Warsaw University of Life Sciences – SGGW
Institute of Economics and Finance
Department of Econometrics and Statistics

QUANTITATIVE METHODS IN ECONOMICS

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH

Volume XXV, No. 2

Warsaw 2024

ISSN 2082-792X
e-ISSN 2543-8565

Department of Econometrics and Statistics WULS – SGGW
(Katedra Ekonometrii i Statystyki SGGW w Warszawie)

Warsaw 2024, Volume XXV, No. 2
Journal homepage: https://qme.sggw.edu.pl

Warsaw University of Life Sciences Press, Nowoursynowska 161, 02-787 Warsaw
Tel. 22 593 55 23
e-mail: wydawnictwo@sggw.edu.pl
wydawnictwo.sggw.edu.pl
Wydawnictwo SGGW
wydawnictwosggw

CONTENTS

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
Tom XXV/2, 2024, s. 43 – 57

Received: 20.03.2024
Accepted: 02.05.2024

ISTOTNY WPŁYW KOBIET W MANAGEMENCIE NA WYNIKI FINANSOWE SPÓŁEK. PRAWDA CZY FAŁSZ NA POLSKIM RYNKU KAPITAŁOWYM?

Krzysztof Kompa https://orcid.org/0000-0001-9538-9589
Dorota Witkowska https://orcid.org/0000-0002-2810-6654
College of Business and Economics
University of Johannesburg, South Africa
e-mail: kkompa@uj.ac.za; mariaw@uj.ac.za

Streszczenie: Celem artykułu jest sprawdzenie czy obecność kobiet we władzach spółek publicznych wpływa na ich wyniki finansowe. Badaniem objęto spółki nieprzerwanie notowane na GPW w Warszawie w latach 2010-2019. Strukturę managementu określano na dzień 30.06. kolejnych lat, a ocenę sytuacji finansowej spółek na koniec roku kalendarzowego. Zastosowano wielowymiarowe syntetyczne miary wektorowe do oceny standingu badanych spółek i zbadano ich korelacje z odsetkiem kobiet w organach statutowych. Wyniki nie potwierdzają pozytywnej zależności pomiędzy wzrostem udziału kobiet w organach statutowych a wynikami finansowymi spółek, bowiem w większości przypadków korelacje są statystycznie nieistotne.

Słowa kluczowe: kobiety w managemencie, parytet, kobiety a wyniki finansowe, wielowymiarowe metody porównawcze

JEL classification: C38, G38, L25, M14

WSTĘP

Potrzeba osiągnięcia równości płci na rynku pracy jest dyskutowana już od kilku dziesięcioleci, zwłaszcza od ONZ-towskiej międzynarodowej dekady kobiet w latach 1975–1985. Kwestia postrzegana jest przy tym przez pryzmat wielu aspektów równych szans i praw, m.in. jednakowego dostępu do wykształcenia, porównywalnej z mężczyznami aktywności zawodowej, równej płacy za równą

https://doi.org/10.22630/MIBE.2024.25.2.5

pracę[1] czy jednakowych możliwości awansowania, w tym zajmowania stanowisk kierowniczych.

Badania wszystkich podmiotów notowanych na GPW przeprowadzane na podstawie danych pochodzących z Notoria Serwis dla wybranych okresów w odniesieniu do rynku głównego, rynku NewConnect oraz obu rynków łącznie, a także z uwzględnieniem wielkości spółek przedstawiono m.in. w pracach Witkowskiej i in. [2017, 2018 i 2019], Kolanko i Wiśniewskiego [2020], Rogalskiej i Wiśniewskiego [2021], a także [Kompa, Witkowska 2023] czy [Kompa, Wiśniewski 2024]. Z badań tych wynika, że w organach statutowych spółek giełdowych ciągle przeważają mężczyźni. Co prawda w latach 2010-2019 nastąpił spektakularny wzrost frakcji kobiet w radach nadzorczych (o 49,6%) i niewielki (15,5-procentowy) wzrost udziału kobiet w zarządach, ale o ile w pierwszym przypadku wzrost ten dotyczył rynku głównego GPW, o tyle w drugim był on bardziej widoczny na NewConnect. Kobiety w organach statutowych wciąż stanowią mniej niż 15% członków zarządów i jedynie w przypadku spółek z NewConnect nieznacznie przekraczają 20% składu osobowego. Kładziemy to na karb specyfiki rynku NC, gdzie notowane są małe, często rodzinne firmy, których kierownictwa obsadzane są zazwyczaj członkami rodziny.

Udział kobiet w zarządzaniu firmami ma jednak ulec zasadniczej zmianie wskutek uchwalenia przez Parlament Europejski w 2022 r. prawa o kwotach w managemencie spółek publicznych. Zasadne wydaje się zatem pytanie czy zmiana ta przyczyni się do poprawy wyników finansowych polskich spółek giełdowych.

W konsekwencji, celem naszych badań jest zbudowanie punktu odniesienia dla analiz implikacji nowego prawa równościowego w biznesie poprzez sprawdzenie czy obecność kobiet w organach statutowych spółek publicznych, nieprzerwanie notowanych w latach 2010–2019 na Giełdzie Papierów Wartościowych w Warszawie, wpływa na sytuację firm, mierzoną w wielowymiarowej przestrzeni wskaźników finansowych. Analizami objęto spółki należące do portfeli indeksów WIG20, WIG30, mWIG40 i sWIG80 w całym okresie badania wykorzystując, analogicznie jak w cytowanej literaturze, dane dostarczone przez Notoria Serwis. Metodologicznie nawiązujemy do badań przedstawionych w pracy [Kompa, Witkowska 2022, 47-58] i – korzystając z tych samych danych[2] – kontynuujemy badania wychodząc poza zagadnienia klasyfikacji spółek pod względem efektywności finansowej. Istotnym *novum* prezentowanego badania jest zastosowanie wielowymiarowych syntetycznych miar wektorowych do oceny zmian standingu spółek w połaczeniu z analizą korelacji tych zmian z odsetkiem kobiet w organach kierowniczych, z uwzględnieniu opóźnienia pomiaru struktury managementu w stosunku do oceny sytuacji finansowej spółek.

[1] Wzmianka o tej zasadzie pojawiła się już w 1957 roku w momencie tworzenia pierwszej Wspólnoty Europejskiej i podpisania Traktatu Rzymskiego [Kupczyk, 2009, s. 116].

[2] https://www.researchgate.net/publication/363456465_Repozytorium_do_artykulu_WEKTOROWA_SYNTETYCZNA_MIARA_EFEKTYWNOSCI_FIRMY_DLA_POLSKICH_SPOLEK_PUBLICZNYCH#fullTextFileContent

PRZEGLĄD LITERATURY

Zapewnienie równości kobiet i mężczyzn w dostępie do stanowisk kierowniczych w gospodarce jest jednym z priorytetów Unii Europejskiej. Dobitnym tego wyrazem jest przyjęta przez Parlament Europejski dyrektywa o parytecie płci w kierownictwie spółek giełdowych[3]. W uzasadnieniu tego prawa twierdzi się m. in., że wprowadzenie parytetu płci zaowocuje lepszym zarządzaniem, co przełoży się na większą rentowność biznesu i lepszą kondycję finansową spółek. Teza ta nie znajduje jednoznacznego rozstrzygnięcia w literaturze przedmiotu, pokazującej zarówno pozytywne, jak i negatywne zależności pomiędzy obydwoma zjawiskami.

Szeroki przegląd takich badań omawiają m. in. Campbell i Minguez-Vera [2008], Carter i in. [2010], Joecks i in. [2013], Lesiewicz [2014], Post i Byron [2015], Kompa i Witkowska [2017, 2018], Shabbir [2018], Sekeroglu i Acar [2019], Witkowska i in. [2019]. Zwraca się przy tym uwagę [Carter i in. 2010] na dwie zasadnicze przesłanki dywersyfikacji menagementu. Po pierwsze, zespoły zdywersyfikowane charakteryzuje szersze spektrum wykształcenia i doświadczenia oraz większa kreatywność. W konsekwencji, zróżnicowanie poglądów członków organów statutowych prowadzi do większej liczby rozwiązań managerskich, chociaż przyczynia się jednocześnie do wydłużonego czasu podejmowania decyzji [Bohdanowicz 2010]. Po drugie, dobrze wykształceni pracownicy, o rozległej wiedzy i szerokich kompetencjach kognitywnych, są niezależnie od płci, rasy, narodowości, różnic kulturowych, etc. predysponowani do pełnienia wysokich funkcji kierowniczych. Wskazuje się również, że zwiększona liczba kobiet w kierownictwie spółek jest dobrze postrzegana przez inwestorów, chociaż zdarzają się powołania kobiet na stanowiska kierownicze wyłącznie po to, by obarczyć je winą za niepowodzenie organizacji – to zjawisko tzw. szklanego klifu, *glass cliff*, opisane przez Ryana i Haslama [2007].

Badania dowodzą, że zarówno w świecie, jak i w Polsce kobiety w nierównym stopniu uczestniczą w zarządzaniu i podejmowaniu decyzji (por. [Kupczyk 2009; Devillard i in. 2013]). Ponadto, kobiety na stanowiskach kierowniczych w większości czują się dyskryminowane, ponieważ ich dostęp do stanowisk jest trudniejszy, mniej niż mężczyźni zarabiają na tych samych stanowiskach i wolniej awansują. Niewiele z nich zauważa poprawę w tym zakresie. Jednocześnie wyróżnia się pewne cechy firm zarządzanych przez kobiety różnicujące je względem tych kierowanych przez mężczyzn [Witkowska i in. 2019, s.179], chociaż wskaźniki przeżywalności dla firm założonych przez kobiety i mężczyzn są zbliżone.

Bogaty przegląd literatury omawiającej z wielu perspektyw przesłanki wzmocnienia roli kobiet na wyższych stanowiskach zarządzania w organizacjach zamieszczono w pracach [Banno i in. 2021; Khatib i in. 2020]. Konkludując, argumenty uzasadniające dzieli się na dwie grupy: społeczne i ekonomiczne

[3] Dyrektywa uchwalona 22.11.2022 r. określa tzw. parytety (kwoty), które powinny zostać osiągnięte do 2026 r.

[Campbell i Mingueza-Vera 2010; Bohdanowicz 2010; Post i Byron 2015]. Wśród argumentów społecznych wymienić należy przede wszystkim dezyderat równego traktowania, tj. zakaz dyskryminacji członków organów statutowych ze względu na płeć. Spośród argumentów ekonomicznych przemawiających za dywersyfikacją organów statutowych wymienia się zrównoważony rozwój firm czy poprawę wyników finansowych implikowaną zwiększeniem się frakcji kobiet. Pozytywny wpływ udziału kobiet w gremiach kierowniczych na wyniki firm potwierdzają m.in. Desvaux i in. [2007], Devillard i in. [2012], Curtis i in. [2012], a także Terjesen i in. [2016], Bennouri i in. [2018], Jyothi i Mangalagiri [2019], Khan i Subhan [2019].

Jednakże w literaturze pojawiają się również opinie, według których udział kobiet we władzach nie ma istotnego wpływu na wyniki przedsiębiorstw (por. [Farrel, Hersch 2005; Bianco i in. 2011; Banno i in. 2021]) lub wskazują, że jest on wręcz negatywny (np. [Adams, Ferreira 2009; Eulerich i in. 2014; Nguyen i in 2020] czy [Ahern, Dittmar 2012]). Ta ostatnia praca dotyczy zmian sytuacji w spółkach norweskich po wprowadzeniu 40% parytetu kobiet w managemencie. Okazało się, że wprowadzenie kwot spowodowało istotny spadek wartości rynkowej spółek oraz znaczny spadek wskaźnika Q Tobina (relacji wartości rynkowej firmy do wartości odtworzeniowej jej aktywów) w kolejnych latach po wprowadzeniu kwot. Z kolei Bohdanowicz [2016] wykazał brak istotnego wpływu frakcji kobiet w radach nadzorczych spółek niefinansowych (notowanych na GPW w Warszawie) na wskaźnik Q Tobina i ROA. Natomiast w przypadku odsetka kobiet w zarządach odnotował brak oddziaływania na ROA oraz statystycznie istotny negatywny wpływ na Q Tobina.

ORGANIZACJA BADANIA

Celem jest ocena wpływu obecności kobiet w organach statutowych na sytuacje spółek notowanych na GPW w Warszawie w latach 2010-2019. W analizach uwzględniono 90 spółek – 73 spółki niefinansowe i 17 spółek finansowych, w tym 8 banków, które były notowane przez cały okres objęty badaniem, tak więc analizowane spółki tworzą próbę longitualną[4] [Kompa, Witkowska 2022]. Wybrane spółki stanowiły 97,8% spółek notowanych pod koniec grudnia 2010 r., ale jedynie 64,54% spółek notowanych w 2019 r.

W kolejnych etapach badań wyznaczono frakcje kobiet w organach statutowych dla wszystkich spółek i lat badania (na dzień 30.06. każdego roku) i wartości wybranych wskaźników finansowych (stan na koniec roku) na podstawie raportów i sprawozdań rocznych w latach 2010–2019, oraz skonstruowano syntetyczne

[4] Dodatkowym warunkiem kwalifikacji spółek do analiz była dostępność pełnej informacji o składach organów statutowych i o ich sprawozdaniach. Pierwotna baza danych została utworzona przez A. Błaszczyk [2021] na podstawie zasobów Nortoria Serwis.

wektorowe mierniki do oceny spółek[5]. W zakresie przyjętych celów pracy zastosowano dwie metody badawcze polegające na sprawdzeniu:

(1) czy spółki o najmniej i najbardziej zróżnicowanych, pod względem udziału kobiet, organach statutowych różnią się istotnie wartościami wybranych wskaźników finansowych – test Cochrana-Coxa, co rozstrzyga kwestię: czy udział kobiet istotnie wpływa na wartości podstawowych wskaźników finansowych;

(2) czy występują korelacje liniowe pomiędzy odsetkiem kobiet w organach kierowniczych, a standingiem finansowym firm, mierzonym syntetyczną miarą efektywności – wykorzystano współczynnik korelacji liniowej Pearsona i test jego istotności.

Na potrzeby tego badania przyjęto, że jeśli dane zjawisko uznaje się za przyczynę innego, to przyczyna jest opóźniona w stosunku do pomiaru zjawiska o 6 lub 18 miesięcy. Innymi słowy przyjęto, że występuje następstwo zdarzeń polegające na wcześniejszej obserwacji odsetka kobiet w managemencie względem pomiaru sytuacji finansowej spółek. Zatem dodatnia korelacja obu zjawisk wskazuje na pozytywny wpływ obecności kobiet w kierownictwie na sytuację spółek.

Tabela 1. Wskaźniki finansowe użyte do badania banków i wszystkich pozostałych firm

Zmienne diagnostyczne banki	Rodzaj	Zmienne diagnostyczne pozostałe	Rodzaj
Wskaźnik kapitału własnego	S	Wskaźnik płynności bieżącej	S
Wskaźnik aktywów przychodowych	S	Wskaźnik płynności szybkiej	S
Produktywność majątku trwałego	S	Wskaźnik ogólnego zadłużenia	D
Koszty działania / Aktywa	D	Wskaźnik rotacji należności	D
Koszty działania / Wynik na działalności bankowej	D	Wskaźnik rotacji zapasów	D
Koszty działania / Dochody z działalności podstawowej	D	Wskaźnik rotacji zobowiązań	D
Współczynnik wypłacalności	S	Wskaźnik rotacji aktywów	S
Wskaźnik płynności MFW	S	EBITDA/aktywa	S
Zmienne diagnostyczne wspólne		Zmienne diagnostyczne wspólne	
Mnożnik zysku P/E	S	Rentowność aktywów	S
Wskaźnik P/BV	S	Wskaźnik rentowności netto	S
Rentowność kapitału własnego	S	Marża zysku operacyjnego	S

Uwaga: S, D – oznaczenia stymulant i destymulant

Źródło: opracowanie własne

W prezentowanym badaniu zastosowano wektorowy miernik rozwoju VSME omówiony w pracy [Kompa, Witkowska 2022] przyjmując, że kondycja firmy w kolejnych latach analizy jest wyrażana wartością tej miary syntetycznej, skonstruowanej w wielowymiarowej przestrzeni wskaźników finansowych.

[5] Przykłady wykorzystania w analizie spółek mierników syntetycznych, opartych na różnicowo transformowanych unormowanych odległościach taksonomicznych ewaluowanych obiektów od hipotetycznego wzorca, można znaleźć m.in. w pracach: [Tarczyński i Tarczyńska-Łuniewska 2018; Kompa 2018; Kompa, Witkowska 2018]. Jednakże stosowanie tych miar rozwoju do porównań obiektów w wieloletnich analizach ma istotne ograniczenia, spowodowane koniecznością wyznaczania wzorców w każdym roku badania.

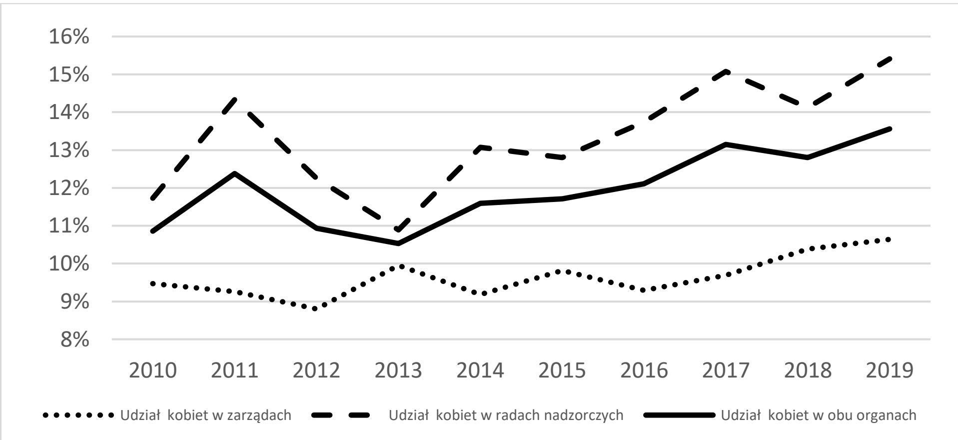
Tabela 1 zawiera listę wybranych wskaźników wraz z informacją czy wskaźnik ma pozytywny czy negatywny wpływ na kondycję spółki.

W związku ze zróżnicowaniem zestawu zmiennych, opisujących spółki o różnym rodzaju działalności, mierniki syntetyczne wyznaczono oddzielnie dla banków i dla pozostałych spółek z podziałem na spółki finansowe inne niż banki oraz spółki niefinansowe, używając po 14 zmiennych, a także mierniki syntetyczne dla wszystkich spółek na podstawie 6 wspólnych zmiennych. Mierniki wyznaczono oddzielnie dla każdego roku analizy 2010-2019 ze wspólnym wzorcem z pierwszego roku badania. Czym wartość miernika większa, tym sytuacja ocenianego obiektu jest lepsza. Miara VSME jest odporna na występowanie obiektów lepszych od wzorca. Może być zatem stosowana do danych przekrojowo-czasowych, bez konieczności zmiany wzorca ustalonego w pierwszym podokresie badania, a obliczone wartości agregatów są porównywalne w całym przedziale czasowym badania.

WYNIKI EMPIRYCZNE

Badania dotyczą wpływu obecności kobiet w kierownictwie na wyniki finansowe spółek. Na podstawie wyznaczonego odsetka kobiet na poszczególnych stanowiskach w analizowanych spółkach (rys. 1) można zauważyć, że w ciągu badanego dziesięciolecia nastąpił znaczący wzrost frakcji kobiet zarówno w zarządach badanych spółek (12%), jak i w radach nadzorczych (31%). Natomiast obecność kobiet na najwyższych stanowiskach tj. prezesów zarządów i przewodniczących rad nadzorczych jest w dalszym ciągu rzadkością i waha się od 1 do 6 prezesek i od 5 do 10 przewodniczących na 90 badanych spółek (czyli odpowiednio od 1,1% do 6,7% oraz od 5,6% do 11,1%).

Rysunek 1. Obecność kobiet w organach kierowniczych badanych spółek



Źródło: obliczenia własne

Zarazem udział kobiet w organach statutowych spółek finansowych jest większy niż w przypadku spółek niefinansowych od 0,22 do 4,11 punktów procentowych w zależności od roku badania. Oprócz tego, w całym badanym okresie odsetek kobiet w kierownictwie mniejszy niż 10% zaobserwowano w ¼ spółek finansowych i niemal ½ spółek niefinansowych. Natomiast przynajmniej 20% udział kobiet w managemencie dotyczy ¼ spółek finansowych i tylko nieco ponad $^1/_6$ spółek niefinansowych. Wśród spółek publicznych znalazło się 7 takich, które nie miały żadnej kobiety w organach statutowych w ciągu całego dziesięciolecia (spośród nich sześć to spółki niefinansowe).

Porównanie wskaźników finansowych spółek o krańcowym zróżnicowaniu udziału kobiet w organach zarządczych

Spośród analizowanych spółek wyróżniono grupę spółek, w których obecność kobiet jednocześnie w obu organach statutowych (biorąc pod uwagę łącznie wszystkie lata badania) była poniżej 10% oraz te, które charakteryzowały się odsetkiem kobiet 20% i więcej. Wyróżniono w ten sposób 38 spółek należących do pierwszej grupy i 17 należących do grupy drugiej, z czego odpowiednio 34 i 13 było spółkami niefinansowymi. Dla obu grup spółek i wszystkich lat badania wyznaczono średnie wartości często używanych w badaniach wskaźników finansowych tj. ROE, ROA, ROS, P/E i P/BV, a następnie porównano czy wyniki osiągane przez spółki o bardziej zróżnicowanym kierownictwie istotnie różnią się od tych, w których organy statutowe są mało zróżnicowane. Wykorzystano test równości dwóch średnich z poprawką Cochrana-Coxa. Statystyki testowe wyznaczono oddzielnie dla spółek finansowych, niefinansowych i wszystkich razem, dla każdego roku oraz dla całego okresu.

Tabela 2. Statystyki t-Studenta dla testu równości dwóch średnich wybranych wskaźników finansowych dla próby badanych spółek niefinansowych (NFin) i próby (Ogółem) wszystkich firm

wskaźniki	P/E		P/BV		ROA		ROS		ROE	
	Kategoria spółki									
Rok	NFin	Ogółem	NFin	Ogółem	NFin	Ogółem	NFin	Ogółem	NFin	Ogółem
2010	1,1890	1,0845	**-3,1174**	-0,2156	-0,1431	0,5227	**-5,5917**	-1,2279	-0,1085	0,9643
2011	1,0243	0,9093	**-3,2786**	-0,7969	0,2246	0,2177	**2,7350**	0,6269	0,4670	0,6775
2012	-0,6678	-0,5899	**-2,1194**	-0,7063	**2,8630**	**1,7568**	**3,7572**	0,8820	**5,4530**	**1,8592**
2013	-1,2650	-1,3253	-0,0816	-0,1344	1,2135	0,8694	**4,3315**	1,0149	**4,6215**	1,0948
2014	**-3,0931**	-0,9425	-0,8644	-0,1754	0,9426	1,1396	**4,3090**	0,9974	**2,5234**	1,6602
2015	-1,0631	-0,8813	-1,6570	-0,3211	**-1,8990**	-0,3935	0,0988	0,0539	**-1,7038**	0,0299
2016	**-4,5297**	-1,0284	-0,9230	-0,3239	**-1,9212**	-0,7984	**-6,6576**	-1,4907	-0,5374	0,4895
2017	0,8862	0,9453	0,0347	0,5619	-1,4366	-1,6883	**-8,2657**	**-1,8738**	-0,5187	-0,6915
2018	0,9074	0,9032	0,6803	0,6190	-1,2389	-0,1052	**-7,6883**	**-1,7192**	-0,7630	0,7120
2019	**1,9492**	0,4001	0,0433	0,1284	**3,0881**	0,9074	**-4,9146**	-1,1431	**-2,0271**	-0,7354
2010-2019	-0,3544	-0,3984	**-1,7459**	-0,7295	-0,0730	0,6811	-1,1198	-0,7464	0,6364	0,7946

Pogrubioną czcionką oznaczono odrzucenie hipotezy zerowej na poziomie istotności 0,05;

Źródło: opracowanie własne

Jak widać w tabeli 2, statystycznie istotne różnice wartości średnich obu grup spółek odnotowano jedynie w 26% przypadków. Przy tym 25 odrzuceń hipotezy zerowej wystąpiło dla współczynników spółek niefinansowych i tylko 4 przypadki odrzuceń, kiedy badano wszystkie spółki. Wynika to z faktu, że wśród spółek niefinansowych częściej (np. w 15 przypadkach na 25 istotnych różnic między średnimi) wykazano niższe wartości wskaźników finansowych generowanych przez spółki, w których w badanym okresie było 20% i więcej kobiet w organach statutowych niż w spółkach o odsetku kobiet w tych organach do 10%. Natomiast spółki finansowe wykazywały częściej sytuację odwrotną[6]. Różnice w osiąganych wynikach w całym okresie badania 2010-2019 są nieistotne na poziomie $\alpha = 0,05$, z wyjątkiem wskaźnika P/BV w przypadku spółek niefinansowych oraz wskaźnika rentowności sprzedaży (ROS) w przypadku spółek finansowych. W analizach prowadzonych dla każdego roku oddzielnie, istotnie lepsze wyniki osiągane przez spółki o zróżnicowanym kierownictwie odnotowano w 12 przypadkach (w tym 10 obserwacji w spółkach niefinansowych dla ROE, ROS i ROA oraz 2 dla wszystkich uwzględnionych w tym badaniu spółek dla ROA i ROE). Natomiast istotnie gorsze w 15,5% wykonanych testów. Trudno jest zatem twierdzić, że zróżnicowane organy statutowe istotnie lepiej/gorzej zarządzają spółkami.

Wpływ frakcji kobiet w managemencie na ocenę sytuacji finansowej spółek

W dalszym postępowaniu analizowano występowanie korelacji między obecnością kobiet w organach statutowych spółek giełdowych a ich standingiem finansowym, mierzonym wektorowymi miernikami syntetycznymi [Kompa, Witkowska 2022; 2023]. Uwzględniono przy tym półroczny i półtora-roczny odstęp pomiędzy pomiarem frakcji kobiet a pomiarem wskaźników finansowych. Badania realizowano na poziomach obu zmiennych n_{it} – frakcja kobiet, $VSMD_{it}$ – wartość wektorowej miary efektywności finansowej oraz na ich przyrostach bezwzględnych i względnych[7] tj. $d_1n_{it} = (n_{it} - n_{i(t-1)})$; $d_1VSMD_{it} = (VSMD_{it} - VSMD_{i(t-1)})$ oraz $d_2VSMD_{it} = d_1VSMD_{it} / VSMD_{i(t-1)}$. Wyniki zestawiono w tabelach 3-7.

Analiza wyników z tabel 3 i 4 jednoznacznie wskazuje na brak zależności między oceną kondycji finansowej spółek a frakcją kobiet w ich organach statutowych. Jedynie bowiem w 6,3% przypadków obliczone współczynniki korelacji liniowej Pearsona były statystycznie istotnie różne od zera. Przy tym ujemna korelacja wystąpiła jedynie dla spółek niefinansowych (w 4 przypadkach na 11).

[6] Pominięto analizę sytuacji w spółkach finansowych w kolejnych latach, ponieważ w każdej grupie spółek było ich po 4. Wyniki dla tych spółek w 10-letnim okresie badania zawarto w pracy [Kompa, Witkowska, 2023].

[7] Współczynniki korelacji obliczano dla takiej samej formy reprezentacji obu zjawisk.

Tabela 3. Współczynniki korelacji liniowej Pearsona pomiędzy frakcją kobiet a efektywnością finansową obliczone dla wszystkich 90 spółek i 73 spółek niefinansowych

Okres pomiaru zmiennych opisujących frakcję kobiet	Miary efektywności finansowej i ich przyrosty bezwzględne							
	$VSMD_{it}$		d_1VSMD_{it}		$VSMD_{it}$		d_1VSMD_{it}	
t	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1
	Wszystkie spółki				Spółki niefinansowe			
2010	-0,0616	0,0316			-0,0522	0,0586		
2011	0,0461	0,0616	-0,1546	-0,0007	0,0089	-0,0712	**-0,2085**	-0,0498
2012	0,0412	0,0402	0,0432	-0,0431	-0,0363	0,0018	0,0732	**-0,2270**
2013	0,1498	**0,2135**	0,0462	0,0835	0,1729	0,0868	0,0264	-0,1877
2014	0,1302	-0,0051	-0,1018	0,0236	-0,0036	-0,0508	**-0,2286**	0,0617
2015	0,0174	-0,0163	-0,0217	-0,0348	-0,0240	-0,0463	-0,0286	0,0020
2016	0,0142	-0,1137	0,0322	-0,0639	-0,0647	-0,1043	0,0123	0,1301
2017	-0,1457	-0,0963	0,0267	0,0767	-0,1032	-0,0969	0,0080	0,1116
2018	-0,0354	-0,0324	0,0541	-0,1592	-0,0880	-0,0467	0,0718	-0,1495
2019	-0,0150		0,0013		-0,0455		0,0365	
2010-2019	0,0068	0,0018	-0,0194	-0,0251	-0,0260	-0,0336	-0,0289	-0,0471

Pogrubieniem oznaczono odrzucenie hipotezy zerowej na poziomie istotności $\alpha = 0{,}05$;
$d_1VSMD_{it} = (VSMD_{it} - VSMD_{i(t-1)})$

Źródło: opracowanie własne

Tabela 4. Współczynniki korelacji liniowej Pearsona pomiędzy frakcją kobiet a efektywnością finansową obliczone dla spółek finansowych

Okres pomiaru zmiennych opisujących frakcję kobiet	Miary efektywności finansowej i ich przyrosty bezwzględne							
	$VSMD_{it}$		d_1VSMD_{it}		$VSMD_{it}$		d_1VSMD_{it}	
t	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1
	Banki				Spółki finansowe pozostałe			
2010	0,2249	-0,4342			0,1769	0,2815		
2011	-0,3762	0,1811	0,2510	-0,1401	0,1241	0,2116	-0,3248	**0,6154**
2012	-0,2369	0,2013	-0,1798	**0,9083**	0,1336	-0,0445	-0,2727	0,0846
2013	0,0637	0,0726	0,0106	**0,6315**	-0,2087	0,0086	-0,2134	**0,5858**
2014	**0,5895**	0,5103	-0,1530	-0,1627	0,1940	0,3966	0,1425	0,0336
2015	**0,6298**	**0,6034**	0,3633	0,5402	0,3014	0,0774	-0,0099	-0,1629
2016	0,3164	0,2761	-0,3738	0,2992	0,1753	-0,1811	0,4628	0,0105
2017	-0,0706	-0,4328	0,4952	0,4718	-0,3495	0,1680	0,0154	0,0040
2018	-0,3168	-0,0900	0,2435	0,3596	0,1783	-0,3279	0,0653	-0,2302
2019	-0,3540		-0,3243		-0,3727		-0,2365	
2010-2019	0,0756	0,01145	-0,0032	0,1141	0,0099	0,0169	-0,0463	-0,0017

Pogrubieniem oznaczono odrzucenie hipotezy zerowej na poziomie istotności $\alpha = 0{,}05$;

Źródło: opracowanie własne

W bankach udział kobiet w obu organach statutowych w całym okresie badania zawierał się w przedziale od 7% do 21% dla poszczególnych spółek, dla pozostałych spółek finansowych od 0 do 56%, a dla spółek niefinansowych od zera do 44%. Przy czym średnie odsetki kobiet w tych trzech grupach spółek wyniosły odpowiednio 15%, 17,3% i 11,6%. Jak zatem widać odsetek kobiet w organach statutowych spółek niefinansowych był najmniejszy.

Tabela 5. Współczynniki korelacji liniowej Pearsona pomiędzy frakcją kobiet a efektywnością finansową obliczone dla wszystkich spółek o niezerowym udziale kobiet w kierownictwie w latach 2012-2018 (44 spółki)

Okres pomiaru zmiennych opisujących frakcję kobiet	Miary efektywności finansowej oraz ich przyrosty bezwzględne i względne					
	$VSMD_{it}$		d_1VSMD_{it}		d_2VSMD_{it}	
t	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1
2012	0,2106	**0,3526**				
2013	**0,4206**	**0,4237**	0,1114	**0,2735**	0,1472	0,0482
2014	**0,3361**	0,0866	-0,1057	-0,0560	-0,0109	-0,0513
2015	0,1223	0,1805	0,1088	-0,1069	0,0960	-0,2059
2016	0,1012	-0,0615	0,0890	0,1397	0,2276	0,0030
2017	-0,0650	-0,0617	-0,0429	**0,2572**	-0,1555	-0,0302
2018	0,0277	-0,0412	0,0999	-0,2229	-0,1093	-0,0075
2019	-0,1003		-0,0403		0,1710	
2012-2019	0,1119	0,1040	0,0341	-0,0017	0,0362	-0,0412

Pogrubieniem oznaczono odrzucenie hipotezy zerowej na poziomie istotności α=0,05; $d_1VSMD_{it} = (VSMD_{it} - VSMD_{i(t-1)})$; $d_2VSMD_{it} = d_1VSMD_{it} / VSMD_{i(t-1)}$

Źródło: opracowanie własne

Wyniki zawarte w tabelach 5 i 6 dotyczą zbiorów danych, utworzonych wskutek eliminacji spółek, w których nie było kobiet (łącznie) w obu organach statutowych przynajmniej w jednym roku badania. Jak widać pominięcie spółek, w których kobiety były nieobecne w kierownictwie przyczyniło się do wzrostu liczby dodatnich istotnych korelacji, z jednej (w 2013 r.) w tabeli 3 do 6 w tabeli 5 (w latach 2012-2014 i 2017 r.) i do 15 w tabeli 6 (w latach 2010-2015 i 2017 r.).

Tabela 6. Współczynniki korelacji liniowej Pearsona pomiędzy frakcją kobiet a efektywnością finansową obliczone dla wszystkich spółek o niezerowej frakcji kobiet w kierownictwie w latach 2010-2019 (34 spółki)

Okres pomiaru zmiennych opisujących frakcję kobiet	Miary efektywności finansowej oraz ich przyrosty bezwzględne i względne					
	$VSMD_{it}$		d_1VSMD_{it}		d_2VSMD_{it}	
t	*t*	*t*+1	*t*	*t*+1	*t*	*t*+1
2010	**0,4568**	**0,4980**				
2011	**0,5675**	**0,3674**	0,0167	-0,1661	**0,4110**	0,0439
2012	0,2421	**0,4491**	0,0053	-0,1181	-0,0535	0,0628
2013	**0,5173**	**0,5152**	0,1087	**0,3562**	0,2485	0,0477
2014	**0,4420**	**0,3310**	-0,0931	-0,0563	0,0415	0,0174
2015	**0,3903**	**0,3696**	**0,2938**	-0,1030	0,1682	-0,2290
2016	0,2259	-0,0024	0,0706	0,1400	0,2338	-0,0323
2017	0,0140	-0,0563	-0,0100	**0,3097**	-0,0778	-0,0497
2018	0,0806	-0,0911	0,1470	**-0,3068**	-0,1105	-0,1206
2019	-0,1933		0,0199		-0,0355	
2010-2019	0,2671	0,2566	0,0422	-0,0228	0,0448	0,0284

Pogrubieniem oznaczono odrzucenie hipotezy zerowej na poziomie istotności α=0,05;

Źródło: opracowanie własne

Pojawiła się też jedna obserwacja wskazująca na negatywny wpływ zmiany odsetka kobiet w 2018 r. (w stosunku do roku poprzedniego) na wynik finansowy.

Z analizy wynika, że wybranie spółek, w których kobiety są obecne w organach statutowych w całym okresie badania wskazuje na ich pozytywny wpływ na standing spółki znacznie częściej niż uwzględnienie w badaniu wszystkich spółek.

Tabela 7. Współczynniki korelacji liniowej Pearsona pomiędzy frakcją kobiet a efektywnością finansową obliczone dla 13 spółek o frakcji kobiet w kierownictwie ≥20%

Okres pomiaru zmiennych opisujących frakcję kobiet	Miary efektywności finansowej i ich przyrosty bezwzględne			
	$VSMD_{it}$		d_1VSMD_{it}	
t	*t*	*t*+1	*t*	*t*+1
2010	**0,3123**	**0,4719**		
2011	**0,5326**	**0,3952**	0,0992	-0,0760
2012	0,1302	**0,4833**	0,0425	-0,3034
2013	**0,4885**	**0,5225**	0,1203	**-0,3140**
2014	**0,4444**	**0,3431**	0,0173	0,4166
2015	**0,4825**	**0,5285**	**-0,3227**	0,1037
2016	0,3915	-0,1526	-0,2836	0,0914
2017	-0,2415	-0,1572	0,1220	0,2803
2018	0,1387	-0,2642	-0,0035	0,1313
2019	-0,2773		**0,3810**	
2010-2019	0,2597	0,2372	0,0241	-0,0369

Pogrubieniem oznaczono odrzucenie hipotezy zerowej na poziomie istotności α=0,05

Źródło: opracowanie własne

Analiza korelacji poziomów obu zjawisk dla spółek o przynajmniej 20% frakcji kobiet w kierownictwie (tabela 7)[8], pokazuje identyczne tendencje jak w przypadku 34 spółek o zerowych frakcjach kobiet. Natomiast, biorąc pod uwagę przyrosty zmiennych, w obu zbiorach tendencje się nie pokrywają, a w latach 2013 i 2015 są wręcz o kierunku przeciwnym.

PODSUMOWANIE

Omawiane wyniki badań dotyczących wpływu obecności kobiet w organach statutowych na wyniki finansowe spółek wykazują brak istotnych zależności między badanami zjawiskami w 81,7% przebadanych przypadków. Warto przy tym odnotować, że statystycznie istotne dodatnie zależności odnotowano w 12,7%, a ujemne w 5,5% przeprowadzonych testów. Przy czym badania realizowane były zarówno na poziomach, jak i na przyrostach zmiennych. W konsekwencji tezę postawioną w tytule rozstrzygamy jako nieprawdziwą – nie potwierdzono istotnego wpływu udziału kobiet w managemencie spółek giełdowych w Polsce na wyniki finansowe tych spółek.

Nasze badanie wpisuje się w światowe trendy badawcze, uzupełniając je o własne koncepcje standaryzacji pomiaru wpływu udziału kobiet w zarządach

[8] jeśli pod uwagę brany był cały okres badania, ale były lata, kiedy kobiet w managemencie nie było.

spółek publicznych na kondycję finansową tych spółek oraz o nowatorskie podejście do pomiaru standingu przedsiębiorstw. Stanowi zarazem bazę odniesienia dla porównań zmian implikowanych nowym prawem równościowym UE po roku 2026.

BIBLIOGRAFIA

Adams R. B., Ferreira D. (2009) Women in the Boardroom and their Impact on Governance and Performance. Journal of Financial Economics, 94, 291-309.

Ahern K. R., Dittmar A. K. (2012) The Change of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. Quarterly Journal of Economics, 127, 137-197.

Banno M., Filippi E., Trento S. (2021) Women in Topic Echelon Positions and their Effects on Sustainability: a Review, Synthesis and Future Research Agenda. Journal of Management and Governance, https://link.springer.com/article/10.1007/s10997-021-09604-7

Bianco M., Ciavarella A., Signoretti R. (2011) Women on Boards in Italy. Quaderni di Finanza, 70, CONSOB COMMISSIONE NAZIONALE PER LE SOCIETA E LA BORSA.

Błaszczyk A. (2021) Obecność kobiet w organach statutowych spółek publicznych w Polsce. Zeszyty Naukowe PTE w Zielonej Górze, 15, 117-134.

Bohdanowicz L. (2010) Kobiety w radach nadzorczych spółek publicznych – pokonywanie barier. Przegląd Corporate Governance, 3(23), 23-33.

Bohdanowicz L. (2016) Własność menedżerska w polskich spółkach publicznych. Wydawnictwo Uniwersytetu Łodzkiego, Łódź.

Bennouri M. C. T., Nagati H., Nekhili M. (2018) Female Board Directorship and Firm Performance: What Really Matters? Journal of Banking & Finance, 88(c), 267-291.

Campbell K., Minguez-Vera A. (2008) Gender Diversity in the Boardroom and Firm Financial Performance. Journal of Business Ethics, 83, 435-451.

Campbell K., Minguez-Vera A. (2010) Female Board Appointments and Firm Valuation: Short and Long-Term Effects. Journal of Management and Governance, 14(1), 37-59.

Carter D.A., D'Souza F., Simkins B.J., Simpson W.G. (2010) The Gender and Ethnic Diversity of US Board Committees and Firm Performance. Corporate Governance: An International Review, 18(5), 396-414.

Curtis M., Schmid Ch., Struber M. (2012) Gender Diversity and Corporate Performance. Research Institute Credit Suisse [on-line:] http://www.calstrs.com/sites/main/files/file-attachments/csri_gender_diversity_and_corporate_performance.pdf [dostęp 10.07.2016]

Desvaux G., Devillard-Hollinger S., Baumgarten P. (2007) Women Matter: Gender Diversity a Corporate Performance Driver. McKinsey&Company [on-line:] http://www.raeng.org.uk/publications/other/women-matter-oct-2007 [dostęp 13.07.2016].

Devillard S., Graven W., Lawson E., Paradise R., Sancier-Sultan S. (2012) Women Matter 2012. Making the Breakthrough. McKinsey&Company [on-line:] http://www.mckinsey.com/features/women_matter [dostęp 12.09.2015].

Devillard S., Sancier S., Werner C., Maller I., Kossoff C. (2013) Women Matter 2013, Gender Diversity in Top Management: Moving Corporate Culture, Moving Boundaries, McKinsey&Company [on-line:] https://www.mckinsey.com/~/media/McKinsey/Fea tured%20Insights/Women%20matter/Addressing%20unconscious%20bias/WomenMatter%202013%20Report%20(8).ashx [dostęp 24.12.2014].

Eulerich M., Velte P., van Uum C. (2014) The Impact of Management Board Diversity on Corporate Performance – an Empirical Analysis for the German Two-Tier System. Problems and Perspectives in Management, 12, 25-39.

Farrel K. A., Hersch P. L. (2005) Additions to Corporate Boards: The Effect of Gender. Journal of Corporate Finance, 11, 85-106.

Joecks J., Pull K., Vetter K. (2013) Gender Diversity in the Boardroom and Firm Financial Performance: What Exactly Constitutes a "Critical Mass"? Journal of Business Ethics, 118(1), 61-72.

Jyothi P., Mangalagiri J. (2019) Would Firm Performance be Better with Women Directors? Evidence from India. The Journal of Business Perspective, 23(2), 180-188.

Khan A.W., Subhan Q.A. (2019) Impact of Board Diversity and Audit on Firm Performance. Cogent Business & Management, 6, 1611719.

Khatib S. F. A., Abdullah D. F., Elamer A. A., Abueid R. (2020) Nudging toward Diversity in the Boardroom: A Systematic Literature Review of Board Diversity of Financial Institutions. Business Strategy and the Environment.

Kolanko Ł., Wiśniewski T. (2020) Coraz więcej kobiet we władzach spółek z warszawskiej giełdy. Parkiet, 7.03.2020, s. 11.

Kompa K. (2018) Parytet płci w kierownictwie spółek giełdowych a zmiany w sytuacji finansowej firm w latach 2010-2013. Przedsiębiorczość i Zarządzanie, 19(3/II).

Kompa K., Wiśniewski T. (2024) Panie w spółkach giełdowych pną się w górę. Parkiet 8.03.2024.

Kompa K., Witkowska D. (2017) Czy obecność kobiet w kierownictwie wpływa na rentowność kapitału własnego spółek notowanych na GPW? Zarządzanie i Finanse, 3.

Kompa K., Witkowska D. (2018) Gender Diversity in the Boardrooms of Public Companies in Poland: Changes and Implications. Montenegrin Journal of Economics, 14(1).

Kompa K., Witkowska D. (2022) Wektorowa syntetyczna miara efektywności firmy dla polskich spółek publicznych. Metody Ilościowe w Badaniach Ekonomicznych, 23(2), 47-58.

Kompa K., Witkowska D. (2023) Impact of Women's Representation in Top Management on Performance of WSE-Listed Companies. [w:] M. Dietl, D. Zarzecki (Eds.), Understanding the Polish Capital Market: from Emerging to Developed. Routledge, New York, s. 280-298.

Kupczyk T. (2009) Kobiety w zarządzaniu i czynniki ich sukcesów. Wyższa Szkoła Handlowa we Wrocławiu, Wrocław.

Lesiewicz E. (2014) Karta różnorodności i jej znaczenie dla polityki równościowej. Przegląd Politologiczny, 2, Uniwersytet im. Adama Mickiewicza w Poznaniu, 195-208,

Nermend K. (2009) Vector Calculus in Regional Development Analysis. Comparative Regional Analysis Using the Example of Poland. Serie: Contributions to Economics. Physica-Verlag Heidelberg (Springer).

Nguyen T. H. H., Elmagrhi M. H., Ntim C. G., Wu Y. (2021) Environmental Performance, Sustainability, Governance and Financial Performance: Evidence from Heavily Polluting Industries in China. Business Strategy and the Environment, 30(5), 2313-2331, http://doi.org/10.1002/bse.2748.

Post C., Byron K. (2015) Women on Boards and Firm Financial Performance: A Meta-Analysis. Academy of Management Journal, 58(5), 1546-1571.

Rogalska T., Wiśniewski T. (2021) Udział kobiet we władzach spółek giełdowych na stabilnym poziomie. Parkiet, 7.03.2021.

Ryan M.K., Haslam S.A. (2007) The Glass Cliff: Exploring the Dynamics Surrounding the Appointment of Women to Precarious Leadership Positions. Academy of Management Review, 32(2), 549-572.

Sekeroglu G., Acar M. (2019) The Effect of Female Board Members on Firm Performance: The Case of Turkey. [in:] M. Zortuk, S. Bayrak Gezdim, H. A. Burhan (Eds.), 4th International Conference on Economics Business Management and Social Sciences. Book of Conference Proceedings and Abstracts, Academia Press Center, Budapest/Hungary.

Shabbir M. S. (2018) Women on Corporate Boards and Firm Performance. Preliminary Results from Italian Listed Companies after Introduction of Gender Quota Law 120/2011. Journal of Business & Finance Affairs, 7(1).

Tarczyński W., Tarczyńska-Łuniewska M. (2018) The Construction of Fundamental Portfolio with the Use of Multivariate Approach. Procedia Computer Science, 126, 2085-2096.

Terjesen S., Couto E.B., Francisco P.M. (2016) Does the Presence of Independent and Female Directors Impact Firm Performance? A Multi-Country Study of Board Diversity. Journal of Management & Governance, 20(3), 447-483.

Witkowska D., Kompa K., Matuszewska-Janica A. (2019) Sytuacja kobiet na rynku pracy: wybrane aspekty. Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Witkowska D., Kompa K., Wiśniewski T. (2017) Kobiety na rynku kapitałowym. Parkiet, 8.03.2017, s. 10.

Witkowska D., Kompa K., Wiśniewski T. (2018) Kobiety w spółkach giełdowych. Parkiet, 5.08.03.2018, s. 7-8.

Witkowska D., Kompa K., Wiśniewski T. (2019) Różnorodność władz spółek giełdowych. Parkiet, 8.03.2019, s. 15.

THE SIGNIFICANT IMPACT OF WOMEN ON THE COMPANIES' FINANCIAL PERFORMANCE. TRUE OR FALSE ON THE POLISH CAPITAL MARKET

Abstract: The purpose of the study is to check whether the presence of women in the authorities of public companies affects their financial performance. The study covered companies continuously listed on the Warsaw Stock Exchange in 2010-2019. The structure of the management was determined as of 30.06. of the following years, and the assessment of the financial standing of the companies at the end of the calendar year. Multidimensional synthetic vector measures were used to assess the standings of the surveyed companies and their correlations with the percentage of women in statutory bodies were examined. The results do not confirm the positive correlation between the increase in the percentage of women in statutory bodies and the financial performance of the companies.

Keywords: women in statutory bodies, public companies, women and financial performance, multivariate comparative methods

JEL classification: C38, G38, L25, M14

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 2, 2024, pp. 58 – 69
Received: 01.04.2024
Accepted: 18.05.2024

EFFECTIVENESS OF VARIABLE SELECTION METHODS FOR MACHINE LEARNING AND CLASSICAL STATISTICAL MODELS

Urszula Grzybowska https://orcid.org/0000-0001-7342-5382
Marek Karwański https://orcid.org/0000-0001-5192-7920
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Instytut Informatyki Technicznej
e-mail: urszula_grzybowska@sggw.edu.pl; marek_karwanski@sggw.edu.pl

Abstract: In line with new international financial supervision directives (IFRS9), banks should look at a new set of analytical tools, such as machine learning. The introduction of these methods into banking practice requires reformulation of business goals, both in terms of the accuracy of predictions and the definition of risk factors. The article compares methods for selecting variables and assigning "importance" in statistical and algorithmic models. The calculations were carried out using the example of financial data classification for loan default. The effectiveness of various machine learning algorithms on selected sets of variables was compared. The results of the analyzes indicate the need to revise the concept of the "importance" of a variable so that it does not depend on the structure of the model.

Keywords: variable selection, machine learning, variable importance

JEL classification: C45, C52, C55

INTRODUCTION

Classical statistical methods have had well established model selection measures and variable significance tests for decades now. Model selection can be done based on AIC or BIC criteria. On other hand variable significance can be obtained in machine learning models.

Machine learning models are better suited for large data sets with many observations.

https://doi.org/10.22630/MIBE.2024.25.2.6

In this paper we present the application of different methods for variable selection using a loan defaults dataset with 88 variables as an example. Our aim is to present and apply contemporary variable selection methods in an economic context and compare it with classical statistical approach. We show the application of methods such as LASSO, Ridge or recursive feature elimination, and the determination of variable importance in algorithmic models: Random Forests, Gradient Boosting, XGBooost and Neural Networks (NN). We also investigate the impact of the number of observations on the variable selection process. For the resulting sets of variables, we compare the effectiveness of the algorithmic methods and logistic regression. In the first section of the paper, we present different ways of selecting variables for the model, together with a review of the literature on this issue. In the second section, we describe the data. In the third chapter, we present computational results on both the selection of variables and the performance of the machine learning models on the considered sets of variables. The last chapter is devoted to conclusions and a summary.

The calculations were performed in Python ver. 3.9.

METHODS

Variable selection related to data dimension reduction is crucial. It allows the elimination of irrelevant, redundant variables, avoids overtraining the model, increases computational speed and allows better interpretation of results. There are many strategies of variable selection and an overview of them can be found in numerous publications (see [Bag et al. 2022, Li et al. 2017; Pudjihartono et al. 2022; Jia et al. 2022; Zebari et al. 2020, Sauerbrei et al. 2020]). The methods can be divided into methods related to the data model so-called wrapper methods (e.g. recursive feature elimination, heuristic methods) or embedded methods (Random Forests, LASSO, Ridge Regression) [Lal et al. 2006] and model-independent filtering methods [Sánchez-Maroño et al. 2007; Hopf 2021] e.g. based on variable correlation or mutual information (MI) measures [Vergara 2014; Gajowniczek et al. 2022].

Statistical modelling

Hypothesis testing is the most common criterion for variable selection in practical statistical modelling problems. Iterative testing of models is performed through forward selection or backward selection algorithms, depending on whether one starts with an empty model or a model with all variables that can be considered. While significance criteria are typically used to include or exclude variables from a model, information criteria focus on selecting a model from a set of plausible models. Including more variables in the model increases the fit of the model. Unfortunately, such a fit is not always desirable, as it leads to an increase in fitting error. Information criteria have been developed to avoid this apparent fitting effect leading to the selection of more complex models. For example, AIC or BIC statistics are used as information criteria.

Variable selection can also be carried out using a strategy based on the so-called regularisation operator (LASSO), which involves imposing additional conditions on the error function when calculating regression coefficients [Hastie et al. 2008; Hastie et al. 2015, p. 32]. LASSO models are widely used in multivariate problems. Regression coefficients estimated by LASSO procedures are biased but may have a smaller mean square total error than by conventional estimation. Because of the loading, their interpretation in explanatory or descriptive models is difficult, and confidence intervals based on resampling procedures such as bootstrap do not reach their stated nominal level [Taylor, Tibshirani 2015].

Variable selection in algorithmic modelling

In the case of algorithmic models, dedicated ways of assigning importance to predictor variables are created. [Elith et al 2008; Adler, Painsky 2022]. Values that measure the weights of variables help in the interpretation of the data, as well as ranking the variables and facilitating the selection of variables into the model. Within machine learning models such as Decision Trees, Random Forests, Gradient Boosting, XGBoost or LightGBM variable importance is measured based on the number of times a variable is selected for a split [Elith et al 2008; Ben Jabeur et al 2023]. CatBoost [Dorogush et al. 2017; Ostroumova et al. 2018] is a relatively new tree based algorithm that has been designed to deal with categorical features. Unlike other tree based algorithms, it does not require categorical features to be one-hot-encoded. As a result it enables easier interpretation of feature importence. Within algorithmic modelling, we can also use one of the wrapper methods, recursive feature elimination (RFE). In this method, smaller and smaller subsets of variables are considered, the least important variables are removed from the current set based on a measure of importance, until a predetermined number of variables is reached [Kohavi, John 1997; Priyatno et al. 2024]. The approach can however be computationally expensive. For neural networks, one way to select variables is the VIANN method based on a modification of Welford's algorithm [De Sa 2019]. To assess the validity of the variable x_s, we use a measure based on the averaged variance of the changes in the weights-parameters of the first hidden layer network connected to the input variable x_s throughout the back-propagation process. This means that the final validity score of the variable will depend on both the final weights-parameters and the variance of them during training. It is assumed that the more the weight $w_{(a,b)}$ of the connection (a, b) varies during the learning phase, the greater the importance of node a in the prediction process. Using VIANN, we need to determine at which learning stages we update the variance. Several options can be considered, due to iteration (after each batch), per epoch or user-defined interval. For simplicity in the paper, we update the variance of the weights at each epoch. Feature importance ranking (FIR) for deep learning models has been described in [Wojtas, Chen 2020].

In addition to the way variables are selected within machine learning methods, there are also hybrid methods that, for example, combine filtering methods with methods based on machine learning approach. Recently, heuristic methods for selecting variables for the model have been used for large sets [Jia et al. 2022].

We will use a hybrid model in our work. We will first eliminate highly correlated variables, quasi-constant variables and variables with high VIF. For the remaining continuous variables, we will select variables using the LASSO method and Ridge Regression in Logistic Regression. Logistic Regression enables regularization that helps to avoid overfitting and can be used for variable selection similarly as it is in linear regression. We also calculate the importance of variables in common machine learning models.

DATA DESCRIPTION

In the research we have used a loan defaults dataset. The set contained 155 572 observations, among them 15802 defaulted. The observations were described by 88 continuous variables. Among the variables 15 were created artificially and had no impact on the target variable Default. The following subsets were used for the calculations:

- Set 1: 250 observations
- Set 2: 500 observations
- Set 3: 1000 observations.

For each set seperately relevant, i.e., influencial features were searched for.

RESULTS

There were 88 continuous variables in the database. In a first step, correlated variables were eliminated, taking a threshold value of 0.9. Variables with a high VIF were also removed, taking a threshold value of 20. Also quasi-constant variables were removed. The remaining variables were taken for further analysis. Further variable selection was done using machine learning methods and also using L1 regularisation (LASSO) and L2 regularisation (Ridge regression, Ridge) in logistic regression, which are embedded methods. The importance of variables was determined by machine learning methods, i.e. using Random Forest (RF), XGBoost algorithm (XGB), Gradient Boosting (GB) and Neural Networks (NN). As a result of a preliminary elimination of correlated variables, varaiables with low variability (quasi-constants), and variables with high VIF, we have obtained the following sets of features: for Set 1 there were 57 explanatory variables distinguished, for Set 2 there were 58 explanatory variables left and for Set 3 there were 49 explanatory variables left.

Statistical feature extraction

A logistic regression model was built for each set of variables in turn, taking into account the Firth correction [Firth 1993; Puhr et al. 2017] for sets 2 and 3. Forward selection for Set 1 performed for 57 distinguished variables left hardly 3 significant variables. The results are shown in Table 1. The AUC for this model is 1 which indicates perfect classification.

Table 1. Results of logistic regression performed for Set 1

Summary of Forward Selection					
Step	Effect Entered	DF	Number In	Score Chi-Square	Pr > ChiSq
1	x38	1	1	155.701	<0.0001
2	x39	1	2	87.8676	<0.0001
3	x34	1	3	6.2232	0.0126

Source: own calculations

Model built for x38 and x39 with Firth correction provided the following results with AUC=0.9988.

Table 2. Results of logistic regression for selected variables in Set 1

Analysis of Penalized Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-12.4444	4.3187	8.3033	0.004
x38	1	1.3209	0.5184	6.4918	0.0108
x39	1	0.9806	0.381	6.6251	0.0101

Source: own calculations

Logit function is of the form $g(x) = -12.4444 + 1.3209\mathrm{x}_{38} + 0.9806\mathrm{x}_{39}$

The probability of default for client with relevant values of x38 and x39 can be evaluated as

$$\pi(x) = \frac{\exp(g(x))}{1+\exp(g(x))}.$$

$\exp(\beta_1) = \exp(1.3187) = 3.75$. This value has an economic interpretation. Namely, the increase of x38 by one unit increases the odds of default almost 4 times.

Model built for Set 2 with 58 variables with Firth amendment provided results presented in Table 3. Only 2 variables were significant. Area under the ROC curve for Set 2 was 0 0.9930, which is very good.

Table 3. Results of logistic regression for selected variables in Set 2

Analysis of Penalized Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-6.732	0.9399	51.2996	<0.0001
x38	1	0.2424	0.0723	11.2429	0.0008
x39	1	1.0281	0.165	38.8434	<0.0001

Source: own calculations

Model built for Set 3 with 49 variables with Firth amendment provided results presented in Table 4. Only 3 variables were significant. For Set 3 AUC was 0.9859.

Table 4. Results of logistic regression for selected variables in Set 3

Analysis of Penalized Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-7.1704	0.7583	89.42	<0.0001
x38	1	0.3636	0.0799	20.7128	0.0001
x39	1	0.4793	0.126	14.4618	<0.0001
x40	1	0.4716	0.1065	19.6009	<0.0001

Source: own calculations

Feature selection by machine learning methods

Preliminary selection left quite large numbers of variables for each set. Therefore, we have performed regularization methods to decrease the numbers of features for further selection. We have also calculated feature importance for preliminary selected sets of features, using Random Forests, Gradient Boosting, XGB and Neural Networks.

Set 1

L1 selection (LASSO) with C=2 distinguished variables x3, x27, x34, x38, x39, x53, var11 and L2 (Ridge) with C=0.01 distinguished: x8, x23, x24, x25, x29, x34, x43, x44, x46, x48, x52, x55, x56, x57, x59, var1, var2, var4, var5, var7, var8, var11, var12, var13, var14, var15. The hyperparameter C has been tuned. For a selected value C=1.5 the following features were extracted: x3, x8, x12, x23, x24, x27, x29, x34, x38, x39,x51, x52, x59, var3, var10, var11, var12, var13,var14. The features that appear in each selected set are x38 and x39.

Set 2

L2 selection with C = 0.01 distinguished the following set of features x8, x20, x23, x24, x25, x30, x34, x39, x40, x43, x44, x51, x54, x56, x57, x59, var4, var7, var9, var11, var12, var13. L1 selection with C=0.1 distinguished x8, x23, x34, x38, x39, x40, x54, x57, x59.

Set 3

L2 (Ridge regularization) with hyperparameter C=1.5 distinguished the following features: x3, x8, x13, x23, x28, x38, x39, x40, x54, var11, var13. L1 with C = 0.01 selected the following features: x23, x34, x38, x39, x40, x43, x44, x54, x57, x59. The features that appear in each selected set are x38, x39 and x40.

We have performed recursive feature elimination for Random Forests. The number of features to be slected was set to 5. The results are shown in Table 5. Feature x38 and x39 appear in each selected set.

Table 5. Features selected by Recursive feature elimination

	Set 1	**Set 2**	**Set 3**
RFE	x34, x38, x39, x52, var3	x38, x39, x40, x46, x54	x23, x38, x39, x40, x54

Source: own calculations

Features distinguished by calculating feature importance in the most popular machine learning algorithms are presented in Table 6. Feature x38 and x39 appear in each set. Additionally, x40 appears in each set for Set 2 and Set 3.

Table 6. Features distinguished by calculating feature importance

Method	**Set1**	**Set2**	**Set3**
RF	x54, x52, x53, x42, x39, x38	x42, x38, x39, x40	x42, x39, x38, x40
GB	x53, x39, x38	x38, x39, x40, x44, x46	x38, x39, x40
XGB	x38, x39	x38, x39, x40, x44, x46	x40, x39, x19, x44, x38

Source: own calculations

Table 7. First 10 features distinguished by calculating relative feature importance in NN

	Set 1		Set2		Set3	
	Name	Relative Importance	Name	Relative Importance	Name	Relative Importance
1	x39	1.0	x39	1.0	x40	1.0
2	x38	0.77	x40	0.75	x38	0.95
3	x19	0.35	x38	0.73	x39	0.85
4	x5	0.27	x18	0.31	x67	0.18
5	x27	0.26	x27	0.29	x34	0.17
6	x18	0.25	x7	0.22	x52	0.15
7	x9	0.24	x5	0.16	x48	0.15
8	x20	0.23	var11	0.16	x3	0.12
9	x47	0.19	var9	0.16	var11	0.12
10	var11	0.19	x3	0.15	x60	0.12

Source: own calculations

Each of the machine learning methods used distinguished different variables, although some variables, e.g., x38 and x39, are repeated in each ranking. It is worth stressing that machine learning methods also distinguished simulated variables (var), although they appear at the end of rankings.

Finally, selected machine learning models performance in terms of accuracy was compared. We have applied Logistic Regression (LR), Random Forests (RF), Gradient Boosting (GB), XGBoost, AdaBoost (AB) and Extra Trees (ET). The results are shown in Table 8. The performance of various models is different on each data set. There is no best model for each set, although ExtraTrees (ET) have the best accuracy for Set 1 and Set 3. One can however notice, that Logistic Regression treated as a machine learning model exhibits the worst performance in all cases.

Table 8 Classification results in terms of accuracy of methods performed for various sets of variables

	Set 1		Set 2		Set 3	
Variables	x34, x38, x39		x38, x39, x40, x46, x54		x23, x38, x39, x40, x54	
	Accuracy	STD	Accuracy	STD	Accuracy	STD
LR	0.935	(0.1026)	0.9525	(0.0425)	0.96	(0.0236)
RF	0.985	(0.0229)	0.9775	(0.0284)	0.9813	(0.0151)
GB	0.98	(0.0245)	0.9675	(0.0372)	0.98	(0.0139)
XGB	0.97	(0.04)	0.9725	(0.0261)	0.9788	(0.0148)
AB	0.985	(0.0229)	0.97	(0.0312)	0.9738	(0.0181)
ET	0.99	(0.03)	0.9725	(0.0236)	0.9825	(0.0127)

Source: own calculations

CONCLUSIONS

One of the basic elements of building models is the selection of appropriate independent variables.

Independent variables are selected to represent expected influences based on: theory (often relatively weak), previous research, and local context (in time and space). In the statistical approach, the main emphasis is placed on the sign, magnitude and statistical significance of the weights for the independent variables. Algorithmic models require different approaches.

The article presents the results of the selection of variables used in practice and dedicated to specific types of models. The work was carried out on a relatively large data set to avoid problems related to the so-called low power effects.

Feature importance computed on large feature sets produced stable results. The number of selected features is small and some of them are repeated in different analyses. It can be stated that up to a certain limit value of the so-called of practical importance, various selection algorithms correctly identify relationships between independent variables and the target variable. The situation is different in the case of a weak relationship, where there is a problem of the so-called multiplicity of data models. This means that prediction accuracy becomes more robust as the set of independent variables changes. Unfortunately, this property of the models makes it much more difficult to correctly interpret the results from a substantive point of view.

Based on the results obtained, it can be concluded that there are large differences between the models. It can therefore be concluded that attempts to reconcile results between different analytical approaches must be carried out very carefully and should take into account the fact that the definitions of "variables of significance" are strongly dependent on the model.

REFERENCES

Adler A. I., Painsky A. (2022) Feature Importance in Gradient Boosting Trees with Cross-Validation Feature Selection. Entropy, 24(5), 687. https://doi.org/10.3390/e24050687

Bag S., Gupta K., Deb S. (2022) A Review and Recommendations on Variable Selection Methods in Regression Models for Binary Data. https://arxiv.org/pdf/2201.06063

Ben Jabeur S., Stef N., Carmona P. (2023) Bankruptcy Prediction using the XGBoost Algorithm and Variable Importance Feature Engineering. Computational Economics, 61, 715-741. https://doi.org/10.1007/s10614-021-10227-1

Breiman L. (2001) Statistical Modeling: The Two Cultures. Statistical Science, 16(3), 199-215.

Dorogush A.V., Ershov V., Gulin A. (2017) CatBoost: Gradient Boosting with Categorical Features Support. Workshop on ML Systems at NIPS 2017.

De Sa C. R. (2019) Variance-Based Feature Importance in Neural Networks. [in:] Kralj Novak P., Šmuc T., Džeroski S. (eds) Discovery Science, Lecture Notes in Computer Science. 11828, Springer, Cham. https://doi.org/10.1007/978-3-030-33778-0_24

Engelmann B., Rauchmeier R. (2011) The Basel II: Risk Parameters. Estimation, Validation, Stress Testing - with Applications to Loan Risk Management. Springer Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-16114-8

Elith J., Leathwick J. R., Hastie T. (2008) A Working Guide to Boosted Regression Trees. Journal of Animal Ecology, 77, 802-813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Firth D. (1993) Bias Reduction of Maximum Likelihood Estimates. Biometrika, 80(10), 27-38. https://doi.org/10.2307/2336755

Gajowniczek K. et al. (2022) HOFS: Higher Order Mutual Information Approximation for Feature Selection in R. SoftwareX, 19, 1-9. https://doi.org/10.1016/j.softx.2022.101148

Hastie T., Tibshirani R., Friedman J. (2008) The Elements of Statistical Learning (2nd ed.). Springer.

Hastie T., Tibshirani R., Wainwright M. (2015) Statistical Learning with Sparsity. The Lasso and Generalizations. New York Chapman & Hall/CRC. https://doi.org/10.1201/b18401

Hopf K., Sascha R. (2021) Filter Methods for Feature Selection in Supervised Machine Learning Applications-Review and Benchmark. arXiv preprint arXiv:2111.12140, 2021

Jia W., Sun M., Lian J. et al. (2022) Feature Dimensionality Reduction: a Review. Complex Intell. Syst., 8, 2663-2693. https://doi.org/10.1007/s40747-021-00637-x

Kohavi R, John G. H. (1997) Wrappers for Feature Subset Selection. Artificial Intelligence, 97(1-2), 273-324. https://doi.org/10.1016/S0004-3702(97)00043-X

Lal T. N., Chapelle O., Weston J., Elisseeff A. (2006) Embedded Methods. [in:] Guyon I., Nikravesh M., Gunn S., Zadeh L. A. (eds) Feature Extraction. Studies in Fuzziness and Soft Computing, 207, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-35488-8_6

Li J., Cheng K., Wang S., Morstatter F. (2017) Feature Selection: A Data Perspective ACM Computing Surveys, 50(6), Article 94, 1-45. https://doi.org/10.1145/3136625

Olden J., Joy M., Death R. (2004) An Accurate Comparison of Methods for Quantifying Variable Importance in Artificial Neural Networks using Simulated Data. Ecological Modelling, 178(3-4), 389-397. https://doi.org/10.1016/j.ecolmodel.2004.03.013

Ostroumova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A. (2017) CatBoost: Unbiased Boosting with Categorical Features. NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems December 2018, Pages 6639–6649. https://arxiv.org/pdf/1706.09516

Priyatno A.,Widiyaningtyas T. (2024) A Systematic Literature Review: Recursive Feature Elimination Algorithms. JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer). 9. 196-207. https://doi.org/10.33480/jitk.v9i2.5015

Pudjihartono N., Fadason T., Kempa-Liehr A. W., O'Sullivan J. M. (2022) A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. Frontiers in Bioinformatics, 2(927312). https://doi.org/10.3389/fbinf.2022.927312

Puhr R, Heinze G, Nold M, Lusa L, Geroldinger A. (2017) Firth's Logistic Regression with Rare Events: Accurate Effect Estimates and Predictions? Stat Med. 36(14), 2302–2317. https://doi.org/10.1002/sim.7273

Sánchez-Maroño N., Alonso-Betanzos, A., Tombilla-Sanromán, M. (2007) Filter Methods for Feature Selection – A Comparative Study. [in:] Yin, H., Tino, P., Corchado, E., Byrne, W., Yao, X. (eds) Intelligent Data Engineering and Automated Learning - IDEAL 2007. IDEAL 2007. Lecture Notes in Computer Science, 4881, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-77226-2_19

Sauerbrei W., Perperoglou A., Schmid M. et al. (2020) State of the Art in Selection of Variables and Functional Forms in Multivariable Analysis-Outstanding Issues. Diagn Progn Res 4, 3. https://doi.org/10.1186/s41512-020-00074-3

Taylor J., Tibshirani R. J. (2015) Statistical Learning and Selective Inference. Proc Natl Acad Sci U S A, 112(25), 7629-34. https://doi.org/10.1073/pnas.1507583112

Wojtas M., Chen K. (2020) Feature Importance Ranking for Deep Learning. 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada. https://doi.org/10.48550/arXiv.2010.08973

Vergara J. R., Estévez P.A. (2014) A Review of Feature Selection Methods based on Mutual Information, Neural Comput & Applic, 24, 175-186. https://doi.org/10.1007/s00521-013-1368-0

Zebari R., Abdulazeez A., Zeebaree D., Zebari D., Saeed, J. (2020) A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction, Journal of Applied Science and Technology Trends, 1(2), 56-70. https://doi.org/10.38094/jastt1224

APPENDIX

Description of variables distinguished in various models

Variable	Description
x3	Ratio of the sum of debit balances at the moment of analysis to the average debit balance in the last 12 months
x5	Linear trend of average monthly payments for the period - 6 months
x7	Increasing trend for average monthly payments for the period - 6 months
x8	Ratio of the sum of balances at the moment of analysis to the average total balance over the last 12 months
x9	Average debit balance over the last 3 months
x13	Average increase in the amount of capital arrears in the last period
x19	Average credit balance over the last 3 months
x20	Average credit balance over the last 6 months
x23	Ratio avg. the amount of overdue capital installments up to avg. amount of the total credit balance (on all customer accounts) in the last month) for - 6 months
x24	Ratio avg. the amount of overdue capital installments up to avg. amount of the total credit balance (on all customer accounts) in the last month) for - 9 months
x25	Ratio avg. the amount of overdue capital installments up to avg. amount of the total credit balance (on all customer accounts) in the last month) for - 12 months
x27	Average time of delay in repayment of installments, determined as the ratio of the sum of days of delay for all installments paid to the number of all installments repaid.) for - 12 months
x28	Average time of delay in repayment of installments, determined as the ratio of the sum of days of delay for all installments paid to the number of all installments repaid.)
x29	Number of overdue accounts as at the date of analysis) for - 6 months
x30	Ratio of the sum of debit balances at the moment of analysis to the average debit balance in the last 12 months
x34	Number of overdue accounts as at the date of analysis) for - 9 months
x38	Sum of amounts from all months of repayments made by the client on credit accounts held by him) for - 3 months
x39	Sum of amounts from all monthly repayments made by the client on account of credit accounts held by him) for - 6 months
x40	Sum of all monthly repayments made by the client on credit accounts held by him/her) for - 9 months
x42	Sum of all monthly repayments made by the client on credit accounts held by him)
x43	Sum of amounts from all months of repayments made by the client on credit accounts held by him) for - 1 month
x44	Sum of amounts from all months of repayments made by the client on credit accounts held by him) for - 1 month
x46	Sum of amounts from all months of repayments made by the client on credit accounts held by him) for - 1 month
x47	The sum of interest arrears at the moment of analysis
x48	Loan repayment ratio, determined as the ratio of the value of all repaid loans to the value of all loans taken/granted) for - 12 months
x51	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 3 months

Variable	Description
x52	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 6 months
x53	Total amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 9 months
x54	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 12 months
x55	Total amounts of all outstanding repayments incurred on all customer credit accounts during the last month)
x56	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 1 month
x57	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 1 month
x59	Sum of the amounts of all overdue repayments incurred on all customer credit accounts in the last month) for - 1 month
x60	Increasing trend in average monthly unused limits on all short-term loans over the last 12 months
x66	Amount of all outstanding receivables falling within the range (over 90 days) (summing over all customer credit accounts))
x67	Average monthly unused limit on all short-term loans over the last 6 months

Artificially created variables have been added to business variables: var1 - var15: independent variables with uniform distribution

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 2, 2024, pp. 70 – 85

Received: 29.04.2024
Accepted: 14.06.2024

POST-MERGER FINANCIAL PERFORMANCE – A STUDY OF HIGH-TECH COMPANIES IN THE UNITED STATES USING ARTIFICIAL NEURAL NETWORKS

Stanislav Tarasov https://orcid.org/0000-0001-7018-9847
Faculty of Economic Sciences
University of Warsaw – UW
e-mail: swtarasow@gmail.com
Bartłomiej Dessoulavy-Śliwiński https://orcid.org/0000-0002-4395-9815
Department of Management and Information Technology
Faculty of Economic Sciences
University of Warsaw – UW
e-mail: bsliwinski@wne.uw.edu.pl

Abstract: The debate about the efficacy of mergers and acquisitions as a growth strategy in terms of ex-post value creation has been developing for decades. This paper aims to create an artificial neural network that examines trends in the financials and marks the potential sources of value creation in the high-tech industry mergers between 2011 and 2021. The findings demonstrate that ANN can be implemented as a highly efficient model for analyzing complex financial events due to its flexibility and lack of prior assumptions about the data.

Keywords: merger, artificial neural network, financial performance, machine learning, high-tech industry

JEL classification: G34, C45

INTRODUCTION

The most common theoretical justification for mergers and acquisitions is that the value of two firms combined is greater than their individual parts (i.e., 2 + 2 = 5) [King et al. 2004]. However, when it comes to the effectiveness of mergers and acquisitions as a growth strategy, researchers have been debating whether M&A create value or destroy it for decades. The literature on M&A performance can be

https://doi.org/10.22630/MIBE.2024.25.2.7

divided into two streams. Originally, the market approach examining stock market performance was dominant among scholars. Rau and Vermaelen [1998] argue that the conclusion regarding M&A value creation for the bidder and target shareholders is based on the results of short-term event studies that find returns to bidders to be small or insignificantly different from zero. For instance, Firth [1980] in a short-term event study reports positive, though small or insignificant abnormal returns using the market approach, and Agrawal et al. [1992] show that half of the acquiror's shareholders can obtain positive abnormal returns. At the same time, Asquith [1983] and André et al. [2004] find short-term negative returns using an event-centered market approach. However, as repeatedly recalled in more recent studies, short-term analyses may not fully reflect the impact of M&A on a business combination. As a result, much attention has also been paid to the long-term effects that M&A have on performance. Agrawal et al. [1992], Anderson and Mandelker [1993], Loughran and Vijh [1997], and Rau and Vermaelen [1998] report long-term statistically significant negative abnormal returns related to post-M&A performance of the bidders. In parallel, Langetieg [1978], Bradley and Jarrell [1988], Frank, Harris and Titman [1991], and Loderer and Martin [1992] do not find significant changes in the long-term post-M&A performance of participating firms. However, it is crucial to consider that there are methodological concerns when it comes to measuring the market performance of a company. The short-term approach, announcement returns studies, may be biased due to price pressure around M&A deals, information asymmetry, or market inefficiencies; while another approach, the computation of long-run abnormal returns, may be biased due to unobserved differences between the firms that merge and those that do not, and consequently due to the inability to accurately estimate where the abnormality of returns starts [Malmendier et al. 2018]. Moreover, if performance measurements are based solely on the market's reaction to a merger announcement or multi-dependent abnormal returns, it makes it problematic to identify the drivers behind the value creation or destruction process of mergers.

Consequently, studies of both short- and long-term operating performance appeared, which generally rely on the accounting approach to measure performance. These studies have attempted to discover the sources of M&A value creation and to determine whether the expected economic gains at announcement are realized. For instance, Ravenscraft and Scherer [1987] conclude that mergers destroy value on average, while Healy, Palepu and Ruback [1992] state that merged firms experience improvements leading to higher operating cash flows compared to their industry benchmark. However, limited data availability, inconsistent sets of performance factors, and potentially mismatched control groups raise concerns about the reliability of results.

Eventually, a conclusion was drawn that large-sample studies – whether following the market or accounting approach – could be unable to capture the richness of the economic effects of mergers and could capture neither the direction of these effects, nor their determinants (see [Kaplan 2000, Shao et al. 2021]). In this study, certain sampling choices are made to address this problem. Initially, a specific subgroup of mergers is considered; only mergers in the High Technology sector are selected because high-tech firms can provide a unique perspective on market value creation, with their long-term performance being related to certain factors that are attributable to virtual network effects [Léger and Quach 2009]. At the same time, selecting a homogeneous sample of companies allows to examine the nature of the deals more accurately (see [Hackbarth and Morellec 2008, Shao et al. 2018]).

It is worth noting that existing studies mainly applied linear methods to analyze mergers that in practice do not necessarily have effects that can be linearly approximated. In the era of big data, machine learning and data mining methods are often being used to analyze financial time series. Artificial neural networks (ANN) are a preferred tool for many predictive data mining applications because of their power and flexibility. To exemplify, Teräsvirta et al. [2005] explore the predictive power of ANNs for macroeconomic series, and Yu et al. [2007] test them in foreign exchange markets. Le and Viviani [2018] perform a comparison of traditional statistical and machine learning methods in predicting bank failure, showcasing the superiority of the latter. Bouteska et al. [2023] develop a focused time-delayed neural network to challenge the nonlinearity in energy commodity price formation. As demonstrated, ANNs are particularly useful in applications where the underlying process is complex.

The purpose of this study is to offer an analysis of post-merger performance, limited to a sample of deals in the high-tech industry, by utilizing ANN models with cross-validation and assessing their accuracy metrics. Additionally, this study marks areas of potential sources of ex-post market performance, analyzing what distinguishes more successful mergers from less successful ones with regard to their market valuation.

DATA

Data selected for the sample

The dataset of mergers is obtained from the Eikon database. The data must meet the following criteria: 1) The announcement and completion years of mergers fall within the period of 2011-2021; 2) Mergers are listed as completed; 3) Mergers are in the High Technology sector; 4) Mergers are between publicly traded companies; 5) Only domestic mergers are included; 6) Overlapping cases (if the acquiror engages in several mergers during the analyzed period) are included.

A deal is classified as high-tech if that is the industry of the target's main economic activity. It is an important filter as it neutralizes the industry-clustering effect in analyzing the structure and efficiency of mergers and the consequent differences in results (see [Mitchell and Mulherin 1996, Andrade et al. 2001, Ahern and Harford 2014]). Moreover, it allows to adequately consider the deals with a conglomerate acquiror. Only mergers between publicly traded companies are taken into consideration, so that the companies have market data and financial statements available for the analyzed period. Domestic mergers are selected for the purpose of avoiding cross-border influence (see [André et al. 2004, Jensen-Vinstrup et al. 2018]). The described approach allows to draw conclusions about the internal effects of mergers by naturally creating an appropriate industry benchmark.

Financial data for all acquired and acquiring companies is obtained from the Bloomberg Terminal. In this study, in order to analyze the effects that mergers have on the market performance of a company, focus is placed on selected financial ratios. The ratios are calculated using the data extracted from financial statements for 6 consecutive years, starting from the last complete fiscal year before a transaction occurs (t, reference point), and for the following five years after the transaction is completed (t+1– t+5).

The ratios used in the study are:

- Profitability ratios – earnings per share (EPS), return on assets (ROA), and return on equity (ROE),
- Liquidity ratio – current ratio (CR),
- Solvency ratio – total debt ratio (TD),
- Market value ratios – capitalization per share (CPS), and price-to-book ratio (P/B).

Two-sample t-tests for performance change significance

Average reference and post-merger financial ratios are compared in pairs (t and t+1, t and t+3, t and t+5) for each sample using paired two-sample t-tests to find whether there are significant changes in financial performance, with the hypotheses being:

- H_0: There is no difference in a US high-tech company's financial performance following a merger.
- H_a: The ex-post financial performance of a US high-tech company changes after engaging in a merger.

The descriptive statistics and results of the tests (at 0.01 significance level) are presented in Table 1. After removing outliers for the reference variables at (t) point in time, the analyzed sample consists of 56 mergers (between a total of 84 companies). To address influential cases for the variables at other points in time (t+1–t+5), 90% winsorization is carried out in order not to be overly exclusive of the observations. Examining the market value ratios, the tests show an insignificant decrease of P/B ratio in the short run, in the following year, and insignificant increases in the longer run, three and five years after the transaction. At the same time, the tests show substantial growth of CPS in the years following the transaction, both in the short and the long run. The comparison of the pre-merger and post-merger profitability ratios shows that returns on equity and assets note statistically significant decreases in the year following the transaction, and EPS show a moderate but insignificant increase five years after the event. These results imply that the analyzed mergers do not bring superior profitability for the business.

Table 1. Descriptive statistics and t-test results for the financial ratios analysis

Panel A. Descriptive statistics

Ratios		(t)		(t+1)		(t+3)		(t+5)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Profitability	EPS	1.385	1.469	1.181	1.854	1.679	1.683	1.870	1.925
	ROA	0.057	0.062	0.034	0.058	0.052	0.049	0.043	0.044
	ROE	0.118	0.152	0.070	0.124	0.130	0.143	0.263	0.583
Liquidity	CR	2.911	1.981	2.636	1.692	2.515	1.640	2.131	1.082
Solvency	TD	0.487	0.183	0.539	0.156	0.567	0.164	0.592	0.202
Market	P/B	3.686	2.713	3.661	2.733	3.875	2.349	5.076	5.446
	CPS	33.123	19.853	34.751	23.363	46.104	30.515	60.927	51.806

Panel B. Test results for the financial ratios

Ratios		(t+1, t)				(t+3, t)				(t+5, t)			
		t-value	Sig. (2-tailed)	Lower	Upper	t-value	Sig. (2-tailed)	Lower	Upper	t-value	Sig. (2-tailed)	Lower	Upper
Profitability	EPS	-1.535	0.131	-0.472	0.063	1.857	0.069	-0.023	0.610	2.563	0.013	0.106	0.864
	ROA	-3.952	0.000	-0.035	-0.011	-0.791	0.432	-0.018	0.008	-1.748	0.086	-0.031	0.002
	ROE	-3.616	0.001	-0.075	-0.021	0.671	0.505	-0.025	0.049	1.828	0.073	-0.014	0.305
Liquidity	CR	-1.409	0.164	-0.667	0.116	-1.736	0.088	-0.854	0.061	-3.091	0.003	-1.286	-0.274
Solvency	TD	3.827	0.000	0.025	0.078	5.650	0.000	0.051	0.108	5.361	0.000	0.066	0.144
Market	P/B	-0.089	0.929	-0.583	0.534	0.624	0.535	-0.419	0.797	1.887	0.064	-0.086	2.866
	CPS	1.314	0.194	-0.855	4.111	5.808	0.000	8.502	17.461	4.972	0.000	16.598	39.010

Source: own calculations

CR shows a stable decline in liquidity of the merged firms, which means that their ability to meet financial obligations with available liquid assets decreases each year following the merger. The TD ratio shows significant increases in leverage each year following the merger, which is an important tool for growth, but also implies greater financial risk for a company. Both trends, especially combined, may indicate a weakening position of the merged firms. At the same time, their market valuation represented as CPS increases substantially through the analyzed period. Hence, hypothesis H_0 is rejected since mergers are found to influence financial performance, specifically by decreasing profitability and liquidity, and increasing solvency and market value of the merged firms.

METHODOLOGY

Methodological background

Over the years, statistical parametric models such as linear regressions with various modifications have been used to analyze merger activity and its effects. With recent technological developments, methods such as artificial neural networks have been frequently used by scholars in various fields. They consist of interconnected neurons capable of pattern recognition, prediction, classification, and learning. Each connection between the neurons has an associated weight that signifies the strength and direction (positive or negative) of the influence that one neuron has on another. ANNs learn by iteratively adjusting weights to predict the correct output for a given set of inputs. The knowledge acquired from the input data is therefore stored in a system of neuron connections called synaptic weights. As compared with conventional statistical models, ANNs have several substantial advantages: they are flexible and adaptive, allowing to analyze data without hypothesizing in advance certain relationships between dependent and independent variables. Consequently, if a linear relationship between the variables is appropriate, an ANN would learn the linear structure and approximate a linear regression, and if a nonlinear relationship is relevant, the model would seek the best model structure fitting the data [IBM 2012].

In recent years, neural networks have been gradually applied to merger forecast research. In their large-sample study, Lee et al. [2020] criticize traditional forecasting methods and use neural networks to account for nonlinearity and complexity in outcome data, developing a failure prediction model for M&A. Specifically, by assessing a "withdrawn takeover prediction model" using a neural network with an enhanced logit activation function, they present the most significant variables based on importance analysis and showcase the superiority of neural networks compared to traditional forecasting techniques. Bi and Zhang [2021] using neural networks provide more insight into the issue by assessing and identifying additional variables that contribute to M&A failure prediction models. Applying neural networks, Zhu and Meng [2021] try to assess and interpret synergy effects by analyzing the rate of changes in the selected financial ratios that represent overall post-M&A performance. Hence, following the tests, we train neural network models to examine the data.

Application of Artificial Neural Networks

Artificial neural networks in this study are created in IBM SPSS Statistics. The architecture used is a multilayer perceptron (MLP) – it is a feedforward ANN with three distinct layers: input, hidden, and output, each comprising several neurons and having activation functions. ANNs also have a bias neuron, which allows them to learn underlying patterns in the data and estimate output. Bias can be viewed as analogous to the error of measurement in linear regression modeling. Activation functions connect the weighted sums of units in a layer to the values of units in the succeeding layer.

The research problem is framed as a classification task aimed at distinguishing between successful and less successful mergers. Therefore, during experimentation phase the error (loss) function used is cross-entropy:

$$L = -\frac{1}{N}\left[\sum_{j=1}^{N}\left[t_j \log(p_j) + (1-t_j)\log(1-p_j)\right]\right], \tag{1}$$

where for N datapoints t_j is the truth value taking the value of 0 or 1, and p_j is the softmax probability for the i^{th} datapoint. The activation function used in the hidden layer is hyperbolic tangent:

$$\gamma(c) = \frac{(e^c - e^{-c})}{(e^c + e^{-c})}, \tag{2}$$

which takes real-valued arguments and transforms them to the range (–1, 1). The activation function used in the output layer is softmax:

$$\gamma(c_k) = \frac{\exp(c_k)}{\sum_j \exp(c_j)}, \tag{3}$$

which takes a vector of real-valued arguments and transforms it to a vector which elements fall in the range (0, 1) and sum up to 1.

The sample is divided into two parts for cross-validation purposes, with approx. 70% of the observations used for training, and the remaining 30% used for testing. The type of training used is batch, which updates synaptic weights only after passing through all training data records and is most useful for smaller datasets. It is commonly favored as it directly minimizes total error, and by its nature is not dependent on case order. The optimalization algorithm used with batch is scaled conjugate gradient (SCG), which is based on the second-order gradient supervised learning procedure. This optimalization algorithm utilizes a trust-region step to scale the step length (learning rate), where the distance for which the model function is trusted is updated at each step [Møller 1993]. The model step is used if it lies within that distance; otherwise, an approximate minimum for the model function on the boundary of the trust region is used, thus contributing to robustness and stability of results.

An important element of each classification task is forecast accuracy validation and quality assessment. Most common measures to test classification effectiveness are accuracy coefficients based on confusion matrix such as F1 Score, which can be effective when False Positive (FP) and False Negative (FN) are equally costly and True Negative (TN) is high, and Matthews Correlation Coefficient (MCC), which is a measure of correlation between predicted classes and basic truth and is superior to F1 Score if the classes are of different sizes [Baldi et al. 2000, Powers 2011, Chicco and Jurman 2020]. A reliable illustration of the models' effectiveness is Receiver Operator Characteristic (ROC) curve, which measures sensitivity and specificity of a classifier, and Area Under the Curve (AUC), which measures the ability of a classifier to distinguish between classes. In this study, the criteria proposed by Department of Math of the University of Utah [n.d.] to interpret

AUC are applied: (A) 0.90 – 1 = excellent; (B) 0.80 – 0.90 = good; (C) 0.70 – 0.80 = fair; (D) 0.60 – 0.70 = poor; (E) 0.50 – 0.60 = fail.

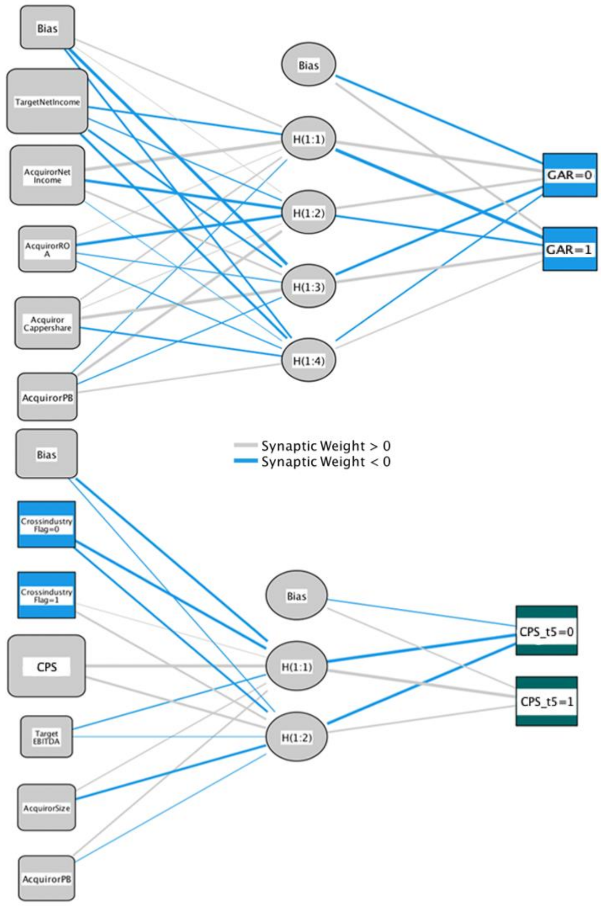
The firms' performance is analyzed in terms of market value (CPS), which is tested to exhibit statistically significant changes of the highest magnitude. In this study, two neural network models are created – the first one is trained to predict the firms' CPS five years after the merger (t+5), and the second one is trained to analyze the geometric average returns (GAR) calculated based on the returns of the firms' CPS from the first to the fifth year after the merger (t+1)–(t+5). Considering that ANNs are effective classification models, CPS (t+5) and GAR are rescaled to binary variables (0 – below the mean, 1 – above the mean in the sample). The purpose of the first model is to predict which business combinations have above-average returns in the long run compared to other mergers in the sample, based on the information available already at the time of the deal's completion. The purpose of the second model is similar, being the prediction of overall cumulated stock performance in the years following the merger.

The set of considered factors is based on the hypothesized impact they can have on stock performance, evidence of which has appeared in M&A studies over the years (e.g. [Cumming et al. 2023] present their bibliometric analysis of key topics in M&A research, including studies on aforementioned factors). As a consequence of limited data availability, an appropriate sample of input variables must be adequate in number to minimize potential underfitting or overfitting of the model. Therefore, independent variable importance analysis (sensitivity analysis) is performed, which computes the importance of each predictor in determining the neural network by investigating the relative contribution of the uncertainty of the input variables on the variability in the output levels [Pianosi and Wagener 2015]. After consecutively adjusting for importance and considering parameter estimates for all available variables during testing, the final sets of five most significant independent variables are selected for each model.

RESULTS

The architecture of the best-performing ANN models is shown in Figure 1. The model for CPS is trained with two neurons, and the model for GAR is trained with four neurons in the hidden layer, excluding Bias. In combination with five input variables, it provides an adequate amount of data for the assessment of each path.

Figure 1. The architecture of the Artificial Neural Networks



* Each model has an input layer (left), a hidden layer with hidden neurons and a bias neuron (middle), and an output layer (right). The top ANN predicts, as an output variable, geometric average returns (GAR) calculated for five years following the merger, and the bottom ANN – capitalization per share (CPS) five years after the merger.

Source: own preparation

Table 2 presents the summary of the ANN models. The model for CPS shows relative errors of 5.7% and 4.8% in the training and testing samples respectively, while the model for GAR reports 7.9% and 11.1% errors respectively. It is worth noting that confusion matrix measures depend considerably on sample size; nevertheless, the models achieved high accuracy in predicting both below and above average cases in both training and testing samples. The models for CPS and GAR report F1 Score of .923 and .918, MCC of .889 and .821, and AUC of .981 and .967 respectively, which is considered excellent according to the accepted criteria.

Table 2. Artificial Neural Networks summary

Panel A. Quality Assessment of ANNs		
Network Type	**MLP - CPS**	**MLP - GAR**
Input Units	5	5
Hidden Layers	1	1
Hidden Neurons (excl. Bias)	2	4
Accuracy	0.946	0.911
Precision	1.000	0.903
Recall	0.857	0.933
F1 Score	0.923	0.918
MCC	0.889	0.821
AUC	0.981	0.967

Panel B. Confusion Matrix					
MLP - CPS			**MLP - GAR**		
Training	**N**	**P**	**Training**	**N**	**P**
N	23	0	N	14	2
P	2	10	P	1	21
Testing	**N**	**P**	**Testing**	**N**	**P**
N	12	0	N	9	1
P	1	8	P	1	7

Source: own calculations

The variable importance analysis results and parameter estimates presented in Table 3 show that the first model estimates the CPS (t) before the merger and Acquiror Size to have the strongest impact on predicting the CPS five years post-merger (t+5). The parameter estimates demonstrate a positive impact of CPS (t) on CPS (t+5) (3.130 in H(1:1) and 1.608 in H(1:2)), and varying influence of Acquiror Size on CPS (t+5) (.539 in H(1:1) and -1.749 in H(1:2)). Cross-industry Flag, Acquiror P/B ratio and Target EBITDA also have a relatively significant influence on CPS (t+5), with a negative impact of cross-industry deals, positive impact of same-industry deals, negative impact of Target EBITDA, and varying influence of Acquiror P/B ratio. These results imply that a high-tech merger involving a relatively smaller bidder and/or a target with lower EBITDA may result in a relatively higher market valuation five years following the merger, which could be attributed to the market putting more value on the growth potential, innovation, or strategic focus rather than the financials.

Table 3. Importance analysis and parameter estimates of the Artificial Neural Networks

Panel A. Importance (sensitivity) analysis and parameter estimates (synaptic weights)					
MLP - CPS					
Predictor		**Hidden Layer**		**Predicted**	
Input Layer		H(1:1)	H(1:2)	CPS(0)	CPS(1)
Importance	(Bias)	-1.749	-0.203		
0.433	CPS (t)	3.130	1.608		
0.170	Acquiror Size	0.539	-1.749		
0.165	Cross-industry Flag (0 = cross-industry)	-2.172	-1.510		
0.165	Cross-industry Flag (1 = same-industry)	0.099	1.367		
0.145	Acquiror P/B	1.153	-0.171		
0.087	Target EBITDA	-0.593	-0.099		
Hidden Layer					
	(Bias)			-0.449	0.709
	H(1:1)			-2.655	2.891
	H(1:2)			-2.406	1.275

MLP - GAR							
Predictor		**Hidden Layer**				**Predicted**	
Input Layer		H(1:1)	H(1:2)	H(1:3)	H(1:4)	GAR(0)	GAR(1)
Importance	(Bias)	0.914	0.071	-2.696	-1.488		
0.267	Target Net Income	-1.672	-0.726	-1.826	-1.966		
0.236	Acquiror Net Income	3.021	-2.565	1.438	-0.073		
0.187	CPS (t)	0.995	0.141	2.869	-1.013		
0.160	Acquiror P/B	-0.461	2.412	-0.620	1.010		
0.150	Acquiror ROA	0.096	-2.661	-0.463	-0.649		
Hidden Layer							
	(Bias)					-1.958	1.899
	H(1:1)					2.699	-3.435
	H(1:2)					1.967	-1.742
	H(1:3)					-2.322	2.685
	H(1:4)					-0.979	0.827

Source: own calculations

Market perception and investor sentiment often play an important role, as investors can perceive smaller, high-growth companies as more agile and capable of achieving strong future performance, which can lead to a more favorable market valuation. At the same time, the results regarding the acquiror's capitalization before the merger imply that the higher it is, the higher the market valuation should be five years post-merger, which could be attributed to investors overextrapolating past company performance and perception when assessing the benefits of a merger and future performance. The Cross-industry Flag coefficients indicate that deals in the same industry account for higher valuation, while cross-industry deals account for lower valuation, which could be explained by certain expectations regarding enhanced market presence, the knowledge of the industry, and expertise in running a business and streamlining operations in that industry – or lack thereof.

The increased number of hidden neurons in the second model suggests more complex relationships between the variables. Sensitivity analysis shows that the profitability of bidders and targets becomes most significant. The influence of all independent variables on GAR varies, having opposite directions in different hidden neurons, except for Target Net Income, which has a negative impact on the dependent variable. In combination with the different influence directions that Acquiror Net Income and Acquiror ROA have on GAR, it is indicated that lower profitability of both sides of mergers contributes to higher market growth in the following years. The managers and large shareholders of the companies with lower profitability could be more prudent before approving a critical transaction that may determine the firm’s future, looking for potential operational improvements, cost synergies, and increased efficiency, thus allowing the merged entity to unlock and realize latent value and growth opportunities that were previously untapped. At the same time, the management of the companies with higher profitability could often look for an opportunity to invest their excess assets and easily acquire similar companies with high profitability, not necessarily considering other aspects of merging and potential implications of such a decision on the future of the business. A generally positive influence of CPS (t) on GAR could indicate that an already established strong market position is a prerequisite for higher market value growth in the following years.

SUMMARY

Using a sample of 56 high-tech mergers between 84 companies in 2011-2021 in the United States, we investigate post-merger performance by applying artificial neural networks and cross-validating the results. The focus of this study is placed on three issues: 1) whether mergers have an impact on the financial performance of the business both in the short and the long run, 2) evaluating the effectiveness of ANNs in describing and finding trends in financial data, 3) discovering the distinguishers of more successful mergers from less successful ones in the sample with regard to their market value.

Firstly, using paired two-sample t-tests, high-tech mergers are found to have a statistically significant impact on the liquidity and solvency conditions of the merged firms, with liquidity decreasing and debt leverage increasing each year for five years following the merger completion. While there is a significant short-term decrease in profitability, the market value represented as capitalization per share increases considerably during the analyzed period.

Secondly, artificial neural networks are developed to analyze the data and classify the mergers into more successful and less successful ones with regard to their market valuation. Trained ANNs predict with high accuracy which mergers will be more successful in terms of market capitalization and average stock returns,

analyzing the directions and magnitude of the connections between input and output variables. Contrary to the coefficients of regression models, ANN weights have intra-variable variation, which provides additional support for the implementation of ANNs to analyze events such as mergers with potential nonlinear impact. The results show that important determinants of the long-term market value growth of the newly merged businesses in the sample are the cross-industry or same-industry nature of the merger, profitability of both parties of the deal, the bidder's size and its capitalization before the merger occurs.

Thirdly, the study provides noteworthy contributions. It shows that ANNs can serve as a highly efficient model for analyzing financial data, including merger performance, due to their flexibility and lack of prior assumptions about the data. A successfully trained ANN on representative datasets can be used for ex-ante forecasts of potential value implications of certain business decisions by inputting new information, for instance about a bidder who previously did not engage in mergers and is only planning to do so with potential targets. Information of this nature can be valuable for academics, as well as managers and consultants, allowing them to make informed strategic and investment decisions. Consequently, future research might consider expanding the scope by adding more transactions classified by industry or other attributes, or training and cross-validating different models on separately classified datasets to capture, for example in textual form, other key trends and events that might take place. In practice, ANNs trained on representative datasets should be regularly retrained on new inputs to adjust for the latest market conditions.

Finally, there are several limitations implying that the results of this study should not be generalized. Even though neural networks are popular and could be highly accurate, financial data can be highly sensitive to market shocks and seasonality. Hence, the modelled results and parameter estimates should be interpreted with caution as they provide only possible explanations of the trends observed in the data, especially when the sample is not representative, and the model used is of a black-box nature. To improve the generalization abilities of ANNs, regularization techniques based on loss function or noise introduction can be considered. Additionally, it should be noted that the number of synaptic weights can become rather large, therefore making their interpretation lack utility. Model-agnostic Explainable Artificial Intelligence (XAI) methods can be used to enhance the interpretability of the black box decision-making process. The complexity of such events as mergers, limitations of the research method, and issues raised in the reviewed literature call for further analysis. Given the high levels of observed M&A activity, which is a data-generative process for these studies, it is vital to use this opportunity to create knowledge.

REFERENCES

Agrawal A., and Jaffe J. F. (2003) Do Takeover Targets Underperform? Evidence from Operating and Stock Returns. The Journal of Financial and Quantitative Analysis, 38(4), 721-746. https://doi.org/10.2307/4126741

Agrawal A., Jaffe J. F., Mandelker G. N. (1992) The Post-Merger Performance of Acquiring Firms: A Re-examination of an Anomaly. Journal of Finance, 47, 1605-1621. https://doi.org/10.1111/j.1540-6261.1992.tb04674.x

Ahern K. P., Harford J. (2014) The Importance of Industry Links in Merger Waves. The Journal of Finance, 69(2), 527-576. https://doi.org/10.1111/jofi.12122

Anderson C., Mandelker G. (1993) Long Run Return Anomalies and the Book-to-Market Effect: Evidence on Mergers and IPOs. Working Paper, University of Pittsburgh.

Andrade G., Mitchell M., Stafford E. (2001) New Evidence and Perspectives on Mergers. Journal of Economic Perspectives, 15(2), 103-120. https://doi.org/10.1257/jep.15.2.103

André P., Kooli M., L'Her J. (2004) The Long-Run Performance of Mergers and Acquisitions: Evidence from the Canadian Stock Market. Financial Management, 33(4), 27-43. https://www.jstor.org/stable/3666327

Asquith P. (1983) Merger Bids, Uncertainty and Stockholder Returns. Journal of Financial Economics, 11, 51-83. https://doi.org/10.1016/0304-405X(83)90005-3

Baldi P., Brunak S., Chauvin Y., Andersen C. A. F., Nielsen H. (2000) Assessing the Accuracy of Prediction Algorithms for Classification: An Overview. Bioinformatics Review, 16(5), 412-424. https://doi.org/10.1093/bioinformatics/16.5.412

Bi W., Zhang Q. (2021) Forecasting Mergers and Acquisitions Failure Based on Partial Sigmoid Neural Network and Feature Selection. PLoS ONE, 16(11), 1-18. https://doi.org/10.1371/journal.pone.0259575

Bouteska A., Hajek P., Fisher B., Abedin M. Z. (2023) Nonlinearity in Forecasting Energy Commodity Prices: Evidence from a Focused Time-Delayed Neural Network. Research in International Business and Finance, 64, 101863. https://doi.org/10.1016/j.ribaf.2022.101863

Bradley M. Jarrell G. A. (1988) Comment. J. C. Coffee, L. Lowenstein, S. Rose-Ackerman, (Eds.), Knights, Raiders, and Targets: The Impact of the Hostile Takeover. Oxford University Press, New York, 252-259. ISBN: 0-19-504405-3

Chicco D., Jurman G. (2020) The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. BMC Genomics, 21(6), 1-13. https://doi.org/10.1186/s12864-019-6413-7

Cumming D., Jindal V., Kumar S., Pandey N. (2023) Mergers and Acquisitions Research in Finance and Accounting: Past, Present and Future. European Financial Management, 29(5), 1-41. https://doi.org/10.1111/eufm.12417

Department of Math of the University of Utah (n.d.). Using the Receiver Operating Characteristic (ROC) Curve to Analyze a Classification Model. Retrieved from http://www.math.utah.edu/~gamez/files/ROCCurves.pdf

Firth M. (1980) Takeovers, Shareholder Returns, and the Theory of the Firm. Quarterly Journal of Economics, 94, 315-347. https://doi.org/10.2307/1884539

Franks J., Harris R., Titman S. (1991) The Postmerger Share-Price Performance of Acquiring Firms. Journal of Financial Economics, 29, 81-96. https://doi.org/10.1016/0304-405X(91)90014-B

Hackbarth D., Morellec E. (2008) Stock Returns in Mergers and Acquisitions. The Journal of Finance, 63(3), 1213-1252. https://doi.org/10.1111/j.1540-6261.2008.01356.x

Healy P. M., Palepu K. G., Ruback R. S. (1992) Does Corporate Performance Improve after Mergers? Journal of Financial Economics, 31, 135-175. https://doi.org/10.1016/0304-405X(92)90002-F

IBM (2012) IBM SPSS Neural Networks – New Tools for Building Predictive Models. Retrieved from https://www.ibm.com/downloads/cas/N7LLA2LB

Jensen-Vinstrup M., Rigamonti D., Wulff J. (2018) European Cross-Border Acquisitions: Long-Run Stock Returns and Firm Characteristics. Journal of Multinational Financial Management, 47, 31-45. https://doi.org/10.1016/j.mulfin.2018.09.003

Kaplan S. N. (2000) Introduction. Steven Kaplan, (Ed.), Mergers and Productivity. University of Chicago Press, Chicago, 1-8. ISBN: 0-226-42431-6

King D. R., Dalton D. R., Daily C. M., Covin J. G. (2004) Meta-analyses of Post-acquisition Performance: Indications of Unidentified Moderators. Strategic Management Journal, 25(2), 187-200. https://doi.org/10.1002/smj.371

Kohers N., Kohers T. (2001) Takeovers of Technology Firms: Expectations vs. Reality. Financial Management, 30(3), 35-54. https://doi.org/10.2307/3666375

Langetieg T. C. (1978) An Application of a Three-Factor Performance Index to Measure Stockholder Gains from Merger. Journal of Financial Economics, 6(4), 365-383. https://doi.org/10.1016/0304-405X(78)90010-7

Le H. H., Viviani J. (2018) Predicting Bank Failure: An Improvement by Implementing a Machine-Learning Approach to Classical Financial Ratios. Research in International Business and Finance, 44, 16-25. https://doi.org/10.1016/j.ribaf.2017.07.104

Lee K., Joo S., Baik H., Han S., In J. (2020) Unbalanced Data, Type II Error, and Nonlinearity in Predicting M&A failure. Journal of Business Research, 109, 271-287. https://doi.org/10.1016/j.jbusres.2019.11.083

Léger P., Quach L. (2009) Post-Merger Performance in the Software Industry: The Impact of Characteristics of the Software Product Portfolio. Technovation, 29(10), 704-713. https://doi.org/10.1016/j.technovation.2009.05.016

Loderer C., Martin K. (1992) Postacquisition Performance of Acquiring Firms. Financial Management, 21(3), 69-79. https://doi.org/10.2307/3666020

Loughran T., Vijh A. M. (1997) Do Long-Term Shareholders Benefit from Corporate Acquisitions. Journal of Finance, 52, 1765-1790. https://doi.org/10.1111/j.1540-6261.1997.tb02741.x

Malmendier U., Moretti E., Peters F. S. (2018) Winning by Losing: Evidence on the Long-run Effects of Mergers. The Review of Financial Studies, 31(8), 3212-3264. https://doi.org/10.1093/rfs/hhy009

Mitchell M. L. Mulherin J. H. (1996) The Impact of Industry Shocks on Takeover and Restructuring Activity. Journal of Financial Economics, 41(2), 193-229. https://doi.org/10.1016/0304-405X(95)00860-H

Møller M. F. (1993) A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. Neural Networks, 6(4), 525-533. https://doi.org/10.1016/S0893-6080(05)80056-5

Pianosi F., Wagener T. (2015) A Simple and Efficient Method for Global Sensitivity Analysis Based on Cumulative Distribution Functions. Environmental Modelling & Software, 67, 1-11. https://doi.org/10.1016/j.envsoft.2015.01.004

Powers D. M. W. (2011) Evaluation: from Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. The Journal of Machine Learning Technologies, 2. https://doi.org/10.48550/arXiv.2010.16061

Rau P. R., Vermaelen T. (1998) Glamour, Value and the Post-Acquisition Performance of Acquiring Firms. Journal of Financial Economics, 49, 223-253. https://doi.org/10.1016/S0304-405X(98)00023-3

Ravenscraft D. J., Scherer F. M. (1987) Mergers, Sell-offs, and Economic Efficiency. Brookings Institution, Washington, D.C. ISBN: 0-8157-7348-X

Shao B., Asatani K., Sakata I. (2018) Categorization of Mergers and Acquisitions in Japan Using Corporate Databases: A Fundamental Research for Prediction. In 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 1523–1527. https://doi.org/10.1109/IEEM.2018.8607408

Shao B., Asatani K., Sasaki H., Sakata I. (2021) Categorization of Mergers and Acquisitions using Transaction Network Features. Research in International Business and Finance, 57, 101421. https://doi.org/10.1016/j.ribaf.2021.101421

Teräsvirta T., Van Dijk D., Medeiros M. C. (2005) Linear Models, Smooth Transition Autoregressions, and Neural Networks for Forecasting Macroeconomic Time Series: a Re-Examination. International Journal of Forecasting, 21(4), 755–774. https://doi.org/10.1016/j.ijforecast.2005.04.010

Town R. J. (1992) Merger Waves and the Structure of Merger and Acquisition Time-Series. Journal of Applied Econometrics, 7(1), 83-100. https://doi.org/10.1002/jae.3950070507.

Yu L., Wang S. Lai K. K. (2007) Foreign-Exchange-Rate Forecasting with Artificial Neural Networks (107). Springer Science & Business Media. ISBN: 978-0-387-71720-3.

Zhu M., Meng Z. (2021) Fuzzy Comprehensive Evaluation Model of M&A Synergy Based on Transfer Learning Graph Neural Network. Complexity, 2021, 1-12. https://doi.org/10.1155/2021/6516722

QUANTITATIVE METHODS IN ECONOMICS
Volume XXV, No. 2, 2024, pp. 86 – 97

Received: 20.03.2024
Accepted: 02.05.2024

ANALYSIS OF REGIONAL DIFFERENTIATION OF POLAND'S AGRICULTURE IN THE YEARS 2016-2022

Agata Binderman-Dubik https://orcid.org/0000-0001-5238-3101
Department of Electrical Engineering, Mathematics and Natural Sciences
University of Gavle, Sweden, signature_3714744577
e-mail: agata.binderman.dubik@hig.se

Abstract: The work is a continuation of a series of works by A. Binderman-Dubik, published in the years 2004-2013, concerning the regional differentiation of Polish agriculture, in which she used two patterns (models): negative and positive. These studies showed the superiority of the two-pattern methods used over the single-pattern methods. Here, methods based on three patterns (negative, positive, indirect) are used. The author uses the method considered by the authors: Binderman Z., Borkowski B., Szczesny W. [2020].

Keywords: agricultural differentiation, synthetic measures, utility function, classification

JEL classification: C38, O18, Q1, R1

INTRODUCTION

Research by many authors has shown that the indicators of the development of Polish agriculture have clearly increased after Polish's accession to the European Union. Nevertheless, despite such a significant increase, the level of agricultural differentiation the provinces is not decreasing, and what is more, it shows an upward tendency. This was shown, among others, by the author's work [Binderman A. 2004, 2007, 2012], which examined the regional differentiation of agriculture in the periods 1989-1998, 1998-2005, and 1998-2010, respectively. The obtained results showed that in relation to the adopted characteristics, regional differentiation is clearly increasing. This paper shows that also in the years 2016-2022 the regional differentiation of agriculture increased. The pandemic and the war in Ukraine have not disrupted this trend. Compared to the economically developed countries of Western Europe [Fogelfors, H. (ed.) 2009; Rabinowicz E. 2020], Poland is a country with a significant agricultural production potential. In addition, the diversity of natural and economic and organizational conditions means that the degree of

https://doi.org/10.22630/MIBE.2024.25.2.8

utilization of the potential of agriculture is regionally differentiation [Harasim 2006, 2009; Krasowicz, Kukuła 2006; Nermend, Miłaszewicz 2016; Poczta, Bartkowiak 2012; Zegar 2003]. In order to analyze complex phenomena, such as the level of development or the potential of agriculture, and to assess voivodships in this respect, it is necessary to consider many factors. The use of the potential of agriculture in the regions is a derivative of the impact of various groups of conditions, both favorable and restrictive.

The analyses show that Polish voivodships have significant resources of basic production factors and relatively favorable natural conditions. One of the main reasons for the low utilization of the potential of agriculture in Poland is the insufficient development of the agri-food industry. A significant part of the agricultural commodity production of the voivodships are raw materials for processing, and not processed products, characterized by a higher share of the so-called added value. The National Agricultural Censuses 2010, 2020 (NAC 2010, NAC 2020) showed that in the period of years 2010-2020 the importance of farms focused on market production increased. In the total number of farms, the number of the largest and smallest units increased, with an increase in the average area of the farm.

METHODS OF CONSTRUCTION FOR INIDICATORS OF DIFFERENTIATION

Ordering composed phenomena characterized in a summary way synthetic (aggregate) variables are used. The substitution of a sequence of many explanatory features by a synthetic variable gives a certain assessment of the phenomenon under study. Pattern methods assume the existence of a hypothetical model object. These methods use appropriately selected diagnostic variables characterizing the studied phenomenon and differ from each other as to the method of normalization of variables and the form of aggregate functions [Cieślak 2023; Hellwig 1968; Kukuła K. 2000; Malina 2004; Młodak 2006; Kisielińska 2021; Nowak 1990, Zeliaś 2000]. In this paper, both two patterns and three patterns will be used at the same time.

Let $X=\Re^n$, $\Re=(-\infty,\infty)$, $n\in \mathrm{N}$, denotes an *n*-dimensional vector space. Consider the problem of ordering $m\in \mathrm{N}$ objects $\mathbf{Q}_1$, $\mathbf{Q}_2$,...,$\mathbf{Q}_m$ by $n\in \mathrm{N}$ variables (features) meant to describe each of them. Without losing the generality of the considerations, let us assume that all features may be considered as stimulant. The symbol $\mathbf{x}_i=\left(x_{i1},x_{i2},...,x_{in}\right)\in \mathrm{X}$, for $i=1,2,...,m$, will denote the vector of values of variables describing the i-th object $\mathbf{Q}_i$. Assume that $\mathcal{W}^*:=\{\mathbf{x}_1, \mathbf{x}_2,...,\mathbf{x}_m\}$ denotes the set of vectors describing the objects $\mathbf{Q}_1$, $\mathbf{Q}_2$,...,$\mathbf{Q}_m$. We say that $\mathbf{x}_i>\mathbf{x}_j$, ($\mathbf{x}_i\geq\mathbf{x}_j$) ($i,j=1,...,m$) if $x_{ik}>x_{jk}\,(x_{ik}\geq x_{jk})$ for $k=1,2,...,n$. On the other hand, $\mathbf{Q}_0$, $\mathbf{Q}_{m+1}$ and $\mathbf{Q}_{sr}$ will denote objects described by vectors with coordinates:

$$x_{0,k} = \min_{1\le i\le m} x_{ik},\ \ x_{m+1,k} = \max_{1\le i\le m} x_{ik},\ \ x_{sr,k} = \frac{1}{m}\sum_{i=1}^{m} x_{ik}\ ;\ k = 1, 2, ..., n.$$

It is obvious that the objects: $\mathbf{Q}_0$ – described by the vector $\mathbf{x}_0$ and $\mathbf{Q}_{m+1}$ – described by the vector $\mathbf{x}_{m+1}$ are respectively not worse or better than the remaining objects $\mathbf{Q}_1$, $\mathbf{Q}_2$,...,$\mathbf{Q}_m$. The components of the vector $\mathbf{x}_{sr}$ are the averages of the components of the vectors under consideration, respectively. Directly from the definition, holds the inequality $\mathbf{x}_0 \le \mathbf{x}_{sr} \le \mathbf{x}_{m+1}$. The three objects $\mathbf{Q}_0$, $\mathbf{Q}_{sr}$ and $\mathbf{Q}_{m+1}$ (perhaps fictitious) can be treated as patterns (extra models) added to the initial, input objects $\mathbf{Q}_1$, $\mathbf{Q}_2$,...,$\mathbf{Q}_m$. Let $d^*(\mathbf{x},\mathbf{y})$ be the Euclid distance between the vectors $\mathbf{x},\mathbf{y} \in \Re^n_+$ and $d^*(\mathbf{x}_0,\mathbf{x}_{m+1})\neq 0$. The most well-known synthetic indicators in the literature built on the basis of patterns intended for ordering objects are following measures:

$$\mu_1(\mathbf{x}) = d^*(\mathbf{x}_0,\mathbf{x}) / d^*(\mathbf{x}_0,\mathbf{x}_{m+1}) \tag{1}$$

$$\mu_2(\mathbf{x}) = 1 - d^*(\mathbf{x}_{m+1},\mathbf{x}) / d^*(\mathbf{x}_0,\mathbf{x}_{m+1}) \tag{2}$$

$$\mu_3(\mathbf{x}) = \frac{\mu_1(\mathbf{x}) + \mu_2(\mathbf{x})}{2} = \frac{1}{2} + \frac{d^*(\mathbf{x}_0,\mathbf{x}) - d^*(\mathbf{x}_{m+1},\mathbf{x})}{2d^*(\mathbf{x}_0,\mathbf{x}_{m+1})} \tag{3}$$

$$\mu_4(\mathbf{x}) = \frac{\mu_1(\mathbf{x})}{1+\mu_1(\mathbf{x}) - \mu_2(\mathbf{x})} = \frac{d^*(\mathbf{x}_0,\mathbf{x})}{d^*(\mathbf{x}_0,\mathbf{x}) + d^*(\mathbf{x}_{m+1},\mathbf{x})}, \tag{4}$$

for $\mathbf{x}\in[\mathbf{x}_0,\mathbf{x}_{m+1}]$.

As you can see, the μ_1 and μ_2 measures use a single model, while the μ_3 and μ_4 measures use two model, expressed as elementary functions of the μ_1 and μ_2 meters. In the paper [Hellwig 1968] a measure based on only single model (best model) is given. The theory and applications of the μ_3 measure are given in the series of works by A. Binderman **[**Binderman A. 2006, 2007, 2011**]** and in the work **[**Binderman Z. 2010**].** The μ_4 measure is related to the TOPSIS method) **[**Hwang, Yoon 1981**]**. In the next part of our discussions, let us normalize the distance $d^*(\mathbf{x},\mathbf{y})$ of the vectors $\mathbf{x},\mathbf{y} \in \Re^n_+$, relative to the assumed model vectors $\mathbf{x}_o$, $\mathbf{x}_{m+1}$, using the formula:

$$d(\mathbf{x},\mathbf{y}) := \frac{d^*(\mathbf{x},\mathbf{y})}{d^*(\mathbf{x}_0,\mathbf{x}_{m+1})}.$$

Then $d(\mathbf{x}_0,\mathbf{x}_{m+1}) = 1$ a formulas (1)-(4) take for $\mathbf{x}\in[\mathbf{x}_0,\mathbf{x}_{m+1}]$ the form:

$$\mu_1(\mathbf{x}) = d(\mathbf{x}_0, \mathbf{x}) \qquad (1a)$$

$$\mu_2(\mathbf{x}) = 1 - d(\mathbf{x}_{m+1}, \mathbf{x}) \qquad (2a)$$

$$\mu_3(\mathbf{x}) = \left[1 + d(\mathbf{x}_0, \mathbf{x}) - d(\mathbf{x}_{m+1}, \mathbf{x})\right]/2 \qquad (3a)$$

$$\mu_4(\mathbf{x}) = d(\mathbf{x}_0, \mathbf{x}) / \left(d(\mathbf{x}_0, \mathbf{x}) + d(\mathbf{x}_{m+1}, \mathbf{x})\right). \qquad (4a)$$

It is easy to see that the considered measures, defined by formulas (1a) - (4a) are standardized in relation to the accepted patterns, i.e. :

$\mu_i(\mathbf{x}_0) = 0$, $\mu_i(\mathbf{x}_{m+1})=1$ and $0 \leq \mu_i(\mathbf{x}_{sr}) \leq 1$ for $i=1,2,3,4$,

where m is the number of objects considered. These measures are standardized utility functions [Panek 2000; Binderman Z. Borkowski, Szczesny 2020]. Note that if a given utility function u induces preference relations in the set of m+3 objects

$$\mathcal{W} := \mathcal{W}^* \cup \{ \mathbf{x}_0, \mathbf{x}_{m+1}, \mathbf{x}_{sr}\} = \{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m, \mathbf{x}_{m+1}, \mathbf{x}_{sr} \},$$

then a composite function $g(u(\mathbf{x}))$, where $g: \Re \to \Re$ is an increasing function, as well as a utility function, generating the same preference relation in the set of objects $\mathcal{W}$ as the function $u(\mathbf{x})$.

Using the above property, it is expedient to normalize the utility function by choosing such a function g that its value for the worst object $\mathbf{x}_0$ is equal to 0, and the value for the best object $\mathbf{x}_{m+1}$ is equal to 1, i.e. that:

1. $g(u(\mathbf{x}_0)) = 0$; 2. $g(u(\mathbf{x}_{m+1})) = 1$.

Using the well-known forms of the $u(\mathbf{x})$ function in our work (realizing that there are infinitely many such functions), we give an example of such a function of one variable $g(u)$, which would satisfy one more condition:

3. $g(u(\mathbf{x}_{sr}))=1/2$.

The simplest function that satisfies these above three conditions is the linear picewise function [Binderman Z., Borkowski, Szczesny 2020]

$$g(u)=\begin{cases} \frac{1}{2\alpha}u & dla\ \ 0 < u < \alpha \\ \frac{(u-1)}{2(1-\alpha)} + 1 & dla\ \ \alpha \leq u \leq 1 \end{cases} \qquad (5),$$

where $\alpha := u(\mathbf{x}_{sr})$.

EMPIRICAL RESEARCH

A set of ten features was selected to describe the regional differentiation of agriculture after the analysis. All the selected variables were stimulant, which means that higher values of these features informed about a higher level of development of the studied phenomenon. The selected variables determine the overall level of agriculture in given years in Poland, at the same time making it possible to show the differences that occur between voivodships. Below is the final list of the ten selected diagnostic variables.

X_1 Share of agricultural area as % of total area.
X_2 Gross domestic product in zł., per 1 inhabitant.
X_3 Sugar beet yield in tons per 1 hectare.
X_4 Stocking density of cattle per 100 hectares of agricultural land.
X_5 Purchase of potatoes in kilograms per 1 hectare of agricultural land.
X_6 Yields of oilseeds in dt per 1 hectare of cultivated area.
X_7 Purchase of fruit from trees in kg per 1 hectare of cultivated area.
X_8 Total purchase value of agricultural products in zł per 1 ha of agricultural land.
X_9 Global crop production per 1 ha of agricultural land in zł, according to the new definition.
X_{10} Capital expenditures in agriculture, forestry and hunting, in zł per 1 ha of agricultural land, current prices.

The data considered in the paper do not include the results of the National Agricultural Census (NAC, NSR - polish), conducted in 2020 by the Central Statistical Office (CSO, GUS -polish). The results of the work are used only by the data of the CSO. It is worth mentioning here that the comparison of the results of NAC 2010 and NAC 2020 shows that the number of farms decreased significantly, while the average total area and agricultural area increased at the same time [CSO NAC 2020]. On farms, the number of livestock per 100 hectares of agricultural land increased in the number of cattle with a marked decrease in the number of pigs. The area of orchards decreased significantly (by approx. 14%).

The data analyzed in the paper can be presented by means of a three-dimensional matrix (*voivodship* $\times$ *value of feature* $\times$ *year*), or by means of a matrix in which each row represents an object represented by the features of a given voivodship in a given year. In the work, the latter method was chosen. Proceeding in this way, $m=16\times7=132$ objects $\mathbf{Q}_1,\mathbf{Q}_2,...,\mathbf{Q}_{132}$ were obtained, each of which was described by $n=10$ features $X_1,X_2,...,X_{10}$. The values of the adopted diagnostic variables for 16 voivodships and for 7 years (2016-2022) formed a matrix $\mathbf{X}$, the matrix with dimensions of 10×132. On the basis of the values assumed by the diagnostic variables for 16 voivodships in the individual 7 years of the studied period, three fixed (static), hypothetical voivodships were created: "minimal" ("the worst" in relation to each voivodship) $\mathbf{Q}_0$, "maximum" ("the best" in relation to each voivodship) Q_{133} and Q_{134} "average", Objects $\mathbf{Q}_0$, $\mathbf{Q}_{133}$ described appropriately by the least, most favorable set of feature values. The "mean" object Q_{134} is described by the average values of the features under consideration between 2016 and 2022. The hypothetical provinces $\mathbf{Q}_0$, $\mathbf{Q}_{133}$, $\mathbf{Q}_{134}$ in this paper will be represented by vectors $\mathbf{x}_0$, $\mathbf{x}_{133}$ and $\mathbf{x}_{134}$ with 10 components each, respectively. These vectors are the benchmark models for the entire period 1998-2010, they determine the cube $[\mathbf{x}_0,\mathbf{x}_{133}]$ in n-dimensional Euclid space $\Re^n_+$, which means that for every $i\in[1,132]$: $\mathbf{x}_i\in[\mathbf{x}_0,\mathbf{x}_{133}]$. In this way, a data matrix was obtained for further analysis, $\mathbf{X}=\left[x_{ij}\right]_{135\times10}$ with 135 rows and 10 columns. Since the selected diagnostic variables had different titers and different orders of magnitude, these variables were normalized. In order to reduce the variables to comparability, zero unitarization was selected and applied from among many types of norming. The choice of the method of normalizing variables was a consequence of the results obtained by the author (Binderman A 2006, 2010). Normalized values for individual variables with m=135, n=10 were calculated according to the formulas [Kukuła K. 2000]:

$$z_{ij}=\left(x_{ij}-x_{0j}\right)/\left(x_{m+1j}-x_{0j}\right) \quad \text{for } 0\le i\le 134,\ 1\le j\le 10 \tag{6}$$

The features transformed in this way, by eliminating the units of measurement, became mutually comparable. The z_{ij} variables transformed by the zero unitarization method (MUZ) take values in the closed interval [0,1]. The transformations made can be symbolically written: $\mathbf{Z}=[z_{ij}]_{135\times10}=\varphi(\mathbf{X})$, where $\mathbf{X}$ is the observation matrix,. After the transformation of the variables, the static pattern vectors are as follows:

$$\mathbf{z}_0=\mathbf{0}:=[0,0,...,0],\ \mathbf{z}_{133}=\mathbf{1}:=[1,1,...,1],$$

i.e. in the created **Z** matrix**,** the first row consists of only zeros, while the penultimate row consists of only ones.

RESEARCH RESULTS

Let $\mathbf{w}_1,\mathbf{w}_2,...,\mathbf{w}_{16}\in\Re^7_+$ $\left(\mathbf{w}_{1s},\mathbf{w}_{2s},...,\mathbf{w}_{16s}\in\Re^7_+\right)$ denote vectors, assigned alphabetically to voivodeships ($\mathbf{w}_1$ – Dolnośląskie,...,$\mathbf{w}_{16}$ – Zachodniopomorskie), whose components are the values of the m(**z**) (m_s(**z**)), measure of development in the individual, seven years of the period 1998-2010, where m(**z**) (m_s(**z**)) is determined by the formula (3a) ((5)). Table 1 shows the results of the research, in which the rows are the coordinates of the vectors $\mathbf{w}_1$, $\mathbf{w}_2$,...,$\mathbf{w}_{16}$ ($\mathbf{w}_{1s}$, $\mathbf{w}_{2s}$,...,$\mathbf{w}_{16s}$).

Table 1. Values of the measures m(**z**), m_S(**z**) for the voivodships

Voivodships	Measure of development	2016	2017	2018	2019	2020	2021	2022
Dolnośląskie	w_1	0.35	0.38	0.35	0.33	0.38	0.41	0.40
	w_{1s}	0.44	0.48	0.45	0.41	0.48	0.51	0.51
Kujawsko-pomorskie	w_2	0.39	0.44	0.38	0.39	0.46	0.47	0.50
	w_{2s}	0.50	0.54	0.48	0.50	0.55	0.56	0.59
Lubelskie	w_3	0.34	0.36	0.37	0.35	0.41	0.41	0.46
	w_{3s}	0.43	0.46	0.47	0.44	0.51	0.51	0.55
Lubuskie	w_4	0.24	0.29	0.17	0.18	0.20	0.24	0.27
	w_{4s}	0.30	0.36	0.21	0.22	0.26	0.31	0.35
Łódzkie	w_5	0.40	0.45	0.42	0.43	0.50	0.50	0.54
	w_{5s}	0.51	0.54	0.52	0.53	0.59	0.58	0.62
Małopolskie	w_6	0.35	0.36	0.35	0.35	0.37	0.39	0.39
	w_{6s}	0.45	0.46	0.45	0.44	0.46	0.49	0.49
Mazowieckie	w_7	0.50	0.52	0.57	0.57	0.63	0.64	0.71
	w_{7s}	0.59	0.61	0.65	0.65	0.70	0.70	0.76
Opolskie	w_8	0.37	0.38	0.39	0.39	0.40	0.48	0.47
	w_{8s}	0.47	0.48	0.50	0.50	0.50	0.57	0.56
Podkarpackie	w_9	0.25	0.23	0.26	0.23	0.26	0.28	0.27
	w_{9s}	0.31	0.29	0.33	0.29	0.33	0.36	0.35
Podlaskie	w_{10}	0.43	0.46	0.44	0.43	0.49	0.55	0.54
	w_{10s}	0.53	0.56	0.54	0.53	0.58	0.63	0.62
Pomorskie	w_{11}	0.36	0.40	0.40	0.41	0.42	0.46	0.47
	w_{11s}	0.46	0.51	0.51	0.52	0.53	0.55	0.56
Śląskie	w_{12}	0.35	0.40	0.33	0.40	0.37	0.43	0.43
	w_{12s}	0.44	0.51	0.42	0.51	0.47	0.53	0.53
Świętokrzyskie	w_{13}	0.33	0.33	0.33	0.32	0.36	0.36	0.39
	w_{13s}	0.42	0.42	0.42	0.41	0.46	0.46	0.50
Warmińsko-mazurskie	w_{14}	0.28	0.35	0.32	0.35	0.37	0.38	0.40
	w_{14s}	0.36	0.44	0.41	0.45	0.48	0.48	0.50
Wielkopolskie	w_{15}	0.61	0.60	0.58	0.59	0.65	0.70	0.75
	w_{15s}	0,68	0.67	0.65	0.66	0.71	0.75	0.79
Zachodnio-pomorskie	w_{16}	0.28	0.25	0.20	0.26	0.31	0.32	0.33
	w_{16s}	0.36	0.32	0.25	0.33	0.40	0.40	0.42
gap	w	0.37	0.37	0.41	0.42	0.44	0.46	0.48
	w_s	0.37	0,38	0.44	0.44	0.45	0.44	0.45
average	w	0.36	0.39	0.37	0.37	0.41	0.44	0.46
	w_s	0.45	0.48	0.45	0.46	0.50	0.53	0.54

Source: Author's own calculations

In Table 1, for example, the components of the vector $\mathbf{w}_1$ are the values of the $m(z)=\mu_{3}(\mathbf{z})$ measure, calculated according to the formula (3a)) for the Dolnośląskie Voivodeship in each year of the period 2016-2022 under consideration. On the other hand, the components of the vector $\mathbf{w}_{1s}$ form the values of the measure $m_S(\mathbf{z})= =g(m(\mathbf{z}))$, calculated according to the formula (5).It should be noted that the values of the range (max-min) given in Table 1 show the growing diversity of voivodships in the years 2016-2022. Differentiation increased by nearly 30% (22%) in this period when determining development coefficients according to the formula (3a) ((5)). The designated measures made it possible to make a ranking of the objects under consideration. As it is easy to see directly from the definition of the functions $m=\mu_3$ and g(m), defined by formulas (3a), (5), it follows accordingly that the function g does not change the relation of preferences in the set of vectors $\boldsymbol{\mathcal{W}}$, induced by the utility function m.

The tables below present the classification of voivodships, according to the obtained measures of the level of development in the years 2016-2022.

Table 2. Classification of voivodships in 2016-2022

Voivodships	2016	2017	2018	2019	2020	2021	2022	2016-2021
Dolnośląskie	9	9	10	12	9	10	10	10
Kujawsko-pomorskie	5	5	7	7	5	6	5	5
Lubelskie	11	11	8	11	7	9	8	9
Lubuskie	16	14	16	16	16	16	16	16
Łódzkie	4	4	4	3	3	4	3	4
Małopolskie	8	10	9	10	12	11	13	11
Mazowieckie	2	2	2	2	2	2	2	2
Opolskie	6	8	6	8	8	5	7	7
Podkarpackie	15	16	14	15	15	15	15	15
Podlaskie	3	3	3	4	4	3	4	3
Pomorskie	7	7	5	5	6	7	6	6
Śląskie	10	6	12	6	11	8	9	8
Świętokrzyskie	12	13	11	13	13	13	12	13
Warmińsko-mazurskie	13	12	13	9	10	12	11	12
Wielkopolskie	1	1	1	1	1	1	1	1
Zachodnio-pomorskie	14	15	15	14	14	14	14	14

Source: Author's own calculations

Analyzing the arrangement of voivodships according to the level of utility, it can be seen that in the analyzed period the Wielkopolskie and Mazowieckie voivodships were at the top of the ranking list. The final places were taken by the Zachodniopomorskie Voivodeship, the Podkarpackie Voivodeship, and the Lubuskie Voivodeship in the last place.

Using the results of calculations presented in Table 1, voivodships were divided into 4 groups, characterized by a similar level of agricultural development. The division of voivodships into classes for the entire period under study was based on the values of the average measure of development. For this purpose, distributive interval series were used, in which the spans of class intervals were equal to approximately one-quarter of the range for the entire period.

To group the voivodships into four groups I, II, III and IV, the upper class boundaries given in Table 3 below were used, in relation to the method of calculating the development rate and years.(the calculations obtained by the function $m=\mu_3$ –

the first four rows, the calculations obtained by the function g(m) – last four rows), the function m, g(m) are defined by formulas (3a), (5)

The results of the grouping of voivodships for the analyzed period are presented in Table 4. Analyzing the regional diversity of agriculture in Poland, four typological groups of voivodships were distinguished.

Table 3. Upper Class Boundaries , according to the method of calculating the synthetic measures in the years 2016-2022

class	2016	2017	2018	2019	2020	2021	2022
IV	0.33	0.32	0.27	0.28	0.31	0.36	0.39
III	0.42	0.42	0.37	0.38	0.42	0.47	0.51
II	0.52	0.51	0.48	0.49	0.53	0.59	0.63
I	0.61	0.60	0.58	0.59	0.65	0.70	0.75
IV	0.396	0.388	0.323	0.333	0.369	0.422	0.457
III	0.489	0.483	0.433	0.443	0.482	0.532	0.568
II	0.583	0.578	0.542	0.554	0.595	0.643	0.680
I	0.676	0.673	0.652	0.664	0.708	0.754	0.792

Source: Author's own calculations

Table 4. Division of voivodships into groups according to the method of calculating the synthetic measures in the years 2016-2022

I	Wielkopolskie, Mazowieckie
II	Łódzkie, Podlaskie, Pomorskie
III	Kujawsko-pomorskie, Dolnośląskie, Małopolskie , Śląskie, Świętokrzyskie, Warmińsko-mazurskie, Lubelskie, Opolskie
IV	Lubuskie, Podkarpackie, Zachodnio-pomorskie
I	Wielkopolskie, Mazowieckie
II	Łódzkie, Podlaskie, Kujawsko-pomorskie, Dolnośląskie, Opolskie
III	Lubelskie, Małopolskie, Śląskie, Świętokrzyskie, Warmińsko-mazurskie, Pomorskie
IV	Lubuskie, Podkarpackie, Zachodnio-pomorskie

Source: Author's own calculations

A synthetic summary of results indicates that Poland is a country clearly differentiated in terms of the level of agricultural development. The change in the method of calculating voivodeship development indices used in the paper, without changing the order, changes their grouping. Regardless of the method of calculating the measure of development, the highest rated group I includes the Wielkopolskie and Mazowieckie voivodships, the second group is Łódzkie and Podlaskie voivodships, the third group is Śląskie, Świętokrzyskie and Warmińsko-Mazurskie groups, and the fourth group is Lubuskie, Podkarpackie and Zachodniopomorskie. The fourth group includes voivodships that are characterized by the lowest level of agricultural development in Poland, according to the adopted characteristics and research methodology. In almost every year of the period under consideration, the lowest level of agricultural development was shown by the Lubuskie voivodeship. Comparing the obtained results with the results of the work [Binderman A. 2007, 2012] it is possible to observe a clear progress and advancement of the following voivodeships: Mazowieckie, Łódzkie, Podlaskie, a slight increase in the following voivodeships: Podkarpackie, Lubuskie, Małopolskie and Zachodnio-Pomorskie.

CONCLUSION

The results of the conducted research showed that the indicators of development of Polish agriculture clearly increased in the studied period 2016 – 2022. Despite this significant increase, the level of diversification of agricultural development in Polish voivodships is not decreasing, and what is more, it shows an upward tendency. In order to confirm this phenomenon, it is advisable to carry out a further, multidimensional analysis of the existing dependencies. The change in the method of calculating voivodeship development indices used in the paper, without changing the order, only changes their grouping.

REFERENCES

Binderman A. (2004) Spatial Differentiation of the Potential of Agriculture in Poland in the Years 1989-1998. R. N. Rol., Series G., 91(1), p. 51, Wieś Jutra, Warsaw.

Binderman A. (2005a) Classification of Polish Voivodships According to the Level of Agricultural Development. AAS, Series G., T.92, Z.1, p. 42, Wieś Jutra, Warsaw.

Binderman A. (2005b) On the Problem of Pattern Selection in the Study of Spatial Differentiation of Agricultural Potential in Poland. QME, V, Warsaw, p. 46.

Binderman A. (2006a) The Use of Utility Functions to Study the Spatial Diversity of Agriculture. R. N. SERiA, Volume VIII, Issue 5, Warsaw-Poznań, p. 5.

Binderman A. (2006b) Classification of Objects Based on Two Standards. EiOGŻ, Scientific Papers of the Warsaw University of Life Sciences, No. 60, Warsaw, p. 25.

Binderman A. (2007a) Multidimensional Analysis of Regional Diversity of Agriculture in Poland. PhD thesis, Warsaw University of Life Sciences.

Binderman A. (2007b) Changes in the Regional Level of Agriculture in Polish in the Years 1998-2005. Roczniki Naukowe SERiA, Volume IX, Warszawa-Kraków.

Binderman A. (2008a) Analysis of Regional Diversity of Agriculture in Polish in 2006. Roczniki Naukowe SERiA, Volume X, Zeszyt 2, Warszawa-Lublin.

Binderman A. (2008b) Application of the Linear and Nonlinear Utility Function to the Study of the Level of Agriculture in Poland. Quantitative Methods in Economics, 9, 29-38.

Binderman A. (2009a) Dynamics of Agricultural Development in Poland after Accession to the European Union. Annals of Agricultural Sciences, SERiA, Volume XI, Issue 3.

Binderman A. (2009b) Dependence of the Assessment of Agricultural Diversity in Poland on Selected Synthetic Measures. Quantitative Methods in Economics, 10(1), 30-41.

Binderman A. (2010) The Effect of Variable Normalization on the Assessment of Regional Agricultural Diversity. Quantitative Methods in Economics, 11(2), 25-38.

Binderman A. (2011) Multi-Criteria Methods of Analysis of the Diversity of Polish Agriculture in 2009. Quantitative Methods in Economics, 12(2), 58-68.

Binderman A. (2012) Regional Differentiation of Agriculture in Poland in the Years 1998-2010. Quantitative Methods in Economics, 13(3), 52-64.

Binderman A. (2013) Dynamics of Regional Differentiation of Agriculture in Poland in the Years 1998-2010. Quantitative Methods in Economics, 14(3), 17-26.

Binderman Z., Borkowski B., Szczesny W. (2020) The Use of Intermediate Standards for the Construction of Synthetic Meters. Quantitative Methods in Economics, 21(3), 157-167.

Binderman A., Krawiec M. (2006) Regional Differentiation of the Level of Agricultural Development in Poland in the Years 2002-2005. Development Potential of Rural Areas in the Aspect of Polish's Accession to the European Union. Szczecin, pp. 39-46.

Cieślak M. (2004) Taxonomic Procedure for Programming Economic Development and Determining the Demand for Qualified Staff, Statistical Review, 1.

Dąbkowski W. (2000) Regional Structures of Polish Agriculture. IRWiR PAN, Warsaw.

Fogelfors H. (ed.) (2009) Strategic Analysis of Swedish Agriculture. Production Systems and Agricultural Landscapes in a Time of Change. Report from the Department of Crop Production Ecology No.10, Swedish University of Agricultural Sciences (SLU), Uppsala.

Harasim A (2001) Indicators for the Assessment of Regional Diversity of Agriculture. Pamiętnik Puławski, 124, Puławy.

Harasim A. (ed.) (2006) Regional Differentiation of Agricultural Production in Poland. Studies and Reports of the Institute of Soil Science and Plant Science – NR Institute, 3.

Harasim A. (ed.) (2009) Selected Elements of Regional Differentiation of Agriculture in Poland. Studies and Reports of the Institute of Soil Science and Plant Cultivation – National Research Institute, 15.

Heller J., Zawisza S. (eds.) (2000) Problems of Regionalization in the Process of Shaping Rural Areas. Wyd. AT-R, Bydgoszcz.

Hellwig Z. (1981) Multivariate Comparative Analysis and its Application in the Study of Multivariate Economic Objects. [in:] Welfe W. (ed.) Methods and Models of Econ.-Mathematics. In Improving the Management of the Socialist Economy. PWE, Warsaw.

Hwang C. L., Yoon K. (1981) Multiple Attribute Decision Making, Methods and Applications. Springer – Verlag, New York.

Kisielińska J. (ed.) (2021) Multivariate Data Analysis in Agricultural Economics. Warsaw University of Life Sciences Press, Warsaw.

Krasowicz S., Kukuła S. (2006) Regional Differentiation of Polish Agriculture in the Light of Research Conducted by the Institute of Soil Science and Plant Cultivation – National Research Institute. Acta Agr. Silv., Econ. Section, 46(2), 73-89.

Kukuła K. (1989) Statistical Structural Analysis and its Application in the Sphere of Production Services for Agriculture. ZN, series spec.: Monographs, No. 89, AE, Kraków 1989.

Kukuła K. (2000) The Method of Zeroed Unitarization. PWN, Warsaw.

Malina A. (2004) Multidimensional Analysis of Spatial Differentiation of the Structure of the Polish Economy by Voivodship. AE, Series Monographs No. 162, Kraków.

Młodak A. (2006) Taxonomic Analysis in Regional Analysis. Difin, Warsaw.

National Agricultural Census 2020, 2022, ed. Central Statistical Office, Warsaw.

Nermend K., Miłaszewicz D. (2016) Diversification of Rural Areas in Poland – a Regional Approach. Quantitative Methods in Economics, 17(1), 72-83.

Nowak E. (1990) Taxonomic Methods in the Classification of Socio-Economic Objects. PWE, Warsaw.

Poczta W., Bartkowiak N. (2012) Regional Differentiation of Agriculture in Poland. Journal of Agribusiness and Rural Development, 1(23), 95-109.

Rabinowicz E. (2020) Sverige i EU efter 25 år – jordbruket och livsmedelsindustrin, Svenska Institutet för Europapolitiska studier, Rapport nr. 7, Stockholm.

Statistical Yearbook of Agriculture 2023 (2024) Statistics Poland, Warsaw.

Zegar J (ed.) (2003) Regional Differentiation of Agriculture. Statistics Poland, Warsaw.

Zeliaś A. (2000) Taxonomic Analysis of Spatial Differentiation of the Standard of Living in Poland in Dynamic Terms. University of Economics, Cracow.