najistotniejszą kwestią badań statystycznych są techniczne aspekty rozpoznania relacji przyczynowych. Kwestie sporne dotyczą na ogół tego, jak odnaleźć najlepszy model analityczny odwzorowujący właściwy

poziom jakości dopasowania różnego typu danych (pozyskiwanych ze skal mocnych bądź skal słabych) ze względu na wzajemne relacje między zmiennymi obserwowalnymi a czynnikami. W zależności od dokładności rozpoznania tych relacji można mówić o tym, że dany model pomiaru przyjmuje postać niezidentyfikowaną, zidentyfikowaną lub nadidentyfikowaną².

Modele czynnikowe (będące elementem szerszego nurtu badań statystycznych, tj. badania współwystępowania) pozwalają nie tylko określić jakość i spójność związków zmiennych, ale także dokonać swoistego rodzaju ilościowej syntezy wniosków w zbyt złożonej strukturze danych. R.L. Gorsuch [1974] podaje np. trzy główne zalety modeli czynnikowych, mianowicie: 1) uproszczoną analizę złożonej macierzy obejmującą wiele wskaźników, 2) możliwość oceny poziomu zmienności związków zachodzących między poszczególnymi wskaźnikami w obrębie badanych czynników, 3) detekcję malejących lub rosnących granicznych obszarów związków, jakie rozdzielają wskaźniki tworzące dany czynnik od pozostałych wskaźników (zmiennych obserwowalnych) i czynników.

3. Konstrukcja modelu czynnikowego na podstawie danych dychotomicznych

Badacze z zakresu marketingu w modelach czynnikowych wykorzystują dane, których pomiar w zakresie cech respondentów dokonuje się na skali porządkowej lub skali interwałowej zgodnie z klasyczną teorią pomiaru. Niewiele osób stosuje w praktyce badawczej modele czynnikowe oparte na danych dychotomicznych, odwołując się m.in. do teorii reakcji na pozycje skali (*item response theory*).

Teoretyczną podbudowę założeń modeli czynnikowych w ujęciu danych dychotomicznych przedstawili w swoich pracach m.in. D. Bartholomew [1980], I. Moustaki i M. Knott [2000], K.G. Jöreskog i I. Moustaki [2001], I. Moustaki i F. Steele [2005]. W procesie konstrukcji modelu czynnikowego ze zmiennymi dychotomicznymi można wykorzystać funkcję logistyczną i model regresji logistycznej w celu określenia zależności każdej z rozpatrywanych zmiennych obser-

² Warto w tym miejscu dodać, że związki te (choćby nawet badane statystycznie za pomocą modelu matematycznego) są zawsze trudne do uznania za wystarczające, tzn. w pełni identyfikowalne. Relacje między płaszczyzną zmiennych obserwowalnych a zmiennymi ukrytymi – czynnikami – opierają się bowiem nie tylko na *stricte* matematycznych zależnościach, ale są uwarunkowane istnieniem teorii abstrakcyjnych (opartych wyłącznie na pojęciach teoretycznych z danej dziedziny aplikacji) oraz teoriami empirycznymi poszukującymi określonych faktów (m.in. poprzez pryzmat działań statystycznych). Stąd rozpoznanie ukrytej struktury danego zjawiska dokonuje się na gruncie teoretycznym (z dokładną znajomością danej dziedziny) oraz na poziomie zjawisk obserwowalnych.

wowalnych (pozycji) wobec generowanego czynnika³ (zob. rys. 1). W funkcji tej dychotomiczna zmienna obserwowalna x_i przyjmuje wartość 1 lub 0 względem danego czynnika f , przy czym wartość oczekiwana zmiennej x_i w ujęciu rozpatrywanego przez badacza czynnika będzie określona tak, że: $P(x_i = 1 | \mathbf{f}) = \pi_i(\mathbf{f})$, gdzie $\pi_i(\mathbf{f})$ oznacza prawdopodobieństwo warunkowe rozpatrywanej zmiennej x_i uzyskującej wartość 1, jeśli zmienna ta odpowiada danemu czynnikowi spośród j -możliwych czynników f_1, \dots, f_j .

Warto w tym miejscu dodać, że model regresji logistycznej, wyznaczany wraz z właściwym składnikiem systematycznym, w którym zachowuje się pod pewnymi względami liniowość [Bartholomew i Knott 1999], jest następujący:

$$\pi_i(x_1, \dots, x_k) = \frac{\exp L}{1 + \exp L}, \quad (1)$$

gdzie:

$$L = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (2)$$

Z kolei składnik losowy to:

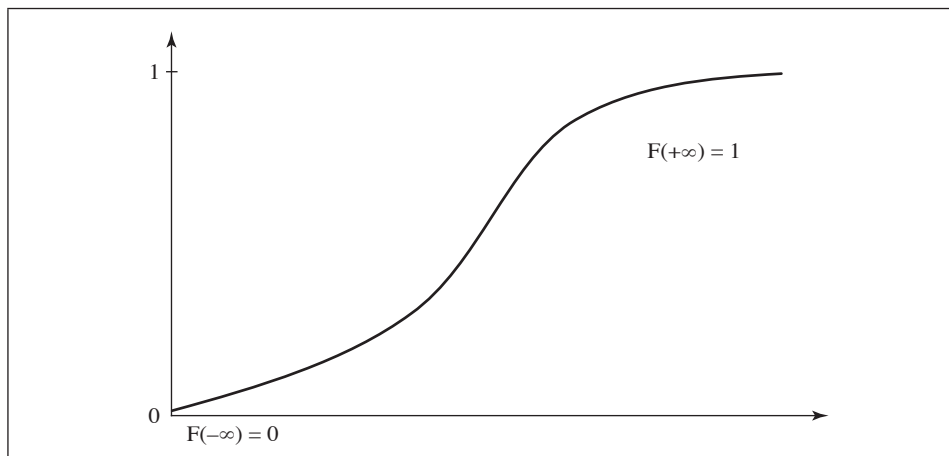
$$y | (x_{i1}, \dots, x_{ik}) \sim \text{Bernoulli}(\pi_i(x_{i1}, \dots, x_{ik})). \quad (3)$$

Składnik systematyczny określa, w jakim stopniu π_i zależy od x , zaś składnik losowy opisuje, jak y zmienia się w obrębie π_i ⁴. Sam rozkład Bernoulliego (zwany też rozkładem dwumianowym)⁵ wyznacza podejście, zgodnie z którym dana zmienna losowa może uzyskać wartość 1 z prawdopodobieństwem π_i i wartość 0 z prawdopodobieństwem równym $(1 - \pi_i)$. W ten sposób model ujęty w zapisie równań (1) i (2) odzwierciedla prawdopodobieństwo tego, że $y = 1$ jest dane funkcją wartości x_i .

³ Przy funkcji logistycznej zakres prawdopodobieństwa odpowiedzi z przedziału $[0, 1]$ przebiega od granicy dolnej $(-\infty)$ do granicy górnej $(+\infty)$ i odpowiada wyrazowi funkcji monotonicznej ze względu na dany rozkład odpowiedzi tworzących określony czynnik f .

⁴ W liniowym modelu regresji składnik losowy jest określany przez rozkład reszt e . W modelu tym zmienna y to zmienna zależna (zmienna objaśniania), a x oznacza zmienną niezależną (zmienną objaśniającą). W modelu regresji logistycznej y jest zmienną opisaną na określonych kategoriach, np. w najprostszym ujęciu zmienna ta może przyjmować wartości binarne typu 0 i 1.

⁵ Rozkład dwumianowy (w Polsce zwany też rozkładem Bernoulliego, choć w krajach anglojęzycznych termin *Bernoulli distribution* odnosi się do rozkładu zero-jedynkowego) to dyskretny rozkład prawdopodobieństwa, opisujący liczbę sukcesów k w ciągu N niezależnych prób, z których każda ma stałe prawdopodobieństwo sukcesu równe p .



Rys. 1. Graficzna prezentacja funkcji logistycznej

Źródło: opracowanie własne.

Funkcja wykładnicza wykorzystana w równaniu (1) $\exp(L) = M$ umożliwia transformację wykładniczą L do postaci logarytmicznej $\log(M) = L$. L określa łączną sumę wyników przy granicy górnej dodatnich odpowiedzi w zakresie rozpatrywanej pozycji (zmiennej obserwowalnej) w ten sposób, że:

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = L = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (4)$$

W ramach dopasowania modelu potrzebna jest jedynie próbka niezależnych obserwacji i zbior zmiennych, jak $\{y_i, x_{i1}, \dots, x_{ik}, i = 1, \dots, n\}$, gdzie wszystkie zmienne x i y będą wzajemnie niezależne. W dalszej kolejności dopasowanie modelu regresji logistycznej przebiega przykładowo za pomocą metody największej wiarygodności. Następnie wyliczane są odpowiednie szacunki α i β , co do których interpretację przeprowadza się w analogiczny sposób jak w standardowym modelu regresji liniowej, przy czym szacunki tego rodzaju dotyczą już wyrażenia $\text{logit}(\pi_i)$, a nie jedynie wyrażenia π_i .

Z powyższych założeń należy wnioskować, że model regresji logistycznej stanowi bardzo ważny punkt odniesienia w początkowej fazie konstrukcji modelu czynnikowego. Sam model logitowy w kontekście proponowanego modelu czynnikowego \mathbf{f} definiuje się już jako:

$$\text{logit} \pi_i(\mathbf{f}) = \log_e \frac{\pi_i(\mathbf{f})}{1-\pi_i(\mathbf{f})} = \alpha_{i0} + \sum_{j=1}^q \alpha_{ij} f_j, \quad (5)$$

gdzie wyrażenie $\pi_i(\mathbf{f})$ to efekt transformacji logitowej, który pozwala sprowadzić model do postaci liniowej w zakresie rozpatrywanych czynników (tj. zmiennych ukrytych). W wyniku przekształcenia modelu (5) otrzymuje się wieloczynnikowy model w następującej formie:

$$\pi_i(\mathbf{f}) = \frac{\exp\left(\alpha_{i0} + \sum_{j=1}^q \alpha_{ij0} f_j\right)}{1 + \exp\left(\alpha_{i0} + \sum_{j=1}^q \alpha_{ij0} f_j\right)}. \quad (6)$$

Jeśli zaś zakładany jest tylko jeden czynnik $j = 1$, wówczas zapis modelu (6) będzie odpowiadać modelowi jednoczynnikowemu w następującej formule [Bock 1972, Moustaki, Jöreskog i Mavridis 2004]:

$$\pi_i(f_1) = \frac{\exp(\alpha_{i0} + \alpha_{i1} f_1)}{1 + \exp(\alpha_{i0} + \alpha_{i1} f_1)}. \quad (7)$$

W literaturze – zgodnie z przesłankami wynikającymi z teorii reakcji na odpowiedzi – model jednoczynnikowy znany jest także pod nazwą modelu dwuparametrycznego. $\pi_i(f_1)$ traktuje się jako prawdopodobieństwo udzielenia przez respondenta odpowiedzi „tak” zgodnie z jego stopniem zdolności i poziomem wiedzy potrzebnej do rozwiązania danego zadania testowego.

4. Interpretacja parametrów modelu

W modelu czynnikowym dla zmiennych dychotomicznych dla każdej pozycji otrzymuje się $q + 1$ parametrów do oszacowania, wartość stałą α_{i0} oraz ładunki czynnikowe $\alpha_{i1}, \dots, \alpha_{iq}$. Im wyższe są wartości ładunków czynnikowych α_{ij} , tym większy jest efekt j -czynnika w zakresie prawdopodobieństwa uzyskania pozytywnej odpowiedzi na daną pozycję. Jednocześnie im wyższa jest wartość α_{ij} dla danej pozycji, tym większa jest różnica w prawdopodobieństwach uzyskania prawidłowej/pozytywnej odpowiedzi pomiędzy dwiema badanymi jednostkami, znajdującymi się w pewnej odległości od siebie na ukrytych wymiarach. W rezultacie łatwiej można pokazać różnice pomiędzy dwiema jednostkami na podstawie ich odpowiedzi względem ocenianej pozycji. Jeśli wartości poszczególnych pozycji (nakładających się na dany czynnik) będą duże (np. wykraczające poza

3 lub 10), to będzie to oznaczać, że mamy do czynienia z bardzo silnym efektem progowym tych pozycji⁶.

Mankamentem modeli czynnikowych dla zmiennych dychotomicznych są surowe ładunki czynnikowe. Ładunki te nie mają umownie przyjętych wartości w zakresie minimalnej (dolnej) i maksymalnej (górnjej) granicy, przez co interpretacja zależności pomiędzy zmiennymi obserwowalnymi a czynnikami jest utrudniona. Wartości takich ładunków nie można więc porównać na wzór współczynników korelacji, które mieszczą się w przedziale od 0 do 1, tak jak ma to miejsce w klasycznej analizie czynnikowej. W związku z powyższym badacz musi je w pierwszej kolejności wystandaryzować⁷.

Warto też dodać, że jeśli ładunki czynnikowe odznaczają się zbyt wysokimi wartościami błędów standardowych, to ich wartości są, po pierwsze, słabo określone przez model i najprawdopodobniej źle przyjęte kryteria doboru próby, a po drugie – świadczą o słabej jakości wygenerowanego czynnika.

5. Wskaźniki dobroci dopasowania danych do modelu czynnikowego

Po zakończeniu etapu związanego z interpretacją parametrów modelu należy ocenić dobroć dopasowania danych dychotomicznych do modelu czynnikowego. W tym celu można skorzystać z dwóch testów. Pierwszy test nosi nazwę globalnego testu dobroci dopasowania danych – G^2 i wyraża się wzorem:

$$G^2 = 2 \sum_{r=1}^{2^p} O(r) \log_e \frac{O(r)}{E(r)}, \quad (8)$$

gdzie r oznacza zaakceptowany przez badacza w analizie wzorzec odpowiedzi, a $O(r)$ i $E(r)$ – obserwowane i oczekiwane częstości według wzorca odpowiedzi r .

Test ten został zaproponowany przez D.J. Bartholomewa i P. Tzamourianego [1999]. Umożliwia on porównanie wartości obserwowanych częstości z wartościami oczekiwanymi. Dopasowanie modelu dokonuje się poprzez obliczenie wartości parametrów (odzwierciedlających położenie względem siebie rozkładów rozpatrywanych zmiennych). Im mniejsze są te wartości, tym lepsze jest dopasowanie modelu (tym większa jest dobroć dopasowania). Małe różnice oznaczają, że

⁶ Wysokie wartości współczynników dyskryminacji świadczą jednocześnie o tym, że krzywe charakterystyczne pozycji odznaczają się stromym nachyleniem ze względu na znaczne różnice w dwóch binarnych wariantach odpowiedzi.

⁷ Zob. punkt dotyczący ilustracji modelu jednoczynnikowego.

związki pomiędzy wszystkimi parami poszczególnych pozycji odpowiedzi zmiennych obserwowalnych są dobrze opisane i przewidziane przez model czynnikowy.

Alternatywnie można także przyjąć test oparty na Pearsonowskiej statystyce chi-kwadrat:

$$\chi^2 = \sum_{r=1}^{2^p} \frac{(O(r) - E(r))^2}{E(r)}. \quad (9)$$

W obu wariantach (8 i 9) jeśli model odzwierciedla poziom dopasowania do danych, to wartości obu statystyk G^2 i χ^2 powinny być porównywalne. Statystyki G^2 i χ^2 traktowane są jako stały element badania rozkładu reszt kategorii zmiennych obserwowalnych. Do analizy danych dychotomicznych można wykorzystać podejście, zgodnie z którym reszty danej zmiennej rozpatrywane będą niezależnie od pozostałych reszt – jako statystyka G^2 i χ^2 z 1 stopniem swobody. Wówczas wartości statystyk G^2 i χ^2 większe niż 4 będą oznaczać słabe dopasowanie danych do modelu. Analiza reszt może być wykorzystana w celu identyfikacji oraz ewentualnej eliminacji tych zmiennych obserwowalnych, które nie znajdują w modelu czynnikowym dobrego dopasowania.

6. Ilustracja modelu jednoczynnikowego w zakresie badania postaw młodzieży wobec hedonistyczno-konsumpcyjnego stylu życia

W ramach ilustracji praktycznej funkcjonalności modelu jednoczynnikowego do analizy wybrano dane empiryczne pochodzące ze studium przeprowadzonego przez autora [Tarka 2010]. Zbiór danych obejmował 232 respondentów oraz 4 pozycje testowe – ujęte w formie dychotomicznych kategorii odpowiedzi: tak (1), nie (0) – z uwagi na następujące stwierdzenia:

1. Zakupy wpływają korzystnie na moje samopoczucie.
2. Jeśli nie stać mnie na zakupy produktów z powodu braku pieniędzy, zaciągam pożyczki.
3. Jeśli dysponuję już pieniędzmi, to najczęściej odwiedzam galerie handlowe.
4. Wydawanie pieniędzy (podczas zakupów w galerii handlowej) sprawia mi ogromną radość.

Przedmiotem analizy były postawy młodzieży wobec hedonistyczno-konsumpcyjnego stylu życia. Obliczenia przeprowadzono w programie LAMI (*Latent Model Interface*) w module Genlat. Jest to program, który umożliwia generowanie czynników na podstawie danych dychotomicznych. Jego słabą stroną jest

ograniczenie w zakresie liczby generowanych czynników. Możliwe jest bowiem wyszczególnienie maksymalnie dwóch czynników⁸.

Tabela 1. Zapis częstości występowania poszczególnych wzorców odpowiedzi

| Wzorec odpowiedzi | Częstości obserwowane | Częstości oczekiwane |
|-------------------|-----------------------|----------------------|
| 1 1 1 1 | 101 | 102,22 |
| 0 0 0 0 | 63 | 61,86 |
| 0 1 1 1 | 14 | 14,12 |
| 0 0 1 1 | 14 | 8,50 |
| 0 0 0 1 | 8 | 10,07 |
| 1 1 1 0 | 7 | 5,83 |
| 0 0 1 0 | 5 | 8,91 |
| 0 1 0 0 | 4 | 3,35 |
| 0 1 1 0 | 7 | 6,26 |
| 1 0 1 1 | 3 | 3,06 |
| 0 1 0 1 | 1 | 0,00 |
| 1 1 0 1 | 3 | 1,40 |
| 1 1 0 0 | 1 | 0,00 |
| 1 0 0 0 | 1 | 0,00 |
| – | 232 | 225,59 |

Źródło: opracowanie na podstawie własnych badań w programie LAMI.

Po zakończeniu obliczeń, na podstawie wyników zamieszczonych w tabeli 1 (przedstawiającej częstości występowania odpowiedzi według określonych wariantów odpowiedzi), można wstępnie określić dwa charakterystyczne wzorce odpowiedzi: 1 1 1 1 („tak”) i 0 0 0 0 („nie”). Oba warianty obejmują w sumie największe liczebności obserwacji. Pozostałe wzorce odpowiedzi stanowią jedynie marginalny wkład informacyjny do konstruowanego modelu. Odsetek odpowiedzi z uwzględnieniem obydwu kategorii odpowiedzi (tj. 1 i 0) według wszystkich czterech pozycji testowych wskazuje na przewagę wartości 1 (zob. tabela 2), co

⁸ W programie w pliku wejściowym zdefiniowano w pierwszej kolejności zakres danych i poziom obliczeń. Polegało to m.in. na tym, że w wierszu pierwszym (oznaczonym N, NPB) określono liczbę respondentów N = 232 oraz liczbę zmiennych obserwowalnych mierzonych na podstawie danych binarnych (4) – (NPB). Z kolei w wierszu drugim, oznaczonym NFAC, INIT, ITER, PREC, SCOR, wyznaczono liczbę czynników (1) do wygenerowania w toku prowadzonej analizy – (NFAC), wybrano procedurę obliczeń w ramach czynnika (INIT), określono liczbę iteracji (maksimum 2000) – (ITER), określono poziom precyzji wyników (0,0000001) – (PREC) oraz przyjęto ustawione przez program domyślnie opcje wydruku wyników – (SCOR).

oznacza, że rozpatrywane 4 pozycje testowe są zbieżne i dostarczają informacji w ramach jednego wspólnego wymiaru. Ładunki czynnikowe tego modelu zaprezentowano w tabeli 3.

Tabela 2. Rozkłady odpowiedzi według poszczególnych pozycji

| Wyszczególnienie | Pozycja testowa 1 | Pozycja testowa 2 | Pozycja testowa 3 | Pozycja testowa 4 |
|------------------|-------------------|-------------------|-------------------|-------------------|
| Kategoria 1 | 0,55 | 0,60 | 0,65 | 0,62 |
| Kategoria 0 | 0,45 | 0,40 | 0,35 | 0,38 |

Źródło: opracowanie na podstawie własnych badań w programie LAMI.

Tabela 3 przedstawia wartości Alpha (surowe ładunki czynnikowe) dla każdej z czterech rozpatrywanych pozycji testowych. Dla przypomnienia – im wyższe są wartości tych ładunków, tym większy jest efekt danego czynnika w zakresie prawdopodobieństwa jego połączenia z daną pozycją testową. W analizowanym przykładzie wartości tych ładunków są bardzo wysokie i dodatnie (zob. kolumnę Alpha(1, I)), przez co jednoznacznie wskazują na jeden wspólny ukryty czynnik.

Tabela 3. Ładunki czynnikowe modelu w zakresie poszczególnych pozycji

| Pozycje | Alpha (0, I) | Błąd standardowy | Alpha (1, I) | Błąd standardowy | $P(X = 1 Z = 0)$ | Standaryzowana Alpha |
|---------|--------------|------------------|--------------|------------------|------------------|----------------------|
| 1 | -0,04 | 0,34 | 6,79 | 0,84 | 0,49 | 0,99 |
| 2 | 1,49 | 0,29 | 6,22 | 0,74 | 0,82 | 0,99 |
| 3 | 2,29 | 0,66 | 5,81 | 0,42 | 0,91 | 0,98 |
| 4 | 1,18 | 0,36 | 5,52 | 0,54 | 0,77 | 0,97 |

Miara dobroci dopasowania danych do modelu – $G = 76,85\%$.

Źródło: opracowanie na podstawie własnych badań w programie LAMI.

W modelach czynnikowych dla zmiennych dychotomicznych surowe ładunki czynnikowe nie podlegają interpretacji, co oznacza, że nie można ich ze sobą w prosty sposób porównać. Ich wartości wykraczają poza umownie przyjęty przedział współczynników, np. korelacji od 0 do 1. Stąd potrzebna jest standaryzacja, w wyniku której otrzymuje się możliwość interpretacji ładunków czynnikowych w zakresie rozpatrywanych pozycji w relacji do badanego czynnika. Ostatnia kolumna w tabeli 3 zawiera wystandaryzowane wartości ładunków czynnikowych, które mieszczą się w przedziale od 0,97 do 0,99 i świadczą o wysokim poziomie ładunków czynnikowych opisujących poziom korelacji między daną pozycją (zmienną obserwowalną) a zmienną ukrytą (czynnikiem).

W praktyce badawczej zdarza się i tak, że duże wartości ładunków czynnikowych dla niektórych pozycji mogą być obciążone dużymi błędami standardowymi, co wynika z rozproszenia wartości odpowiedzi, jakie respondenci przydzielają poszczególnym pozycjom testowym. Wartości te w konsekwencji wpływają negatywnie na sam konstruowany czynnik⁹. Z zaprezentowanego przykładu wynika, że taką pozycją jest m.in. pozycja nr 1 „Zakupy wpływają korzystnie na moje samopoczucie”, której błąd standardowy wyniósł 0,84 przy wartości $\text{Alpha}(1, 1)$ równej 6,79. Ponadto w przypadku pozycji nr 1 wartość oszacowanego prawdopodobieństwa odpowiedzi ($P(X = 1|Z = 0)$) wyniosła jedynie 0,49. Oszacowane prawdopodobieństwa określają w tym wypadku, na ile jednostki badane w próbie są zgodne w zakresie pozytywnych odpowiedzi względem czterech rozpatrywanych pozycji.

Ostatecznie, kierując się miarą dobroci dopasowania danych do modelu, tj. rozpatrując wskaźnik G (który osiągnął poziom 76,85%), można stwierdzić, że jakość tego modelu w układzie czterech pozycji testowych kształtuje się na dobrym poziomie. Przeanalizowane w ten sposób pozycje należy traktować jako odpowiednie składowe czynnika charakteryzującego „hedonistyczno-konsumpcyjny styl życia” młodzieży.

7. Podsumowanie

Cechy poszczególnych jednostek badawczych mogą być obserwowalne i nieobserwowalne. Pierwsze wyróżniają się poprzez fakt, że można je zmierzyć, drugie natomiast dotyczą ukrytych postaw, opinii, emocji i podświadomych wyborów jednostek, przez co ich bezpośredni pomiar jest niemożliwy. Analiza zmiennych ukrytych (a tym samym rozpoznanie zjawisk rynkowych o charakterze ukrytym) jest trudnym, ale też bardzo pożądanym przedsięwzięciem. Sama analiza oparta na modelach czynnikowych dla zmiennych dychotomicznych jest bardzo istotna dla dalszego rozwoju badań i analiz marketingowych. Jest ona szczególnie przydatna wobec rosnącej liczby danych jakościowych (pochodzących z takich obszarów analiz, jak pomiar reakcji konsumentów na reklamę, deklaracji konsumenckich, wzorców postaw i zachowań). Dane tego typu dotyczą coraz częściej pomiaru cech jednostek z wykorzystaniem skal słabych, tj. skal porządkowych i skal nominalnych. Zaprezentowany w pracy model czynnikowy dla zmiennych dychotomicznych można zatem uznać za niezwykle przydatny w opisie nieobserwowalnych zjawisk rynkowych na gruncie tradycyjnego pomiaru, tj. skali nominalnej.

⁹ Czynnik taki jest niespójny i tym samym słabo zdefiniowany przez model.

Literatura

- Bartholomew D.J. [1980], *Factor Analysis of Categorical Data*, „Journal of the Royal Statistical Society”, vol. 42, nr 3.
- Bartholomew D.J., Knott M. [1999], *Latent Variable Models and Factor Analysis*, wyd. 2, Arnold, London.
- Bartholomew D.J., Tzamourani P. [1999], *The Goodness-of-fit of Latent Trait Models in Attitude Measurement*, „Sociological Methods and Research”, vol. 27, nr 4.
- Bock R.D. [1972], *Estimating Item Parameters and Latent Ability When Responses Are Scored in Two or More Nominal Categories*, „Psychometrika”, vol. 37, nr 1.
- Gorsuch R.L. [1974], *Factor Analysis*, W.B. Saunders Co., Toronto.
- Jöreskog K.G., Moustaki I. [2001], *Factor Analysis of Ordinal Variables: A Comparison of Three Approaches*, „Multivariate Behavioral Research”, vol. 36, nr 3.
- Moustaki I., Jöreskog K.G., Mavridis D. [2004], *Factor Model for Ordinal Variables with Covariate Effects on the Manifest and Latent Variables: A Comparison LISREL and IRT Approaches*, „Structural Equation Modeling”, vol. 11, nr 4.
- Moustaki I., Knott M. [2000], *Generalized Latent Trait Models*, „Psychometrika”, vol. 65, nr 3.
- Moustaki I., Steele F. [2005], *Latent Variable Models for Mixtures of Categorical and Duration Responses, with an Application to Fertility Preferences and Family Planning in Bangladesh*, „Statistical Modelling”, vol. 5, nr 4.
- Sagan A. [2000], *Wybrane problemy identyfikacji i pomiaru struktur ukrytych*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, nr 543, Kraków.
- Sikorska I. [2012], *Analiza zmiennych ukrytych [w:] Zaawansowane metody analiz statystycznych*, red. E. Frątczak, Szkoła Główna Handlowa w Warszawie, Warszawa.
- Tarka P. [2010], *Latent Variable Models – Issues on Measurement and Finding Exact Constructs*, „Przegląd Statystyczny”, nr 4.

Factor Analysis for Binary Data in the Analysis of the Latent Structure of Market Phenomena

The article describes a number of theoretical and practical issues pertaining to the factor analysis model and its application in the field of marketing research. The first part discusses the dependence of marketing research on the theory of measurement e.g., the course of research development in the sphere of latent events analysis. The next part looks at factorial modeling in research on consumer attitudes and opinions, followed by a discussion of the possibilities for binary data analysis using the latent trait model. Finally, using empirical data, the author presents a one-factor model to measure youth (young consumers) attitudes towards a hedonistic-consumer lifestyle.

Keywords: factor analysis, latent variables, binary data, nominal scale.

Mirosława Kaczmarek

Katedra Badań Rynku i Usług
Uniwersytet Ekonomiczny w Poznaniu

Problemy stosowania skali Likerta w pomiarze jakości użytkowej serwisów internetowych

Streszczenie

W badaniu jakości użytkowej (użyteczności) serwisów internetowych wykorzystywane są różne rodzaje metod, np. testy użytkownika, *eyetracking*, *clicktracking*. Nadal jednak ważną metodą badań jest wywiad osobisty, a narzędziem badawczym – kwestionariusz, szczególnie wtedy, gdy pomiar jakości prowadzony jest z uwzględnieniem normy ISO 9241-11:1998. Utworzona skala pomiaru jakości użytkowej serwisu internetowego powinna z jednej strony odwoływać się do wyznaczników użyteczności wynikających z normy, takich jak: skuteczność, wydajność oraz satysfakcja użytkownika, a z drugiej – uwzględniać złożony charakter architektury serwisu.

Artykuł prezentuje rezultaty analizy wymiarowości skali z zastosowaniem metody głównych składowych oraz ocenę rzetelności wyodrębnionych skal jednowymiarowych na podstawie współczynnika α -Cronbacha.

Słowa kluczowe: jakość użytkowa, rzetelność skali, bankowość internetowa, analiza czynnikowa.

1. Wprowadzenie

Ocena jakości użytkowej serwisu internetowego może być przeprowadzana w dwojaki sposób: za pomocą metod, które w bezpośredni sposób pozwalają dokonać pomiaru zachowania użytkownika danego serwisu, takich jak *eyetracking*,

clicktracking czy testy użytkownika, oraz za pomocą metod, których celem jest zbieranie opinii i pomiar satysfakcji osób korzystających z danego serwisu. W tym drugim przypadku ważną metodą badań pozostaje wywiad osobisty, a narzędziem badawczym – kwestionariusz, szczególnie gdy pomiar jakości funkcjonowania serwisu prowadzony jest z uwzględnieniem normy ISO 9241-11:1998 *Wskazówki dotyczące użyteczności* [ISO 9241-11:1998...].

Celem artykułu jest określenie możliwości zastosowania skali Likerta w pomiarze jakości użytkowej serwisu internetowego banku. Obok wyodrębnienia kategorii nominalnych, pozwalających mierzyć postawy wobec określonego zjawiska, równie istotne jest przyporządkowanie utworzonym stwierdzeniom odpowiedniej skali ocen. Na ogół stosuje się do tego celu 5- lub 7-stopniową porządkową skalę ocen. W przypadku stwierżeń określających postawy wobec serwisu transakcyjnego banku zastosowanie takiego rozwiązania wydaje się nie do końca prawidłowe, na co wskazują definicje postawy. Do najczęściej przytaczanych należą definicje strukturalne, w myśl których postawa to „względnie trwała dyspozycja jednostki do określonego zachowania się wobec danego przedmiotu, wynikająca z poglądów, uczuć i dążeń danej jednostki odnoszących się do przedmiotu postawy” [Turowski 1993, s. 50; Hilgard 1972, s. 834; Mika 1972, s. 65; Nowak 1973, s. 23]. Postawy mogą mieć zróżnicowany zakres, przyjmują określony kierunek i intensywność, a ich cechą charakterystyczną jest trwałość w odniesieniu do danego przedmiotu, tzn. niewielka zmienność w czasie (w przypadku jej braku byłoby to nastawienie).

W odniesieniu do serwisu internetowego banku jako przedmiotu postawy możliwe jest wystąpienie sytuacji, w której respondentowi będzie trudno wypowiedzieć się na temat jego funkcjonalności na skutek braku wiedzy o zasadach jego budowy i sposobie działania. Jest to spowodowane bardzo zróżnicowaną znajomością pojęć z zakresu technologii informacyjnej i bankowości internetowej – znajomość ta jest wyraźnie słabsza wśród osób starszych, rzadziej korzystających z Internetu i posiadających mniejsze doświadczenie związane z obsługą kont bankowych online. Ta sytuacja powinna być uwzględniona na etapie konstrukcji skali pomiarowej. Powstaje zatem pytanie, jak powinien postąpić badacz, stykając się w przypadku części respondentów z problemem pomiaru postaw wobec wybranych aspektów jakości użytkowej serwisu bankowego – czy zupełnie zrezygnować z danej kategorii skali, a jeśli tak, jak wpłynie to na jej rzetelność? Innym pytaniem, które nasuwa się w kontekście prowadzenia analizy danych, jest pytanie o sposób traktowania braku odpowiedzi. Ten problem nabiera znaczenia, gdy w analizie rzetelności skali ma być stosowana metoda głównych składowych i brak danych nawet dla jednej zmiennej wyklucza dany przypadek z analizy.

2. Elementy serwisu internetowego i wyznaczniki jego jakości użytkowej

Budowa skali Likerta przebiega w kilku etapach. Pierwszym jest utworzenie stwierdzeń dotyczących postaw wobec mierzonego problemu [Sagan 2004, s. 89], a więc w tym przypadku – jakości użytkowej (inaczej użyteczności) serwisu internetowego.

Serwis internetowy banku jest kompleksową strukturą charakteryzowaną za pomocą takich elementów, jak zawartość informacyjna, *design* i *layout*, struktura i nawigacja oraz komunikacja z użytkownikiem. Zawartość serwisu internetowego banku jest określana przez ogół informacji udostępnianych klientom na poszczególnych stronach witryny bankowej, które podobnie jak strony w innych serwisach oferujących dostęp do usług online powinny być przede wszystkim zwięzłe i komunikatywne. *Design* i *layout* (prezentacja graficzna) określają wygląd serwisu internetowego – jest on kształtowany przez układ i wielkość różnych elementów składających się na szatę graficzną serwisu, takich jak: kolory, krój i rozmiar czcionek, symbole, znaki graficzne oraz zdjęcia. Struktura serwisu internetowego, czyli układ i powiązania pomiędzy pojedynczymi stronami serwisu, powinna pozwolić użytkownikowi na odnalezienie określonych informacji w konkretnym miejscu. Poza przejrzystą strukturą istotne znaczenie ma intuicyjna nawigacja serwisu internetowego, a więc łatwość poruszania się po nim.

Równie wieloaspektową kategorią jest jakość użytkowa (użyteczność) – jak najczęściej w języku polskim tłumaczy się termin *usability* [Michalski 2008, s. 72–73]. Pojęcie to zostało zdefiniowane przez Międzynarodową Organizację Standardyzacyjną w normie *Wskazówki dotyczące użyteczności* [ISO 9241-11:1998...]. Zgodnie z tą definicją podstawowymi wyznacznikami użyteczności są: skuteczność – zgodność uzyskanych efektów z zamierzonymi, wydajność – relacja poniesionych nakładów (zużytych zasobów) do uzyskanych efektów oraz satysfakcja – zadowolenie użytkownika wynikające z używania produktu.

Efektom rozwoju Internetu jako kanału komunikacji z użytkownikami było pojawienie się w literaturze przedmiotu terminu *web usability* oraz wyodrębnienie wyznaczników użyteczności odnoszących się bezpośrednio do witryny internetowej, popularnie nazywanych heurystykami Nielsena [Nielsen 1993, s. 26]. Nauczenie się obsługi interfejsu witryny internetowej powinno być łatwe, tak aby użytkownik, który wejdzie na nią po raz pierwszy, mógł wykonać swoje zadanie (*learnability*). Ponadto serwis internetowy powinien być zaprojektowany w sposób zapewniający wysoką wydajność użytkownika podczas realizacji zadania (*efficiency*), obsługa serwisu powinna być łatwa do zapamiętania (*memorability*), a błędy, które użytkownik może popełnić w trakcie korzystania z witryny, powinny być nieliczne i łatwe do wyeliminowania (*errors*). Obsługa witryny inter-

netowej powinna być także wygodna, tak aby podczas interakcji użytkownik był zadowolony (*satisfaction*).

Zaprezentowana w tej krótkiej charakterystyce złożoność zagadnienia, jakim jest jakość użytkowa serwisu internetowego, wydaje się uzasadniać zastosowanie skali Likerta do pomiaru postaw klientów korzystających z usług bankowych poprzez serwis transakcyjny.

3. Zastosowanie skali Likerta w badaniach ilościowych jakości użytkowej serwisu internetowego banku

Potwierdzeniem trafności wyboru skali Likerta w ewaluacji jakości użytkowej serwisu internetowego banku są pozytywne doświadczenia badaczy związane z zastosowaniem tej skali do oceny użyteczności oprogramowania komputerowego. Skala Likerta była wykorzystywana w kwestionariuszach utworzonych przez Human Factors Research Group działającą na Uniwersytecie w Cork. Należą do nich: *Computer User Interface Satisfaction Inventory* [Kirakowski i Corbett 1988, s. 329–338], *Software Usability Measurement Inventory* [Kirakowski i Corbett 1993, s. 210–212] oraz *Web Analysis and Measurement Inventory* [Kirakowski, Claridge i Whitehand 1998]. W ramach każdego z tych narzędzi wyodrębnione zostały subskale pozwalające ocenić poziom satysfakcji użytkowników w zakresie różnych aspektów jakości użytkowej, np. wyglądu, wydajności czy łatwości obsługi. Warto przywołać także trzy inne kwestionariusze stosowane do ewaluacji oprogramowania, które w swojej konstrukcji bezpośrednio nawiązują do wytycznych sformułowanych w normie ISO 9241 w części 10. Są to: *ISONORM 9241/10* [Prümper i Anft 1993, s. 145–156; Prümper 1997 s. 253–262], *IsoMetrics* [Gediga, Hamborg i Düntsch 1999, s. 151–164] oraz *ErgoNorm* [Dzida *et al.* 2000]. Przydatność w badaniu jakości użytkowej oprogramowania skal zastosowanych w tych narzędziach została potwierdzona analizą rzetelności przeprowadzoną metodą α -Cronbacha¹.

Na potrzeby oceny jakości serwisu internetowego banku realizowanej z wykorzystaniem kwestionariusza wywiadu przygotowana została skala złożona z 17 pozycji. Poszczególne stwierdzenia dotyczące jakości użytkowej były oceniane nie za pomocą pięciostopniowej skali porządkowej, jak to ma zwykle miejsce

¹ Ocena rzetelności dotyczyła subskal utworzonych dla siedmiu kategorii wymagań, które zgodnie z zapisami normy powinno spełniać oprogramowanie, takich jak: możliwość oddziaływania na nie (sterowalność) podczas pracy, wsparcie użytkownika w realizacji zadań, zrozumiałość użytych pojęć, symboli i objaśnień, łatwość uczenia się dostępnych w oprogramowaniu funkcji, zgodność z oczekiwaniami i przyzwyczajeniami użytkownika, tolerancja popełnianych przez użytkownika błędów oraz możliwość dostosowania do jego indywidualnych potrzeb.

w przypadku skali z werbalnie oznaczonymi stopniami, lecz skali z sześcioma wariantami odpowiedzi. Zastosowana skala miała następującą postać: 1 – całkowicie nie zgadzam się, 2 – nie zgadzam się, 3 – ani się nie zgadzam, ani się zgadzam, 4 – zgadzam się, 5 – całkowicie się zgadzam, 6 – nie wiem. Celowość dodania wariantu „nie wiem” potwierdził przeprowadzony wcześniej eksperyment, w którym 45 respondentów najpierw ustosunkowało się do stwierdzeń dotyczących jakości użytkowej serwisu internetowego banku za pomocą 5-stopniowej skali porządkowej, a po upływie godziny ta sama grupa respondentów udzielała odpowiedzi, korzystając z kwestionariusza, w którym skalę ocen rozszerzono o wariant „nie wiem”. Osoby, które w drugiej wersji kwestionariusza wskazały wariant „nie wiem”, w pierwszym kwestionariuszu wskazywały relatywnie najczęściej odpowiedź „raczej nie zgadzam się”, a następnie „raczej zgadzam się” i „ani się nie zgadzam, ani się zgadzam”.

Rezultaty eksperymentu zostały uwzględnione w konstrukcji kwestionariusza w badaniu przeprowadzonym w gospodarstwach domowych w województwie wielkopolskim. Badanie zrealizowano w okresie maj–czerwiec 2012 r. metodą wywiadu osobistego na próbie 280 gospodarstw domowych korzystających z bankowości internetowej. Największy odsetek wskazań dla wariantu „nie wiem” wystąpił w przypadku stwierdzenia „Istnieje możliwość dostosowania wyglądu serwisu do potrzeb klienta” (wybrała go ponad połowa respondentów). Z tego względu stwierdzenie to zostało wyeliminowane z dalszej analizy. Odpowiedź „nie wiem” na poziomie 8–12% ogółu udzielonych odpowiedzi odnotowano także w odniesieniu do takich stwierdzeń, jak: „Sposób logowania do serwisu internetowego banku jest w pełni bezpieczny”, „W serwisie dostępna jest szczegółowa informacja o ofercie i warunkach oferowanych usług”, „Wskazówki, które można wyświetlić dla poszczególnych usług bankowych są przydatne” oraz „Wyświetlane komunikaty o błędach są jednoznaczne i zrozumiałe”. W przypadku tych stwierdzeń odpowiedź „nie wiem” została potraktowana jako brak danych wynikający z braku wiedzy respondenta, a tym samym braku wystarczających kompetencji do zajęcia stanowiska (określenia postawy) wobec postawionego problemu (stwierdzenia). Innymi słowy, jeśli respondent nie miał do czynienia ze wskazówkami dotyczącymi korzystania z usług i operacji bankowych, nie mógł ocenić ich przydatności, a więc udzielić miarodajnej odpowiedzi. Formalnie więc w utworzonej bazie danych wariant „nie wiem” został potraktowany jako brak danych. Analizę głównych składowych, a następnie rzetelności skali przeprowadzono w dwóch wariantach: po eliminacji braków danych (wariant 1), poprzez zastąpienie braków danych wartościami uzyskanymi w wyniku zastosowania metody k -najbliższych sąsiadów (wariant 2).

Analiza głównych składowych pozwala określić, w ilu wymiarach jest mierzona jakość użytkowa serwisów internetowych banków, czyli jakie subskale można wyodrębnić w ramach konstruktów składającego się z 16 pozycji. Uwzględ-

Tabela 1. Analiza wymiarowości skali z uwzględnieniem głównych składowych

| Pozycje skali | Eliminacja odpowiedzi „nie wiem” | | | | Metoda k-najbliższych sąsiadów | | | |
|--|----------------------------------|-------------|-------------|-----------|--------------------------------|-------------|-------------|-------------|
| | czynnik 1 | czynnik 2 | czynnik 3 | czynnik 4 | czynnik 1 | czynnik 2 | czynnik 3 | czynnik 4 |
| Operacje bankowe w serwisie wykonuje się za pomocą kilku kliknięć | 0,74 | 0,05 | 0,25 | 0,31 | 0,74 | 0,23 | 0,07 | 0,25 |
| Operacje bankowe w serwisie wykonuje się bardzo szybko | 0,70 | 0,26 | 0,31 | 0,29 | 0,75 | 0,29 | 0,16 | 0,22 |
| Operacje bankowe w serwisie wykonuje się bardzo łatwo | 0,87 | 0,21 | 0,13 | 0,11 | 0,84 | 0,13 | 0,21 | 0,10 |
| Stosowane znaki, ikony i symbole są komunikatywne | 0,63 | 0,36 | 0,35 | -0,24 | 0,47 | -0,07 | 0,56 | 0,31 |
| Wielkość i krój czcionki pozwalają bez problemu odczytać informacje na stronie | 0,63 | 0,31 | 0,33 | 0,00 | 0,58 | 0,02 | 0,44 | 0,35 |
| Zakres usług w serwisie internetowym banku zaspokaja moje potrzeby | 0,25 | 0,62 | 0,08 | 0,45 | 0,34 | 0,55 | 0,26 | 0,15 |
| W serwisie dostępna jest szczegółowa informacja o ofercie i warunkach usług | 0,06 | 0,56 | 0,35 | 0,29 | -0,02 | 0,45 | 0,37 | 0,43 |
| Szata graficzna serwisu bankowego jest estetyczna i przyjazna dla oka | 0,21 | 0,81 | 0,09 | 0,08 | 0,11 | 0,32 | 0,80 | 0,01 |
| Układ elementów na stronie serwisu jest spójny i logiczny | 0,36 | 0,65 | 0,36 | -0,02 | 0,35 | 0,11 | 0,71 | 0,24 |
| Wskazówki, które można wyświetlić dla usług bankowych, są przydatne | 0,18 | 0,26 | 0,65 | 0,12 | 0,11 | 0,16 | 0,37 | 0,54 |
| Informacje i opisy usług bankowych są napisane zrozumiałym językiem | 0,25 | 0,31 | 0,67 | 0,20 | 0,29 | 0,25 | 0,23 | 0,65 |
| Układ menu serwisu bankowego pozwala łatwo odnaleźć poszczególne usługi | 0,40 | 0,30 | 0,64 | 0,09 | 0,51 | 0,09 | 0,30 | 0,56 |

| Pozycje skali | Eliminacja odpowiedzi „nie wiem” | | | | Metoda <i>k</i> -najbliższych sąsiadów | | | |
|---|----------------------------------|-----------|-------------|-------------|--|-------------|-----------|-------------|
| | czynnik 1 | czynnik 2 | czynnik 3 | czynnik 4 | czynnik 1 | czynnik 2 | czynnik 3 | czynnik 4 |
| Wyświetlane komunikaty o błędach są zrozumiałe i jednoznaczne | 0,23 | -0,03 | 0,78 | 0,08 | 0,25 | 0,05 | -0,06 | 0,76 |
| Sposób logowania do serwisu internetowego banku jest bezpieczny | 0,15 | 0,18 | 0,18 | 0,84 | 0,16 | 0,83 | 0,03 | 0,08 |
| Nawigacja po serwisie banku jest całkowicie intuicyjna | 0,48 | 0,19 | 0,41 | 0,37 | 0,39 | 0,36 | 0,24 | 0,23 |
| Skumulowany odsetek wyjaśnianej wariancji | 46,98 | 54,78 | 61,16 | 67,42 | 43,14 | 50,63 | 57,16 | 63,27 |

Objaśnienie: pogrubioną czcionką oznaczono ładunki czynnikowe o najwyższej wartości dla danej pozycji skali (pominięto przypadki, gdy stwierdzenie nie jest specyficzne dla żadnego z czynników).

Źródło: opracowanie własne.

niając kryterium Kaisera i stosując rotację Varimax, wyodrębniono cztery czynniki określające cztery subskałe (tabela 1). Należy zastrzec, że klasyczna metoda głównych składowych jest oparta na analizie współczynników korelacji Pearsona, co wymaga stosowania pomiaru na skali przedziałowej lub ilorazowej. Niemniej jednak dopuszcza się stosowanie tej skali pomiaru także dla skal porządkowych 5- i 7-stopniowych [Sagan 2004, s. 89]. Na podstawie analizy wartości ładunków czynnikowych można zidentyfikować pozycje skali, które są najsłabiej powiązane z pozostałymi, mierzącymi tę samą właściwość, a więc określić, które z nich można usunąć ze skali. Chodzi o te stwierdzenia, które nie są specyficzne dla żadnego czynnika i tak samo silnie korelują z dwoma lub większą liczbą czynników, oraz takie, które słabo korelują ze wszystkimi czynnikami.

W analizowanym przykładzie lepsze rezultaty uzyskano dla wariantu 1 – stwierdzenie, które należy usunąć ze skali, to: „Nawigacja po serwisie banku jest całkowicie intuicyjna”. Mankamentem tego rozwiązania jest jednak trudność precyzyjnego nazwania utworzonych subskał. Pierwsza z nich obejmuje zarówno stwierdzenia dotyczące realizacji operacji bankowych, jak i elementy odnoszące się do szaty graficznej. Równie eklektyczna jest druga subskała, na którą składają się stwierdzenia dotyczące struktury serwisu, szaty graficznej oraz zakresu usług i informacji o nich. Stosunkowo najspójniejsza treściowo jest trzecia subskała, obejmująca stwierdzenia dotyczące sposobu prezentacji informacji w serwisie.

W przypadku zastosowania metody k -najbliższych sąsiadów (wariant 2) można wskazać aż pięć stwierdzeń, które nie są specyficzne dla żadnego z czynników. Oprócz wymienionej powyżej intuicyjnej nawigacji są to takie stwierdzenia, jak: „Stosowane w serwisie bankowym znaki, ikony i symbole są komunikatywne”, „Wielkość i krój czcionki pozwalają bez problemu odczytać informacje na stronie”, „Układ menu serwisu bankowego pozwala łatwo odnaleźć poszczególne usługi”, „W serwisie dostępna jest szczegółowa informacja o ofercie i warunkach usług”.

Na podstawie przeprowadzonej oceny wymiarowości skali w kolejnym etapie przeprowadzono analizę rzetelności wyodrębnionych skal jednowymiarowych (tabela 2). W tym przypadku próbowano uzyskać odpowiedź na pytanie, czy utworzone na podstawie analizy głównych składowych cztery subskałe są jednorodne, a więc w sposób rzetelny mierzą określoną właściwość. Eliminacja ze skali aż pięciu stwierdzeń w wariantcie analizy z zastosowaniem metody k -najbliższych sąsiadów spowodowała, że przeprowadzenie analizy rzetelności skali stało się zasadne jedynie dla wariantu 1, zakładającego usunięcie z bazy danych przypadków z odpowiedzią „nie wiem”.

W ocenie rzetelności skali stosowane są generalnie dwa podejścia. Pierwsze polega na powtarzaniu pomiaru na tej samej próbie badanych w niezbyt długim odstępie czasu (metoda *test-retest*) i badaniu stopnia zgodności uzyskanych wyników. W drugim podejściu, które znalazło zastosowanie w niniejszym przy-

Tabela 2. Rezultaty analizy rzetelności skali metodą α -Cronbacha

| Pozycje (stwierdzenia) w ramach subskali | Wartość α , gdy pozycja usunięta |
|--|---|
| Subskala 1 – α -Cronbacha 0,866 | |
| Operacje bankowe w serwisie wykonuje się za pomocą kilku kliknięć | 0,841 |
| Operacje bankowe w serwisie wykonuje się bardzo szybko | 0,828 |
| Operacje bankowe w serwisie wykonuje się bardzo łatwo | 0,819 |
| Stosowane w serwisie bankowym znaki, ikony i symbole są komunikatywne | 0,842 |
| Wielkość i krój czcionki pozwalają bez problemu odczytać informacje na stronie | 0,857 |
| Subskala 2 – α -Cronbacha 0,757 | |
| Szata graficzna serwisu bankowego jest estetyczna i przyjemna dla oka | 0,679 |
| Układ elementów na stronie serwisu jest spójny i logiczny | 0,674 |
| W serwisie dostępna jest szczegółowa informacja o ofercie i warunkach usług | 0,747 |
| Zakres usług w serwisie internetowym banku zaspokaja moje potrzeby | 0,704 |
| Subskala 3 – α -Cronbacha 0,784 | |
| Wskazówki, które można wyświetlić dla usług bankowych, są przydatne | 0,745 |
| Informacje i opisy usług bankowych są napisane zrozumiałym językiem | 0,699 |
| Układ menu serwisu bankowego pozwala łatwo odnaleźć poszczególne usługi | 0,687 |
| Wyświetlane komunikaty o błędach są zrozumiałe i jednoznaczne | 0,790 |

Źródło: opracowanie własne.

kładzie, analizuje się wewnętrzną spójność skali na podstawie współczynnika α -Cronbacha, który opiera się na zestawieniu współczynników korelacji wszystkich pozycji skali z ogólnym wynikiem skali [Sagan 2004, s. 93–95]. Przyjmuje się, że za rzetelną można uznać skalę, dla której wartość współczynnika α -Cronbacha wynosi 0,7. Oceniając rezultaty analizy rzetelności tą metodą, należy jednak pamiętać, że na jej wartość duży wpływ mają liczba ocenianych pozycji oraz wielkość próby badawczej. Im więcej znajdzie się w kwestionariuszu stwierdzeń oraz im większa będzie próba badawcza, tym wyższej wartości α -Cronbacha można się spodziewać [Böhner i Wanke 2004, s. 45–46].

Z analizy rzetelności przeprowadzonej dla utworzonych skal jednowymiarowych wynika, że charakteryzują się one dużą spójnością, co potwierdzają wysokie wartości współczynnika α -Cronbacha. Ponadto usunięcie poszczególnych pozycji w ramach utworzonych subskali nie podniosłoby wartości tego współczynnika. Utworzone subskale uzyskały dobre parametry rzetelności, pomimo że na podstawie wiedzy o elementach kształtujących strukturę serwisu internetowego można by oczekiwać nieco innego „składu stwierdzeń” w ramach prezentowanych subskali.

4. Podsumowanie

Niezależnie od tego, jaka metoda oceny rzetelności skali zostanie wybrana w celu zapewnienia odpowiednio wysokiej jakości badań, autor powinien starać się zawsze krytycznie spojrzeć na utworzone narzędzie badawcze. W analizowanym przykładzie wyniki analizy rzetelności można by uznać za satysfakcjonujące, gdyby nie problem z precyzyjnym nazwaniem utworzonych subskal. Na etapie budowy skali dyskusyjnym rozwiązaniem jest z kolei stosowanie we wzorcu reakcji respondentów wariantu „nie wiem”, a następnie eliminacja z analizy przypadków, w których wystąpiła taka odpowiedź. Takie działanie jest nieekonomiczne i prowadzi do zmniejszenia próby badawczej, a w konsekwencji najczęściej także do zachwiania jej struktury. Również działanie mające na celu zachowanie liczebności próby poprzez zastępowanie braków danych np. wartością średnią czy, jak w analizowanym przykładzie, wartościami uzyskanymi z zastosowaniem metody *k*-najbliższych sąsiadów nie jest dobrym rozwiązaniem, co potwierdziła analiza wymiarowości skali i uzyskane metodą głównych składowych wartości ładunków czynnikowych. Prezentowany przykład pokazuje zatem, że dążąc do optymalizacji tworzonej skali służącej do pomiaru postaw w odniesieniu do jakości użytkowej serwisów internetowych, na równi z kryteriami *stricte* statystycznymi należy stawiać merytoryczną znajomość problematyki badawczej.

Literatura

- Böhner G., Wanke M. [2004], *Postawy i ich zmiana*, tłum. J. Radzicki, Gdańskie Wydawnictwo Psychologiczne, Gdańsk.
- Dzida W. et al. [2000], *Gebrauchstauglichkeit von Software ErgoNorm: Ein Verfahren zur Konformitätsprüfung von Software auf der Grundlage von DIN EN ISO 9241 Teile 10 und 11*, Bundesanstalt für Arbeitsschutz und Arbeitsmedizin, Dortmund.
- Gediga G., Hamborg K.-C. [1999], *IsoMetrics: Ein Verfahren zur Evaluation von Software nach ISO 9241-10* [w:] *Evaluationsforschung*, red. H. Holling, G. Gediga, Hogrefe Verlag, Göttingen.
- Gediga G., Hamborg K.-C., Düntsch I. [1999], *The IsoMetrics Usability Inventory: An Operationalisation of ISO 9241-10 Supporting Summative and Formative Evaluation of Software Systems*, „Behaviour and Information Technology”, vol. 18, nr 3.
- Hilgard E.R. [1972], *Wprowadzenie do psychologii*, Wydawnictwo PWN, Warszawa.
- ISO 9241-11:1998, *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) – Part 11: Guidance on Usability*.
- Kirakowski J., Claridge N., Whitehand R. [1998], *Human Centered Measures of Success in Web Site Design*, <http://research.microsoft.com/en-us/um/people/marycz/hfweb98/kirakowski/> (dostęp:10.10.2012).

- Kirakowski J., Corbett M. [1988], *Measuring User Satisfaction* [w:] *People and Computers IV. Proceedings of HCI '88*, red. D.M. Jones, R. Winder, Cambridge University Press, Cambridge.
- Kirakowski J., Corbett M. [1993], *SUMI: The Software Usability Measurement Inventory*, „British Journal of Educational Technology”, vol. 24, nr 3.
- Michalski R. [2008], *Jakość użytkowa w procesie wytwarzania oprogramowania*, „Badania Operacyjne i Decyzje”, nr 4.
- Mika S. [1972], *Wstęp do psychologii społecznej*, Wydawnictwo PWN, Warszawa.
- Nielsen J. [1993], *Usability Engineering*, Morgan Kaufmann, San Francisco.
- Nowak S. [1973], *Pojęcie postawy w teoriach i stosowanych badaniach społecznych* [w:] *Teorie postaw*, red. S. Nowak, Wydawnictwo PWN, Warszawa.
- Prümper J. [1997], *Der Benutzerfragebogen ISONORM 9241/10: Ergebnisse zur Reliabilität und Validität* [w:] *Software-Ergonomie '97. Usability Engineering: Integration von Mensch-Computer-Interaktion und Software-Entwicklung*, red. R. Liskowsky, B.M. Velichkovsky, W. Wünschmann, B.G. Teubner Verlag, Stuttgart.
- Prümper J., Anft M. [1993], *Die Evaluation von Software auf Grundlage des Entwurfs zur internationalen Ergonomie-Norm 9241 Teil 10 als Beitrag zur partizipativen Systemgestaltung: Ein Fallbeispiel* [w:] *Software-Ergonomie '93. Von der Benutzungsoberfläche zur Arbeitsgestaltung*, red. K.H. Rödiger, B.G. Teubner Verlag, Stuttgart.
- Sagan A. [2004], *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Turowski J. [1993], *Socjologia. Małe struktury społeczne*, Towarzystwo Naukowe KUL, Lublin.

Issues in the Application of a Likert Scale to Measure Website Usability

Various types of methods can be used in website usability research, including user testing, eye tracking, and click tracking. However, the interview remains an important research method, and the questionnaire an important research tool, especially when quality measurement is being done in the context of ISO 9241-11:1998. A website's usability measurement scale should refer to utility determinants concerning a standard – such as effectiveness, efficiency and user satisfaction in the use of a given service. It should also take into account the complexity of the site's architecture.

The article presents the results of scale dimensionality analysis with the application of factorial analysis with reference to the scale measuring the usability of web services offered by banks. The subsequent step is an analysis of reliability of separated, one-dimensional scales based on the α -Cronbach method.

Keywords: usability, reliability of the scale, Internet banking, factorial analysis.

Małgorzata Sej-Kolasa
Mirosława Sztemberg-Lewandowska
Katedra Ekonometrii i Informatyki
Uniwersytet Ekonomiczny we Wrocławiu

Funkcjonalna analiza głównych składowych w badaniu zmian liczby studentów w czasie w krajach europejskich

Streszczenie

Analiza głównych składowych (PCA) polega na transformacji zmiennych pierwotnych w zbiór nowych wzajemnie ortogonalnych zmiennych, zwanych głównymi składowymi. Funkcjonalna analiza głównych (FPCA) składowych ma zalety klasycznej analizy głównych składowych, dodatkowo umożliwia analizę danych o charakterze dynamicznym. Podstawową różnicą między tymi dwiema metodami jest rodzaj danych: PCA bazuje na danych wielowymiarowych, natomiast FPCA na danych funkcjonalnych. Danymi funkcjonalnymi są krzywe i trajektorie, czyli ciąg indywidualnych obserwacji, a nie pojedyncza obserwacja. Celem artykułu jest pokazanie możliwości wykorzystania funkcjonalnej analizy głównych składowych do badania zjawisk opisanych danymi wzdłużnymi (*longitudinal data*). Przykład wykorzystania tej metody omówiony w artykule opiera się na analizie zmiany liczby studentów w czasie w wybranych krajach europejskich. Możliwości wizualizacyjne metody pozwalają na porównanie krajów i wyodrębnienie obserwacji odstających.

Słowa kluczowe: dane funkcjonalne, dane wzdłużne, funkcjonalna analiza głównych składowych, szkolnictwo wyższe.

1. Wprowadzenie

Analiza głównych składowych (PCA) polega na transformacji zmiennych pierwotnych w zbiór nowych wzajemnie ortogonalnych zmiennych, zwanych głównymi składowymi. Metodę tę wykorzystuje się do konstrukcji map percepcji, stosuje się ją także na etapie doboru zmiennych jako metodę redukcji danych. PCA umożliwia ponadto opis zjawisk z punktu widzenia nowych kategorii społeczno-ekonomicznych zdefiniowanych przez czynniki główne (zob. [Harman 1975, s. 154–160; Hair *et al.* 1998]).

Klasyczna analiza głównych składowych nie jest jednak odpowiednia do analizy danych o charakterze dynamicznym. Na analizę takich danych pozwala m.in. dynamiczna analiza głównych składowych lub funkcjonalna analiza głównych składowych.

Dynamiczne modele czynnikowe umożliwiają uzyskanie syntetycznej informacji o kształtowaniu się zmienności dużego zbioru danych, wykorzystywane są do konstruowania prognoz (głównie krótkookresowych) i monitorowania zjawisk. Istota metody sprowadza się do agregacji dużej liczby potencjalnych zmiennych objaśniających do kilku wzajemnie niezależnych czynników, które następnie wykorzystywane są do prognozowania wybranej zmiennej. Równanie prognozy opisujące zależności pomiędzy zmiennymi prognozowanymi a czynnikami ma zwykle postać liniową. Oprócz czynników w równaniu tym mogą wystąpić ich opóźnienia, a także składniki o charakterze autoregresyjnym.

Funkcjonalna analiza głównych składowych (FPCA) pozwala określić naturę danych, kształt trajektorii w czasie. Zarówno klasyczna, jak i funkcjonalna analiza głównych składowych pozwalają dokonać rzutu wielowymiarowych danych na przestrzeń o dużo mniejszym wymiarze przy jednoczesnym zachowaniu maksimum informacji (w tym przypadku zmienności danych). Analizowana metoda polega na znalezieniu takich składowych, których iloczyn skalarny z danymi daje maksymalną zmienność. Pierwsza składowa wyjaśnia najwięcej zmienności, druga jest prostopadła do pierwszej i wyjaśnia maksymalnie dużo z pozostałej części zmienności danych. Podstawową różnicą między tymi dwiema metodami jest rodzaj danych: PCA bazuje na danych wielowymiarowych, natomiast FPCA na danych funkcjonalnych [Ramsay i Silverman 2005, Ramsay, Hooker i Graves 2009].

Celem artykułu jest analiza zmiany liczby studentów w czasie w krajach europejskich. Badanie obejmuje nie tylko tendencję, ale także tempo zmian liczby studentów w czasie. Zastosowano funkcjonalną analizę głównych składowych, której możliwości wizualizacyjne pozwalają na porównanie krajów i wyodrębnienie obserwacji odstających.

Obliczenia zostały wykonane w programie R (v.2.15.0, pakiet *fda* i *fda.usc*).

2. Metoda badawcza

Danymi funkcjonalnymi są krzywe i trajektorie, czyli ciąg indywidualnych obserwacji, a nie jak zwykle pojedyncza obserwacja. Choć dane funkcjonalne często są wyrażone w czasie (zależą od czasu), to ich zakres i cel są zupełnie inne niż szeregów czasowych. Analiza szeregów czasowych ma na celu modelowanie lub prognozowanie danych, natomiast funkcjonalna analiza danych pozwala badać naturę danych, kształt trajektorii w czasie [Ingrassia i Costanzo 2005].

Dane funkcjonalne posiadają realizacje dyskretne. Niech $\mathbf{y}_i = ((y_i(t_1), y_i(t_2), \dots, y_i(t_p)))$ będzie próbkowym pomiarem zmiennej Y w czasie t_1, t_2, \dots, t_p dla i -tej jednostki ($i = 1, 2, \dots, n$). Dane y_i nazywane są surowymi danymi funkcjonalnymi (*raw functional data*). Dane te przekształca się zgodnie z procedurami wygładzającymi, np. za pomocą liniowych kombinacji K znanych funkcji bazowych, na odpowiednią funkcję $x_i(t)$, która jest właściwą postacią funkcjonalną danych. Zbiór $\mathbf{X}_t = ((x_1(t), x_2(t), \dots, x_n(t)))$ nazywany jest funkcjonalnym zbiorem danych (*functional dataset*) [Daniele 2006, Hall, Müller i Wang 2006].

Techniki statystyczne dla funkcjonalnych danych zakładają, że funkcje opisujące dane należą do przestrzeni Hilberta: są funkcjami rzeczywistymi określonymi na przedziale domkniętym, całka kwadratów tych funkcji jest skończona (tzn. norma funkcji jest skończona).

Klasyczna analiza głównych składowych służy do eksploracji zmienności w wielowymiarowym zbiorze danych. Wykorzystując wartości własne macierzy wariancji dla danych PCA, za jej pomocą wyznacza się składowe, które wyjaśniają zmienność w obserwowanym zbiorze danych. Dla każdej składowej głównej ustala się ładunki czynnikowe na wszystkich zmiennych, określające wariancję wyjaśnioną przez daną składową [Harman 1975].

W przypadku funkcjonalnej analizy głównych składowych każda główna składowa wyrażona jest przez funkcję wagową głównych składowych (*principal component weight function*), inaczej nazwaną funkcją własną (*eigenfunction*) $\xi_j(t)$ zależną od czasu [Daniele 2006]. Funkcja własna maksymalizuje wariancję funkcji głównych składowych:

$$v(t, s) \stackrel{\text{def}}{=} \frac{1}{n-1} \sum_{i=1}^n \{x_i(t) - \bar{x}(t)\} \{x_i(s) - \bar{x}(s)\}. \quad (1)$$

Analogicznie do klasycznej PCA problemem w funkcjonalnej analizie głównych składowych jest rozkład wariancji funkcji:

$$v(t, s) = \sum_j \lambda_j \xi_j(t) \xi_j(s), \quad (2)$$

gdzie λ_j , $\xi_j(t)$ spełniają równanie własne:

$$\langle v(s), \xi_j \rangle = \lambda_j \xi_j(s), \quad (3)$$

a wartości własne są dodatnie i niemalejące:

$$\lambda_j \stackrel{\text{def}}{=} \int_T \xi_j(t) v(t, s) \xi_j(s) dt ds. \quad (4)$$

Funkcje własne spełniają warunek:

$$\int_T \xi_j^2(t) dt = 1 \quad \text{oraz} \quad \int_T \xi_j(t) \xi_i(t) dt = 0 \quad (i < j). \quad (5)$$

Wyniki głównych składowych dla i -tego obiektu w zbiorze danych są zdefiniowane następująco:

$$w_i^{(j)} \stackrel{\text{def}}{=} \langle x_i, \xi_j \rangle = \int_T \xi_j(t) x_i(t) dt. \quad (6)$$

Funkcje własne określają główne składowe zmienności między próbkowymi funkcjami x_i [Ingrassia i Costanzo 2005, Hall, Müller i Wang 2006].

3. Przebieg badania

Spadek liczby ludności oraz starzenie się społeczeństw skutkują wieloma niekorzystnymi zmianami o charakterze ekonomicznym i społecznym. Demograficzne tsunami wpływa również na sytuację szkolnictwa wyższego. Od kilku lat w szkołach wyższych w większości krajów Europy liczba studentów spada, co znacząco wpływa na ograniczenie możliwości rozwoju szkolnictwa wyższego.

W celu uniknięcia bezpośredniego porównywania liczby studentów w przeprowadzonych analizach wykorzystano zmienną będącą stosunkiem liczby studentów w danym roku do liczby studentów w pierwszym badanym roku.

Badaniem objęto kraje europejskie, dane pochodziły z baz Eurostatu. Początkowo zakładano uwzględnienie wszystkich krajów Europy, jednak brak danych dla niektórych państw spowodował konieczność zawężenia zakresu czasowego (lata 2000–2009) oraz przestrzennego (28 państw). Z tego powodu wśród państw poddanych analizie nie znalazła się m.in. Francja (tabela 1).

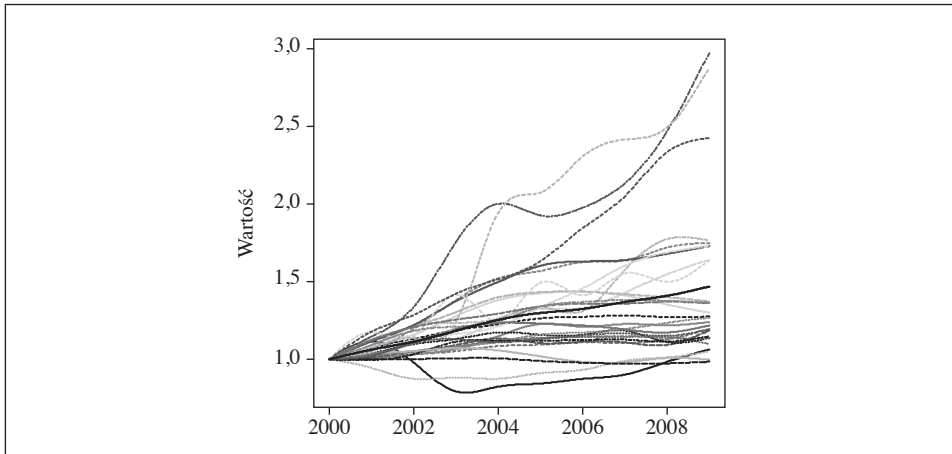
Tabela 1. Zmiana liczby studentów w czasie

| Kraj | 2000 ^a | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|-----------------|-------------------|------|------|------|------|------|------|------|------|------|
| Austria | 1,00 | 1,12 | 0,98 | 0,79 | 0,82 | 0,84 | 0,87 | 0,90 | 0,98 | 1,06 |
| Belgia | 1,00 | 1,01 | 1,03 | 1,05 | 1,08 | 1,09 | 1,11 | 1,11 | 1,13 | 1,19 |
| Bułgaria | 1,00 | 0,95 | 0,87 | 0,88 | 0,87 | 0,91 | 0,93 | 0,99 | 1,01 | 1,05 |
| Cypr | 1,00 | 1,15 | 1,34 | 1,75 | 2,00 | 1,93 | 1,98 | 2,13 | 2,47 | 2,98 |
| Czechy | 1,00 | 1,03 | 1,12 | 1,13 | 1,26 | 1,33 | 1,33 | 1,43 | 1,55 | 1,64 |
| Dania | 1,00 | 1,01 | 1,03 | 1,07 | 1,15 | 1,23 | 1,21 | 1,23 | 1,22 | 1,24 |
| Estonia | 1,00 | 1,08 | 1,13 | 1,19 | 1,22 | 1,26 | 1,27 | 1,28 | 1,27 | 1,28 |
| Finlandia | 1,00 | 1,03 | 1,05 | 1,08 | 1,11 | 1,13 | 1,14 | 1,14 | 1,15 | 1,10 |
| Hiszpania | 1,00 | 1,00 | 1,00 | 1,01 | 1,01 | 0,99 | 0,98 | 0,97 | 0,97 | 0,98 |
| Holandia | 1,00 | 1,03 | 1,05 | 1,08 | 1,11 | 1,16 | 1,17 | 1,19 | 1,24 | 1,27 |
| Irlandia | 1,00 | 1,04 | 1,10 | 1,13 | 1,17 | 1,16 | 1,16 | 1,19 | 1,11 | 1,14 |
| Islandia | 1,00 | 1,05 | 1,20 | 1,38 | 1,52 | 1,57 | 1,63 | 1,64 | 1,72 | 1,75 |
| Litwa | 1,00 | 1,12 | 1,22 | 1,37 | 1,50 | 1,60 | 1,63 | 1,64 | 1,68 | 1,73 |
| Łotwa | 1,00 | 1,13 | 1,21 | 1,30 | 1,40 | 1,43 | 1,44 | 1,42 | 1,40 | 1,37 |
| Macedonia | 1,00 | 1,09 | 1,21 | 1,24 | 1,26 | 1,34 | 1,31 | 1,58 | 1,77 | 1,77 |
| Malta | 1,00 | 1,18 | 1,15 | 1,42 | 1,25 | 1,50 | 1,41 | 1,55 | 1,50 | 1,64 |
| Niemcy | 1,00 | 1,01 | 1,05 | 1,09 | 1,13 | 1,10 | 1,11 | 1,11 | 1,09 | 1,19 |
| Norwegia | 1,00 | 1,00 | 1,03 | 1,11 | 1,12 | 1,12 | 1,12 | 1,13 | 1,11 | 1,15 |
| Polska | 1,00 | 1,12 | 1,21 | 1,26 | 1,29 | 1,34 | 1,36 | 1,36 | 1,37 | 1,36 |
| Portugalia | 1,00 | 1,04 | 1,05 | 1,07 | 1,06 | 1,02 | 0,98 | 0,98 | 1,01 | 1,00 |
| Rumunia | 1,00 | 1,18 | 1,29 | 1,42 | 1,51 | 1,63 | 1,84 | 2,05 | 2,33 | 2,43 |
| Słowacja | 1,00 | 1,06 | 1,12 | 1,16 | 1,21 | 1,33 | 1,46 | 1,60 | 1,69 | 1,73 |
| Słowenia | 1,00 | 1,09 | 1,18 | 1,21 | 1,25 | 1,34 | 1,37 | 1,38 | 1,38 | 1,36 |
| Szwecja | 1,00 | 1,03 | 1,10 | 1,20 | 1,24 | 1,23 | 1,22 | 1,19 | 1,17 | 1,22 |
| Turcja | 1,00 | 1,08 | 1,14 | 1,24 | 1,94 | 2,07 | 2,31 | 2,42 | 2,49 | 2,88 |
| Węgry | 1,00 | 1,08 | 1,16 | 1,28 | 1,38 | 1,43 | 1,44 | 1,41 | 1,35 | 1,30 |
| Wielka Brytania | 1,00 | 1,02 | 1,11 | 1,13 | 1,11 | 1,13 | 1,15 | 1,17 | 1,15 | 1,19 |
| Włochy | 1,00 | 1,02 | 1,05 | 1,08 | 1,12 | 1,14 | 1,15 | 1,15 | 1,14 | 1,14 |

^a 2000 r. – rok bazowy.

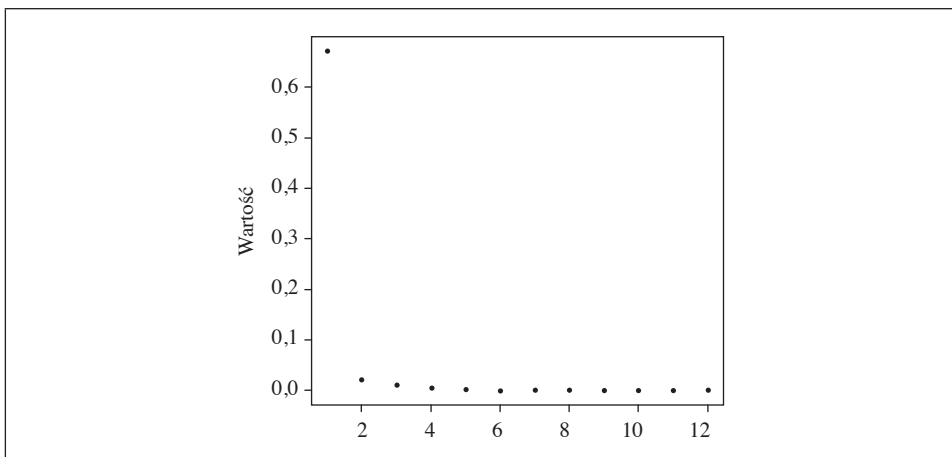
Źródło: opracowanie własne na podstawie [Eurostat Statistics... 2012].

Wielkości zmian liczby studentów w badanym okresie dla wybranych państw europejskich przedstawia rys. 1. Pogrubiona krzywa oznacza średnią dla badanych państw. Rok 2000 był rokiem bazowym, w związku z czym dla tego roku wskaźnik dla wszystkich państw przyjmuje wartość 1.



Rys. 1. Wielkości zmian liczby studentów w czasie

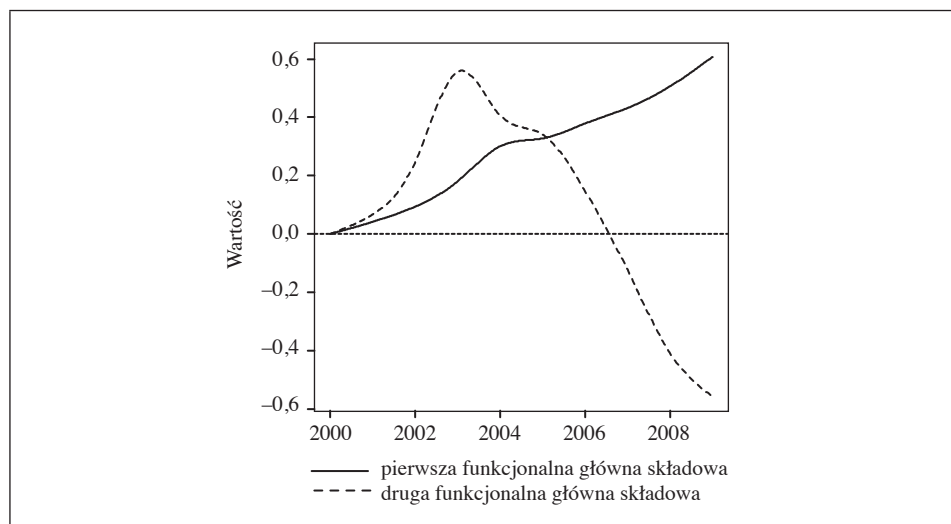
Źródło: opracowanie własne z wykorzystaniem programu R.



Rys. 2. Wykres osypiska

Źródło: opracowanie własne z wykorzystaniem programu R.

Wykres pokazuje, jak zmienia się liczba studentów w czasie. Średnia liczba studentów w badanym okresie rośnie, jednak trudno jest wskazać, w których krajach tempo zmian jest podobne. W tym celu przeprowadzono funkcjonalną analizę głównych składowych.



Rys. 3. Wyodrębnione funkcjonalne główne składowe

Źródło: opracowanie własne z wykorzystaniem programu R.

Na podstawie wykresu osypiska (rys. 2) ustalono liczbę funkcji składowych, a następnie wyodrębniono te funkcje (rys. 3).

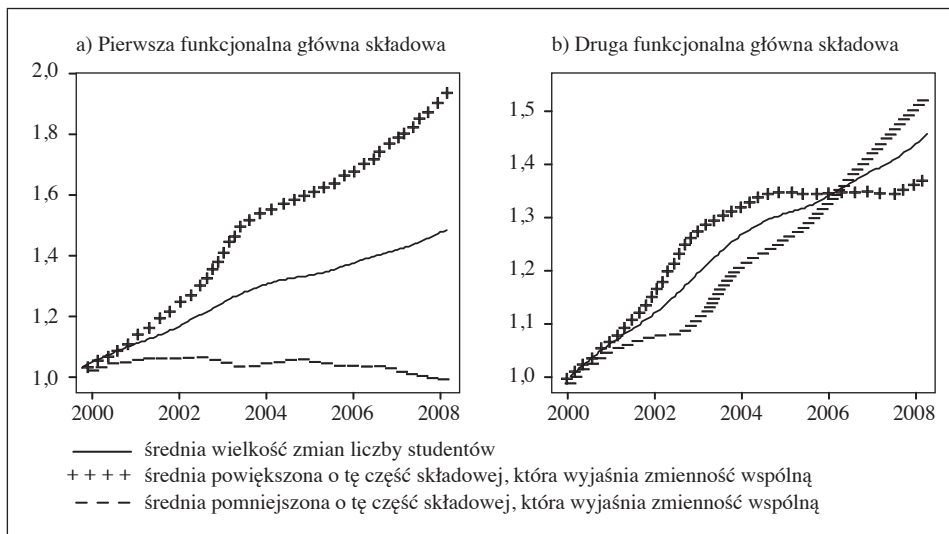
Pierwsza funkcjonalna główna składowa, na rys. 3 oznaczona linią ciągłą, wyjaśnia 85% zmienności wspólnej, natomiast druga, oznaczona linią przerywaną – 7%.

Podobnie jak w klasycznej analizie czynnikowej główne składowe należy zinterpretować, jednak w przypadku danych funkcjonalnych interpretacja jest trudniejsza. Praktyczne wyjaśnienie funkcjonalnych głównych składowych ułatwiają wykresy odchylenia każdej ze składowych od średniej (rys. 4).

Z rys. 4a wynika, że pierwsza składowa odpowiada za ogólną tendencję. Przyjmuje wartości dodatnie (rys. 3), zatem dla danego kraju dodatni ładunek na tej składowej oznacza, że krzywa opisująca wielkość zmian liczby studentów leży powyżej średniej.

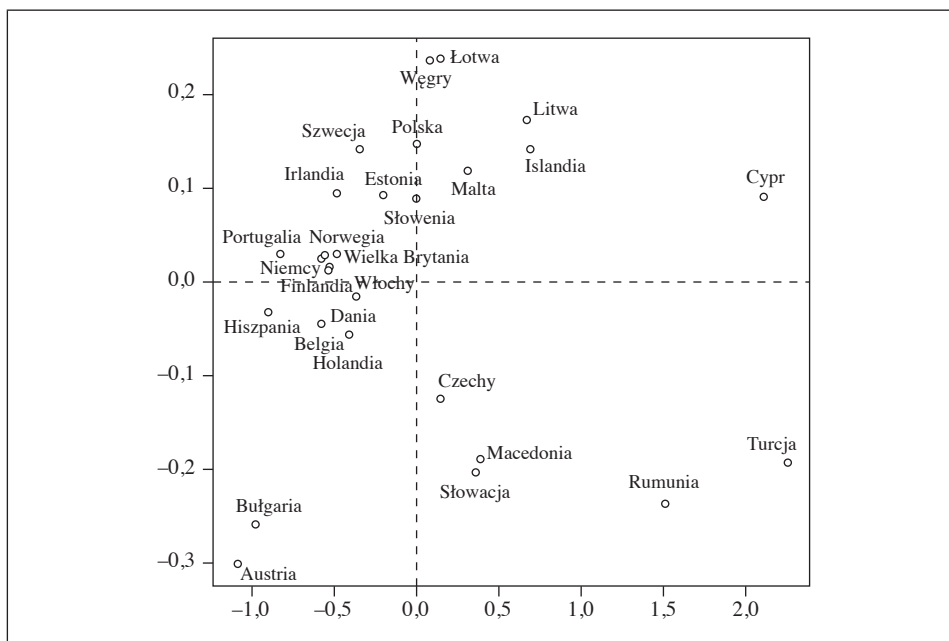
Druga składowa pokazuje tendencje w pierwszych latach w odniesieniu do lat ostatnich („początek kontra koniec”) – porównuje zatem okres do 2007 r. z okresem po 2007 r. (rys. 4b). Dodatni ładunek na drugiej składowej oznacza, że tempo zmian na początku badanego okresu było większe od średniej, natomiast na końcu tego okresu tempo było mniejsze od średniej.

Funkcjonalna analiza czynnikowa pozwala na wizualizację danych umożliwiającą porównanie badanych obiektów. Rys. 5 zawiera rzut danych na płaszczyznę wyznaczoną przez dwie funkcjonalne główne składowe.



Rys. 4. Odchylenia funkcjonalnych głównych składowych od średniej

Źródło: opracowanie własne z wykorzystaniem programu R.



Rys. 5. Rzut obiektów na płaszczyznę wyznaczoną przez dwie funkcje składowe

Źródło: opracowanie własne z wykorzystaniem programu R.

Wśród krajów znacząco odbiegających od pozostałych pod względem tempa i kierunku zmian liczby studentów znalazły się m.in.: Turcja, Cypr, Rumunia, Bułgaria i Austria.

Najbardziej od pozostałych państw różni się Turcja, która ma wysoki ładunek na pierwszej składowej (tempo zmian powyżej średniej) i niski (ujemny) na drugiej, co oznacza, że tempo zmian w końcowym okresie było większe niż w początkowym. Szybkie tempo zmian liczby studentów można tłumaczyć odmienną niż w pozostałych krajach sytuacją demograficzną – w Turcji systematycznie wzrasta liczba osób młodych, co jest ewenementem na skalę Europy.

Cypr – podobnie jak Turcja – wykazuje tempo zmian powyżej średniej (wysoki ładunek na pierwszej składowej), ale tempo zmian w początkowym okresie było wyższe niż w końcowym (dodatni ładunek na drugiej składowej). Tempo zmian powyżej średniej częściowo wyjaśnia duża liczba studiujących tam obcokrajowców – jak wynika z danych Eurostatu, wskaźnik umiędzynarodowienia uczelni wynosi na Cyprze ponad 28% i jest jednym z najwyższych w Europie.

W Rumunii zmiany liczby studentów mają podobny charakter jak w Turcji: odnotowuje się tempo zmian powyżej średniej, wyższe w końcowym okresie. Wynika to z sytuacji demograficznej – w badanym okresie osoby w wieku „studenckim” zgodnie z danymi Eurostatu stanowiły ponad 20% społeczeństwa Rumunii.

Bułgaria i Austria są w podobnej sytuacji pod względem omawianych składowych. Ujemny ładunek na pierwszej składowej oznacza, że tempo zmian liczby studentów jest poniżej średniej, większe tempo zmian miało miejsce na końcu badanego okresu (ujemny ładunek na drugiej składowej).

W Polsce tempo zmian liczby studentów jest średnie, w początkowym okresie zaobserwowano większe zmiany niż w końcowym.

4. Wnioski

Na podstawie przeprowadzonych analiz można wysnuć wnioski zarówno o charakterze teoretycznym, dotyczące metody, jak i o charakterze aplikacyjnym.

Podobnie jak klasyczna analiza głównych składowych FPCA pozwala na wizualizację zjawiska (co znacznie ułatwia analizę), wykazując nietrywialnie zależności, które trudno wykryć w inny sposób. Istotną zaletą obu metod jest redukcja danych przy zachowaniu maksimum informacji.

Funkcjonalna analiza głównych składowych, wzbogacając możliwości klasycznej analizy głównych składowych, pozwala na analizę danych o charakterze dynamicznym – pokazuje nie tylko tendencję, ale i tempo zmian w czasie.

Zastosowana metoda, określając tempo zmiany liczby studentów, umożliwia wskazanie krajów podobnych oraz znalezienie krajów różniących się od pozostałych ze względu na przyjęte kryterium analiz.

Braki danych, dotyczące głównie zakresu przestrzennego, mogą wpływać na otrzymane wyniki badań. Badaniem należałoby objąć wszystkie kraje. Uwzględnienie w badaniu pominiętych państw (np. Francji) mogłoby spowodować konieczność zmiany interpretacji funkcjonalnych składowych głównych, a tym samym zmianę wniosków na temat sytuacji poszczególnych państw.

Analiza otrzymanych wyników w celu wyjaśniania przyczyn sytuacji określonych państw nie była przedmiotem zainteresowania i powinna być przedmiotem odrębnych badań.

Literatura

- Daniele M. [2006], *Functional Principal Components Analysis to Study Environmental Data*, http://www.sis-statistica.it/files/pdf/atti/Spontanee%202006_677-680.pdf (dostęp: 5.12.2013).
- Eurostat Statistics* [2012], http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database?_piref458_1209540_458_211810_211810.node_code=educ_enr18, (dostęp: 30.10.2012).
- Hair J.F. *et al.* [1998], *Multivariate Data Analysis with Readings*, Prentice-Hall, New York.
- Hall P., Müller H.G., Wang J.L. [2006], *Properties of Principal Component Methods for Functional and Longitudinal Data Analysis*, „The Annals of Statistics”, vol. 34, nr 3.
- Harman H. [1975], *Modern Factor Analysis*, The University of Chicago Press, Chicago.
- Ingrassia S., Costanzo G.D. [2005], *Functional Principal Component Analysis of Financial Time Series [w:] New Developments in Classification and Data Analysis*, red. M. Vichi *et al.*, Springer, Berlin.
- Ramsay J.O., Hooker G., Graves S. [2009], *Functional Data Analysis with R and MATLAB*, Springer, New York.
- Ramsay J.O., Silverman B.W. [2005], *Functional Data Analysis*, Springer, New York.

The Analysis of Changes over Time in the Number of Students Using Functional Principal Component Analysis in European Countries

Principal component analysis (PCA) transforms an original set of variables into a new orthogonal set called principal components. Functional principal component analysis (FPCA) has the same advantages as classical principal component analysis while also enabling the analysis of dynamic data. The main difference between them is that PCA is based on multidimensional data and FPCA is based on functional data. The functional data are curves, surfaces or anything else varying over a continuum. They are not a single observation. The main aim of the paper is to show the usefulness of applying functional principal component analysis in order to analyse longitudinal data. The paper presents an

example of how this method has been used based on the analysis of changes in the number of students (over time) in chosen European countries. Visualisation of the results makes it possible to compare countries and detect outliers.

Keywords: functional data, longitudinal data, functional principal component analysis, higher education.

Iwona Olejnik

Robert Skikiewicz

Katedra Badań Rynku i Usług
Uniwersytet Ekonomiczny w Poznaniu

Metoda k -średnich w segmentacji emerytów na podstawie priorytetów życiowych*

Streszczenie

W artykule porównano wyniki segmentacji emerytów metodą k -średnich na podstawie zestawu zmiennych behawioralnych oraz zestawu zmiennych behawioralnych i demograficzno-ekonomicznych. Posługując się zestawem cech behawioralnych, otrzymano segmenty emerytów wykazujące znaczne różnice w zakresie priorytetów życiowych oraz stosunkowo niewielkie różnice w zakresie profili demograficzno-ekonomicznych. Z kolei w przypadku segmentacji opartej na połączonym zestawie zmiennych behawioralnych oraz demograficzno-ekonomicznych większy wpływ na ostateczny wynik grupowania wywierały cechy demograficzno-ekonomiczne. Segmenty emerytów stały się łatwiej identyfikowalne, jednak różnice w zakresie cech behawioralnych uległy zmniejszeniu.

Słowa kluczowe: segmentacja, metoda k -średnich, emeryci, badania ilościowe.

* Artykuł powstał w ramach realizacji projektu badawczego NCN pt. *Zachowania oszczędnościowe i finansowe zabezpieczenie emerytalne w gospodarstwach domowych – determinanty, postawy, modele* (nr UMO-2012/05/B/HS4/04183).

1. Wprowadzenie

Wobec starzenia się społeczeństwa i jednocześnie coraz dłuższej średniej długości życia coraz liczniejszą grupę konsumentów stanowią ludzie starsi. Badania i analizy zachowań konsumentów w zdecydowanej większości dotyczą jednak osób postrzeganych jako atrakcyjne dla przedsiębiorców – będących co najwyżej w średnim wieku i usyskujących średnie lub wysokie dochody. Jednocześnie osoby starsze (w niniejszym opracowaniu uznaje się za takie ludzi w wieku przedemerytalnym i emerytalnym) wciąż postrzegane są jako segment generalnie homogeniczny, którym zainteresowanych jest niewielu menedżerów¹. Można jednak postawić pytanie, czy jest to grupa wewnątrznie jednorodna.

Celem poznawczym artykułu jest próba wskazania segmentów emerytów² posiadających zbliżone priorytety życiowe oraz ich identyfikacja na podstawie głównych cech demograficzno-ekonomicznych. Celem metodycznym artykułu jest natomiast porównanie efektów segmentacji metodą *k*-średnich na podstawie zmiennych behawioralnych oraz zestawu zmiennych behawioralnych i demograficzno-ekonomicznych.

2. Segmentacja emerytów – założenia wstępne

Osoby starsze w procedurze segmentacji najczęściej ujmowane są jako jednolita grupa konsumentów. Coraz częściej jednak wiek emeryta zapisany w metryce urodzenia nie ma istotnego wpływu na jego zachowania, a starość postrzegana jest nie jako cecha fizyczna, ale stan umysłu. Jednocześnie doświadczenie i relatywnie ustabilizowany tryb życia emerytów, a także rosnąca średnia długość życia oraz coraz większa „otwartość na świat” powodują, że potrzeby i zachowania osób starszych po zakończeniu aktywności zawodowej ewoluują. Można zatem postawić hipotezę, że osoby starsze stanowią grupę niejednorodną, a zatem możliwe jest przeprowadzenie ich pogłębionej segmentacji³. Obok cech demograficznych

¹ Za osobę w wieku przedemerytalnym w artykule uznaje się pracownika, który objęty jest ochroną przedemerytalną (obejmującą 4 lata przed osiągnięciem ustawowego wieku emerytalnego), tj. do 2012 r. co najmniej 56-letnią kobietę i co najmniej 61-letniego mężczyznę. Z analiz wyłączono osoby w wieku przedemerytalnym, które mogły przejść na emeryturę w obniżonym wieku emerytalnym (np. pracujących w służbach mundurowych).

² W niniejszym artykule emeryci będą zamiennie nazywani osobami starszymi.

³ Przykłady segmentacji osób starszych znaleźć można m.in. w pracach: [Bombol i Słaby 2011, s. 103 i nast.; Olejnik 2012, s. 109–115; Sudbury i Simcock 2009, s. 251–262].

czy ekonomicznych⁴ [Kaczmarczyk 2007, s. 77–78; Sudbury i Simcock 2009, s. 251–262] w przypadku emerytów istotne znaczenie mogą mieć w szczególności: stan zdrowia, priorytety życiowe wyrażone przez potrzeby, postawy, style życia, wyznawane wartości oraz zachowania [Jurek 2012, s. 165–167].

Celem weryfikacji możliwości zastosowania procedury segmentacji wobec osób starszych na podstawie zmiennych demograficzno-ekonomicznych oraz priorytetów życiowych w artykule wykorzystano wyniki badań przeprowadzonych w Katedrze Badań Marketingowych Uniwersytetu Ekonomicznego w Poznaniu w 2011 r.⁵

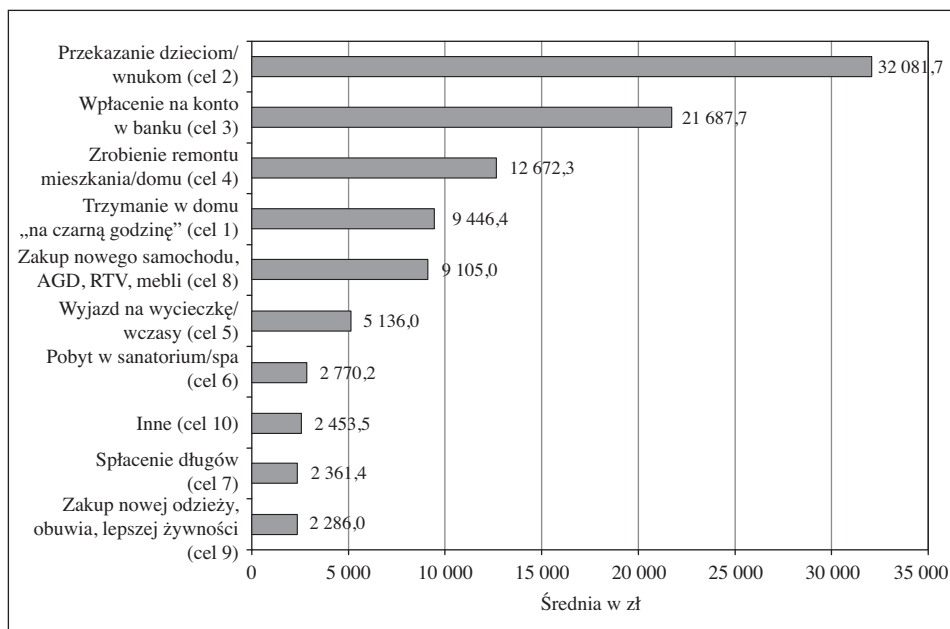
Jako podstawę segmentacji osób starszych przyjęto zatem działania lub cele o pierwszoplanowym dla nich znaczeniu w sytuacji, kiedy niespodziewanie wygraliby 100 tys. zł – w opracowaniu określono je jako priorytety życiowe. Wśród dziesięciu kategorii rangowej skali sumowanych ocen, pomiędzy które osoby starsze miały za zadanie podzielić tę hipotetyczną wygraną (100 tys. zł), znalazły się: trzymanie pieniędzy w domu „na czarną godzinę” (cel 1), przekazanie dzieciom/wnukom (cel 2), wpłacenie na konto w banku (cel 3), zrobienie remontu mieszkania/domu (cel 4), wyjazd na wycieczkę/wczasy (cel 5), pobyt w sanatorium/spa (cel 6), spłacenie długów (cel 7), zakup nowego samochodu, sprzętu AGD, RTV, mebli (cel 8), zakup nowej odzieży, obuwia, lepszej żywności (cel 9), inne⁶ (cel 10). Deklarowane cele, na które zostałaby przeznaczona wygrana mogą świadczyć bowiem z jednej strony o niezaspokojonych potrzebach, z drugiej zaś – o stylu życia oraz uznawanych wartościach.

⁴ Przykładowo bardzo odmienna w stosunku do innych segmentów rynku jest struktura wydatków z budżetu gospodarstwa domowego emerytów (np. brak wydatków związanych z utrzymaniem dzieci czy wykonywaniem pracy zawodowej) oraz posiadanie na ogół własnego, nieobciążonego kredytem hipotecznym i wyposażonego w różne sprzęty mieszkania lub domu.

⁵ Badania zrealizowane zostały w Wielkopolsce metodą wywiadu bezpośredniego z udziałem ankietowanych. Przeprowadzono je wśród 633 respondentów – emerytów, osób będących na wcześniejszej emeryturze oraz osób pobierających zasiłek lub świadczenie przedemerytalne. Dobór próby miał charakter losowy na etapie wyboru miejscowości, w których realizowano badania, oraz charakter kwotowy na etapie doboru poszczególnych respondentów (uwzględniono dane statystyczne dotyczące wieku i wykształcenia osób w wieku przedemerytalnym oraz emerytów). Średnia wieku osób w próbie wyniosła 69 lat (minimum 37 lat, maksimum 94 lata). 67% osób badanych stanowiły kobiety, co piąta osoba legitymowała się wykształceniem podstawowym, 26% – zawodowym, 37% – średnim, a 17% – wyższym. Pod względem wielkości gospodarstw domowych w próbie dominowały gospodarstwa dwuosobowe – 43% (1-osobowe stanowiły 30%, 3 i więcej-osobowe – 27%), a średnie dochody wynosiły 1570 zł (minimum 548 zł, maksimum 7 tys. zł).

⁶ Wśród innych przeznaczeń wygranej podanych przez osoby biorące udział w badaniu pojawiły się takie odpowiedzi, jak: cele charytatywne (fundacja dla dzieci, na kościół, dla biednych i potrzebujących, na rzecz zwierząt), zakup nieruchomości (dom nad jeziorem, działka), inwestycje finansowe (obligacje skarbowe), usługi medyczne (w tym zakup leków), nowy pomnik na cmentarzu, gry losowe.

Jak wynika z przeprowadzonych badań (rys. 1), największą część wygranej emerycy przekazałby dzieciom lub wnukom (średnio ponad 32 tys. zł), wpłaciliby na konto w banku (prawie 22 tys. zł) lub przeprowadzili remont mieszkania/domu (prawie 13 tys. zł). Hierarchia ta wskazuje z jednej strony na bardzo prorodzinne postawy osób starszych, z drugiej zaś świadczy o istnieniu potrzeby tworzenia i posiadania oszczędności również w wieku przedemerytalnym lub emerytalnym (wpłacenie na konto w banku lub trzymanie w domu „na czarną godzinę”) – pomimo niezaspokojenia innych potrzeb [Psychologia ekonomiczna... 2004, s. 550]⁷.



Rys. 1. Przeznaczenie wygranej przez osoby starsze – średnie wartości

Źródło: opracowanie na podstawie przeprowadzonych badań.

Znaczna dywersyfikacja średnich kwot hipotetycznej wygranej na różne cele w badanej grupie respondentów (jak również analiza tabel kontyngencji) stanowiła podstawę podjęcia decyzji o wykorzystaniu priorytetów życiowych wyrażonych przez podział wygranej w procedurze segmentacji osób starszych.

⁷ Co ciekawe, w hierarchii wartości Polaków [Diagnoza społeczna... 2011, s. 220] „dzieci” znalazły się dopiero na trzecim miejscu, zaś „pieniądze” na piątym. Ponadto, uwzględniając fakt, że w systemie wartości Polaków na pierwszym miejscu znalazło się „zdrowie”, zaskakujące jest przeznaczenie przez emerytów relatywnie niskiej kwoty z ewentualnej wygranej na wydatki pozwalające na poprawienie lub utrzymanie dobrego stanu zdrowia – niecałe 2,8 tys. zł.

3. Zastosowana procedura segmentacji

W klasycznym ujęciu procedury segmentacji z zastosowaniem metod analiz wielowymiarowych na początku dokonuje się wyboru zmiennych segmentacyjnych (często są to cechy behawioralne) oraz zmiennych deskryptywnych służących do profilowania (cechy demograficzne, ekonomiczne, geograficzne, psychograficzne). Następnie na podstawie zestawu zmiennych segmentacyjnych z zastosowaniem wybranej metody analitycznej w ramach zbiorowości wyodrębniane są segmenty. W końcowym etapie segmenty te są charakteryzowane na podstawie służących do ich identyfikacji zmiennych deskryptywnych [Rószkiewicz 2011, s. 16; Sagan 2004, s. 159–161].

Po dokonaniu podziału jednostek na segmenty nierzadko pojawia się problem ograniczonej ich identyfikacji na podstawie zmiennych deskryptywnych, w szczególności cech demograficzno-ekonomicznych. Wiąże się to z tym, że w podobny sposób mogą zachowywać się respondenci o różnych cechach, a respondenci posiadający takie same cechy mogą zachowywać się odmiennie [Fatuła 2010, s. 159–160]. Z tego względu podkreślane są zalety segmentacji prowadzonej na podstawie zestawu zmiennych, zawierającego zarówno zmienne behawioralne, jak i demograficzno-ekonomiczne. Podejście takie umożliwia łatwiejszą identyfikację segmentów, które równocześnie pod względem cech demograficzno-ekonomicznych stają się wyrazistsze [Rondan-Cataluna, Sanchez-Franco i Villarejo-Ramos 2010, s. 840–843].

W dalszej części badań do segmentacji emerytów zastosowana została metoda *k*-średnich, należąca do najczęściej stosowanych metod analizy skupień. Metoda ta należy do grupy metod optymalnego rozdziału. Jej istota sprowadza się do iteracyjnego przyporządkowywania wszystkich jednostek do *a priori* zdefiniowanej liczby *k* segmentów [Rószkiewicz 2002, s. 246]. W wielu publikacjach poświęconych porównaniom efektywności metod segmentacji jako jedna z najlepszych wskazywana jest metoda *k*-średnich. Metoda ta charakteryzuje się w szczególności większą trafnością łączenia jednostek w skupienia od metod hierarchicznych [Blashfield 1976, s. 377–388; Rujasiri i Chomtee 2009, s. 378–388].

Ze względu na prowadzenie segmentacji na podstawie danych ilościowych i jakościowych wykorzystany został wariant metody *k*-średnich dostępny w module Data Mining w pakiecie Statistica 10. Dodatkową zaletą tego rozwiązania jest możliwość zastosowania *V*-krotnego sprawdzianu krzyżowego w celu zautomatyzowanego ustalenia optymalnej liczby segmentów [Migut 2004, s. 26].

Segmentację emerytów metodą *k*-średnich przeprowadzono w dwóch ujęciach – na podstawie zestawu zmiennych wyrażających ich priorytety życiowe oraz na podstawie tego zestawu powiększonego o cztery zmienne demograficzno-ekonomiczne (wiek, wykształcenie, zawód wykonywany przed przejściem na emeryturę

oraz dochód). Pomiar priorytetów życiowych na rangowej skali sumowanych ocen pozwala na potraktowanie reprezentujących je zmiennych jako mierzonych na skali ilorazowej. Zmienne demograficzno-ekonomiczne zastosowane w segmentacji prezentują cechy respondentów na jakościowym poziomie pomiaru.

W celu sprowadzenia do porównywalności ilościowych zmiennych segmentacyjnych przeprowadzono ich standaryzację. W procedurze segmentacji opartej wyłącznie na zmiennych ilościowych użyto metrykę euklidesową, zaś w przypadku wariantu opartego na zmiennych ilościowych i jakościowych zastosowano odległość miejską [Balicki 2009, s. 215–216].

4. Ocena wyników segmentacji

Segmentacja przeprowadzona w wariancie klasycznym (wariant segmentacji I) – wyłącznie na podstawie dziesięciu ilościowych zmiennych prezentujących priorytety emerytów – pozwoliła na wyłonienie trzech segmentów. Test ilorazu wariancji wykonany dla zmiennych segmentacyjnych wskazuje, że różnice pomiędzy średnimi wartościami tych zmiennych w analizowanych segmentach można uznać za statystycznie istotne (tabela 1). Następnie na etapie profilowania zweryfikowano, czy cztery wstępnie wybrane cechy demograficzno-ekonomiczne – wiek, wykształcenie, zawód wykonywany przed przejściem na emeryturę oraz dochód – różnicują segmenty w sposób statystycznie istotny⁸. Wyniki testu niezależności chi-kwadrat (tabela 2) potwierdziły to w przypadku dwóch cech – wieku

Tabela 1. Wyniki testu ilorazu wariancji dla zmiennych ilościowych

| Zmienne ilościowe | Wariant I | | Wariant II | | Wariant III | |
|-------------------|----------------|-------------|----------------|-------------|----------------|-------------|
| | statystyka F | wartość p | statystyka F | wartość p | statystyka F | wartość p |
| Cel 1 | 46,45 | 0,000000 | 9,50 | 0,000004 | 5,50 | 0,000996 |
| Cel 2 | 594,87 | 0,000000 | 8,09 | 0,000028 | 9,39 | 0,000005 |
| Cel 3 | 672,98 | 0,000000 | 13,08 | 0,000000 | 6,80 | 0,000166 |
| Cel 4 | 27,12 | 0,000000 | 3,57 | 0,013927 | – | – |
| Cel 5 | 13,33 | 0,000002 | 4,54 | 0,003723 | 3,06 | 0,028006 |
| Cel 6 | 12,06 | 0,000007 | 0,95 | 0,414243 | – | – |
| Cel 7 | 12,90 | 0,000003 | 3,15 | 0,024671 | – | – |
| Cel 8 | 45,16 | 0,000000 | 15,53 | 0,000000 | 10,79 | 0,000001 |
| Cel 9 | 36,94 | 0,000000 | 1,87 | 0,134325 | – | – |
| Cel 10 | 12,42 | 0,000005 | 0,07 | 0,975169 | – | – |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

⁸ Przykład zastosowania podobnej procedury z wykorzystaniem testów statystycznych do analizy danych uzyskanych z prób kwotowych – zob. [Sudbury i Simcock 2009].

i wykształcenia. Oznacza to, że identyfikacja segmentów na podstawie pozostałych cech – zawodu i dochodu – byłaby trudna.

Tabela 2. Wyniki testu niezależności chi-kwadrat dla zmiennych jakościowych

| Zmienne jakościowe | Wariant I | | Wariant II | | Wariant III | |
|--------------------|-------------|------------------|-------------|------------------|-------------|------------------|
| | chi-kwadrat | wartość <i>p</i> | chi-kwadrat | wartość <i>p</i> | chi-kwadrat | wartość <i>p</i> |
| Wiek | 48,64 | 0,00000 | 169,67 | 0,0000 | 173,9281 | 0,0000 |
| Wykształcenie | 17,37 | 0,00802 | 441,21 | 0,0000 | 416,0499 | 0,0000 |
| Dochód | 12,01 | 0,06175 | 561,41 | 0,0000 | 574,1179 | 0,0000 |
| Zawód | 22,84 | 0,11797 | 175,79 | 0,0000 | 186,1484 | 0,0000 |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

Kolejny, II wariant segmentacji został oparty na rozszerzonym zestawie zmiennych segmentacyjnych, obejmującym zarówno ilościowe zmienne ukazujące priorytety życiowe emerytów, jak i zmienne demograficzno-ekonomiczne. Następnie przeprowadzono test niezależności chi-kwadrat, który potwierdził statystycznie istotne różnice między segmentami dla wszystkich czterech zmiennych jakościowych. Z kolei test ilorazu wariancji dla niektórych zmiennych ilościowych nie potwierdził statystycznie istotnych różnic. W związku z tym w sposób sekwencyjny zmienne statystycznie nieistotne były eliminowane z zestawu zmiennych segmentacyjnych aż do osiągnięcia stanu, w którym pozostały wyłącznie zmienne, dla których po ponownym przeprowadzeniu segmentacji oba testy wskazywały na występowanie statystycznie istotnych różnic między segmentami. W ten sposób zredukowano zestaw zmiennych segmentacyjnych do pięciu zmiennych ilościowych oraz czterech zmiennych jakościowych (wariant segmentacji III).

Ze względu na występowanie wśród zmiennych segmentacyjnych w wariancie II zmiennych, dla których nie potwierdzono statystycznie istotnych różnic między segmentami, wyniki te nie mogą być traktowane jako ostateczne. W konsekwencji w dalszej części artykułu zostaną scharakteryzowane segmenty uzyskane metodą *k*-średnich: wariant I (bez zmiennych demograficzno-ekonomicznych w zestawie zmiennych segmentacyjnych) i wariant III (z uwzględnieniem zmiennych demograficzno-ekonomicznych w zestawie zmiennych segmentacyjnych).

5. Charakterystyka uzyskanych segmentów – warianty I i III

Segmentacja metodą *k*-średnich w wariancie I pozwoliła wyodrębnić trzy grupy emerytów wyraźnie różniące się pod względem priorytetów (tabela 3). Najliczniejszy był segment I_C , który obejmował 39,3% respondentów, następny pod względem liczności był segment I_A , w skład którego wchodziło 33,8%

emerytów, zaś najmniej liczny segment I_B skupiał 26,9% osób starszych. Emeryci z segmentu I_A średnio nieco ponad 62 tys. zł z potencjalnej wygranej 100 tys. zł byli gotowi przekazać dzieciom lub wnukom. Z kolei segment I_B skupiał osoby, które największą część wygranej – przeciętnie nieco ponad 54 tys. zł – ulokowałyby w banku. W ostatnim z otrzymanych segmentów, I_C , znalazły się osoby, które nie miały aż tak sprecyzowanych priorytetów życiowych, przy czym ponad 60% potencjalnej wygranej przeznaczyłyby na cztery cele: remont domu lub mieszkania, darowiznę dla dzieci lub wnuków, przechowanie w domu „na czarną godzinę” oraz zakup nowego samochodu lub innych dóbr trwałego użytku.

Tabela 3. Podział wygranej 100 tys. zł w segmentach uzyskanych metodą k -średnich – wariant I

| Zmienne ilościowe | Segment I_A | Segment I_B | Segment I_C |
|-------------------|---------------|---------------|---------------|
| Cel 1 | 7 084 | 2 689 | 15 792 |
| Cel 2 | 62 233 | 17 611 | 16 788 |
| Cel 3 | 9 636 | 54 129 | 10 411 |
| Cel 4 | 8 396 | 10 355 | 17 692 |
| Cel 5 | 3 699 | 3 808 | 7 181 |
| Cel 6 | 1 944 | 1 821 | 4 067 |
| Cel 7 | 903 | 793 | 4 581 |
| Cel 8 | 4 267 | 7 033 | 14 426 |
| Cel 9 | 1 305 | 970 | 3 949 |
| Cel 10 | 537 | 789 | 5 113 |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

Identyfikacja segmentów uzyskanych w ramach wariantu I, które są wyraziste pod względem priorytetów, nie jest łatwa nawet na podstawie cech demograficznych, takich jak wiek i wykształcenie, dla których stwierdzono statystycznie istotne różnice między segmentami (tabela 4). W przypadku wszystkich segmentów dominującą grupą są osoby z wykształceniem średnim. Nieco korzystniej prezentują się możliwości identyfikacji na podstawie wieku. W segmencie I_A dominują osoby w wieku 66–80 lat, w segmencie I_B są to osoby mające 66–75 lat, zaś w segmencie I_C najliczniejszą grupę stanowią emeryci w wieku do 65 lat.

Segmentacja przeprowadzona w wariantcie III (na podstawie zestawu uwzględnianego zmienne demograficzno-ekonomiczne) pozwoliła na wyodrębnienie czterech grup emerytów, przy czym nie były one tak wyraziste pod względem priorytetów życiowych jak w wariantcie I. Najwięcej emerytów znalazło się

Tabela 4. Struktura segmentów uzyskanych metodą k-średnich – wariant I dla cech profilujących (w %)

| Zmienne profilujące | Segment I _A | Segment I _B | Segment I _C |
|---------------------|------------------------|------------------------|------------------------|
| Wiek | | | |
| Do 60 lat | 8,6 | 17,0 | 20,3 |
| 61–65 lat | 13,9 | 13,3 | 24,2 |
| 66–70 lat | 22,0 | 23,6 | 16,8 |
| 71–75 lat | 21,1 | 29,1 | 19,1 |
| 76–80 lat | 21,5 | 12,1 | 9,8 |
| 81 i więcej lat | 12,9 | 4,8 | 9,8 |
| Wykształcenie | | | |
| Podstawowe | 23,8 | 12,6 | 21,1 |
| Zawodowe | 26,7 | 19,8 | 29,3 |
| Średnie | 33,8 | 46,1 | 34,0 |
| Wyższe | 15,7 | 21,6 | 15,6 |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

w segmencie III_B – 28,7%. Segment III_A obejmował 27,0% respondentów. Kolejny pod względem liczności był segment III_C skupiający 24,7% emerytów, a relatywnie najmniej osób starszych zostało włączonych w skład segmentu III_D – 19,6%. W przypadku trzech segmentów (III_A, III_C oraz III_D) największą część wygranej emeryci przekazaliby dzieciom lub wnukom – od 32 tys. zł do niemal 40 tys. zł (tabela 5). W ostatnim z czterech wyodrębnionych segmentów (III_B) priorytetem okazało się ulokowanie środków w banku – na ten cel zostałyby przeznaczone 27,5 tys. zł. Dalsze różnice między segmentami można dostrzec, analizując priorytety życiowe (tabela 5).

Tabela 5. Podział wygranej 100 tys. zł w segmentach uzyskanych metodą k-średnich – wariant III

| Zmienne ilościowe | Segment III _A | Segment III _B | Segment III _C | Segment III _D |
|-------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| Cel 1 | 12 242 | 7 609 | 5 879 | 10 966 |
| Cel 2 | 34 682 | 24 245 | 39 872 | 32 387 |
| Cel 3 | 15 585 | 27 574 | 21 677 | 21 975 |
| Cel 5 | 5 011 | 6 463 | 5 093 | 3 326 |
| Cel 8 | 7 063 | 14 052 | 7 324 | 7 951 |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

Tabela 6. Struktura segmentów uzyskanych metodą k -średnich – wariant III dla cech jakościowych (w %)

| Zmienne ilościowe | Segment III _A | Segment III _B | Segment III _C | Segment III _D |
|--|--------------------------|--------------------------|--------------------------|--------------------------|
| Wiek | | | | |
| Do 60 lat | 12,8 | 23,4 | 15,4 | 10,2 |
| 61–65 lat | 14,1 | 36,1 | 5,1 | 12,0 |
| 66–70 lat | 40,9 | 15,2 | 11,0 | 12,0 |
| 71–75 lat | 6,7 | 6,3 | 39,7 | 42,6 |
| 76–80 lat | 14,8 | 12,0 | 17,6 | 13,0 |
| 81 i więcej lat | 10,7 | 7,0 | 11,0 | 10,2 |
| Wykształcenie | | | | |
| Podstawowe | 12,1 | 1,3 | 10,3 | 73,1 |
| Zawodowe | 65,1 | 12,0 | 11,0 | 13,0 |
| Średnie | 12,8 | 60,1 | 60,3 | 7,4 |
| Wyższe | 10,1 | 26,6 | 18,4 | 6,5 |
| Zawód wykonywany przed przejściem na emeryturę | | | | |
| Robotnik, gospodyni domowa | 55,0 | 26,6 | 30,9 | 47,2 |
| Technik, inżynier | 8,7 | 13,9 | 12,5 | 4,6 |
| Kadra ekonomiczna, nauczyciel | 8,1 | 20,3 | 22,8 | 5,6 |
| Pracownik administracji państwowej | 5,4 | 15,8 | 13,2 | 1,9 |
| Wolny zawód | 4,7 | 16,5 | 13,2 | 0,9 |
| Rolnik | 7,4 | 0,6 | 1,5 | 33,3 |
| Biznesmen | 10,1 | 6,3 | 5,9 | 4,6 |
| Inne | 0,7 | 0,0 | 0,0 | 1,9 |
| Dochód | | | | |
| Do 1000 zł | 19,5 | 12,0 | 7,4 | 76,9 |
| 1001–1400 zł | 59,1 | 8,9 | 17,6 | 13,9 |
| 1401–2000 zł | 10,7 | 10,1 | 70,6 | 5,6 |
| 2000 i więcej zł | 10,7 | 69,0 | 4,4 | 3,7 |

Źródło: opracowanie własne na podstawie przeprowadzonych analiz.

Bardzo wyraziste są otrzymane w ramach wariantu III segmenty emerytów pod względem cech demograficzno-ekonomicznych. Segment III_A stanowią głównie osoby z wykształceniem zawodowym, pobierające emeryturę niższą od średniej. Emeryci ci przed przejściem na emeryturę należeli do grup zawodowych, takich jak robotnicy oraz gospodynie domowe, i najczęściej są w wieku 66–70 lat. Z kolei segment III_B obejmuje przede wszystkim emerytów z wykształ-

ceniem średnim, którzy otrzymują w przeważającej części świadczenia emerytalne powyżej średniego poziomu. Najczęściej wcześniej byli oni robotnikami, gospodyniami domowymi, nauczycielami lub reprezentowali kadry ekonomiczne, a ponadto są relatywnie najmłodszy – największą część segmentu stanowią osoby w wieku do 65 lat. Segment III_C składa się z osób, które najczęściej legitymują się średnim wykształceniem i otrzymują emerytury nieco niższe lub nieznacznie wyższe od średniej. Osoby te wcześniej były robotnikami, gospodyniami domowymi, nauczycielami lub też były zaliczane do kadr ekonomicznych, przy czym relatywnie duży odsetek stanowią emeryci powyżej 70 roku życia. Ostatni segment, III_D, tworzą emeryci, którzy w znacznej części posiadają wykształcenie podstawowe i uzyskują najniższe emerytury, nieprzekraczające 1000 zł. Przed przejściem na emeryturę najczęściej byli oni robotnikami, gospodyniami domowymi lub rolnikami, znaczna ich część ma 71 lat i więcej.

6. Podsumowanie

Emeryci są grupą zróżnicowaną pod względem priorytetów życiowych, jednak segmenty wyróżnione wyłącznie na podstawie zmiennych ukazujących te priorytety są trudne do jednoznacznego zidentyfikowania na podstawie cech deskryptywnych. Otrzymane wyniki segmentacji potwierdzają, że uwzględnienie zmiennych demograficzno-ekonomicznych w zestawie zmiennych segmentacyjnych umożliwia uzyskanie segmentów, których identyfikacja na podstawie tych cech będzie łatwiejsza niż w przypadku klasycznej procedury segmentacji opartej na cechach behawioralnych i potraktowania cech demograficzno-ekonomicznych jako profilujących. Prowadzi to jednak również do osłabienia wpływu zmiennych behawioralnych na wyniki grupowania, przez co wyodrębnione segmenty różnią się w mniejszym stopniu od siebie pod względem behawioralnym. Wyniki te wskazują, że włączenie zmiennych demograficzno-ekonomicznych do zestawu zmiennych segmentacyjnych nie przynosi wyłącznie korzyści w postaci łatwiejszej identyfikacji segmentów i nie pozostaje bez wpływu na wynik segmentacji, a nawet może w znacznym stopniu zmienić ostateczny skład segmentów.

Literatura

- Balicki A. [2009], *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
- Blashfield R.K. [1976], *Mixture Model Tests of Cluster Analysis: Accuracy of Four Agglomerative Hierarchical Methods*, „The Psychological Bulletin”, vol. 83, nr 3.

- Bombol M., Słaby T. [2011], *Konsument 55+ wyzwaniem dla rynku*, Oficyna Wydawnicza SGH, Warszawa.
- Diagnoza społeczna. Warunki i jakość życia Polaków* [2011], red. J. Czapiński, T. Panek, Rada Monitoringu Społecznego, Warszawa.
- Fatuła D. [2010], *Zachowania polskich gospodarstw domowych na rynku finansowym*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
- Jurek Ł. [2012], *Ekonomia starzejącego się społeczeństwa*, Difin, Warszawa.
- Kaczmarczyk S. [2007], *Zastosowania badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa*, PWE, Warszawa.
- Migut G. [2004], *Jak znaleźć grupy podobnych klientów, czyli metody segmentacji*, <http://www.statsoft.pl/czytelnia/finanse/pdf/04segment.pdf> (dostęp: 5.10.2013).
- Olejnik I. [2012], *Zachowania emerytów w Polsce – determinanty i segmentacja* [w:] *Zachowania konsumenckie – badania, uwarunkowania, różnice*, t. 1, „Handel Wewnętrzny”, maj–czerwiec.
- Psychologia ekonomiczna* [2004], red. T. Tyszka, GWP, Gdańsk.
- Rondan-Cataluna F.J., Sanchez-Franco M.J., Villarejo-Ramos A.F. [2010], *Searching for Latent Class Segments in Technological Services*, „The Service Industries Journal”, vol. 30, nr 6.
- Rószkiewicz M. [2002], *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa.
- Rószkiewicz M. [2011], *Analiza klienta*, SPSS Polska, Kraków.
- Rujasiri P., Chomtee B. [2009], *Comparison of Clustering Techniques for Cluster Analysis*, http://kasetsartjournal.ku.ac.th/kuj_files/2009/a0912091419338281.pdf (dostęp: 5.12.2013).
- Sagan A. [2004], *Badania marketingowe: podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Sudbury L., Simcock P. [2009], *A Multivariate Segmentation Model of Senior Consumers*, „Journal of Consumer Marketing”, vol. 26, nr 4.

The *k*-means Method in the Segmentation of Retirees on the Basis of Life Priorities

The article compares the results of the segmentation of retirees done using the *k*-means method on the basis of a set of behavioural variables and a combined set of behavioural, demographic and economic variables. On the basis of the behavioural set, segments of retirees which showed significant differences in the area of life priorities were obtained. Between these groups there were relatively small differences in their demographic and economic profile. In the case of the segmentation done on the basis of the combined set of variables, the results of segmentation were influenced to a greater extent by the demographic and economic variables. The segments of retirees became more easily identifiable, but the differences in the area of behavioural features decreased.

Keywords: segmentation, *k*-means method, retirees, quantitative research.