

**Przegląd  
literatury naukowej  
w analizach współczesnych  
wyzwań gospodarczych**

**Prace uczestników  
programu Młody  
Naukowiec SGH**



REDAKCJA NAUKOWA **RADOSŁAW MALIK**

# **Przegląd literatury naukowej w analizach współczesnych wyzwań gospodarczych**

**Prace uczestników  
programu Młody  
Naukowiec SGH**

**Recenzje**

Michał Choraś

Wojciech Mazurczyk

**Redakcja językowa polska**

Anna Gaudy-Piątek

**Redakcja językowa angielska**

Mirosław Szymański

© Copyright by Szkoła Główna Handlowa w Warszawie, Warszawa 2022

Wszelkie prawa zastrzeżone. Kopiowanie, przedrukowywanie i rozpowszechnianie całości lub fragmentów niniejszej publikacji bez zgody wydawcy zabronione.

Wydanie I

**ISBN 978-83-8030-560-1**

Oficyna Wydawnicza SGH – Szkoła Główna Handlowa w Warszawie

02-554 Warszawa, al. Niepodległości 162

[www.wydawnictwo.sgh.waw.pl](http://www.wydawnictwo.sgh.waw.pl)

e-mail: [wydawnictwo@sgh.waw.pl](mailto:wydawnictwo@sgh.waw.pl)

**Projekt i wykonanie okładki**

Magdalena Limbach

**Skład i łamanie**

DM Quadro

**Druk i oprawa**

[volumina.pl](http://volumina.pl) Daniel Krzanowski

ul. Ks. Witolda 7–9

71–063 Szczecin

tel. 91 812 09 08

e-mail: [druk@volumina.pl](mailto:druk@volumina.pl)

Zamówienie 78/VII/22

# Spis treści

<b>Wprowadzenie</b> .....	7
Radostław Malik	
<b>Teoria oczekiwań diagnostycznych w literaturze ekonomicznej</b> .....	11
Paweł Galiński	
<b>Strategiczna polaryzacja polityczna – perspektywa informacji, niepewności i wartościowości</b> .....	35
Anna Halewska	
<b>Wpływ czynników ESG/CSR na koszt kapitału przedsiębiorstw europejskich: przegląd literatury</b> .....	49
Piotr Hościłowicz	
<b>Aktywne i pasywne strategie inwestycyjne na rynkach akcyjnych – przegląd systematyczny literatury naukowej</b> .....	69
Mateusz Izdebski	
<b>Status społeczno-ekonomiczny a wyniki w nauce w szkołach ponadpodstawowych – systematyczny przegląd literatury</b> .....	89
Jan Kroszka	
<b>Nieracjonalne oczekiwania na rynkach finansowych – przegląd literatury finansów behawioralnych</b> .....	111
Joanna Krysta	
<b>Zakres oraz determinanty ujawnień dotyczących wartości firmy w sprawozdaniach finansowych według MSR/MSSF – systematyczny przegląd literatury</b> .....	133
Michał Siczek	
<b>Analiza sentymentu w języku chińskim z wykorzystaniem metod uczenia głębokiego – systematyczny przegląd literatury</b> .....	147
Marcin Żurek	



# Wprowadzenie

Radosław Malik

Niniejsza monografia naukowa jest zbiorem ośmiu rozdziałów przygotowanych przez studentów Szkoły Głównej Handlowej w Warszawie, uczestników trzeciej edycji programu „Młody Naukowiec SGH”. Teksty zebrane w monografii stanowią systematyczne przeglądy literatury, które podsumowują stan badań naukowych w wybranych przez uczestników obszarach. W tekstach dokonano syntezy dotychczasowych osiągnięć naukowych, co stanowi ważny punkt wyjścia do dalszych, empirycznych badań. Dzięki adekwatnie sformułowanym pytaniom badawczym i precyzyjnemu wyborowi literatury, analizy zawarte w tej monografii mogą stanowić inspirację do dalszych badań naukowych podejmowanych przez autorów tych tekstów oraz innych badaczy.

Zamierzeniem programu „Młody Naukowiec SGH” jest rozwój kompetencji naukowych utalentowanych słuchaczy studiów licencjackich i magisterskich, którzy pragną realizować swoje aspiracje badawcze jeszcze na tym etapie oraz kontynuować je w przyszłości. Za miarę sukcesu programu „Młody Naukowiec SGH” można uznać znaczny zakres tematyczny przeglądów literatury zgromadzonych w niniejszej monografii, co dowodzi jak różnorodne i często nowatorskie zainteresowania badawcze mają uczestnicy. Rozdziały zawarte w monografii stanowią często pierwsze teksty naukowe przygotowane przez autorów, a charakterystyka tych analiz wskazuje na znaczne możliwości kontynuowania zainicjowanych badań.

W pierwszym rozdziale monografii Paweł Galiński podjął tematykę teorii oczekiwań diagnostycznych, która, odnosząc się m.in. do zjawisk negacji ryzyka i nadreakcji na szoki, podejmuje próbę pogłębiania zrozumienia nieracjonalności w prognozach podmiotów gospodarczych. W rozdziale przekonująco pokazano jak teoria oczekiwań diagnostycznych, przy wykorzystaniu dorobku psychologii i zastosowaniu narzędzi matematycznych, pomaga wyjaśnić nieracjonalne zachowania podmiotów gospodarczych oraz złożone procesy w całych gospodarkach.

Drugi rozdział, napisany przez Annę Halewską, podejmuje interesujące zagadnienie badawcze strategicznej polaryzacji politycznej, pokazując, jak w dotychczasowych badaniach zjawisko to było analizowane z wykorzystaniem teorii gier. Tekst

autorki stanowi udane podsumowanie stanu wiedzy tego problemu w ujęciu perspektywy informacji, niepewności i wartościowości oraz ma znaczny potencjał do praktycznego zastosowania wyników przeprowadzonej analizy.

Trzeci rozdział monografii, którego autorem jest Piotr Hościłowicz, dotyczy aktualnej tematyki znaczenia czynników społecznej odpowiedzialności biznesu w procesie alokacji kapitału w Europie. Interesujące wnioski z dokonanego przeglądu literatury mają istotne walory aplikacyjne, wskazujące na znaczenie tych czynników w kreowaniu wartości przedsiębiorstwa oraz pokazujące brak ich jednoznacznego wpływu na koszt pozyskania kapitału dla europejskich przedsiębiorstw.

W czwartym rozdziale Mateusz Izdebski dokonał systematycznego przeglądu literatury, dotyczącego porównania aktywnych i pasywnych strategii inwestycyjnych na rynku akcji. Przeprowadzona analiza prowadzi do ciekawych wniosków, które wskazują m.in. na wzrost popularności strategii pasywnych i wiążące się z tym następstwa dla uczestników rynku oraz ogólnie niższą zyskowność strategii aktywnych, które jednak pozostają popularne wśród uczestników rynku ze względu na ich powodzenie w określonych okolicznościach.

Piąty rozdział monografii opracowany przez Jana Kroszkę podejmuje istotne społecznie zagadnienie zależności między statusem społeczno-ekonomicznym dzieci a ich wynikami w nauce w państwach wysokorozwiniętych. Przegląd literatury dotyczący tego zagadnienia, które w ostatnich latach koncentruje znaczną uwagę badaczy, prowadzi do ważnych wniosków mających praktyczne implikacje dla różnych grup aktorów związanych z systemem edukacji oraz wskazuje na obiecujące kierunki dalszych badań.

W szóstym rozdziale monografii Joanna Krysta przeanalizowała implikacje rozluźnienia założeń o racjonalnych oczekiwaniach w dziedzinie finansów behawioralnych. Wyniki analizy przeprowadzonej w tekście mogą mieć liczne biznesowe i gospodarcze zastosowania, wskazując m.in. na powszechność nieracjonalnych oczekiwań, analizując ich źródła oraz podkreślając ich znaczenie dla decyzji inwestycyjnych i oczekiwań dotyczących stopy zwrotu.

W siódmym rozdziale opracowanym przez Michała Siczka przedmiotem analizy jest zagadnienie utraty wartości firmy, ze szczególnym uwzględnieniem tematu testów na utratę wartości przedsiębiorstwa. Autor rozdziału wskazuje m.in., że zakres ujawnień dotyczących wartości firmy w sprawozdaniach finansowych przygotowanych zgodnie z międzynarodowymi standardami rachunkowości jest często niekompletny, co może prowadzić do istotnych następstw gospodarczych. Istotnym walorem tego rozdziału jest wskazanie kierunków dalszych badań o potencjalnie dużym znaczeniu dla rozwoju tej tematyki badawczej.



W ostatnim, ósmym rozdziale niniejszej monografii Marcin Żurek przeprowadził systematyczny przegląd literatury na nowatorski temat metod uczenia głębokiego, wykorzystywanych do analizy sentymentu w tekstach w języku chińskim. Autor zidentyfikował obiecujące kierunki dalszych badań nad analizą sentymentu w języku chińskim, które wskazują na potrzebę rozwoju wiedzy m.in. w obszarach metod reprezentacji tekstu chińskiego, analiz porównawczych modeli oraz badań parametrów poszczególnych modeli.

Teksty zgromadzone w monografii łączą zastosowanie metody przeglądu literatury, która została wykorzystana do analiz różnorodnej i często nowatorskiej tematyki badawczej. Wyniki przeglądów literatury mają istotne znaczenie dla systematyzacji wiedzy w analizowanych obszarach oraz stanowią zapowiedź dalszych badań naukowych. Poza walorami badawczymi teksty wskazują na potencjał i naukowe ambicje uczestników programu „Młody Naukowiec SGH”.

Jako redaktor naukowy tej monografii chciałbym podziękować osobom, które wspierały uczestników trzeciej edycji programu „Młody Naukowiec SGH”. Pani Prorektor ds. nauki dr hab. Agnieszce Chłoń-Domińczak za inspirujący przykład i motywowanie uczestników do rozwoju ambicji naukowych. Opiekunom naukowym uczestników programu za radę i pomoc, jaką służyli autorom w czasie przygotowania rozdziałów. Autorom warsztatów rozwijających kompetencje naukowe uczestników programu za inspiracje i hojne dzielenie się swoim doświadczeniem. Pani Annie Ulatowskiej z działu nauki SGH za opiekę nad uczestnikami i pomoc udzielaną im we wszelkich sprawach organizacyjno-administracyjnych.



# Teoria oczekiwań diagnostycznych w literaturze ekonomicznej

Paweł Galiński

## Streszczenie

Kryzys finansowy 2007–2008 zmotywował ekonomistów do stworzenia teorii oczekiwań diagnostycznych (DE, *Diagnostic Expectations*), która jest w stanie wytłumaczyć nieracjonalność w prognozach tworzonych przez podmioty gospodarcze przejawiającą się negacją ryzyka (*neglect of risk*) oraz nadreakcją na szoki (*overreaction*). Celem przeglądu literatury jest podsumowanie założeń teorii DE oraz omówienie wyników, których ona dostarcza. W artykule podsumowano informacje z 12 wyselekcjonowanych artykułów, a w szczególności pokazano, że: 1) dane empirycznie bezpośrednio wykazują, że podmioty gospodarcze podejmują decyzje na podstawie nieracjonalnych oczekiwań, 2) źródła tej nieracjonalności doszukiwać się możemy w heurystyce reprezentatywności, 3) teoria oczekiwań diagnostycznych jest głęboko osadzona w psychologii, przy zachowaniu matematycznej elegancji i prostoty, oraz że 4) oczekiwania diagnostyczne są w stanie pomóc nam zrozumieć zachowanie poszczególnych podmiotów, jak i całej gospodarki.

**Słowa kluczowe:** oczekiwania diagnostyczne, nieracjonalne oczekiwania, nadreakcja, negacja ryzyka, heurystyka reprezentatywności, kryzys finansowy

## 1. Wstęp

### 1.1. Wprowadzenie do tematu przeglądu literatury

Do czasu wybuchu kryzysu finansowego 2007–2008 wśród ekonomistów dominowało przekonanie, że podmioty gospodarcze kształtują swoje oczekiwania na temat przyszłości w sposób racjonalny. Nowoczesne pojęcie racjonalności zdefiniował Muth [1961] w ramach swojej hipotezy racjonalnych oczekiwań (REH, *Rational Expectations Hypothesis*). Zakłada ona, że każdy z nas przy podejmowaniu decyzji kieruje się

oczekiwaniami co do przyszłości, które to mają być konstruowane w sposób racjonalny, tzn., że przy ich tworzeniu korzystamy ze wszystkich dostępnych informacji, a sama prognoza jest statystycznie optymalna i pozbawiona systematycznych błędów. Warto przy tym jednak zaznaczyć, że ekonomiści są świadomi tego, iż ludzki proces wnioskowania nie jest idealny, oraz że ludzie popełniają błędy przy podejmowaniu decyzji i konstruowaniu oczekiwań co do przyszłości. Niemniej założenie racjonalnych oczekiwań (RE, *Rational Expectations*) oznacza, że błędy te są na tyle mało istotne, że nie wpływają znacząco na działalność gospodarstw domowych, przedsiębiorstw czy całej gospodarki. Lucas Jr. [1972] argumentuje dodatkowo, że RE pozwalają na ustrukturyzowanie dyskusji na temat zachowania podmiotów gospodarczych, zatem warto je stosować. Z tych też powodów hipoteza racjonalnych oczekiwań jest do dzisiaj powszechnie używana, co wynika również z tego, że RE w prosty sposób mogą zostać zaimplementowane do modelowania ekonomicznego [m.in. Iacoviello, 2005; Brzoza-Brzezina, Kolasa, Makarski, 2021; Mukhin, 2022].

Zachowanie rynków finansowych przed i po kryzysie 2007–2008 pokazało jednak, że ludzki mechanizm kształtowania oczekiwań jest obciążony wieloma błędami poznawczymi, a przede wszystkim nadreakcją na szoki, które w istotny sposób wpływają na gospodarkę. W amerykańskiej gospodarce od końca lat 90. narastała bańka cenowa na rynku nieruchomości. Wzrost cen wynikał z faktu, że banki udzielały wielu kredytów hipotecznych gospodarstwom domowym, a kredyty te były następnie sprzedawane<sup>1</sup> z zyskiem na rynkach finansowych [Gennaioli, Shleifer, 2018]. Wraz z udzielaniem kredytów rosły ceny mieszkań, co przyczyniało się do zaciągania kolejnych kredytów przez gospodarstwa domowe, które to jeszcze bardziej napędzały wzrost cen. W ten sposób powstało zamknięte koło wzrostu cen, które sprawiło, że w okresie 1996–2008 ceny nieruchomości w USA wzrosły o 200% [Shiller, 2015]. Pomimo pojawiających się informacji o niewypłacalności sporej części kredytobiorców oraz o istnieniu bańki cenowej, rynki finansowe, w tym inwestorzy, zdawali się nie być świadomi ryzyka, jakie narastało w gospodarce. Ryzyko to zaczęło się materializować w 2006 r. wraz z brakiem spłaty sporej części kredytów hipotecznych, co w konsekwencji doprowadziło do paniki w systemie finansowym, która spowodowała bankructwo wielu instytucji finansowych, a ostatecznie globalną recesję.

Dziesięć lat po kryzysie Uniwersytet Chicago przeprowadził ankietę wśród zachodnich ekonomistów<sup>2</sup> [Gennaioli, Shleifer, 2018], z której wynika, że ich zda-

<sup>1</sup> W postaci instrumentów pochodnych np. typu CDO (*Collateralized Debt Obligation*) [Gennaioli, Shleifer 2018].

<sup>2</sup> W ankiecie udział wzięli m.in. Daron Acemoglu (Harvard), William Nordhaus (Yale) i Richard Tayler (Chicago).

niem drugą najważniejszą przyczyną kryzysu finansowego, zaraz po „zawodnym nadzorze nad systemem finansowym”, była „negacja ryzyka” w następstwie błędów poznawczych, takich jak: 1) nad optymizm (*overoptimism*), 2) nadmierna pewność siebie (*overconfidence*) czy 3) ignorowanie prawdopodobieństwa wystąpienia skrajnie negatywnych zdarzeń (*base rate neglect*).

Odkrycie, że błędy poznawcze mają istotny wpływ nie tylko na rynki finansowe, ale i na całą gospodarkę, skłoniło ekonomistów do pracy nad nową teorią oczekiwań, która tłumaczyłaby nieracjonalność w ludzkim procesie rozumowania. Jedną z dominujących teorii nieracjonalności oczekiwań jest tzw. podejście „uczenia się” (*learning*)<sup>3</sup>. Zakłada ono, że każdy podmiot gospodarczy zachowuje się jak ekonometryk, który przy pomocy danych historycznych tworzy prognozy statystyczne na temat przyszłych wartości zmiennych (np. poziomu cen czy konsumpcji), i na podstawie tych prognoz podejmuje decyzje. Podejście to stosowane jest w szczególności w modelowaniu ekonomicznym i pozwala na uzyskanie znacznie większego dopasowania modeli do danych empirycznych. Co więcej, podejście „uczenia się” jest w stanie endogenicznie wygenerować silniejsze reakcje podmiotów gospodarczych na szoki, co jest znacząco bliższe wynikom badań empirycznych [Brzoza-Brzezina, Suda 2021; Eusepi, Preston, 2011]. Niemniej istotnym problemem tego podejścia jest fakt, że nie jest ono silnie osadzone w psychologii i przez to nie jest w stanie od podstaw wytłumaczyć nam źródeł występowania błędów poznawczych w procesach gospodarczych.

Najbardziej obiecującą teorią wydaje się być w tym kontekście teoria oczekiwań diagnostycznych, która charakteryzuje się silnym osadzeniem w psychologii oraz matematyczną prostotą i elegancją. Twórcy teorii oczekiwań diagnostycznych (DE, *Diagnostic Expectations*) oparli ją na odkryciach ekonomistów behawioralnych Daniela Kahnemana i Amosa Tverskiego, którzy wprowadzili do literatury pojęcie heurystyki [Kahneman, Tversky, 1972; Tversky, Kahneman, 1983]. Heurystyka to uproszczony proces wnioskowania, którym na co dzień posługują się ludzie. Korzystamy z heurystyk przy podejmowaniu większości codziennych decyzji, ponieważ wymaga to od nas mniej skupienia, a sama decyzja wydawana jest szybciej. Istnienia heurystyk dowiedziono eksperymentalnie i, co istotne, udało się udowodnić, że prowadzą one do występowania błędów poznawczych. Teoria DE bazuje na jednej z heurystyk – heurystyce reprezentatywności – i zakłada, że podmioty gospodarcze przy konstruowaniu prognoz przywiązują nieproporcjonalnie dużo uwagi do najbardziej „dostępnych” (reprezentatywnych) informacji. Heurystykę reprezentatywności

<sup>3</sup> Wśród innych przykładów analizy nieracjonalności oczekiwań możemy wyróżnić również podejście „krótkowzroczności” (*short seeing*) zakładające, że trafność oczekiwań spada w kolejnych, bardziej oddalonych okresach [Gabaix, 2020].

udało się opisać z wykorzystaniem aparatu matematycznego i sprowadzić do prostego równania oczekiwań diagnostycznych, co pozwoliło na wdrożenie DE do istniejących już modeli ekonomicznych. Wyniki uzyskane w modelach z DE są w stanie wyjaśnić nam wiele niezrozumiałych dotychczas mechanizmów, do których dochodzi na rynkach finansowych i w całej gospodarce [Bordalo, Gennaioli, Kwon, Shleifer, 2021b; Gennaioli, Shleifer, 2018].

W poniższym przeglądzie literatury omówiono zagadnienie oczekiwań diagnostycznych, tj. teorię, która jest w stanie potencjalnie wyjaśnić mechanizm podejmowania decyzji przez podmioty gospodarcze na podstawie nieracjonalnych oczekiwań. W tym celu przedstawiono zagadnienie DE od psychologicznych podstaw po implementację do modeli ekonomicznych oraz przedstawiono wyniki uzyskane przez naukowców korzystających z DE w swoich badaniach. Ze względu na fakt, że teoria DE jest nową koncepcją<sup>4</sup>, a samo zagadnienie odbiega od klasycznej teorii ekonomii, w poniższym przeglądzie literatury posłużono się przykładami mającymi na celu pomóc w budowie odpowiedniej intuicji, kluczowej dla zrozumienia DE<sup>5</sup>.

Artykuł podzielono na pięć części. W części drugiej bardziej szczegółowo przedstawiono przykłady występowania nieracjonalnych oczekiwań w danych ekonomicznych. W części trzeciej wyjaśniono pojęcie heurystyki reprezentatywności oraz sposób matematycznego wyprowadzenia na jej podstawie oczekiwań diagnostycznych, wraz z omówieniem własności DE. W czwartej części podsumowano wyniki uzyskane przez naukowców z wdrożenia DE do modeli ekonomicznych. W ostatniej części przedstawiono krótką dyskusję na temat DE wraz z propozycjami kolejnych kierunków badań na ten temat.

## 1.2. Metodyka doboru literatury

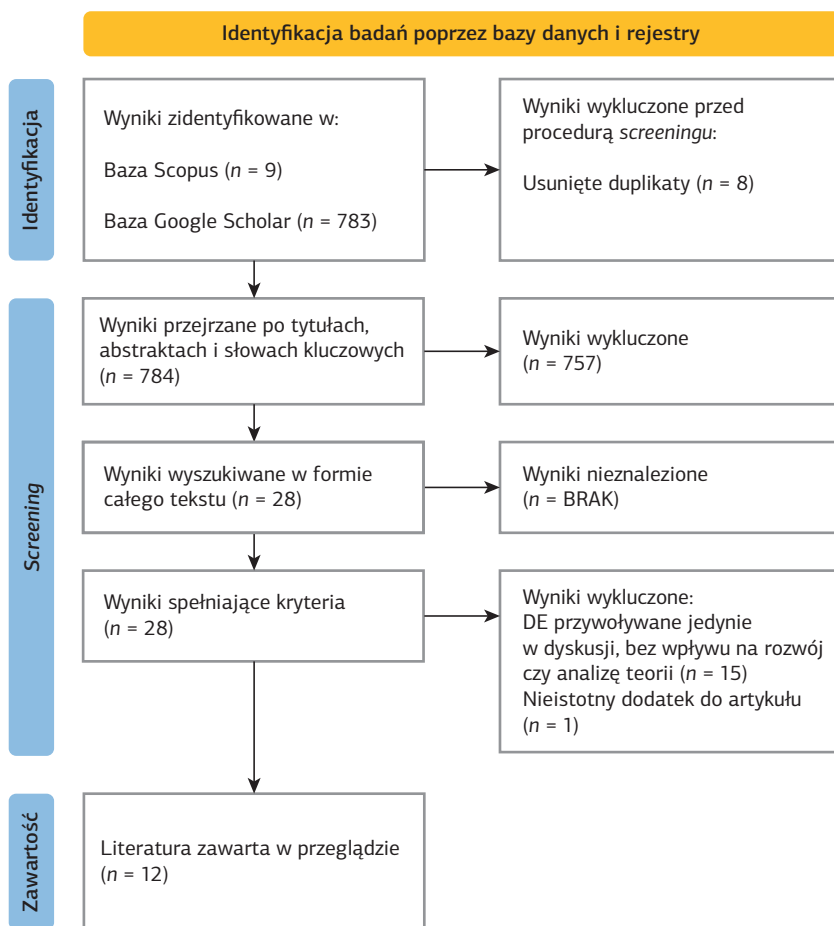
Źródła wybrane do przeglądu literatury zostały dobrane zgodnie z manifestem PRISMA 2021 [Page i in., 2021]. W fazie identyfikacji skorzystano z bazy Scopus, w ramach której dla hasła „Diagnostic Expectations” oraz przy zawężeniu wyszukiwania do nauk ekonomicznych uzyskano 9 rezultatów. Ze względu na to, że teoria DE jest teorią nową, zdecydowano się skorzystać także z bazy Google Scholar, która wyszukuje również nierecenzowane artykuły naukowe, w tym względnie niedawno opublikowane prace NBER (*National Bureau of Economic Research*), gdzie pojawia się duża część nowych artykułów na temat DE. Baza Google Scholar na hasło „Diagno-

<sup>4</sup> Pierwszy raz pojęcie oczekiwań diagnostycznych zastosowano w Bordalo, Gennaioli, Shleifer [2018].

<sup>5</sup> Identyczne podejście stosowane jest w artykułach omawianych w przeglądzie literatury, a część przykładów została bezpośrednio z nich zaczerpnięta.

stic Expectations” dostarczyła 784 rezultaty<sup>6</sup>. Podczas procesu *screeningu* udało się wyłączyć większość literatury oraz uzyskać 12 artykułów do przeglądu (rysunek 1).

**Rysunek 1.** Metodyka doboru literatury do przeglądu na podstawie manifestu PRISMA



Źródło: opracowanie własne na podstawie Page i in. [2021].

Teoria DE jest jednak silnie związana z tematem heurystyki reprezentatywności oraz analizą przyczyn kryzysów finansowych. W związku z tym autorzy w znalezionych 12 artykułach powszechnie opierają swój proces wnioskowania na pracach naukowców, którzy nie korzystali z określenia DE, dlatego też badania te nie pojawiają

<sup>6</sup> Google Scholar nie umożliwia filtrowania rezultatów ze względu na dziedzinę nauki oraz nie pozwala ograniczać rezultatów jedynie do słów kluczowych czy abstraktów, przez co uzyskano tak dużą liczbę rezultatów.

się w bazach, a są silnie związane z omawianym tematem. Co więcej, pojęcie DE pojawiło się w literaturze dopiero w 2017 r. i wiele pionierskich prac związanych z tym zagadnieniem nie korzystało z początku z tego pojęcia<sup>7</sup>. Dlatego też w celu przedstawienia teorii DE w sposób kompletny, w przeglądzie odwołano się również do części literatury cytowanej w znalezionych 12 artykułach. Poza czynnikami ograniczającymi, przedstawionymi powyżej, przy procesie doboru literatury nie napotkano na istotne problemy związane przykładowo z brakiem dostępu do artykułów.

## 2. Dowody występowania nieracjonalnych oczekiwań w danych empirycznych

Nieracjonalność oczekiwań występuje na wielu poziomach gospodarki. W związku z tym tę część przeglądu literatury podzielimy na omówienie przykładów występowania nieracjonalnych oczekiwań w danych na poziomie: 1) przedsiębiorstw, 2) inwestorów oraz 3) całego systemu finansowo-gospodarczego. Na koniec, w celu podsumowania, posłużymy się przykładem kryzysu finansowego 2007–2008.

### 2.1. Nieracjonalność oczekiwań w działalności przedsiębiorstw

Firmy podejmują decyzje biznesowe na podstawie własnych prognoz, które są jednak zawyżane (zaniżane) z powodu nad optymizmu (nad pesymizmu). Gennaioli, Shleifer i Vishny [2015] przeanalizowali dane pochodzące z ankiety systematycznie przeprowadzanej wśród dyrektorów finansowych (CFO, *Chief Financial Officer*) dużych amerykańskich firm. W ramach tej ankiety CFO przedstawiali swoje oczekiwania co do przyszłych zysków oraz planów inwestycyjnych. Okazało się, że byli oni nadmiernie optymistyczni co do przyszłych dochodów swoich firm w okresie dobrych wyników finansowych oraz nadmiernie pesymistyczni w momencie, gdy wyniki były poniżej oczekiwań. Nieracjonalność w tej sytuacji polegała na tym, że ich prognozy były silnie zawyżone (zaniżone) w stosunku do rzeczywistych wartości, w przypadku gdy sytuacja była dobra (zła).

Co więcej, autorzy badania wykazali, że błędy w oczekiwaniach są na tyle systematyczne, że możemy je przewidzieć, co stoi w sprzeczności z hipotezą racjonalnych oczekiwań. W innym badaniu Bordalo, Gennaioli, Shleifer i Terry [2021a] wykazali ilościowo, że błędy w prognozach CFO można przewidywać na podstawie

<sup>7</sup> Początków prac nad teorią DE doszukiwać się można w pracy Gennaioli i Shleifer [2010], czyli na przynajmniej sześć lat przed zastosowaniem pojęcia „oczekiwania diagnostyczne”.



takich zmiennych, jak: 1) wysokość zysku w obecnym okresie, 2) poziom inwestycji w danym okresie, 3) poziom zaciągniętego kredytu. Dodatkowo wyniki sugerują, że:

- firmy konstruujące wysokie prognozy zysków w następnym roku częściej osiągały niższy wynik finansowy, niż zakładały;
- firmy z wyższymi zyskami w danym roku są bardziej niezadowolone z wyniku finansowego w roku następnym;
- firmy inwestujące w danym okresie mają z założenia bardziej pozytywne prognozy.

## 2.2. Nieracjonalność oczekiwań wśród inwestorów

Analogiczny problem występuje wśród inwestorów, którzy przy dokonywaniu decyzji inwestycyjnych polegają na swoich oczekiwaniach, które również obciążone są systematycznym błędem. Greenwood i Shleifer [2014] pokazali na podstawie dużego badania przeprowadzonego wśród amerykańskich inwestorów, że jeśli inwestorzy są pozytywnie nastawieni co do przyszłych stóp zwrotu, to są skłonni zainwestować znacznie więcej. Odkrycie to jest ważne dlatego, że pokazuje, iż inwestorzy podejmują decyzje również na podstawie swoich własnych oczekiwań, a nie tylko samych obiektywnych analiz statystycznych. Dodatkowo autorom badania udało się wykazać, że inwestorzy, którzy odnotowywali wysokie stopy zwrotu w przeszłości, przewidują średnio, że ich przyszłe zwroty będą wyższe aniżeli okazują się być naprawdę. Wynika z tego, że również inwestorzy ulegają nieracjonalności w postaci nad optymizmu.

Błędy poznawcze mają kluczowy wpływ na sukces inwestorów, ponieważ ci, którzy za bardzo ulegają nad optymizmowi, odnotowują niższe stopy zwrotu z inwestycji. Za przykład może posłużyć zjawisko zaobserwowane przez La Portę [1996], który odkrył, że bardziej optymistyczni inwestorzy, tj. ulegający nad optymizmowi, odnotowywali niższe stopy zwrotu z inwestycji aniżeli inwestorzy bardziej ostrożni i pesymistyczni. Zjawisko to wyjaśnili dokładniej Bordalo, Gennaioli, Ma i Shleifer [2020], którzy badali stopy zwrotu z dwóch różnych typów portfeli inwestycyjnych: 1) portfela inwestującego w spółki z niską długookresową stopą zwrotu oraz 2) portfela inwestującego w spółki z wysoką długookresową stopą zwrotu. Okazało się, że wyższe stopy zwrotu z inwestycji w dany portfel osiągnęli inwestorzy inwestujący w akcje mniej dochodowych spółek (średnio 15%), aniżeli ci inwestujący w spółki z wysoką stopą wzrostu (średnio 3%). Zdaniem autorów mechanizm stojący za omawianym zjawiskiem polega na tym, że inwestorzy obserwujący wysokie zyski spółek na plus przewartościliwali ceny ich akcji i w momencie, gdy rynek zweryfikował, że cena portfela jest zawyżona, to w konsekwencji cena akcji spadła, a inwestorzy osiągnęli niskie stopy zwrotu. Analogicznie, inwestorzy, którzy byli pesymistycznie

nastawieni do stóp zwrotu niskodochodowych spółek, po czasie spostrzegli się, że cena akcji jest zaniżona, co doprowadziło do jej wzrostu, a oni mogli sprzedać swój portfel z większym zyskiem.

### 2.3. Nieracjonalność oczekiwań a stabilność gospodarki

Badania pokazują, że nieracjonalne prognozy wśród podmiotów gospodarczych są jedną z podstawowych przyczyn narastania baniek na rynkach finansowych. Gdy ryzyko udzielanych kredytów utrzymuje się na niskim poziomie, banki prognozują, że nie zmieni się ono istotnie w przyszłości. Ze względu na to, że zgodnie z prognozami ryzyko kredytowe nie wzrasta, banki udzielają ich coraz więcej, napędzając tym samym bańkę kredytową [Greenwood, Hanson, 2013]. Dane pokazują jednak, że wraz z nadmierną ekspansją kredytową, ryzyko naturalnie wzrasta, a banki nie biorą tego pod uwagę właśnie z powodu nieracjonalnych prognoz [Baron, Xiong, 2017; Fahlenbrach, Prilmeier, Stulz, 2018].

Bańki na rynku kredytowym są istotnie oraz pozytywnie skorelowane z wystąpieniem recesji w gospodarce. Wraz z nadmiernym wzrostem ilości kredytu na rynku, gospodarka zaczyna się przegrzewać (*overheat*), a sam kredyt staje się coraz bardziej ryzykowny dla banków. Co więcej, ten negatywny proces jest istotnie skorelowany z wybuchem kryzysu finansowego lub recesji w przeciągu następnych dwóch lub trzech lat [López-Salido, Stein, Zakrajsek, 2017]. Jordà, Schularick i Taylor [2015] oraz Greenwood, Hanson, Shleifer i Sørensen [2020] stawiają przy tym tezę, że bańka na rynku kredytowym jest dobrym predyktorem nadchodzącego kryzysu. Zdaniem autorów mechanizm wyjaśniający tę zależność polega na tym, że w pewnym momencie ryzyko kredytowe, które narastało wraz z rozwojem bańki, prowadzi do bankructw firm oraz niewypłacalności gospodarstw domowych, które to powodują panikę na rynkach finansowych, a w konsekwencji w całej gospodarce, co ostatecznie doprowadza do recesji.

Podsumowując, z przedstawionych powyżej badań wynika, że podmioty gospodarcze kształtują swoje oczekiwania w sposób nieracjonalny, podchodząc do prognoz często za bardzo optymistycznie lub pesymistycznie. Kluczowym pytaniem pozostaje: jaki mechanizm odpowiada za to, że ludzie postępują w ten nieracjonalny sposób? Tak jak pokazano w tym rozdziale, odpowiedź na to pytanie jest kluczowa nie tylko ze względu na działalność konkretnych firm czy inwestorów, ale przede wszystkim dla zrozumienia zjawisk, jakimi są kryzysy finansowe i gospodarcze.

### 3. Oczekiwania Diagnostyczne

#### 3.1. Heurystyka reprezentatywności

Występowanie błędów poznawczych w danych ekonomicznych tłumaczone jest przez teorię oczekiwań diagnostycznych (DE), która bazuje na heurystyce. Badania Daniela Kahnemana i Amosa Tverskiego (dalej KT) z lat 70. XX w. wykazały, że ludzie przy wydawaniu sądów systematycznie odchodzą od bayesowskiego rozkładu prawdopodobieństwa, co nazwali heurystykami [Tversky, Kahneman, 1974]. Heurystyki to uproszczony proces rozumowania i podejmowania decyzji, który działa na zasadzie tzw. reguły kciuka, i z którego na co dzień korzysta każdy z nas. Zjawisko to jest dla ludzi co do zasady korzystne, ponieważ pozwala nam na szybkie podejmowanie przybliżenie dobrych decyzji, zamiast dogłębnej analizy każdej sytuacji. KT wyróżniają trzy heurystyki, jednak w kontekście wyjaśnienia nieracjonalności oczekiwań kluczowa jest jedna – heurystyka reprezentatywności – którą jej odkrywcy określają jako ludzką tendencję do szacowania prawdopodobieństwa na podstawie podobieństwa (*similarity*). KT podkreślają, że heurystyka ta jest tym silniejsza, im bardziej reprezentatywna jest cecha, na której temat tworzymy osąd. Zgodnie ze stworzoną przez nich definicją „cecha jest reprezentatywna, jeśli jest silnie diagnostyczna<sup>8</sup>, tzn., jej relatywna częstość występowania w diagnozowanej grupie jest znacząco wyższa niż w grupie referencyjnej” [Tversky, Kahneman, 1983].

Dla zbudowania intuicji związanej z heurystyką reprezentatywności, która będzie istotna dla zrozumienia dalszej części przeglądu literatury, posłużymy się dwoma przykładami:

- *Miejsce Austrii w Europie*: Wśród respondentów Austria była postrzegana jako bardziej podobna do Szwecji niż do Węgier, w momencie gdy w pytaniu o to wspomniało się również o Polsce. Jednocześnie Austria staje się postrzegana jako bardziej podobna do Węgier niż do Szwecji, w momencie gdy w pytaniu przytoczy się Norwegię. Przedstawiona nieracjonalność decyzji spowodowana jest właśnie heurystyką reprezentatywności. Otóż Austria dzieli wspólne cechy ze Szwecją (kraj „Zachodu”<sup>9</sup>) oraz Węgrami (kraj położony w Europie Centralnej). Dlatego też w momencie, gdy Węgry zestawiono się z Polską (kraj ówczesnie komunistyczny), to na zasadzie skojarzenia respondenci przywiązali większą wagę do tego, że Austria wyznaje „zachodni” model wartości i gospodarki. Tak samo, gdy zestawimy Szwecję z Norwegią, respondenci na zasadzie skojarzenia

<sup>8</sup> Od stwierdzenia o diagnostyczności cechy pochodzi nazwa oczekiwań diagnostycznych.

<sup>9</sup> Badanie zostało przeprowadzone w latach 80. XX w., gdy Polska i Węgry były jeszcze krajami komunistycznymi.

przywiązują większą wagę do położenia geograficznego, które jest bliższe Węgrom [Tversky, 1977].

- *Diagnoza lekarska*: Zapytano lekarzy o określenie stanu zdrowia pacjenta na podstawie pozytywnego wyniku testu na bardzo rzadką chorobę. Intuicyjne, jeśli choroba jest rzadka, to wzrasta prawdopodobieństwo uzyskania niewiążącego wyniku testu (*false positive*). Widać to we wzorze Bayesa:

$$P(\text{chory} | \text{pozytywny wynik}) = \frac{P(\text{pozytywny wynik} | \text{chory}) P(\text{chory})}{P(\text{pozytywny wynik})},$$

który pokazuje, że wraz ze spadkiem prawdopodobieństwa, że pacjent jest rzeczywiście chory (choroba jest rzadka), równie mocno spada prawdopodobieństwo tego, że test pokazał prawidłowy wynik. W badaniu okazało się jednak, że lekarze nadmiernie sugerowali się wynikiem testu (*overreacted*) i podawali znacząco większe prawdopodobieństwo zachorowania, aniżeli wynikałoby to z obiektywnej reguły Bayesa i danych statystycznych. Mechanizm takiej decyzji wynikał z tego, że lekarze skojarzyli, że chory pacjent to pacjent z pozytywnym wynikiem testu, co ponownie spowodowane zostało przez zastosowanie heurystyki reprezentatywności [Casscells, Schoenberger, Graboys, 1978].

### 3.2. Formalny model heurystyki reprezentatywności

Nicola Gennaioli i Andrei Shleifer (dalej GS) zaproponowali matematyczny model heurystyki reprezentatywności, który stanowi punkt wyjścia dla oczekiwań diagnostycznych. W modelu rozpatrujemy jednostkę podejmującą decyzję na temat cechy  $C$  w grupie  $G$ . Zakładamy, że prawdziwy rozkład tej cechy to  $f(C = c | G)$ . GS [2010] definiują reprezentatywność cechy  $C$  w grupie  $G$  jako:

$$\frac{f(C = c | G)}{f(C = c | -G)},$$

gdzie  $-G$  to stosowna grupa porównawcza. Zgodnie z definicją przedstawioną na wstępie, cecha jest tym bardziej reprezentatywna, im relatywnie częściej występuje w grupie  $G$  w porównaniu z grupą  $-G$ . Wynika to z faktu, że ludzki umysł stara się oszczędzać energię przy podejmowaniu decyzji i dlatego przypisuje wyższą wagę do najbardziej reprezentatywnych cech, ponieważ to te cechy najłatwiej jest nam w umyśle przywołać.

Dla zilustrowania tej koncepcji ponownie przytoczmy przykład *diagnozy lekarskiej*. W tym przypadku cecha  $C$  to stan zdrowia pacjenta (zdrowy lub chory),  $G$  to grupa pacjentów z pozytywnym wynikiem testu, a  $-G$  to cała populacja. Załóżmy,

że rozkład prawdopodobieństwa  $C$  wygląda zgodnie z informacjami przedstawionymi w tabeli 1.

**Tabela 1.** Przykładowy rozkład prawdopodobieństwa cechy zdrowie (w %)

	C = chory	C = zdrowy
G = pacjenci z pozytywnym wynikiem testu	60	40
-C = cała populacja	5	95

Źródło: opracowanie własne.

Na podstawie modelu GS okazuje się, że bardziej reprezentatywną cechą jest *chory*, i to z tego powodu badani lekarze znacznie częściej przypisywali negatywną diagnozę, aniżeli wynikałoby to z prawdziwego rozkładu tej cechy.

$$\frac{P(\text{chory} | G)}{P(\text{chory} | -G)} = \frac{60\%}{5\%} = 12 > \frac{P(\text{zdrowy} | G)}{P(\text{zdrowy} | -G)} = \frac{40\%}{95\%} \approx 0,4$$

Jak potwierdziły badania, model reprezentatywności GS jest w stanie wytłumaczyć takie błędy poznawcze, jak „zaniedbanie rozkładu bazowego” (*base rate neglect*) ale przede wszystkim pokazuje, jak ludzie konstruują nadreakcyjne diagnozy [Bordalo, Coffman, Gennaioli, Shleifer, 2016].

### 3.3. Wyprowadzenie oczekiwań diagnostycznych

Logika stojąca za modelem oczekiwań diagnostycznych jest taka sama, jak w przypadku modelu reprezentatywności, tylko że teraz pod uwagę bierzemy procesy gospodarcze (np. wzrost gospodarczy, wzrost wynagrodzeń), dla których cechą reprezentatywną jest wartość zmiennej z poprzedniego okresu. Posługując się modelem zaproponowanym przez Bordala, Gennaioli i Shleifera [2018] założmy, że podmiotem w naszym modelu jest prognosta ekonomiczny, który obserwuje czas w sposób dyskretny  $t = 0, 1, \dots$ . Stan gospodarki w momencie  $t$  reprezentuje zmienna losowa  $X_t$ , która zachowuje się zgodnie z procesem AR(1):  $X_t = \rho X_{t-1} + \epsilon_t$ , gdzie  $\epsilon_t \sim N(0, \sigma^2)$ , a  $\rho \in [0, 1]$ . W takiej sytuacji podmiot przy konstrukcji prognozy rozkładu cechy  $X_{t+1}$  bierze pod uwagę wartość tej cechy w okresie  $t$  (tj.  $X_t = \widehat{X}_t$ )<sup>10</sup>. Innymi słowy, tak jak w przypadku lekarza, który musi podjąć decyzję o diagnozie pacjenta na podstawie pozytywnego wyniku testu, tak prognosta ekonomiczny

<sup>10</sup>  $\widehat{X}_t$  oznacza obserwowaną realizację cechy w okresie  $t$ .

musi oszacować rozkład  $X_{t+1}$  na podstawie wartości  $X_t$ , która reprezentuje całą jego wiedzę o gospodarce.

Jeśli prognosta zachowuje się w pełni racjonalnie, to do skonstruowania prognozy skorzysta z prawdziwego bayesowskiego rozkładu prawdopodobieństwa danego jako  $f(X_{t+1}|X_t = \hat{X}_t)$ . Jednakże prognosta, który ulega heurystyce reprezentatywności, zacznie przeważać swoją prognozę o najbardziej reprezentatywne informacje. W tym przypadku reprezentatywną informacją jest to, co wydarzyło się w okresie poprzedzającym prognozę.

Dla przykładu założmy, że jesteśmy analitykiem w *startupie*, który ma za zadanie stworzyć prognozę przychodu na kolejny miesiąc, i aby to zrobić, korzysta z informacji o tym, jak przychód firmy kształtował się w przeszłości. Dodajmy do tego jednak fakt, że w ostatnim miesiącu firma, po wielu miesiącach ze względnie małymi przychodami, odnotowała nagle wzrost przychodów na poziomie 300% z miesiąca na miesiąc. Zgodnie z heurystyką reprezentatywności w tym przypadku analityk przypisze nieracjonalnie wysoką wagę do wydarzenia, które miało miejsce ostatnio, ponieważ pełni dla niego rolę „diagnozy”, tak jak wynik testu w przypadku lekarzy. W konsekwencji prognoza wysokiego przychodu jest bardziej reprezentatywna aniżeli prognoza stabilnego przychodu:

$$\frac{P(\text{wysoki przychód}_{t+1} | \text{wysoki przychód}_t)}{P(\text{wysoki przychód}_{t+1} | \rho \times \text{niski przychód}_{t-1})} >$$

$$> \frac{P(\text{niski przychód}_{t+1} | \text{wysoki przychód}_t)}{P(\text{wysoki przychód}_{t+1} | \rho \times \text{niski przychód}_{t-1})}$$

co w konsekwencji prowadzi do nieracjonalnego zawyżenia prognozy. Proces ten tłumaczy zjawisko tego, że wiele *startupów* po kilku okresach szybkiego wzrostu zaczyna postrzegać siebie jako potencjalnie nowe Google [Bordalo i in., 2021b].

Formalnie reprezentatywność  $X_{t+1}$  dana jest przez:

$$\frac{f(\hat{X}_{t+1}|X_t = \hat{X}_t)}{f(\hat{X}_{t+1}|X_t = \rho \hat{X}_{t-1})}$$

Na podstawie powyższej formuły, podmiot podejmujący decyzje ma z tyłu głowy wiedzę, że prawdziwy rozkład  $X_{t+1}$  to  $f(\hat{X}_{t+1}|X_t = \hat{X}_t)$ . Jednakże po zaobserwowaniu informacji  $\Omega_t = X_t$  (np. wzrost dochodów od 300%), automatycznie zaczyna przypisywać on większe szanse wystąpienia wydarzenia, którego prawdopodobieństwo najbardziej wzrosło na skutek informacji  $\Omega_t$ . Ograniczenia powstałe w wyniku stosowania heurystyki sprawiają jednocześnie, że podmiot przeszacowuje wpływ

informacji na swoją prognozę i w konsekwencji nieracjonalnie zawyża (zaniża) on prognozę, jeśli informacja była pozytywna (negatywna).

Sformalizowana funkcja rozkładu zmiennej prognozowanej przez podmiot ulegający heurystyce reprezentatywności prezentuje się następująco:

$$f_t^\theta(\hat{X}_{t+1}) = f(\hat{X}_{t+1} | X_t = \hat{X}_t) \left[ \frac{f(\hat{X}_{t+1} | X_t = \hat{X}_t)}{f(\hat{X}_{t+1} | X_t = \rho \hat{X}_{t-1})} \right]^\theta \frac{1}{Z},$$

gdzie  $Z$  to stała zapewniająca, że  $f_t^\theta(\hat{X}_{t+1})$  całkuje się do jedności, a  $\theta \in [0, +\infty]$  to parametr reprezentatywności określający jak bardzo podmiot ulega heurystyce. Gdy  $\theta = 0$ , podmiot tworzy prognozy w pełni racjonalnie, a gdy  $\theta > 0$ , to ulega heurystyce. Z badań empirycznych wynika, że wartość tego parametru jest zbliżona do  $\theta \approx 1$  [Bordalo i in. 2020].

W formie funkcji liniowej, która to pozwala na łatwe wdrożenie DE do modeli ekonomicznych, DE mają następującą postać:

$$\mathbb{E}_t^\theta(X_{t+1}) = \mathbb{E}_t(X_{t+1}) + \theta[\mathbb{E}_t(X_{t+1}) - \mathbb{E}_{t-1}(X_{t+1})].$$

Pokazuje ona, że oczekiwania diagnostyczne  $\mathbb{E}_t^\theta(X_{t+1})$  nadmiernie reagują (*over-react*) na informacje otrzymane w okresie pomiędzy  $t$  a  $t-1$  poprzez wyrażenie  $\theta[\mathbb{E}_t(X_{t+1}) - \mathbb{E}_{t-1}(X_{t+1})]$ . Mechanizm ten jednak nie implikuje, że podmioty tworzą prognozy na podstawie łączenia racjonalnych oczekiwań. Tak jak sugerują Bordalo i in. [2018], progności raczej przypisują większą wagę do informacji, która dotarła w ostatnim okresie, a która to nadmiernie wpływa na przewidywania prognozy, co jest spójne z badaniami Kahnemana [2011].

### 3.4. Własności oczekiwań diagnostycznych

Własności oczekiwań diagnostycznych wyznaczane są przez wielu autorów i w związku z tym zebrano je wszystkie w jednym miejscu. Część własności może być nieintuicyjna, w związku z czym dodano do nich krótkie wytłumaczenie.

*Własność 1.* DE charakteryzują się addytywnością pod warunkiem, że dotyczą przyszłych wartości [L'Huillier, Singh, Yoo, 2021].

$$\mathbb{E}_t^\theta(X_{t+r} + Y_{t+s}) = \mathbb{E}_t^\theta(X_{t+r}) + \mathbb{E}_t^\theta(Y_{t+s}), \quad r, s \geq 1$$

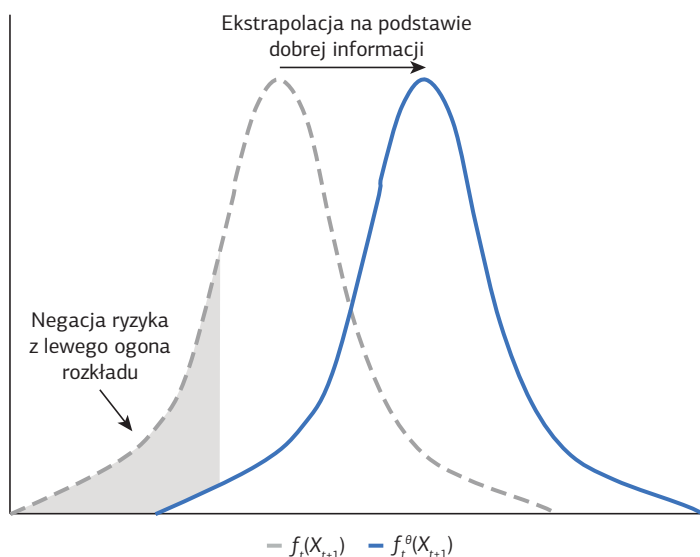
*Własność 2.* Podmiot korzystający z DE popełnia błąd poznawczy, jakim jest negacja ryzyka (*neglect of risk*) [Bordalo i in., 2018].

- *Dobra informacja:* Dla  $\mathbb{E}_t(X_{t+1}) > \mathbb{E}_{t-1}(X_{t+1})$  podmiot przypisuje większe prawdopodobieństwo wystąpieniu pozytywnego zjawiska w przyszłości (wyolbrzymia

prawy ogon rozkładu), przy jednoczesnym zlekceważeniu (*neglect*) prawdopodobieństwa wystąpienia negatywnego zjawiska (lewy ogon rozkładu).

- *Zła informacja*: Dla  $\mathbb{E}_t(X_{t+1}) < \mathbb{E}_{t-1}(X_{t+1})$  podmiot przypisuje większe prawdopodobieństwo wystąpieniu negatywnego zjawiska w przyszłości (wyolbrzymia lewy ogon rozkładu), przy jednoczesnym zlekceważeniu (*neglect*) prawdopodobieństwa wystąpienia pozytywnego zjawiska (prawy ogon rozkładu).

## Rysunek 2. Negacja ryzyka i ekstrapolacja



Źródło: Bordalo i in. [2018].

Wykres pokazuje subiektywny rozkład cechy w okresie  $t + 1$ . Linią ciągłą przedstawiono rozkład,  $f_t^\theta(X_{t+1})$  bazujący na DE po pojawieniu się pozytywnej informacji w okresie  $t$ ; linią przerywaną przedstawiono ten sam rozkład,  $f_t(X_{t+1})$ , tylko że bazujący na oczekiwaniach racjonalnych.

Ilustrując to, wróćmy do przykładu analityka w *startupie*. W momencie otrzymania pozytywnej informacji, że przychody wzrosły o 300%, analityk jest praktycznie przekonany, że nic złego się w kolejnym okresie nie stanie (nie dojdzie np. do drastycznego spadku przychodu), a jednocześnie zacznie przypisywać nieproporcjonalnie duże prawdopodobieństwo temu, że te ponadprzeciętne przychody będą się dalej utrzymywać. Reprezentację tego zjawiska w formie graficznej zobaczyć można na rysunku 2.

*Własność 3.* Siła, z jaką podmiot neguje ryzyko, zależy od zmienności prognozowanej zmiennej zgodnie ze wzorem:



$$\frac{\partial}{\partial X_{t+1}} \ln \left[ \frac{f(\hat{X}_{t+1} | X_t = \hat{X}_t)}{f(\hat{X}_{t+1} | X_t = \rho \hat{X}_{t-1})} \right] = \frac{\mathbb{E}_t(X_{t+1}) - \mathbb{E}_{t-1}(X_{t+1})}{\sigma_X^2}.$$

Powyższy wzór pokazuje, że im mniejsza jest wariancja zmiennej  $X$ , tym większa jest siła negacji ryzyka, co jednocześnie prowadzić może do nadmiernej pewności siebie, przykładowo wśród inwestorów [Bordalo i in., 2018].

*Własność 4.* Jeśli  $X_t$  zachowuje się zgodnie z procesem AR(1) o szokach  $\epsilon \sim N(0, \sigma^2)$ , to prognoza DE w okresie  $t$  co do wartości zmiennej  $X_{t+T}$  wyrażona jest wzorem:

$$\mathbb{E}_t^\theta(X_{t+T}) = \mathbb{E}_t(X_{t+T}) + \theta[\mathbb{E}_t(X_{t+T}) - \mathbb{E}_{t-1}(X_{t+T})].$$

*Własność 5.* Na podstawie DE można przewidzieć wartość rewizji prognoz, gdzie błędy te wynoszą średnio zero:

$$\mathbb{E}_{t-1}[\mathbb{E}_t^\theta(X_{t+T})] = \mathbb{E}_{t-1}(X_{t+T}).$$

Z powyższej własności wynika, że, uśredniając, prognozy DE są racjonalne, a działanie heurystyki reprezentatywności ma charakter jedynie krótkookresowy. Własność ta pokazuje również, dlaczego Bordalo i in. [2020] mogli przewidzieć wartość błędów prognozy.

*Własność 6.* Teoria DE jest odporna na krytykę *Lucasa* [1976], ponieważ podmiot w ramach DE stara się w rzeczywistości prognozować parametr persystencji  $\rho$ , który to nie jest ustalany w sposób mechaniczny, i który może się zmieniać w zależności od wprowadzonej polityki gospodarczej.

## 4. Wyniki uzyskane za pomocą oczekiwań diagnostycznych

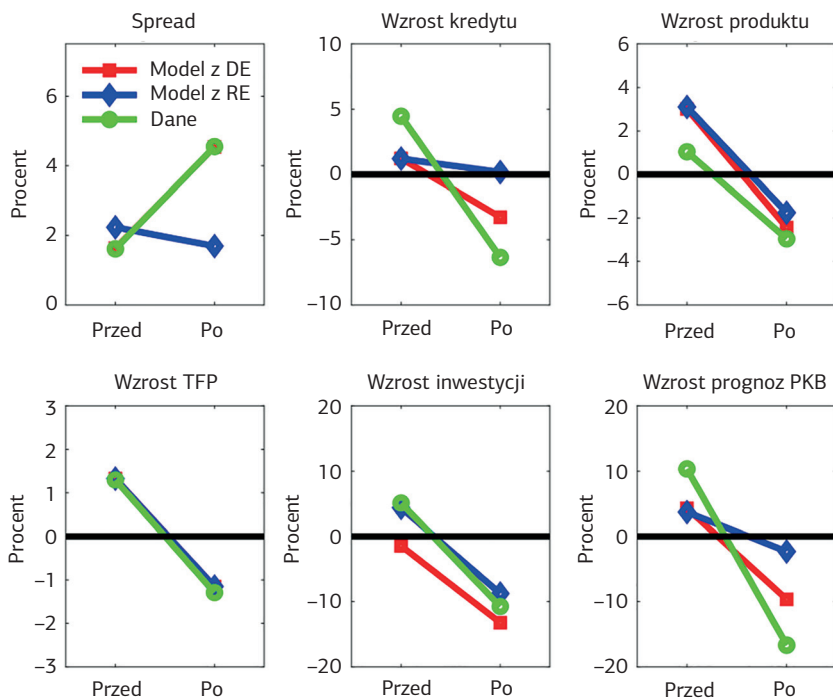
W ostatnim rozdziale przeglądu literatury podsumowano wyniki wdrożenia DE do modeli ekonomicznych typu *Real Business Cycle* (RBC) oraz innych modeli typu DSGE (*Dynamic Stochastic General Equilibrium*), jak również omówiono zdolność DE do wygenerowania cyklu *Kindlebergera* [Kindleberger, 1978] oraz przedstawiono jak DE może nam pomóc w przypadku analizy kryzysu COVID-19.

### 4.1. Modele *Real Business Cycle* (RBC)

Wprowadzenie DE w miejsce RE w modelach cyklu koniunkturalnego (RBC) pozwala na endogeniczne odtworzenie mechanizmu nadreakcji na szoki (*overreaction*). Bordalo i in. [2021a] wdrożyli DE do modelu RBC z szokami produktywności (TFP) oraz rynkiem kredytowym. Z rozwiązania modelu wynika, że po wystąpieniu dobrej

informacji w modelu (pozytywny szok TFP) diagnostyczne firmy stają się nadmiernie optymistyczne, dlatego też zaczynają wyżej wyceniać swój własny kapitał, więcej inwestować oraz zaciągać więcej kredytu, co napędza ekspansję kredytową w modelu. Analogiczne zjawisko występuje w przypadku pojawienia się złej informacji (negatywny szok TFP), która prowadzi do nadmiernego pesymizmu. Twórcy modelu porównali te wyniki z modelem RBC z RE, gdzie omawiana nadreakcja na szoki nie występowała. Dodatkowo autorzy wyestymowali oba modele i pokazali, że model z DE jest w stanie wyjaśnić znacznie więcej zmienności w danych. Oznacza to, że za pomocą DE można z poziomu samych firm w modelu wygenerować cykle kredytowe w skali całej gospodarki, co pokrywa się z danymi empirycznymi przytoczonymi w części pierwszej [López-Salido i in., 2017].

**Rysunek 3.** Kryzys Finansowy 2007–2008: Model DE vs. Model RE



Uwagi: Powyższy rysunek przedstawia jak zmieniły się wartości wybranych zmiennych makroekonomicznych w gospodarce Stanów Zjednoczonych pomiędzy okresem sprzed kryzysu, tj. przed upadkiem Lehman Brothers a okresem po kryzysie. Źródło: Bordalo i in. [2021a].

Autorzy badania sprawdzili również, czy DE są w stanie wytłumaczyć silną reakcję gospodarki na bankructwo banku Lehman Brothers (LB) w 2008 r., które to oficjalnie zapoczątkowało globalną recesję. Jak pokazuje rysunek 3, model z DE wyzna-

czył identyczną reakcją ryzyka kredytowego na upadek banku, jaką obserwujemy w danych empirycznych. Co więcej, porównując reakcje innych zmiennych na kryzys (bankructwo LB), okazuje się, że model z DE generuje wyniki bardziej zbliżone do danych empirycznych aniżeli model z RE. Pokazuje to, że zastosowanie DE wprowadza do modelu zawansowany mechanizm ludzkiego procesu wnioskowania, który ma istotne przełożenie na rezultaty uzyskiwane w ramach modelu, w momencie gdy tej cechy nie posiadają klasyczne modele z oczekiwaniami racjonalnymi. Podobne rezultaty do tych przedstawionych osiągnęli inni ekonomiści, którzy również wdrożyli DE do modeli typu RBC [L’Huillier i in., 2021; Bordalo i in., 2018].

## 4.2. Inne modele typu DSGE

Na wdrożeniu oczekiwań diagnostycznych zyskują również inne modele typu DSGE (*Dynamic Stochastic General Equilibrium*), dla których to wykazano, że wraz z wdrożeniem DE model zaczyna generować znacznie większe fluktuacje w reakcji na szoki stochastyczne aniżeli modele stosujące RE [L’Huillier i in., 2021]. Źródłem tego zjawiska, tak jak w przypadku modeli RBC, jest mechanizm nadreakcji na szoki, występujący jedynie w wariancie modelu z DE [L’Huillier i in., 2021]. Wpływ DE na występowanie większej zmienności, szczególnie zmiennych finansowych (kredyt, ceny mieszkań) potwierdzili także Favero, Melone i Tamoni [2021]. Co więcej, autorzy L’Huillier i in. [2021] wyestymowali zarówno model z DE, jak i RE, pokazując, że pierwszy z nich jest w stanie wytłumaczyć znacząco więcej zmienności w danych empirycznych.

## 4.3. Oczekiwania diagnostyczne a teoria baniek spekulacyjnych Kindlebergera [1978]

W innym badaniu Bordalo i in. [2021b] skonstruowali kompleksowy model symulujący proces tworzenia się baniek spekulacyjnych, który odwzorował mechanizm zgodny z teorią Kindlebergera [1978], który twierdził, że bańki narastają w sposób ewidentny/oczywisty (*self-evident*). Model ten zakłada, że podmioty w gospodarce podejmują swoje decyzje na podstawie zarówno oczekiwań diagnostycznych, jak i cen fundamentalnych<sup>11</sup>. Dodatkowo autorzy dopuszczają w modelu możliwość spekulacji, tj. zakupu aktywa w celu gry na wzrost jego ceny. Wraz z zastosowaniem kilku mechanizmów obserwowanych w trakcie kryzysów finansowych, udało się odtworzyć

<sup>11</sup> Zgodnie z mechanizmem cen fundamentalnych inwestorzy w modelu w momencie przeprowadzenia transakcji dowiadują się jaka jest cena fundamentalna danego aktywa na podstawie „zszumionych sygnałów” z rynku oraz poprzednich cen tego aktywa [Grossman, 1976; Grossman, Stiglitz, 1976]

bańkę spekulacyjną zgodną z teorią Kindlebergera [1978]. Zgodnie z tą teorią oraz modelami Bordalo i in. [2018] oraz Chodorowa-Reicha, Gurena, McQuade'a [2021] proces powstawania bańki spekulacyjnej składa się z trzech faz:

- 1) *Faza powstania*: W fazie pierwszej gospodarka doświadcza innowacji, która generuje serię dobrych informacji. Informacje te prowadzą następnie do wystąpienia nadmiernego optymizmu zgodnie z teorią DE. Optymizm ten dociera również do inwestorów, którzy zgłaszają zawyżony popyt na aktywa, co prowadzi do istotnego wzrostu ich cen.
- 2) *Faza szybkiego wzrostu*: Druga faza związana jest z napędzaniem wzrostu cen, a w konsekwencji wzrostem bańki spekulacyjnej. Inwestorzy wraz ze wzrostem cen stają się coraz bardziej optymistyczni co do ich przyszłej wartości, co prowadzi do dalszego wzrostu popytu oraz dalszego wzrostu cen. Przedstawiony mechanizm zamkniętego koła napędzany jest nie tylko przez nad optymizm wygenerowany przez DE, ale jest również stymulowany przez fakt uczenia się z cen (*learning from prices*) oraz możliwość spekulacji.
- 3) *Faza kryzysu*: W trzeciej fazie dochodzi do kryzysu na rynku, który spowodowany jest również przez DE. Kryzys nie rozpoczyna się jednak z powodu pojawienia się złej informacji, ale ponieważ dobra informacja, która napędzała nad optymizm inwestorów, zaczyna zanikać. Wraz z zanikaniem informacji inwestorzy stają się mniej podatni na heurystykę reprezentatywności, przez co lepiej wyceniają aktywa i zaczynają je wyprzedawać, co szybko redukuje cenę danego aktywa, która wraca do swoich wartości fundamentalnych, zwiększając tym samym ryzyko wystąpienia kryzysu.

#### 4.4. Oczekiwania diagnostyczne a kryzys COVID-19

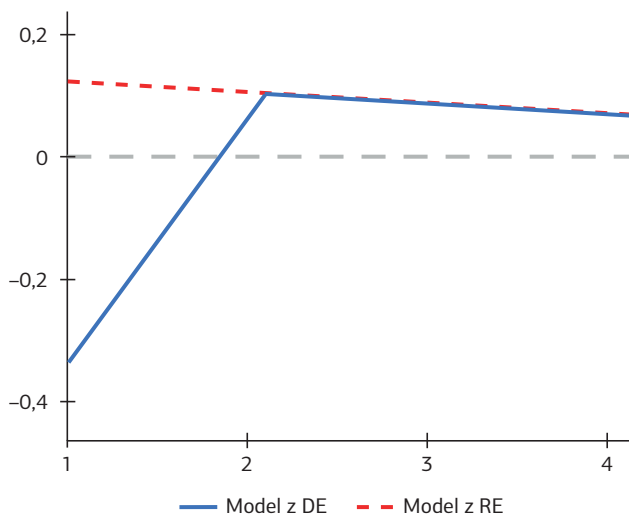
Ciekawym wątkiem związanym z DE jest to, że może pomóc nam wyjaśnić część mechanizmów zachowywania się gospodarki z czasu kryzysu COVID-19. Silny rozwój badań na temat oczekiwań diagnostycznych przypadł na okres światowej pandemii COVID-19, co skłoniło naukowców do zbadania również tego kryzysu. Badania L'Huilliera i in. [2021] prowadzą do dwóch istotnych wniosków:

- 1) W trakcie pandemii w większości państw zastosowano całkowity *lockdown*, który zamroził działalność gospodarczą i w konsekwencji znacząco obniżył łączną produktywność czynników produkcji (negatywny szok TFP). Doprowadziło to do silnego spadku konsumpcji i produkcji krajowej. Okazuje się, że w modelu DSGE z racjonalnymi oczekiwaniami, na scenariusz, który zakłada krótki negatywny szok produktywności (tak jak w rzeczywistości), gospodarka reaguje wręcz odwrotnie, tj. wzrostem konsumpcji oraz produktu krajowego. Tak jak

pokazuje rysunek 4, wynik zgodny z danymi empirycznymi uzyskujemy dopiero przy zastosowaniu oczekiwań diagnostycznych w modelu. Wynika to z faktu, że model z DE generuje mechanizm, w ramach którego podmioty stają się skrajnie pesymistyczne po negatywnym szoku produktywności, co sprawia, że ograniczają konsumpcję (można powiedzieć, że przygotowują się na „gorsze czasy”). W modelu z RE podobny mechanizm nie występuje, a spadek produktywności prowadzi do wzrostu zatrudnienia (w wyniku zwiększenia podaży pracy), a w konsekwencji do wzrostu PKB.

- 2) Oczekiwania diagnostyczne okazują się być również przydatne przy planowaniu polityki gospodarczej. Na przykładzie kryzysu COVID-19 autorzy pokazali, że w modelu DSGE z oczekiwaniami diagnostycznymi podmioty reagują silniej na pozytywne szoki fiskalne, aniżeli w modelach z oczekiwaniami racjonalnymi, ponieważ DE nie ograniczają wartości mnożnika polityki fiskalnej do poziomu 1, tak jak w przypadku modeli z RE. Ponownie wynik ten spowodowany jest mechanizmem nadreakcji, który sprawia, że podmioty bardziej reagują na szoki, a w tym przypadku pozytywny szok fiskalny. Z tego powodu DE wykazuje potencjał, aby wyjaśnić dlaczego, zgodnie z badaniami empirycznymi, mnożnik polityki fiskalnej dla wielu krajów pozostaje na poziomie ok. 1,5, pomimo iż klasyczne modele ekonomiczne zakładają, że jego wartość powinna być niższa o 1 [Nakamura, Steinsson, 2014].

**Rysunek 4.** Reakcja PKB na negatywny szok TFP



Źródło: L'Huillier i in. [2021].

## 5. Dyskusja i podsumowanie

Przykład zastosowania oczekiwań diagnostycznych do analizy kryzysu COVID-19 pozostawiono na koniec, aby zasygnalizować bardzo istotną kwestię. Otóż pomimo tego że teoria oczekiwań diagnostycznych powstała w celu wytłumaczenia zachowań inwestorów oraz rynków finansowych, to znajduje ona również istotne zastosowanie przy analizie niefinansowych zagadnień, a dodatkowo może być ona przydatnym narzędziem przy projektowaniu polityki gospodarczej. Warto jednak zauważyć, że teoria oczekiwań diagnostycznych pozostaje dalej teorią młodą i została dotychczas wykorzystana do analizy ograniczonej liczby zjawisk gospodarczych. W tym kontekście należy pamiętać, że DE mogą dostarczać wyników niezgodnych z danymi empirycznymi w obszarach wychodzących poza analizę zjawisk na rynkach finansowych, i w tym kontekście należy podchodzić do zagadnienia DE z ostrożnością, dopóki zagadnienie to nie zostanie lepiej poznane.

Celem powyższego przeglądu literatury było pokazanie, że: 1) dane empiryczne bezpośrednio wykazują, że podmioty gospodarcze podejmują decyzje na podstawie nieracjonalnych oczekiwań, charakteryzujących się nadreakcją na szoki informacyjne; 2) źródła tego zjawiska doszukiwać się możemy w psychologicznych podstawach tworzenia sądów przez ludzi, a w szczególności w heurystyce reprezentatywności; 3) teoria oczekiwań diagnostycznych jest osadzona głęboko w psychologii przy zachowaniu matematycznej elegancji i prostoty, oraz że 4) oczekiwania diagnostyczne są nam w stanie pomóc zrozumieć zarówno zachowanie inwestorów, jak i całej gospodarki.

Teoria oczekiwań diagnostycznych jest względnie młoda i ma potencjał do tego, aby wyjaśnić nam wiele nowych procesów występujących w gospodarce, a w konsekwencji pomóc nam lepiej zrozumieć, jak gospodarka reaguje na różnego rodzaju szoki zewnętrzne. Z drugiej jednak strony należy zaznaczyć, że DE nie jest w stanie w pełni odpowiedzieć nam na pytania dotyczące nieracjonalności oczekiwań. Wynika to z faktu, iż zakres oddziaływania DE odpowiada jedynie za ograniczony obszar ludzkiego procesu podejmowania decyzji. Przykładowo nawet po zastosowaniu DE zakładamy, że podmiot gospodarczy przy okazji każdej podejmowanej decyzji bierze pod uwagę wszystkie dostępne informacje na temat wszystkich nieskończenie wielu okresów. Założenie takie jest naturalnie niezgodne z odkryciami psychologii [Gabaix, 2020]. Dlatego też, zdaniem autora, teoria oczekiwań diagnostycznych powinna być traktowana jedynie jako ważna część całego zbioru teorii na temat nieracjonalności oczekiwań w ekonomii.

Z tego też powodu interesującym kierunkiem dalszych badań nad DE jest zastosowanie jej wraz z innymi teoriami nieracjonalnych oczekiwań. W tym kontekście

ciekawym byłaby implementacja zarówno teorii DE, jak i „behawioralnej krótko-wzroczności” zaproponowanej przez Gabaix [2020]. Na podstawie literatury wiemy, że obie te teorie pomagają nam zrozumieć zachowanie gospodarki i w adekwatny sposób opisują nieracjonalne zachowania podmiotów gospodarczych, dlatego też wspólne ich zastosowanie może przynieść jeszcze większe korzyści.

Do analizy oczekiwań diagnostycznych można podejść również od zupełnie innej strony. Otóż głównym wnioskiem z teorii DE jest fakt, że podmioty gospodarcze za bardzo skupiają się na ostatnich okresach przy konstruowaniu prognoz, co w konsekwencji prowadzi do wielu problemów w trakcie kryzysów gospodarczych. Dlatego też zasadnym jest zadać pytanie, czy istnieją sposoby, aby ograniczyć wpływ nieracjonalnych oczekiwań na gospodarkę. Jest to obszar badań, który intensywnie się rozwija, na co wskazują Kahneman, Sibony i Sunstein [2021], którzy sami przeprowadzają podobne analizy, proponując strategie ograniczania wpływu błędów poznawczych na podejmowane przez nas decyzje. Niemniej rekomendacje te odnoszą się przede wszystkim do skali mikroekonomicznej, a pytanie o możliwe rozwiązanie tego problemu w skali makroekonomicznej pozostaje dalej otwarte.

## Bibliografia

---

- Baron, M., Xiong, W. (2017). Credit Expansion and Neglected Crash Risk, *The Quarterly Journal of Economics*, 132(2), s. 713–764. DOI: 10.1093/qje/qjx004.
- Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A. (2016). Stereotypes, *The Quarterly Journal of Economics*, 131(4), s. 1753–1794. DOI: 10.1093/qje/qjw029.
- Bordalo, P., Gennaioli, N., Kwon, S.Y., Shleifer, A. (2021b). Diagnostic Bubbles, *Journal of Financial Economics*, 141(3), s. 1060–1077. DOI: 10.1016/j.jfineco.2020.06.019.
- Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A. (2020). Overreaction in Macroeconomic Expectations, *American Economic Review*, 110(9), s. 2748–2782. DOI: 10.1257/aer.20181219.
- Bordalo, P., Gennaioli, N., Porta, R.L., Shleifer, A. (2019). Diagnostic Expectations and Stock Returns, *The Journal of Finance*, 74(6), s. 2839–2874. DOI: 10.1111/jofi.12833.
- Bordalo, P., Gennaioli, N., Shleifer, A. (2018). Diagnostic Expectations and Credit Cycles, *The Journal of Finance*, 73(1), s. 199–227. DOI: 10.1111/jofi.12586.
- Bordalo, P., Gennaioli, N., Shleifer, A., Terry, S.J. (2021a). Real Credit Cycles. Working Paper 28416, National Bureau of Economic Research.
- Brzoza-Brzezina, M., Kolasa, M., Makarski, K. (2021). Monetary Policy and COVID-19. IMF Working Papers, 2021(274).
- Brzoza-Brzezina, M., Suda, J. (2021). Are DSGE Models Irreparably Flawed?, *Bank i Kredyt*, 52(3), s. 227–252.

- Casscells, W., Schoenberger, A., Graboys, T.B. (1978). Interpretation by Physicians of Clinical Laboratory Results, *New England Journal of Medicine*, 299(18), s. 999–1001. DOI: 10.1056/NEJM197811022991808.
- Chodorow-Reich, G., Guren, A.M., McQuade, T.J. (2021). The 2000s Housing Cycle with 2020 Hindsight: A Neo-Kindlebergerian View. Working Paper 29140, National Bureau of Economic Research.
- Eusepi, S., Preston, B. (2011). Expectations, Learning, and Business Cycle Fluctuations, *American Economic Review*, 101(6), s. 2844–2872. DOI: 10.1257/aer.101.6.2844.
- Fahlenbrach, R., Prilmeier, R., Stulz, R.M. (2018). Why Does Fast Loan Growth Predict Poor Performance for Banks?, *The Review of Financial Studies*, 31(3), s. 1014–1063.
- Favero, C.A., Melone, A., Tamoni, A. (2021). Monetary Policy and Bond Prices with Drifting Equilibrium Rates and Diagnostic Expectations. Working Paper. DOI: 10.2139/ssrn.3704241.
- Gabaix, X. (2020). A Behavioral New Keynesian Model, *American Economic Review*, 110(8), s. 2271–2327. DOI: 10.1257/aer.20162005.
- Gabaix, X., Koijen, R.S. (2021). In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis. Working Paper 28967, National Bureau of Economic Research.
- Gennaioli, N., Ma, Y., Shleifer, A. (2016). Expectations and Investment, *NBER Macroeconomics Annual*, 30(1), s. 379–431. DOI: 10.1086/685965.
- Gennaioli, N., Shleifer, A. (2010). What Comes to Mind, *The Quarterly Journal of Economics*, 125(4), s. 1399–1433. DOI: 10.1162/qjec.2010.125.4.1399.
- Gennaioli, N., Shleifer, A. (2018). *A Crisis of Beliefs: Investor Psychology and Financial Fragility*. Princeton University Press.
- Gennaioli, N., Shleifer, A., Vishny, R. (2015). Neglected Risks: The Psychology of Financial Crises. *American Economic Review*, 105(5), s. 310–314. DOI: 10.1257/aer.p20151091.
- Greenwood, R., Hanson, S.G. (2013). Issuer Quality and Corporate Bond Returns, *The Review of Financial Studies*, 26(6), s. 1483–1525. DOI: 10.1093/rfs/hht016.
- Greenwood, R., Hanson, S.G., Shleifer, A., Sørensen, J.A. (2020). Predictable Financial Crises. Working Paper 27396, National Bureau of Economic Research. DOI: 10.1111/jofi.13105.
- Greenwood, R., Shleifer, A. (2014). Expectations of Returns and Expected Returns, *The Review of Financial Studies*, 27(3), s. 714–746. DOI: 10.1093/rfs/hht082.
- Grossman, S. (1976). On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information, *The Journal of Finance*, 31(2), s. 573–585. DOI: 10.2307/2326627.
- Grossman, S.J., Stiglitz, J.E. (1976). Information and Competitive Price Systems, *The American Economic Review*, 66(2), s. 246–253.
- Iacoviello, M. (2005). House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle, *American Economic Review*, 95(3), s. 739–764. DOI: 10.1257/0002828054201477.
- Jordà, Ò., Schularick, M., Taylor, A.M. (2015). Leveraged Bubbles, *Journal of Monetary Economics*, 76, s. 1–20. DOI: 10.1016/j.jmoneco.2015.08.005.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.
- Kahneman, D., Sibony, O., Sunstein, C.R. (2021). *Noise: A Flaw in Human Judgment*. Little, Brown.
- Kahneman, D., Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness, *Cognitive Psychology*, 3(3), s. 430–454. DOI: 10.1016/0010-0285(72)90016-3.



- Kindleberger, C.P. (1978). Manias, Panics, and Rationality, *Eastern Economic Journal*, 4(2), s. 103–112.
- L’Huillier, J.P., Singh, S.R., Yoo, D. (2021). Incorporating Diagnostic Expectations into the New Keynesian Framework. Working Papers 339, University of California, Davis, Department of Economics.
- La Porta, R. (1996). Expectations and the Cross-Section of Stock Returns, *The Journal of Finance*, 51(5), s. 1715–1742. DOI: 10.2307/2329535.
- López-Salido, D., Stein, J.C., Zakrajšek, E. (2017). Credit-Market Sentiment and the Business Cycle, *The Quarterly Journal of Economics*, 132(3), s. 1373–1426. DOI: 10.1093/qje/qjx014.
- Lucas Jr., R.E. (1972). Expectations and the Neutrality of Money, *Journal of Economic Theory*, 4(2), s. 103–124. DOI: 10.1016/0022-0531(72)90142-1.
- Lucas Jr., R.E. (1976). Econometric Policy Evaluation: A Critique. W: *The Phillips Curve and Labor Markets: Carnegie-Rochester Conference Series on Public Policy*, K. Brunner, A.H. Meltzer (Eds.). American Elsevier Pub.Co. DOI: 10.1016/S0167-2231(76)80003-6.
- Mukhin, D. (2022). An Equilibrium Model of the International Price System, *American Economic Review*, 112(2), s. 650–688. DOI: 10.1257/aer.20181550.
- Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements, *Econometrica: Journal of the Econometric Society*, s. 315–335. DOI: 10.2307/1909635.
- Nakamura, E., Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions, *American Economic Review*, 104(3), s. 753–792. DOI: 10.1257/aer.104.3.753.
- Page, M.J., Moher, D., Bossuyt, P.M., Boutron, I., Hoffmann, T.C. i in. (2021). PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews, *BMJ*, 372. DOI: 10.1136/bmj.n160.
- Shiller, R.J. (2015). *Irrational Exuberance*. Princeton University Press.
- Tversky, A. (1977). Features of Similarity. *Psychological Review*, 84(4), s. 327. DOI: 10.1142/9789813231573\_0004.
- Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, 185(4157), s. 1124–1131. DOI: 10.1142/9789814417358\_0015.
- Tversky, A., Kahneman, D. (1983). Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment, *Psychological Review*, 90(4), s. 293.

## Diagnostic Expectations in the Economic Literature

### Summary

The financial crisis of 2007–2008 motivated economists to create the theory of diagnostic expectations (DE) which can explain non-rational economic forecasts made by economic agents reflected in the neglect of risk as well as overreaction to shocks. The aim of the literature review is to summarise the assumptions of the DE theory and to discuss the results it provides. This article summarises information from 12 selected

articles and shows that (I) according to empirical data the expectations of economic agents are non-rational, (II) the source of this lack of rationality can be found in the representativeness heuristic, (III) the theory of diagnostic expectations is deeply rooted in psychology with mathematical elegance and simplicity maintained, and that (IV) diagnostic expectations can help us understand the behaviour of each economic agent and the economy as a whole.

**Keywords:** diagnostic expectations, non-rational expectations, overreaction, neglect of risk, representativeness heuristic, financial crisis

# Strategiczna polaryzacja polityczna – perspektywa informacji, niepewności i wartościowości

Anna Halewska

## Streszczenie

Tekst stanowi systematyczny przegląd literatury mający na celu analizę zjawiska strategicznej polaryzacji politycznej. W analizie skoncentrowano się na syntezie prób teoretycznego modelowania strategicznej polaryzacji politycznej z wykorzystaniem osiągnięć teorii gier. Jak pokazuje praktyka, popularny model downsowski nie jest w stanie, w swojej klasycznej formie, wyjaśnić ułożenia wielu scen politycznych. Tekst podsumowuje próby odpowiedzi naukowców na ten problem. Wśród nich szczególną rolę odgrywają: wartościowanie kandydatów na podstawie ich cech charakterystycznych, trudności w przepływie informacji między partiami politycznymi a wyborcami oraz niepewność głosujących co do przyszłych zachowań polityków. W przeglądzie literatury zastosowano metodę PRISMA. Po przeanalizowaniu 128 publikacji naukowych przeprowadzono syntezę dokonań z 14 artykułów naukowych o największej kontrybucji do badanego zagadnienia.

**Słowa kluczowe:** polaryzacja polityczna, informacja, niepewność, wartościowość

## 1. Wstęp

Nie bez powodu wybory nazywane są świętem demokracji – to dni, w których każdy obywatel może dać wyraz swoim politycznym przekonaniom przez udzielenie poparcia kandydatom, których uważa za najlepszych. Jednak nierzadko okazuje się, że wyniki wyborów wydają się nie odzwierciedlać postaw najgęściej występujących w społeczeństwie. Jak pokazują badania, politycy często prezentują pozycje bardziej skrajne niż większość z ich wyborców [Fiorina, 2017]. Czy można zatem

mówić o rzeczywistej reprezentacji w organach legislacyjnych i wykonawczych? Co doprowadziło do takiej sytuacji?

Zdaniem badaczy działających w latach 50. ubiegłego wieku, w sytuacji jednowymiarowej pod względem polityki, gdzie istnieje nieparzysta liczba głosujących i posiadają oni preferencje jednoszczytowe, pozycja medianowa zapewnia zwycięstwo Condorceta<sup>1</sup> [Jakubowski, 2005, s. 54]. Przekonanie to było motywowane faktem, iż zbliżając się do centrum, kandydaci są w stanie pozyskać większą liczbę głosów, a tym samym nieoptymalne byłoby dla nich uplasowanie swoich stanowisk na krańcach rozkładu. Ostatecznie obie strony powinny znaleźć się w identycznym, centrystycznym położeniu, na środku rozkładu preferencji osób głosujących [Black, 1958; Downs, 1957]. Tak powstał przełomowy model Blacka-Downsa, który, przy pewnych założeniach, pozwolił wnioskować o nieuniknionej konwergencji kandydatów. Jak jednak zauważyli krytycy, przewidywania te nie odpowiadają rzeczywistości [Green, Shapiro, 1994].

W praktyce głosujący najczęściej mogą opowiadać się po stronie jednej z wielu różnorodnych platform politycznych, w większości nieprzypominających tego, co w społeczeństwie uznaje się za podejście umiarkowane. Pozostając wciąż w środowisku wymiarowym, nakreślić można przestrzeń dla konkurencji wyborczej, na której każdy członek populacji zaznacza swój punkt idealny, który jest wypadkową jego poglądów politycznych. Punkty te wyznaczają również preferencje subpopulacji opowiadających się za każdą ze stron. To, co w takim środowisku modelowym interesujące, to sytuacja, gdzie platformy kandydatów, w sytuacji rywalizacji partyjnej czy też między poszczególnymi politykami, są bardziej oddalone od centrum rozkładu preferencji wyborców, wyznaczone przez ich medianę. Sytuacja, w której rozwarstwienie między politykami jest większe niż rozwarstwienie między ich wyborcami, jest o tyle interesująca z punktu widzenia politologii i teorii gier, co ryzykowna z perspektywy szerokiego horyzontu czasowego: jak demonstruje literatura, pozycje wyborców mogą adaptować pozycje kandydatów i stawać się coraz bardziej do nich podobne [Kollman, Miller, Page, 1992], co w efekcie prowadzić może do bardziej spolaryzowanego społeczeństwa.

Nie chcąc dopuścić do postępującej polaryzacji, można podejmować próby projektowania systemów głosowania w ten sposób, aby uniknąć nadmiernej polaryzacji politycznej (*excessive polarization*), wykraczającej poza polaryzację poglądów społeczeństwa. Aby tego dokonać, należy wpięrow poznać okoliczności, które mogą doprowadzić do takiego ustawienia aktorów sceny politycznej, w ujęciu czysto teo-

<sup>1</sup> Zwycięzca Condorceta to kandydat, który w zestawieniu z każdą z osób startujących w wyścigu wyborczym z osobna, jest preferowany przez większość głosujących [McLean, McMillan, Monroe, 2013].

retycznym. Celem artykułu jest analiza tekstów naukowych opisujących modele teoretyczne dotyczące czynników wpływających na strategiczną polaryzację polityczną. Uwzględnione zostaną głównie wyniki tych prac, jako że wprowadzenie do niniejszego artykułu równań i zmiennych, jakim przyglądali się badacze, wydłużyłoby go znacząco, nie zwiększając jednocześnie jego wartości. Artykuł ma być pomocny dla tych, którzy chcą móc rozpoznawać czynniki, które wpływają na strategiczną polaryzację polityczną. Niniejsza publikacja poświęcona jest trzem z tych bodźców: informacji, niepewności i cechom kandydatów. Aby zrealizować cel analizy, postawiono następujące pytania badawcze:

- w jaki sposób przekazywanie informacji między członkami społeczeństwa wpływa na strategiczną polaryzację polityczną?
- jaką rolę w strategicznej polaryzacji sceny politycznej odgrywa niepewność?
- jakie cechy kandydatów przyczyniają się do ich polaryzacji?

Artykuł składa się z czterech części. W pierwszej przedstawiono metodę badawczą wykorzystaną w analizie. W drugim rozdziale zaprezentowano wyniki przeprowadzonej analizy. Trzeci rozdział to ich dyskusja, omówienie każdego z uwzględnionych artykułów i dokonanie syntezy dotychczasowych badań naukowych. Praca zakończona jest podsumowaniem oraz prezentacją kierunków dalszych badań.

## 2. Metoda badawcza

Wśród wielu metod przeprowadzania systematycznego przeglądu literatury, jedną z najpopularniejszych jest PRISMA – *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*. Metoda ta zawiera trzy etapy, które w swojej końcowej fazie mają pozwolić na wyróżnienie tekstów najlepiej odpowiadających na postawiony problem badawczy. Najważniejszym narzędziem pierwszej części tego procesu – identyfikacji – jest meta baza danych, w przypadku której wpisanie odpowiednich słów kluczowych prowadzi do pierwszego zestawienia tekstów, jakim należy się przyjrzeć. Bazą użytą do przeprowadzenia tego badania jest Web of Science – wybrana ze względu na dużą liczbę rekordów (171 milionów) [Clarivate, 2022] i szeroki zakres uwzględnianych dyscyplin (256). Po fazie identyfikacji następuje ocena kwalifikowalności, gdzie autor przeglądu literatury dokonuje krytyki abstraktów, a następnie również głównej treści artykułów. Na tym etapie następują odrzucenia ze względu na niedopasowanie do tematu, jak również stosowaną metodę badawczą. Proces kończy się decyzją, która pozwala na określenie które publikacje znajdują się w przeglądzie literatury.

Przedstawiona metoda posiada naturalnie swoje ograniczenia. Wśród takowych wymienić można możliwość użycia również innych baz tekstów naukowych lub mniej specyficznej frazy: pozwoliłoby to na pozyskanie większej liczby artykułów, natomiast podaje się w wątpliwość czy podniosłoby to wartość analizy ze względu na powtarzalność schematów i pomysłów już na poziomie tekstów znalezionych za pomocą wybranej metody.

### 3. Wyniki przeglądu literatury

W celu odpowiedzi na pytania badawcze użyto słów kluczowych w języku angielskim (zarówno amerykańskim, jak i brytyjskim), uznając jego nadrzędną rolę we współczesnej nauce. Artykuł, aby znalazł się w grupie prac rozważanych do uwzględnienia w przeglądzie, musiał spełnić cztery wymienione poniżej kryteria. Pierwsze dotyczyło tematyki, co zostało ocenione przez kryterium zawarcia w tytule, abstrakcie i słowach kluczowych jednego ze sformułowań z kolumny pierwszej w tabeli 1. Ponadto, aby uzyskać poszukiwany format rywalizacyjny między kandydatami na tle wyborczym, za ważną uznano obecność w wyżej wymienionych obszarach publikacyjnych jednego z wyrazów z kolumny drugiej. Ostatnimi elementami kwalifikującymi tekst były odpowiednia data publikacji (nie wcześniej niż lata 90. XX w.) oraz metoda – szeroko pojęty model (z racji, że użycie frazy „model teoretyczny” spowodowałoby ominięcie zbyt wielu wyników, jako że nie jest to termin powszechnie stosowany w literaturze).

**Tabela 1.** Struktura frazy użytej do pozyskania listy tekstów do fazy identyfikacji

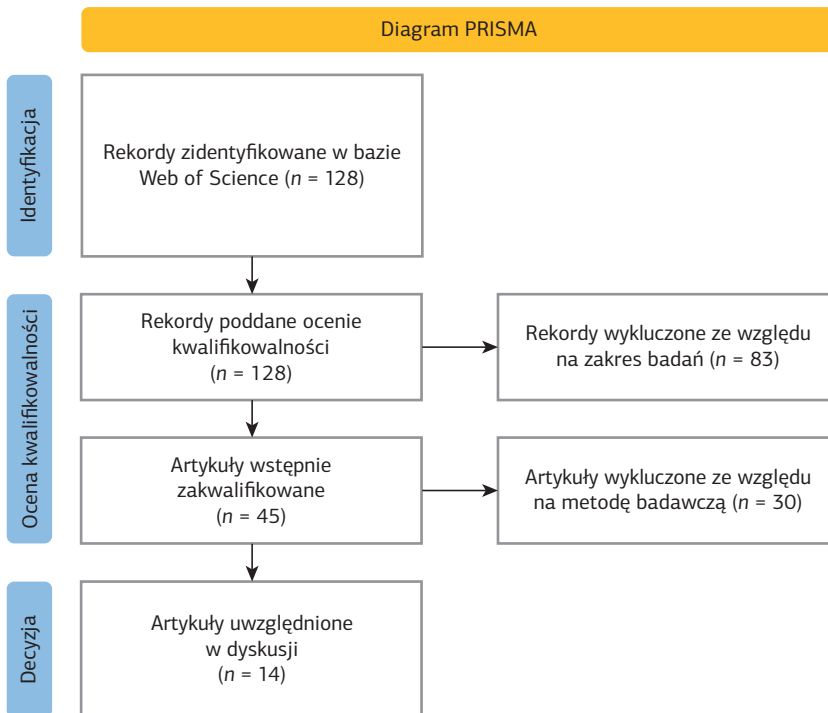
Temat przewodni	Kontekst	Metoda	Okres publikacji
Political polarization (political polarisation)	voter	model	od 1.01.1990 do 1.01.2022
Party polarization (party polarisation)	voters		
Policy polarization (policy polarisation)	election		
Political divergence	elections		
Party divergence			

Źródło: opracowanie własne.

Do fazy identyfikacji, na podstawie powyżej nakreślonych wyników wyszukiwania, zakwalifikowano 128 pozycji. Tym samym pierwszy etap zakończono z wynikiem 128 artykułów i przystąpiono do wstępnej analizy i oceny kwalifikowalności, polegającej na przyjrzeniu się abstraktom, tytułom i słowom kluczowym każdego z tek-

stów. W tej fazie zestawienia najważniejszym kryterium odrzucenia była tematyka niebędąca przedmiotem tego badania, a w szczególności próba uchwycenia efektów polaryzacji politycznej zamiast jej przyczyn, jak również próby odpowiedzi na pytanie, czy polaryzacja zachodzi, a jeśli nie – dlaczego. Wśród niezakwalifikowanych tekstów znalazły się również artykuły badające polaryzację wyłącznie pośród społeczeństw, nie wyróżniając warstwy politycznej, jak i analizujące polaryzację w krajach niedemokratycznych, a tym samym nieprzeprowadzających wyborów, których spełnianie kryterium wiarygodności uznano za niewątpliwe. Do odrzucenia dochodziło także w przypadkach, gdy artykuł był sam w sobie przeglądem literatury, a nie badaniem oryginalnym. Odrzucano także artykuły ze względu na metodę – głównie dotyczyło to artykułów opierających się na empirii, opisujących modele ekonometryczne regresji, gdzie polaryzacja występowała jako zmienna wyjaśniana bądź wyjaśniająca. Choć niniejszy przegląd literatury jest motywowany i napędzany zjawiskami empirycznymi, warto podkreślić, że poczynienie takich kroków było kluczowe ze względu na cel badawczy, który dotyczy wyłącznie teoretycznych modeli.

**Rysunek 1.** Struktura frazy użytej do pozyskania listy tekstów do fazy identyfikacji



Źródło: opracowanie własne przygotowane za pomocą narzędzi PRISMA.

Podsumowując, jak zaprezentowano na rysunku 1, do pierwszej fazy analizy, zwanej fazą identyfikacyjną, zaliczono 128 rekordów, z czego żadne nie były duplikatami. Następnie przystąpiono do wstępnej kwalifikacji, w wyniku której wyeliminowano 83 artykuły z powodu nieodpowiedniego zakresu badań, co przedstawiono już w poprzednim rozdziale dotyczącym metodologii. Wśród 45 artykułów wstępnie zakwalifikowanych, dwie trzecie wykluczono ze względu na zastosowaną metodę badawczą – powodem była najczęściej praca z danymi, a nie budowanie modelu teoretycznego, co stanowi cel niniejszego przeglądu literatury. Ostatecznie wśród artykułów uwzględnionych w dyskusji znalazło się 14 pozycji.

## 4. Dyskusja wyników badań

Klasyczna teoria medianowego wyborcy głosi, że w jednowymiarowym środowisku politycznym, gdzie istnieje pełna informacja o preferencjach wyborców i ustawieniach kandydatów, a wybory wygrywane są większością głosów, stały zestaw polityków motywowanych jedynie chęcią wygrania wyborów znajdzie się w równowadze jedynie będąc w centrum preferencji głosujących. Zatem wnioski o konwergencji kandydatów do środka politycznego spektrum można wyciągać tylko pod pewnymi założeniami. W odpowiedzi na zupełnie inne ustawienia spotykane na rzeczywistych scenach politycznych, badacze podejmują się analizy założeń przyjmowanych przez model klasyczny, czego dowodzi poniższa dyskusja wyników przeglądu literatury.

Esponda i Pouzo [2019], w celu wyjaśnienia polaryzacji partyjnej, modelują sposób przetwarzania informacji przez wyborców, zwracając uwagę na popartą badaniami tezę politologii o niepełnej racjonalności jednostek [Convers, 2000], które w efekcie głosując, nagradzają lub karzą polityków i ich partie na podstawie ich przeszłych decyzji. Jedyną zmianą, jaką wprowadzają do klasycznego modelu downsowskiego, jest posiadanie przez wyborców prywatnej informacji o stanie świata – taką informacją może być np. wiadomość o recesji czy ożywieniu gospodarczym. Po jej otrzymaniu decydują na kogo oddać swój głos: partię, która wprowadzi w życie politykę lewicową lub prawicową, czy też partię prowadzącą politykę neutralną. Ponieważ oceniają polityków na podstawie faktycznych działań, a nie kontrfaktycznych dywagacji, partie mają motywację do zajmowania stanowisk skrajnych, które pozwalają na bardziej widoczne działania. Zatem prywatny sygnał informacyjny połączony z niepełną racjonalnością wyborców przyczynia się do zmiany w ustawieniu kandydatów na bardziej spolaryzowane pozycje.

Jak zauważa Tajika [2021], inaczej jest z sygnałem pochodzącym od dobrze poinformowanych polityków. Partie w tym modelu wybierają działania, które później



są skonfrontowane ze stanem rzeczywistym świata. Wyborcy są niedoskonale poinformowani, ponieważ otrzymują sygnał o stanie rzeczywistym świata, który tylko z pewnym określonym zewnętrznym prawdopodobieństwem jest poprawny. W tym modelu po stronie wyborców niepełna informacja idzie w parze z preferencją do ryzyka, której występowanie w kontekście konfliktów ideologicznych jest udokumentowane empirycznie [Kendall, Nannicini, Trebbi, 2015]. Głosujący podejmuje decyzje na podstawie nie tylko pozycji kandydatów, ale też otrzymanego sygnału i jego jakości. Tajika dowodzi, że poniżej pewnego poziomu dostarczanych informacji spolaryzowane działania wygrywają ze *status quo*, preferowanym przez medianowego wyborcę, a dobrobyt społeczny jest zminimalizowany. Model jest następnie rozszerzany po wprowadzeniu informacji o stanie świata, także po stronie kandydatów. Politycy wybierający swoje stanowiska po otrzymaniu sygnału powiększają pośrednio stan wiedzy wyborców. Wysoka precyzyjność tych sygnałów może nawet doprowadzić do zaniknięcia równowagi z polaryzacją i wygranej *status quo*. To sugeruje, że lepiej poinformowani politycy zwiększają dobrobyt społeczny, nawet jeśli nie taki był ich oryginalny cel.

W przeciwieństwie do Tajiki, Zakharov [2009] uznaje wyborców za agentów nieskłonnych do ryzyka. W jego ujęciu politycy konkurują, by zdobyć głosy niezdecydowanych za pomocą swoich platform i cech charakterystycznych. Te ostatnie to często powtarzający się wątek w ekonomii politycznej i politologii, określane terminem wartościowania (*valence*). Składają się na nie cechy charakterystyczne kandydatów, niezależne od ideologii czy partii do której należą. Takimi cechami mogą być efektywność, doświadczenie czy popularność. Tę ostatnią do modelu wprowadza Adams [1999], dowodząc nie tylko, że kreuje ona równowagi inne niż te w punkcie medianowego wyborcy, ale też, że kandydaci, którzy tracą na popularności, najczęściej zajmują ekstremalne stanowiska polityczne.

Schofield [2007] bada wpływ wartościowości kandydatów na polaryzację, odchodząc także od innego założenia modelu klasycznego – jednowymiarowości politycznej. W rzeczywistości bowiem sprowadzenie wszystkich stanowisk partyjnych do jednego ustawienia na osi, która ma określać polityczne położenie partii, jest w większości niewykonalne – politycy mogą być konserwatywni w pewnych kwestiach a liberalni w innych. Schofield modeluje sytuację, w której występują dwa wymiary polityczne, nieredukowalne do siebie nawzajem. Choć i wtedy partie wysoce wartościowalne umiejscawiają się blisko pozycji medianowego wyborcy, to aby wygrać wybory, muszą wchodzić w koalicję z partiami o niższym wartościowaniu, których stanowiska znajdują się na ogonach preferencji głosujących.

W modelu Zakharova [2009] wartościowość może być wykreowana za pomocą środków finansowych. Ze względu na wspomnianą już niską skłonność do ryzyka,

wyborcy są bardziej wrażliwi na zmianę w jakości kandydatów, jeśli zajmują oni mniej spolaryzowane pozycje. Zatem dla polityka o platformie bardziej radykalnej, inwestycja w jakość jest bardziej wartościowa – zwiększenie swojej wartości w oczach wyborców, co wymaga zaangażowania finansowego, przekłada się na większy przychód krańcowy w postaci większej liczby głosów. Kandydaci w modelu Zakharova polaryzują się, aby zmniejszyć koszty osiągnięcia pewnego poziomu jakościowego wizerunku w oczach głosujących. Polaryzują się ponadto tym bardziej, im więcej jest wyborców niezdecydowanych, nieprzywiązanych do żadnej z partii.

Inwestowanie w celu zwiększenia prawdopodobieństwa swojej wygranej występuje także w modelu Galeottiego i Mattozziego [2011], gdzie społeczeństwo dowiaduje się o stanowiskach kandydatów na dwa sposoby: przez losowo adresowane reklamy i przez komunikację międzyludzką. Jeśli reklamy są wystarczająco kosztowne, partie zajmują ekstremalne stanowiska i nie przekazują ze swojej strony informacji społeczeństwu. W przeciwnym wypadku politycy są mniej spolaryzowani, a wolumen przekazywanych przez nich informacji zależy od tego, jak bardzo skomunikowani są między sobą wyborcy – im bogatsze są ich sieci komunikacyjne, tym mniej informacji decydują się przekazać kandydaci. Według Galeottiego i Mattozziego takie ustawienie zwiększa prawdopodobieństwo posiadania przez głosujących fałszywych przekonań na temat sceny politycznej, a tym samym zwiększa szansę ekstremistów na wygraną. W bardziej wyrafinowanej wersji modelu reklama i komunikacja nie mają przypadkowych odbiorców – występuje hermetyczność środowiska, a partie decydują się kierować reklamy do niezdecydowanych wyborców, którzy są jednocześnie im ideologicznie najbliżsi. Analiza takiego środowiska dowodzi, że duża komunikacja między głosującymi o podobnych do siebie przekonaniach zwiększa prawdopodobieństwo wybrania kandydata ekstremalnego. W przeciwieństwie więc do wniosków wyciągniętych przez Tajikę [2021] na temat komunikacji między kandydatami i wyborcami, która doprowadza do zmniejszenia polaryzacji, według Galeottiego i Mattozziego zupełnie inny wpływ mają przepływy horyzontalne między samymi wyborcami.

Przepływy informacyjne między politykami a społeczeństwem modelują także Carrillo i Castanheira [2008]. Podkreślają również rolę wartościowania budowanego za pomocą środków finansowych. W tym modelu jednak wyborcy obserwują tę jakość tylko z pewnym prawdopodobieństwem, nie są doskonale poinformowani, a pozyskane informacje wpływają na ich wyborcze decyzje. Carrillo i Castanheira dowodzą, że jeśli to prawdopodobieństwo poinformowania jest bardzo niskie, element ten staje się nieistotny i model sprowadza się do klasycznego modelu downsowskiego. Jeśli jest natomiast bardzo wysokie, kandydaci konkurują między sobą w wymiarze jakościowym, ale równowaga wciąż wskazuje na konwersję. Do pola-

ryzacji dochodzi w przypadku pośrednim: ponieważ kandydaci mają jedynie częściową szansę na to, że ich wysiłek zostanie zaobserwowany, preferują ustawienie się na pozycjach oddalonych od centrum, sygnalizując tym samym głosującym, że będą politykami o wysokiej jakości. Wybory mogą wygrać dzięki zaufaniu wyborców w wysoką, oczekiwaną wartość prezentowanej jakości. Zaobserwować wtedy można sytuację symetrycznej lub asymetrycznej polaryzacji.

Rolę zaangażowania finansowego kandydatów podkreślają też Grosser i Palfrey [2014], powielając je z formą biernego uczestnictwa w wyborach przez agentów takich samych, jak wyborcy. Jak donoszą Grosser i Palfrey [2014], z tego prawa korzystają wyłącznie obywatele o przekonaniach powyżej pewnego poziomu radykalności, co prowadzi do polaryzacji sceny politycznej. Jest ona tym wyższa, im droższe z perspektywy obywatela wejście do polityki. Polaryzacja również rośnie wraz ze wzrostem awersji do ryzyka wśród społeczeństwa. Ostatecznie do zmniejszenia polaryzacji prowadzą większe korzyści ze sprawowania urzędu.

Dellis i Oak [2016] badają, jak zmienia się środowisko polityczne powstałe z kandydatów-obywateli w sytuacji odejścia od założenia ordynacji większościowej na rzecz głosowania zatwierdzającego (*approval voting*)<sup>2</sup>. Jeśli wyborcy mogą głosować na więcej niż jednego z egzogenicznego zestawu kandydatów, szanse na wygraną centrowego kandydata rosną, ponieważ dodatkowe głosy są oddawane na niego. Nie jest tak jednak w przypadku gdy liczba opcji może rosnąć wraz z liczbą możliwych do oddania głosów. W takiej sytuacji efekt, który sprzyjać miał politykowi o umiarkowanej platformie, zanika, ponieważ dodatkowe głosy są po prostu oddawane na dodatkowych kandydatów, których stanowiska nie muszą koniecznie znajdować się w niedalekiej odległości od centrum. Tak powstać może spolaryzowana scena polityczna w postaci kilku silnych środowisk politycznych. W kontraście do tego sposobu oddawania głosów stoi zasada głosu alternatywnego, którą autorzy Dellis, Gauthier-Belzile i Oak [2017] zbadali pod kątem jej potencjału zmniejszenia stopnia polaryzacji sceny politycznej<sup>3</sup>.

Inne rozróżnienie między kandydatami wprowadzają Berger, Munger i Potthoff [2000]. Modyfikują model klasyczny przez wprowadzenie niepewności wyborców co do pozycji dwóch kandydatów o odmiennej pozycji: osoby do tej pory zajmującej

<sup>2</sup> Głosowanie zatwierdzające (*approval voting*) to system wyborczy, w którym każdy wyborca może wybrać („zatwierdzić”) dowolną liczbę kandydatów, a zwycięzcą jest kandydat zatwierdzony przez największą liczbę wyborców [Brams, Fishburn, 1978].

<sup>3</sup> Nie jest tematem niniejszej pracy badanie metod zmniejszania spotkania polaryzacji politycznej. Uznano jednak za stosowne zaznaczyć, że zasada głosowania alternatywnego polega na przypisaniu przez wyborców kandydatom miejsc w rankingu. Wygrywa polityk z największą liczbą miejsc pierwszych. W razie remisu eliminowany jest kandydat z najmniejszą liczbą miejsc pierwszych i proces się powtarza [Dellis i in., 2017, s. 566].

urząd i kogoś, kto chce go przejąć. Jeśli ta niepewność jest większa w stosunku do nowego polityka, może on wygrać wybory, zajmując pozycję oddaloną od centrum – istnieje więc potencjał do strategicznej polaryzacji politycznej.

Podobne rozróżnienie wprowadza Redmond [2017], nadając jednak pretendentowi do urzędu przewagę zamiast niekorzyści, pod postacią możliwości późniejszego ogłoszenia swojej platformy, po zaobserwowaniu platformy przeciwnika, który obecnie zajmuje dane stanowisko. Dzięki temu nowy kandydat może przystąpić do politycznej rywalizacji z większą ilością informacji niż jego konkurent. Tak modelowani kandydaci, motywowani chęcią wprowadzenia danej polityki, a nie wygraną, oddalają się od medianowego wyborcy, a jest to tendencja tym silniejsza, im mniej wartościowości posiada w oczach wyborców kandydat z zewnątrz.

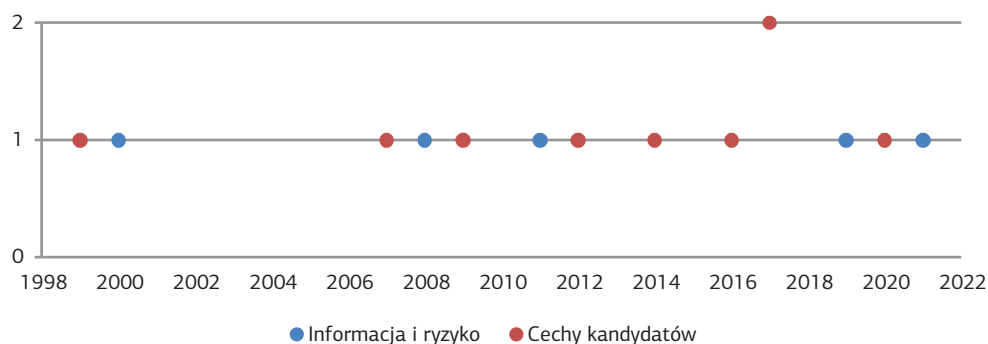
Karakas i Mitra [2020] podobnie decydują się przypisać kandydatom cechy na podstawie ich roli na scenie politycznej. Wyróżniają kandydata pochodzącego z elit politycznych i kandydata z zewnątrz, z czego ten drugi jest postrzegany jako bardziej skłonny do zrealizowania obietnicy dużej politycznej zmiany. W konsekwencji zajmuje bardziej ekstremalne pozycje, a ponadto jego szanse na wygranę wyborów mogą zależeć od stopnia nierówności społecznych – przykładowo w sytuacji wyborów dotyczących podwyższania podatków.

Rolę cech niezmiennych i wartościowości dla teorii medianowego wyborcy badali też Krasa i Polborn [2012]. Dowodzą oni, że można odejść od klasycznego założenia o homogeniczności kandydatów i nadal uzyskać równowagę z konwergencją. Inaczej jest, kiedy użyteczność wyborcy z ustawienia platformy danego kandydata i jego cech niezmiennych nie jest addytywnie separowalna – mając do czynienia z dwoma kandydatami deklarującymi politykę X, wyborca woli polityka pierwszego, ale jeśli ci sami kandydaci deklarują jednocześnie politykę Y, głosujący preferuje polityka drugiego, właśnie ze względu na cechy charakterystyczne. Jeśli takie preferencje występują w środowisku wyborczym, w równowadze politycy odchodzą od środka rozkładu w celu wygrania wyborów, a ich platformy są bardziej spolaryzowane.

Połączeniem koncepcji niepewności i cech niezmiennych polityków jest model Polborna i Snydera [2017], gdzie kandydaci walczą o wygraną w swoich dystryktach, ale są ograniczani działaniami partii, do której należą. Głosujący natomiast kierują się w swoich wyborach cechami indywidualnymi kandydata oraz pozycją partii, do której on należy – co odzwierciedlać ma większą wiedzę obywateli na temat stanowisk partyjnych w porównaniu do przekonań konkretnych polityków. Opierając się na statystykach porównawczych wykonanych na podstawie ich teorii, wnioskuje o dwóch potencjalnych powodach zwiększenia polaryzacji między partiami: zmniejszona niepewność co do wyników wyborów i zmniejszone znaczenie cech charakterystycznych kandydatów dla decyzji wyborców.

Wśród 14 przeanalizowanych tekstów aż 9 zostało napisanych przez przynajmniej jednego autora mającego afiliację na amerykańskim uniwersytecie. Tak wysoki odsetek może być wyjaśniony względnie wysokim stopniem polaryzacji społeczeństwa amerykańskiego w porównaniu do krajów takich jak Niemcy, Kanada czy Wielka Brytania [Boxell, Gentzkow, Shapiro, 2020]. Publikujący na temat strategicznej polaryzacji politycznej w zakresie informacji, ryzyka i cech kandydatów, pochodzą także z Japonii, Rosji, Kanady i Australii. Zaskakujące może być niewielkie zaangażowanie europejskich instytucji naukowych w analizę obszaru, co może wskazywać na potencjalną lukę badawczą.

**Rysunek 2.** Informacja i ryzyko oraz cechy kandydatów (publikacje według daty wydania)



Źródło: opracowanie własne.

Jak zaprezentowano na rysunku 2, zainteresowanie tematem polaryzacji politycznej z powodów związanych z przekazem informacji i cechami kandydatów jako tematem publikacji pozostało stabilne od 2006 r. – od tego czasu ukazywał się średnio jeden artykuł rocznie, co jawi się jako duża liczba, biorąc pod uwagę wąskość zakresu niniejszego przeglądu literatury.

## 5. Podsumowanie

Polaryzacja polityczna to zjawisko towarzyszące wielu współczesnym procesom demokratycznym. Jej występowanie można wyjaśniać na wiele sposobów, spośród których niniejszy artykuł przybliżył kilka. Wyjaśniono w nim rolę przekazywania wiadomości między obywatelami, a także strumieni informacyjnych płynących od polityków. Rozpatrzono w jakim stosunku do polaryzacji politycznej ma się niepewność inicjowana brakiem dostatecznej jakości informacji. Zwrócono także uwagę

na rolę stosunku wobec ryzyka, zarówno po stronie biernych, jak i aktywnych członków sceny politycznej. Przytoczono cechy kandydatów na stanowiska polityczne, które są w stanie doprowadzić do dywergencji od preferencji medianowego wyborcy, wśród nich wartościowość i wolumen politycznej przeszłości. Artykuł może dzięki temu służyć osobom pragnącym działać na rzecz przeciwstawiania się polaryzacji.

Autorzy literatury przedmiotu w większości nie są zgodni co do tego, jakie czynniki wpływają na polaryzację polityczną. Przekazywanie informacji między członkami społeczeństwa może na nią wpływać zarówno pozytywnie, jak i negatywnie, zależnie od kierunków przepływu informacji. Powtarzającym się wątkiem jest finansowanie kanałów informacyjnych kandydatów. To znajduje swoje odzwierciedlenie w innych częściach literatury przedmiotu polaryzacji politycznej, gdzie stawia się na nacisk na rolę partii i aktywistów finansujących polityków. Znaczenie ma również jakość sygnałów informacyjnych – z tego powodu decydujące w wielu ujęciach jest podejście obywateli do ryzyka. Brak jednak konsensusu badawczego co do tego jak to podejście wygląda, czy członkowie społeczeństwa preferują ryzyko, czy raczej go unikają. Wśród naukowców brak też jest zgody co do roli cech kandydatów, a w szczególności ich wartościowości – cech niezwiązanych z poglądami, a raczej ze skutecznością polityka. Przywiązanie wyborców do tych cech, w połączeniu z innymi elementami modelu, może zarówno zwiększać, jak i zmniejszać stopień polaryzacji politycznej.

Niniejszy przegląd literatury badał jedynie wąski fragment literatury na temat przyczyn strategicznej polaryzacji politycznej. Znacznym ograniczeniem było rozpatrywanie jedynie modeli teoretycznych. Ponadto można zaobserwować oprócz informacji, niepewności i cech kandydatów, kilka trendów wartych uwagi w kontekście dalszych badań. Jednym z nich jest analiza funkcji użyteczności głosujących, która pozwala na przyjrzenie się jak same preferencje głosujących, nie wskazujące na polaryzację bezpośrednio, mogą wywoływać ją w sposób pośredni – przykładem może być troska o to jak kandydat poradzi sobie na scenie politycznej. Innym terminem często pojawiającym się w literaturze przedmiotu jest przynależność partyjna i organizacja sceny politycznej. Stanowiska kandydatów są w tym ujęciu przedstawiane jako wypadkowa sił na nich działających: aktywistów, szefów partii i zagrożeń w postaci wstąpienia do wyścigu osób trzecich. Jakkolwiek interesujące, te trendy nie były centrum niniejszego przeglądu, ale mogą stanowić kierunek dalszych badań. Na koniec warto zauważyć, że literatura przedmiotu jak do tej pory nie poruszyła bezpośrednio wpływu zjawisk aktualnych, takich jak pandemia czy wybuch wojny – co stanowi potencjalną lukę badawczą.

## Bibliografia

---

- Adams, J. (1999). Policy Divergence in Multicandidate Probabilistic Spatial Voting, *Public Choice*, 100(1), s. 103–122. DOI: 10.1023/A:1018301007077.
- Berger, M.M., Munger, M.C., Potthoff, R.F. (2000). The Downsian Model Predicts Divergence, *Journal of Theoretical Politics*, 12(2), s. 228–240. DOI: 10.1177/0951692800012002005.
- Black, D. (1958). *The Theory of Committees and Elections*. Norwell: Kluwer Academic Publishers. DOI: 10.1007/978-94-009-4225-7.
- Boxell, L., Gentzkow, M., Shapiro, J.M. (2020). Cross-country Trends in Affective Polarization. Working Paper 26669, National Bureau of Economic Research. DOI: 10.3386/w26669.
- Brams, S.J., Fishburn, P.C. (1978). Approval Voting, *The American Political Science Review*, 72(3), s. 831–847. DOI: 10.2307/1955105.
- Carrillo, J.D., Castanheira, M. (2008). Information and Strategic Political Polarisation, *The Economic Journal*, 118(530), s. 845–874. DOI: 10.1111/j.1468-0297.2008.02152.x.
- Clarivate (2022). <https://clarivate.com/webofsciencelibrary/solutions/webofscience-platform/> (dostęp: 6.06.2022).
- Convers, P.E. (2000). Assessing the Capacity of Mass Electorates, *Annual Review of Political Science*, 3(1), s. 331–353. DOI: 10.1146/annurev.polisci.3.1.331.
- Dellis, A., Gauthier-Belzile, A., Oak, M. (2017). Policy Polarization and Strategic Candidacy in Elections under the Alternative-Vote Rule, *Journal of Institutional and Theoretical Economics*, 173(4), s. 565–585. DOI: 10.1628/093245616X14799744684614.
- Dellis, A., Oak, M. (2016). Multiple Votes, Multiple Candidacies and Polarization, *Social Choice and Welfare* 46(1), s. 1–38. DOI: 10.1007/s00355-015-0900-9.
- Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper & Row.
- Esponda, I., Pouzo, D. (2019). Retrospective Voting and Party Polarization, *International Economic Review*, 60(1), s. 157–186. DOI: 10.1111/iere.12348.
- Fiorina, M.P. (2017). *Unstable Majorities: Polarization, Party Sorting, and Political Stalemate*. Hoover Press. DOI: 10.1002/polq.12880.
- Galeotti, A., Mattozzi, A. (2011). “Personal Influence”: Social Context and Political Competition, *American Economic Journal: Microeconomics*, 3(1), s. 307–327. DOI: 10.1257/mic.3.1.307.
- Green, D., Shapiro, I. (1994). *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven: Yale University Press. DOI: 10.1017/S0008423900019132.
- Grosser, J., Palfrey, T.R. (2014). Candidate Entry and Political Polarization: An Antimedial Voter Theorem, *American Journal of Political Science*, 58(1), s. 127–143. DOI: 10.1111/ajps.12032.
- Ingberman, D., Villani, J. (1993). An Institutional Theory of Divided Government and Party Polarization, *American Journal of Political Science*, s. 429–471. DOI: 10.2307/2111380.
- Jakubowski, M. (2005). Teoria wyboru społecznego. W: *Teoria wyboru publicznego* (s. 46–68), J. Wilkin (red.). Warszawa: Wydawnictwo Naukowe SCHOLAR.
- Karakas, L.D., Mitra, D. (2020). Inequality, Redistribution, and the Rise of Outsider Candidates, *Games and Economic Behavior*, 124, s. 1–16. DOI: 10.1016/j.geb.2020.07.012.
- Kendall, C., Nannicini, T., Trebbi, F. (2015). How Do Voters Respond to Information? Evidence from a Randomized Campaign, *American Economic Review*, 105(1), s. 322–353. DOI: 10.1257/aer.20131063.

- Kollman, K., Miller, J.A.H., Page, S.E. (1992). Adaptive Parties in Spatial Elections, *The American Political Science Review*, 4(26), s. 929–937. DOI: 10.2307/1964345.
- Krasa, S., Polborn, M.K. (2012). Political Competition between Differentiated Candidates, *Games and Economic Behavior*, 76(1), s. 249–271. DOI: 10.1016/j.geb.2012.05.003.
- McCarty, N. (2018). Pivotal Politics, Partisan Polarization, and Policy Predictability, *The Journal of Politics*, 80(3), s. 1076–1081. DOI: 10.1086/697949.
- McLean, I.S., McMillan, A., Monroe, B.L. (Eds.). (2013). *The Theory of Committees and Elections by Duncan Black and Committee Decisions with Complementary Valuation by Duncan Black and RA Newing*. Springer Science & Business Media. DOI: 10.1007/978-94-011-4860-3.
- Polborn, M.K., Snyder Jr, J.M. (2017). Party Polarization in Legislatures with Office-motivated Candidates, *The Quarterly Journal of Economics*, 132(3), s. 1509–1550. DOI: 10.1093/qje/qjx012.
- Redmond, P. (2017). Incumbent-challenger and Open-seat Elections in a Spatial Model of Political Competition, *Public Choice*, 170(1), s. 79–97. DOI: 10.1007/s11127-016-0383-y.
- Schofield, N. (2007). The Mean Voter Theorem: Necessary and Sufficient Conditions for Convergent Equilibrium, *The Review of Economic Studies*, 74(3), s. 965–980. DOI: 10.1111/j.1467-937X.2007.00444.x.
- Tajika, T. (2021). Polarization and Inefficient Information Aggregation under Strategic Voting. *Social Choice and Welfare*, 56(1), s. 67–100. DOI: 10.1007/s00355-020-01270-2.
- Zakharov, A.V. (2009). A Model of Candidate Location with Endogenous Valence, *Public Choice*, 138(3–4), s. 347–366. DOI: 10.1007/s11127-008-9362-2.

## Strategic Political Polarisation from the Perspective of Information, Uncertainty and Valence

### Summary

The text provides a systematic literature review aimed at analysing the phenomenon of strategic political polarisation. The analysis focuses on the synthesis of theoretical attempts to model the strategic political polarisation with the application of game theory. In practice, the popular Downsian model is, in its classical form, unable to explain the settings of most political scenes. This article summarises the way this problem was tackled in the past. Among many answers, there are three aspects that played a key role: valence assigned to non-political candidate values, imperfect information among political parties and voters as well as uncertainty on the part of voters about the future behaviour of politicians. The PRISMA method was used in the literature review. After the analysis of 128 research publications, a synthesis of 14 research articles with the greatest contribution to the discussed issue was conducted.

**Keywords:** political polarisation, political divergence, valence, uncertainty, information



# Wpływ czynników ESG/CSR na koszt kapitału przedsiębiorstw europejskich: przegląd literatury

Piotr Hościłowicz

## Streszczenie

Regulacje europejskie oraz związane z nimi zachowania uczestników rynku wpływają na wzrost znaczenia czynników ESG/CSR w procesie alokacji kapitału. Celem przeglądu literatury jest synteza obecnego stanu wiedzy na temat charakteru oraz kierunku wpływu czynników ESG/CSR na koszt kapitału firm europejskich. Badanie przeprowadzono metodą systematycznego przeglądu literatury PRISMA, do którego zaliczono piętnaście publikacji naukowych, które ukazały się w latach 2012–2021. Analiza artykułów przedstawia, że działania ESG/CSR nie mają jednoznacznie określonego kierunku oraz charakteru wpływu na kształtowanie się kosztu kapitału. Przeprowadzona ocena wskazuje jednak na istotność czynników ESG/CSR w kształtowaniu kosztu kapitału, który jest ważnym elementem kreowania wartości przedsiębiorstwa. Co więcej, informacje te są kluczowe dla wszystkich uczestników otoczenia rynkowego: kadry zarządzającej, inwestorów, dawców kapitału, regulatorów oraz innych zaangażowanych interesariuszy.

**Słowa kluczowe:** ESG, CSR, koszt kapitału, wycena, stopa dyskontowa

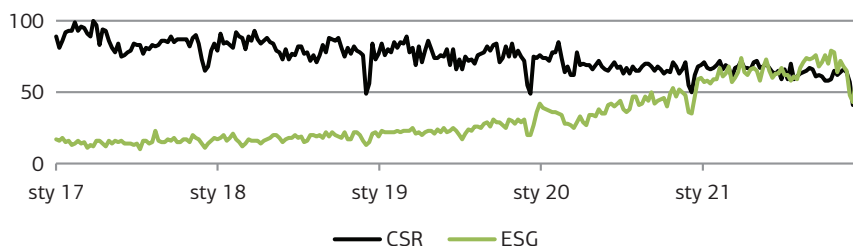
## 1. Podstawy teoretyczne ESG/CSR na płaszczyźnie kosztu kapitału

Uwzględnianie czynników ESG (*Environmental, Social, Governance*) jest obecnie istotnym trendem wpływającym na proces alokacji kapitału, m.in. przez inwestorów instytucjonalnych. Jak podaje agencja Bloomberg [2021], prognozuje się, że inwestycje w aktywa ESG globalnie przekroczą 53 tryliony dolarów w 2025 r. Stanowiąc to będzie ponad jedną trzecią z 140,5 trylionów dolarów aktywów zarządzanych przez fundusze inwestycyjne. Obecnie ponad 3000 inwestorów instytucjonalnych oraz dostawców usług prowadzi działalność zgodnie z *Principles of Responsible Investment*

(PRI) [Gillan, Koch, Starks, 2021]. PRI jest zobowiązaniem do uwzględniania czynników ESG/CSR podczas analiz inwestycyjnych oraz w procesie decyzyjnym. Co więcej, w samej Unii Europejskiej dyrektywa Parlamentu Europejskiego z 15 kwietnia 2014 nakłada na około 6000 największych firm w UE wymóg ujawniania informacji dotyczących polityki ESG/CSR [Komisja Europejska, 2014].

W literaturze przedmiotu oprócz terminu ESG wykorzystywane są również inne terminy celem określenia koncepcji zrównoważonego rozwoju przedsiębiorstw o nieprecyzyjnie określonej, a jednak istotnej roli w ich funkcjonowaniu [Benlemlih, 2017]. Powszechnie wykorzystywane terminy to m.in.: ESG (*Environmental, Social, Governance*), CSR (*Corporate Social Responsibility*), CSP (*Corporate Social Performance*), *Corporate Environmental Responsibility*, *Firm Sustainability* lub SRI (*Socially Responsible Investing*). Co więcej, różne nazwy używane są jako synonimy poszczególnych części składowych inwestycji ESG. Dla przykładu termin *environment* analizujący obszar podejmowania korzystnych decyzji środowiskowych może być alternatywnie określony jako *sustainable*, *green* czy *eco* [Daugaard, 2020]. Rysunek 1 wskazuje jak w ostatnich pięciu latach (2017–2021 YTD) zwrot ESG jest coraz powszechniej wykorzystywany, generując zbliżone zainteresowanie do terminu CSR, odnotowującego trend spadkowy.

**Rysunek 1.** Zainteresowanie pojęciami ESG oraz CSR



Źródło: opracowanie własne na podstawie Google Trends (stan na 17 stycznia 2022 r.).

Grupa Banku Światowego (*The World Bank*) [2004] definiuje CSR jako „zobowiązanie przedsiębiorstw do przyczyniania się do zrównoważonego rozwoju gospodarczego poprzez współpracę z pracownikami, ich rodzinami, społecznością lokalną i całym społeczeństwem w celu poprawy jakości życia w sposób, który jest korzystny zarówno dla przedsiębiorstw, jak i dla rozwoju ogółem”. W skrócie CSR można rozumieć jako działania firmy na rzecz wspierania zrównoważonego rozwoju poprzez dobre praktyki biznesowe [Cui, Jo, Na, 2018]. ESG natomiast definiowane jest jako schemat i system uwzględniający czynniki środowiskowe (E, *environmental*), społecznej odpowiedzialności (S, *social*) oraz ładu korporacyjnego (G, *governance*) [Li,

K. Wang, Sueyoshi, D. Wang, 2021]. Analizując powyższe definicje, ESG można rozumieć jako schemat/system do oceny działań CSR firmy.

Jedną z kluczowych różnic pomiędzy przytoczonymi terminami jest podejście do uwzględniania ładu korporacyjnego. ESG analizuje ten czynnik odrębnie, natomiast CSR włącza ład korporacyjny pośrednio poprzez czynniki środowiskowe oraz społecznej odpowiedzialności. Oznacza to, że ESG jest terminem szerszym niż CSR [Gillan i in., 2021]. Pomimo pewnych różnic ze względu na powszechne użycie wielu z przytoczonych wyżej wyrażeń jako synonimów, w artykule postanowiono stosować pojęcia ESG oraz CSR, zamiennie stosując termin ESG/CSR. Przyjęte podejście zaczerpnięte zostało od Gillana i in. [2021], będących autorami jednego z najnowszych przeglądów literatury w obszarze społecznie odpowiedzialnych inwestycji.

Badania analizujące ESG/CSR poruszają wiele różnych tematów od decyzji finansowych, zachowań konsumentów czy etykę biznesu [Benlemlih, 2017]. Fundamentalnym pytaniem pozostaje jednak nadal, czy (oraz jeśli tak, to w jaki sposób) ESG/CSR przyczynia się do kreowania wartości [Cheng, Ioannou, Serafeim, 2014]. Innymi słowy, jaka jest zależność pomiędzy rezultatami społecznej odpowiedzialności firmy (*Corporate Social Performance*, CSP), a jej wynikami finansowymi (*Corporate Financial Performance*, CFP). Jest to istotne, ponieważ zarządzający przedsiębiorstwem chcą wiedzieć w jaki sposób ESG/CSR może zwiększyć zwroty finansowe, wpłynąć na pozostałe decyzje finansowe [Benlemlih, 2017] oraz w jaki sposób włączyć ESG/CSR do procesów inwestycyjnych [Bassen, Busch, Friede, 2015].

Zakres, w jakim przedsiębiorstwa przynoszą korzyści lub szkodzą dobrobytowi społecznemu, jest tematem rosnącego zainteresowania [Gillan i in., 2021]. Wybrane poprzednie badania nie udzielały jednak jednoznacznych odpowiedzi. Badanie Margolisa, Elfenbein i Walsh [2009] analizujące ponad 200 artykułów na temat relacji pomiędzy CSR i CFP dowiodło, że aż 59% badanych artykułów nie wykazało istotnej zależności. Implikuje to potrzebę dalszych analiz.

Przedmiotem przedstawionego w artykule przeglądu literatury jest wpływ czynników ESG/CSR na koszt kapitału przedsiębiorstw europejskich. Nacisk na koszt kapitału wynika z jego istotnej roli w kształtowaniu wartości przedsiębiorstwa obok generowanych przepływów pieniężnych. Poziom kosztu kapitału warunkuje w wielu przypadkach dalszą dynamikę rozwoju firmy. Przykładowo niższy koszt pozyskania kapitału dłużnego zmniejsza obciążenia odsetkowe przedsiębiorstwa.

Poprzednio zidentyfikowane przekrojowe przeglądy literatury dotyczące wpływu ESG/CSR na finansowe aspekty funkcjonowania firm [Gillan i in., 2021; Benlemlih, 2017] nie wyróżniały potencjalnego wpływu czynników geograficznych. Co więcej, próby firm przyjęte do badania najczęściej nie dotyczyły przedsiębiorstw europejskich. Na osiem analizowanych przez Benlemliha badań [2017], sześć opartych było

na próbach ze Stanów Zjednoczonych [Sharfman, Fernando, 2008; Bauer, Hann, 2010; El Ghouli, Guedhami, Kwok, Mishra, 2011; Dhaliwal, Li, Tsang, Yang, 2011; Harjoto, Jo, 2014; Goss, Roberts, 2011]. Pozostałe dwa artykuły odnosiły się do bazy firm chińskich [Ye, Zhang, 2011] oraz próby międzynarodowej [Girerd-Potin, Jimenez-Garcès, Louvet, 2014].

Wyróżnienie czynnika geograficznego znajduje podstawę we wcześniejszych badaniach. Cai, Pan i Statman [2016] podają, że czynniki specyficzne dla danego kraju są bardziej istotne niż wewnętrzna charakterystyka firmy w wyjaśnianiu różnic w uzyskiwanych wynikach *Corporate Social Performance* (CSP). Autorzy zwracają jednak uwagę, że globalizacja może potencjalnie zmniejszyć istotność czynników krajowych. Wynika to z ułatwionego dostępu firm do rynków finansowych i instytucji w innych krajach.

Celem przeglądu literatury jest zatem uzyskanie odpowiedzi na następujące pytania:

- 1) Czy czynniki ESG/CSR wpływają na koszt kapitału przedsiębiorstw europejskich?
- 2) Jaki jest charakter zależności między czynnikami ESG/CSR a kosztem kapitału przedsiębiorstw europejskich?

W następnym rozdziale opisano metodę badawczą wykorzystaną do przeprowadzenia systematycznego przeglądu literatury. W rozdziale trzecim przedstawiono zagregowane wyniki analizy. Czwarty rozdział zawiera szczegółową analizę treści poszczególnych pozycji. Ostatnią częścią tekstu jest podsumowanie przeglądu wraz ze wskazaniem potencjalnych kierunków przyszłych badań.

## 2. Metoda badawcza

W badaniu wykorzystano metodę PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) będącą metodą systematycznego przeglądu literatury. PRISMA 2020 [Page i in., 2021] składa się z listy kontrolnej obszarów do uwzględnienia w przeglądzie (obejmującej siedem sekcji z 27 pozycjami) oraz diagramem przepływu (*flow diagram*) ilustrującym schemat wyszukiwania i selekcji artykułów. Kluczowymi zaletami przeprowadzania przeglądu literatury w systematyczny sposób są m.in.: 1) synteza obecnego stanu wiedzy w danej dziedzinie, 2) identyfikacja ograniczeń badań oraz 3) wskazanie przyszłych kierunków badań.

Do bazy danych włączone zostały artykuły naukowe opublikowane w języku angielskim, które analizowały wpływ czynników ESG/CSR na koszt kapitału przedsiębiorstw europejskich. Kluczowym parametrem było wykorzystanie w badaniu próby przedsiębiorstw funkcjonujących jedynie w Europie. W bazie uwzględniono

również publikacje, które analizują jedynie konkretny kraj lub kraje europejskie (np. Francję). Co więcej, w badaniu uwzględniono pozycje analizujące koszt kapitału własnego, koszt kapitału obcego lub średni ważony koszt kapitału (WACC). Do bazy danych zostały włączone artykuły opublikowane w dowolnym horyzoncie czasowym. Z analizy wyłączono rozdziały w monografiach naukowych/książkach.

Do tworzenia bazy artykułów wykorzystano stronę Web of Science (WoS), będącą bazą abstraktowo-bibliometryczną. Podczas wyszukiwania wykorzystano również narzędzie Scopus. Ze względu jednak na pełne pokrycie wyników z bazy Scopus w bazie Web of Science ostatecznie posługiwano się jedynie bazą Web of Science. Ostatnie wyszukiwanie zostało przeprowadzone 4 stycznia 2022 r. W celu zidentyfikowania potencjalnych dodatkowych badań przeanalizowano pozycje cytowane oraz bibliografie artykułów uwzględnionych w przeglądzie literatury.

Podstawą wyszukiwania były słowa kluczowe wykorzystane przez Daugaarda [2020] w systematycznym przeglądzie literatury dotyczącym nowych motywów w inwestycjach ESG/CSR. Ostatecznie wykorzystano następujące kombinacje słów kluczowych: 1) "ESG" AND "cost of\*", 2) "CSR" AND "cost of\*", 3) "corporate social responsibility" AND "cost of\*" oraz 4) "social\* responsibil\*" AND "cost of\*".

Wyniki wyszukiwania zostały wyeksportowane oraz analizowane przy użyciu narzędzia MS Excel. Jedna osoba była odpowiedzialna za selekcję artykułów do dalszej analizy na podstawie tytułu artykułu oraz abstraktu. Kolejny etap selekcji polegał na pobraniu pełnych tekstów artykułów. Po zapoznaniu się z treścią zdecydowano czy ostatecznie spełniają kryteria włączenia do analizy.

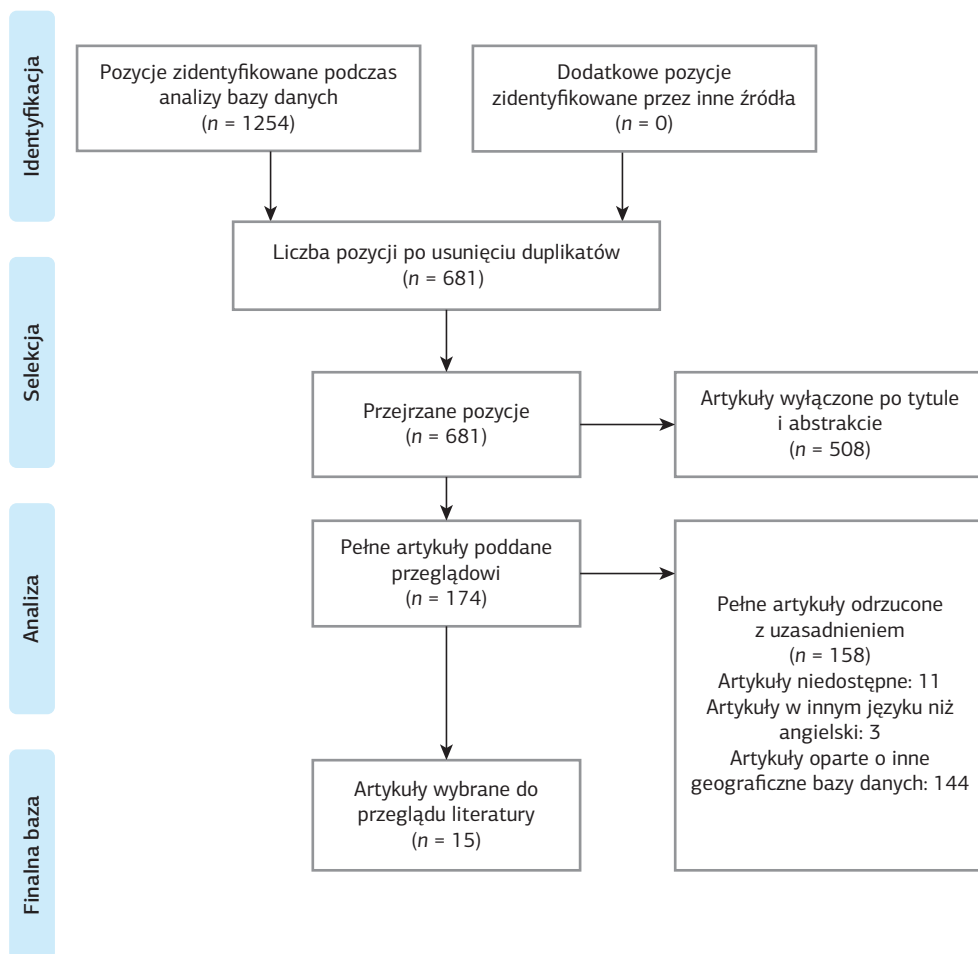
### 3. Wyniki

Wyszukiwanie wykazało 1254 pozycje. Nie zidentyfikowano dodatkowych prac z innych źródeł niż baza Web of Science (m.in. analiza bibliografii wybranych artykułów). Po usunięciu 573 duplikatów analizie poddanych zostało 681 pozycji. Z grupy tej na podstawie kryteriów włączenia/wyłączenia odrzucono 508 artykułów na podstawie tytułu oraz abstraktu. Na ostatnim etapie analizy pełnych tekstów zostały 174 pozycje. Do jedenastu artykułów nie uzyskano dostępu. Trzy artykuły zostały napisane w innym języku niż angielski. Największą grupę odrzuconych pozycji (144) stanowiły źródła wykorzystujące bazę firm z innego regionu geograficznego niż Europa. Były to kolejno bazy: azjatyckie (60 artykułów), międzynarodowe (47 artykułów), USA oraz Kanady (29 artykułów), południowoamerykańskie (4 artykuły), afrykańskie (2 artykuły), australijskie (1 artykuł) oraz bliskowschodnie (1 artykuł). Ostatecznie do przeglądu literatury zakwalifikowano 15 artykułów.

Rysunek 2 przedstawia diagram PRISMA, obrazujący proces identyfikacji oraz selekcji publikacji.

Analiza roku publikacji artykułów potwierdza zwiększone zainteresowanie tematyką ESG/CSR. W samym 2021 r. opublikowanych zostało 8 artykułów, co stanowi 50% finalnej bazy. Najwcześniejsza zidentyfikowana praca została opublikowana w 2012 r. Co więcej, wyłączając rok 2019, kiedy zostały opublikowane trzy artykuły, w każdym z lat 2012–2020 ukazał się maksymalnie jeden artykuł rocznie.

**Rysunek 2.** Diagram selekcji artykułów PRISMA



Źródło: opracowanie własne na podstawie PRISMA Flow Diagram.

Wydawnictwo Wiley opublikowało największą liczbę artykułów w analizowanej bazie (pięć artykułów). Inne wydawnictwa opublikowały po trzy prace (MPDI

oraz Emerald) lub dwie (Elsevier oraz Springer). Czasopisma wydawnictwa Wiley charakteryzują się również wysoką jakością mierzoną wskaźnikiem Journal Impact Factor bazy danych Web of Science. Wskaźnik ten pokazuje ile razy średnio każdy artykuł opublikowany w danym magazynie był cytowany dwa lata po jego publikacji. Dwa z czasopism Wiley, *Business Strategy and the Environment* oraz *Corporate Social Responsibility and Environmental Management* uzyskały wartość wskaźnika na poziomie kolejno 10,302 oraz 8,741. Obie wartości wskazują na wysoką jakość magazynów. W tabeli 1 przedstawiono podsumowanie wydawnictw, magazynów oraz ilości publikacji.

**Tabela 1.** Podsumowanie wydawnictw oraz magazynów publikacji artykułów

Magazyn	Wydawnictwo	Ilość
<i>Sustainability</i>	MPDI	3
<i>Business Strategy and the Environment</i>	Wiley	2
<i>Corporate Social Responsibility and Environmental Management</i>	Wiley	2
<i>Business Ethics, the Environment &amp; Responsibility</i>	Wiley	1
<i>Corporate Governance: The International Journal of Business in Society</i>	Emerald	1
<i>Critical Perspectives on Accounting</i>	Elsevier	1
<i>European Management Journal</i>	Emerald	1
<i>International Journal of Accounting &amp; Information Management</i>	Elsevier	1
<i>Journal of Management Control</i>	Springer	1
<i>Management Decision</i>	Emerald	1
<i>Review of Managerial Science</i>	Springer	1

Źródło: opracowanie własne.

Tabela 2 przedstawia zagregowane informacje na temat wpływu czynników ESG/CSR na kształtowanie kosztu kapitału przedsiębiorstw europejskich. Negatywną relację oznaczono znakiem „-”. Oznacza to, że wzrost aktywności ESG/CSR przyczynia się do spadku kosztu kapitału. Relację pozytywną oznaczono znakiem „+”. W przypadku występowania czynników ograniczających wpływ zastosowano symbol „\*”. Spośród piętnastu badań, sześć analizowało koszt kapitału własnego, sześć kapitału obcego oraz trzy średni ważony koszt kapitału (WACC).

Na piętnaście zidentyfikowanych artykułów, dziesięć wykazało negatywną relację pomiędzy czynnikami ESG/CSR a kosztem kapitału firmy. Należy zaznaczyć jednak, że trzy publikacje wskazały na pozytywną relację przy występowaniu pewnych kryteriów ograniczających skalę wpływu. W dwóch przypadkach badały one średni ważony koszt kapitału (WACC) oraz w jednym przypadku koszt kapitału obcego. Co więcej, dwa źródła nie oceniły jednoznacznie kierunku zależności.

**Tabela 2.** Zestawienie wpływu czynników ESG/CSR na koszt kapitału

Artykuł	Koszt kapitału	Wpływ
Chouaibi, Rossi, Zouari, 2012	własny	(-)
Reverte, 2012	własny	(-)
La Rosa, Liberatore, Mazzi, Terzani, 2018	własny	(-)
Gerged, Matthews, Elheddad, 2021	własny	(+)/(-)
Martnez-Ferrero, Ruiz-Barbadillo, Guidi, 2021	własny	(-)
Michaels, Gruning, 2017	własny	(-)
Ahmed, Eliwa, Power, 2019	obcy	(-)
Hamrouni, Uyar, Boussaada, 2019	obcy	(+)/(-)
Drago, Carnevale, 2020	obcy	(+)*
Bacha, Ajina, Ben Saad, 2021	obcy	(-)
Eliwa, Aboud, Saleh, 2021	obcy	(-)
Kordaschia, 2021	obcy	(-)
Carp, Pavalioia, Afrasinei, Georgescu, 2019	WACC	(+)*
Gjergji, Vena, Sciascia, Cortesi, 2020	WACC	(+)*
Mariani, Pizzutilo, Caragnano, Zito, 2021	WACC	(-)

Źródło: opracowanie własne.

## 4. Dyskusja

### 4.1. Analiza treści artykułów badających koszt kapitału własnego

Chouaibi, Rossi i Zouari [2021] badali wpływ ESG/CSR oraz wynagrodzenia zarządzających oparte na osiągnięciu zrównoważonych celów na koszt kapitału własnego. Autorzy wykorzystali pierwotną bazę 407 firm francuskich należących do indeksu ESG przez sześć lat, w latach 2015–2020. Szczegółowe informacje na temat przedsiębiorstw zostały pozyskane przy użyciu baz danych ASSET4 (DataStream) oraz I/B/E/S. Z próby firm wyłączono instytucje finansowe ze względu na specyfikę podejścia do rachunkowości. Finalna próba składała się z 154 podmiotów, dając razem 924 obserwacje. Autorzy przetestowali trzy hipotezy. Pierwsza hipoteza zakładająca, że ESG/CSR ma negatywny wpływ na koszt kapitału własnego została potwierdzona oraz jest istotna statystycznie. Autorzy tłumaczą zależność przez istotność praktyk ESG/CSR przy podejmowaniu decyzji inwestycyjnych oraz podejście świadomych zjawiska inwestorów. Druga hipoteza oparta na założeniu, że wynagrodzenie zarządcze oparte na celach zrównoważonego rozwoju ma negatywny wpływ na koszt kapitału własnego, nie została potwierdzona. Występuje negatywna,



jednak nieistotna statystycznie zależność. Trzecia hipoteza badała, czy COVID-19 pozytywnie wpływa na koszt kapitału własnego. Ostatnia hipoteza została potwierdzona, będąc istotną statystycznie. Autorzy oceniają, że rynki finansowe wezmą pod uwagę nowy typ ryzyka związany z COVID-19, przyczyniając się tym samym się do wzrostu kosztów finansowania przedsiębiorstw. Dwoma głównymi ograniczeniami badania są: 1) ograniczenie próby do podmiotów francuskich oraz 2) brak testów na endogeniczność (np. relację pomiędzy działaniami ESG/CSR, COVID-19 a kosztem kapitału własnego), które mogą być badane w obu kierunkach.

Reverte [2012] potwierdził statystycznie istotną negatywną relację pomiędzy ratingami raportowania ESG/CSR a kosztem kapitału własnego. Autor dodatkowo wsparł tezę, że zależność ta jest silniejsza, jeśli przedsiębiorstwo funkcjonuje w sektorze wrażliwym na aspekty środowiskowe. Koszt kapitału własnego spada o 88 punktów bazowych między najniższym a najwyższym kwintylem w rankingu CSR. Próba składała się z 35 publicznie notowanych hiszpańskich firm wchodzących w skład indeksu IBEX35 (35 największych podmiotów w Hiszpanii pod względem kapitalizacji rynkowej) w latach 2003–2008. Dane dotyczące raportowania CSR/ESG pochodziły z raportów OCSR (*Observatory on Corporate Social Responsibility*). Badanie sugeruje, że koszt kapitału własnego jest istotnym kanałem, przez który rynek uwzględnia raportowanie ESG/CSR. Autor wskazuje, że raportowanie ESG/CSR może przyczynić się do precyzyjniejszego odzwierciedlenia wartości rynkowej przedsiębiorstwa. Potencjalnym ograniczeniem badania jest sytuacja, w której raportowanie jakości ESG/CSR może być przybliżeniem innych czynników, takich jak jakość kadry zarządzającej, i prowadzić do mylących wyników. Związane jest to z tzw. *halo effect*, czyli błędem poznawczym, w którym postrzeganie jednej zmiennej jest zniekształcone/zastępowane przez łatwą dostępność innej zmiennej.

La Rosa, Liberatore, Mazzi i Terzani [2018] badali związek pomiędzy ESG/CSR a kosztem kapitału własnego. Próba badawcza obejmowała 227 brytyjskich podmiotów notowanych na London Stock Exchange w latach 2010–2014. Wykorzystaną bazą danych była Thomson Reuters DataStream. Badania potwierdziły, że firmy z wyższymi poziomami ESG/CSR charakteryzują się niższym kosztem kapitału własnego. Wyniki wskazują, że przedsiębiorstwa posiadające agendy ESG/CSR mogą uzyskać przewagę konkurencyjną poprzez zmniejszanie negatywnego wpływu swojej działalności na społeczeństwo lub otaczające środowisko, przy jednoczesnym obniżeniu kosztów finansowania kapitałem własnym w Wielkiej Brytanii. Sugerują, że inwestorzy premiuje „dobre” spółki ESG/CSR, oczekując relatywnie mniejszych zwrotów finansowych z akcji, co przekłada się na mniejszy koszt kapitału własnego. Dalsze badania mogą skupić się na poznaniu czynników wpływających na decyzje

firm, celem zaangażowania w praktyki ESG/CSR. Potencjalnie istotne może być również przeprowadzenie badań międzykrajowych oraz odkrycie przyczyn oczekiwania przez inwestorów mniejszych zwrotów z inwestycji w podmioty charakteryzujące się potwierdzoną aktywnością w obszarze ESG/CSR.

Gerged, Matthews i Elheddad [2021] analizowali wpływ przymusowego raportowania emisji gazów cieplarnianych na koszt kapitału własnego. Próba składała się z 406 spółek notowanych publicznie w latach 2012–2016 na FTSE350 (350 największych spółek o największej kapitalizacji na London Stock Exchange). Wyniki wskazują, że wysoki poziom raportowania emisji gazów cieplarnianych jest negatywnie skorelowany z kosztem kapitału własnego. W pewnym momencie jednak, określonym jako punkt zwrotny (*turning point*), każdy wzrost raportowania emisji gazów cieplarnianych przyczynia się do wzrostu kosztu kapitału. Oznacza to, że nieliniowa zależność pomiędzy zmiennymi przyjmuje kształt litery „U”. Przed brytyjskimi regulacjami z czerwca 2013, wymagającymi od wszystkich publicznie notowanych podmiotów raportowania emisji gazów cieplarnianych, punkt zwrotny został obliczony na 1,711 tys. ton. Po wprowadzeniu regulacji jest to 1,38 tys. ton. Autorzy sugerują, że ryzyko emisji gazów cieplarnianych i przymusowe ich raportowanie są powiązane w sposób, który pozwala firmom o wyższym poziomie ryzyka emisji gazów cieplarnianych być efektywniejszymi podmiotami pod względem ich ujawniania/raportowania. Kluczowym ograniczeniem badania jest brak uwzględnienia zarządzania ryzykiem emisji gazów cieplarnianych i kosztu ich kontroli na koszt kapitału, co może być potencjalnym istotnym czynnikiem dla uczestników rynku.

Martínez-Ferrero, Ruiz-Barbadillo oraz Guidi [2021] badali w jaki sposób wiarygodność i dokładność raportowania ESG/CSR zwiększa zaufanie rynku poprzez wpływ na koszt kapitału. Próba badania składała się z 1580 rocznych obserwacji 353 firm europejskich z 20 krajów publicznie notowanych w latach 2012–2018, pobranych z bazy Thomson Reuters Eikon. Wyniki potwierdziły, że zapewnienie jakości raportowania ESG/CSR zwiększa wiarygodność i pozytywnie wpływa na zaufanie rynku, obniżając koszt kapitału. Zjawisko to jest związane m.in. ze zmniejszeniem asymetrii informacji. Autorzy zbadali również wpływ przekształceń raportów ESG/CSR na dokładność raportowania. Jeśli przekształcenia związane są z aktualizacją definicji, zakresem lub kalkulacjami metodologicznymi, to zwiększają one dokładność raportowania, co przekłada się na wzrost zaufania rynku. Podawanie do informacji publicznej błędów lub pominięć w poprzednich raportach ESG/CSR jest jednak karane i może skutkować wzrostem kosztu kapitału. Najistotniejsze ograniczenia badania dotyczą braku uwzględnienia materialności przekształcenia raportu. Dodatkowo występować mogą również inne determinanty oprócz zapewnienia jakości ESG/CSR wpływające na poziom zaufania rynkowego.

Michaels i Gruning [2017] potwierdzili, że raportowanie ESG/CSR jest istotnie negatywnie skorelowane z asymetrią informacji oraz kosztem kapitału. Próba badania składała się z 264 niemieckich przedsiębiorstw. Badania dodatkowo wspierają tezę, że firmy funkcjonujące w branżach wrażliwych na czynniki ESG/CSR charakteryzują się wyższym kosztem kapitału. W odniesieniu do korzyści finansowych, raportowanie ESG/CSR może być postrzegane jako szczególny rodzaj dobrowolnego ujawniania informacji. Wyniki wspierają tezę, że ujawnianie informacji na temat ESG/CSR wykracza poza cel zaspokajania różnych roszczeń interesariuszy i służy jako „licencja na działalność” (*license to operate*), co może generować istotne obniżenie kosztów dla firm. Implikuje to możliwość kreowania wartości w długim terminie, pozytywnie wpływając na osiągnięte wyniki. Autorzy jako ograniczenia badania wskazują potencjalne omińnięcie dodatkowych czynników, które mogą mieć wpływ na konstrukty teoretyczne. Dodatkowo przyjęty przekrojowy charakter badania (*cross-sectional design*) może istotnie ograniczyć wyniki, jako konsekwencja możliwego, jednak nieokreślonego przesunięcia w czasie pomiędzy raportowaniem ESG/CSR, asymetrią informacji a kosztem kapitału.

#### 4.2. Analiza treści artykułów badających koszt kapitału obcego

Ahmed, Eliwa i Power [2019] badali wpływ aktywności CSP na koszt oraz dostęp do kapitału obcego. Próba składała się z niefinansowych firm europejskich publicznie notowanych w latach 2005–2012 (razem 1228 rocznych obserwacji firm). Badania wskazują na negatywną relację pomiędzy aktywnością CSP a kosztem kapitału obcego oraz pozytywną relację pomiędzy CSP a ratingiem kredytowym. Wyniki sugerują, że firmy o wyższym CSP (co potwierdza ich społecznie odpowiedzialne podejście) stały się synonimem nie tylko mocniejszego wizerunku przedsiębiorstwa, ale również tworzą oczekiwania dotyczące wyników finansowych oraz zmniejszają ekspozycję na ryzyko. Co więcej, im większy poziom jakości CSP, tym większa dźwignia finansowa. Statystyczna zależność pomiędzy jakością CSP, kosztem kapitału obcego oraz dźwignią finansową jest jednak jedynie istotna podczas okresów stabilności ekonomicznej. W okresach niestabilności ekonomicznej autorzy sugerują, że pożyczkodawcy ignorują czynniki ESG/CSR, skupiając się jedynie na aspektach wyników finansowych. Badanie nie brało pod uwagę zrównoważonego rozwoju środowiskowego oraz czynników ładu korporacyjnego. Co więcej, przyszłe badania powinny dać odpowiedź, czy zwroty ekonomiczne uzyskane dzięki CSP w formie obniżonego kosztu kapitału obcego i zwiększonej dźwigni przewyższają koszty inwestycji koniecznej do implementacji CSP.

Hamrouni, Uyar i Boussaada [2019] badali w jaki sposób raportowanie ESG/CSR wpływa na koszt kapitału obcego. Próba składała się z 80 francuskich przedsiębiorstw notowanych w indeksie SBF 120 w latach 2010–2015 (z wyłączeniem regulowanych usług komunalnych oraz instytucji finansowych). Do uzyskania danych wykorzystano bazę Bloomberg. Wyniki wskazują, że ogólne raportowanie ESG/CSR wpływa negatywnie na koszt kapitału obcego. O ile ujawnianie informacji środowiskowych jest skorelowane negatywnie, o tyle ujawnianie informacji związanych ze społeczeństwem – pozytywnie. Informacje dotyczące ładu korporacyjnego nie mają istotnego wpływu na koszt kapitału obcego. Autorzy podają, że raportowanie informacji środowiskowych jest jedyną płaszczyzną ESG/CSR, która ma wpływ na decyzje instytucji udzielających finansowania, celem udzielenia korzystnych warunków. Pozytywna korelacja informacji społecznych może wynikać z marnowania lub kosztownej konsumpcji ograniczonych zasobów firmy przez zarządzających. Badanie ograniczone jest brakiem uwzględnienia ograniczeń umownych i zobowiązań, potencjalnie występujących w kontraktach dłużnych. Co więcej, analiza została zawężona do jednego kraju, w konkretnym okresie, wyłączając pewne sektory.

Drago i Carnevale [2020] badali w jaki sposób ocena ESG/CSR wpływa na spready kredytów konsorcjalnych (*syndicated loan*). Próba badania składała się z 1101 kredytów udzielonych 297 firmom niefinansowym z Unii Europejskiej w latach 2006–2015. Autorzy potwierdzili, że rating ESG/CSR ma znaczący wpływ na spready kredytów. Nisko oceniane firmy ESG/CSR płacą wyższe spready kredytowe niż lepiej oceniane podmioty, jednak wysokie oceny ESG/CSR nie zawsze są doceniane przez kredytodawców. Korzyści z posiadania wysokiej oceny ESG/CSR są istotne jedynie w przypadku firm zlokalizowanych w krajach, które zwracają dużą uwagę na kwestie zrównoważonego rozwoju. Skutkuje to sytuacją, w której inwestycje zakładające pozyskanie wysokiej oceny ESG/CSR są wynagradzane jedynie, kiedy przedsiębiorstwo funkcjonuje w otoczeniu skupiającym się na czynnikach ESG/CSR. Co więcej, analiza autorów sugeruje, że ratingi ESG/CSR są determinantami drugiego rzędu spreadów kredytowych, które są brane pod uwagę dopiero po ocenie „tradycyjnych” fundamentów firmy (m.in. danych księgowych oraz ratingów kredytowych).

Bacha, Ajina i Ben Saad [2021] sprawdzali, jaki jest efekt ESG/CSR na koszt kapitału obcego oraz czy jakość audytu wpływa na koszt kapitału obcego, który jest ponoszony przez firmy odpowiedzialne społecznie. Próba przyjęta do badania składała się z 87 francuskich spółek niefinansowych publicznie notowanych w latach 2005–2016. Autorzy wykorzystali również bazę danych DataStream od Thomson Reuters-ASSET4. Badanie dowiodło, że działania ESG/CSR są negatywnie skorelowane z kosztem kapitału obcego. Autorzy wskazują, że banki utożsamiają ESG/CSR ze wzrostem transparentności. Co więcej, wyższe oceny przyznane przez agencje

ratingowe ESG/CSR również pełnią rolę gwarancyjną dla dawców kapitału, w szczególności odpowiedzialnych społecznie banków. Dalsze badanie potwierdziło również, że pozytywny efekt ESG/CSR na koszt kapitału jest bardziej obserwowalny, jeśli dana firma korzysta z wysokiej jakości usług audytorskich (w szczególności spółek z tzw. grupy Big4 – Deloitte, PwC, E&Y oraz KMPG). Dalsze eksperymenty powinny skupić się na wynagrodzeniu nieaudytowym (*non-audit fees*). Wysoki stosunek wynagrodzenia nieaudytowego do całkowitego wynagrodzenia za audyt może potencjalnie wzmocnić relację bieglego rewidenta z klientem i osłabić jego niezależność, wpływając tym samym na jakość badania.

Eliwa, Aboud i Saleh [2021] badali, czy instytucje udzielające finansowania wynagradzają firmy za zaangażowanie w oraz publikację czynników ESG/CSR niższym kosztem kapitału obcego. Próba składała się z 6018 podmiotów niefinansowych z 15 krajów Unii Europejskiej (Austrii, Belgii, Danii, Finlandii, Francji, Niemiec, Grecji, Irlandii, Włoch, Holandii, Portugalii, Norwegii, Hiszpanii, Szwecji i Wielkiej Brytanii), przy użyciu baz danych Thomson Reuters Asset4, Bloomberg oraz Thomson Reuters DataStream. Norwegia została uwzględniona ze względu na stosowanie tych samych standardów rachunkowości oraz regulacji. Wyniki wskazują, że podmioty europejskie mogą korzystać ze zwiększonych poziomów zaangażowania i publikacji ESG/CSR, co przekłada się na niższy koszt kapitału obcego. Autorzy zwracają uwagę, że czynniki rynkowe, reprezentowane przez instytucje finansujące, odgrywają znaczącą rolę w istotności oraz wiarygodności zaangażowania i ujawniania ESG/CSR. Interesującym odkryciem jest, że raportowanie ESG/CSR ma nie tylko taki sam wpływ na koszt kapitału obcego, ale również działa jako substytut dla jakości ESG/CSR. Implikuje to brak rozróżnienia przez instytucje finansujące pomiędzy raportowaniem ESG/CSR a jego jakością. Dodatkowo autorzy udowodnili, że wpływ praktyk ESG/CSR na koszt kapitału własnego jest większy w krajach zorientowanych na interesariuszy (*stakeholder-oriented countries*), m.in. tam, gdzie społeczeństwo obywatelskie odgrywa istotną rolę. Badanie ograniczone jest 1) wykorzystaniem źródeł wtórnych w postaci zewnętrznych baz danych, 2) miarami ESG/CSR wykorzystanymi w badaniu oraz 3) wykorzystaniem miernika odsetek od długu do średniego rocznego poziomu długu oprocentowanego do obliczenia kosztu długu. Powyższy wskaźnik potencjalnie może zniekształcić wyniki, jeśli firma manipuluje poziomem długu blisko końca roku.

Kordsachia [2021] analizował związek pomiędzy ESG/CSR a marginalnym kosztem kredytu. Próba składała się z 778 podmiotów europejskich z wyłączeniem instytucji finansowych w latach 2014–2017 (całkowita liczba 2117 obserwacji). Autor wykorzystał wtórne bazy Thomson Reuters DataStream/ASSET4 publicznie notowanych instrumentów. Badanie dowiodło, że firmy proaktywne na polu ESG/CSR

są wynagradzane niższym kosztem kapitału obcego. Co więcej, pozytywny wpływ ESG/CSR na koszt kapitału obcego jest statystycznie i ekonomicznie ważniejszy, kiedy podmioty doświadczają problemów finansowych (prawdopodobieństwo upadłości). Czynniki związane z radą zarządzającą sugerują zmniejszanie wymagań odsetkowych kredytodawców. Statystyczną istotność potwierdziły niezależność, wielkość oraz zróżnicowanie ze względu na płeć rady zarządzającej. Autorzy argumentują, że te cechy rady związane są z lepszą interakcją z interesariuszami oraz mniejszą skłonnością do podejmowania ryzyka. Należy zwrócić uwagę jednak, że działania ESG/CSR niekoniecznie muszą przynieść wartość dla udziałowców, którzy oceniają ryzyko inaczej niż kredytodawcy. Główne ograniczenia badania wskazane przez autorów to: 1) wyłączenie firm z sektora finansowego oraz 2) wykorzystanie zewnętrznych ratingów dostarczanych przez Thomson Reuters ASSET4 (pomimo że jest to często wykorzystywane podejście), ponieważ ESG/CSR oraz zasadność raportowania ESG/CSR nie są bezpośrednio obserwowalne.

### 4.3. Analiza treści artykułów badających średni ważony koszt kapitału

Carp, Pavaloaia, Afrasinei i Georgescu [2019] weryfikowali wpływ raportowania ESG/CSR na rozwój firm, jako rezultat adoptowania działań odpowiedzialnych środowiskowo oraz społecznie. Próba składała się z 59 spółek w Rumunii publicznie notowanych w latach 2012–2017. Hipoteza autorów zakładała, że raportowanie ESG/CSR istotnie wpływa na koszt kapitału (WACC). Wyniki wskazują, że podmioty raportujące informacje ESG/CSR w formie zarówno opisowej, jak i skwantyfikowanej cechują się wyższym WACC. Zrównoważone działania interpretowane są jako czynniki przyczyniające się do zmniejszenia wydajności, które generują niepewność zdolności do spłaty kredytu, co związane jest z dodatkowymi kosztami. Szczegółowa analiza różnych szeregów wartości WACC sugeruje, że negatywny wpływ aktywności ESG/CSR jest obserwowalny jedynie w przypadku podmiotów cechujących się relatywnie niskim poziomem WACC w porównaniu do średniej badanej próby (znajdujące się w pierwszym i drugim kwartylu). W przypadku podmiotów z największym WACC (trzeci kwartył) relacja nie jest istotna statystycznie, co zaznacza ograniczony wpływ raportowania ESG/CSR na wzrost WACC. Autorzy jako ograniczenia wskazują skupienie się na firmach notowanych na Giełdzie Papierów Wartościowych w Bukareszcie (BSE) oraz wykorzystaniu jedynie jakościowych zmiennych do oceny raportowania ESG/CSR.

Gjergji, Vena, Sciascia i Cortesi [2020] badali wpływ raportowania ESG/CSR na kształtowanie się WACC w segmencie małych i średnich przedsiębiorstw. Autorzy wskazują, że w dużym kontraście do danych z dużych podmiotów, raportowanie

ESG/CSR w segmencie SME przyczynia się do wzrostu kosztu kapitału. Małe i średnie przedsiębiorstwa nie są w stanie pokryć kosztów upowszechniania informacji ESG/CSR powiązanych korzyściami. Co więcej, odrębna analiza przeprowadzona przez autorów trzech komponentów (*Environmental, Social, Governance*) wskazuje, że jedynie pierwszy wymiar (*Environmental*) ma znaczenie dla uczestników rynku. Nie dotyczy to jednak rodzinnych SME, których WACC obniża się wraz ze wzrostem raportowania ESG/CSR. Finalna próba składała się z 87 podmiotów notowanych na włoskiej giełdzie AIM Italia (z wyłączeniem instytucji finansowych lub niespełniających innych kryteriów, m.in. ilości danych). Badanie ograniczone jest przez wykorzystanie jedynie próby firm włoskich oraz danych z jednego roku (2018). Co więcej, nie wskazano, które płaszczyzny i charakterystyki firmy rodzinnej wpływają na odmienne kształtowanie kosztu kapitału niż SME.

Mariani, Pizzutilo, Caragnano i Zito [2021] określali wpływ polityk środowiskowych na średni ważony koszt kapitału (WACC), celem zrozumienia reakcji rynków kapitałowych na zobowiązania środowiskowe przedsiębiorstw i efektywność w redukowaniu emisji dwutlenku węgla. Próba składała się z 577 spółek na STOXX Europe publicznie notowanych w latach 2014–2018. Autorzy potwierdzili silną zależność pomiędzy wysokim zaangażowaniem w zmniejszanie oddziaływania na środowisko a zmniejszonym poziomem WACC. Wyniki wskazują, że zwiększenie o 10 punktów bazowych wyniku emisyjności/ *emission score* (rozumianego jako Thomson Reuters Emission Score, liczonego w skali od 0 do 100, gdzie 0 oznacza brak zaangażowania i efektywności w redukowaniu emisji środowiskowych), przyczynia się do obniżki WACC o 4,4 punktów bazowych dużych europejskich podmiotów. Dodatkowo wpływ pozytywnego zaangażowania środowiskowego na WACC przedsiębiorstw jest istotny zarówno dla branż wysoce-, jak i niskoemisyjnych.

## 5. Podsumowanie oraz kierunki przyszłych badań

Odpowiadając na pierwsze pytanie badawcze, należy zwrócić uwagę, że przeprowadzony przegląd potwierdził, iż czynniki ESG/CSR wpływają na koszt kapitału przedsiębiorstw europejskich. Oznacza to, że elementy ESG/CSR są istotnym kanałem wpływającym na kreowanie wartości przedsiębiorstwa. Informacje te są ważne dla wszystkich uczestników otoczenia rynkowego, od kadry zarządzającej po inwestorów, dawców kapitału, regulatorów oraz innych zaangażowanych interesariuszy. Należy jednak zwrócić uwagę, że raportowanie może być postrzegane jako substytut dla jakości ESG/CSR. W tej sytuacji wzrasta rola instytucji ratingowych oraz firm audytorskich. Co więcej, poszczególni interesariusze różnią się w postrzeganiu

działań ESG/CSR. Wynika to z innych potrzeb oraz motywacji do wykorzystania danych niefinansowych. W przypadku kredytodawców może być to ograniczenie ekspozycji na ryzyko. Inwestorzy mogą być jednak bardziej zainteresowani maksymalizacją osiągniętej stopy zwrotu. Istotne jest zwrócenie uwagi, że raportowanie ESG/CSR przekłada się na projekcję wydajności przedsiębiorstwa. Firmy powinny być świadome, że nie w każdym przypadku inwestorzy pozytywnie postrzegają te dodatkowe informacje. Konieczne jest wtedy wyjaśnienie relacji korzyści do kosztów implementacji projektów ESG/CSR oraz w jaki sposób podejmowane działania przyczyniają się do wzrostu przewagi konkurencyjnej.

W przypadku odpowiedzi na drugie pytanie, analiza artykułów przedstawia, że działania ESG/CSR nie mają jednoznacznego kierunku oraz charakteru wpływu na kształtowanie się kosztu kapitału przedsiębiorstw europejskich. Poszczególne elementy ESG/CSR nie charakteryzują się identyczną skalą oddziaływania. W niektórych przypadkach jedynie czynnik środowiskowy może być istotny dla interesariuszy. W ekstremalnym przypadku może doprowadzić jedynie do skupienia się na minimalizacji negatywnego wpływu środowiskowego przy zaniechaniu działań w obszarze społecznej odpowiedzialności i ładu korporacyjnego. Brak holistycznego podejścia do ESG/CSR nie jest jednak podejściem optymalizującym zrównoważony rozwój przedsiębiorstwa.

Należy zauważyć, że wpływ wskaźników ESG/CSR jest silniejszy, jeśli spełnione są dwa kryteria. Po pierwsze, dana spółka funkcjonuje w otoczeniu zwracającym uwagę/świadomym czynników ESG/CSR. Po drugie, wielkość spółki również odgrywa istotną rolę, ponieważ skala działalności determinuje relację pomiędzy uzyskanymi korzyściami do poniesionych kosztów wynikających z raportowania ESG/CSR. W przypadku przedsiębiorstw nieposiadających wystarczających zasobów może przełożyć się to na rezygnację z raportowania informacji niefinansowych, zwiększając asymetrię informacji. Przyszłe regulacje powinny skupić się na zachęcaniu firm do ujawniania danych ESG/CSR, celem m.in. precyzyjniejszego określenia wartości rynkowej przedsiębiorstw. Co więcej, istotna jest motywacja przedsiębiorstw, które nie funkcjonują w branży lub kraju zwracającym uwagę na kwestie zrównoważonego rozwoju. Potencjalne kierunki przyszłych badań powinny m.in. skupić się na analizie porównawczej pomiędzy regionami oraz krajami Europy. Wynika to nie tylko ze zróżnicowania poziomu rozwoju gospodarczego czy specyfiki rynku kapitałowego. Również legislacje na poziomie krajowym mogą stanowić istotny element różnicujący charakter oraz kierunek kształtowania się zmiennych. Dodatkowo dalsze analizy szczegółowo rozpatrujące poszczególne czynniki ESG/CSR (środowisko naturalne, społeczeństwo oraz ład korporacyjny) i ich wpływ na koszt kapitału pozwoliłyby na możliwość podejmowania precyzyjniejszych decyzji przez



zarządzających. Wynika to potencjalnie z możliwości zrównoważenia i maksymalizacji wartości wszystkich interesariuszy. Co więcej, kwantyfikacja wpływu ESG/CSR (szczególnie na koszt kapitału własnego) dostarczyłaby precyzyjniejszych informacji do porównania korzyści i kosztów zaangażowania w społecznie odpowiedzialne aktywności. Potrzebne są również dalsze badania skupiające się na instytucjach finansowych, ponieważ ze względu na ich specyfikę była to grupa często wyłączana. Istotne będą też kolejne analizy skupiające się na jakości audytu. Szczególnie ważnym będzie określenie w jaki sposób przekształcanie raportu lub potencjalna manipulacja publikowanymi danymi może wpłynąć na zaufanie rynku nie tylko do konkretnej spółki, ale również do danego sektora ogółem. W tym przypadku jakość audytu powiązana jest również z sytuacją, w której większość ocen o ESG/CSR formułowana jest na podstawie zewnętrznych ratingów. Zbadanie różnic pomiędzy ratingiem, raportowaniem firmy oraz danymi pierwotnymi pomoże ocenić jaka jest skala asymetrii informacji lub przekształcenia faktycznych danych pomiędzy wiedzą spółki a interesariuszami. Podsumowując, przyszłe badania w obszarze ESG/CSR pozwolą na określenie, czy uczestnicy rynku odbierają inaczej informacje o nieprawidłowym raportowaniu i przekazywaniu informacji finansowych, a inaczej informacji o charakterze niefinansowym.

## Bibliografia

---

- Ahmed, A.H., Eliwa, Y., Power, D.M. (2019). The Impact of Corporate Social and Environmental Practices on the Cost of Equity Capital: UK Evidence, *International Journal of Accounting and Information Management*, 27, s. 425–441. DOI: 10.1108/IJAIM-11-2017-0141.
- Bacha, S., Ajina, A., Ben Saad, S. (2021). CSR Performance and the Cost of Debt: Does Audit Quality Matter?, *Corporate Governance*, 21, s. 137–158. DOI: 10.1108/CG-11-2019-0335.
- Bassen A., Busch, T., Friede, G. (2015). ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies, *Journal of Sustainable Finance & Investment*, 5, s. 210–233. DOI: 10.1080/20430795.2015.1118917.
- Bauer, R., Hann, D. (2010). *Corporate Environmental Management and Credit Risk*, Maastricht University. DOI: 10.2139/ssrn.1660470.
- Benlemlih M. (2017). Corporate Social Responsibility and Firm Financing Decisions: A Literature Review, *Journal of Multinational Financial Management*, 42–43(3), s. 1–10. DOI: 10.1016/j.mulfin.2017.10.004.
- Bloomberg (2021). *ESG Assets May Hit \$ 53 Trillion by 2025, a Third of Global AUM*, <https://www.bloomberg.com/professional/blog/esg-assets-may-hit-53-trillion-by-2025-a-third-of-global-aum/> (dostęp: 20.11.2021).
- Cai, Y., Pan, C.H., Statman, M. (2016). Why Do Countries Matter so Much in Corporate Social Performance?, *Journal of Corporate Finance*, 41, s. 591–609. DOI: 10.1016/j.jcorpfin.2016.09.004.

- Carp, M., Pavaloaia, L., Afrasinei, M.B., Georgescu, I.E. (2019). Is Sustainability Reporting a Business Strategy for Firm's Growth? Empirical Study on the Romanian Capital Market, *Sustainability*, 11. DOI: 10.3390/su11030658.
- Cheng, B.T., Ioannou, I., Serafeim, G. (2014). Corporate Social Responsibility and Access to Finance, *Strategic Management Journal*, 35, s. 1–23. DOI: 10.1002/smj.2131.
- Chouaibi, Y., Rossi, M., Zouari, G. (2021). The Effect of Corporate Social Responsibility and the Executive Compensation on Implicit Cost of Equity: Evidence from French ESG Data, *Sustainability*, 13. DOI: 10.3390/su132011510.
- Cui, J., Jo, H., Na, H. (2018). Does Corporate Social Responsibility Affect Information Asymmetry?, *Journal of Business Ethics*, 148, s. 549–572. DOI: 10.1007/s10551-015-3003-8.
- Daugaard, D. (2020). Emerging New Themes in Environmental, Social and Governance Investing: A Systematic Literature Review, *Accounting and Finance*, 60, s. 1501–1530. DOI: 10.1111/acfi.12479.
- Dhaliwal, D., Li, O., Tsang, A., Yang, G. (2011). Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting, *The Accounting Review*, 86, s. 59–100. DOI: 10.2308/accr.00000005.
- Drago, D., Carnevale, C. (2020). Do CSR Ratings Affect Loan Spreads? Evidence from European Syndicated Loan Market, *Sustainability*, 12. DOI: 10.3390/su12187639.
- El Ghouli, S., Guedhami, O., Kwok, C., Mishra, D. (2011). Does Corporate Social Responsibility Affect the Cost of Capital?, *Journal of Banking & Finance*, 35, s. 2388–2406. DOI: 10.1016/j.jbankfin.2011.02.007.
- Eliwa, Y., Aboud, A., Saleh, A. (2021). ESG Practices and the Cost of Debt: Evidence from EU Countries, *Critical Perspectives on Accounting*, 79. DOI: 10.1016/j.cpa.2019.102097.
- Gerged, A.M., Matthews, L., Elheddad, M. (2021). Mandatory Disclosure, Greenhouse Gas Emissions and the Cost of Equity Capital: UK Evidence of a U-shaped Relationship, *Business Strategy and the Environment*, 30, s. 908–930. DOI: 10.1002/bse.2661.
- Gillan, S.L., Koch, A., Starks, L.T. (2021). Firms and Social Responsibility: A Review of ESG and CSR Research in Corporate Finance, *Journal of Corporate Finance*, 66. DOI: 10.1016/j.jcorpfin.2021.101889.
- Girerd-Potin, I., Jimenez-Garcès, S., Louvet, P. (2014). Which Dimensions of Social Responsibility Concern Financial Investors?, *Journal of Business Ethics*, 121, s. 559–576. DOI: 10.1007/s10551-013-1731-1.
- Gjergji, R., Vena, L., Sciascia, S., Cortesi, A. (2020). The Effects of Environmental, Social and Governance Disclosure on the Cost of Capital in Small and Medium Enterprises: The Role of Family Business Status, *Business Strategy and the Environment*, 30, s. 683–693. DOI: 10.1002/bse.2647.
- Goss, A., Roberts, G.S. (2011). The Impact of Corporate Social Responsibility on the Cost of Bank Loans, *Journal of Banking & Finance*, 35, s. 1794–1810. DOI: 10.1016/j.jbankfin.2010.12.002.
- Hamrouni, A., Uyar, A., Boussaada, R. (2019). Are Corporate Social Responsibility Disclosures Relevant for Lenders? Empirical Evidence from France, *Management Decision*, 58, s. 267–279. DOI: 10.1108/MD-06-2019-0757.

- Harjoto, M.A., Jo, H. (2014). Legal vs. Normative CSR: Differential Impact on Analyst Dispersion, Stock Return Volatility, Cost of Capital, and Firm Value, *Journal of Business Ethics*, 128, s. 1–20. DOI: 10.1007/s10551-014-2082-2.
- Komisja Europejska (2014). *Disclosure of Non-Financial Information: Europe's Largest Companies to Be More Transparent on Social and Environmental Issues*, [https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT\\_14\\_291](https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_14_291) (dostęp: 20.11.2021).
- Kordsachia, O. (2021). A Risk Management Perspective on CSR and the Marginal Cost of Debt: Empirical Evidence from Europe, *Review of Managerial Science*, 15, s. 1611–1643. DOI: 10.1007/s11846-020-00392-2.
- La Rosa, F., Liberatore, G., Mazzi, F., Terzani, S. (2018). The Impact of Corporate Social Performance on the Cost of Debt and Access to Debt Financing for Listed European Non-Financial Firms, *European Management Journal*, 36, s. 519–529. DOI: 10.1016/j.emj.2017.09.007.
- Li, T.T., Wang, K., Sueyoshi, T., Wang, D.R.D. (2021). ESG: Research Progress and Future Prospects, *Sustainability*, 13. DOI: 10.3390/su132111663.
- Margolis, J.D., Elfenbein, H.A., Walsh, J.P. (2009). Does it Pay to Be Good...And Does it Matter? A Meta-Analysis of the Relationship between Corporate Social and Financial Performance. DOI: 10.2139/ssrn.1866371.
- Mariani, M., Pizzutilo, F., Caragnano, A., Zito, M. (2021). Does it Pay to Be Environmentally Responsible? Investigating the Effect on the Weighted Average Cost of Capital, *Corporate Social Responsibility and Environmental Management*, 28, s. 1854–1869. DOI: 10.1002/csr.2164.
- Martínez-Ferrero, J., Ruiz-Barbadillo, E., Guidi, M. (2021). How Capital Markets Assess the Credibility and Accuracy of CSR Reporting: Exploring the Effects of Assurance Quality and CSR Restatement Issuance, *Business Ethics, the Environment & Responsibility*, 30, s. 551–569. DOI: 10.1111/beer.12355.
- Michaels, A., Gruning, M. (2017). Relationship of Corporate Social Responsibility Disclosure on Information Asymmetry and the Cost of Capital, *Journal of Management Control*, 28, s. 251–274. DOI: 10.1007/s00187-017-0251-z.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C. i in. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews, *BMJ*, 372(71). DOI: 10.1136/bmj.n71.
- Reverte, C. (2012). The Impact of Better Corporate Social Responsibility Disclosure on the Cost of Equity Capital, *Corporate Social Responsibility and Environmental Management*, 19, s. 253–272. DOI: 10.1002/csr.273.
- Sharfman, M.P., Fernando, C.S. (2008). Environmental Risk Management and the Cost of Capital, *Strategic Management Journal*, 29, s. 569–592. DOI: 10.1002/smj.678.
- The World Bank (2004). *Public Sector Roles in Strengthening Corporate Social Responsibility: Taking Stock*, s. 3.
- Ye, K., Zhang, R. (2011). Do Lenders Value Corporate Social Responsibility? Evidence from China, *Journal of Business Ethics*, 104, s. 197–206. DOI: 10.1007/s10551-011-0898-6.

# The Impact of ESG/CSR Factors on the Cost of Capital of European Enterprises. A Literature Review

## Summary

European regulations and the related behaviour of market participants increase the importance of ESG/CSR factors in the capital allocation process. The aim of literature review is to synthesise the current state of knowledge of the nature and direction of impact of ESG/CSR factors on the cost of capital of European firms. The study was conducted with the PRISMA systematic literature review method. Fifteen research studies published in 2012–2021 were ultimately included in the review. The analysis of articles indicates that ESG/CSR measures do not have a clearly defined direction and character of impact on the cost of capital. The literature review, however, indicates the importance of ESG/CSR factors in shaping the cost of capital, which is an important element of the company value creation. Importantly, this information is relevant to all participants in the market environment: executives, investors, capital providers, regulators and other stakeholders involved.

**Keywords:** ESG, CSR, cost of capital, valuation, discount rate

# Aktywne i pasywne strategie inwestycyjne na rynkach akcyjnych – przegląd systematyczny literatury naukowej

Mateusz Izdebski

## Streszczenie

Rynek akcyjny jest istotną częścią rynku finansowego. Jego kapitalizacja na koniec 2020 r. wyniosła niespełna 94 biliony dolarów amerykańskich. Celem niniejszego opracowania jest analiza badań dotyczących optymalnych strategii inwestycyjnych na tym niezwykle istotnym rynku. Znaczący udział strategii aktywnych jak i pasywnych na rynku akcyjnym stawia pytania o skuteczność tych alternatywnych rozwiązań. Niniejsze badanie poszukuje odpowiedzi na cztery pytania badawcze. W ramach przeglądu badań, dokonanego za pomocą metodologii PRISMA, analizie poddano 26 wyselekcjonowanych artykułów. Kluczowe wnioski: 1) strategie aktywne osiągają przeciętnie stopy zwrotu niższe od benchmarku, 2) prowadzi to do wzrostu popularności strategii pasywnych, 3) strategie aktywne istnieją m.in. dlatego, że osiągają lepsze wyniki w określonych momentach cyklu, w określonych sektorach i strategiach, 4) inwestorzy podejmują decyzje w oparciu o inne czynniki niż stopa zwrotu. Ponadto autorzy argumentują, iż menedżerowie często stosują obie strategie jednocześnie, a rozwój strategii pasywnych wpływa także na inne aspekty rynku niż stopy zwrotu, m.in. koncentrację branży, zmienność rynku, wysokość opłaty.

**Słowa kluczowe:** inwestowanie pasywne, inwestowanie aktywne, alokacja kapitału, efektywność rynku

## 1. Wstęp

Historia rynku akcyjnego na najbardziej rozwiniętych obecnie rynkach, tj. amerykańskim, japońskim i państw zachodnioeuropejskich sięga kilku stuleci, jednak dopiero w XX w. nastąpił znaczny wzrost znaczenia rynku akcyjnego na rynkach finansowych, ale także w gospodarkach poszczególnych państw. Przez większość tego okresu inwestorzy stosowali strategię aktywnej selekcji spółek, która miała zapewnić maksymalizację stopy zwrotu. Wraz z pojawieniem się teorii portfelowej [Markowitz, 1952] oraz teorii efektywności rynków [Fama, 1970] naukowe podstawy zyskała strategia alternatywna, której fundamentalnym założeniem jest próba pasywnego odwzorowania szerokiego zestawu spółek zwanego indeksem bądź benchmarkiem rynkowym. Jednakże gwałtowny rozwój funduszy pasywnych nastąpił dopiero na początku XXI w. – w 2009 r. fundusze pasywne stanowiły już 18% całego rynku funduszy inwestycyjnych, których kapitalizacja globalna wynosiła 8,5 bilionów USD. Po dekadzie udział tychże funduszy wzrósł do ok. 40% w rynku o kapitalizacji nieco ponad 21 bilionów dolarów [Anadu, Kruttli, McCabe, Osambela, 2020].

Niniejsza praca stanowi podsumowanie badań dotyczących strategii pasywnych i aktywnych. Jej fundamentalnym zagadnieniem jest próba weryfikacji czy tak znaczące tempo rozwoju funduszy pasywnych pozostaje zgodne z wynikami badań naukowych w tymże obszarze. W trakcie analizy postawiono następujące pytania badawcze, na które próbę odpowiedzi podjęto w niniejszym opracowaniu:

- 1a) Czy badania dotyczące stóp zwrotu osiąganych przy pomocy strategii pasywnych oraz aktywnych uzasadniają znaczący wzrost udziału strategii pasywnych w rynku?
- 1b) Jeżeli strategię pasywne osiągają przeciętnie lepsze stopy zwrotu, to co uzasadnia fakt, iż wciąż ponad połowa światowego kapitału na rynku akcji wykorzystuje podejście aktywne?
- 2) Czy równoczesne wykorzystanie obu strategii skutkuje poprawą osiąganych wyników inwestycyjnych?
- 3) Na jakie aspekty rynku, inne niż stopa zwrotu, wpływa rozwój strategii pasywnych?

W trakcie analizy badań naukowych dotyczących aktywnych i pasywnych strategii inwestycyjnych nie zidentyfikowano przeglądu systematycznego adresującego kwestie przedstawione w powyższych pytaniach badawczych. Obszar badawczy jest niezwykle istotny dla kształtowania współczesnej gospodarki, a zwłaszcza rynków finansowych, ponadto posiada długą historię niezwykle ważnych i przełomowych badań, które na przestrzeni kilkudziesięciu lat skutkowały zmianą percepcji opła-

całości poszczególnych strategii inwestycyjnych. Kolejne badania w tym obszarze stanowią wkład do fragmentu problemu, toteż z reguły badania w swych opisach teoretycznych odnoszą się do opracowań bezpośrednio z nimi związanych. Autorzy tychże badań, powołując się na publikacje je poprzedzające, podejmują próbę weryfikacji, polemiki bądź udoskonalenia i rozszerzenia istniejących wniosków. Wszystkie te czynniki przyczyniają się do faktu, iż niniejsza praca, będąca przeglądem systematycznym, pozwala zapisać istotną lukę, jaką był brak systematycznego podsumowania badań odpowiadających na przedstawione pytania badawcze. W dalszej części artykułu przedstawiono metodę badawczą zastosowaną w tymże opracowaniu, wyniki tejże analizy oraz dyskusję wyników wyselekcjonowanych badań stanowiącą próbę odpowiedzi na postawione pytania badawcze.

## 2. Metoda badawcza

Przegląd systematyczny jest metodą badawczą i procesem identyfikacji oraz krytycznej oceny odpowiednich badań, a także zbieraniem i analizą danych z tychże badań [Moher, Liberati, Tetzlaff, Altman, 2009]. Celem tego rodzaju przeglądu jest zidentyfikowanie wszystkich dowodów empirycznych, które spełniają uprzednio zdefiniowane kryteria włączenia, aby odpowiedzieć na określone pytanie badawcze lub hipotezę.

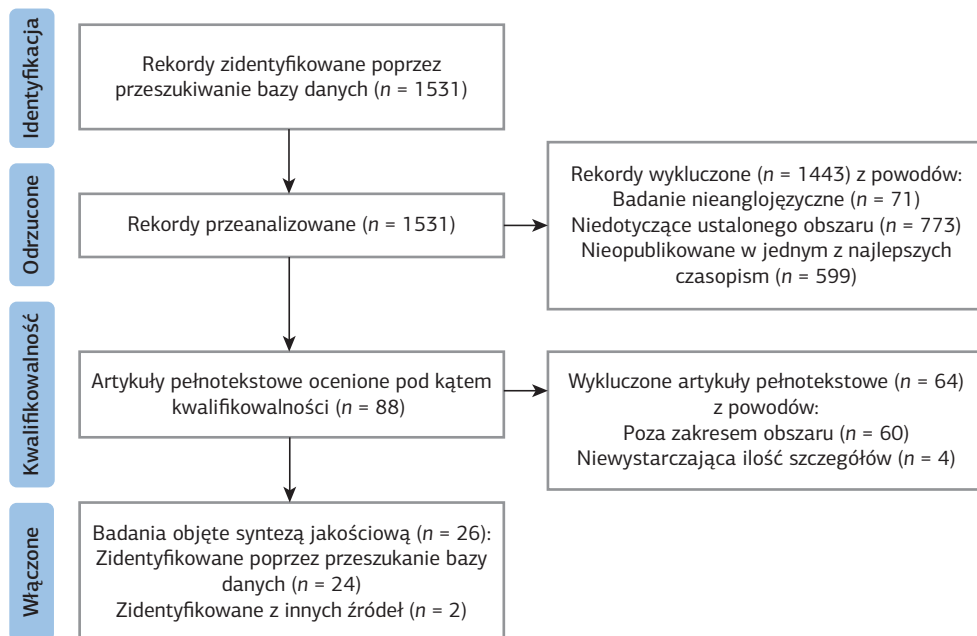
Aby zapewnić przejrzystość i powtarzalność, do systematycznego przeglądu literatury zastosowano podejście zawarte w deklaracji PRISMA. Podejście systematyczne zostało opracowane w ramach nauk medycznych jako sposób na syntezę wyników badań w sposób systematyczny, przejrzysty i powtarzalny. Dzięki zastosowaniu jednoznacznych i systematycznych metod podczas przeglądania artykułów i wszystkich dostępnych dowodów, można zminimalizować stronniczość, dostarczając w ten sposób wiarygodnych wyników, na podstawie których możliwym staje się wyciągnięcie wniosków i podejmowanie decyzji [Moher i in., 2009]. Zastosowanie tejże metodologii niesie za sobą istotne korzyści i możliwości potencjalnego wkładu w dorobek badanego obszaru. Możliwym jest m.in. ustalenie, czy zakładane oddziaływanie jest spójne i trwałe pomiędzy poszczególnymi badaniami oraz jakie mogą być dalsze kierunki badań, aby potwierdzić lub odrzucić przyjęte wnioski. Narzędzie, jakim jest przegląd systematyczny, może być pomocne w celu wykazania które zmienne na poziomie badania lub próby wpływają na badane zjawisko, np. czy badania przeprowadzone w jednym otoczeniu kulturowym dają znacząco różne wyniki niż badania przeprowadzone w innych kontekstach kulturowych [Davis, Mengersen, Bennett, Mazerolle, 2014].

### 3. Wyniki analizy

Przeglądu systematycznego dokonano przy użyciu bazy danych Web of Science, zawierającej dane dotyczące publikacji oraz cytowań dla wielu dyscyplin akademickich. Analiza została przeprowadzona w dniu 29 grudnia 2021 r. Aby zidentyfikować możliwie najszerszy zakres publikacji dotyczących tematyki niniejszego opracowania, przygotowano następujące zapytanie: (“index\*” OR “benchmark\*” OR “ETF\*” OR “passive\*”) AND “active\*” AND (“invests” OR “investo\*” OR “investin\*” OR “investme\*”).

Poniższy diagram przedstawia kolejne kroki analizy wraz z informacją o zidentyfikowanej liczbie badań, które podlegają włączeniu bądź wykluczeniu z dalszej analizy.

**Rysunek 1. Wyniki przeglądu systematycznego**



Źródło: opracowanie własne.

Opracowane zapytanie wymaga, aby wyniki zawierały co najmniej jeden element z każdej z trzech kategorii słów w dowolnym z pól (tytuł, temat, abstrakt, treść artykułu). Pierwsza grupa dotyczy pojęć związanych ze strategiami pasywnymi, kolejna dotyczy strategii aktywnych, ostatnia natomiast wymaga słowa wskazującego na aspekty związane z inwestowaniem. Zapytanie to pozwoliło na uzyskanie 1531 badań spełniających te kryteria, w kolejnym etapie wykluczono opracowania



nieangielskojęzyczne. Następnie określono, iż badania należące muszą do jednej z pięciu dziedzin naukowych związanych z finansami, ekonomią lub wykorzystaniem matematyki w naukach społecznych. W celu zapewnienia najwyższej jakości analizowanych badań wykluczono wszystkie badania, które nie znajdują się na liście 50 najlepszych czasopism naukowych w obszarze finansów według zestawienia SCImago Journal Rank. Ostatnim krokiem była analiza treści opracowań, w której wyniku wyselekcjonowano 24 badania odpowiadające na zdefiniowane pytania badawcze. Zestaw analiz rozszerzono o dwa badania, które nie zostały opublikowane w jednym z najlepszych czasopism naukowych, jednak liczba cytowań oraz użyteczność wniosków wynikających z tychże prac powoduje, iż stanowią one istotny wkład w dorobek tego obszaru naukowego.

#### **4. Wzrost popularności strategii pasywnych w obliczu badań naukowych dotyczących stóp zwrotu**

Jednym z pierwszych i niezwykle istotnych dla kształtowania percepcji strategii pasywnych jest badanie empiryczne opublikowane w 1995 r. [Malkiel, 1995]. Opracowanie to powstało, gdy większość ówczesnych publikacji wskazywała, iż fundusze zarządzane w sposób aktywny osiągają wyniki lepsze aniżeli benchmark przed, a w niektórych przypadkach także po uwzględnieniu kosztów, oraz że przewagi te są trwałe. Wnioski były sprzeczne ze znaną już wówczas hipotezą efektywności rynków Famy i badaniem Jensena [1968]. Wykorzystanie unikalnego zbioru danych, zawierającego zwroty ze wszystkich funduszy istniejących w każdym roku dla lat 1971–1991 pokazuje, że fundusze radziły sobie gorzej od portfeli benchmarków zarówno po kosztach zarządzania, jak i brutto [Malkiel, 1995]. Badanie zwraca również uwagę na fakt, iż błąd przetrwania może być ważniejszy, aniżeli dotychczas zakładano. Trwałość wyników widoczna w latach 70., w latach 80. uległa załamaniu. Wnioskiem z badania jest stwierdzenie, iż aktywne zarządzanie na ogół nie zapewnia nadzwyczajnych zysków (przewyższających benchmark) i ma tendencję do generowania większych obciążeń podatkowych dla inwestorów, więc wnioski wynikające z badań Famy i Jensena dotyczące przewagi pasywnego zarządzania zostają potwierdzone i umocnione.

Specyficzną kategorią spółek notowanych na giełdzie papierów wartościowych są REITy, a więc spółki będące funduszami inwestującymi w rynek najmu nieruchomości. W badaniu tegoż rynku autorzy Kallberg, Liu i Trzcinka [2000] przyglądają się stopom zwrotu z tego rodzaju funduszy. W przeciwieństwie do większości badań funduszy inwestycyjnych, w tymże badaniu średnia i mediana alfa (przed

uwzględnieniem kosztów) są dodatnie i znaczące, jeśli zastosuje się benchmarki z badań, do których odnosili się autorzy. Badacze wskazują na fakt, iż zarządzający funduszami REIT mogą zwiększać wartość poprzez aktywne zarządzanie portfelem, zwłaszcza gdy rynek nieruchomości jest w fazie spadkowej. Wyniki te uzasadniane są istnieniem trwale cennych informacji prywatnych, które wydają się być bardziej prawdopodobne na rynku nieruchomości, który jest rynkiem znacznie mniej płynnym i podlegającym silniejszym wstrząsom agregatowym aniżeli pozostałe rynki aktywów finansowych. Ponadto badanie dostarcza wniosków, iż obrót funduszu oraz jego wielkość są pozytywnie skorelowane z alfą funduszu, natomiast wydatki nie są z nią znacząco związane.

Autorzy badania *Luck versus Skill in the Cross-Section of Mutual Fund Returns* [Fama, French, 2010] dochodzą do wniosku, iż łączny portfel amerykańskich funduszy akcyjnych zarządzanych aktywnie jest zbliżony do benchmarku rynkowego, toteż dodatnie koszty zarządzania przekładają się na niższe zwroty dla inwestorów. Przed kosztami alfa inwestorów pasywnych wynosi 0, badacze uznają więc za Sharpe'ę [1991], że rynek inwestorów aktywnych musi również być grą o sumie zerowej przed kosztami, a więc dodatnia alfa jednego inwestora uzyskiwana jest kosztem ujemnej alfy innego. Badanie dostarcza także ważnych wniosków dotyczących testów trwałości alf wśród zarządzających funduszami. Dowodzą, że testy te opierają się na krótkoterminowych stopach zwrotu, gdzie istotnym czynnikiem jest losowość. Zastosowanie alternatywnej metody opartej o symulację bootstrap dla danych z lat 1984–2006 dostarcza wniosków, iż niewielu zarządzających jest w stanie osiągać dodatnie alfy po uwzględnieniu kosztów, a istotnym czynnikiem obok umiejętności pozostaje szczęście. Wyniki symulacji dostarczają również silnych dowodów umiejętności zarządzających, zarówno negatywnych, jak i pozytywnych.

Bogle [2014] przedstawia próbę oszacowania, jak wysokie obciążenie dla stóp zwrotów z funduszy inwestycyjnych stanowią wydatki inwestycyjne, w tym nie tylko dotychczas powszechnie stosowane wskaźniki kosztów, ale także koszty transakcji funduszu, obciążenia sprzedażowe, przetrzymywanie gotówki, koszty nieefektywności podatkowej i szkodliwe zachowania inwestorów. Autor, stosując konserwatywne podejście do oszacowanych wielkości kosztów, dochodzi do wniosku, że choć w krótkim okresie koszty mogą sprawiać wrażenie nieznaczących i mających niewielki wpływ, to w horyzoncie inwestycji w ramach planów emerytalnych, uniknięcie tychże kosztów pozwoliłoby na osiągnięcie majątku wyższego o ok. 65%.

Przedstawione dotychczas badania w większości sugerują wyższość strategii pasywnych. Do odwrotnych wniosków dochodzą Gallagher, Harman, Schmidt i Warren [2017]. Dotychczasowe badania z reguły skupiały się na najbardziej dojrzałym

rynku amerykańskim, natomiast badacze rozszerzają zbiór danych na rynek globalnych funduszy akcyjnych. Analiza wyników odnotowanych przez 143 globalne fundusze akcyjne w latach 2002–2012 pozwala dojść do wniosku, iż przeciętny zarządzający akcjami przewyższa benchmark o 1,2–1,4%. Autorzy dokonują dodatkowo dekompozycji nadwyżkowych stóp zwrotów funduszy na wkład wynikający z selekcji akcji, selekcji kraju inwestycji oraz wpływ zmian na rynku walutowym. Głównym źródłem nadwyżki jest selekcja akcji, niewielki dodatni wpływ wykazuje wybór krajów, natomiast efekty walutowe są mieszane. Badacze przedstawiają także wniosek, że fakt, iż zarządzający mogą uzyskiwać znaczące nadwyżki na rynkach wschodzących, wynikać może z wpływu efektywności lub segmentacji rynku na zdolności aktywnych zarządzających do osiągania lepszych wyników. Inwestowanie aktywne jest warte rozważenia zwłaszcza przez inwestorów, którzy posiadają dostęp do funduszy rynków globalnych, których opłaty wynoszą poniżej 1%.

Kolejnym badaniem istotnym dla percepcji wyników osiągniętych zarówno przez strategie aktywne, jak i pasywne, było badanie autorstwa Crane'a i Crotty'ego [2018]. Autorzy wskazują, że gdy metody dotychczas stosowane do pomiaru umiejętności menedżerskich doboru portfeli w przypadku przekroju funduszy aktywnych zastosuje się do funduszy pasywnych, to testy sugerują istnienie tychże umiejętności. Umiejętności te są trwałe i, co zaskakujące, mają podobną proporcję jak w przypadku funduszy aktywnych. Rozproszenie wyników funduszy pasywnych wynika z niejednorodności benchmarków oraz niejednorodności błędów śledzenia. Wykorzystanie rozkładów stóp zwrotu funduszy indeksowych jako zestawu funduszy kontrfaktycznych do oceny przyrostu wyników w rozkładzie funduszy aktywnych powoduje, że jakkolwiek przewaga w wynikach strategii aktywnych staje się znacznie mniejsza, aniżeli wskazywałoby odniesienie do średnich wyników funduszy indeksowych. Ponadto autorzy badania dostrzegają, że przewaga najlepszych funduszy aktywnych nad najlepszymi pasywnymi jest znacznie mniejsza niż przewaga najgorszych indeksowych nad najgorszymi aktywnymi. Oznacza to, że żaden inwestor charakteryzujący się awersją do ryzyka nie powinien preferować losowania z rozkładu funduszy aktywnych w stosunku do losowania z rozkładu funduszy indeksowych.

Song [2020] przedstawia natomiast implikacje wynikające z przepływu kapitału do funduszy aktywnych na podstawie ich przeszłych stóp zwrotu. Autor zauważa, że przepływy funduszy wzajemnych reagują pozytywnie na stopy zwrotu związane z ekspozycją na wspólne czynniki, np. czynniki wskazane przez Famę i Frencha [2010]. Mylenie skutków ekspozycji na czynniki systematyczne z umiejętnościami zarządzania skutkuje zjawiskiem przeinwestowania, tzn. nadmiernym gromadzeniem aktywów pod zarządzaniem, co prowadzi do znacząco negatywnych przyszłych stóp zwrotu po uwzględnieniu opłat. Efekt ten jest na tyle silny i powszechny, że

podzbiór ten odpowiada za ujemne łączne zwroty wszystkich aktywnych funduszy inwestycyjnych akcyjnych.

W badaniu z 2020 r. [Leippold, Rueegg, 2020] po raz kolejny podjęto próbę oceny alfa aktywnych i indeksowych funduszy inwestycyjnych na próbie funduszy globalnych. Badanie wykorzystuje ponad 60 tys. funduszy akcyjnych i o stałym dochodzie w latach 1992–2016. Autorzy dokonują rozróżnienia na inwestorów instytucjonalnych i detalicznych, jako że istnieją znaczne różnice w kosztach, korzyściach skali oraz asymetrii informacji pomiędzy tymi dwiema grupami. Badacze wskazują, iż znacząco ujemne alfy po opłatach w przypadku funduszy akcyjnych uzyskano jedynie w kategorii „US Equity Large Cap Blend” dla funduszy instytucjonalnych. W pozostałych przypadkach autorzy nie identyfikują dowodów odrzucających hipotezę, iż aktywne inwestowanie to gra o sumie zerowej po opłatach. Powyższe stwierdzenia są zbliżone do wniosków z badania *Global Equity Fund Performance: An Attribution Approach* [Gallagher i in., 2017] oraz poprzedzających je badań rynku amerykańskiego. Przeciętni aktywni zarządzający mają mniejszą ekspozycję na ryzyko rynkowe, natomiast preferują inwestycje w instrumenty, które dobrze prosperują w dobrych czasach, ale mogą mieć znaczną stratę w bessie. Analizując portfele aktywnych funduszy, autorzy sugerują aktywnym inwestorom detalicznym unikanie funduszy o wysokich opłatach oraz niewielkich stratach w przeszłości. Odrzucenie aktywnych funduszy inwestycyjnych o najniższych stopach zwrotu za minione 12 miesięcy może znacząco poprawić wyniki inwestora detalicznego, jednakże skupienie się na funduszach o najwyższych stopach zwrotu nie może znacząco poprawić wyników.

Rok później także Gerakos, Linnainmaa i Morse [2021] dochodzą do wniosków, które nie wspierają tezy o wyższości inwestycji pasywnych nad aktywnymi. Choć w przypadku tego badania, dotyczącego aktywnie zarządzanych rachunków instytucjonalnych w latach 2000–2012, znaczącą rolę w kształtowaniu wniosków odegrały dodatnie alfy osiągnięte na rynku papierów dłużnych, które nie stanowią elementu niniejszego przeglądu. Należy jednak wspomnieć o tymże badaniu, gdyż autorzy, wychodząc od założenia, iż powszechna opinia o aktywnym zarządzaniu jako o grze o sumie ujemnej po opłatach nie jest uzasadniana przez ogromnych rozmiarów przepływy na rachunkach instytucjonalnych. Autorzy stwierdzają, iż w trakcie 12 analizowanych lat średnia stopa zwrotu z dolara instytucjonalnego uzyskana dzięki pośrednictwu przewyższała stopę zwrotu z rynku o 134 punkty bazowe, co oznacza, że przeciętny dolar nieinstytucjonalny osiągnął wynik gorszy od rynkowego o 0,55 p.p. Autorzy dowodzą także, że przeciętny instytucjonalny zarządzający aktywami jest wykwalifikowany, ale wzrost rynku funduszy ETF prawdopodobnie osłabia przewagę tej grupy zarządzających.

## 5. Wyniki dotyczące przeciętnych stóp zwrotu a trwałość, znacząca pozycja strategii aktywnych

Jednym z pierwszych badań, stanowiących próbę uzasadnienia szybkiego rozwoju oraz utrzymującego się znaczącego udziału aktywnego zarządzania na rynkach finansowych było badanie Grubera [1996]. Proponowane uzasadnienie opiera się o charakterystykę wyceny otwartych funduszy wzajemnych, które to sprzedawane i kupowane są po cenie uwzględniającej jedynie wartość aktywów netto funduszu. Oznacza to, że zdolności zarządzania mogą nie być wyceniane. W kontrze do otwartych funduszy wzajemnych przedstawione zostają fundusze zamknięte, których wycena odbywa się w sposób pozwalający określić zdolności zarządzających do osiągnięcia wyników trwale lepszych bądź słabszych od indeksu. W okresie analizy (1985–1994) zauważono trwałość wyników. Inwestorzy świadomi tego faktu nabywali umiejętności zarządzających niejako „za darmo”, gdyż ceny jednostek funduszy wzajemnych nie odzwierciedlały tychże umiejętności. Świadomi klienci osiągnęli więc dodatkowy zysk, który nie byłby możliwy do osiągnięcia w przypadku inwestycji w fundusze zamknięte. Umiejętności zarządzających mogły być jednakże wyceniane w sposób pośredni, poprzez opłaty. Gruber dowodzi, że nie jest to prawdą, fundusze o najlepszych stopach zwrotu nie pobierały średnio wyższych opłat oraz opłaty te nie rosły wraz z dobrymi wynikami. Pomimo trwałości wyników, znaczna część kapitału pozostaje w rękach funduszy o niskich stopach zwrotu. Autor uzasadnia ten fakt istnieniem dwóch rodzajów klientów określanym mianem klienteli wyrażonej i klienteli pokrzywdzonej (*sophisticated clientele, disadvantaged clientele*). Przy czym druga kategoria dzieli się na trzy grupy: a) inwestorów niewyszukanych, którzy inwestują w oparciu o inne czynniki, np. reklamę i porady brokerów, b) inwestorów w niekorzystnej sytuacji instytucjonalnej, np. inwestujących poprzez konta emerytalne ograniczone przez plan, którego są częścią, c) inwestorów w niekorzystnej sytuacji podatkowej, dla których ze względów podatkowych optymalnym wyborem jest pozostawienie funduszy w gorzej prosperujących funduszach.

W badaniu z 2001 r. autorzy Baks, Metrick i Wachter [2001] również poszukują uzasadnienia tak dużych aktywów inwestowanych przy użyciu strategii aktywnych. Istniejące wówczas badania w większości dowodziły, iż ogół funduszy nie pokonuje benchmarków, jednak na przełomie wieków opublikowano badania, które wykazywały możliwą trwałość dodatnich alf najlepszych zarządzających. Badacze starają się odpowiedzieć na pytanie, czy wszyscy inwestorzy powinni unikać strategii aktywnych. Przyjmują hipotezę zerową, mówiącą iż żaden zarządzający nie posiada umiejętności. Następnie ustalają, iż dane nie przeczą tej hipotezie, co w konsekwencji powinno skutkować preferowaniem strategii pasywnych względem aktywnych.

Autorzy pokazują jednak, że krok 2 nie implikuje kroku 3, oraz że istniejące dowody nie są wystarczająco silne, aby odróżnić postawioną hipotezę od bliskich alternatyw, z których niektóre implikują duże inwestycje w strategię aktywne. Ponadto autorzy podkreślają, iż na decyzje inwestycyjne wpływ mają nie tylko dane statystyczne, ale też opłaty, podatki i ograniczenia krótkiej sprzedaży. Strata na opłatach przy braku umiejętności jest ograniczona, a potencjalny zysk znacznie większy, co może uzasadniać skłonność części inwestorów do ponoszenia kosztów opłat.

Kolejnym znaczącym badaniem w kontekście zadanego pytania badawczego jest opracowanie Avramova i Wermersa [2006]. Autorzy wskazują, iż poszczególni zarządzający funduszami znacznie częściej osiągają stopy zwrotu powyżej benchmarku w określonym momencie cyklu gospodarczego, niż uśredniając w każdych warunkach. Inwestor dostosowujący strategię do cyklu może osiągnąć trwale dodatnią alfę. Dotychczas zakładano, że umiejętności zarządzających lub ich braki są trwałe, w konsekwencji zarządzający osiągający świetne wyniki w poszczególnych momentach cyklu pozostawali niezauważeni. W celu określenia faz cyklu badacze stosują odczyty czterech zmiennych makroekonomicznych. Avramov i Wermers dowiedli, że prawdopodobieństwo pokonania rynku przez menadżerów w silnym stopniu zależy od tychże zmiennych. 1/3 zarządzających wykazała trwałość tejszej zdolności, natomiast jedynie 10% uzyskuje dodatnią alfę w całym okresie analizy (1980–2002). Hipotetyczny portfel zbudowany na podstawie tej strategii uzyskał średniorocznie 8 p.p. powyżej benchmarku. Portfel zdominował również strategię aktywne o podobnej charakterystyce. Zastosowanie w praktyce jest ograniczone ze względu na koszty częstego dostosowania portfela funduszy, ponadto złożoność obliczeń powoduje, że strategia przedstawia wyższą użyteczność w przypadku inwestorów instytucjonalnych.

Kenneth R. French w badaniu *Presidential Address: The Cost of Active Investing* przedstawia cztery elementy składające się na opłaty, wydatki i koszty handlowe, które społeczeństwo ponosi, aby inwestować na giełdzie amerykańskiej w sposób aktywny, w porównaniu do tego, co zostałyby zapłacone, gdyby wszyscy inwestowali pasywnie [French, 2008]. Składniki te to: a) opłaty oraz wydatki ponoszone przez inwestorów w związku z funduszami inwestycyjnymi, w tym funduszami otwartymi, zamkniętymi i ETF, b) koszty, jakie ponoszą inwestorzy instytucjonalni związane z zarządzaniem inwestycjami, c) opłaty inwestorów ponoszone w związku z funduszami hedgingowymi, d) koszty, jakie ponoszą wszyscy inwestorzy w związku z zawieraniem transakcji. Autor, stosując założenie o braku transferu netto w przypadku portfeli pasywnych, oszacowuje, że inwestor, który posiada pasywny portfel rynkowy, uzyskuje w latach 1980–2006 wynik lepszy od średniej ważonej wartością wszystkich aktywnych i pasywnych inwestorów o 0,67 p.p. rocznie. Przy zało-

zeniu, że oczekiwana realna stopa zwrotu z amerykańskich akcji wynosi ok. 6,7% rocznie, skapitalizowany koszt inwestowania aktywnego w 2008 r. wyniósł ok. 10% ówczesnej kapitalizacji rynku. French [2008] przedstawia także własne uzasadnienie dla istnienia przepływów do funduszy aktywnych, co, jak dowodzi autor, jest grą o sumie ujemnej. Pierwszym z powodów jest błędne postrzeganie możliwości inwestycyjnych, do czego przyczyniają się osoby zawodowo zajmujące się rynkiem akcyjnym. Kolejnymi argumentami są nadmierna pewność siebie inwestorów, chęć posiadania unikalnego portfela inwestycyjnego oraz posiadanie przez niektórych inwestorów umiejętności pozwalających osiągać stopy zwrotu powyżej benchmarku.

Cremers i Petajisto [2009] pogłębili analizę funduszy aktywnych poprzez wprowadzenie miernika „Active Share”, reprezentującego udział portfela, który różni się od udziałów indeksu referencyjnego. W porównaniu do znanego już wówczas wskaźnika błędu śledzenia (*tracking error*), wskaźnik ten kładzie większy nacisk na selekcję akcji. Wykorzystanie nowego wskaźnika pozwala w sposób znaczący prognozować wyniki funduszu w stosunku do odpowiadającego mu benchmarku. Fundusze o najwyższym wskaźniku aktywnego udziału osiągają lepsze wyniki od swoich benchmarków zarówno przed, jak i po kosztach, podczas gdy fundusze o najniższym wskaźniku osiągają gorsze wyniki po kosztach. W przeciwieństwie do tego aktywne zarządzanie mierzone błędem odwzorowania nie przewiduje wyższych zysków. Zależność pomiędzy aktywnym udziałem a stopami zwrotu funduszy skorygowanymi o benchmark istnieje dla wszystkich wielkości funduszy, ale jest ona silniejsza dla mniejszych funduszy. Autorzy znajdują również silne dowody na trwałość wyników w przypadku funduszy o najwyższym aktywnym udziale, nawet po uwzględnieniu kontroli momentum. Fundusze o najwyższym „Active Share”, najmniejszej kapitalizacji i najlepszych wynikach jednorocznych w okresie analizy (1980–2003) osiągnęły wyniki lepsze od swoich benchmarków o 6,5% rocznie po odliczeniu opłat i kosztów.

W kolejnej próbie zracjonalizowania słabych przeciętnych wyników inwestowania aktywnego zastosowano podział na inwestorów detalicznych bezpośrednio inwestujących w fundusze, którzy podążają za zwrotami po opłatach skorygowanymi o ryzyko, oraz inwestorów korzystających ze wsparcia brokerów [Del Guercio, Reuter, 2014]. Autorzy wskazują, iż zachowanie pierwszej grupy inwestorów prowadzi do aktywniejszego zarządzania mierzonego wskaźnikiem „Active Share” i zwiększenia alfy, oraz że istnieje niewiele dowodów na to, że aktywnie zarządzane fundusze bezpośrednio osiągają wyniki gorsze od indeksów. Jako uzasadnienie faktu, iż pomimo słabych wyników praktycznie wszystkie aktywa w segmencie sprzedaży brokerskiej pozostają aktywnie zarządzane, autorzy proponują stwierdzenie, że inwestorzy ci nie są świadomi średnio gorszych wyników, oraz że istnieje konflikt

interesów pomiędzy brokerami a klientami. Choć podkreślają także, że biorąc pod uwagę dyskomfort odczuwany przez wielu inwestorów przy podejmowaniu decyzji finansowych, nie jest jasne, czy klienci byłiby w stanie lepiej inwestować bez swoich brokerów.

Podobnie jak wspomniani Crane i Crotty [2018], również Polkovnichenko, Wei i Zhao [2019] dokonali analizy rozkładów stóp zwrotów z konkurencyjnych strategii. W swym badaniu wprowadzają rozróżnienie strategii aktywnych na strategię „value” oraz „growth”. Dowodzą, że mimo iż średnie stopy zwrotu osiągnęte przez fundusze aktywne są niższe od benchmarków, to istnieje grupa inwestorów, dla których wybór strategii aktywnej jest racjonalny. Autorzy zauważyli bowiem, że fundusze aktywne typu „value” lepiej zabezpieczają ryzyko spadku, natomiast fundusze aktywne „growth” lepiej wychwytyją potencjał wzrostowy. Inwestorzy, których preferencje znajdują się w ogonach rozkładu stosunku do ryzyka, powinni więc przeważać strategie aktywne. Hipoteza ta posiada silne wsparcie empiryczne, m.in. przepływy do aktywnych funduszy „value” i „growth” są związane kolejno z zapotrzebowaniem inwestorów na ochronę przed spadkami i preferencjami w zakresie potencjału wzrostu. Wyniki te są silniejsze dla funduszy, które w przeszłości wykazywały pożądaną cechę oraz dla tych, które charakteryzują się wyższym udziałem aktywnego zarządzania. Porównanie przepływów do funduszy emerytalnych oraz nieemerytalnych także potwierdza tę hipotezę. Fundusze emerytalne wykazują silniejszą wrażliwość na preferencje dotyczące ochrony przed spadkami, natomiast mniejszą wrażliwość na preferencje dotyczące potencjału wzrostu.

Badanie z 2020 r. [Pedraza, Pulga, Vasquez, 2020] przedstawia jeszcze jeden argument potwierdzający, że wnioski uzyskane w badaniach rynku amerykańskiego, dotyczące przewagi pasywnego inwestowania, nie zawsze obowiązują. Autorzy zbadali wyniki inwestorów zagranicznych na kolumbijskim rynku akcji, dochodząc do wniosku, że łączne słabe wyniki inwestorów zagranicznych na tym rynku można w dużej mierze przypisać funduszom pasywnym, które płacą wyższe ceny za zakup akcji oraz otrzymują mniej, gdy je sprzedają. Zjawisko to nasila się w dniach, w których fundusze obracają wieloma akcjami w tym samym kierunku i dokonują dużych transakcji w pobliżu zamknięcia rynku, co prowadzi do gorszych wyników skorygowanych o ryzyko. Autorzy zaznaczają, że choć analiza dotyczy jednego kraju, to podkreśla koszty mechanicznych strategii handlowych na małych rynkach. Przewidują także, że podobne ustalenia są prawdopodobne w innych krajach rozwijających się lub w przypadku nie płynnych akcji na rynkach rozwiniętych. W takich przypadkach koszty inwestowania indeksowego mogą przewyższać jego korzyści.



## 6. Potencjał jednoczesnego wykorzystania obu strategii w celu poprawy osiągniętych wyników inwestycyjnych

Stopy zwrotu z benchmarku modelu wyceny z reguły obliczane są bez uwzględnienia jakichkolwiek kosztów związanych z wdrażaniem podstawowych strategii inwestycyjnych. Intuicyjnie i powszechnie jako najbliższe substytuty dla tych hipotetycznych benchmarków uznaje się fundusze indeksowe. Pastor i Stambaugh [2003] dowodzą jednak, że podczas gdy hipotetyczne stopy zwrotu z indeksów benchmarkowych i niebenchmarkowych uznane zostaną za niedostępne dla inwestycji, nie muszą istnieć dla nich bliskie substytuty w uniwersum funduszy aktywnych i pasywnych. Szacują, że inwestorzy w pełni wierzący w modele Famy-Frencha [1993] czy Carharta [1997] osiągają maksymalny wskaźnik Sharpe'a na poziomie 66% i 54% tego, co mogliby osiągnąć poprzez bezpośrednią inwestycję w benchmarki tych modeli. Autorzy badania pokazują, że inwestorzy pasywni, którzy nie dopuszczają możliwości istnienia umiejętności zarządzających, a poszukują wysokiego wskaźnika Sharpe'a, mogą preferować niektóre fundusze aktywne, ponieważ fundusze te mogą być lepszymi substytutami benchmarków niż istniejące fundusze pasywne.

Arytmetyka Sharpe'a implikuje, że skoro aktywni menadżerowie tracą po opłatach, inwestorzy powinni przesunąć pieniądze z funduszy aktywnych w kierunku pasywnych, jednakże również inwestorzy pasywni dokonują częstych transakcji w celu utrzymania portfeli ważonych rynkiem [Pedersen, 2018]. Ponadto, jak dowiedli Pedraza, Pulga, Vasquez [2020], inwestorzy mogą dokonywać tych transakcji po znacząco mniej korzystnych cenach. Pedersen podkreśla, że wzrost funduszy pasywnych, a więc mniej aktywnego zarządzania, oznacza mniej efektywne rynki, co ułatwia czerpanie zysków z aktywnego zarządzania. Pedersen wnioskuje, że odsetek inwestorów pasywnych nadal będzie rósł, ale wyniesie mniej niż 100%, jako że obie grupy oddziałują na siebie wzajemnie. Kolejny wniosek to stwierdzenie, że rynek kapitałowy jest grą o sumie dodatniej, gdzie spółki korzystają z dostępu do rynku, inwestorzy pasywni z taniego dostępu do zdywersyfikowanych funduszy, a inwestorzy aktywni ze swoich wysiłków w zakresie gromadzenia informacji poprzez potencjalnie wyższe wyniki. Aktywni zarządzający mogą być warci dodatknych opłat, gdyż pomagają w efektywnej alokacji zasobów, natomiast ważną rolę ekonomiczną zarządzających w sposób pasywny jest zapewnienie dostępu do rynków po niskich kosztach.

Aktywnie zarządzane fundusze inwestycyjne korzystają z funduszy ETF. Sherrill, Shirley i Stark [2020] podejmują próbę pomiaru potencjalnych korzyści wynikających z takich działań. Autorzy w oparciu o pracę Cremersa i Petajisto [2009] dokonują podziału funduszy ETF na dwie kategorie: 19 funduszy benchmarkowych

oraz fundusze niebenchmarkowe. Fundusze stosujące benchmarkowe fundusze ETF odnotowują zmniejszenie posiadanych środków pieniężnych, szczególnie w okresach dużych przepływów, oraz niższy błąd śledzenia. Gdy fundusze inwestycyjne utrzymują gotówkę, uzyskują stopę zwrotu wolną od ryzyka, co powoduje, że ich zwroty pozostają w tyle za benchmarkiem. Wykorzystanie funduszy ETF pozwala zmniejszyć wpływ stopy wolnej od ryzyka. Pozycja w funduszach ETF niebędących funduszami benchmarkowymi generuje unikalny zestaw zwrotów. Choć pozycja w ETF niebędącym wzorcowym nadal jest pasywną inwestycją, to ponieważ zyski te różnią się od wzorca, autorzy klasyfikują ją jako aktywną inwestycję. Dowodzą także, że pozycje te poprawiają wyniki dużych funduszy inwestycyjnych inwestujących w akcje o małej kapitalizacji, jednocześnie zmniejszając ryzyko portfela.

Poprzednie badanie wskazywało na wykorzystanie funduszy pasywnych w strategiach aktywnych, kolejne natomiast przedstawia wnioski, iż większość funduszy ETF to aktywne inwestycje w formie (przeznaczone do generowania alfy) lub funkcji (służą jako elementy budulcowe portfeli aktywnych) [Easley, Michayluk, O'Hara, Putnins, 2021]. W przypadku ETFów o wysokiej aktywności (mierzonej Activeness Index) występuje zdecydowana większość obrotów ETF. Autorzy dochodzą do wniosku, że ogólna aktywność inwestycyjna w przekroju różnych instrumentów nie zmieniła się znacząco w ostatnich 20 latach. ETFy aktywne w formie charakteryzują się dodatnią wrażliwością przepływów na wyniki, ponadto fundusze te pobierają najwyższe opłaty wśród tej kategorii funduszy oraz wyróżniają się wysokim obrotem w ramach portfela. Natomiast ETFy aktywne w funkcji mają bardziej skoncentrowane udziały, mniejszy obrót w ramach portfela, ale większy obrót na rynku wtórnym. Fundusze bardziej aktywne zyskują popularność względem mniej aktywnych, co prowadzi do wzrostu presji na obniżkę opłat w całej branży zarządzania inwestycjami. Rosnąca aktywność funduszy ETF może także łagodzić obawy o szkodliwość tychże funduszy w procesie ustalania cen i osiągnięcia efektywności rynków.

## **7. Wpływ rozwoju strategii pasywnych oraz towarzyszących temu trendowi zjawisk na aspekty rynku inne niż stopy zwrotu**

Analiza zmian zachodzących na rynkach akcyjnych w trakcie ostatnich trzech dekad, a przede wszystkim rozwój strategii pasywnych, nie ogranicza się jedynie do wnioskowania na temat stóp zwrotu. Badanie opublikowane w 2005 r. [Peress, 2005] weryfikuje, czy zmiany w kosztach uczestnictwa w giełdzie mogą wyjaśniać długoterminowy wzrost liczby akcjonariuszy na rynku amerykańskim. Choć odsetek amerykańskich gospodarstw domowych posiadających akcje wzrósł z 6%

w roku 1952 do 49% w 1998, autor rozpatruje także potencjalne przyczyny faktu, iż połowa gospodarstw nie posiada żadnych akcji. Koszty uczestnictwa podzielone zostały na dwie grupy: a) koszty pozyskiwania informacji rynkowych, b) koszty wejścia, a więc łączne opłaty, prowizje, czas wypełnienia formularzy podatkowych, otwierania rachunków czy zrozumienia zasad działania rynku akcyjnego. Wnioski z tejże publikacji są następujące: spadający koszt informacji nie może wyjaśnić obserwowanego wzrostu udziału w rynku akcji, natomiast spadający koszt wejścia może wyjaśniać wzrost liczby akcjonariuszy. Ponadto spadek kosztów wejścia najprawdopodobniej przyczynia się także do takich zjawisk, jak spadająca premia za inwestycje w akcje, rosnące wariacje stóp zwrotu oraz boom inwestowania pasywnego w stosunku do aktywnego.

Choć liczba gospodarstw domowych posiadających akcje rośnie, to w ciągu ostatnich kilku dekad część rynku akcji będąca w bezpośrednim posiadaniu osób fizycznych znacząco się zmniejszyła. Towarzyszącymi temu trendami są: wzrost znaczenia strategii indeksowych, przesunięcie aktywnych menedżerów w kierunku niższych opłat (średnie opłaty uiszczane na rzecz funduszy aktywnych spadły z 2,3% w 1980 r. do 1% w 2012 r.) oraz inwestowania bardziej zbliżonego do indeksów (spadek błędu śledzenia z 2,2% w 1980 r. do 1,2% w 2006 r. oraz aktywnego udziału z 38% do 23% w latach 1980–2011). Stambaugh [2014] osoby fizyczne bezpośrednio inwestujące na rynku akcyjnym określa mianem „handlarzy szumem”. Grupa ta powoduje odchylenia od stopnia optymalnego, określanego mianem efektywności cenowej rynków. Badanie dowodzi, że aktywne zarządzanie koryguje większość błędnych wycen. Ze względu na istnienie kosztów transakcyjnych i ryzyka idiosynkratycznego, część błędów pozostaje nieskorygowana, co pozwala aktywnym inwestorom osiągać dodatnią alfę. Przy mniejszej liczbie błędnych wycen, spowodowanej wpływem inwestorów indywidualnych do funduszy wzajemnych i indeksowych, aktywne zarządzanie musi charakteryzować się mniejszym zasięgiem i potencjałem do generowania nadzwyczajnych stóp zwrotu.

Cremers, Ferreira, Matos, Starks [2016] podejmują próbę powiązania indeksowania i aktywnego zarządzania w branży funduszy inwestycyjnych na świecie, poprzez testowanie hipotezy, że wzrost indeksowania i występowania tanich funduszy indeksowych na rynku wpływa na strukturę konkurencyjną rynków funduszy inwestycyjnych poprzez zmuszenie aktywnie zarządzanych funduszy do zwiększenia ich aktywnego udziału (tj. większego odchylenia od benchmarku) i obniżenia opłat. Dowody pozwalają autorom potwierdzić słuszność tej hipotezy. Na rynkach charakteryzujących się silniejszą konkurencją ze strony funduszy wyraźnie indeksowych, fundusze aktywne stosują bardziej zróżnicowane strategie, aby wygenerować dodatnią alfę oraz pobierać niższe opłaty. Natomiast w krajach, gdzie inwestorzy mają

ograniczone możliwości inwestowania w tanie fundusze indeksowe, wielu zarządzających aktywnymi funduszami osiąga niskie wartości „Active Share”, pobiera wyższe opłaty i osiąga gorsze wyniki. Podstawową implikacją wyników jest stwierdzenie, że rozwój funduszy pasywnych zwiększa konkurencję w branży zarządzania aktywami.

Znaczące przejście od aktywnych do pasywnych strategii inwestycyjnych, które nastąpiło zwłaszcza w okresie ostatnich dwóch dekad, powoduje wzrost zainteresowania badaczy, zarządzających aktywami, inwestorów oraz organów regulacyjnych, wpływem tychże zmian na stabilność finansową rynków. Anadu i in. [2020] analizują ten wpływ poprzez przełożenie na: a) płynność funduszy i ryzyko umorzenia, szczególnie w sektorach funduszy wzajemnych i ETF, b) zmienność rynku aktywów, c) koncentrację branży zarządzania aktywami oraz d) zmiany zwrotów z aktywów i płynności. Autorzy badania dochodzą do wniosku, że zmiana ta wpływa na część zagrożeń w sposób wzmacniający, a część w sposób osłabiający. Niektóre strategie pasywne zwiększają bowiem zmienność rynku, a rozwój strategii pasywnych wpływa także na zwiększenie koncentracji branży. Z drugiej strony trend ten zmniejsza pewne ryzyka płynności i umorzenia. Natomiast dowody co do powiązania indeksowania z ruchem zwrotów z aktywów i płynnością są mieszane.

Rozwój strategii pasywnych pozwolił także empirycznie potwierdzić część hipotez badawczych. Roussanov, Ruan i Wei [2021] badają wpływ wydatków w kategorii „marketing i dystrybucja” na wielkość funduszy. Badacze stwierdzają, że marketing jest prawie tak samo ważny, jak wyniki i opłaty, ponadto wydatki w tej kategorii stanowią około jednej trzeciej kosztów aktywnego zarządzania w branży funduszy inwestycyjnych. Autorzy opracowania dowodzą, iż wyeliminowanie marketingu znacząco poprawia dobrobyt inwestorów, kapitał przesuwa się w kierunku funduszy o niższych opłatach, w konsekwencji czego konkurencja obniża opłaty. Aktywa funduszy zarządzanych w sposób aktywny zmniejszają się, a alokacja kapitału jest ściślej dostosowana do umiejętności. Spadek kosztów poszukiwania inwestorów prowadzi do zmniejszenia wydatków marketingowych i opłat za zarządzanie, a także przejścia w kierunku inwestowania pasywnego.

## 8. Podsumowanie

W niniejszej pracy dokonano analizy alternatywnych strategii inwestycyjnych na rynku akcyjnym, jakimi są strategie aktywne oraz pasywne. Analiza wykonana została w formie przeglądu systematycznego, w oparciu o metodologię PRISMA. Wnioski pozwalają na udzielenie odpowiedzi na cztery pytania badawcze przedstawione we wstępie.

Reasumując, stanowisko, iż strategie aktywne osiągają wyniki gorsze od benchmarków oraz funduszy pasywnych, znajduje potwierdzenie w większości badań akademickich, co stanowi odpowiedź na pierwsze z pytań badawczych. Lata publikacji opracowań wspierających to stwierdzenie sugerują, iż rosnąca popularność strategii pasywnych jest konsekwencją wniosków z tychże badań. W dalszej części pracy podjęto próbę uzasadnienia istnienia strategii aktywnych przy założeniu prawdziwości wniosków dotyczących pierwszego pytania badawczego. Autorzy analizowanych badań dowodzą m.in., że skuteczność strategii aktywnych jest wyższa na rynkach słabiej rozwiniętych, o niższej płynności i efektywności niż rynek amerykański. Ponadto aktywni inwestorzy są skuteczniejsi na rynku funduszy REIT, w określonych fazach cyklu makroekonomicznego, czy wówczas gdy wskaźnik „Active Share” rośnie. Kolejne z pytań postawionych we wstępie dotyczyło możliwości połączenia obu strategii w celu zwiększenia skuteczności planów inwestycyjnych przyjmowanych przez zarządzających. Badania dowodzą, iż w rzeczywistości różniczenie nie jest binarne, strategie te nie są w pełni alternatywnymi, a inwestorzy często wykorzystują obie strategie równocześnie. Dotychczasowe badania w tym obszarze nie osiągają konsensusu co do możliwości uzyskania trwale wyższej stopy zwrotu w wyniku zastosowania kombinacji tych dwóch strategii. Jednakże podejście to pozwala zarządzającym przezwycięzać trudności związane z brakami dostatecznej płynności, oferty produktowej czy wysokimi kosztami na poszczególnych rynkach. Ostatnie z pytań badawczych dotyczyło analizy wpływu rozwoju strategii pasywnych na rynek funduszy inwestycyjnych w zakresie charakterystyk innych niż stopy zwrotu. Wśród wniosków z badań poddanych opracowaniu wskazać należy, że wzrost popularności strategii pasywnych przekłada się m.in. na zwiększanie się konkurencyjności, zmniejszanie się opłaty i kosztów wejścia na rynek, choć potencjalnie rośnie także zmienność oraz koncentracja branży.

Historia analiz skuteczności strategii aktywnych i pasywnych na rynkach akcyjnych sięga połowy XX w. Dorobek naukowy w analizowanym obszarze badawczym stanowi bogaty zestaw niezwykle istotnych opracowań, jednakże w przyszłości z perspektywy rynku polskiego szczególnie wartościowym byłaby weryfikacja wniosków z badania rynku kolumbijskiego. Poziom rozwoju polskiego rynku akcyjnego jest zbliżony do badanego rynku. Zasadnym byłaby więc weryfikacja wniosków mówiących o tym, że koszty inwestowania indeksowego na rynkach o niższej płynności mogą przewyższać korzyści z tej formy inwestycji.

Kolejnym obszarem o potencjalnej możliwości dalszej analizy byłoby rozszerzenie badania o zestaw transakcji z wielu rynków o różnym stopniu płynności i kapitalizacji. Pedraza, Pulga i Vasquez stawiają bowiem hipotezę, iż koszty strategii pasywnych mogą być znaczące w przypadku mniej płynnych rynków oraz papierów

wartościowych. Celem tychże analiz byłaby próba oszacowania zależności pomiędzy korzyściami wynikającymi ze strategii pasywnych oraz płynności i kapitalizacji rynków, a także poszczególnych papierów wartościowych. Wnioski z takiego badania mogłyby mieć przełomowe znaczenie dla kształtowania polityki inwestycyjnej globalnych funduszy inwestycyjnych, które w swoich strategiach łączą strategie aktywne i pasywne.

## Bibliografia

---

- Anadu, K., Kruttli, M., McCabe, P., Osambela E. (2020). The Shift from Active to Passive Investing: Risks to Financial Stability?, *Financial Analyst Journal*, 76(4), s. 23–39. DOI: 10.1080/0015198X.2020.1779498.
- Avramov, D., Wermers, R. (2006). Investing in Mutual Funds when Returns Are Predictable, *Journal of Financial Economics*, 81(2), s. 339–377. DOI: 10.1016/j.jfineco.2005.05.010.
- Baks, K.P., Metrick, A., Wachter, J. (2001). Should Investors Avoid All Actively Managed Mutual Funds? A Study in Bayesian Performance Evaluation, *Journal of Finance*, 56(1), s. 45–85. DOI: 10.1111/0022-1082.00319.
- Bogle, J.C. (2014). The Arithmetic of “All-In” Investment Expenses, *Financial Analyst Journal*, 70(1), s. 13–21. DOI: 10.2469/faj.v70.n1.1.
- Carhart, M.M. (1997). On Persistence in Mutual Fund Performance, *Journal of Finance*, 52(1), s. 57–82. DOI: 10.2307/2329556.
- Crane, A.D., Crotty, K. (2018). Passive versus Active Fund Performance: Do Index Funds Have Skill?, *Journal of Financial and Quantitative Analysis*, 53(1), s. 33–64. DOI: 10.1017/S0022109017000904.
- Cremers, K.J.M., Petajisto, A. (2009). How Active Is Your Fund Manager? A New Measure That Predicts Performance, *Review of Financial Studies*, 22(9), s. 3329–3365. DOI: 10.1093/rfs/hhp057.
- Cremers, M., Ferreira, M.A., Matos, P., Starks, L. (2016). Indexing and Active Fund Management: International Evidence, *Journal of Financial Economics*, 120(3), s. 539–560. DOI: 10.1016/j.jfineco.2016.02.008.
- Davis, J., Mengersen, K., Bennett, S., Mazerolle, L. (2014). Viewing Systematic Reviews and Meta-Analysis in Social Research through Different Lenses, *SpringerPlus*, 511(3). DOI: 10.1186/2193-1801-3-511.
- Del Guercio, D., Reuter, J. (2014). Mutual Fund Performance and the Incentive to Generate Alpha, *Journal of Finance*, 69(4), s. 1673–1704. DOI: 10.1111/jofi.12048.
- Easley, D., Michayluk, D., O’Hara, M., Putnins, T.J. (2021). The Active World of Passive Investing, *Review of Finance*, 25(5), s. 1433–1471. DOI: 10.1093/rof/rfab021.
- Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance*, 25(2), s. 383–423. DOI: 10.2307/2325486.

- Fama, E.F., French, K.R. (1993). Common Risk Factors in the Returns on Stock and Bonds, *Journal of Financial Economics*, 33(1), s. 3–56. DOI: 10.1016/0304-405X(93)90023-5.
- Fama, E.F., French, K.R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns, *Journal of Finance*, 65(5), s. 1915–1947. DOI: 10.1111/j.1540-6261.2010.01598.x.
- French, K.R. (2008). Presidential Address: The Cost of Active Investing, *Journal of Finance*, 63(4), s. 1537–1573. DOI: 10.1111/j.1540-6261.2008.01368.x.
- Gallagher, D.R., Harman, G., Schmidt, C.H., Warren, G.J. (2017). Global Equity Fund Performance: An Attribution Approach, *Financial Analyst Journal*, 73(1), s. 56–71. DOI: 10.2469/faj.v73.n1.1.
- Gerakos, J., Linnainmaa, J., Morse, A. (2021). Asset Managers: Institutional Performance and Factor Exposures, *Journal of Finance*, 76(4), s. 2035–2075. DOI: 10.1111/jofi.13026.
- Gruber, M.J. (1996). Another Puzzle: The Growth in Actively Managed Mutual Funds, *Journal of Finance*, 51(3), s. 783–810. DOI: 10.2307/2329222.
- Jensen, M.C. (1968). The Performance of Mutual Funds in the Period 1945–1964, *Journal of Finance*, 23(2), s. 389–416. DOI: 10.1111/j.1540-6261.1968.tb00815.x.
- Kallberg, J.G., Liu, C.L., Trzcinka, C. (2000). The Value Added from Investment Managers: An Examination of Funds of REITs, *Journal of Financial and Quantitative Analysis*, 35, s. 387–408. DOI: 10.2307/2676210.
- Leippold, M., Rueegg, R. (2020). How Rational and Competitive Is the Market for Mutual Funds?, *Review of Finance*, 24(3), s. 579–614. DOI: 10.1093/rof/rfz011.
- Malkiel, G.B. (1995). Returns from Investing in Equity Mutual Funds 1971 to 1991, *Journal of Finance*, 50(2), s. 549–572. DOI: 10.2307/2329419.
- Markowitz, H.M. (1952). Portfolio Selection, *Journal of Finance*, 7(1), s. 77–91. DOI: 10.1111/j.1540-6261.1952.tb01525.x.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement, *International Journal of Surgery*, 8(5), s. 336–341. DOI: 10.1371/journal.pmed.1000097.
- Pastor, L., Stambaugh, R.F. (2003). Investing in Equity Mutual Funds, *Journal of Financial Economics*, 63(3), s. 351–380. DOI: 10.1016/S0304-405X(02)00065-X.
- Pedersen, L.H. (2018). Sharpening the Arithmetic of Active Management, *Financial Analyst Journal*, 74(1), s. 21–36. DOI: 10.2469/faj.v74.n1.4.
- Pedraza, A., Pulga, F., Vasquez, J. (2020). Costly Index Investing in Foreign Markets, *Journal of Financial Markets*, 51. DOI: 10.1016/j.finmar.2019.10050.
- Peress, J. (2005). Information vs. Entry Costs: What Explains US Stock Market Evolution?, *Journal of Financial and Quantitative Analysis*, 40(3), s. 563–594. DOI: 10.1017/S0022109000001873.
- Polkovnichenko, V., Wei, K.D., Zhao, F. (2019). Cautious Risk Takers: Investor Preferences and Demand for Active Management, *Journal of Finance*, 74(2), s. 1025–1075. DOI: 10.1111/jofi.12747.
- Roussanov, N., Ruan, H.X., Wei, Y.H. (2021). Marketing Mutual Funds, *Review of Financial Studies*, 34(6), s. 3045–3094. DOI: 10.1093/rfs/hhaa095.

- Sharpe, W.F. (1991). The Arithmetic of Active Management, *Financial Analyst Journal*, 47(1). DOI: 10.2469/faj.v47.n1.7.
- Sherrill, D.E., Shirley, S.E., Stark, J.R. (2020). ETF Use among Actively Managed Mutual Fund Portfolios, *Journal of Financial Markets*, 51. DOI: 10.1016/j.finmar.2019.100529.
- Song, Y. (2020). The Mismatch between Mutual Fund Scale and Skill, *Journal of Finance*, 75(5), s. 2555–2589. DOI: 10.1111/jofi.12950.
- Stambaugh, R.F. (2014). Presidential Address: Investment Noise and Trends, *Journal of Finance*, 69(4), s. 1415–1453. DOI: 10.1111/jofi.12174.

## Active and Passive Investment Strategies on Equity Markets. A Systematic Literature Review

### Summary

The stock market is an important part of the financial market. Its capitalisation at the end of 2020 was just under 94 trillion US dollars. The aim of this paper is to analyse the research of optimal investment strategies in this important market. A significant share of both active and passive strategies in the stock market raises questions about the effectiveness of these alternative solutions. This study seeks answers to four research questions. Using the PRISMA methodology to review the research, 26 selected articles were analysed. Key findings: (I) active strategies achieve lower average rates of return than the benchmark, (II) this leads to an increased popularity of passive strategies, (III) active strategies exist, among other reasons, because they bring better results at certain points in the cycle, in certain sectors and strategies, (IV) investors make decisions based on other factors, not the rate of return. In addition, the authors argue that managers often use both strategies at the same time, and that the development of passive strategies also affects other aspects of the market than rates of return, such as industry concentration, market volatility and volume of fees.

**Keywords:** passive investing, active investing, capital allocation, market efficiency



# Status społeczno-ekonomiczny a wyniki w nauce w szkołach ponadpodstawowych – systematyczny przegląd literatury

Jan Kroszka

## Streszczenie

Status społeczno-ekonomiczny dzieci jest w literaturze często wymieniany jako jeden z kluczowych czynników kształtujących ich wyniki w nauce oraz szanse odniesienia sukcesu w szkole. Celem tego artykułu jest opis współczesnych badań poświęconych tej zależności wśród młodzieży w wieku nastoletnim w państwach wysoko rozwiniętych. Przy wykorzystaniu metody PRISMA wyodrębniono 17 prac badawczych, które następnie poddano systematycznej analizie. Podczas analizy zwrócono uwagę przede wszystkim na: sposób, w jaki badania edukacyjne operacjonalizują status społeczno-ekonomiczny; zmienne pośredniczące opisywaną zależność oraz implikacje dla polityk publicznych. Autorzy analizowanych prac wyróżniają liczne mechanizmy, które przyczyniają się do powstania różnic w wynikach w nauce oraz nierówności edukacyjnych. Wśród nich są te związane z kształtem systemu edukacji, politykami edukacyjnymi, szkołą oraz jej relacją z uczniem, rodzicami, a także samymi dziećmi. Pozwala to na przedstawienie propozycji działań dla różnych aktorów wewnątrz publicznego systemu edukacji – decydentów politycznych, dyrektorów szkół, nauczycieli oraz rodziców. Systematyczny przegląd literatury wskazuje również na rosnącą w ostatnich latach popularność tego zagadnienia w naukach społecznych.

**Słowa kluczowe:** status społeczno-ekonomiczny, wyniki edukacyjne, osiągnięcia edukacyjne, nierówności edukacyjne, PRISMA

## 1. Wstęp

Do najważniejszych celów publicznego systemu oświaty należy zapewnienie wszystkim obywatelom równych szans w zakresie edukacji. Realizacja tego zadania nabiera szczególnego znaczenia w sytuacji pogłębiających się nierówności społecznych we współczesnych społeczeństwach. Nierówny dostęp do edukacji może kształtować nierówności w późniejszych etapach życia – np. na rynku pracy [Pokropek, Borgonovi, Jakubowski, 2015, s. 16]. W związku z tym istotnym zagadnieniem badawczym staje się poszukiwanie strukturalnych przyczyn różnic w osiągnięciach edukacyjnych dzieci oraz młodzieży z różnych środowisk [OECD, 2019, s. 15].

Jednym z wiodących kierunków badań w tym zakresie staje się analiza zależności pomiędzy statusem społeczno-ekonomicznym gospodarstw domowych, z których pochodzą dzieci, a ich wynikami w nauce. Raport Organizacji Współpracy Gospodarczej i Rozwoju (OECD) wskazuje, że status społeczno-ekonomiczny jest najbardziej wiarygodnym predyktorem szkolnego sukcesu, odpowiadającym przeciętnie za 12% wariacji w wynikach międzynarodowych testów piętnastolatków [OECD, 2019, s. 50]. Podobne wnioski można znaleźć również w metaanalizach badających to zagadnienie [Sirin, 2005].

W ostatnich latach – m.in. w następstwie kolejnych publikacji OECD – stale rośnie zainteresowanie badaczy tematyką nierówności edukacyjnych powiązanych z pochodzeniem społecznym. Ekonomiści oraz socjologowie zwracają uwagę na rolę systemu edukacji w reprodukcji nierówności społecznych [Sadura, 2017, s. 78], a także opisują mechanizmy, które wpływają na kształtowanie się różnic w osiągnięciach edukacyjnych w zależności od pozycji społecznej. Rosnąca popularność tego tematu prowadzi do potrzeby powstania systematycznego przeglądu współczesnych badań opisujących zależność pomiędzy statusem społeczno-ekonomicznym a wynikami dzieci w wieku szkolnym. Celem poniższej pracy jest realizacja tego zadania.

### 1.1. Pytania badawcze

Podczas analizy artykułów poruszających temat nierówności edukacyjnych oraz zależności pomiędzy statusem społeczno-ekonomicznym a osiągnięciami edukacyjnymi dzieci poszukiwane będą odpowiedzi przede wszystkim na trzy pytania badawcze: *Pytanie 1.* W jaki sposób w badaniach edukacyjnych operacjonalizowany jest status społeczno-ekonomiczny?

*Pytanie 2.* Jakie zmienne pośredniczące w relacji pomiędzy statusem społeczno-ekonomicznym a osiągnięciami edukacyjnymi są wyróżniane przez badaczy?

*Pytanie 3.* Do jakich wniosków na temat wpływu zmiennych związanych z podłożem społecznym na wyniki w nauce dochodzą autorzy współczesnych badań?

*Pytanie 3a.* Jakie implikacje dla polityk publicznych wynikają z tych wniosków?

## 2. Metoda badawcza

W celu dokonania wyboru artykułów analizowanych w tej pracy skorzystano z procedury opisanej przez autorów metody PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*). PRISMA jest narzędziem, które ma na celu ułatwienie badaczom przeprowadzenia transparentnych oraz kompletnych meta-analiz i systematycznych przeglądów literatury. Ta procedura jest powszechnie stosowana przede wszystkim w naukach medycznych, jednak może być użyta również w badaniach społecznych oraz edukacyjnych [Page i in., 2021]. Autorzy przedstawiają listę elementów, które powinny zostać zawarte w prawidłowo przeprowadzonym systematycznym przeglądzie literatury<sup>1</sup>. Należą do nich m.in. dokładnie opisany proces wyszukiwania artykułów oraz szczegółowe kryteria ich odrzucenia lub uwzględnienia w dalszej analizie.

### 2.1. Strategia wyszukiwania tekstów

Do zidentyfikowania artykułów uwzględnionych w dalszej analizie wykorzystano elektroniczną anglojęzyczną bazę danych Web of Science (WoS). WoS należy do największych naukowych baz danych pozwalających badaczom na wyszukanie podstawowych informacji na temat ponad stu milionów artykułów naukowych. Procesu wyszukiwania artykułów w WoS dokonano w grudniu 2021 r. W jego trakcie autor poszukiwał artykułów poruszających jednocześnie dwa zagadnienia: 1) status społeczno-ekonomiczny (SSE; *Socio-Economic Status – SES*) oraz 2) wyniki edukacyjne. Dokładna fraza użyta do wyszukiwania tekstów została określona następująco: (“socio-economic status” OR “ses”) AND (“educational achievements” OR “educational performance” OR “education achievements” OR “education results” OR “educational results” OR “educational outcomes” OR “education outcomes”) NOT (“tertiary education” OR “higher education” OR “postsecondary education” OR “university”), gdzie spójniki AND, OR oraz NOT oznaczały odpowiednio koniunkcję, alternatywę oraz wykluczanie. Zidentyfikowane artykuły miały zawierać określone frazy w tytule, abstrakcie lub słowach kluczowych.

<sup>1</sup> Pełną listę można znaleźć na stronie internetowej: <http://www.prisma-statement.org/> (dostęp: 2.01.2022).

## 2.2. Kryteria wyboru oraz wykluczenia artykułów

Kluczowy element procedury PRISMA stanowi dokładne określenie kryteriów wyboru oraz wykluczenia (*inclusion and exclusion criteria*) artykułów, które zostaną poddane analizie w ramach systematycznego przeglądu literatury. W tej pracy najważniejszym kryterium wyboru jest przedstawienie zależności pomiędzy SSE dziecka (jego rodziny) a jego wynikami w nauce. Co ważne, opisana zależność dotyczy SSE na poziomie indywidualnym, a więc z analizy wykluczone są badania, które analizują różnice w osiągnięciach edukacyjnych jedynie na poziomie szkół lub regionów.

Kolejnym kryterium jest wiek badanych osób: przegląd literatury uwzględnia jedynie badania dotyczące opisywanej zależności na poziomie szkół ponadpodstawowych – wśród młodzieży w wieku nastoletnim. Wynika to z faktu wysokiego znaczenia osiągnięć edukacyjnych na tym etapie edukacji dla dalszego powodzenia w życiu [Pokropek i in., 2015] oraz dużej liczby badań edukacyjnych przeprowadzanych przez międzynarodowe organizacje wśród nastolatków<sup>2</sup>. Ze względu na oczekiwaną przynajmniej częściową porównywalność analizowanych systemów edukacyjnych, przegląd bierze pod uwagę jedynie badania prowadzone w krajach o wysokim poziomie rozwoju społeczno-ekonomicznego<sup>3</sup>. Analiza nie uwzględnia również badań kładących szczególny nacisk na sytuację migrantów lub mniejszości etnicznych w systemie edukacji [zob. Ledwith, Reilly, 2014; Kim, Mok, Seidel, 2020] – jako przyczynę autor podaje w tym przypadku występowanie potencjalnie innych czynników kształtujących relację między SSE a osiągnięciami edukacyjnymi.

Ostatecznie w przeglądzie uwzględnione zostały jedynie artykuły naukowe z ostatnich dwóch dekad (od 2001 r.) – wynika to z zamiaru opisanego współczesnego stanu wiedzy na ten temat, jak również z faktu istnienia wyczerpujących metaanaliz syntetyzujących wcześniejsze badania [zob. White, 1982; Sirin, 2005]. Wykluczone zostają również artykuły napisane w języku innym niż angielski.

## 2.3. Proces wyboru artykułów

Proces wyboru artykułów do dalszej analizy składa się z trzech etapów. Podczas pierwszego – identyfikacji (*identification*) – baza Web of Science została przeszukana przy wykorzystaniu podanej wyżej frazy w celu wyodrębnienia odpowiednich artykułów. Następnie, w trakcie fazy wstępnej analizy (*screening*) autor przeczytał

<sup>2</sup> Największym z nich jest PISA – *Programme for International Student Assessment*. Dokładny opis tego badania znajduje się w części artykułu poświęconej charakterystyce analizowanych prac.

<sup>3</sup> Do tej grupy autor kwalifikuje państwa bardzo wysoko rozwinięte według Wskaźnika Rozwoju Społecznego (*Human Development Index, HDI*) – takie, dla których HDI jest większe niż 0,8 (UNDP, 2020).

abstrakty zidentyfikowanych artykułów, sprawdzając, czy badają one analizowane zjawisko oraz czy spełniają określone kryteria. W ostatniej fazie – ocenie kwalifikowalności (*eligibility*) – wybrane artykuły zostały przeczytane w całości i ostatecznie ocenione pod kątem kwalifikowalności do systematycznego przeglądu literatury (w tej fazie wykluczone zostały artykuły niezawierające wystarczających szczegółów oraz napisane w języku innym niż angielski).

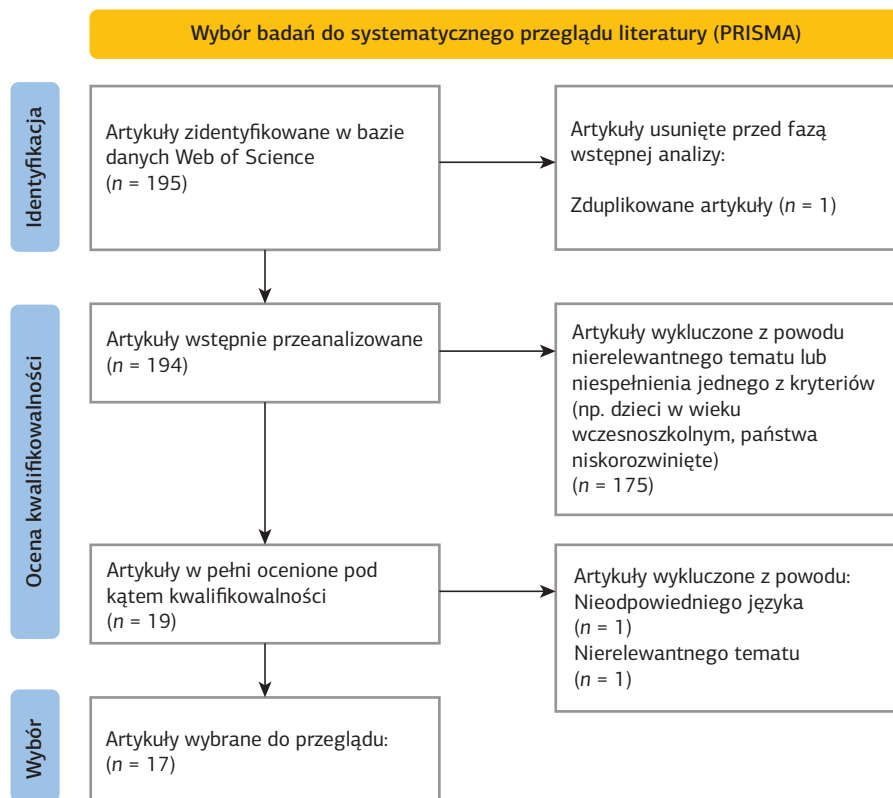
## 2.4. Ograniczenia metody

PRISMA jest narzędziem pozwalającym na systematyczny oraz uporządkowany przegląd literatury, jednak jej zastosowanie wiąże się z pewnymi ograniczeniami. Po pierwsze, użycie tej procedury może spowodować nieuwzględnienie w analizie pewnych istotnych badań, które w abstrakcie lub słowach kluczowych zawierają jedynie synonimy słów wybranych jako kluczowe (np. zamiast „status społeczno-ekonomiczny” – „kapitał społeczny i ekonomiczny”). Ponadto do wyszukania artykułów użyto jedynie jednej bazy danych (Web of Science), co może również stanowić pewne ograniczenie liczby znalezionych badań. Dodatkowo elektroniczne bazy danych, takie jak Web of Science, charakteryzują się nadreprezentacją czasopism naukowych z wybranych państw [Mongeon, Paul-Hus, 2016], co może przekładać się na mniejsze zróżnicowanie regionalne analizowanych badań.

## 3. Wyniki przeglądu literatury o wpływie statusu społeczno-ekonomicznego na osiągnięcia w nauce

Podczas pierwszego etapu wyboru badań do dalszej analizy zidentyfikowano w bazie danych Web of Science 195 artykułów zawierających wyszukiwane frazy. Po usunięciu zduplikowanych prac badawczych pozostały 194 artykuły, których streszczenia zostały przeanalizowane. Spośród nich 175 artykułów zostało wykluczonych z dalszej analizy. Duża część z tej grupy nie odpowiadała poruszanemu w tej pracy tematowi badawczemu, a inne zostały odrzucone z powodu niespełnienia jednego z kryteriów wyboru oraz wykluczenia (np. przedmiotem badania były dzieci w wieku wczesnoszkolnym albo z krajów niskorozwiniętych). W kolejnej fazie procedury PRISMA oceniono 19 wybranych artykułów pod kątem kwalifikowalności. Dwa z nich zostały wykluczone z powodu nieodpowiedniego języka lub analizowanych zmiennych. Pozostałe 17 artykułów wybrano do systematycznego przeglądu literatury. Pełna procedura wyboru prac naukowych została przedstawiona na rysunku 1.

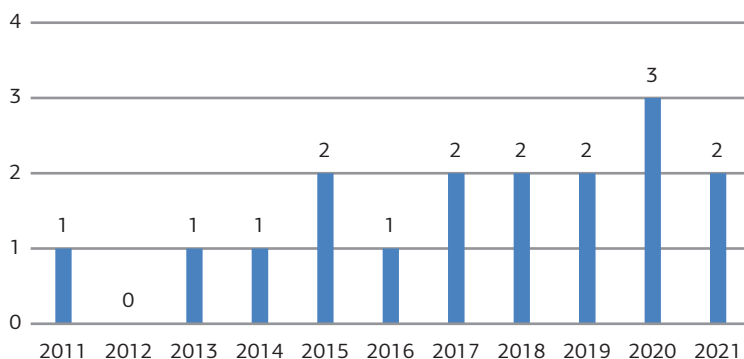
**Rysunek 1.** Wybór badań do systematycznego przeglądu literatury (metodą PRISMA)



Źródło: opracowanie własne na podstawie Page i in. [2021].

### 3.1. Charakterystyka wybranych badań

Rysunek 2 przedstawia rozkład badań wybranych do analizy według roku ich publikacji. Mimo że pod uwagę brane były prace naukowe opublikowane od początku XXI w., w trakcie procedury PRISMA odnaleziono odpowiednie artykuły jedynie z ostatnich dziesięciu lat. Dodatkowo ponad połowa finalnie analizowanych artykułów została opublikowana po 2017 r. Może to świadczyć o tym, że zagadnienie badania zależności pomiędzy statusem społeczno-ekonomicznym a osiągnięciami edukacyjnymi w ostatnich latach staje się coraz popularniejsze wśród badaczy. Przyczyną tego stanu rzeczy mogą być ostatnie raporty OECD [2016, 2019] podkreślające znaczenie tego zjawiska.

**Rysunek 2.** Rozkład analizowanych artykułów według roku ich publikacji

Źródło: opracowanie własne.

Serwis Web of Science pozwala na pogrupowanie wyszukanych artykułów według dyscyplin badawczych, do których należą. Zgodnie z tym zestawieniem ponad połowa wybranych prac zalicza się do badań edukacyjnych (*education & educational research*). Warto jednak zauważyć, że aż pięć wyodrębnionych artykułów zostało przyporządkowanych przez autorów do psychologii. Wśród analizowanych artykułów znajdują się również te należące do nauk społecznych, ekonomii oraz pediatrii, a dwa zostały określone jako międzydyscyplinarne (*interdisciplinary/multidisciplinary*). Powyższa analiza kategorii wyodrębnionych artykułów wskazuje wyraźnie na interdyscyplinarność badanego zjawiska.

Warto zwrócić uwagę na fakt, że zdecydowana większość wybranych prac naukowych opiera się na analizie danych empirycznych. Jedyne wyjątek stanowi w tym przypadku badanie Shanks i Robinson [2013], które przedstawia teoretyczny model zależności pomiędzy ekonomicznymi oraz kulturowymi zasobami a osiągnięciami edukacyjnymi uczniów. Ta praca nie precyzuje wieku dzieci, jednak zarysowany w niej model powiązania zmiennych wydaje się wysoce adekwatny dla analizowanego zagadnienia.

Tabela 1 (w załączniku) przedstawia inne charakterystyki analizowanych artykułów: m.in. źródło danych, a także wiek oraz kraj badanych uczniów. Ponad jedna trzecia rozpatrywanych prac korzystała z danych z badania PISA (*Programme for International Student Assessment*). PISA jest przeprowadzanym przez Organizację Współpracy Gospodarczej i Rozwoju (OECD) cyklicznym badaniem, którego celem jest zmierzenie wiedzy oraz umiejętności 15-latków w zakresie czytania, matematyki oraz nauk przyrodniczych. Ostatnia edycja badania odbyła się w 2018 r., jednak analizowane artykuły bazują na danych z jego poprzednich edycji. W 2012 r. w badaniu wzięło udział ponad 60 państw, w tym wszystkie kraje członkowskie OECD. Zdecydowana większość uczestników badania należała do państw wysoko rozwiniętych.

Drugie popularne źródło danych na temat osiągnięć edukacyjnych uczniów stanowi przeprowadzane cyklicznie od 1995 r. badanie TIMSS (*Trends in International Mathematics and Science Study*). W przeciwieństwie do badania PISA, TIMSS jest badaniem częściowo wzdłużnym, a więc jego grupa respondentów w kolejnych badaniach częściowo się powtarza. Pozwala to na dokonywanie analiz porównawczych w szerszym wymiarze czasowym [Broer, Bai, Fonseca, 2019]. W badaniu biorą udział uczniowie czwartej oraz ósmej klasy – wyodrębnione prace korzystają z danych dotyczących tych drugich, czyli dzieci mających przeciętnie 13 lat. TIMSS jest przeprowadzane głównie w krajach o wysokim stopniu rozwoju, jednak ich grupa w kolejnych edycjach się zmienia – jedynie 13 z nich uczestniczyło w co najmniej pięciu spośród sześciu edycji w latach 1995–2015 [Broer i in., 2019].

Badacze nierówności edukacyjnych korzystają również z prowadzonego w Australii badania LSAC (*Longitudinal Study of Australian Children*), a także stanowiącego jego część programu NAPLAN (*National Assessment Program – Literacy and Numeracy*). Jego zaletę stanowi fakt, że jest to badanie kohortowe, a więc pozwala śledzić losy uczestników w kolejnych latach. Wybrane na potrzeby tego przeglądu literatury badania analizują wyniki edukacyjne uczniów podczas piątej fali badania – kiedy mieli 12–13 lat.

Innymi źródłami danych wykorzystanymi w analizowanych artykułach były zbliżone wzdłużne badania przeprowadzane w Stanach Zjednoczonych (ELS – *Education Longitudinal Study*, NLSAH – *National Longitudinal Study of Adolescent Health*) oraz Norwegii (spisy powszechnie). W ich ramach analizowano wyniki edukacyjne uczniów z klas 7–12, a więc w wieku 12–17 lat. Wyjątek w tym zakresie stanowi badanie Frønesa [2016], w którym osiągnięcia edukacyjne uczniów szkół ponadpodstawowych są mierzone przez ich oceny na zakończenie tego poziomu edukacji.

#### 4. Dyskusja wyników przeglądu literatury

W części pracy poświęconej dyskusji wyników systematycznego przeglądu literatury główny nacisk zostanie położony na trzy pytania badawcze – 1) dotyczące operacjonalizacji SSE, 2) zmiennych pośredniczących oraz 3) najważniejszych wniosków badaczy.



#### 4.1. Operacjonalizacja statusu społeczno-ekonomicznego

Pomimo faktu że status społeczno-ekonomiczny (SSE) jest bardzo często wykorzystywanym konstruktem w licznych badaniach z zakresu nauk społecznych (również w badaniach edukacyjnych), trudno odnaleźć w literaturze jedną, spójną oraz szczegółową definicję tego konceptu, która pokrywałaby założenia większości badaczy. Najogólniej można go określić jako pozycję społeczną oraz zasoby (materialne, społeczne, kulturowe), którymi dysponuje jednostka lub gospodarstwo domowe. Jednak już metaanaliza White'a [1982, s. 27] pokazała, że do operacjonalizacji tego konceptu wykorzystywane są bardzo różne zmienne – w analizowanych przez niego 143 badaniach autor odnalazł ponad 70 zmiennych mających pełnić rolę wskaźników SSE.

Określenie sposobu mierzenia SSE jest szczególnie istotne, w związku z faktem, że jego nieodpowiednia operacjonalizacja może prowadzić do przeszacowania (lub odwrotnie – niedoszacowania) związku między pozycją społeczną uczniów a ich osiągnięciami edukacyjnymi [White, 1982, s. 27; O'Connell, 2019].

Tabela 1 przedstawia w jaki sposób analizowane w tej pracy badania operacjonalizowały status społeczno-ekonomiczny. Najczęściej występującymi wskaźnikami SSE były: poziom wykształcenia rodziców (został uwzględniony we wszystkich siedemnastu badaniach), wykonywany zawód przez rodziców (dziewięć badań) oraz dochody gospodarstwa domowego (siedem badań). Stosunkowo często pojawiała się również zmienna mierząca kapitał kulturowy rodziny – liczba książek w domu (pięć badań). Autorzy do pomiaru SSE stosowali również zmienne mierzące materialne ułatwienia w nauce, takie jak posiadanie przez dziecko komputera, biurka, połączenia internetowego lub własnego pokoju do nauki (łącznie pięć badań) oraz finansowa możliwość uczestnictwa w dodatkowych zajęciach (jedno badanie). Warto zwrócić uwagę na fakt, że wszystkie powyższe wskaźniki są zmiennymi obiektywnymi – jedynie Henry, Cavanagh i Oetting [2011] użyli w tym celu również zmiennej mierzącej subiektywną (odczuwaną przez badanych) sytuację materialną ich gospodarstwa domowego.

Zdecydowana większość analizowanych badań traktuje SSE jako zmienną ukrytą (*latent variable*), a więc na potrzeby analizy łączy ze sobą różne aspekty tego konceptu. Innymi słowy, traktuje SSE jako połączenie wymienionych powyżej zmiennych. Doskonały przykład stanowi tutaj przygotowywany przez badanie PISA Wskaźnik Ekonomicznego, Społecznego i Kulturowego Statusu (ESCS), który zostaje bezpośrednio użyty w dwóch badaniach, a osiem innych prac wykorzystuje zbliżoną do niego metodę. ESCS powstaje przez zestawienie poziomu wykształcenia rodziców, ich zawodu oraz posiadanych przez gospodarstwo domowe zasobów (takich jak

komputer lub biurko do nauki), a także zastosowanie metody statystycznej – analizy głównych składowych (*principal component analysis*) [OECD, 2014].

Warto jednak podkreślić również inną perspektywę badawczą, którą proponuje O’Connell (2019). Badacz sugeruje, że zastosowana przez większość analizowanych prac metoda operacjonalizacji statusu społeczno-ekonomicznego może być błędna. Wskazuje na niewystarczająco wysoką korelację pomiędzy zmiennymi wchodzącymi w skład ESCS lub podobnych wskaźników mierzących SSE. W związku z tym jego zdaniem korzystanie z takich zagregowanych wskaźników może prowadzić do nieodpowiednich wniosków w kontekście siły wpływu SSE na osiągnięcia edukacyjne. Autor proponuje zastąpienie konceptu SSE przez użycie w modelu pojedynczych, łatwo mierzalnych zmiennych (takich jak np. dochód gospodarstwa domowego).

Na koniec można zwrócić uwagę na inny aspekt operacjonalizacji SSE – czas. Zdecydowana większość analizowanych prac traktuje opisywane powyżej zmienne w sposób statyczny – mierzy je w momencie przeprowadzania badania. Jednakże badania wzdłużne (*longitudinal studies*) pozwalają na pomiar SSE badanych w dłuższym okresie. Tę możliwość wykorzystują Frønes [2016], który wśród zmiennych niezależnych w modelu umieszcza m.in. wiek matki w momencie urodzenia pierwszego dziecka<sup>4</sup> oraz liczbę przepracowanych lat przez rodziców dziecka w wieku od 25 do 40 lat<sup>5</sup>, a także van Zwieten i in. [2021] – w ich badaniu wskaźnik mierzący SSE agreguje wykształcenie, zawód oraz dochód rodziców w czterech kolejnych fazach badania (kiedy dzieci miały od 4–5 do 10–11 lat).

## 4.2. Zmienne pośredniczące

Jednym z najważniejszych celów badań edukacyjnych poruszających temat zależności między SSE a wynikami w nauce jest znalezienie i opisanie mechanizmów kształtujących tę relację. Do często wykorzystywanych przez badaczy w tym celu narzędzi należy analiza ścieżek (*path analysis*), która wpływ SSE dzieli na bezpośredni oraz zapośredniczony przez inne zmienne [Harwell, Zhao, 2021; Pokropek i in., 2015]. Taka analiza pozwala na wyodrębnienie zmiennych pośredniczących opisywaną zależność oraz zaproponowanie adekwatnych implikacji dla aktorów wewnątrz systemu edukacji (nauczycieli, rodziców, politycznych decydentów).

Międzynarodowe badania porównawcze sprzyjają czynieniu takich propozycji na poziomie polityki publicznej. Autorzy analizowanych badań sugerują, że to,

<sup>4</sup> Autor sugeruje, że posiadanie dziecka w wieku nastoletnim w wysoko rozwiniętym kraju (Norwegia) wiąże się z niskim statusem społeczno-ekonomicznym.

<sup>5</sup> W przypadku matek – w wieku od 25 do 35 lat. Założeniem autora jest, że rodzice osób z rodzin o niskim statusie społeczno-ekonomicznym częściej w tym okresie nie pracowali.

w jaki sposób system publicznej edukacji jest zbudowany, stanowi mediator w relacji pomiędzy SSE uczniów a ich osiągnięciami w nauce [Rowley, Edmunds, Dufur, Jarvis, Silveira, 2020; Ferraro, Pöder, 2018; Broer i in., 2019; Gustafsson, Nilsen, Hansen, 2018]. Broer i in. [2019] proponują dwa wymiary, na podstawie których można analizować systemy edukacji: stopień ich heterogeniczności oraz centralizacji. Pierwszy z nich odnosi się do różnic finansowych pomiędzy szkołami wewnątrz danego systemu edukacji, drugi – do poziomu autonomii pojedynczych szkół. Ferraro i Pöder [2018], bazując na danych z kwestionariuszy wypełnianych przez włoskie szkoły w badaniu PISA, analizują wpływ polityk edukacyjnych na poziomie mikro (*school-level policies* – np. liczba konkurencyjnych szkół w bezpośredniej okolicy szkoły lub podział uczniów na klasy według umiejętności) na relację pomiędzy SSE a wynikami uczniów.

Znaczenie szkoły w kształtowaniu poziomu nierówności edukacyjnych jest podkreślane również w innych badaniach. Gustafsson i in. [2018] wśród zmiennych pośredniczących zależność pomiędzy SSE nastolatków a ich osiągnięciami w nauce wymieniają ilość oraz subiektywną (ocenianą przez uczniów) jakość nauczania w szkole, klimat nacisku na sukces w szkole (*School Emphasis on Academic Success*), klimat przestrzegania zasad w szkole (*Safe and Orderly Climate*) oraz przeciętny poziom SSE wewnątrz szkoły. Autorzy sugerują, że niskie wartości tych zmiennych (np. niska jakość nauczania) mogą pogłębiać różnice w wynikach pomiędzy uczniami z rodzin o różnym statusie społeczno-ekonomicznym<sup>6</sup>. Takashiro [2017], opierając się na bioekologicznej teorii Bronfenbrennera [Bronfenbrenner, Morris, 1998], również uwzględnia w swoim modelu SSE szkoły jako zmienną pośredniczącą. Jego analiza nie potwierdza jednak hipotezy o mediującym wpływie tej zmiennej.

Zbliżone podejście do wspomnianych powyżej prac prezentują Schmidt, Burroughs, Zoido i Houang [2015], którzy przewidują, że analizowana zależność jest pośredniczona przez możliwość nauki (*Opportunity to Learn*, OTL). OTL jest mierzonym w ramach badania PISA konstruktem, który oznacza świadomy kontakt uczniów z formalnymi metodami matematycznymi. Autorzy tego badania sugerują, że słabsze wyniki w matematycznych testach młodzieży o niższym statusie społeczno-ekonomicznym wynikają z różnic w poznanym przez nich w szkole materiale.

Kolejnym aspektem wymienianym w analizowanej literaturze jako potencjalnie pośredniczący wpływ statusu społeczno-ekonomicznego ucznia na jego wyniki

<sup>6</sup> Warto w tym miejscu zaznaczyć, że przedmiotem analizy nie jest w tym przypadku wpływ SSE szkoły na osiągnięcia edukacyjne, a znaczenie SSE szkoły jako mediatora w relacji pomiędzy indywidualnym SSE a wynikami w nauce. Innymi słowy, interesującym nas wnioskiem nie jest jedynie to, że uczniowie w szkołach o wysokim przeciętnym poziomie SSE osiągają lepsze wyniki niż ci chodzący do szkół o jego niższym poziomie, ale również to, że w szkołach o przeciętnie niskim SSE wpływ pozycji społecznej na wyniki w nauce jest większy.

w nauce jest relacja szkoła – uczeń. Benner i Wang [2014] w roli mediatora w swoim modelu umieszczają poczucie samotności w szkole oraz poczucie emocjonalnego związku z nią. Z kolei Harwell i Zhao [2021] jako zmienne pośredniczące traktują wsparcie nauczyciela oraz relację uczeń – nauczyciel.

Na bezpośrednie znaczenie rodziców w kształtowaniu zależności między pozycją społeczną uczniów a ich wynikami w nauce wskazują Henry i in. [2011]. W swoim modelu powiązania zmiennych jako mediatory umieszczają zmienne mierzące troskę rodziców o edukację dzieci (*parent concerned*) oraz zaangażowanie rodziców w szkołę ich dzieci (*parent involved*). Pierwsza grupa zmiennych obejmuje subiektywne (oceniany przez dzieci) poziom potencjalnego zmartwienia rodziców z powodu niepowodzenia dziecka w szkole (np. otrzymania słabej oceny albo wagarowania), a druga – uczestnictwo rodziców w zebraniach klasowych oraz innych wydarzeniach szkolnych. Również Shanks i Robinson [2013] podkreślają znaczenie rodziców w odpowiedzi na to pytanie badawcze, jednak w sposób bardziej pośredni – w przedstawionym przez nich modelu zmienną pośredniczącą jest poziom stresu, na jaki narażone jest dziecko. Na tą zmienną wpływają – powiązane ze statusem społeczno-ekonomicznym – czynniki takie jak zaangażowanie rodziców, ich zachowanie oraz deprywacja materialna. Zgodnie z hipotezą autorów działają w tym przypadku dwa efekty – efekt rodzinnego stresu, w ramach którego rodziny o niższym poziomie SSE są narażone na większy stres, oraz efekt rodzinnej inwestycji, który zakłada, że rodziny o wyższym poziomie SSE więcej inwestują w rozwój swoich dzieci<sup>7</sup>.

Ostatnią grupę zmiennych, które są wyróżniane przez badaczy jako mediujące zależność pomiędzy pozycją społeczną dzieci a ich osiągnięciami w nauce, stanowią te związane z samymi dziećmi – ich poziom zaangażowania oraz motywacji. Tomaszewski, Xiang i Western (2020) korzystają w tym celu z konstruktów zaangażowania ucznia, na który składają się zmienne mierzące zaangażowanie behawioralne (np. czytanie w domu lub wagarowanie), zaangażowanie afektywne (np. poczucie emocjonalnego związku ze szkołą, zainteresowanie szkołą) oraz zaangażowanie kognitywne (np. nastawienie na osiągnięcie sukcesu w szkole). Z kolei Rogelberg, Starrett, Irvin i DiStefano [2021] jako zmienne pośredniczące analizują wiarę we własne możliwości, motywację do nauki oraz zainteresowanie nauką.

Warto jeszcze zwrócić uwagę na fakt, że niektóre z analizowanych prac w roli mediatorów umieszczają zmienne, które przez innych badaczy zostały uznane za mierniki statusu społeczno-ekonomicznego – np. dostęp do komputera do nauki w domu [Pokropek i in., 2015; Harwell, Zhao, 2021]. Potwierdza to wniosek O’Con-

<sup>7</sup> Poprzez inwestycje rozumie się w tym przypadku nie tylko te finansowe, ale również takie związane z ludzkim, kulturowym oraz społecznym kapitałem. Np. inwestycją w rozwój dzieci jest pomoc rodziców w odrobieniu lekcji lub kupienie dzieciom książki zamiast zabawki.

nella [2019] mówiący o tym, że przy analizie oraz porównywaniu badań poświęconych nierównościom edukacyjnym kluczowe znaczenie ma zwrócenie uwagi na sposób operacjonalizacji statusu społeczno-ekonomicznego w tych badaniach.

### 4.3. Konkluzje oraz implikacje dla polityk publicznych

Na początek warto podkreślić, w jaki sposób analizowane badania rozumieją „osiągnięcia edukacyjne”, a więc do czego konkretnie odnosić się będą poniższe wnioski. W tabeli 1 przedstawiono sposób operacjonalizacji tego konstruktu – wybór zmiennej zależnej w analizowanym modelu. Większość badań opiera się na wynikach międzynarodowych standaryzowanych testów mierzących umiejętności czytania, znajomość matematyki oraz nauk przyrodniczych (PISA, TIMSS). Inne podejście prezentują Benner i Wang [2014] oraz Rogelberg i in. [2021], którzy wykorzystują bardziej subiektywny (zależny od szkoły) miernik osiągnięć edukacyjnych – oceny uczniów. Interesująca – szczególnie w kontekście rozpatrywania wykształcenia jako kluczowej determinanty późniejszej pozycji społecznej [Pokropek i in., 2015] – jest również analiza modelu z prawdopodobieństwem przedwczesnego zakończenia edukacji (*dropout*) jako zmienną zależną [Henry i in., 2011; Frønes, 2016].

Zdecydowana większość autorów analizowanych prac dochodzi do wniosku, że status społeczno-ekonomiczny dzieci należy do najważniejszych czynników kształtujących ich osiągnięcia w nauce [zob. Tomaszewski i in., 2020]. Istotne są zarówno bezpośrednie efekty pozycji społecznej gospodarstwa domowego, jak i efekty zapośredniczone przez opisane wcześniej zmienne – te drugie są często silniejsze [zob. Pokropek i in., 2015; Harwell, Zhao, 2021]. Wyjątek stanowią w tym przypadku badania Marksa [2017] oraz O’Connella [2019]. Ich autorzy dowodzą, że wpływ SSE na osiągnięcia w nauce jest jedynie przeciętny, a co za tym idzie – często przeszacowywany.

Dzięki wykorzystaniu międzynarodowych baz danych autorzy mogą porównać ze sobą stopień analizowanego zjawiska (wpływu SSE na wyniki w nauce) w różnych państwach. Gustafsson i in. [2018, s. 27] konkludują, że w niektórych krajach systemy edukacji działają w sposób kompensacyjny (wspierają dzieci z rodzin o niższej pozycji społecznej), a w innych – antykompensacyjny (pogłębiają nierówności w edukacji). Broer i in. [2019, s. 77] zauważają, że zdecentralizowane systemy edukacji przeciętnie charakteryzują się tendencją malejącą w poziomie nierówności<sup>8</sup>. Interesujący wniosek stanowi również obserwacja, że kompensacyjny charakter

<sup>8</sup> Oznacza to, że w takich systemach edukacji zależność pomiędzy statusem społeczno-ekonomicznym a wynikami w nauce stopniowo maleje.

systemów edukacji łączy się z przeciętnie wyższymi wynikami uczniów [Gustafsson i in., 2018; Broer i in., 2019]. Innymi słowy, te systemy edukacji, w których poziom nierówności jest niższy, osiągają przeciętnie wyższe wyniki w międzynarodowych testach (biorąc pod uwagę wszystkich uczniów).

Połączenie dwóch powyższych wniosków – dużego znaczenia statusu społeczno-ekonomicznego jako predyktora osiągnięć w nauce oraz międzynarodowych różnic w stopniu opisywanej zależności – prowadzi do pytania o przykłady polityk publicznych mogących kształtować tę zależność. Van Zwieten i in. [2021] piszą o potrzebie bezpośrednich interwencji wspierających dzieci z rodzin o niższym poziomie SSE. Potencjalnym polem takiej interwencji może być np. ułatwianie takim dzieciom integracji z ich szkołą [Benner, Wang, 2014]. Innymi rozwiązaniami mogą być wpływanie na sposób oceniania oraz organizacji lekcji [Tomaszewski i in., 2020] lub zwiększanie kompetencji nauczycieli oraz organizowanie dodatkowych zajęć [Rogelberg i in., 2021]. Pokropek i in. [2015, s. 16] zwracają uwagę na fakt, że takie interwencje mogą być szczególnie efektywne w zmniejszaniu pośrednich efektów statusu społeczno-ekonomicznego. Do istotnych wniosków prowadzi również analiza Ferraro i Pöder [2018], która pokazuje, że wysoka konkurencja pomiędzy szkołami zwiększa nierówności edukacyjne (efekt antykompensacyjny), jednocześnie nie poprawiając przeciętnych wyników uczniów, z kolei zwiększenie autonomii szkół może prowadzić do poprawy wyników bez powiększania nierówności.

## 5. Podsumowanie

Systematyczna analiza artykułów naukowych na temat zależności pomiędzy statusem społeczno-ekonomicznym a wynikami w nauce wśród nastoletniej młodzieży może prowadzić do wielu interesujących wniosków. Zauważalna jest rosnąca popularność tego zagadnienia w badaniach edukacyjnych, która odpowiada szerszym tendencjom w naukach społecznych.

Pierwsze pytanie badawcze postawione w tekście dotyczyło sposobu operacjonalizacji statusu społeczno-ekonomicznego w badaniach edukacyjnych. Na podstawie analizy udało się ustalić, że najczęściej wykorzystywanymi w tym celu zmiennymi są: sytuacja materialna gospodarstwa domowego, a także poziom wykształcenia oraz sytuacja zawodowa rodziców. Autorzy analizowanych prac często korzystają również z wieloaspektowych wskaźników, które agregują wymienione powyżej wskaźniki cząstkowe. Warto jednak zwrócić uwagę na to, że takie podejście jest krytykowane przez niektórych badaczy.

Do najważniejszych wniosków wynikających z analizowanych artykułów należy również wskazanie oraz opisanie ścieżek, za pośrednictwem których status społeczno-ekonomiczny może oddziaływać na szansę osiągnięcia sukcesu w szkole, co stanowi odpowiedź na drugie z pytań badawczych.

Zestawienie ze sobą tych prac pozwala na kompleksową analizę mechanizmów kształtujących nierówności na tle społeczno-ekonomicznym występujące w systemach edukacji. Autorzy analizowanych badań wyróżniają liczne zmienne pośredniczące w opisywanej relacji – wśród nich są te związane z politykami publicznymi (stopień heterogeniczności oraz centralizacji systemu edukacji, a także polityki edukacyjne w skali mikro), ze szkołą (ilość oraz jakość nauczania w szkole, klimat szkoły, relacja szkoła – uczeń), rodzicami (zaangażowanie oraz troska rodziców związane ze szkołą, inwestycje w rozwój dzieci) oraz samymi dziećmi (zaangażowanie oraz motywacja do nauki, wiara we własne możliwości).

Trzecie pytanie badawcze odnosiło się do najważniejszych konkluzji oraz implikacji dla polityk publicznych, które zostały wyróżnione przez autorów badań. Po przeanalizowaniu tych prac można stwierdzić, że status społeczno-ekonomiczny jest jednym z kluczowych czynników kształtujących osiągnięcia edukacyjne nastolatków.

Kolejnym cennym rezultatem tych badań jest zwrócenie uwagi na polityki publiczne, które potencjalnie mogą wpływać na kształt analizowanej zależności. Odpowiednie decyzje aktorów w systemie edukacji (decydentów politycznych, dyrektorów szkół, nauczycieli, rodziców) mogą przyczynić się do ograniczenia negatywnego wpływu niskiej pozycji społecznej dzieci na prawdopodobieństwo ich sukcesu w szkole. Analizowane prace wyróżniają działania mogące pełnić taką rolę – są nimi np. ułatwianie dzieciom integracji ze szkołą, wpływanie na sposób organizacji lekcji, a także zwiększanie kompetencji nauczycieli.

Wydaje się, że powyższe badania mogą mieć duże znaczenie również z polskiej perspektywy. Wynika z nich, że polski system edukacji wypada stosunkowo dobrze w międzynarodowych porównaniach efektywności oraz względnej równości nauczania [Schmidt i in., 2015; Rowley i in., 2020]. Autorzy socjologicznych oraz ekonomicznych badań wskazują jednak na istotność statusu społeczno-ekonomicznego dla edukacyjnych wyborów polskiej młodzieży [zob. Sadura, 2017; Smulczyk, Dolata, Pokropek, 2019]. Wyróżnione w analizowanych pracach mechanizmy oddziaływania pozycji społecznej na wyniki w nauce mogą stanowić cenną inspirację dla polskich badaczy. Wydaje się, że zastosowanie metod oraz konceptualizacji zaproponowanych przez autorów tych prac do analizy polskiego systemu edukacji może być interesującym kierunkiem dalszych badań.

W przyszłych badaniach warto również zwrócić uwagę na nowy aspekt analizowanej zależności, który pojawił się po 2020 r. Pandemia koronawirusa oraz czasowe

przejście szkół na nauczanie zdalne mogą przyczynić się do znacznego pogłębienia nierówności edukacyjnych. W takim trybie nauczania szkoła ma mniejsze szanse na wyrównanie różnic w kapitale kulturowym, który dzieci wynoszą z domu. Ponadto kolejnym problemem, który pojawił się w związku z przejściem na edukację zdalną, jest zjawisko wykluczenia cyfrowego. Wydaje się, że istotnym kierunkiem dla przyszłych badań w zakresie zależności pomiędzy statusem społeczno-ekonomicznym a wynikami edukacyjnymi będzie analiza tych czynników związanych z nauczaniem zdalnym.

Inny interesujący kierunek przyszłych badań może stanowić analiza wpływu opisywanych w tej pracy zmiennych społeczno-ekonomicznych na alternatywne miary nabytych w szkole kompetencji. Taką miarą może być, uważany powszechnie za kluczową umiejętność we współczesnym świecie, poziom wiedzy finansowej (*financial literacy*) [OECD, 2017, s. 15]. Od 2012 r. ten aspekt edukacji pojawia się również w badaniach PISA, co pozwala na uwzględnienie tego zagadnienia w przyszłych badaniach nierówności edukacyjnych.

## Bibliografia

---

- Benner, A.D., Wang, Y. (2014). Demographic Marginalization, Social Integration, and Adolescents' Educational Success, *Journal of Youth and Adolescence*, 43(10), s. 1611–1627. DOI: 10.1007/s10964-014-0151-6.
- Broer, M., Bai, Y., Fonseca, F. (2019). Socioeconomic Inequality and Educational Outcomes: Evidence from Twenty Years of TIMSS, *Springer Nature*. DOI: 10.1007/978-3-030-11991-1.
- Bronfenbrenner, U., Morris, P.A. (1998). The Ecology of Developmental Process. W: *Handbook of Child Psychology: Theoretical Models of Human Development*, W. Damon, R. Lerner (Eds.). New York: John Wiley & Sons Inc. DOI: 10.1002/9780470147658.
- Ferraro, S., Pöder, K. (2018). School-Level Policies and the Efficiency and Equity Trade-Off in Education, *Journal of Policy Modeling*, 40(5), s. 1022–1037. DOI: 10.1016/j.jpolmod.2017.11.001.
- Frønes, I. (2016). The Absence of Failure: Children at Risk in the Knowledge Based Economy, *Child Indicators Research*, 9(1), s. 247–260. DOI: 10.1007/s12187-015-9309-3.
- Gustafsson, J.E., Nilsen, T., Hansen, K.Y. (2018). School Characteristics Moderating the Relation between Student Socio-economic Status and Mathematics Achievement in Grade 8. Evidence from 50 Countries in TIMSS 2011, *Studies in Educational Evaluation*, 57, s. 16–30. DOI: 10.1016/j.stueduc.2016.09.004.
- Harwell, M.R., Zhao, Q. (2021). An Empirical Example of Capturing the Impact of SES on Student Achievement Using Path Analysis, *International Journal of Educational Research*, 105, 101715. DOI: 10.1016/j.ijer.2020.101715.



- Henry, K.L., Cavanagh, T.M., Oetting, E.R. (2011). Perceived Parental Investment in School as a Mediator of the Relationship between Socio-economic Indicators and Educational Outcomes in Rural America, *Journal of Youth and Adolescence*, 40(9), s. 1164–1177. DOI: 10.1007/s10964-010-9616-4.
- Kim, Y., Mok, S.Y., Seidel, T. (2020). Parental Influences on Immigrant Students' Achievement-Related Motivation and Achievement: A Meta-Analysis, *Educational Research Review*, 30. DOI: 10.1016/j.edurev.2020.100327.
- Ledwith, V., Reilly, K. (2014). Fringe Benefits? Educational Experiences of Migrant and Non-migrant Youth in the Urban-rural Fringe of Galway City, Ireland, *Journal of Rural Studies*, 36, s. 219–225. DOI: 10.1016/j.jrurstud.2014.07.003.
- Marks, G.N. (2017). Is SES Really That Important for Educational Outcomes in Australia? A Review and Some Recent Evidence, *The Australian Educational Researcher*, 44(2), s. 191–211. DOI: 10.1007/s13384-016-0219-2.
- Mongeon, P., Paul-Hus, A. (2016). The Journal Coverage of Web of Science and Scopus: A Comparative Analysis, *Scientometrics*, 106(1), s. 213–228. DOI: 10.1007/s11192-015-1765-5.
- O'Connell, M. (2019). Is the Impact of SES on Educational Performance Overestimated? Evidence from the PISA Survey, *Intelligence*, 75, s. 41–47. DOI: 10.1016/j.intell.2019.04.005.
- OECD (2014). PISA 2012 Technical Report, PISA, OECD Publishing, Paris.
- OECD (2016). PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA, OECD Publishing, Paris.
- OECD (2017). PISA 2015 Results (Volume IV): Students' Financial Literacy, PISA, OECD Publishing, Paris.
- OECD (2019). PISA 2018 Results (Volume II): Where All Students Can Succeed, PISA, OECD Publishing, Paris.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C. i in. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews, *BMJ*, 372. DOI: 10.1016/j.ijisu.2021.105906.
- Pokropek, A., Borgonovi, F., Jakubowski, M. (2015). Socio-Economic Disparities in Academic Achievement: A Comparative Analysis of Mechanisms and Pathways, *Learning and Individual Differences*, 42, s. 10–18. DOI: 10.1016/j.lindif.2015.07.011.
- Rogelberg, S.L., Starrett, A., Irvin, M.J., DiStefano, C. (2021). Examining Motivation Profiles within and across Socioeconomic Levels on Educational Outcomes, *International Journal of Educational Research*, 109. DOI: 10.1016/j.ijer.2021.101846.
- Rowley, K.J., Edmunds, C.C., Dufur, M.J., Jarvis, J.A., Silveira, F. (2020). Contextualising the Achievement Gap: Assessing Educational Achievement, Inequality, and Disadvantage in High-income Countries, *Comparative Education*, 56(4), s. 459–483. DOI: 10.1080/03050068.2020.1769928.
- Sadura P. (2017). *Państwo, szkoła, klasy*. Warszawa: Wydawnictwo Krytyki Politycznej.
- Schmidt, W.H., Burroughs, N.A., Zoido, P., Houang, R. T. (2015). The Role of Schooling in Perpetuating Educational Inequality: An International Perspective, *Educational Researcher*, 44(7), s. 371–386. DOI: 10.3102/0013189X15603982.

- Shanks, T.R.W., Robinson, C. (2013). Assets, Economic Opportunity and Toxic Stress: A Framework for Understanding Child and Educational Outcomes, *Economics of Education Review*, 33, s. 154–170. DOI: 10.1016/j.econedurev.2012.11.002.
- Sirin, S.R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research, *Review of Educational Research*, 75(3), s. 417–453. DOI: 10.3102/00346543075003417.
- Smulczyk, M., Dolata, R., Pokropek, A. (2019). Selekcja na progu szkoły ponadgimnazjalnej: merytokracja czy statusowy determinizm?, *Kwartalnik Pedagogiczny*, 64, s. 216–236. DOI: 10.5604/01.3001.0013.8470.
- Takashiro, N. (2017). A Multilevel Analysis of Japanese Middle School Student and School Socio-economic Status Influence on Mathematics Achievement, *Educational Assessment, Evaluation and Accountability*, 29(3), s. 247–267. DOI: 10.1007/s11092-016-9255-8.
- Tomaszewski, W., Xiang, N., Western, M. (2020). Student Engagement as a Mediator of the Effects of Socio-economic Status on Academic Performance among Secondary School Students in Australia, *British Educational Research Journal*, 46(3), s. 610–630. DOI: 10.1002/berj.3599.
- UNDP (2020). Human Development Report 2020: The Next Frontier – Human Development and the Anthropocene, United Nations Development Programme, New York.
- White, K.R. (1982). The Relation between Socioeconomic Status and Academic Achievement, *Psychological Bulletin*, 91(3), s. 461. DOI: 10.1037/0033-2909.91.3.461.
- van Zwieten, A., Teixeira-Pinto, A., Lah, S., Nassar, N., Craig, J.C., Wong, G. (2021). Socioeconomic Status During Childhood and Academic Achievement in Secondary School, *Academic Pediatrics*, 21(5), s. 838–848. DOI: 10.1016/j.acap.2020.10.013.

## Załączniki

Tabela 1. Charakterystyka artykułów analizowanych w przeglądzie literatury

Badanie	Źródło danych	Wiek badanych	Kraje	Operacjonalizacja SSE	Zmienna pośrednicząca	Zmienna zależna
O'Connell, 2019	PISA 2015	15	65, głównie członkowie OECD	dochód gospodarstwa domowego, poziom wykształcenia rodziców	-	wyniki testów PISA z matematyki oraz czytania
Gustafsson i in., 2018	TIMSS 2011	13	50, głównie wysoko rozwinięte państwa	liczba książek w domu, poziom wykształcenia rodziców, posiadane zasoby do nauki (własny pokój do nauki, połączenie internetowe)	SSE szkoły, jakość oraz ilość nauczania, klimat nacisku na sukces w szkole, klimat przestrzegania zasad w szkole	wyniki testu TIMSS z matematyki
Tomaszewski i in., 2020	LSAC, NAPLAN	12-13	Australia	poziom wykształcenia rodziców, dochód gospodarstwa domowego, prestiż zawodu rodziców	zaangażowanie ucznia	wyniki testów NAPLAN z czytania oraz matematyki
Pokropek i in., 2015	PISA 2012	15	członkowie OECD (33)	poziom wykształcenia rodziców, prestiż zawodu rodziców	dostęp do ekonomicznych, kulturowych oraz edukacyjnych środków (np. do komputera do nauki w domu)	wyniki testów PISA z matematyki, czytania oraz nauk przyrodniczych
Marks, 2017	LSAC, NAPLAN	12-13	Australia	poziom wykształcenia rodziców, dochód gospodarstwa domowego, prestiż zawodu rodziców	-	wyniki testów NAPLAN z czytania oraz matematyki
Rogelberg i in., 2021	ELS 2002	15-16	USA	poziom wykształcenia rodziców, dochód gospodarstwa domowego, prestiż zawodu rodziców	poziom wewnętrznej motywacji oraz oceny własnych umiejętności	ocena z matematyki w 12. klasie, ukończenie liceum, kontynuacja edukacji po liceum
Harwell, Zhao, 2021	PISA 2012	15	USA	poziom wykształcenia rodziców, prestiż zawodu rodziców	kapitał materialny (dostępne środki do nauki), kapitał społeczny (wsparcie nauczycieli, relacje uczeń – nauczyciel)	wyniki testu PISA z matematyki
Rowley i in., 2020	PISA 2012	15	42 wysoko rozwinięte państwa	Wskaźnik Ekonomicznego, Społecznego i Kulturowego Statusu (ESCS)	system edukacji	wyniki testu PISA z matematyki

Badanie	Źródło danych	Wiek badanych	Kraje	Operacjonalizacja SSE	Zmienna pośrednicząca	Zmienna zależna
Ferraro, Pöder, 2018	PISA 2012	15	Włochy	poziom wykształcenia rodziców	system edukacji (polityki publiczne na poziomie szkolnym)	wyniki testu PISA z matematyki
Takashiro, 2017	TIMSS 2003	13	Japonia	poziom wykształcenia rodziców, posiadanie komputera w domu, liczba książek w domu, uczestnictwo w dodatkowych zajęciach z matematyki	SES szkoły	wyniki testu TIMSS z matematyki
van Zwieten i in., 2021	LSAC, NAPLAN	12–13	Australia	poziom wykształcenia rodziców, zawod rodziców (w dziecięcym wieku badanych)	–	wyniki testów NAPLAN z czytania oraz matematyki
Shanks, Robinson, 2013	brak (model teoretyczny)	–	–	poziom wykształcenia rodziców, dochód gospodarstwa domowego, kapitał	poziom stresu	wyniki w nauce
Benner, Wang, 2014	NLSAH	12–18	USA	poziom wykształcenia rodziców, zawod rodziców	samotność w szkole, poczucie związku ze szkołą	przeciętne oceny, liczba lat uczestnictwa w edukacji
Schmidt i in., 2015	PISA 2012	15	62, głównie członkowie OECD	Wskaźnik Ekonomicznego, Społecznego i Kulturowego Statusu (ESCS)	szansa poznania materiału w szkole	wyniki testu PISA z matematyki
Henry i in., 2011	Stanley i in., 2008	11–14	USA	poziom wykształcenia rodziców, subiektywna sytuacja materialna gospodarstwa domowego	zaangażowanie rodziców w szkołę	prawdopodobieństwo dropoutu (nieukończenia szkoły)
Frones, 2016	kohortowe spisy powszechne w Norwegii	21–23	Norwegia	relatywne ubóstwo, nastoletnia matka, poziomy wykształcenia rodziców, liczba przepracowanych lat przez rodziców po skończeniu 25 r.ż.	–	prawdopodobieństwo dropoutu (nieukończenia szkoły), oceny ucznia na zakończenie szkoły ponadpodstawowej
Broer i in., 2019	TIMSS 1995–2015	13	13, głównie wysoko rozwinięte państwa	liczba książek w domu, poziom wykształcenia rodziców, posiadane zasoby do nauki (komputer, własne biurko)	system edukacji	wyniki testów TIMSS z matematyki oraz nauk przyrodniczych

Źródło: opracowanie własne.

# Socio-Economic Status and Academic Performance in Secondary Schools. A Systematic Literature Review

## Summary

The socioeconomic status of children is often discussed in the literature as a one of the key factors affecting their academic performance and chances of success at school. The aim of this article is to describe contemporary research on this dependence among adolescents in developed countries. Using the PRISMA method, 17 research papers were selected and made subject to systematic analysis. The main focus of the analysis was the way in which educational research operationalises the socio-economic status, intermediary variables of the described dependence and implications for public policies. The authors of the analysed works distinguish numerous mechanisms which contribute to the emergence of differences in learning outcomes as well as educational inequalities. They include those related to the conditions of the education system, educational policies, school and its relationship with the students, parents and among children themselves. This allows for indicating actions for a variety of actors within the public education system – policy makers, school principals, teachers and parents. A systematic literature review also shows the growing popularity of this issue in social sciences in recent years.

**Keywords:** socio-economic status, educational results, educational achievements, educational inequalities, PRISMA



# Nieracjonalne oczekiwania na rynkach finansowych – przegląd literatury finansów behawioralnych

Joanna Krysta

## Streszczenie

Artykuł bada implikacje rozluźnienia założenia racjonalnych oczekiwań, będącego standardowo przyjmowanym założeniem w modelach badających różne zjawiska ekonomiczne. Korzystając z literatury z dziedziny finansów behawioralnych, artykuł próbuje odpowiedzieć na pytania dotyczące powodów posługiwania się przez racjonalnie zachowujących się uczestników rynku nieracjonalnymi oczekiwaniami przy podejmowaniu decyzji inwestycyjnych. Charakteryzuje on także owe heterogeniczne oczekiwania graczy rynkowych oraz bada wpływ nieracjonalnych oczekiwań na ostateczne stopy zwrotu na rynkach finansowych. Opisane zniekształcenia i błędy poznawcze prowadzące do obciążenia oczekiwań, źródła nieracjonalnych oczekiwań oraz ich powszechność, uzewnętrzniająca się m.in. w ich udokumentowanym istotnym wpływie na stopę zwrotu, sugerują, iż w pełni racjonalny uczestnik rynku powinien uwzględniać je przy kształtowaniu racjonalnych oczekiwań i podejmowaniu decyzji inwestycyjnych. Z kolei uczestnik rynku ulegający w sposób nieświadomy którymś spośród opisanych heurystyk ma okazję do dokonania rewizji swoich oczekiwań oraz próby kontroli wpływu zniekształceń poznawczych na podejmowane przez siebie decyzje inwestycyjne.

**Słowa kluczowe:** nieracjonalne oczekiwania, teoria perspektywy, zniekształcenia poznawcze, zjawisko momentum, overconfidence

## 1. Wstęp – badanie implikacji odrzucenia założenia racjonalnych oczekiwań

Racjonalność podmiotów podejmujących decyzje w gospodarce jest jednym z powszechnie przyjmowanych założeń przez konstruujących modele służące

wyjaśnieniu różnych zjawisk ekonomicznych [Mankiw, 1990]. To założenie składa się z dwóch części.

Pierwsza część odnosi się do **racjonalności oczekiwań** (hipotezy racjonalnych oczekiwań w kształtowaniu oczekiwań [Haltiwanger, Waldman, 1985]). Zakłada się, iż podmioty gospodarcze znają obiektywny rozkład prawdopodobieństwa, który generuje wartości zmiennych losowych [Manski, 2004], takich jak np. stopy zwrotu na rynkach finansowych. Racjonalność oczekiwań wymaga także stosowania przez podmioty gospodarcze reguły Bayesa do aktualizacji własnych oczekiwań za każdym razem, gdy posiadają oni nową informację mogącą wpłynąć na dany zwrot.

Druga część zakłada kierowanie się przez graczy rynkowych racjonalnością w podejmowaniu decyzji (tzw. **racjonalne zachowanie**). Wiąże się ono zazwyczaj z podejmowaniem decyzji optymalizujących funkcje celu, czy też maksymalizujących użyteczność decydenta, na podstawie dostępnych informacji i oczekiwań dotyczących przedmiotu decyzji, w ramach ograniczeń w postaci, np. różnego rodzaju środków dostępnych decydentowi (takich jak czas, środki finansowe czy inne szeroko pojęte zasoby) [np. Harsanyi, 1980].

Racjonalność w rozumieniu obu części definicji jest pożądanym przez ekonomistów założeniem, znacznie ułatwia bowiem modelowanie i analizę zjawisk gospodarczych [Kahneman, 1994]. W szczególności, przy dodatkowym założeniu, że wszyscy gracze rynkowi posługują się identycznym rozkładem początkowym (*common prior*), często przyjmowanym w modelach z teorii gier [Morris, 1995], utrzymywanie założenia występowania racjonalnych oczekiwań znacznie ułatwia analizę zjawisk gospodarczych. Analogicznie jest w przypadku spełnienia założenia racjonalnego zachowania: zapewnia ono bowiem, do pewnego stopnia, przewidywalność zachowań i oczekiwań graczy rynkowych oraz umożliwia ekonomistom modelowanie podejmowanych przez podmioty gospodarcze decyzji i wyborów, na podstawie posiadanych informacji dotyczących ich preferencji i zasobów [Harsanyi, 1980].

Nieracjonalność może być rozumiana jako zaprzeczenie przytoczonej wyżej definicji racjonalności. W rozumieniu pierwszej jej części, nieracjonalne oczekiwania będą się wiązały z wszelkimi, różnorodnymi oczekiwaniami utrzymywanymi przez graczy rynkowych, które niekoniecznie (lub nieumiejętnie) będą aktualizowane przy pomocy reguły Bayesa [De Bondt, Thaler, 1987]. Widać zatem, iż tutaj zbiór dopuszczalnych oczekiwań jest dużo bogatszy niż w przypadku ograniczenia się do oczekiwań racjonalnych.

W niniejszym artykule skupiono się na pierwszej kwestii, a więc kwestii różnorodności oczekiwań graczy rynkowych, zakładając ich racjonalne zachowanie (a więc spełnienie drugiej części wyżej przytoczonej definicji racjonalności). Tak więc celem tego artykułu jest zaprezentowanie przeglądu literatury finansów behawioralnych



dotyczącej istnienia i wpływu heterogenicznych oczekiwań na stopy zwrotu na rynkach finansowych, przy założeniu, że wszyscy gracze rynkowi kierują się racjonalnością przy podejmowaniu ostatecznych decyzji. Takie podejście jest przyjmowane w wielu pracach badawczych [np. Dominitz, Manski, 2010; Haltiwanger, Waldman, 1985; Dave, 2011]. Aby zrealizować wyżej postawiony cel, artykuł analizuje historię badań empirycznych z dziedziny finansów behawioralnych w celu opisanie oraz analizy wyników tych badań, a także przedstawienia zjawisk udokumentowanych w literaturze klasycznej finansów behawioralnych. Artykuł przedstawia również wyniki współczesnych badań z tej dziedziny.

Wiele zjawisk obserwowanych na rynkach finansowych zdaje się naruszać założenie racjonalności oczekiwań, co odzwierciedlają licznie udokumentowane przypadki obciążenia oczekiwań graczy rynkowych (*behavioral bias* [np. Tversky, Kahneman, 1974; Cohen, 1981]). Spełnienie założenia racjonalności oczekiwań nie dopuszcza także częstej realizacji takich zdarzeń, jak np. bańki informacyjne (*market bubbles*), czy zachowania stadne u inwestorów (*investors' herding*) – to założenie implikuje bowiem zerowe prawdopodobieństwo ich wystąpienia. W ciągu ostatnich kilku dekad doszło jednak do kilku załamań na rynkach finansowych, np. spowodowanych przewartościowaniem dolara amerykańskiego w latach 80., międzynarodowym krachem na giełdzie z roku 1987, czy załamaniem rynku Bitcoina z początku 2018 r., które były najprawdopodobniej wynikiem używania analizy technicznej i/lub nadmiernych spekulacji dotyczących fundamentów rynku przez dużą część uczestników rynku finansowego [Frankel, Froot, 1986; Shiller, 1987; Taskinsoy, 2019]. Warto wobec tego przyrzeć się obecnym badaniom empirycznym i literaturze dotyczącej nieracjonalności oczekiwań podmiotów gospodarczych.

Analizując różne oczekiwania uczestników rynku co do kształtowania się przyszłych stóp zwrotu z aktywów, warto przytoczyć słynną w literaturze finansów hipotezę informacyjnej efektywności rynku finansowego (HIERF) Eugene'a Famy [Fama, 1960]. Hipoteza ta postuluje, iż ceny aktywów na rynkach finansowych odzwierciedlają ich (zmienną w czasie i aktualizowaną na skutek pojawienia się nowych informacji nieoczekiwanych wcześniej przez ogół graczy rynkowych) wartość fundamentalną adekwatną dla danego momentu w czasie. Stanowi ona konsensus uczestników rynku dotyczący wysokości strumienia zdyskontowanych przyszłych przychodów finansowych generowanych przez owe aktywo [Fama, 2021; Fama, 1976]. Oznacza to, iż wszystkie dostępne w danym momencie informacje i oczekiwania uczestników rynku co do przyszłych zdarzeń wpływających na wysokość cen aktywów są przeciętnie uwzględnione w tej cenie w czasie rzeczywistym<sup>1</sup>. Wszystkie

<sup>1</sup> Z opóźnieniem będącym samym w sobie również zmienną losową.

wahania wokół wartości fundamentalnej analizowanego aktywa finansowego są natomiast z natury losowe, a zatem niemożliwe do przewidzenia<sup>2</sup> [Fama, 1995].

Fama [1995] podsumowuje konsensus badaczy naukowych co do wykazywania przez aktywa trajektorii przypominających błądzenie losowe, a więc w szczególności procesu charakteryzującego się brakiem występowania zależności pomiędzy cenami danego aktywa w czasie. Wyniki te sugerują, iż kompletowanie portfela inwestycyjnego, posługując się historią cen aktywów, zwaną w literaturze finansów analizą techniczną<sup>3</sup>, do szacowania przyszłych zwrotów finansowych, nie powinno w długim okresie przynosić wyników lepszych niż losowy dobór aktywów do portfela.

Ze względu na wyniki empiryczne zdające się zaprzeczać poprawności analizy technicznej oraz brak osiągania systematycznych zysków przez posługujących się nią inwestorów [Taylor, Allen, 1992], w literaturze przedmiotu zwykło się ją traktować sceptycznie [Malkiel, 1985; Sharpe, 1985]. Wyniki niektórych badań empirycznych wykazujące formułowanie oczekiwań co do kierunku zmian cen aktywów, bazowane m.in. na analizie technicznej, zdają się sugerować nieracjonalność oczekiwań uczestników rynku [np. Froot, Frankel, 1989].

Warto zadać pytanie dotyczące powodów bazowania przez niektórych uczestników rynku swoich decyzji inwestycyjnych na nieracjonalnych oczekiwaniach. W jaki sposób uzasadniają oni te oczekiwania i dlaczego uznają je za słuszne? Jakie rodzaje nieracjonalnych oczekiwań występują wśród inwestorów finansowych i czy są one wystarczająco powszechne, aby mieć istotny wpływ na stopy zwrotu na rynkach finansowych? Jeśli faktycznie na rynkach finansowych występuje liczna grupa graczy posiadających nieracjonalne oczekiwania, jakie decyzje inwestycyjne powinien zatem podejmować gracz posiadający racjonalne oczekiwania, a więc próbujący uwzględnić w swoich decyzjach inwestycyjnych wszystkie czynniki mogące wpłynąć na wysokość stóp zwrotu?

Przedstawiony przegląd literatury jest zorganizowany w następujący sposób. W drugim rozdziale zostały zaprezentowane sposoby wykorzystywane przez podmioty gospodarcze do racjonalizacji nieracjonalnych oczekiwań, opisane na podstawie literatury klasycznej finansów behawioralnych oraz najnowszych kierunków badań z tej dziedziny, które rozwinęły się w ciągu ostatnich kilkunastu lat. W trzecim rozdziale opisane zostały przykłady występujących nieracjonalnych oczekiwań oraz prawdopodobne źródła ich pochodzenia, zgodnie z literaturą finansów beha-

<sup>2</sup> Oznacza to w praktyce, iż w dowolnym momencie czasu każdy uczestnik rynku, bez względu na cechującą go liczbę lat doświadczenia w inwestowaniu, ma około 50% szans na odgadnięcie kierunku zmian cen dowolnej akcji, natomiast praktycznie niemożliwe jest systematyczne osiąganie w ten sposób zysków finansowych.

<sup>3</sup> Polegającą na szukaniu wzorów i zależności w szeregach czasowych przeszłych cen aktywów oraz na próbie tłumaczenia przyszłych cen aktywów ich przeszłymi trajektoriami [Taylor, Allen, 1992].

wioralnych. W czwartym rozdziale skupiono się na istotności nieracjonalnych oczekiwań w praktyce, a więc ich wpływie na stopy zwrotu na rynkach finansowych<sup>4</sup>. Może on służyć także jako uzasadnienie zasadności badania nieracjonalnych oczekiwań wśród graczy rynkowych oraz zajmowania się tym tematem. Artykuł został zakończony podsumowaniem, gdzie wskazane zostały również potencjalne kierunki badań dotyczących przedstawionego tematu.

## 2. Sposoby racjonalizacji nieracjonalnych oczekiwań przez graczy rynkowych

We wstępie do artykułu zostały przedstawione powody sugerujące brak racjonalności oczekiwań graczy rynkowych oraz zadane pytania dotyczące implikacji, jakie niesie za sobą odrzucenie założenia racjonalnych oczekiwań. Ten rozdział, stanowiący preludeum do analizy zjawiska nieracjonalnych oczekiwań, skupi się na wymienieniu i opisanu prawdopodobnych powodów, dla których gracze rynkowi maksymalizujący swoje stopy zwrotu (postępujący racjonalnie), korzystają mimo wszystko z nieracjonalnych oczekiwań.

W tym rozdziale przedstawione zostały sposoby uzasadniania przez podmioty gospodarcze poprawności swoich nieracjonalnych oczekiwań. Opisane zostały kolejno zjawiska błędu confirmacji, efektu fałszywego konsensusu (zwanego także efektem fałszywej zgody), iluzji kontroli oraz *overconfidence*, udokumentowane w literaturze finansów behawioralnych oraz licznych badaniach empirycznych. Takie podejście do inwestowania, przede wszystkim, gdy gracz rynkowy podejmujący decyzję inwestycyjną nie liczy się ze zdaniem pozostałych uczestników rynku (nie uwzględnia w niej rynkowych sentymentów lub uwzględnia jedynie informacje zgodne z jej/jego przekonaniem), może prowadzić do strat finansowych w przypadku zainwestowania – gdy wśród pozostałych uczestników rynkowych przeważa przekonanie o negatywnej stopie zwrotu w przyszłym okresie lub niewykorzystanych możliwości zarobku – gdy występuje przeważające przekonanie o pozytywnym zwrocie [Forsythe, Nelson, Neumann, Wright, 1992]. Warto zatem przeanalizować udokumentowane w literaturze finansów behawioralnych źródła takich zachowań.

<sup>4</sup> Aby nieracjonalne oczekiwania uczestników rynku wpłynęły w sposób statystycznie istotny na cenę, konieczne jest występowanie wystarczająco licznej grupy podmiotów gospodarczych posiadających takowe oczekiwania. Wynika to chociażby ze standardowego założenia znikomego wpływu decyzji indywidualnych graczy rynkowych (jednostek stanowiących nieskończenie mały udział wśród wszystkich uczestników rynku) na stopy zwrotu na rynkach finansowych (tzw. *price taking hypothesis*, hipoteza najprawdopodobniej sformułowana po raz pierwszy w Cournot, 1838).

Błąd konfirmacji (*confirmation bias*) odnosi się do ludzkiej tendencji do szukania jedynie takich informacji, które utwierdzą podmioty gospodarcze w poprawności wyznawanych przez nie przekonań [Hilton, 2001]. Unikane są natomiast informacje, które zaprzeczają ich poprawności. Z eksperymentu zaprezentowanego w publikacji Wasona i Johnsona-Lairda [1972] wynika, że ludzie mają tendencję do uznawania poprawności różnych stwierdzeń na podstawie wyszukiwania jedynie takich informacji, które mogłyby świadczyć o ich poprawności. Nie sprawdzają natomiast, czy występują wyjątki od reguł uznanych przez nich za słuszne. W kontekście rynków finansowych, gracze rynkowi przekonani o własnej nieomyślności mogą mieć tendencję do szukania tylko takich informacji, np. na różnego rodzaju forach internetowych, które potwierdzą ich własne przekonania, oraz do unikania lub nieszukania informacji, które świadczyłyby przeciwnie [Hilton, 2001; Forsythe i in., 1992]. Takie zachowanie z kolei może sprawić, że w przypadku pojawienia się nowej informacji świadczącej przeciwnie do przekonania danego podmiotu, zostanie ona zignorowana i nieuwzględniona przy podejmowaniu przez nią/niego decyzji ekonomicznej, a więc niezainwestowania, gdy wśród pozostałych graczy rynkowych będzie istnieć przeważające przekonanie o pozytywnym zwrocie lub zainwestowania w przypadku obecnego wśród pozostałych graczy rynkowych silnego przeważającego przekonania o ujemnej stopie zwrotu [Hilton, 2001]. Zostało natomiast udokumentowane [np. w Forsythe i in., 1992], że jest bardziej prawdopodobne, iż gracze rynkowi niepodlegający tej tendencji albo umiejący świadomie sobie z nią radzić, osiągną dodatnie stopy zwrotu.

Efekt fałszywego konsensusu lub efekt fałszywej zgody (*false consensus bias*) wiąże się z ludzką tendencją do ślepej wiary we własne przekonania, która charakteryzuje się przeszacowywaniem liczby osób, które posiadają takie same przekonania lub preferencje, jak oni sami [Hilton, 2001]. W sytuacji inwestycyjnej może ona prowadzić do pochopnych decyzji inwestycyjnych – gdy dany gracz rynkowy jest przekonany, iż decydując się na zakup albo sprzedaż danego aktywa finansowego należy do większości podmiotów rynkowych, która również podjęła tę samą decyzję. Dany uczestnik rynku może założyć w ten sposób nieomyślność swoich przekonań i w oparciu o nie podjąć swoją decyzję inwestycyjną. Eksperyment przedstawiony w publikacji Rossa, Greene'a, House'a [1977] bada tę tendencję. Na podstawie przeprowadzonych ankiet dotyczących potencjalnych sytuacji oraz autentycznych okoliczności, przed jakimi zostali postawieni uczestnicy badania, wykazano przeważające przekonanie o posiadaniu przez większość osób tych samych przekonań, co osoba badana, i nieposiadaniu przekonań nienależących do niej.

Iluzja kontroli (*illusion of control*), zwana także skłonnością ludzi do myślenia magicznego, jest heurystyką udokumentowaną po raz pierwszy u Langer [1975],

która opisuje ludzką tendencję do przeszacowywania możliwości kontroli zdarzeń losowych. Takie osoby mogą być przekonane o posiadaniu pewnego rodzaju zdolności magicznych, które umożliwiają im kontrolę nad daną sytuacją. Według Langer i Roth [1975] efekt ten jest silniejszy, gdy dana osoba często podejmuje danego rodzaju decyzje lub gdy analizowana jest sytuacja uwzględniająca konkurencję. Osoby rzucające kostką do gry mogą być np. przekonane o własnej umiejętności wpływania na wynik rzutu. W eksperymencie opisanym u Rothbarta, Snydera [1970] pierwsza część uczestników miała założyć się o wynik rzutu przed jego wykonaniem, a pozostała część – po rzucie. Wyniki wskazują na większą pewność badanych z pierwszej grupy co do poprawności odgadnięcia wyniku rzutu, w porównaniu do tych z grupy drugiej. Natomiast uczestnicy rynków finansowych mogą być zdania, iż są w stanie kontrolować przyszłe stopy zwrotu, co może skutkować u nich bazowaniem na tym przekonaniu swoich decyzji inwestycyjnych. Eksperyment udokumentowany w publikacji Heatha i Tversky'ego [1991] stawiał uczestników badania przed wyborem pomiędzy dwiema sytuacjami. W obu poprzez losowanie wyselekcjonowano akcje pewnej spółki, a poprawna odpowiedź po wybraniu jednej z sytuacji była gratyfikowana za pomocą niewielkiej nagrody pieniężnej. W pierwszej sytuacji należało przewidzieć, czy kurs akcji w dniu jutrzejszym spadnie, czy wzrośnie. W drugiej natomiast, bez sprawdzania informacji w żadnym źródle – odgadnąć, czy kurs tych samych akcji wczoraj wzrósł, czy spadł. Mimo iż obie sytuacje przedstawiają realizację zmiennych losowych, uczestnicy badania preferowali odgadnąć jutrzejszy kurs akcji, a więc sytuację pierwszą.

*Overconfidence* jest heurystyką badaną oraz zdefiniowaną w literaturze finansów behawioralnych na trzy różne, nietożsame sposoby<sup>5</sup>. U Moore'a i Healy'ego [2008] zostały one nazwane kolejno *overestimation*, *overplacement* i *overprecision*. Zgodnie z Moorem i Healym [2008], *overestimation* odnosi się do systematycznego przeszacowywania własnych umiejętności. *Overplacement* cechuje podmioty gospodarcze, które przeszacowują własne umiejętności względem innych, sądząc, iż są od nich lepsze. *Overprecision* natomiast odnosi się do nadmiernego ludzkiego przeświadczenia, iż zna się właściwą odpowiedź na dane pytanie lub też, że posiadane przekonania są właściwe. W badaniach empirycznych dotyczących zjawiska *overprecision*, uczestnikom badania zadawane są pytania wymagające oszacowania x-procentowego przedziału ufności, np. 90-procentowego<sup>6</sup>. Przykładowe pytanie może dotyczyć podania wysokości lub długości danego obiektu geograficznego, jak np. szczyt górski

<sup>5</sup> Moore i Healy [2008] postulują, iż definicje te w ogólności odnoszą się do różnych zjawisk, a więc ich pomiar powinien być przeprowadzany osobno.

<sup>6</sup> Jest to taki przedział, dla którego uczestnik badania jest w 90% pewny, że w nim mieści się prawidłowa odpowiedź.

czy rzeka. Wyniki badań świadczą, iż podawane przedziały są przeważnie zbyt wąskie, ponieważ właściwa odpowiedź znajduje się w nich w mniej niż 50% przypadków – zamiast w 90% [np. Soll, Klayman, 2004].

Wszystkie trzy sposoby identyfikacji i pomiaru zjawiska *overconfidence* są badane formalnie przez Moore'a, Healy'ego [2008]. W pierwszym kroku zakładają oni przeprowadzenie pewnego testu lub zadania, w którym badane podmioty zdobywają wynik znajdujący się na pewnej skali. W drugim natomiast zakładają posiadanie przez podmioty gospodarcze subiektywnego rozkładu prawdopodobieństwa dotyczącego własnego wyniku osiągniętego w owym teście lub zadaniu, oraz odpowiadającego mu rozkładu dotyczącego wyników innych osób. Oba te rozkłady są zdefiniowane na całym przedziale możliwych do osiągnięcia wyników.

Zgodnie z Moorem i Healy'em [2008] mówimy, iż dana osoba podlega heurystyce *overestimation*, jeśli wartość oczekiwana subiektywnego rozkładu prawdopodobieństwa opisującego własny wynik testu jest wyższa niż wynik rzeczywiście osiągnięty przez daną osobę w zadaniu. *Overplacement* zachodzi natomiast wtedy, gdy wartość oczekiwana własnego wyniku jest wyższa od wartości oczekiwanej innych uczestników badania, a przekonanie to nie jest wyjaśniane za pomocą różnicy w rzeczywistych wynikach. *Overprecision* jest urzeczywistnione w subiektywnym rozkładzie prawdopodobieństwa własnego wyniku posiadającego niższą wariancję, niż wskazywałyby na to rzeczywiste wyniki testu danego uczestnika badania<sup>7</sup>.

W kontekście rynków finansowych przekonanie o własnej nieomyślności, a więc, że jest się w stanie przewidzieć przyszłe stopy zwrotu, w dodatku dużo lepiej, niż jest to możliwe w rzeczywistości, i dużo lepiej niż inni gracze rynkowi, może prowadzić do nieprzemyślanych decyzji inwestycyjnych, czy nawet baniek informacyjnych [np. Plous, 1993; Camerer, Lovallo, 1999; Odean, 1999]. Racjonalność oczekiwań wiążąca się z (obiektywną) analizą sentymentów rynkowych, bayesowską aktualizacją prawdopodobieństwa osiągnięcia pozytywnego wyniku na inwestycji oraz analizą fundamentalną firmy, w której aktywa dany podmiot gospodarczy zamierza zainwestować, zmniejsza bowiem prawdopodobieństwo popełnienia błędu, a w rezultacie osiągnięcia straty na inwestycji [Hilton, 2001].

Ostatnim zjawiskiem z literatury klasycznych finansów behawioralnych opisanym w tym rozdziale jest efekt zakotwiczenia (*anchoring effect* [Tversky, Kahneman, 1974]). Opisuje on ludzką tendencję do bycia zakotwiczonym w przypadkowej liczbie, którą dany podmiot gospodarczy zobaczył przed podjęciem danej decyzji lub udzieleniem odpowiedzi na dane pytanie. Oznacza to, iż decyzje i odpowiedzi takiej osoby są obciążone w stronę tej przypadkowej liczby, nie mającej jednak nic

<sup>7</sup> Oznacza to np., że po wielokrotnym przeprowadzeniu podobnego testu powstałby empiryczny rozkład prawdopodobieństwa wyników danej osoby, posiadający wyższą wariancję.

wspólnego z zadaniem pytaniem. W sytuacji inwestycyjnej, po usłyszeniu lub zobaczeniu pewnej losowej liczby, np. na pewnym forum internetowym, decydent podlegający efektowi zakotwiczenia będzie miał tendencję do bazowania swojej decyzji inwestycyjnej na tej liczbie, która nie będzie jednak uzasadniona w żaden dodatkowy sposób. Wyniki badania empirycznego opisanego u Stephana i Kiella [2017] wskazują, iż efekt zakotwiczenia występuje także wśród profesjonalnych inwestorów na giełdzie. W owym badaniu grupy inwestorów zostały poproszone o przewidzenie, czy przyszła wartość indeksu rynkowego będzie wyższa, czy niższa niż pewna losowa liczba, przy czym liczby te różniły się pomiędzy grupami – tak samo, jak później uzyskane różnice pomiędzy oszacowaniami w obu grupach okazały się statystycznie istotne.

Należy również zwrócić uwagę na badania dotyczące błędów poznawczych pochodzące z ostatnich lat. Przykładowo autorzy Sadi, Asl, Rostami, A. Gholipour, F. Gholipour [2011] wykazali silny związek pomiędzy cechami osobowości i błędami poznawczymi, wpływającymi na wielkość odchylenia oczekiwań inwestorów od oczekiwań racjonalnych. Stanowi zatem wskazówkę dla inwestorów chcących kształtować swoje oczekiwania w sposób racjonalny. Analizowane są takie błędy poznawcze, jak: *overconfidence bias*, heurystyka dostępności<sup>8</sup>, eskalacja zaangażowania<sup>9</sup>, efekt pewności wstecznej<sup>10</sup> i błąd losowości<sup>11</sup>. Wykorzystany został również tzw. model Wielkiej Piątki (*Big Five Model*), który opisuje osobowość inwestorów za pomocą m.in. pięciu głównych składowych, do których należy określenie stopnia ekstrawersji inwestorów (*extraversion*), stopnia ich ugodowości (*agreeableness*), sumienności (*conscientiousness*), neurotyzmu (*neuroticism*) i otwartości na doświadczenia (*openness to experience*). Przeprowadzone badanie analizuje występowanie oraz znaki korelacji pomiędzy poszczególnymi błędami poznawczymi a cechami osobowości inwestorów. Wykazuje m.in. występowanie bezpośredniego związku pomiędzy otwartością inwestorów a efektem pewności wstecznej i *overconfidence bias*.

<sup>8</sup> *Availability bias*, zjawisko to występuje, gdy zawyżane jest prawdopodobieństwo wystąpienia zdarzeń losowych łatwiej przywoływalnych do świadomości, np. w przypadku występowania na ich temat łatwo dostępnych danych, przy jednoczesnym zaniżaniu prawdopodobieństwa wystąpienia tych zdarzeń, które trudniej jest sobie wyobrazić lub dostrzec [np. Kim, Nofsinger, 2008].

<sup>9</sup> *Escalation of commitment*, zjawisko występujące w przypadku gdy kontynuowane są działania i zwiększane jest zaangażowanie osoby w decyzję, która może być zła, ze względu na poczucie odpowiedzialności za ewentualną porażkę oraz chęć udowodnienia, że początkowa decyzja nie była błędna [np. Gholipour, Kehtarnavaz, Briggs, Devous, Gopinath, 2007].

<sup>10</sup> *Hindsight bias*, zjawisko pojawiające się w przypadku gdy inwestor jest ex-post przekonany o możliwości trafniejszego przewidzenia realizacji pewnego zdarzenia losowego, które wystąpiło w przeszłości, co może prowadzić np. do tworzenia przez niego związków przyczynowo-skutkowych pomiędzy dwoma niezależnymi zdarzeniami [np. Fischhoff, Beyth, 1975].

<sup>11</sup> *Randomness bias*, zjawisko wpływu przesądów na postrzeganie zdarzeń losowych, które, różniąc się wśród inwestorów, mogą wpływać na różne postrzeganie szczęśliwego trafu i zbiegu okoliczności, oddziałując zatem w różny sposób na decyzje inwestycyjne [np. Gholipour i in., 2007].

### 3. Nieracjonalne oczekiwania graczy rynkowych

W poprzednim rozdziale opisane zostały sposoby racjonalizacji nieracjonalnych oczekiwań uczestników rynku. W tym zaś omówione zostaną owe nieracjonalne oczekiwania (często wzajemnie się wykluczające) opisane w literaturze finansów behawioralnych oraz próby wyjaśnienia źródeł ich pochodzenia. Nieracjonalne oczekiwania można podzielić na oczekiwania dotyczące zarówno spadku, jak i wzrostu cen aktywów po ich wcześniejszym spadku albo wzroście. Opisane zostały sposoby uzasadniania przez graczy rynkowych wszystkich tych czterech możliwości. W tym celu przywołano teorię perspektywy [Kahneman, Tversky, 1979], wymagającą dłuższego wprowadzenia, oraz teorię *momentum* [Asness, Moskowitz, Pedersen, 2013].

Opis różnorodnych oczekiwań skupił się na perspektywie inwestora, który podejmuje decyzję dotyczącą inwestycji lub jej braku, polegającej na zajęciu pozycji długiej<sup>12</sup>. Odpowiadająca temu podejściu analiza inwestorów alokujących swoje aktywa w pozycjach krótkich jest analogiczna i została z tego powodu pominięta.

W celu przeanalizowania występującego na rynkach finansowych przekonania, iż w najbliższej przyszłości trend cenowy aktywa finansowego ulegnie odwróceniu, a więc nastąpi odbicie ceny po wcześniejszym spadku albo po poprzedzającym wzroście ceny, nastąpi jej spadek, można posłużyć się teorią perspektywy [Kahneman, Tversky, 1979]. Autorzy tej teorii podważają poprawność teorii oczekiwanej użyteczności, która skupia się na ewaluacji i porządkowaniu możliwych wyborów cechujących się niepewnością lub ryzykiem dotyczącym ostatecznego wyniku zdarzenia losowego<sup>13</sup>, zgodnie z preferencjami podmiotów gospodarczych [Mas-Colell, Whinston, Green, 1995]. W kontekście rynków finansowych można ten problem decyzyjny interpretować jako polegający na wyborze inwestycji w aktywo albo zbiór aktywów finansowych z pewnej listy dostępnych alternatyw. Zgodnie z tą teorią ewaluacja alternatywnych inwestycji polega na zestawieniu ze sobą oczekiwanej użyteczności wyniku finansowego wiążącego się z każdą z możliwych inwestycji, a następnie wybraniu spośród nich takiej, która generuje najwyższą oczekiwaną użyteczność [Mas-Colell i in., 1995].

Przyjmuje się, że uczestnik rynku jest gotowy wziąć udział w niepewnej co do wyniku finansowego inwestycji, jeśli wartość oczekiwana poziomu jego ostatecznego majątku przewyższa użyteczność pewnego (tzn. deterministycznego) poziomu majątku w przypadku rezygnacji z inwestycji. Warto zauważyć, że dziedzina funkcji użyteczności zawiera zatem jedynie ostateczną pozycję finansową decydenta.

<sup>12</sup> Tzn. zakupie pewnego aktywa albo portfela aktywów finansowych.

<sup>13</sup> Opisanego pewnym zdefiniowanym subiektywnie lub obiektywnie na zbiorze tzw. konsekwencji rozkładem prawdopodobieństwa, zwanym w literaturze ekonomii *loterią*.



Powszechnym założeniem występującym w literaturze ekonomii (w teorii oczekiwanej użyteczności), jest założenie mówiące o cechowaniu się przez podmioty gospodarcze awersją do ryzyka. Jest ona obrazowana przez nierówność Jensena [Mas-Colell i in., 1995]:

$$\int u(x)dF(x) \leq u\left(\int x dF(x)\right) \quad \forall_{F(\cdot)},$$

gdzie  $F: \mathbb{R} \rightarrow [0, 1]$  jest dowolną dystrybuantą rozkładu prawdopodobieństwa (finalnego) majątku. Preferencje spełniające tę nierówność cechują się dodatnią<sup>14</sup> różnicą pomiędzy oczekiwaną użytecznością ze zdegenerowanej loterii przynoszącej deterministyczny wynik będący średnią wartością pieniężną, jaką otrzymuje uczestnik loterii, a oczekiwaną użytecznością wynikającą z uczestnictwa w loterii. Innymi słowy, osoba cechująca się awersją do ryzyka woli zatrzymać opłatę za uczestnictwo w loterii (jeśli wynosi ona tyle, ile można średnio w niej wygrać), zamiast wziąć w niej udział, który wiązałby się z zaakceptowaniem ryzyka towarzyszącego nieznanemu wynikowi loterii (oraz niepożądaną przez decydenta potencjalnej straty). Taki osobnik ceni sobie zatem pewność i unika (czuje awersję do) sytuacji niepewnych lub ryzykownych [Kahneman, Tversky, 1979]. Nierówność Jensena definiuje wklęsłe funkcje użyteczności.

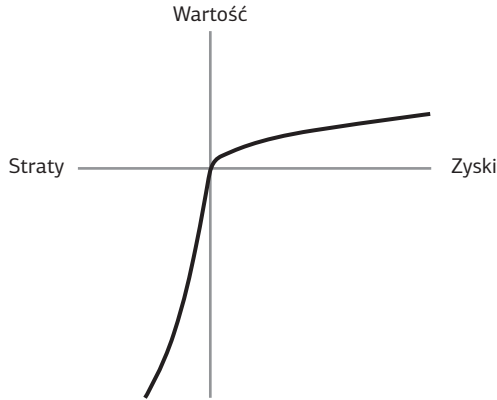
W przeciwieństwie do teorii oczekiwanej użyteczności, teoria perspektywy skupia się na relatywnym wobec obecnej w danym momencie pozycji finansowej decydenta (przed inwestycją) poziomie majątku po realizacji zmiennej losowej będącej wynikiem inwestycji [Kahneman, Tversky, 1979]. Dotyczy ona perspektywicznego spojrzenia na stan majątku decydenta, relatywnie wobec jego wysokości w okresie poprzednim. Uwzględnia zatem zmiany w majątku (zarówno straty, jak i zyski finansowe), a nie wysokość majątku finalnego. Dziedzina funkcji wartości będącej odpowiednikiem klasycznej funkcji użyteczności zawiera wobec tego zarówno wartości dodatnie, jak i ujemne.

W przeciwieństwie do funkcji użyteczności, która jest przeważnie wklęsła na całej swojej dziedzinie, krzywizna funkcji wartości różni się w zależności od tego, czy analizowane są straty, czy zyski finansowe. Funkcja ta jest monotonicznie rosnąca na całej swojej dziedzinie liczb rzeczywistych, jednakże jest ona wypukła w dziedzinie strat i wklęsła – w dziedzinie zysków. Dodatkowo jej nachylenie w dziedzinie strat jest większe niż w dziedzinie zysków.

<sup>14</sup> Dokładniej – nieujemną.

Poniższy wykres przedstawia przykładową funkcję wartości.

**Rysunek 1. Funkcja wartości**



Źródło: opracowanie własne na podstawie Kahneman, Tversky, 1979, rys. 3.

Takie zachowanie funkcji, która cechuje się występowaniem awersji do ryzyka w dziedzinie zysków oraz skłonnością do podejmowania ryzyka w dziedzinie strat, jest tłumaczone niechęcią do dalszego utrzymywania pozycji dłużych po osiągnięciu zysku – w związku ze strachem związanym z możliwą stratą, która ze względu na większą stromość funkcji w dziedzinie strat byłaby bardziej odczuwalna przez podmiot gospodarczy niż potencjalny zysk. Skłonność do dalszego utrzymywania pozycji dłużych po osiągniętej stracie wyjaśniana jest natomiast odwrotnie – występującym zwiększającym się znieczuleniem co do kolejnych krańcowych strat. Wiąże się to z przewagą krańcowej użyteczności z potencjalnego zysku nad krańcową użytecznością z potencjalnej straty. Przy założeniu równego prawdopodobieństwa osiągnięcia zysku i straty w każdym momencie czasu (zgodnie z procesem błędzenia losowego), oznacza to, iż uczestnik rynku będzie się obawiał utrzymania pozycji dłużej po uzyskaniu dodatniego zwrotu ze swojej inwestycji, ze względu na przeważającą negatywną użyteczność (*disutility*) związaną z potencjalną stratą. Będzie natomiast utrzymywał pozycję długą po uzyskaniu ujemnego zwrotu, co wynikać będzie z przeważającej (dodatniej) użyteczności związanej z potencjalnym zyskiem.

Po przedstawieniu teorii perspektywy, dokonana zostanie teraz próba przeanalizowania źródła nieracjonalnych oczekiwań uczestników rynku co do zmiany trendu cenowego. W kontekście teorii perspektywy uczestnik rynku będzie nadmiernie pesymistycznie podchodził do potencjalnej zmiany trendu ze wzrostowego na spadkowy (będzie subiektywnie przeważał prawdopodobieństwo tego zdarzenia w porównaniu do rzeczywiście świadczących o tym przesłanek), co sprawi, iż zaniecha

on inwestycji w takie aktywo, będąc przekonanym, że owa zmiana faktycznie nastąpi. Natomiast zbyt optymistyczne podejście będzie się wiązało z nadmiernym przekonaniem o zmianie trendu cenowego ze spadkowego na wzrostowy, będzie więc skutkowało podjęciem decyzji o inwestycji w takie aktywo.

W szczególności przeświadczenie to należy uważać za nieracjonalne ze względu na podparcie go jedynie o przeszłą historię cenową aktywa, bez przeprowadzenia jakiegokolwiek analizy ekonomicznej czy analizy wartości fundamentalnej [Taylor, Allen, 1992]. W literaturze finansów powszechnie przyjmuje się, iż ceny aktywów finansowych są procesem błędzącym losowo, empirycznie wykazującym zatem własność braku pamięci (*memorylessness* [np. Fama, 1995]), co wiąże się z kolei z brakiem mocy predykcyjnej przeszłych trendów cenowych – ze względu na niezależność cen w poszczególnych okresach. Fama [1995] podsumowuje wyniki badań empirycznych z dziedziny finansów behawioralnych, iż w większości przypadków zwracają one współczynniki korelacji cen w poszczególnych okresach statystycznie istotnie nieróżniące się od zera.

W celu pokazania źródła skłonności uczestników rynku do obstawiania utrzymywania się trendów rynkowych w najbliższej przyszłości, można posłużyć się wspomnianą wcześniej teorią *momentum* [Asness i in., 2013; Jegadeesh, Titman, 1993]. Strategia inwestycyjna oparta na zjawisku *momentum* bazuje na przekonaniu o utrzymywaniu się trendów cenowych w czasie. Wiarę w *momentum* można oprzeć na badaniu przeprowadzonym w pracy Jegadeesha i Titmana [1993]. Bada ono stopy zwrotu kilku portfeli rynkowych złożonych z akcji firm osiągających wcześniej dodatnie stopy zwrotu i, osobno, portfeli złożonych z akcji odnotowujących w poprzednich okresach zwroty ujemne. Wyniki badania wskazują na utrzymywanie się trendów cenowych w krótkim i średnim okresie, a więc dodatnich korelacji pomiędzy cenami aktywa w tych okresach, zmniejszających się w miarę wydłużania się okresu inwestycyjnego.

Przekonanie o istnieniu zjawiska *momentum* może skłaniać uczestników rynku do inwestowania w akcje, które przyniosły w poprzednich okresach dodatnie zwroty finansowe oraz unikania akcji, które wcześniej osiągały stopy ujemne. Decyzja inwestycyjna będzie wówczas oparta na nadmiernym optymizmie, a więc wierze w utrzymanie się dodatniego trendu cenowego aktywa w najbliższych okresach. W przypadku wiary w utrzymanie się ujemnego trendu cenowego, nadmierny, nieracjonalny pesymizm będzie wiązał się z niechęcią do zakupu (lub skłonnością do zajmowania pozycji krótkich) aktywów, które w poprzednich okresach przynosiły straty.

Również takie oczekiwanie można uważać za nieracjonalne, ponieważ zakłada ono zależność cen aktywów w czasie (a więc poprawność analizy technicznej). Nawet jeśli założy się niezerową korelację cen w czasie (co jest zgodne np. z wynikami

badania przeprowadzonego w pracy Jegadeesha, Titmana [1993]), problem polega na empirycznie wykazanej losowości długości okresu, w którym utrzymuje się dany trend [np. Fama, 1965]. Jeśli natomiast założyć się, że analizowane aktywo finansowe błędzi losowo, oznaczać to będzie, że utrzymujący się trend cenowy ulegnie odwróceniu w zupełnie losowych, niemożliwych do (systematycznego) przewidzenia momentach [Fama, 1995]. Warto tutaj wymienić występujące w literaturze finansów behawioralnych testy, które badają moment odwrócenia trendu cenowego (testy znaków (*run test* lub *sign test*) [np. Fama, 1965]). Wyniki tych badań przeważnie potwierdzają hipotezę mówiącą o błędzeniu losowym cen akcji.

Wśród przeprowadzonych w ostatnich latach badań dotyczących nieracjonalnych oczekiwań, można wyróżnić m.in. Chhabra, De [2012]. Analizuje ono wpływ wcześniejszych strat inwestycyjnych na przyszłe decyzje inwestycyjne. Postuluje występowanie większego wpływu znaku zwrotu finansowego na przekonanie o dochodowości inwestycji (dodatni – świadczy o korzystnym wyniku, ujemny – o wyniku niekorzystnym), niż wielkość zainwestowanych środków finansowych. Zatem fakt osiągnięcia dodatniego lub ujemnego wyniku finansowego z inwestycji (bez względu na odległość na osi liczb rzeczywistych pomiędzy wartością osiągniętego wyniku a zerem) powoduje relatywnie większą zmianę w skłonności do ponownego zainwestowania środków niż wielkość środków finansowych, które wygenerowały ów dodatni albo ujemny zwrot. Papadovasilaki, Guerrero, Sundali i Stone [2015] analizują wpływ wcześniejszych zysków i strat na następnie podejmowane decyzje inwestycyjne przez podmioty rynkowe. Przeprowadzone badania wykazały silną dodatnią korelację pomiędzy wcześniejszymi wynikami inwestycyjnymi i później podejmowanymi decyzjami związanymi z inwestycjami w aktywa finansowe.

#### **4. Analiza wpływu nieracjonalnych oczekiwań na stopy zwrotu na rynkach finansowych**

W poprzednich dwóch rozdziałach opisane zostały różne oczekiwania uczestników rynku oraz próby wyjaśnienia źródeł ich pochodzenia. Wymienione zostały także sposoby, w jakie uczestnicy rynku próbują uzasadniać swoje nieracjonalne oczekiwania. Naturalnie nasuwającym się pytaniem w tym momencie jest, czy istnienie nieracjonalnych oczekiwań graczy rynkowych wpływa w istotny sposób na wysokość stóp zwrotu na rynkach finansowych.

Literatura finansów behawioralnych dokumentuje zjawisko *momentum*, a więc występującej po chwili ukazania się nowych wiadomości powodujących zmiany wartości fundamentalnych aktywów finansowych, opóźnionej i/lub przedłużonej

reakcji cen [Lo, MacKinlay, 1990; Jegadeesh, Titman, 1993] oraz krótko- i średniookresowo występujące nadmierne reakcje rynku (na pojawiające się, ogólnodostępne nowe informacje) powodujące nadmierne zmiany cen akcji w stosunku do ich nowych wartości fundamentalnych, które są następnie zredukowane w długim okresie [De Bondt, Thaler, 1987].

De Bondt i Thaler [1987] badają zjawisko nadmiernej reakcji rynku i podsumowują wyniki badań i ankiet empirycznych, podając, iż wśród wielu inwestorów występuje tendencja do przykładania nadmiernej wagi do pojawiających się nowych informacji (dotyczących np. wyników finansowych firm). Powoduje ona następnie wśród tych uczestników rynku nieumiejętne aktualizowanie ich subiektywnych rozkładów prawdopodobieństwa stóp zwrotu z inwestycji, które skutkuje w występującym zbyt optymistycznym lub zbyt pesymistycznym konsensusie rynkowym. Oznacza to, iż na skutek pojawienia się nowych informacji, zmiany cen aktywów są większe w wartości bezwzględnej niż zmiany ich wartości fundamentalnych, powodowane tymi informacjami [De Bondt, Thaler, 1987]. W przypadku pojawienia się powszechnie dostępnych nowych pozytywnych informacji, a więc bardziej optymistycznych, niż wcześniej wskazywał na to konsensus rynkowy, zjawisko nadmiernej reakcji rynku spowoduje krótko- i średniookresowe zawyżenie cen aktywów w stosunku do ich nowych wartości fundamentalnych. Odwrotnie w przypadku nowych negatywnych informacji. Wyniki badań empirycznych wskazują na powodowane wówczas nadmiernym pesymizmem zjawisko nadmiernej reakcji rynku, które wiąże się z krótko- i średniookresowym zaniżeniem cen aktywów w stosunku do ich nowych wartości fundamentalnych [De Bondt, Thaler, 1987].

Badanie przeprowadzone przez De Bondta i Thalara [1987] zdaje się potwierdzać występowanie zjawiska nadmiernej reakcji rynku. Postuluje ono, iż po nadmiernych reakcjach rynku następuje rewizja oczekiwań i (zgodnie z HIERF) powrót cen do ich nowych wartości fundamentalnych. W owym badaniu zbudowane zostały dwa portfele, każdy z nich złożony z 50 aktywów firm notowanych na nowojorskiej giełdzie papierów wartościowych. W pierwszym portfelu znajdowały się akcje firm, które w latach 1926–1982 osiągnęły najwyższe stopy zwrotu, natomiast w drugim – te, które osiągnęły w tym okresie największe straty. Wyniki wskazują, iż portfel złożony z wcześniej najbardziej stratnych aktywów osiągnął średnio o 31,9% wyższe stopy zwrotu, niż portfel złożony w wcześniej najbardziej zyskownych aktywów.

Zjawisko *momentum* mówi natomiast o utrzymujących się w czasie trendach cenowych powodowanych pojawiającymi się na rynkach nowymi informacjami. Z jednej strony jest ono wyjaśniane występującym wśród inwestorów opóźnieniem w analizie nowo pojawiających się informacji, a zatem również opóźnionymi reakcjami (decyzjami inwestycyjnymi) uczestników rynkowych [Lo, MacKinlay, 1990].

Z drugiej strony może być spowodowane również (nieracjonalnym) przekonaniem części podmiotów gospodarczych o utrzymaniu się w najbliższym czasie trendów wzrostowych lub spadkowych [Asness i in., 2013]. Jeśli grupa uczestników rynkowych przekonana o takim zachowaniu ceny danego aktywa rynkowego jest wystarczająco duża, w agregacie będą oni mieli widoczny wpływ na jego cenę<sup>15</sup>. Jeśli na rynku finansowym występuje liczna grupa inwestorów przekonanych (zbyt optymistycznie) o utrzymaniu się dodatniego trendu cenowego, przez swoje decyzje inwestycyjne o zakupie danego aktywa finansowego przyczynią się oni rzeczywiście do dalszego utrzymania tego trendu. Z drugiej strony występowanie dużej grupy inwestorów cechujących się nadmiernie pesymistycznym przekonaniem o utrzymaniu się ujemnego trendu cenowego spowoduje, iż owi uczestnicy rynku swoimi decyzjami o braku inwestycji lub poprzez obstawianie pozycji krótkich, przyczynią się do faktycznego utrzymania takiego trendu.

Wśród najnowszych badań wpływu nieracjonalnych oczekiwań na stopy zwrotu na rynkach finansowych, wyróżnia się w szczególności temat tzw. zachowania stadnego inwestorów (*investors' herding*). Badania tego zjawiska szukają źródeł odchyleń cen aktywów finansowych od ich wartości fundamentalnych w opieraniu decyzji indywidualnych inwestorów w większym stopniu na występujących na rynkach wspólnych informacjach i oczekiwaniach (niekoniecznie racjonalnych), niż na posiadanych przez nich prywatnych informacjach i oczekiwaniach [Zahera, Bansal, 2018; Shiller, 2000; Kahneman, Tversky, 1979]. Ze względu na charakteryzującą zjawisko zachowania stadnego powszechność takiego samego sposobu budowania indywidualnych oczekiwań wśród inwestorów, ceny aktywów finansowych odchylają się bardziej od swoich wartości fundamentalnych, niż np. gdyby w oparciu o nie decyzje podejmował relatywnie niewielki udział uczestników rynku.

Feldman i Lepori [2016] dzielą inwestorów posiadających nieracjonalne oczekiwania na dwie grupy. U pierwszej, którą nazywa inwestorami introspekcyjnymi (*introspective investors*), nieracjonalne oczekiwania mają źródło we własnych wynikach finansowych i w przekonaniu o własnej nieomyślności (tłumaczonej np. przy pomocy zjawiska *overconfidence*). Drugą grupę nazywa inwestorami agregującymi (*aggregating investors*). Charakteryzuje ich jako bazujących swoje nieracjonalne oczekiwania na kolektywnych wynikach finansowych innych inwestorów (co jest dokumentowane w literaturze finansów behawioralnych np. jako zachowania stadne). Na podstawie modelowania rynku finansowego z wykorzystaniem podmiotów rynkowych podzielonych na racjonalnych inwestorów, grupę inwestorów

---

<sup>15</sup> Wiszniewska-Matyszkiewicz [2008] nazywa to zjawisko samospełniającymi się przekonaniem (*self-verifying beliefs*).

nieracjonalnych (posiadających nieracjonalne oczekiwania) oraz na grupę inwestorów posiadających oczekiwania będące mieszanką dwóch poprzednich grup, z uwzględnieniem dodatkowego podziału inwestorów posiadających nieracjonalne oczekiwania na inwestorów introspekcyjnych i agregujących, Feldman i Lepori [2016] sugerują, iż jedynie inwestorzy agregujący, a nie natomiast inwestorzy introspektywni, mają istotny wpływ na stopy zwrotu na rynkach finansowych. Zatem zgodnie z tymi wynikami jedynie ta pierwsza grupa powinna być uwzględniana w modelach zawierających rozróżnienie podmiotów rynkowych ze względu na posiadane przez nich różne rodzaje oczekiwań.

## 5. Podsumowanie

W tym artykule skoncentrowano się na analizie implikacji odrzucenia założenia racjonalności oczekiwań graczy rynkowych, przyjmowanego standardowo w modelach badających zjawiska ekonomiczne. Dokonana została próba opisu występujących w literaturze finansów behawioralnych najważniejszych powodów utrzymywania przez graczy rynkowych nieracjonalnych oczekiwań wobec kształtowania się cen aktywów na rynkach finansowych. Omówione zostały również te różne oczekiwania i sposoby tłumaczenia źródeł ich pochodzenia opisane w literaturze klasycznej oraz współczesnej przedmiotu. Na końcu został przedstawiony wpływ, jaki mogą mieć nieracjonalne oczekiwania na stopę zwrotu aktywów finansowych.

Przedstawiony w tym artykule przegląd literatury dotyczącej heurystyk i obciążenia oczekiwań graczy rynkowych wskazuje na duży potencjał badawczy analizy nieracjonalnych oczekiwań. W stosunku do powszechnie zakładanej w modelach racjonalności oczekiwań podmiotów gospodarczych, nieracjonalność oczekiwań powoduje bowiem mniejszą przewidywalność decyzji graczy rynkowych. Modele ekonomiczne uwzględniające tę heterogeniczność oczekiwań umożliwiają formułowanie wniosków na podstawie badania zachowania zarówno graczy rynkowych stosujących regułę Bayesa do aktualizacji swoich subiektywnych oczekiwań dotyczących wysokości przyszłych stóp zwrotu na rynkach finansowych, jak i tych graczy, którzy formułują swoje oczekiwania na różne inne sposoby, ulegając licznym heurystykom i zniekształceniom poznawczym (jak np. te opisane w niniejszym artykule).

Jeśli na rynkach finansowych inwestuje wystarczająco liczna grupa podmiotów gospodarczych kierujących się nieracjonalnymi oczekiwaniami (np. tymi samymi strategiami inwestycyjnymi z analizy technicznej), będą one systematycznie odzwierciedlone w cenach aktywów [Wiszniewska-Matyszkiewicz, 2008]. Nasuwa się wów-

czas pytanie dotyczące właściwego postępowania gracza rynkowego, który chciałby adekwatnie uwzględnić wpływ graczy posiadających nieracjonalne oczekiwania na stopy zwrotu aktywów, w celu maksymalizacji swoich wyników finansowych.

Gracz rynkowy posiadający racjonalne oczekiwania powinien być świadomy występowania wśród inwestorów zarówno graczy posiadających racjonalne, jak i nieracjonalne oczekiwania. Racjonalne oczekiwania powinny uwzględniać bowiem wszystkie oczekiwania wpływające na wysokość przyszłych stóp zwrotu: zarówno te uzasadnione wartością fundamentalną analizy przyszłych poziomów cen aktywów, własne oczekiwania dotyczące pojawiających się w przyszłości nowych informacji wpływających na ceny aktywów oraz oczekiwania graczy rynkowych co do oczekiwań pozostałych graczy rynkowych, dotyczących przyszłych poziomów cen aktywów finansowych<sup>16</sup>.

Decyzja inwestycyjna *homo oeconomicusa* podjęta jedynie w oparciu o kalkulację wartości fundamentalnej, przy założeniu racjonalności oczekiwań wszystkich pozostałych uczestników rynkowych (oparta na założeniu, iż oni wszyscy szacują w podobny sposób wartość fundamentalną, kierując się racjonalnymi oczekiwaniami), może okazać się wówczas sama w sobie nieracjonalna (ignorowałaby wówczas obecność licznej grupy inwestorów, którzy posiadają nieracjonalne oczekiwania i nie korzystają z reguły Bayesa lub nie uwzględniają oczekiwań innych uczestników rynku przy podejmowaniu własnych decyzji inwestycyjnych).

Nieracjonalne oczekiwania mogą powodować krachy na giełdzie, jeśli trend spadkowy cen utrzyma się przez wystarczająco długi czas, w szczególności, jeśli jest on wzmacniany nadmiernym pesymizmem wystarczająco licznej grupy inwestorów [Taylor, Allen, 1992; Wiszniewska-Matyszek, 2008]. W przyszłych badaniach dotyczących nieracjonalnych oczekiwań oraz przy konstruowaniu modeli ekonomicznych próbujących uwzględnić występowanie graczy rynkowych posiadających nieracjonalne oczekiwania, warto zastanowić się nad znalezieniem sposobów przekonania takich graczy do rozpoczęcia korzystania z reguły Bayesa przy aktualizacji rozkładu prawdopodobieństwa przyszłych stóp zwrotu oraz do adekwatnego uwzględniania wszystkich nowych informacji przy podejmowaniu decyzji inwestycyjnych (a nie tylko takich, które np. potwierdzają poprawność przekonań decydenta ulegającego błędowi konfirmacji).

Jeśli brak korzystania przez niektórych graczy rynkowych z reguły Bayesa jest ugruntowany w uleganiu heurystykom i zniekształceniom poznawczym lub ignorowaniu wszelkich informacji przeciwnych do posiadanych przez dany podmiot go-

---

<sup>16</sup> Tzw. *sentymenty rynkowe, intencjonalność* lub *przekonania drugiego rzędu (second order beliefs)* [np. Li, Shah, 2017].



spodarczy przekonani (a nie w ograniczonej umiejętności przetwarzania informacji), przekonanie takiego gracza o korzyściach finansowych (np. wyższym oczekiwanym zwrotem finansowym) wiążących się z posiadaniem racjonalnych oczekiwań mogłoby zmniejszyć prawdopodobieństwo drastycznych spadków cen na giełdzie i przyczynić się do skrócenia oczekiwanego czasu trwania recesji gospodarczych<sup>17</sup>.

## Bibliografia

---

- Asness, C.S., Moskowitz, T.J., Pedersen, L.H. (2013). Value and Momentum Everywhere, *The Journal of Finance*, 68(3), s. 929–985. DOI: 10.1111/jofi.12021.
- Camerer, C., Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach, *American Economic Review*, 89(1), s. 306–318. DOI: 10.1257/aer.89.1.306.
- Chhabra, R., De, S. (2012). Are Out-of-Pocket Costs Overweighted Relative to Opportunity Costs? A Disposition Effect-Based Investigation. A Disposition Effect-Based Investigation (January 15, 2012). DOI: 10.2139/ssrn.2022986.
- Cohen, L.J. (1981). Can Human Irrationality Be Experimentally Demonstrated?, *Behavioral and Brain Sciences*, 4(3), s. 317–331. DOI: 10.1017/S0140525X00009092.
- Cournot, A.A. (1838). *Research on the Mathematical Principles of Wealth Theory*. L. Hachette.
- Czerwonka, M., Gorlewski, B. (2012). *Finanse behawioralne*. Warszawa: Oficyna Wydawnicza SGH.
- Dave, C. (2011). Are Investment Expectations Rational, Adaptive or Regressive?, *Economic Inquiry*, 49(1), s. 212–225. DOI: 10.1111/j.1465-7295.2010.00305.x.
- De Bondt, W.F., Thaler, R.H. (1987). Further Evidence on Investor Overreaction and Stock Market Seasonality, *The Journal of Finance*, 42(3), s. 557–581. DOI: 10.1111/j.1540-6261.1987.tb04569.x.
- Dominitz, J., Manski, C.F. (2010). Measuring and Interpreting Expectations of Equity Returns, *Journal of Applied Econometrics*, 26(3), s. 352–370. DOI: 10.1002/jae.1225.
- Fama, E.F. (1960). Efficient Market Hypothesis (doctoral thesis).
- Fama, E.F. (1965). The Behavior of Stock-Market Prices, *The Journal of Business*, 38(1), s. 34–105.
- Fama, E.F. (1976). *Foundations of Finance*. Basic books.
- Fama, E.F. (1995). Random Walks in Stock Market Prices, *Financial Analysts Journal*, 51(1), s. 75–80. DOI: 10.2469/faj.v51.n1.1861.
- Fama, E.F. (2021). Efficient Capital Markets a Review of Theory and Empirical Work. W: *The Fama Portfolio – Selected Papers of Eugene F. Fama* (s. 76–121), J.H. Cochrane, T.J. Moskowitz (Eds.). University of Chicago Press.
- Feldman, T., Lepori, G. (2016). Asset Price Formation and Behavioral Biases, *Review of Behavioral Finance*, 8(2), s. 137–155. DOI: 10.1108/RBF-05-2015-0020.

<sup>17</sup> Np. hipoteza Hayeka [Smith, 1982] postuluje, iż rynek może funkcjonować efektywnie nawet jeśli większość uczestników rynkowych posiada nieracjonalne oczekiwania.

- Fischhoff, B., Beyth, R. (1975). I Knew It Would Happen: Remembered Probabilities of Once – Future Things, *Organizational Behavior and Human Performance*, 13(1), s. 1–16. DOI: 10.1016/00305073(75)90002 1.
- Forsythe, R., Nelson, F., Neumann, G.R., Wright, J. (1992). Anatomy of an Experimental Political Stock Market, *The American Economic Review*, s. 1142–1161.
- Frankel, J.A., Froot, K.A. (1986). Understanding the US Dollar in the Eighties: The Expectations of Chartists and Fundamentalists, *Economic Record*, 62(1), s. 24–38.
- Froot, K.A., Frankel, J.A. (1989). Forward Discount Bias: Is It an Exchange Risk Premium?, *The Quarterly Journal of Economics*, 104(1), s. 139–161. DOI: 10.2307/2937838.
- Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K. (2007). Brain Functional Localization: A Survey of Image Registration Techniques, *IEEE Transactions on Medical Imaging*, 26(4), s. 427–451. DOI: 10.1109/TMI.2007.892508.
- Haltiwanger, J., Waldman, M. (1985). Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity, *The American Economic Review*, 75(3), s. 326–340.
- Harsanyi, J.C. (1980). Advances in Understanding Rational Behavior. W: *Essays on Ethics, Social Behavior, and Scientific Explanation* (s. 89–117), G. Eberlein, W. Leinfellner (Eds.). Dordrecht: Springer. DOI: 10.1007/978-94-010-9327-9\_6.
- Heath, C., Tversky, A. (1991). Preference and Belief: Ambiguity and Competence in Choice under Uncertainty, *Journal of Risk and Uncertainty*, 4(1), s. 5–28. DOI: 10.1007/BF00057884.
- Hilton, D.J. (2001). The Psychology of Financial Decision-Making: Applications to Trading, Dealing, and Investment Analysis, *The Journal of Psychology and Financial Markets*, 2(1), s. 37–53. DOI: 10.1207/S15327760JPFM0201\_4.
- Jegadeesh, N., Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *The Journal of Finance*, 48(1), s. 65–91. DOI: 10.1111/j.1540-6261.1993.tb04702.x.
- Kahneman, D. (1994). New Challenges to the Rationality Assumption, *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 150(1), s. 18–36.
- Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, (47)2, s. 263–292.
- Kim, K.A., Nofsinger, J.R. (2008). Behavioral Finance in Asia, *Pacific-Basin Finance Journal*, 16 (1–2), s. 1–7. DOI: 10.1016/j.pacfin.2007.04.001.
- Langer, E.J. (1975). The Illusion of Control, *Journal of Personality and Social Psychology*, 32(2), s. 311. DOI: 10.1037/0022-3514.32.2.311.
- Langer, E.J., Roth, J. (1975). Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task, *Journal of Personality and Social Psychology*, 32(6), s. 951. DOI: 10.1037/0022-3514.32.6.951.
- Li, Q., Shah, S. (2017). Learning Stock Market Sentiment Lexicon and Sentiment-Oriented Word Vector from Stocktwits. W: *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)* (s. 301–310). DOI: 10.18653/v1/K17-1031.
- Lo, A.W., MacKinlay, A.C. (1990). When Are Contrarian Profits Due to Stock Market Overreaction?, *The Review of Financial Studies*, 3(2), s. 175–205. DOI: 10.1093/rfs/3.2.175.
- Malkiel, B. (1985). *A Random Walk Down Wall Street* (4th ed.). New York: Norton.

- Mankiw, N.G. (1990). A Quick Refresher Course in Macroeconomics. Working Paper 3256, National Bureau of Economic Research. DOI: 10.3386/w3256.
- Manski, C.F. (2004). Measuring Expectations, *Econometrica*, 72(5), s. 1329–1376. DOI: 10.1111/j.1468-0262.2004.00537.x.
- Mas-Colell, A., Whinston, M.D., Green, J.R. (1995). *Microeconomic Theory* (vol. 1). New York: Oxford University Press.
- Moore, D.A., Healy, P.J. (2008). The Trouble with Overconfidence, *Psychological Review*, 115(2), s. 502. DOI: 10.1037/0033-295X.115.2.502.
- Morris, S. (1995). The Common Prior Assumption in Economic Theory, *Economics & Philosophy*, 11(2), s. 227–253. DOI: 10.1017/S0266267100003382.
- Odean, T. (1999). Do Investors Trade Too Much?, *American Economic Review*, 89(5), s. 1279–1298. DOI: 10.1257/aer.89.5.1279.
- Papadovasilaki, D., Guerrero, F., Sundali, J., Stone, G. (2015). How Important Are Early Investment Experiences on Subsequent Investment Decisions? A Laboratory Experiment on Asset Allocation, *Managerial Finance* 41(6), s. 582–590.
- Plous, S. (1993). *The Psychology of Judgment and Decision Making*. McGraw-Hill Book Company.
- Ross, L., Greene, D., House, P. (1977). The “False Consensus Effect”: An Egocentric Bias in Social Perception and Attribution Processes, *Journal of Experimental Social Psychology*, 13(3), s. 279–301. DOI: 10.1016/0022-1031(77)90049-X.
- Rothbart, M., Snyder, M. (1970). Confidence in the Prediction and Postdiction of an Uncertain Outcome, *Canadian Journal of Behavioural Science/ Revue canadienne des sciences du comportement*, 2(1), s. 38. DOI: 10.1037/h0082709.
- Sadi, R., Asl, H.G., Rostami, M.R., Gholipour, A., Gholipour, F. (2011). Behavioral Finance: The Explanation of Investors’ Personality and Perceptual Biases Effects on Financial Decisions, *International Journal of Economics and Finance*, 3(5), s. 234–241. DOI: 10.5539/ijef.v3n5p234.
- Sharpe, W.F. (1985). *Investments* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Shiller, R. (1987). Investor Behavior in the October 1987 Stock Market Crash: Survey Evidence. Working Paper 2446, National Bureau of Economic Research. DOI: 10.3386/w2446.
- Shiller, R.C. (2000). Irrational Exuberance, *Philosophy and Public Policy Quarterly*, 20(1), s. 18–23.
- Smith, V.L. (1982). Markets as Economizers of Information: Experimental Examination of the “Hayek Hypothesis”, *Economic Inquiry*, 20(2), s. 165–179. DOI: 10.1111/j.1465-7295.1982.tb01149.x.
- Soll, J.B., Klayman, J. (2004). Overconfidence in Interval Estimates, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), s. 299. DOI: 10.1037/0278-7393.30.2.299.
- Stephan, E., Kiell, G. (2017). Decision Processes in Professional Investors: Does Expertise Guard against Judgmental Biases?, *Zeitschrift für Arbeitswissenschaft*, 71(3), s. 142–156. DOI: 10.1007/s41449-017-0061-z.
- Taskinsoy, J. (2019). Bitcoin: The Longest Running Mania – Tulips of the 21st Century?, *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3505953.
- Taylor, M.P., Allen, H. (1992). The Use of Technical Analysis in the Foreign Exchange Market, *Journal of International Money and Finance*, 11(3), s. 304–314. DOI: 10.1016/0261-5606(92)90048-3.

- Tversky, A., Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, 185(4157), s. 1124–1131. DOI: 10.1126/science.185.4157.1124.
- Wason, P.C., Johnson-Laird, P.N. (1972). *Psychology of Reasoning: Structure and Content* (vol. 86). Harvard University Press.
- Wiszniewska-Matyszek, A. (2008). Stock Market as a Dynamic Game with Continuum of Players, *Control and Cybernetics*, 37(3), s. 617–647.
- Zahera, S.A., Bansal, R. (2018). Do Investors Exhibit Behavioral Biases in Investment Decision Making? A Systematic Review, *Qualitative Research in Financial Markets*, 10(2), s. 210–251. DOI: 10.1108/QRFM-04-2017-0028.

## Irrational Expectations in Financial Markets. Behavioural Finance Literature Review

### Summary

The article examines the implications of relaxed assumption of rational expectations, which is a standard assumption imposed in models analysing economic phenomena. Based on the literature on behavioural finance, the article attempts to investigate the reasons why rationally behaving market players have irrational expectations when making investment decisions; it also characterises their heterogeneous expectations and examines the impact of irrational expectations on the final rates of return in financial markets. The described phenomena of cognitive distortions and biases resulting in biased expectations, and sources of irrational expectations and their universality, reflected in their documented significant impact on the rate of return in financial markets, suggest that fully rational market players should take them into account when developing their rational expectations or making investment decisions. On the other hand, market players subconsciously succumbed to any of the described heuristics have an opportunity to revise their expectations and try to control the impact of these cognitive distortions on their investment decisions.

**Keywords:** irrational expectations, prospect theory, behavioural bias, momentum investing, overconfidence

# Zakres oraz determinanty ujawnień dotyczących wartości firmy w sprawozdaniach finansowych według MSR/MSSF – systematyczny przegląd literatury

Michał Siczek

## Streszczenie

Celem rozdziału jest dokonanie krytycznej analizy literatury dotyczącej wartości firmy oraz testów na utratę wartości według MSR/MSSF. Przeanalizowano 18 tekstów naukowych wyodrębnionych za pomocą metody naukowej PRISMA. W szczególności niniejszy rozdział dotyczy zakresu ujawnień dotyczących wartości firmy w sprawozdaniach finansowych oraz determinant tychże ujawnień i skutków w odniesieniu do rynków kapitałowych. Analiza artykułów naukowych zakończona jest dyskusją wyników, w której okazało się, że zakres ujawnień dotyczących wartości firmy w sprawozdaniach finansowych, mimo obowiązujących międzynarodowych regulacji rachunkowości, jest wąski i niekompletny. Ponadto w niniejszym opracowaniu wskazano potencjalne obszary, które stanowią obiekt dalszych badań, tj. porównanie zakresu ujawnień dotyczących wartości firmy pomiędzy poszczególnymi kontynentami czy opracowanie wskaźnika opisującego ich zakres i szczegółowość.

**Słowa kluczowe:** wartość firmy, ujawnienia, test na utratę wartości firmy, rachunkowość, sprawozdanie finansowe

## 1. Wstęp

Wartość firmy jest jednym z najważniejszych oraz najbardziej kontrowersyjnych tematów we współczesnej rachunkowości i sprawozdawczości finansowej. Jedną z pierwszych definicji wartości firmy pochodzi z 1810 r.: „Wartość firmy, która jest przedmiotem sprzedaży, jest niczym innym, jak prawdopodobieństwem, że starzy klienci powrócą na stare miejsce nawet wtedy, gdy starego kupca już nie będzie” [Kamela-Sowińska, 1996, s. 14]. Ujęcie to zatem odnosi się do wartości firmy jako do czynnika gwarantującego niezawodność przedsiębiorstwa oraz zapewniającego lojalność jego klientów.

Zgodnie z międzynarodowymi regulacjami w zakresie rachunkowości, wartość firmy jest definiowana jako „składnik aktywów reprezentujący przyszłe korzyści ekonomiczne powstające z aktywów nabytych w ramach połączenia jednostek, których nie można pojedynczo zidentyfikować ani osobno ująć” [IFRS, 2022a]. Należy zauważyć, że Międzynarodowe Standardy Rachunkowości/ Międzynarodowe Standardy Sprawozdawczości Finansowej (MSR/MSSF) są obecnie stosowane w 166 jurysdykcjach. Niektóre z państw zamiast implementacji regulacji zawartych w MSR/MSSF dostosowały swoje krajowe standardy rachunkowości w taki sposób, aby były z nimi zgodne (np. Chiny, Australia) [IFRS, 2022b].

Unifikacja standardów rachunkowości z pewnością przyczynia się do zapewnienia porównywalności sprawozdań finansowych na całym świecie, co jest szczególnie ważne dla inwestorów. Są oni zainteresowani ujawnieniami informacyjnymi zawartymi w tych dokumentach. To właśnie te dane są użyteczne dla inwestorów oraz innych podmiotów żywo zainteresowanych sprawozdawczością finansową podmiotów gospodarczych [IFRS, 2022b].

Ujawnienia informacyjne dotyczące wartości firmy oraz odpisów z tytułu utraty wartości firmy zawarte są w MSR 36 *Utrata wartości aktywów*, MSR 38 *Wartości niematerialne* oraz MSSF 3 *Połączenia jednostek gospodarczych*. Proces przeprowadzania testów na utratę wartości firmy zastąpił zasady rachunkowości, które nakazywały dokonywanie systematycznych odpisów amortyzacyjnych wartości firmy w czasie.

Nowe rozwiązanie, czyli konieczność przeprowadzania testów na utratę wartości firmy oraz ujawniania informacji z tego tytułu, zostało wprowadzone od 1 stycznia 2005 r. dla skonsolidowanych sprawozdań finansowych. Zatem w miejsce powtarzalnego rachunkowego procesu, w którym osądy i szacunki były w znacznym stopniu ograniczone, wprowadzono metodę opartą w głównej mierze właśnie na subiektywnych założeniach oraz osądach menedżerów. Z tego względu wartość firmy w sprawozdaniach finansowych może być podatna na manipulacje ze strony kadry zarządzającej [Giner, Pardo, 2015, s. 36–37; Jordan, Clark, 2015, s. 166].

Celem tego rozdziału jest analiza tekstów naukowych w odniesieniu do ujawnień dotyczących wartości firmy oraz odpisów z tytułu utraty wartości firmy. Ponadto tekst ten jest próbą podsumowania głównych determinant ujawnień dotyczących wartości firmy, a także ich ewentualnych skutków w odniesieniu do rynków kapitałowych.

Sformułowano również następujące pytania badawcze:

- 1) Czy przedsiębiorstwa sporządzające sprawozdania finansowe zgodnie z MSSF spełniają wymogi ujawnieniowe dotyczące wartości firmy i odpisów z tytułu utraty wartości firmy?
- 2) Jakie są główne czynniki, które wpływają na zakres ujawnień informacji w sprawozdaniach finansowych według MSSF dotyczących wartości firmy oraz odpisów z tytułu utraty wartości firmy?

W pierwszej części rozdziału przedstawiono metodę badawczą wraz z prezentacją jej kolejnych etapów. W kolejnej sekcji zaprezentowano wyniki badań. Została ona podzielona na dwie części. W pierwszej z nich opisany został aspekt zgodności sprawozdań finansowych przedsiębiorstw w kontekście wartości firmy z MSR/MSSF, natomiast w drugiej skupiono się na determinantach oraz skutkach ujawnień dotyczących wartości firmy w sprawozdaniach finansowych. Dyskusja wyników umożliwiła krytyczne podsumowanie efektów badań. Natomiast w podsumowaniu wskazano potencjalne obszary, które mogą stanowić obiekt dalszych badań nad wartością firmy.

## 2. Metoda badawcza

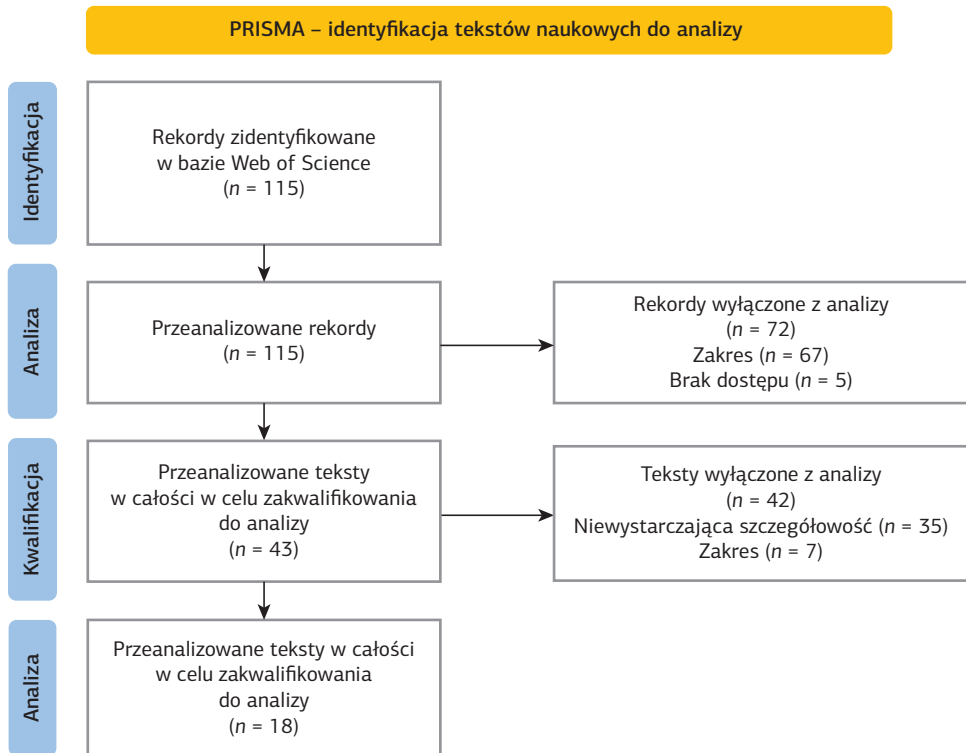
PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) jest metodą badawczą używaną szczególnie często w opracowywaniu systematycznego przeglądu literatury. Składa się z kilku etapów. Pierwszym z nich jest odpowiedni dobór słów kluczowych, które zostaną użyte do przeprowadzenia wyszukiwania w bazie tekstów naukowych. Zidentyfikowane rekordy są następnie analizowane i kwalifikowane pod względem zgodności z tematem badawczym. Eliminacja poszczególnych rekordów może nastąpić z przyczyny niewystarczającego zakresu lub braku dostępu do pełnego tekstu naukowego. Następnie teksty zakwalifikowane do kolejnego etapu są dogłębnie analizowane. Część z nich może być odrzucona z różnorodnych powodów, np. niewystarczającej szczegółowości. Ostatecznie grupa zakwalifikowanych tekstów naukowych na tym etapie jest analizowana. Analiza polega na wyodrębnieniu kluczowych wyników badań i podsumowaniu ich w systematycznym przeglądzie literatury.

W niniejszym rozdziale zdecydowano o użyciu metody PRISMA, ponieważ charakteryzuje się ona wysoką skutecznością oraz jakością. Ustrukturyzowany schemat działania pozwala na efektywne przeprowadzenie analizy tekstów naukowych. Ponadto dzięki niej czytelnicy mają możliwość dokonania oceny mocnych i słabych stron przeprowadzonego systematycznego przeglądu literatury, a także są zaznajomieni z jego ograniczeniami.

### 3. Wyniki

Przegląd literatury został wykonany za pomocą metody PRISMA. Przeszukano bazę Web of Science, gdzie jedynym kryterium były słowa kluczowe w związku AND: “goodwill” (wartość firmy), “disclosure” (ujawnienie), “accounting” (rachunkowość). Wyszukiwanie nie zostało ograniczone do konkretnych pól (np. tytuł, abstrakt itp.).

**Rysunek 1.** Wyniki systematycznego przeglądu literatury z wykorzystaniem diagramu PRISMA



Źródło: opracowanie własne.



Dzięki założeniom początkowym udało się wyodrębnić 115 tekstów, z których ostatecznie do analizy wybrano 18. Pozostałe z nich nie spełniały wymogów dotyczących zakresu badań lub nie zapewniały wystarczającej pojemności informacyjnej, która przewidziana została w celu głównym niniejszego opracowania. Szczegóły przedstawiono na rysunku 1.

W tej sekcji opisane zostaną wyniki analizy zakwalifikowanych dzięki PRISMA 18 tekstów naukowych. Podzielono ją na dwie części. W pierwszej z nich opisany zostanie aspekt zgodności sprawozdań finansowych z wymogami ujawnieniowymi MSR/MSSF w kontekście wartości firmy, natomiast druga część ma na celu sprawdzenie głównych determinant oraz skutków ujawnień dotyczących wartości firmy w sprawozdaniach finansowych według MSR/MSSF.

### 3.1. Zgodność z wymogami ujawnieniowymi MSR/MSSF

Jedno z badań, które skoncentrowało się na przesłedzeniu ujawnień dotyczących wartości firmy dotyczyło 36 przedsiębiorstw notowanych na giełdzie w Malezji (z indeksu FTSE Bursa Malaysia Index), które po raz pierwszy zastosowały MSR 36 w 2006 r. Okazało się, że 21 firm spełniła wymogi techniczne dotyczące ujawniania informacji, natomiast pozostałe 15 z nich nie wywiązało się należycie z tego obowiązku [Carlin, Finch, Laili, 2009, s. 87].

Badanie sprawdzające zgodność raportów finansowych z wymogami ujawnieniowymi przeprowadzone na próbie 50 spółek giełdowych z Australii wykazało, że w 2006 oraz 2007 r. odpowiednio 43 oraz 45 firm przyporządkowało wartość firmy do ośrodków generujących środki pieniężne. Ponadto większość spółek, tj. 39 w 2006 r. oraz 40 w 2007 r. ujawniło, że metodą wykorzystaną do oszacowania wartości odzyskiwalnej wartości firmy było oszacowanie wartości użytkowej. Natomiast w badanych latach odpowiednio trzy i cztery przedsiębiorstwa nie ujawniły użytej metody. Ponadto odnośnie do stopy dyskontowej użytej do oszacowania przepływów pieniężnych, większość spółek (23 w 2006 oraz 2007 r.) użyło konkretnej wartości, odpowiednio 6 i 9 użyło przedziału stóp dyskontowych, a odpowiednio 6 i 5 spółek nie ujawniło informacji na ten temat [Carlin, Finch, 2010 s. 270–272]. Kwestię ujawnień dotyczących wartości firmy australijskich spółek przeanalizowano także w innym badaniu, w którym wzięto pod uwagę 164 spółki oraz ich sprawozdania finansowe w 2006 r. Okazało się, że przedsiębiorstwa różniły się stopniem ujawnień co do testów na utratę wartości firmy oraz w znaczącej większości nie były zgodne z wymogami ujawnieniowymi wymaganymi przez MSR 36. Aż 111 spółek nie ujawniło w swoich sprawozdaniach finansowych informacji o stopie wzrostu użytej w testach na utratę wartości [Carlin, Finch, 2011, s. 386–387].

Wyniki badań przeprowadzonych także na spółkach australijskich wykazały, że zakres ujawnień dotyczących szacunków i osądów zastosowanych w testach na utratę wartości firmy zgodnie z MSR 36 jest niejednorodny. Przede wszystkim zależy on od branży, w jakiej działa jednostka, oraz od fazy cyklu życia, w której się znajduje. Okazało się, że poziom ujawnień w jednostkach, które są w środku cyklu życia przedsiębiorstwa, jest bardziej szczegółowy w porównaniu do jednostek, które są w początkowych oraz końcowych fazach cyklu życia. Powyższa prawidłowość ma większe zastosowanie w odniesieniu do mniejszych firm [Kabir, Su, Rahman, 2020, s. 20–21].

Podobne wnioski do tych zaprezentowanych przez Carlina i Fincha [2011] zostały wykazane w innym badaniu, gdzie okazało się, że przedsiębiorstwa w odniesieniu do wartości firmy stosują różne modele wyceny, a stopy dyskontowe używane do oszacowania przepływów pieniężnych znacznie się różnią pomiędzy poszczególnymi podmiotami. Ponadto w tym badaniu okazało się, że niektóre firmy mogą manipulować szacunkami i założeniami dotyczącymi wartości firmy i odpisów z tytułu utraty wartości, w celu uniknięcia konieczności zmniejszenia bilansowej wartości firmy [Comiskey, Mulford, 2010, s. 765].

Wyniki badań przeprowadzonych na przedsiębiorstwach wchodzących w skład indeksu S&P Europe 350 w latach 2008–2011 wykazały, że przedsiębiorstwa te znacznie różnią się zakresem ujawnień dotyczących wartości firmy w swoich sprawozdaniach finansowych. Duże rozbieżności zostały zanotowane w przypadku ujawnień dotyczących osądów zarządu oraz oczekiwań co do przyszłości, tj. oszacowania wartości użytkowej, wartości odpisów z tytułu utraty wartości firmy w podziale na segmenty operacyjne oraz ilościowego opisu dotyczącego wartości firmy rozpoznawanej podczas fuzji i przejęć [Mazzi, André, Dionysiou, Tsalavoutas, 2016, s. 300–301].

Celem badania przeprowadzonego na spółkach z Włoch oraz Wielkiej Brytanii było prześledzenie ujawnień dotyczących odpisów z tytułu utraty wartości firmy. Okazało się, że zaledwie 27 ze 110 przedsiębiorstw dokonało ww. odpisów. Ponadto zakres ujawnień dotyczących testów na utratę wartości firmy został określony w badaniu jako niski, bowiem spółki te osiągnęły średni wynik na poziomie 5,5 (w skali od 1 do 10, gdzie 10 oznaczało najwyższy poziom ujawnień). Ponadto okazało się, że spółki z Włoch ujawniają więcej informacji w notach do sprawozdania finansowego w sytuacji, kiedy nastąpiła faktyczna utrata wartości firmy. Dla porównania, w przypadku spółek brytyjskich nie zaobserwowano takiej relacji. Dodatkowo spółki włoskie z niższym wskaźnikiem ROE ujawniają mniejszy zakres informacji na temat utraty wartości firmy. Z drugiej strony badania na spółkach z Wielkiej Brytanii także wykazały taki związek, jednakże nie tak silny, jak w przypadku spółek z Włoch [D'Alauro, 2013, s. 780–781].

Perspektywę badawczą polegającą na porównaniu dwóch podejść do wartości firmy i utraty jej wartości, tj. systematycznej amortyzacji oraz przeprowadzania testów na utratę wartości przyjęto w celu zbadania hiszpańskich spółek. Okazało się, że metoda testu na utratę wartości (zaprezentowana w MSR 36) w przypadku tych podmiotów nie zapewnia wierniejszej informacji na temat faktycznej utraty wartości firmy. Pomimo faktu, że testowanie wartości firmy jest rozwiązaniem bardziej elastycznym oraz pokazuje w lepszy sposób sytuację finansową przedsiębiorstwa, nie zaobserwowano w przypadku hiszpańskich spółek jego prawidłowego wdrożenia, co wynikało przede wszystkim z niechęci ujawniania wrażliwych informacji oraz założeń w sprawozdaniach finansowych [Cavero Rubio, Amorós Martínez, Collazo Mazón, 2020, s. 241–242].

Badanie przeprowadzone na 141 podmiotach notowanych na rynkach regulowanych w Paryżu, Frankfurtach, Madrycie oraz w Mediolanie wykazało, że zakres obowiązkowych ujawnień przewidzianych w MSR 36 jest niski. Okazało się, że zaledwie 27% analizowanych podmiotów ujawniło wszystkie niezbędne informacje w sprawozdaniu finansowym wymagane przez międzynarodowe regulacje rachunkowości. Badacze szczególnie zwrócili uwagę na niski odsetek podmiotów, które ujawniły tzw. analizę wrażliwości (zaledwie 50,38% badanych jednostek). Ponadto okazało się, że występują znaczne rozbieżności w ujawnieniach, biorąc pod uwagę kraj, w którym działa dane przedsiębiorstwo. Dodatkowo większość przedsiębiorstw (97,60%) ujawniło, że do kalkulacji wartości odzyskiwalnej została użyta wartość użytkowa [Devalle, Rizzato, 2012, s. 107].

Próbie zbadania zgodności z wymogami ujawnieniowymi w zakresie utraty wartości firmy podjęto także na przykładzie 83 przedsiębiorstw notowanych na rynku regulowanym w Johannesburgu. Okazało się, że większość z nich (78 przedsiębiorstw) ujawniło wartość firmy alokowaną na tzw. ośrodki wypracowujące środki pieniężne. Badacze zwrócili jednakże uwagę na jakość tych ujawnień, która w niektórych przypadkach okazała się być na niskim poziomie. Ponadto okazało się, że zdecydowana większość przedsiębiorstw (55) nie ujawniła bądź podała szacunkowe informacje na temat głównych założeń dotyczących oszacowania przepływów pieniężnych. Badacze określili poziom ujawnień (zgodnie z par. 134 MSR 36) jako niski, pomimo faktu, że zdecydowana większość badanych spółek charakteryzowała się znaczną istotnością wartości firmy w sprawozdaniach finansowych. [Day, 2019, s. 59–60].

Badania ujawnień dotyczących testów na utratę wartości firmy zostały także przeprowadzone na przykładzie przedsiębiorstw z Singapuru. Okazało się, że zgodność obowiązkowych ujawnień (zawartych w MSR 36) wśród badanych jednostek jest zaskakująco niska, szczególnie w obliczu rozwiniętej gospodarki, jaką jest Singa-

pur. Badanie objęło 168 przedsiębiorstw. Okazało się, że liczba podmiotów w pełni zgodnych z MSR 36 w latach 2005–2007 wyniosła odpowiednio 95, 120 oraz 125 jednostek. Zauważa się tendencję rosnącą, jednakże na tle wszystkich badanych przedsiębiorstw jest to nadal niski odsetek [Carlin i in., 2009, s. 237].

W ramach badania przeprowadzonego na podstawie 76 przedsiębiorstw notowanych na rynku regulowanym w Atenach (które rozpoznały wartość firmy w swoich sprawozdaniach finansowych), badano wpływ tychże ujawnień na wycenę przedsiębiorstw. Badacze wykazali, że zakumulowana wartość firmy w bilansach spółek ma pozytywny wpływ na wycenę tychże firm na giełdzie wtedy i tylko wtedy, kiedy wykazują one wysoki stopień zgodności z ujawnieniami dotyczącymi wartości firmy wymaganymi przez MSR. Z drugiej strony, w przypadku kiedy przedsiębiorstwa wykazywały niski stopień zgodności z MSR oraz wartość firmy nie stanowiła istotnej pozycji na tle sprawozdania finansowego jako całości, inwestorzy nie wiązali z tą pozycją sprawozdawczą swoich decyzji inwestycyjnych oraz nie oczekiwali potencjalnych korzyści ekonomicznych z tego tytułu w przyszłości [Baboukardos, Rimmel, 2014, s. 12–13].

### **3.2. Determinanty oraz skutki ujawnień dotyczących wartości firmy w sprawozdaniu finansowym**

W odniesieniu do ujawnień dotyczących odpisów z tytułu utraty wartości firmy, badanie przeprowadzone na przedsiębiorstwach z Baltic Stock Exchange (w latach 2005–2013) wykazało, że firmy, które dokonały zmiany kadry menedżerskiej częściej ujawniają ww. odpisy. Podobna sytuacja występowała w przypadku firm, które charakteryzowały się wysokim stopniem zadłużenia. Ponadto wyniki tego badania wskazują, że nowi menedżerowie są bardziej skłonni do ujawniania wyższych odpisów z tytułu utraty wartości w porównaniu do menedżerów, którzy piastują te stanowiska w firmie od wielu lat. Interesującym wnioskiem okazało się także to, że przedsiębiorstwa, które wykazywały niesatysfakcjonujące wyniki finansowe, były bardziej skłonne do ujawniania wyższych odpisów z tytułu utraty wartości firmy [Sapkauskienė, Leitoniene, Vainiusienė, 2016, s. 427]. Wniosek przeciwstawny do powyższego został zaprezentowany w badaniu przeprowadzonym na hiszpańskich spółkach. Okazało się, że im lepsze wyniki finansowe przedsiębiorstwa, tym wyższe odpisy z tytułu utraty wartości zostały ujęte w sprawozdaniach finansowych [Cavero Rubio i in., 2020, s. 224]. Ponadto kwestię wpływu czynników ludzkich na decyzje w zakresie założeń oraz ujawnień dotyczących przeprowadzania testów z tytułu utraty wartości firmy prześledzono w badaniu przedsiębiorstw na podstawie danych z bazy Compustat. Okazało się, że analitycy finansowi poprzez moni-

torowanie terazniejszych wyników finansowych przedsiębiorstwa oraz przeszłych tendencji, wpływają istotnie na założenia testów na utratę wartości firmy oraz decyzje menedżerów w sprawie ich akceptacji. Ponadto wyniki badań sugerują, że w przypadku kiedy menedżerowie nie zdecydują się na przeprowadzenie testów na utratę wartości firmy (przez co nie znajdują się one w sprawozdaniu finansowym wraz z ujawnieniami), przedsiębiorstwa w których pracują osiągają gorsze wyniki na rynku kapitałowym [Ayres, Campbell, Chyz, Shipman, 2019, s. 1247].

Inne badanie przeprowadzone na niemieckich spółkach giełdowych w latach 2010–2015 wykazało, że aspekty takie jak interpretacja menedżerów oraz zastosowanie przez nich standardów MSR są powiązane z informacjami ujawnionymi przez nich w sprawozdaniach finansowych przedsiębiorstw. Ponadto okazało się, że wyniki przedsiębiorstw, a także wysokość odpisów z tytułu utraty wartości firmy są powiązane z jakością ujawnianych informacji w sprawozdaniach finansowych. Co ciekawe, badacze wskazali także na fakt, że wysokość raportowanej wartości firmy nie jest powiązana z jakością ujawnień jej dotyczących [Lazar, Velte, 2018, s. 327].

W badaniu, które skupiło się na ujawnieniach dotyczących odpisów z tytułu utraty wartości firmy, okazało się, że jednostki, które zapewniają dodatkowe ujawnienia niewymagane przez MSR, w kolejnych okresach sprawozdawczych dokonują odpisów z tytułu utraty wartości firmy. Ponadto wyniki badań sugerują, że dodatkowe ujawnienia dotyczące testów z tytułu utraty wartości firmy zapewniają dodatkową użyteczność dla inwestorów, którzy nie postrzegają wtedy sytuacji danej jednostki w sposób negatywny [Chen, Gu, 2019, s. 33].

Kwestię ujawnień przedsiębiorstwa w zakresie wartości firmy w odniesieniu do współpracy z podmiotami zewnętrznymi przedstawiono w dwóch badaniach. W jednym z nich okazało się, że ujawnienie w sprawozdaniu finansowym faktu o skorzystaniu w danym okresie sprawozdawczym z pomocy niezależnego eksperta dotyczącej przeprowadzenia testów na utratę wartości firmy powoduje, że ujawniane odpisy z tytułu utraty wartości firmy są wyższe niż w kolejnych latach [Gietzmann, Wang, 2020, s. 50]. Ponadto kwestia ujawnień dotyczących wartości firmy została także zbadana w odniesieniu do instytucji audytu zewnętrznego. Okazało się, że zakres ujawnień dotyczących wartości firmy w sprawozdaniu finansowym jest skorelowany dodatnio z wysokością opłat za audyt zewnętrzny. Jest to spowodowane zwiększonym ryzykiem badania związanego z zaistnieniem ujawnień dotyczących wartości firmy, opartych na znaczących osądach i szacunkach [Chen, Gu, 2019, s. 17].

## 4. Dyskusja wyników

Analiza wybranych tekstów naukowych wykazała, że zakres ujawnień dotyczących wartości firmy oraz testów na utratę wartości firmy w sprawozdaniach spółek jest niekompletny oraz niejednorodny. Oznacza to, że podmioty gospodarcze nie wywiązują się z obowiązków jakie nakładają na nich MSR/MSSF w odniesieniu do zapewnienia odpowiedniej informacji dla interesariuszy sprawozdań finansowych. Zauważa się, że większość badań empirycznych koncentruje się na analizie sprawozdań finansowych. Były one przeprowadzane w odniesieniu do jednostek notowanych na rynkach regulowanych w różnych krajach i kontynentach.

Kolejnym ograniczeniem przeanalizowanych badań naukowych była istotność wartości firmy w sprawozdaniach finansowych danych jednostek. Nieliczne badania brały pod uwagę wagę pozycji sprawozdawczej „wartość firmy” w sprawozdaniach finansowych w kontekście obecności ujawnień w notach dodatkowych [Ayres i in., 2019; Chen, Gu, 2019; Day, 2019; Kabir i in., 2020; Mazzi i in., 2016]. Większość studiów naukowych skupiło się na kryterium „zero-jedynkowym”, tj. czy ujawnienia są obecne, czy ich brak. Opracowanie uniwersalnego wskaźnika dotyczącego zakresu ujawnień pozwoliłoby na głębsze zbadanie tematu. Zatem jest to potencjalny obszar dalszych badań.

Ponadto zauważa się, że badania naukowe zaprezentowane w przeanalizowanych tekstach różniły się zakresem czasowym. Część z nich dotyczyła jednego lub dwóch okresów sprawozdawczych [Carlin, Finch, 2010; Devalle, Rizzato, 2012; Day, 2019], natomiast w pozostałych wzięto pod uwagę szerszy zakres [Mazzi i in., 2016; Cavero Rubio i in., 2020; Kabir i in., 2020]. Zatem porównywalność badań ze względu na ewolucję ujawnień w danym okresie jest ograniczona.

Należy podkreślić, że metoda badań różniła się w odniesieniu do poszczególnych analizowanych tekstów naukowych. Dokonywano zarówno analiz ekonometrycznych [Baboukardos, Rimmel, 2014; Mazzi i in., 2016; Sapkauskiene i in., 2016; Ayres i in., 2019; Chen, Gu, 2019; Kabir i in., 2020], jak i subiektywnej oceny obecności i jakości ujawnień w sprawozdaniach finansowych. W szczególności ostatnia metoda jest podatna na interpretacje ze strony badaczy. Oznacza to, że mogli oni dokonywać oceny obecności oraz jakości ujawnień dotyczących wartości firmy i testów na utratę wartości w różnorodny sposób, która z pewnością była podatna na subiektywne osądy i opinie.

Zauważa się, że wąski zakres ujawnień obserwowany jest w przypadku podmiotów z różnych kontynentów, tj. Europy, Australii, Afryki i Azji. Jednakże w przeanalizowanych tekstach naukowych nie znaleziono badania, które porównuje zakres ujawnień dotyczących wartości firmy i testów na utratę wartości firmy w zależności

od kontynentu, na którym działa jednostka gospodarcza. Oznacza to, że kwestia ta stanowi potencjalny obszar dalszych badań.

Ponadto niniejszy przegląd literatury skupił się także na analizie determinant ujawnień oraz ich zakresu, a także ich wpływu na rynki kapitałowe. Okazało się, że na poziom oraz jakość ujawnień w znacznym stopniu wpływa kadra menedżerska, analitycy finansowi oraz osoby odpowiedzialne za sporządzenie sprawozdań finansowych [Lazar, Velte, 2018; Ayres i in., 2019]. Ponadto zakres ujawnień nie był obojętny dla inwestorów, którzy na podstawie pozycji sprawozdawczej „wartość firmy” opierali swoje decyzje inwestycyjne [Chen, Gu, 2019].

## 5. Podsumowanie

Celem tego rozdziału było podsumowanie tekstów naukowych odnoszących się do wartości firmy oraz testów na utratę wartości firmy w sprawozdaniach finansowych przedsiębiorstw. Zakres ujawnień dotyczących wartości firmy zgodnie z MSR/MSSF okazał się w większości przypadków niekompletny – jest to odpowiedź na pierwsze pytanie badawcze.

Ponadto wpływ na pojemność informacyjną sprawozdań finansowych w tym zakresie mają także menedżerowie, analitycy finansowi i osoby odpowiedzialne za sporządzenie sprawozdań finansowych. Udało się zatem odpowiedzieć na drugie pytanie badawcze postawione na początku tego rozdziału. Wyniki analizy sugerują, że w przedsiębiorstwach konieczne jest wdrożenie odpowiednich systemów kontroli wewnętrznej oraz elementów nadzoru korporacyjnego, w celu wyeliminowania wpływu osób odpowiedzialnych za sporządzenie sprawozdania finansowego na zakres ujawnień dotyczących wartości firmy. W istotny sposób przyczyni się to do obiektywnego przedstawiania informacji na temat tej pozycji sprawozdawczej.

Przyczyny wąskiego zakresu ujawnień dotyczących wartości firmy mogą wynikać m.in. z braku wiedzy osób sporządzających sprawozdanie finansowe, braku zrozumienia reguł sprawozdawczości finansowej czy twierdzenia o niskiej wartości not objaśniających dla interesariuszy przedsiębiorstw. Czynniki te wpływają w istotny sposób na zmniejszenie pojemności informacyjnej sprawozdań finansowych. Wskutek tego inwestorzy, którzy są użytkownikami sprawozdań finansowych, nie są wyposażeni w odpowiedni zestaw informacji będący podstawą podejmowania decyzji inwestycyjnych. Tekst ten stanowi zatem wstęp do dalszych badań nad analizą ujawnień dotyczących wartości firmy oraz testów na utratę wartości firmy, w szczególności biorąc pod uwagę perspektywę inwestorów.

Przeprowadzony przegląd literatury przyczynia się do usystematyzowania stanu wiedzy na temat zgodności ujawnień dotyczących wartości firmy w sprawozdaniach finansowych. Jednakże przeprowadzona analiza cechuje się pewnymi ograniczeniami. Głównym z nich jest fakt, że teksty naukowe pochodziły z jednej bazy, tj. Web of Science. Niewykluczone jest, że skorzystanie z większej liczby baz danych umożliwiłoby dokładniejsze opisanie faktycznego stanu rzeczy oraz bardziej holistyczne ujęcie tematu. Ponadto analiza została przeprowadzona zgodnie z metodą PRISMA oraz przy użyciu tzw. słów kluczowych. Możliwe zatem jest, że użycie innych słów kluczowych spowodowałoby objęcie analizą dodatkowych artykułów naukowych, które stanowiłyby wartość dodaną dla tego opracowania.

Niniejszy tekst może stanowić także pomocne narzędzie dla regulatorów, którzy zajmują się projektowaniem standardów rachunkowości, ich wdrożeniem, a także kontrolą ich obecności i kompletności. Wąski zakres ujawnień dotyczący wartości firmy oraz testów na utratę wartości firmy może wskazywać na konieczność podjęcia niezbędnych kroków w celu poprawy tej sytuacji wśród spółek notowanych na rynkach regulowanych.

Wyniki przeanalizowanych w niniejszym rozdziale badań wskazują, że zakres ujawnień dotyczących wartości firmy jest często niekompletny. Zatem tekst ten może skłaniać do refleksji na temat obecnych regulacji MSR/MSSF w tym zakresie. Prawdopodobne jest, że rewizja i uaktualnienie wymaganych ujawnień dotyczących wartości firmy przyczyniłyby się do zwiększenia wartości użytkowej i transparentności sprawozdań finansowych w tym zakresie.

## Bibliografia

---

- Ayres, D.R., Campbell, J.L., Chyz, J.A., Shipman, J.E. (2019). Do Financial Analysts Compel Firms to Make Accounting Decisions? Evidence from Goodwill Impairments, *Review of Accounting Studies*, 24(4), s. 1214–1251. DOI: 10.1007/S11142-019-09512-0.
- Baboukardos, D., Rimmel, G. (2014). Accounting Forum Goodwill under IFRS: Relevance and Disclosures in an Unfavorable Environment, *Accounting Forum*, 38, s. 1–17. DOI: 10.1016/j.accfor.2013.11.001.
- Carlin, T.M., Finch, N. (2010). Resisting Compliance with IFRS Goodwill Accounting and Reporting Disclosures: Evidence from Australia, *Journal of Accounting & Organizational Change*, 6(2), s. 260–280. DOI: 10.1108/18325911011048781.
- Carlin, T.M., Finch, N. (2011). Goodwill Impairment Testing under IFRS: A False Impossible Shore?, *Pacific Accounting Review*, 23(3), s. 368–392. DOI: 10.1108/01140581111185544.



- Carlin, T.M., Finch, N., Khairi, K.F. (2010). FRS36 and Post-Transition Compliance Quality among Singapore Firms, *Asian Review of Accounting*, 18(3), s. 221–244. DOI: 10.1108/13217341011089630.
- Carlin, T.M., Finch, N., Laili, N.H. (2009). Goodwill Accounting in Malaysia and the Transition to IFRS: A Compliance Assessment of Large First Year Adopters, *Journal of Financial Reporting & Accounting*, 7(1), s. 75–104.
- Cavero Rubio, J.A., Amorós Martínez, A., Collazo Mazón, A. (2020). Economic Effects of Goodwill Accounting Practices: Systematic Amortisation versus Impairment Test, *Spanish Journal of Finance and Accounting*, 50(2), s. 224–245. DOI: 10.1080/02102412.2020.1778376.
- Chen, V.Y.S., Gu, F. (2019). On the Usefulness of Supplemental Disclosure about Fair Value Measurement: Evidence for Goodwill Impairment, *Journal of Accounting Review*, 69, s. 1–41. DOI: 10.6552/JOAR.201907\_(69).0001.
- Chen, V.Y.S., Keung, E.C., Lin, I.M. (2019). Disclosure of Fair Value Measurement in Goodwill Impairment Test and Audit Fees, *Journal of Contemporary Accounting & Economics*, 15(3), s. 1–20. DOI: 10.1016/J.JCAE.2019.100160.
- Comiskey, E.E., Mulford, C.W. (2010). Goodwill, Triggering Events, and Impairment Accounting, *Managerial Finance*, 36(9), s. 746–767. DOI: 10.1108/03074351011064636.
- D’Alauro, G. (2013). The Impact of IAS 36 on Goodwill Disclosure: Evidence of the Write-Offs and Performance Effects, *Intangible Capital*, 9(3), s. 754–799. DOI: 10.3926/IC.415.
- Day, C. (2019). Goodwill Impairment Testing Disclosures – South African Compliance in 2018, *South African Journal of Accounting Research*, 34(1), s. 45–62. DOI: 10.1080/10291954.2019.1668120.
- Devalle, A., Rizzato, F. (2012). The Quality of Mandatory Disclosure: The Impairment of Goodwill. An Empirical Analysis of European Listed Companies, *Procedia Economics and Finance*, 2, s. 101–108. DOI: 10.1016/S2212-5671(12)00069-X.
- Gietzmann, M., Wang, Y. (2020). Goodwill Valuations Certified by Independent Experts: Bigger and Cleaner Impairments?, *Journal of Business Finance & Accounting*, 47(1–2), s. 27–51. DOI: 10.1111/JBFA.12411.
- Giner, B., Pardo, F. (2015). How Ethical Are Managers’ Goodwill Impairment Decisions in Spanish-Listed Firms?, *Journal of Business Ethics*, 132(1), s. 21–40. DOI: 10.1007/S10551-014-2303-8/FIGURES/1.
- IFRS (2022a). *IFRS 3 – Business Combination*, <https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2021/issued/part-a/ifrs-3-business-combinations.pdf> (dostęp: 9.01.2022).
- IFRS (2022b). *Who Uses IFRS Standards?*, <https://www.ifrs.org/use-around-the-world/use-of-ifrs-standards-by-jurisdiction/> (dostęp: 9.01.2022).
- Jordan, C.E., Clark, S.J. (2015). Do Canadian Companies Employ Big Bath Accounting When Recording Goodwill Impairment?, *International Journal of Economics and Finance*, 7(9), s. 159–168. DOI: 10.5539/ijef.v7n9p159.
- Kabir, H., Su, L., Rahman, A. (2020). Firm Life Cycle and the Disclosure of Estimates and Judgments in Goodwill Impairment Tests: Evidence from Australia, *Journal of Contemporary Accounting & Economics*, 16(3), s. 1–24. DOI: 10.1016/J.JCAE.2020.100207.

- Kamela-Sowińska, A. (1996). *Wartość firmy*. Warszawa: PWN.
- Lazar, L., Velte, P. (2018). Determinants of Mandatory Goodwill Disclosure: The Case of Impairment Testing in Germany, *International Journal of Managerial and Financial Accounting*, 10(4), s. 301–330. DOI: 10.1504/IJMFA.2018.095941.
- Mazzi, F., André, P., Dionysiou, D., Tsalavoutas, I. (2016). Compliance with Goodwill-Related Mandatory Disclosure Requirements and the Cost of Equity Capital, *Accounting and Business Research*, 47(3), s. 268–312. DOI: 10.1080/00014788.2016.1254593.
- Sapkauskienė, A., Leitonienė, S., Vainiusienė, E. (2016). Disclosure of Goodwill Impairment in the Baltic States, *Engineering Economics*, 27(4), s. 417–429. DOI: 10.5755/J01.EE.27.4.153.

## The Scope and Determinants of Disclosures of Goodwill in Financial Statements according to IAS/IFRS. A Systematic Literature Review

### Summary

The aim of the chapter is to critically analyse the literature on goodwill and impairment tests according to IAS/IFRS. Eighteen scholarly texts were analysed with the PRISMA scientific method. In particular, this chapter deals with the scope of disclosures of goodwill in financial statements as well as the determinants of such disclosures and their effects on capital markets. The analysis of scholarly articles concludes with a discussion of the results, which turned out to indicate that the scope of disclosures of goodwill in financial statements, despite the applicable international accounting regulations, is narrow and incomplete. In addition, this study points to potential areas for further research, i.e. a comparison of the scope of disclosures of goodwill between continents or the development of an indicator describing their scope and detail.

**Keywords:** goodwill, disclosures, goodwill impairment test, accounting, financial statement

# **Analiza sentymentu w języku chińskim z wykorzystaniem metod uczenia głębokiego – systematyczny przegląd literatury**

Marcin Żurek

## **Streszczenie**

Celem publikacji jest systematyczny przegląd literatury naukowej z lat 2018–2021 na temat metod uczenia głębokiego wykorzystywanych do analizy sentymentu w tekstach w języku chińskim. W niniejszym artykule podjęto próbę odpowiedzi na pięć pytań badawczych. Przegląd systematyczny wykonano zgodnie ze wskazaniem metody PRISMA. W ramach wyszukiwania zidentyfikowano 52 artykuły, spośród których 18 przeczytano w pełnej długości, a 11 włączono do przeglądu. Na podstawie analizy tekstów stwierdzono, że we wszystkich badaniach został zaproponowany nowy model sieci neuronowej do klasyfikacji sentymentu, w tym w siedmiu badaniach przeanalizowano jego parametry. W czterech badaniach wykorzystano nową metodę reprezentacji tekstu. W trzech badaniach stworzono zasoby tekstowe do analizy sentymentu. W trzech badaniach odnotowano inne cele.

W analizowanych modelach zastosowano reprezentację tekstu w języku chińskim na różnych poziomach: kreski, klucze, znaki, słowa, zdania. W badaniach opisano modele oparte na konwolucyjnych sieciach neuronowych, rekurencyjnych sieciach neuronowych lub kombinacjach obu typów sieci. We wszystkich badaniach wyniki osiągnięte przez modele poddano ocenie. Modele osiągały przy określonych zbiorach, parametrach i warunkach testowych dokładność między 82,24% a 96,34%, natomiast wskaźnik F1 mieścił się w przedziale 79,93–99,25%. Z uwagi na różne zbiory danych wykorzystane w badaniach, w niniejszej publikacji nie wskazano architektury o najlepszych wynikach. Zidentyfikowano kierunki przyszłych badań nad analizą sentymentu w języku chińskim, do których należą: metody reprezentacji tekstu chińskiego, porównanie wyników

modeli, analiza parametrów poszczególnych modeli, wkład w tworzenie zbiorów zasobów tekstowych w języku chińskim.

**Słowa kluczowe:** analiza sentymentu, uczenie głębokie, język chiński, sieci neuronowe, reprezentacja tekstu chińskiego

## 1. Wprowadzenie do klasyfikacji sentymentu w języku chińskim

Platformy takie jak Facebook i Twitter w Stanach Zjednoczonych czy Weibo w Chinach w ciągu ostatnich lat zyskały na popularności i użytkowane są przez dużą liczbę internautów. Jednym z głównych celów, do których wykorzystuje się media społecznościowe, są interakcje międzyludzkie [Whiting, Williams, 2013, s. 366]. W efekcie dane zgromadzone na tych stronach stanowią zbiór opinii społecznych. Oprócz platform typu social media, w Internecie istnieją też inne portale, na których zbierane są opinie konsumenckie, w tym m.in. sklepy e-commerce, bazy filmów czy strony internetowe do rezerwacji hoteli. Duża ilość gromadzonych danych sprawia, że analiza dostępnych zasobów wykonywana przez człowieka staje się niemożliwa, co stwarza zapotrzebowanie na metody do zautomatyzowanej i skomputeryzowanej eksploracji tekstów [Zhang, Wang, Liu, 2018].

Analiza sentymentu to zadanie przetwarzania języka naturalnego (NLP) pozwalające na ekstrakcję opinii jego autora. Najczęściej opinie dzielone są na kategorie: pozytywne, neutralne, negatywne [Guerini, Gatti, Turchi, 2013]. Współcześnie analiza sentymentu jest powszechnym przedmiotem badań naukowych w dziedzinie przetwarzania języka naturalnego (NLP). W ramach badań nad klasyfikacją sentymentu dominują dwa główne podejścia: oparte na słownikach oraz oparte na uczeniu maszynowym [Drus, Khalid, 2019, s. 708], a także podejście hybrydowe. Metody związane z uczeniem maszynowym wykorzystują algorytmy w celu ekstrakcji i wykrywania opinii, natomiast słowniki zliczają pozytywne i negatywne słowa w zbiorze danych. Do analizy sentymentu wykorzystuje się także uczenie głębokie oparte na sieciach neuronowych [Tang, Qin, Liu, 2015] jako szczególną kategorię uczenia maszynowego.

Ponieważ analiza sentymentu jest zadaniem klasyfikacyjnym uczenia maszynowego, w literaturze naukowej wyróżnia się tzw. klasyfikatory (*classifiers*), które pozwalają na ocenę poprawności (jakości) wykonanej klasyfikacji. Do popularnych i często wykorzystywanych klasyfikatorów należą dokładność (*accuracy*) oraz F1 (*F1-Score*). Ich interpretację wyjaśnić można za pomocą tzw. tablicy pomyłek, przedstawionej w tabeli 1.

**Tabela 1.** Tablica pomyłek

		Klasa rzeczywista	
		Pozytywna	Negatywna
Klasa predykowana	Pozytywna	Prawdziwie pozytywna (TP)	Fałszywie pozytywna (FP)
	Negatywna	Fałszywie negatywna (FN)	Prawdziwie negatywna (TN)

Źródło: opracowanie własne.

Dokładność jest to odsetek obserwacji, które zostały sklasyfikowane do poprawnej klasy [Pereira, Mitchell, Botvinick, 2009] i wyrażana jest wzorem:

$$Dokładność = \frac{TP + TN}{TP + TN + FP + FN}.$$

Natomiast wskaźnik F1 jest średnią harmoniczną klasyfikatorów [Goutte, Gausier, 2005]:

$$Precision = \frac{TP}{TP + FP}$$

oraz

$$Recall = \frac{TP}{TP + FN},$$

które posiadają interpretację statystyczną. Precision jest miarą dokładności rozpoznania w obrębie predykowanej klasy, natomiast recall opisuje dokładność rozpoznania w obrębie klasy rzeczywistej.

W klasyfikacji sentymentu opinie badane są m.in. w takich obszarach, jak aplikacje mobilne [Guzman, Maalej, 2014], produkty e-commerce [Fang, Zhan, 2015], poglądy polityczne [Park, Ko, Kim, Liu, Song, 2011] czy wiadomości [Moreo, Romero, Castro, Zurita, 2012]. Najwięcej publikacji w tej dziedzinie dostępnych jest w języku angielskim. W mniejszej liczbie dostępne są również publikacje poruszające problem analizy sentymentu w języku chińskim [Peng, Cambria, Hussain, 2017], arabskim [Oueslati, Cambria, HajHmida, Ouneilli, 2020], polskim [Lula, Wójcik, 2011] i wielu innych [Abbasi, Chen, Salem, 2008].

Celem niniejszej publikacji jest przegląd metod opartych o uczenie głębokie wykorzystywanych do analizy sentymentu w języku chińskim. Uczenie głębokie osiąga satysfakcjonujące rezultaty i jest perspektywiczne pod względem rozwoju wiedzy naukowej w tej dziedzinie. Poruszany problem jest istotny, ponieważ systematyczne przeglądy literatury naukowej dotyczącej klasyfikacji sentymentu skupiają się głównie na tekstach angielskich. Analiza sentymentu pod pewnymi względami różni

się między językiem chińskim a angielskim i innymi językami europejskimi [Peng, Ma, Li, Cambria, 2018], co przekłada się na potrzebę prowadzenia dodatkowych badań naukowych nad jego komputerowym przetwarzaniem.

W języku chińskim nie występują przerwy pomiędzy wyrazami, a całe zdanie składa się z jednakowo od siebie oddalonych grafemów, nazywanych znakami. W związku z tym przed rozpoczęciem przetwarzania tekstu należy przeprowadzić prawidłową segmentację zdań na słowa [Xue, 2003]. Przykładowo, zdanie w języku chińskim [我害怕他要害我] oznacza [Boję się, że on chce mnie skrzywdzić]. Poprawna segmentacja wyrazów wygląda następująco: [我/害怕/他/要/害我] – w przełożeniu na język polski [Ja/boję się/on/chce/mnie skrzywdzić]. W tym zdaniu istnieje jednak druga, niepoprawna możliwość segmentacji na wyrazy: [我/害怕/他/要害/我], która w tym przypadku nie ma sensu, ponieważ oznaczałaby [Ja/boję się/on/główny punkt/ja].

Charakterystyczne w języku chińskim jest również to, że język ten przechowuje cechy semantyczne na poziomie niższym niż słowo – na poziomie znaków oraz na poziomie tzw. kluczy (j. angielski – *Chinese radicals*, j. chiński – 部首). Cechy semantyczne na poziomie znaków można zilustrować za pomocą następującego przykładu: chińskie słowo 旅游 oznacza „podróż” i składa się z dwóch znaków – znak 旅 oznacza „wycieczka, podróż”, natomiast 游 oznacza „chodzić, podróżować”. Najmniejsza jednostka o cechach semantycznych w języku chińskim to tzw. klucz (zwany również pierwiastkiem), których jest łącznie 214. Na przykład klucz 女 oznacza „kobieta”, i może być częścią znaku 妈 występującego w słowie 妈妈, oznaczającego „mama”. W końcu każdy klucz składa się z kresek (*strokes*), jednak te nie posiadają już cech semantycznych. Taka konstrukcja języka chińskiego stanowi podstawę do badań w dziedzinie reprezentacji tekstu w języku chińskim na różnych poziomach [Yin, Wang, P. Li, R. Li, Wang, 2016].

Język chiński cechuje się inną składnią, wobec tego analiza składniowa przyjmuje również inną formę niż w językach europejskich [Peng i in., 2018]. Ponadto istnieje fonetyczny sposób zapisu języka chińskiego – tzw. pinyin (j. chiński – 拼音). W języku chińskim jeden znak odpowiada zawsze jednej wymawianej sylabie. Na przykład słowo 人 (człowiek) można zapisać jako „ren”, a analogicznie dłuższe, czterosylabowe słowo, np. 人工智能 znaczące „sztuczna inteligencja”, można zapisać jako „ren gong zhi neng”. Zapis fonetyczny pinyin jest wykorzystywany w Chinach m.in. do pisania na klawiaturze.

Opisane powyżej cechy języka chińskiego stanowią wprowadzenie do dalszej części dyskusji i mają znaczny wpływ na analizę sentymentu w języku chińskim, a w szczególności na metody wektorowej reprezentacji tekstów w języku chińskim.

Na podstawie wstępnego przeglądu literatury naukowej z lat 2018–2021 dotyczącej metod uczenia głębokiego wykorzystywanych do analizy sentymentu w języku chińskim, wskazano pięć pytań badawczych dla niniejszej publikacji:

- 1) Jakie są cele badań nad klasyfikacją sentymentu przy wykorzystaniu uczenia głębokiego w tekstach w języku chińskim?
- 2) Na jakich zbiorach danych tekstowych w języku chińskim przeprowadza się analizę sentymentu przy wykorzystaniu uczenia głębokiego?
- 3) Jakie metody są wykorzystywane do wektorowej reprezentacji tekstu chińskiego?
- 4) Jakie modele i ich architektury głębokich sieci neuronowych są wykorzystywane do analizy sentymentu w chińskich tekstach?
- 5) Jakie wyniki osiągają najskuteczniejsze modele w zadaniu klasyfikacji sentymentu w języku chińskim?

Dalsza część publikacji ustrukturyzowana jest w następujący sposób: w pierwszej części wskazano i opisano metody badawcze wykorzystane w niniejszym przeglądzie, w tym kryteria wyszukiwania i selekcji powiązanych badań, sposób zestawienia najważniejszych informacji z publikacji i opis metody wykorzystanej do podsumowania publikacji ze względu na cel badania. W drugiej części zaprezentowano wyniki selekcji badań i podstawowe informacje o publikacjach wybranych do dalszej analizy. W trzeciej części przeprowadzono dyskusję na temat wyników niniejszego przeglądu literatury z uwzględnieniem wyżej wymienionych pytań badawczych. Na końcu wskazano ograniczenia w niniejszym badaniu i potencjalne cele przyszłych badań w dziedzinie analizy sentymentu w języku chińskim oraz podsumowano wnioski z niniejszej publikacji.

## 2. Metody badawcze

Niniejszy systematyczny przegląd literatury został wykonany zgodnie ze wskazaniami metody PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) [Page i in., 2021]. Diagram PRISMA przedstawiono na rysunku 1.

### 2.1. Strategia wyszukiwania powiązanych badań

Do znalezienia powiązanych badań wykorzystano bazę publikacji Web of Science (webofscience.com) z wynikami aktualnymi na dzień 5 grudnia 2021. Wyszukiwanie przeprowadzono za pomocą następującej frazy: Chinese (Title) AND ai OR “artificial intelligence” OR “machine learning” OR nlp OR “natural language processing”

OR “deep learning” OR “big data” OR “data science” (Topic) AND “sentiment analysis” (Topic) AND 2018 OR 2019 OR 2020 OR 2021 (Publication Years).

## 2.2. Kryteria selekcji badań

Do przeglądu systematycznego wybrano publikacje, które spełniały następujące kryteria: 1) przedmiotem publikacji była analiza sentymentu dla tekstów w języku chińskim; 2) analiza sentymentu była głównym tematem publikacji, a nie tylko jedną z metod analizy informacji (np. jeśli celem badania była ocena jakości hoteli w Hongkongu, a metodą badawczą była analiza sentymentu opinii turystów oceniających hotele na portalu internetowym, to ta publikacja została wyłączona z analizy); 3) do analizy sentymentu w publikacji wykorzystano uczenie głębokie; 4) publikacja była dostępna do przestudiowania w pełnej długości. Publikacje, które spełniały wyżej wymienione kryteria, zostały następnie przeanalizowane w całości, żeby ocenić, czy ich treść pasuje do przeglądu systematycznego. W badaniu wzięto pod uwagę tylko publikacje z lat 2018–2021, co jest uzasadnione z uwagi na aktualność tematu i poruszanych zagadnień.

## 2.3. Proces selekcji badań

Do wyszukania i selekcji tytułów oraz streszczeń zostały wykorzystane funkcjonalności bazy Web of Science. Po przeanalizowaniu tytułów i streszczeń z przeglądu wyłączono badania niespełniające kryteriów selekcji. Następnie podjęto próbę uzyskania dostępu do pozostałych artykułów w pełnej długości. Badania następnie przeanalizowano w pełnej długości i ponownie podjęto decyzję o wyłączeniu badań niespełniających kryteriów selekcji. Pozostałe badania zostały wykorzystane do analizy w niniejszym systematycznym przeglądzie literatury. Cały proces wyszukiwania i wyboru powiązanych badań został przeprowadzony przez jednego autora.

## 2.4. Zestawienie najważniejszych informacji z publikacji

Podczas analizy tekstów w pełnej długości sporządzono spis publikacji w arkuszu kalkulacyjnym Excel, w którym zawarto podstawowe informacje dotyczące każdego z artykułów. Atrybuty zestawione w przedstawionej poniżej tabeli 2 wybrano jako najistotniejsze z punktu widzenia przeglądu systematycznego, a na ich podstawie przeprowadzono dalszą analizę.



**Tabela 2.** Atrybuty badań uwzględnione w najważniejszych informacjach o badaniach

Typ danych	Opis
Tytuł	Tytuł publikacji
Rok	Rok publikacji
Cel	Cel publikacji
Metoda reprezentacji tekstu	Sposób reprezentacji tekstu wykorzystany w modelu zaproponowanym przez autorów
Sieć	Rodzaj architektury sieci neuronowej wykorzystanej w modelu
Źródło danych	Źródło danych (Weibo, opinie o produktach itp.)
Ilość danych	Ilość rekordów danych (np. liczba opinii, liczba zdań itp.)
Kategorie sentymentu	Czy model identyfikuje sentyment (pozytywny, neutralny, negatywny), czy konkretne emocje
Liczba porównanych modeli	Liczba modeli, z którymi został porównany model opisywany w publikacji
Dokładność	Dokładność modelu
F1	Miara F1 dla modelu
Dodatkowe informacje	Dodatkowe informacje, które uwzględniono podczas analizy

Źródło: opracowanie własne.

Każdy z przedstawionych atrybutów wybrano z konkretnego powodu. Tytuł ma na celu ułatwić czytelnikowi przegląd systemu dostępu do artykułów źródłowych. Rok publikacji pozwala na szybkie uzyskanie informacji na temat trendów i kierunków badań nad analizą sentymentu w języku chińskim na przestrzeni lat. Kolejne sześć kategorii: cel, metoda reprezentacji tekstu, sieć, źródło danych, ilość danych oraz kategorie sentymentu zapewniają informacje o szczegółach analizowanych publikacji. Informacje wynikające z tych kategorii są podstawą do dalszej analizy oraz kategoryzacji publikacji ze względu na cel i kierunek badania. Atrybuty: liczba porównanych modeli, dokładność i F1 pozwalają zrozumieć efektywność modelu zaproponowanego przez autorów danej publikacji. Natomiast w ramach ostatniej kategorii – dodatkowe informacje – zidentyfikowano i zapisano wszystkie informacje, które uznano za przydatne w niniejszym przeglądzie systematycznym. Oprócz wyżej wymienionych atrybutów, sporządzono zestawienie schematów poszczególnych architektur głębokich sieci neuronowych, które pozwala na obszerniejsze zrozumienie opisywanych zagadnień.

## 2.5. Podział publikacji na kategorie ze względu na cel badania

Podczas dalszej analizy publikacje podzielono na dodatkowe kategorie wskazujące na cel poszczególnych badań. Taki podział ułatwił odpowiedź na pytanie badawcze na temat kierunków badań naukowych w dziedzinie analizy sentymentu w języku chińskim za pomocą uczenia głębokiego. W ramach celów analizowanych publikacji wyodrębniono cztery główne kategorie:

- 1) propozycja nowej architektury modelu wykorzystywanego do analizy sentymentu w języku chińskim,
- 2) propozycja nowej metody wektorowej reprezentacji słów w języku chińskim,
- 3) stworzenie zasobów tekstowych w języku chińskim wykorzystywanych do analizy sentymentu (np. zbiór tekstów z nowej branży, słowniki tematyczne),
- 4) pozostałe cele.

W ramach pierwszej kategorii wyodrębniono jedną subkategorię: 1.1. porównanie parametrów modelu w celu optymalizacji jego działania. Do tej subkategorii zostały włączone badania, które spełniają zarówno pierwszą kategorię, jak i subkategorię. Każdy z tekstów został przydzielony do jednej lub większej liczby kategorii.

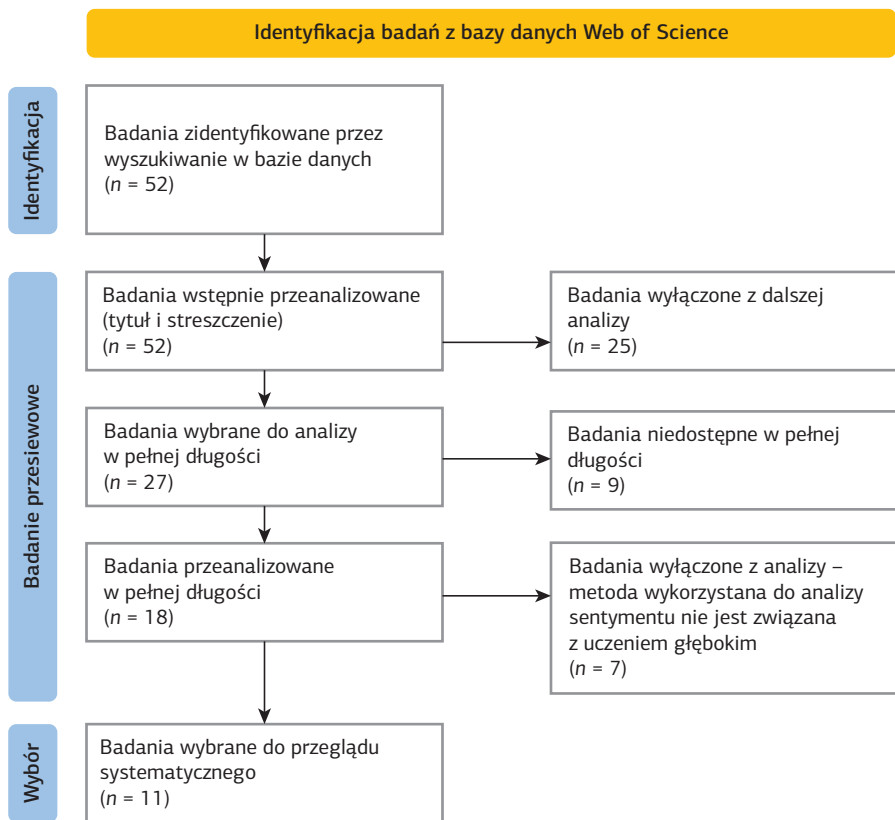
## 3. Wyniki

### 3.1. Badania wybrane do przeglądu

Diagram wizualizujący proces wyboru publikacji do analizy przedstawiono na zamieszczonym poniżej rysunku 1.

W wyniku wyszukiwania w bazach danych zidentyfikowano 52 publikacje. Spośród nich 25 zostało wyeliminowanych z dalszego badania poprzez analizę tytułów i streszczeń. Dziewięć publikacji odrzucono, ponieważ nie uzyskano dostępu do artykułu w pełnej długości. Na końcu 18 publikacji przeanalizowano w pełnej długości, a spośród nich 11 wybrano do przeglądu systematycznego. Siedem publikacji odrzucono w ostatnim etapie selekcji, ponieważ metoda wykorzystana do analizy sentymentu nie była związana z uczeniem głębokim (odrzucono np. publikacje wykorzystujące tradycyjne metody uczenia maszynowego lub metody oparte o słowniki).

**Rysunek 1. Diagram PRISMA**



Źródło: opracowanie własne.

### 3.2. Charakterystyka analizowanych badań

W zamieszczonej w załączniku tabeli 5 przedstawiono zbiór badań wraz z najważniejszymi informacjami o każdym z nich, na podstawie którego przeprowadzono dalszą analizę. Spośród badań, które wykorzystano w niniejszym przeglądzie systematycznym, cztery zostały opublikowane w 2018 r., jedno w 2019 r., pięć w 2020 r. i jedno w 2021 r. Tabela 3 przedstawiona poniżej prezentuje podział badań ze względu na ich cel, co pozwala zrozumieć jakie są obecne kierunki badań nad analizą sentymentu z wykorzystaniem uczenia głębokiego w języku chińskim.

We wszystkich badaniach zaproponowano nową architekturę modelu opartego na uczeniu głębokim, służącego do analizy sentymentu w języku chińskim. W siedmiu publikacjach spośród nich autorzy dodatkowo przeanalizowali wpływ parametrów modelu na jego wydajność lub dokładność. W czterech publikacjach

zaproponowano nową metodę wektorowej reprezentacji tekstu w języku chińskim. Szczegóły i analiza tych metod zostaną przedstawione w dalszej części niniejszego przeglądu systematycznego. Celem trzech badań było stworzenie zasobów tekstowych w języku chińskim, które wykorzystano do publikacji autorskiej, jak również udostępniono w formie open-source na poczet przyszłych badań. Trzy publikacje skupiły się na dodatkowych celach: 1) porównaniu wydajności i skuteczności modeli pomiędzy językiem chińskim i angielskim i/lub 2) wykorzystaniu emotikon do analizy sentymentu w języku chińskim.

**Tabela 3.** Kierunki badań nad analizą sentymentu w języku chińskim

Cel badania	Liczba badań
Propozycja architektury modelu do analizy sentymentu w języku chińskim	11
W tym: Badania parametrów modelu w celu doboru optymalnych wartości	7
Propozycja nowej metody wektorowej reprezentacji tekstu w języku chińskim	4
Stworzenie zasobów tekstowych w języku chińskim wykorzystywanych do analizy sentymentu	3
Inne	3

Źródło: opracowanie własne.

Modele w dziesięciu badaniach polegały na wyodrębnieniu sentymentu z podziałem na kategorie negatywny/neutralny/pozytywny, natomiast w jednym z badań model określał jedną z siedmiu emocji dominujących w tekście (np. smutek, szczęście itp.).

Analiza sentymentu została przeprowadzona na różnych zbiorach danych. Ich zestawienie zostało przedstawione w znajdującej się poniżej tabeli 4. W niektórych badaniach posłużono się zbiorami wykorzystującymi więcej niż jeden zbiór danych.

Analiza sentymentu w języku chińskim przy wykorzystaniu metod uczenia głębokiego najczęściej była przeprowadzana na opiniach produktów ze stron e-commerce (dziewięć razy), następnie na krótkich tekstach z portali społecznościowych (Twitter, Weibo, Plurk – cztery razy) oraz na podstawie zbiorów zawierających opinie o hotelach (trzy razy). Osobno wydzielono kategorię, gdzie analizie sentymentu zostały poddane konwersacje w języku chińskim.

Proponowane przez autorów badań modele w analizowanych publikacjach były porównywane w każdym wypadku z modelami stanowiącymi punkt odniesienia w liczbie od 2 do 17 modeli w zależności od badania. Porównywano m.in. dokładność i wskaźnik F1. Zaproponowane przez badaczy w publikacjach modele osiągały na podanych zbiorach, parametrach i warunkach testowych dokładność między 82,24% a 96,34%, natomiast wskaźnik F1 mieścił się w przedziale 79,93–99,25%.

**Tabela 4.** Zbiory danych wykorzystywanych do kwalifikacji sentymentu w analizowanych badaniach

Źródła danych		Liczba zbiorów
<b>Opinie na stronach e-commerce</b>		<b>9</b>
W tym:	Restauracje	2
	Telefony	1
	Laptopy	2
	Samochody	2
	Sprzęt fotograficzny	1
	Książki	1
<b>Opinie o hotelach</b>		<b>3</b>
<b>Mikroblogi</b>		<b>4</b>
W tym:	Weibo	2
	Twitter	1
	Plurk	1
<b>Inne</b>		<b>3</b>
W tym:	Teksty finansowe	1
	NLP&CC2013	1
	SMP-ECISA 2019	1
<b>Konwersacje w j. chińskim</b>		<b>2</b>
w tym:	Książki do nauki języka angielskiego	1
	Tencent, QQ, Wechat	1

Źródło: opracowanie własne.

## 4. Dyskusja

### 4.1. Ocena jakości modeli oraz przegląd wykorzystanych zbiorów danych

W pozycji Zhou, Li, B. Xu, J. Xu, Yang [2017] model trenowano i testowano na krótkich konwersacjach w języku chińskim, z których podjęto próbę ekstrakcji sentymentu. Wskaźnik F1 dla zaproponowanego modelu wyniósł 80,3%. Dla następujących publikacji: Cheng, Liu, Li, Zhong, Feng, 2020; Lai, Zhang, Han, Zhou, Wang 2020; Peng i in., 2018; Rao, Huang, Jiang, Deverajan, Patan, 2021; Su, C.H. Chen, T.Y. Chen, Cheng, 2020; Wang, Hou, Li, Zhang, 2020; Xiao i in., 2018; Xu, Zhang, Xin, Yang, 2019; Yang, Li, Wang, Sherratt, 2020; Zhou, Lu, Dai, Wang, Xiao, 2019, modele zostały trenowane i testowane na krótkich tekstach, takich jak opinie

o produktach czy teksty z Weibo lub Twittera. Badane modele osiągnęły następujące miary jakości: dokładność w przedziale 82,24–96,34%; wskaźnik F1 w przedziale 79,93–99,25%. W każdym z wyżej wymienionych badań wyniki modelu porównywano z modelami stanowiącymi punkt odniesienia, które oparte były o znane architektury, takie jak CNN (np. w Cheng i in., 2020; Xu i in., 2019) lub LSTM (np. w Lai i in., 2020; Su i in., 2020). Modele proponowane w publikacjach objętych analizą osiągały lepsze wyniki w porównaniu do punktów odniesienia, które wybierano na podstawie dostępnych badań naukowych i aktualnej wiedzy naukowej w tej dziedzinie. Na tej podstawie można wyciągnąć wniosek, że przy określonych parametrach oraz na określonych zbiorach treningowych i testowych, naukowcom udaje się uzyskiwać coraz wyższą dokładność, wykorzystując metody uczenia głębokiego do analizy sentymentu w języku chińskim.

Oprócz różnej tematyki, wykorzystywane zbiory danych miały również różne wielkości – od 2162 [Zhou i in., 2017] do 115 708 elementów [Xiao i in., 2018], co może skłaniać do wniosku, że jakość modeli nie powinna być porównywana, a zamiast tego być analizowana osobno dla każdego modelu.

Wyniki uzyskane w analizowanych artykułach nie są gorsze niż modele analizujące sentyment w tekstach w języku angielskim (np. w Yadav, Vishwakarma, 2020; Do, Prasad, Maag, Alsadoon, 2019).

Modele wykorzystywane w badaniach uwzględnianych w niniejszym przeglądzie były trenowane i testowane na zbiorach danych z różnych źródeł (tabela 4). Część z badań czerpie z publicznie dostępnych zbiorów danych (np. Lai i in., 2020 – NL-P&CC2013 lub Cheng i in., 2020 – ChnSentiCorp, a część wnosi nowe zbiory tekstów. Na podstawie tabeli 4 można stwierdzić, że zakres tematyczny tekstów jest ograniczony, a jedynie dwie publikacje wykorzystują dane o innej tematyce niż opinie o produktach konsumpcyjnych, hotelach, restauracjach lub blogi z mediów społecznościowych (są to Rao i in., 2021 – teksty finansowe, oraz Zhou i in., 2017 – krótkie konwersacje w języku chińskim z komunikatorów oraz stron do nauki języka). W związku z tym w przyszłości badania skupione na tworzeniu nowych zbiorów danych o branżowej tematyce w języku chińskim oraz wykorzystaniu tych danych do klasyfikacji opinii mogą stanowić cenny wkład w rozwój analizy sentymentu jako dziedziny nauki.

Podsumowując dyskusję na temat wyników osiągniętych przez poszczególne modele, brakuje podstaw do jednoznacznej oceny, który model osiąga najlepsze wyniki w analizie sentymentu, ponieważ każdy z nich uczony i testowany był na innym zbiorze danych i przy różnych parametrach. Jakość każdego z modeli powinna być oceniana osobno, a bezpośrednie ich porównanie wymaga dalszych badań.

## 4.2. Kierunki badań nad analizą sentymentu w języku chińskim

W celu przeanalizowania kierunków badań nad metodami uczenia głębokiego wykorzystywanymi do analizy sentymentu w języku chińskim, przyjrano się celom poszczególnych badań i przedstawiono je w tabeli 3. We wszystkich 11 publikacjach zaproponowano model o nowej architekturze, będący wkładem do aktualnego stanu wiedzy naukowej. W czterech publikacjach: Cheng i in. [2020], Peng i in. [2018], Rao i in. [2021] oraz Wang i in. [2020] autorzy dodatkowo zaproponowali nowe sposoby reprezentacji tekstu w postaci wektorowej. W publikacjach Lai i in. [2020], Rao i in. [2021] oraz Zhou i in. [2017] stworzono nowe zasoby tekstowe do analizy.

## 4.3. Metody uczenia głębokiego wykorzystywane do analizy sentymentu w języku chińskim

Jak wskazano powyżej, celem każdego z badań było stworzenie nowego modelu architektury sieci neuronowej do analizy sentymentu w języku chińskim. W dalszej części niniejszego przeglądu przeanalizowano rodzaje sieci neuronowych wykorzystane w analizowanych badaniach oraz częstotliwość ich występowania w analizowanej literaturze naukowej.

Do typów sieci neuronowych wykorzystywanych do analizy sentymentu zaliczane są m.in. konwolucyjne sieci neuronowe (CNN) i rekurencyjne sieci neuronowe (RNN) [Do i in., 2019]. Konwolucyjne sieci neuronowe są efektywne oraz często wybierane przez naukowców zajmujących się przetwarzaniem języka naturalnego, ponieważ pozwalają na ekstrakcję najważniejszych cech n-gramowych z danych wejściowych, które następnie stanowią reprezentację do podejmowania kolejnych zadań klasyfikacyjnych [Do i in., 2019]. Rekurencyjne sieci neuronowe (RNN) odpowiednio podejmują sekwencyjną naturę języka, a fakt, że wynik z RNN jest zależny od poprzednich obliczeń sprawia, że sieci rekurencyjne odpowiednio wyłapują zależności kontekstualne w języku [Do i in., 2019]. Popularne odmiany RNN to Long Short Term Memory (LSTM) oraz Gated Recurrent Unit (GRU) [Goldberg, 2016].

Pośród badań ujętych w niniejszym przeglądzie, pięć posiada architektury bazujące na konwolucyjnych sieciach neuronowych, trzy oparte na rekurencyjnych sieciach neuronowych (LSTM lub Bi-LSTM), natomiast trzy to modele hybrydowe, w których na poszczególnych etapach wykorzystuje się zarówno CNN, jak i BNN (BiGRU, LSTM). W modelach pojawiają się również mechanizmy uwagi [Rao i in., 2021; Su i in., 2020; Yang i in., 2020; Zhou i in., 2017]. Szczegółowe informacje

o wynikach poszczególnych sieci zamieszczono w tabeli 5. Analizowana próba nie daje podstaw do stwierdzenia, które rodzaje sieci neuronowych osiągają najlepsze wyniki w zadaniu klasyfikacji sentymentu.

Spośród 11 badań z niniejszego przeglądu systematycznego, w siedmiu [Lai i in., 2020; Peng i in., 2018; Rao i in., 2021; Wang i in., 2020; Xiao i in., 2018; Yang i in., 2020; Zhou i in., 2019] oprócz samej propozycji architektury i badania jakości modelu wykorzystywanego do analizy sentymentu w języku chińskim, wykonano również badanie parametrów modelu w celu doboru optymalnych wartości, co zapewnia wyższą jakość badania oraz rzetelność analizy.

#### 4.4. Metody wykorzystywane do wektorowej reprezentacji tekstu w języku chińskim

Z uwagi na opisaną we wstępie niniejszego rozdziału charakterystykę języka chińskiego, w jego komputerowym przetwarzaniu dużą uwagę przywiązuje się do segmentacji wyrazów oraz do reprezentacji tekstu w postaci wektorów. W analizowanych badaniach naukowcy podjęli różne próby reprezentacji języka chińskiego. W pracy Cheng i in. [2020] wykorzystano połączenie modeli *cw2vec* oraz SAT, gdzie *w2vec* reprezentuje chińskie słowa jako sekwencje kresek (*strokes*), wykorzystuje więc właściwości morfologiczne języka, a SAT to model oparty na modelu Skip-Gram, który dobrze obejmuje znaczenie wyrazów. Pokazano więc, że włączenie niskopoziomowej reprezentacji opartej na kreskach pozwala zwiększyć miary jakości modelu. Przedstawienie tekstu w języku chińskim za pomocą wektorów stworzonych na różnych poziomach szczegółowości zostało wykorzystane również w publikacji Peng i in. [2018] do analizy sentymentu opartej na aspekcie, gdzie wykorzystano cechy semantyczne słów, znaków oraz kluczy. W pracy Wang i in. [2020] zaprezentowano model, w którym wykorzystano reprezentację wektorową na trzech poziomach: znaków, słów i zdań – tzw. Hierarchical Knowledge Enhancement Representation Learning. Oceniono, że wielopoziomowa reprezentacja tekstu wywiera pozytywny wpływ na wyniki (dokładność i F1) modelu.

W badaniach ujętych w przeglądzie podjęto również próbę reprezentacji tekstu chińskiego przetworzonego na zapis fonetyczny pinyin, co jednak nie przełożyło się na wyższą dokładność klasyfikacji sentymentu (84,11% bez pinyin, 80,4% z pinyin) [Rao i in., 2021]. Uzasadniono to faktem, że w języku chińskim występują liczne homofony – np. sylaba „ma” może być zapisana za pomocą wielu różnych znaków chińskich, a większość z nich ma inne znaczenie, co prowadzi do braku możliwości jednoznacznego połączenia zgłoski z jej znaczeniem.



#### 4.5. Tworzenie zasobów tekstowych do analizy na poczet przyszłych badań

Do celów trzech z jedenastu publikacji zaliczono stworzenie nowych zasobów tekstowych w języku chińskim (zbiorów danych tekstowych), które zostały wykorzystane zarówno w badaniu, jak i udostępnione w formie open-source na poczet przyszłych publikacji. W ramach tych badań udostępniono zbiory z krótkimi tekstami finansowymi [Rao i in., 2021], tekstami (postami) z Weibo [Lai i in., 2020], a także zbiór konwersacji ze strony do nauki języka chińskiego oraz prywatnych komunikatorów takich jak Tencent, QQ i Wechat [Zhou i in., 2017].

Zapotrzebowanie na efektywną klasyfikację sentymentu rośnie, zatem wkład do zasobów tekstowych wydaje się zasadnym kierunkiem badań. W szczególności specyficzne teksty domenowe (branżowe) na tematy finansowe, polityczne i inne stwarzają potencjał do badań i trenowania bardziej wszechstronnych modeli do klasyfikacji sentymentu.

#### 4.6. Inne kierunki badań nad analizą sentymentu z wykorzystaniem uczenia głębokiego

W ramach analizy kierunków badań nad analizą sentymentu w tekstach chińskich z wykorzystaniem uczenia głębokiego, zidentyfikowano trzy artykuły, których cele sklasyfikowano jako „inne”. W pracy Su i in. [2020] przedmiotem badania było wykorzystanie emotikon do klasyfikacji sentymentu za pomocą konwolucyjnej sieci neuronowej z mechanizmem uwagi. Jest to tematyka relatywnie rzadko podejmowana w literaturze naukowej w kontekście klasyfikacji sentymentu i może inicjować nowatorski kierunek badawczy. W dwóch publikacjach: Rao i in. [2021] oraz Peng i in. [2018] wyniki osiągnięte przez modele w zadaniu klasyfikacji sentymentu na chińskich tekstach porównano z wynikami tych samych modeli przy zadaniu klasyfikacji sentymentu w tekstach angielskich. Rao i in. [2021] zaproponowali nowy model i porównali z siedmioma modelami stanowiącymi punkt odniesienia (CNN, CNN-BiLSTM, RCNN-LSTM i in.). Relacje pomiędzy skutecznością modeli w wykonaniu zadania były podobne dla wszystkich analizowanych zbiorów danych, niezależnie od języka (chińskiego lub angielskiego). Pokazano więc, że na badanych zbiorach tekstowych i przy ustalonych parametrach modeli, język tekstów, na których wykonano klasyfikację sentymentu, nie wpływał na różnice w wynikach osiągniętych przez modele. Inaczej ujmując, udowodniono, że dobry model odpowiednio klasyfikuje sentyment zarówno dla języka chińskiego, jak i angielskiego (choćby wyniki niekoniecznie są identyczne), a gorsze modele osiągają gorsze wyniki dla obu języków.

## 5. Podsumowanie

Wraz ze wzrostem liczby użytkowników Internetu i rosnącą popularnością mediów społecznościowych, a także stron e-commerce i innych platform do wymiany opinii w sieci, zwiększa się ilość rekordów danych, które można poddać przetwarzaniu w celu klasyfikacji sentymentu wypowiedzi. Analiza sentymentu skupia uwagę naukowców głównie w przypadku tekstów w języku angielskim, jednak występuje również zapotrzebowanie na inne języki, w tym azjatyckie. Charakterystyka języka chińskiego, w tym struktura, składnia i zapis, różni się od języka angielskiego, dlatego w literaturze naukowej poświęca się mu dodatkową uwagę.

W niniejszym przeglądzie literatury podjęto próbę odpowiedzi na następujące pytania badawcze dotyczące metod uczenia głębokiego wykorzystywanych do analizy sentymentu w języku chińskim: 1) Jakie są cele i kierunki badań nad klasyfikacją sentymentu przy wykorzystaniu uczenia głębokiego w tekstach w języku chińskim? 2) Na jakich zbiorach danych tekstowych w języku chińskim przeprowadza się analizę sentymentu przy wykorzystaniu uczenia głębokiego? 3) Jakie metody są wykorzystywane do wektorowej reprezentacji tekstu chińskiego? 4) Jakie modele i ich architektury głębokich sieci neuronowych są wykorzystywane do analizy sentymentu w chińskich tekstach? 5) Jakie wyniki osiągają najskuteczniejsze modele w zadaniu klasyfikacji sentymentu w języku chińskim?

Podczas niniejszego przeglądu literatury udało się uzyskać odpowiedź na pierwsze pytanie badawcze poprzez zidentyfikowanie celów i kierunków badań. Przeanalizowano 11 publikacji, których przedmiotem była analiza sentymentu w języku chińskim przy wykorzystaniu metod opartych o uczenie głębokie. We wszystkich badaniach zaproponowane zostały modele o nowej architekturze osiągające lepsze wyniki w porównaniu do modeli stanowiących punkt odniesienia. W siedmiu badaniach wzięto pod uwagę dodatkowo parametry modelu w celu optymalizacji jego wyników, co wskazuje na rzetelność badań. Przedmiotem czterech badań była reprezentacja chińskiego tekstu na wielu poziomach (od kresek po zdania) i stwierdzono, że tego typu podejście zwiększa dokładność modeli. W trzech badaniach zidentyfikowano cel publikacji polegający na stworzeniu zasobów tekstowych w formie open-source na poczet przyszłych badań. W trzech publikacjach zanotowano inne cele lub kierunki.

W ramach niniejszej publikacji przeanalizowano i w formie tabelarycznej przedstawiono następujące dane: 1) zbiory danych tekstowych w języku chińskim wykorzystywane do analizy sentymentu, 2) architektury sieci neuronowych wykorzystywane do analizy sentymentu, 3) metody wektorowej reprezentacji tekstu w języku chińskim. Następnie przeprowadzono dyskusję o wynikach tej analizy. Podsumowanie to pozwoliło na odpowiedź na pytania badawcze nr 2, 3 i 4.

Próbując odpowiedzieć na ostatnie pytanie badawcze, zestawiono w tabeli wyniki poszczególnych modeli, uwzględniając miary: dokładność oraz F1. Modele osiągały przy określonych zbiorach, parametrach i warunkach testowych dokładność między 82,24%, a 96,34%, natomiast wskaźnik F1 mieścił się w przedziale 79,93–99,25%. Z uwagi na różnorodność danych testowych (pochodzenie, ilość, jakość) brakuje podstaw do bezpośredniego porównania wyników osiąganych przez proponowane w artykułach modele. Ocena jakości modeli powinna być wykonywana ostrożnie, osobno dla każdego z nich. W związku z tym nie wskazano jednoznacznie modelu o najwyższej skuteczności w zadaniu klasyfikacji sentymentu w języku chińskim według obecnego stanu wiedzy naukowej. W ramach przyszłych badań zasadna wydaje się próba wskazania modelu o najlepszych wynikach, co będzie wymagało systematycznego przeglądu literatury połączonego z porównaniem dostępnych modeli na tych samych zbiorach danych oraz przy identycznych warunkach testowych.

Oprócz odpowiedzi na wyżej wymienione pytania badawcze, w ramach niniejszego systematycznego przeglądu literatury podjęto również próbę wskazania kierunków przyszłych badań, które wydają się zasadne w celu rozwoju klasyfikacji sentymentu w chińskich tekstach jako dziedziny naukowej. Analizując publikacje, stwierdzono, że przetwarzanie języka chińskiego oraz angielskiego za pomocą sztucznych sieci neuronowych różni się z uwagi na charakterystykę obu języków (zapis, składnia, gramatyka). Na przykład do analizy sentymentu w chińskich tekstach stosowana jest wielopoziomowa reprezentacja tekstu wykorzystująca cechy semantyczne kluczy, znaków i słów. Podjęcie próby porównania architektur sieci neuronowych wykorzystywanych do analizy tekstów w różnych językach stanowi temat na odrębne pytanie badawcze i może zostać omówione w ramach kolejnych publikacji. W większości spośród przeanalizowanych prac wykorzystano dostępne zbiory danych, w tym przeważnie z portali społecznościowych lub opinie o produktach z Internetu. Wkład w postaci tworzenia nowych zasobów tekstowych, zarówno obszernych, jak i tych o wąskiej, wyspecjalizowanej tematyce, wydaje się być słusznym kierunkiem dalszych badań, ponieważ stwarza szanse do trenowania modeli na bardziej zróżnicowanych danych, co może prowadzić do ich wszechstronniejszego zastosowania. Większa uwaga może zostać również poświęcona wpływowi zmian parametrów modeli na ich wyniki. Potencjalnie istotnym kierunkiem dalszych badań wydaje się być wektorowa reprezentacja tekstu w języku chińskim na wielu poziomach, ponieważ takie podejście pozwalało modelom osiągać wyższe wyniki w porównaniu do architektur stanowiących punkt odniesienia. W niniejszym badaniu przeanalizowano tylko metody wykorzystujące uczenie głębokie, zatem w ramach przyszłych systematycznych przeglądów literatury na temat analizy sentymentu

w języku chińskim można podjąć próbę podsumowania również metod opartych o słowniki i tradycyjne formy uczenia maszynowego oraz hybrydowych.

Istnieje możliwość wskazania praktycznych implikacji uzyskanych wyników badań. Naukowcy oraz organizacje tworzące modele do klasyfikacji sentymentu w języku chińskim do rozwiązania praktycznych lub teoretycznych problemów mogą wykorzystać zestawienie stosowanych modeli w celu doboru odpowiedniego algorytmu do zadań klasyfikacyjnych. Natomiast zebrane zasoby tekstowe umożliwiają trenowanie (*pre-training*) własnych sztucznych sieci neuronowych. Publikacja agreguje również aktualny stan wiedzy naukowej w dziedzinie klasyfikacji sentymentu w języku chińskim, stanowiąc materiał do dalszej analizy dla badaczy zajmujących się tą tematyką.

## Bibliografia

---

- Abbasi, A., Chen, H., Salem, A. (2008). Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums, *ACM Transactions on Information Systems*, 26(3), s. 1–34. DOI: 10.1145/1361684.1361685.
- Cheng, Y., Liu, C., Li, Y., Zhong, L., Feng, Y. (2020, June). A New Text Sentiment Analysis Method Based on Chinese Morphological Features and HowNet., *Journal of Physics: Conference Series*, 1575(1), IOP Publishing. DOI: 10.1088/1742-6596/1575/1/012101.
- Do, H.H., Prasad, P.W.C., Maag, A., Alsadoon, A. (2019). Deep Learning for Aspect-based Sentiment Analysis: A Comparative Review, *Expert Systems with Applications*, 118, s. 272–299. DOI: 10.1016/j.eswa.2018.10.003.
- Drus, Z., Khalid, H. (2019). Sentiment Analysis in Social Media and Its Application: Systematic Literature Review, *Procedia Computer Science*, 161, s. 707–714. DOI: 10.1016/j.procs.2019.11.174.
- Duric, A., Song, F. (2012). Feature Selection for Sentiment Analysis Based on Content and Syntax Models, *Decision Support Systems*, 53(4), s. 704–711. DOI: 10.1016/j.dss.2012.05.023.
- Fang, X., Zhan, J. (2015). Sentiment Analysis Using Product Review Data, *Journal of Big Data*, 2(1), s. 1–14. DOI: 10.1186/s40537-015-0015-2.
- Goldberg, Y. (2016). A Primer on Neural Network Models for Natural Language Processing, *Journal of Artificial Intelligence Research*, 57, s. 345–420. DOI: 10.1613/jair.4992.
- Goutte, C., Gaussier, E. (2005, March). A Probabilistic Interpretation of Precision, Recall and F-score, with Implication for Evaluation. W: European Conference on Information Retrieval (s. 345–359). Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-540-31865-1\_25.
- Guerini, M., Gatti, L., Turchi, M. (2013). Sentiment Analysis: How to Derive Prior Polarities from SentiWordNet. DOI: 10.48550/arXiv.1309.5843.
- Guzman, E., Maalej, W. (2014, August). How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews. W: *2014 IEEE 22nd International Requirements Engineering*

- Conference RE*, Karlskrona, Sweden, August 25–29, 2014 (s. 153–162), T. Gorschek, R.R. Lutz (Eds.). IEEE. DOI: 10.1109/RE.2014.6912257.
- Lai, Y., Zhang, L., Han, D., Zhou, R., Wang, G. (2020). Fine-Grained Emotion Classification of Chinese Microblogs Based on Graph Convolution Networks, *World Wide Web*, 23(5), s. 2771–2787. DOI: 10.1007/s11280-020-00803-0.
- Lula, P., Wójcik, K. (2011). Sentiment Analysis of Consumer Opinions Written in Polish, *Economics and Management*, 16(1), s. 1286–1291.
- Moreo, A., Romero, M., Castro, J.L., Zurita, J.M. (2012). Lexicon-based Comments-oriented News Sentiment Analyzer System, *Expert Systems with Applications*, 39(10), s. 9166–9180. DOI: 10.1016/j.eswa.2012.02.057.
- Oueslati, O., Cambria, E., HajHmida, M.B., Ounelli, H. (2020). A Review of Sentiment Analysis Research in Arabic Language, *Future Generation Computer Systems*, 112, s. 408–430. DOI: 10.1016/j.future.2020.05.034.
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C. i in. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews, *BMJ*, 372. DOI: 10.1016/j.ijisu.2021.105906.
- Park, S., Ko, M., Kim, J., Liu, Y., Song, J. (2011, March). The Politics of Comments: Predicting Political Orientation of News Stories with Commenters' Sentiment Patterns. W: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (s. 113–122). DOI: 10.1145/1958824.1958842.
- Peng, H., Cambria, E., Hussain, A. (2017). A Review of Sentiment Analysis Research in Chinese Language, *Cognitive Computation*, 9(4), s. 423–435. DOI: 10.1007/s12559-017-9470-8.
- Peng, H., Ma, Y., Li, Y., Cambria, E. (2018). Learning Multi-Grained Aspect Target Sequence for Chinese Sentiment Analysis, *Knowledge-Based Systems*, 148, s. 167–176. DOI: 10.1016/j.knsys.2018.02.034.
- Pereira, F., Mitchell, T., Botvinick, M. (2009). Machine Learning Classifiers and fMRI: A Tutorial Overview. *Neuroimage*, 45(1), s. 199–209. DOI: 10.1016/j.neuroimage.2008.11.007.
- Rao, D., Huang, S., Jiang, Z., Deverajan, G.G., Patan, R. (2021). A Dual Deep Neural Network with Phrase Structure and Attention Mechanism for Sentiment Analysis, *Neural Computing and Applications*, s. 1–12. DOI: 10.1007/s00521-020-05652-6.
- Su, Y.J., Chen, C.H., Chen, T.Y., Cheng, C.C. (2020). Chinese Microblog Sentiment Analysis by Adding Emoticons to Attention-Based CNN, *Journal of Internet Technology*, 21(3), s. 821–829.
- Tang, D., Qin, B., Liu, T. (2015). Deep Learning for Sentiment Analysis: Successful Approaches and Future Challenges, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(6), s. 292–303. DOI: 10.1002/widm.1171.
- Wang, H., Hou, M., Li, F., Zhang, Y. (2020). Chinese Implicit Sentiment Analysis Based on Hierarchical Knowledge Enhancement and Multi-Pooling, *IEEE Access*, 8, 126051–126065. DOI: 10.1109/ACCESS.2020.3008874.
- Whiting, A., Williams, D. (2013). Why People Use Social Media: A Uses and Gratifications Approach, *Qualitative Market Research: An International Journal*, 16(4), s. 362–369. DOI: 10.1108/QMR-06-2013-0041.

- Xiao, Z., Li, X., Wang, L., Yang, Q., Du, J., Sangaiah, A.K. (2018). Using Convolution Control Block for Chinese Sentiment Analysis, *Journal of Parallel and Distributed Computing*, 116, s. 18–26. DOI: 10.1016/j.jpdc.2017.10.018.
- Xu, F., Zhang, X., Xin, Z., Yang, A. (2019). Investigation on the Chinese Text Sentiment Analysis Based on Convolutional Neural Networks in Deep Learning, *Computers, Materials & Continua*, 58(3), s. 697–709. DOI: 10.32604/cmc.2019.05375.
- Xue, N. (2003, February). Chinese Word Segmentation as Character Tagging, *International Journal of Computational Linguistics & Chinese Language Processing*, 8(1), s. 29–48: Special Issue on Word Formation and Chinese Language Processing.
- Yadav, A., Vishwakarma, D.K. (2020). Sentiment Analysis Using Deep Learning Architectures: A Review, *Artificial Intelligence Review*, 53(6), s. 4335–4385. DOI: 10.1007/s10462-019-09794-5.
- Yang, L., Li, Y., Wang, J., Sherratt, R.S. (2020). Sentiment Analysis for E-commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning, *IEEE Access*, 8, 23522–23530. DOI: 10.1109/ACCESS.2020.2969854.
- Yin, R., Wang, Q., Li, P., Li, R., Wang, B. (2016). Multi-granularity Chinese Word Embedding. W: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (s. 981–986). DOI: 10.18653/v1/D16–1100.
- Zhang, L., Wang, S., Liu, B. (2018). Deep Learning for Sentiment Analysis: A Survey, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4). DOI: 10.1002/widm.1253.
- Zhou, J., Lu, Y., Dai, H.N., Wang, H., Xiao, H. (2019). Sentiment Analysis of Chinese Microblog Based on Stacked Bidirectional LSTM, *IEEE Access*, 7, 38856–38866. DOI: 10.1109/ACCESS.2019.2905048.
- Zhou, Y., Li, C., Xu, B., Xu, J., Yang, L. (2017, November). Constructing a Chinese Conversation Corpus for Sentiment Analysis, *National CCF Conference on Natural Language Processing and Chinese Computing*, s. 579–590. Springer, Cham. DOI: 10.1007/978-3-319-73618-1\_48.

## Załączniki

**Tabela 5. Najważniejsze informacje o badaniach**

Lp.	Tytuł	Rok	Cel	Metoda wektorowej reprezentacji tekstu	Architektura sieci	Źródło danych	Ilość danych	Kategorie sentymentu	Liczba porównanych modeli	Skuteczność	F1	Dodatkowe informacje
1.	A New Text Sentiment Analysis Method Based on Chinese Morphological Features and HowNet	2020	1. Propozycja nowej metody wektorowej reprezentacji tekstu 2. Propozycja architektury modelu do analizy sentymentu	1. cw2vec 2. SAT (Sememe Attention over Target Model)	Konwulcyjna sieć neuronowa (Double-Channel Convolutional Neural Network Based on Strokes and Sememes)	1. Opinie o hotelach (Tan Songbo Hotel Review Corpus Data Set – CnnSentiCorp) 2. Opinie o samochodach	1. Brak danych 2. Zbiór treningowy 63000; zbiór testowy 7000	Pozytywne/ Neutralne/ Negatywne	5	1. 91,83% 2. 96,34%	Brak danych	Brak
2.	Fine-Grained Emotion Classification of Chinese Microblogs Based on Graph Convolution Networks	2020	1. Propozycja architektury modelu do analizy sentymentu 2. Stworzenie zbioru tekstów (mikroblogi) wykorzystywanych do analizy sentymentu na poczet przyszłych badań	Brak informacji	Bi-LSTM + Graph Convolutional Network + Percentile Pooling	1. Dane NLP&CC2013 2. Mikroblogi z Weibo	1. Zbiór treningowy: 4338; zbiór testowy: 2172 2. Zbiór treningowy: 15664	Emocje: szczęście, smutek, lubienie (like); złość, zniechęcenie, strach, zaskoczenie	11	Brak danych	Mikro F1: 82,32% Makro F1: 79,93%	1. Parametry modelu zostały dodatkowo przeanalizowane w celu doboru optymalnych wartości
3.	Learning Multi-Grained Aspect Target Sequence for Chinese Sentiment Analysis	2018	1. Propozycja architektury modelu do analizy sentymentu opartej na aspekcie (aspect target sentence) 2. Propozycja nowej metody wektorowej reprezentacji tekstu	1. Addaptive embedded learning na celu poznanie kontekstu wewnątrz zdania dla każdej jednostki sekwencji celu na poziomie kreski, znaku i wyrazu 2. Early/late fusion (ma wektorów słów, znaków i wyrazów)	1. Addaptive embedded learning 2. LSTM + softmax	1. Opinie o laptopach 2. Opinie o samochodach 3. Opinie o sprzęcie fotograficznym 4. Opinie o telefonach 5. Opinie o restauracjach (język angielski) 6. Opinie o laptopach (język angielski)	1–4. Zbiór 6582 opinii 5–6. Zbiór 2309 opinii	Pozytywne/ Neutralne/ Negatywne	7	89,38% lub 89,35% (w zależności od parametrów)	Brak danych	1. Parametry modelu zostały dodatkowo przeanalizowane w celu doboru optymalnych wartości 2. Jakość modelu została sprawdzona również dla języka angielskiego 3. Wykorzystano w analizie strukturę języka chińskiego: podział zdania na słowa, znaki, klucze 4. Wykorzystanie reprezentacji zdań na poziomie kreski nie pomogło w uzyskaniu wyższej skuteczności modelu

Ip.	Tytuł	Rok	Cel	Metoda wektorowej reprezentacji tekstu	Architektura sieci	Źródło danych	Ilość danych	Kategorie sentymentu	Liczba porównanych modeli	Skuteczność	F1	Dodatkowe informacje
4.	A Dual Deep Neural Network with Phrase Structure and Attention Mechanism for Sentiment Analysis	2021	1. Propozycja architektury modelu do analizy sentymentu opartej na aspekcie tekstu 2. Stworzenie zbioru tekstów (branża finansowa) wykorzystywanych do analizy sentymentu na początek przyszłych badań 3. Propozycja nowej metody wektorowej reprezentacji tekstu	1. Reprezentacja pinyin – Pypinyin 2. Segmentacja wyrazów – Jieba, wektoryzacja tekstu – Word2vec 3. Parse trees – wektoryzacja wyrażen POS i Syntax Tree	1. Self attention 2. CNN 3. Dropout 4. Batch Normalization 5. Bi-LSTM 6. Połączenie poziomu słów i wyrazen 7. Konkatenacja, dense 8. Softmax	1. Teksty finansowe 2. Opinie z platformy do zamawiania jedzenia 3. Blogi z Twitera (język angielski)	1. 18000 tekstów 2. 11987 opinii 3. 10000 lub 30000 tekstów (2 warianty)	Poztywne/ Neutralne/ Negatywne	7	82,2–89,28% w zależności od parametrów i zbioru	Brak danych	1. Parametry modelu zostały dodatkowo przetestowane w celu doboru optymalnych wartości 2. Podjęto próbę wykorzystania pinyin w modelu 3. Jakość modelu została sprawdzona również dla języka angielskiego
5.	Chinese Microblog Sentiment Analysis by Adding Emoticons to Attention-Based CNN	2020	1. Propozycja architektury modelu do analizy sentymentu	1. word2vec	1. Mechanizm atencji z emotikonami 2. CNN (konwulcyjna sieć neuronowa) 3. Max Pooling	1. Wiadomości SMS z platformy Plurk 3. Blogi z Twitera	Zbiór treningowy: 10000 wiadomości; zbiór testowy: 3000 wiadomości	Poztywne/ Neutralne/ Negatywne	3	85,80%	86,60%	1. Badanie wykorzystania emotikon w analizie sentymentu
6.	Chinese Implicit Sentiment Analysis Based on Hierarchical Knowledge Enhancement and Multi-Pooling	2020	1. Propozycja architektury modelu do analizy sentymentu 2. Propozycja nowej metody wektorowej reprezentacji tekstu	1. Hierarchical Knowledge Enhancement Representation learning: character level, region level, global level	1. CNN, Multi Pooling	1. Zbiór danych SMP-ECISA2019. Tematyka: Weibo, strony turystyczne, o produktach	Liczba tekstów; zbiór treningowy: 12664; zbiór testowy: 4591 Liczba komentarzy; zbiór treningowy: 14774; zbiór testowy: 5143	Poztywne/ Neutralne/ Negatywne	13	85,20%	89,10%	1. Parametry modelu zostały dodatkowo przetestowane w celu doboru optymalnych wartości 2. Wykorzystano reprezentacje języka chińskiego na poziomie znaku, słowa i zdania
7.	Using Convolution Control Block for Chinese Sentiment Analysis	2018	1. Propozycja architektury sentymentu opartego o convolution control block	1. Algorytm skip gram	1. Convolution control block 2. Max Pooling 3. Tiering 4. Average Pooling 5. Softmax	1. Opinie o hotelach (MioChnCorp)	Zbiór treningowy: 115708; zbiór testowy: 38570; zbiór weryfikacyjny: 38570	Poztywne/ Neutralne/ Negatywne	2	Brak danych	F1 positive: 92,93% F1 negative: 92,38%	1. Parametry modelu zostały dodatkowo przetestowane w celu doboru optymalnych wartości



Lp.	Tytuł	Rok	Cel	Metoda wektorowej reprezentacji tekstu	Architektura sieci	Źródło danych	Ilość danych	Kategorie sentymentu	Liczba porównanych modeli	Skuteczność	F1	Dodatkowe informacje
8.	Investigation on the Chinese Text Sentiment Analysis Based on Convolutional Neural Networks in Deep Learning	2019	1. Propozycja architektury modelu do analizy sentymentu	1. Algorytm skip gram	CNN (Konwulcyjna sieć neuronowa)	1. Opinie o hotelach (Tan Songbo Hotel Review Corpus Data Set – ChnSentiCorp)	Zbiór treningowy: 4800; Zbiór testowy: 1200	Pozytywne/Neutralne/Negatywne	2	Brak danych	99,25%	Brak
9.	Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning	2020	1. Propozycja architektury modelu do analizy sentymentu	1. BERT (Bidirectional Encoder Representatios Transformers)	1. CNN 2. Pooling 3. BiGRU 4. Attention	1. Opinie o książkach ze strony Dangdang	Zbiór 100 000 opinii o książkach	Pozytywne/Neutralne/Negatywne	6	93,50%	93,30%	1. Parametry modelu zostały dodatkowo przeanalizowane w celu doboru optymalnych wartości 2. Do analizy sentymentu został wykorzystany również leksykon słów
10.	Sentiment Analysis of Chinese Microblog Based on Stacked Bidirectional LSTM	2018	1. Propozycja architektury modelu do analizy sentymentu	1. CBOW (Continue Bag-of-Words)	1. Stacked Bi-LSTM 2. Softmax	1. Blogi z Weibo	Zbiór 3000 blogów	Pozytywne/Neutralne/Negatywne	5	87,6–89,9% w zależności od parametrów	Brak danych	1. Parametry modelu zostały dodatkowo przeanalizowane w celu doboru optymalnych wartości
11.	Constructing a Chinese Conversation Corpus for Sentiment Analysis	2018	1. Stworzenie zbioru konwersacji chińskich na poszczególnych etapach tego oraz przyszłych badań 2. Analiza skuteczności znanych modeli na przykładzie zebranego zbioru	1. Mechanizm uwagi na poziomie słów i znaków	Attention-Based BLSTM + softmax (conversation-level)	1. Konwersacje ze stron do nauki języka angielskiego 2. Konwersacje z komunikatorów (Tencent, QQ, WeChat)	2162 konwersacje	Pozytywne/Neutralne/Negatywne	18	Brak danych	80,30%	Brak

Źródło: opracowanie własne.

# Sentiment Analysis in Chinese Based on Deep Learning Methods. A Systematic Literature Review

## Summary

The aim of this publication is a systematic review of research in 2018–2021 on deep learning methods used for sentiment analysis in Chinese texts. The research attempted to answer five research questions. The systematic review was carried out in accordance with the PRISMA guidelines. The database search identified 52 studies, 18 of which were screened at full length and 11 were selected for systematic review. In all these studies, a new model of neural network for sentiment analysis was proposed, including seven studies which analysed model parameters. In four studies a new text embedding method was applied. Three studies created Chinese text corpora for sentiment analysis. In three studies additional research objectives were identified. In the analysed models, Chinese texts were represented on different levels: strokes, radicals, characters, words, sentences. The models were based on Convolutional Neural Networks, Recurrent Neural Networks or combinations of both. In all the studies, the performance of models was subject to assessment. With the specific datasets, parameters and testing conditions, the models achieved the accuracy between 82.24% and 96.34% and F1 score between 79.93% and 99.25%. Since different datasets were used, the architecture with the best performance was not selected in this research. In this systematic review, directions of further research have been identified and include: Chinese text embedding, comparison of model performance, analysis of parameters of models, contribution to Chinese text corpora development.

**Keywords:** sentiment analysis, deep learning, the Chinese language, neural networks, Chinese text embedding