



Anna Golec 

Uniwersytet Gdański, Wydział Zarządzania, Katedra Finansów Przedsiębiorstw
anna.golec@ug.edu.pl

Ocena skuteczności modelu Beneisha w wykrywaniu manipulacji w sprawozdaniach finansowych

Streszczenie: Celem artykułu jest ocena, czy model Beneisha może stanowić użyteczne narzędzie do wykrywania manipulacji wynikami finansowymi, które prowadziły do wydania negatywnej opinii biegłego rewidenta lub odmowy jej wydania w polskich spółkach kapitałowych. Badaniem objęto 24 pary przedsiębiorstw z głównego rynku Giełdy Papierów Wartościowych w Warszawie oraz z rynku alternatywnego New Connect. Z przeprowadzonych analiz wynika, że przy punkcie granicznym $-2,22$ model poprawnie identyfikował 67% manipulatorów i 75% niemanipulatorów. Dokładność modelu wzrastała z 71% do 75% wraz z przesuwaniem punktu odcięcia do $-1,98$. Kolejną obserwacją był fakt, że duże zmiany w wartościach *M-Score* okazały się lepszym kryterium oceny. Klasyfikacja podmiotów na podstawie 35% zmiany wskaźnika rok do roku pozwoliła zwiększyć dokładność grupowania do 85%.

Słowa kluczowe: model Beneisha, *M-Score*, manipulacje wynikami finansowymi, Polska, rynek kapitałowy

JEL: M42, G30

1. Wprowadzenie

Rozwój nowoczesnych narzędzi informatycznych spowodował daleko idące zmiany w procesie inwestowania na rynkach kapitałowych. Skróceniu uległ czas rozprzestrzeniania się informacji oraz realizacji zleceń. Uczestnicy rynku muszą mierzyć się z coraz większą ilością danych, między innymi dlatego na popularności zyskują narzędzia wykorzystujące algorytmy sztucznej inteligencji. Osiąganie ponadprzeciętnych stóp zwrotu na rynku akcji wymagać będzie zatem umiejętności wyselekcjonowania właściwych spółek w krótkim czasie, wspieranej przez odpowiednie narzędzia. Aby jednak zapewnić poprawne działanie tych narzędzi, dane wykorzystywane w analizie powinny być wiarygodne i właściwie odzwierciedlać rzeczywiste zdarzenia. Jednymi z najczęściej wykorzystywanych informacji w analizie finansowej są dane zawarte w sprawozdaniu finansowym, a elementem zapewnienia ich wiarygodności jest obowiązek jego badania przez biegłego rewidenta. Chociaż od spółek publicznych oczekuje się najwyższych standardów w zakresie sprawozdawczości, to zdarzają się przypadki, w których audytorzy zgłaszają zastrzeżenia lub odmawiają wydania opinii, co negatywnie rzutuje na wiarygodność informacji zawartych w sprawozdaniu takiej spółki. Opracowanie skutecznego i prostego w użyciu narzędzia, które dodatkowo pozwalałoby ocenić jakość zaprezentowanych wyników będących podstawą dalszych analiz, mogłoby zatem istotnie wspomóc proces decyzyjny. Największe korzyści z jego zastosowania mogliby osiągnąć interesariusze, którzy z jednej strony nie mają dostępu do wewnętrznych informacji, a z drugiej strony muszą poddawać analizie dużą liczbę spółek w krótkim czasie i przy ograniczonych nakładach finansowych, np. analitycy i doradcy inwestycyjni czy sami inwestorzy, jak również, w pewnym zakresie, instytucje nadzorujące rynek kapitałowy. Próby stworzenia takich narzędzi trwają od lat osiemdziesiątych XX wieku, a ich przykładem, który poddawany jest szerokiej dyskusji, jest tzw. model Beneisha¹. Narzędzie stało się popularnym elementem w ramach tzw. rachunkowości śledczej (*forensic accounting*). W ciągu ostatniej dekady autorzy z różnych krajów starali się zweryfikować, czy mimo upływu czasu oraz różnic w obszarach działalności, jurysdykcji podatkowej, standardach sprawozdawczości narzędzie to sprawdza się w identyfikacji spółek dokonujących manipulacji, czy nie. Niniejszy artykuł wpisuje się właśnie w tę dyskusję.

Celem artykułu jest ocena skuteczności modelu Beneisha dla spółek notowanych na polskim rynku giełdowym i alternatywnym oraz możliwości jego wykorzystania jako narzędzia wspierającego procesy decyzyjne inwestorów na rynku kapitałowym.

1 Zgodnie z informacjami zawartymi w Google Scholar na 9 września 2018 r. każda z publikacji M. Beneisha dotycząca wspomnianego modelu doczekała się ponad 650 cytowań.

Opracowanie składa się z pięciu części. W pierwszej zaprezentowano główne obszary i motywy dokonywania manipulacji wynikami finansowymi. Część druga zawiera omówienie głównych narzędzi służących weryfikacji wiarygodności informacji zawartych w sprawozdaniach finansowych. Kolejne fragmenty zawierają: przegląd badań innych autorów wykorzystujących model Beneisha, założenia i omówienie metodycznych aspektów przeprowadzonego badania oraz prezentację jego wyników. Całość zamyka podsumowanie, w którym zawarto główne wnioski i ograniczenia w interpretacji uzyskanych rezultatów w odniesieniu do postawionej hipotezy.

2. Manipulacje w sprawozdaniach finansowych

Problem niskiej wiarygodności danych jest nierozzerwalnie związany ze sprawozdawczością. W powszechnej świadomości pojawia się on zazwyczaj przy okazji skandali finansowych, jak to miało miejsce w przypadku Enronu, Toshiba czy Amber Gold, jednak nie można ograniczać go tylko do oszustw finansowych. Manipulacje wynikami na mniejszą skalę mogą prowadzić do podjęcia suboptymalnych decyzji inwestycyjnych i, choć znacznie mniej spektakularne, przynosić równie negatywne konsekwencje dla efektywności systemu finansowego. Dokonywanie manipulacji może odbywać się na różnych poziomach, również w ramach dopuszczalnych ram prawnych, co zaprezentowano w tabeli 1.

Przedmiotem zainteresowania w niniejszej publikacji jest sytuacja, w której dochodzi do daleko posuniętej księgowej manipulacji danymi ilościowymi, która istotnie wypacza wyniki, jednak nie do złamania prawa powodującego odpowiedzialność karną.

W literaturze wiele uwagi poświęcono zarówno motywom dokonywania oszustw i manipulacji, jak również metodom ich dokonywania. Rozbudowaną systematykę motywów zaprezentował M. El Diri (2018), co przedstawiono na rysunku 1.

Inną klasyfikację prezentują H. A. Hashim, Z. Salleh i A. M. Ariff (2013), którzy dzielą motywy na altruistyczne, spekulacyjne i wynikające z presji. Z kolei H. Schilit i J. Perler (2010) wskazują, że istotną rolę wśród motywów manipulowania wynikami finansowymi odgrywają chciwość (*greed factor*) i strach (*fear factor*) i wyróżniają siedem głównych obszarów, w których istnieje pole do nadużyć (tabela 2)

Przy budowaniu narzędzi, których zadaniem jest wykrywanie manipulacji wynikami, należy zatem przykładać szczególną uwagę do wymienionych obszarów.

Tabela 1. Różne poziomy manipulacji dotyczącej sprawozdań finansowych

	Zarządzanie oczekiwaniami (<i>Expectation management</i>)	Manipulacja wrażeniem (<i>Impression management</i>)	Realna manipulacja wynikiem (<i>Real earning management</i>)	Rachunkowa manipulacja wynikami (<i>Accrual earning management</i>)	Oszustwo księgowe (<i>Fraud</i>)
Istota	Udostępnianie lub ukrywanie informacji, których celem jest uzyskanie korzystnego benchmarku	Celowe działania zmierzające do wywarcia wpływu na interpretację wyników finansowych	Decyzje ekonomiczne oddziałujące na rzeczywiste przepływy finansowe	Świadoma ingerencja w dane finansowe, zniekształcająca je w porównaniu z podejściem neutralnym	Zniekształcenie lub fałszowanie danych dokonywane z naruszeniem prawa
Przykład	Publikacja prognoz	Zmiana skali na wykresie	Zmiana harmonogramu inwestycji	Zmiana wielkości szacunkowych	Uwzględnienie fikcyjnych zdarzeń
Czas względem realnych zdarzeń wpływających na wynik	<i>Ex-post</i>	<i>Ex-post</i>	<i>Ex-ante</i>	<i>Ex-ante</i> i <i>ex-post</i>	Na każdym etapie lub niezależnie od realnych zdarzeń
Czas względem publikacji sprawozdania	<i>Ex-ante</i>	Równoległe lub <i>ex-post</i>	<i>Ex-ante</i>	<i>Ex-ante</i>	<i>Ex-ante</i>
Naruszenie prawa/zasad rachunkowości	Nie	Nie	Nie	Nie/na granicy	Tak
Obszar manipulacji	Narracja, dane ilościowe	Narracja, opis jakościowy, grafika	Dane ilościowe	Dane ilościowe	Dane ilościowe oraz opis jakościowy

Źródło: opracowanie własne na podstawie El Diry, 2018

Teoria kontraktów	Teoria ograniczonej racjonalności	Teoria asymetrii informacyjnej
↓	↓	↓
Motywy kontraktowe	Motywy związane z rynkiem kapitałowym	Motywy zewnętrzne
<ul style="list-style-type: none"> – Wynagrodzenie menedżerów – Zmiana prezesa – Zdolności menedżerskie – Corporate Governance – Kredyty i klauzule restrykcyjne – Specyfika przedsiębiorstwa 	<ul style="list-style-type: none"> – Giełda – Emisja akcji – Debiut giełdowy lub notowania krzyżowe – Fuzje i przejęcia – Insider trading – Wykupy menedżerskie – Benchmark – Analitycy 	<ul style="list-style-type: none"> – Branża – Dywersyfikacja branżowa – Regulacje – Środowisko polityczne i specyficzne polityki krajowe – Standardy rachunkowości – Zagadnienia podatkowe – Konkurencja – Dostawcy i odbiorcy

Rysunek 1. Motywy kreowania wyników finansowych w ramach wybranych ujęć teoretycznych
 Źródło: opracowanie własne na podstawie El Diri, 2018

Tabela 2. Siedem obszarów dokonywania manipulacji w sprawozdaniach finansowych

Lp.	Obszary manipulacji wynikiem	Przykłady działań
1	Zbyt wczesne ujmowanie przychodów	Ujęcie całego przychodu w momencie, gdy świadczenia wobec odbiorcy się nie zakończyły Ujęcie przychodu mimo braku przeniesienia ryzyk na klienta
2	Zawyżanie przychodów	Manipulacja założeniami i danymi szacunkowymi Ujmowanie rabatów od dostawców jako przychodu
3	Podwyższanie wyniku finansowego jednorazowymi zdarzeniami	Sprzedaż aktywów niedowartościowanych w bilansie Niewyodrębnianie zdarzeń jednorazowych
4	Przesuwanie bieżących kosztów na inne okresy	Zbyt szybkie/wolne ujęcie amortyzacji Nieprawidłowe kapitalizowanie nakładów Przyspieszanie bądź wstrzymywanie się z dokonaniem zasadnego odpisu
5	Nieujmowanie wszystkich zobowiązań	Nieprawidłowe ujęcie zobowiązań warunkowych Wykorzystywanie transakcji klasyfikowanych jako zobowiązania pozabilansowe
6	Przesuwanie bieżącego dochodu na przyszłe okresy	Zawiązanie/rozwiązanie rezerw Dokonanie/odwrócenie odpisów
7	Przesuwanie przyszłych kosztów na bieżący okres	Uwzględnianie nakładów jako koszty zamiast kapitalizowania ich

Źródło: Schilit, Perler, 2010

3. Narzędzia wspomagające ocenę wiarygodności informacji zawartych w sprawozdaniu finansowym

Wiarygodność informacji zawartych w sprawozdaniu finansowym jest fundamentem zaufania na rynku kapitałowym, dlatego istnieje szereg elementów o charakterze instytucjonalnym, których zadaniem jest zapewnienie tego wymogu. Można tu wymienić regulacje definiujące standardy rachunkowości i sprawozdawczości, jak również przepisy wprowadzające sankcje za ich naruszanie. Równie ważnym elementem jest obowiązek badania sprawozdania przez biegłego rewidenta. Należy jednak podkreślić, że standardy wyznaczają dopuszczalne ramy elastyczności w ujęciu zdarzeń gospodarczych i niekiedy trudno zarysować jednoznaczną granicę między korzystaniem z tej elastyczności w dobrej lub złej wierze.

System zapewnienia wiarygodności informacji zawartych w sprawozdaniach finansowych nie działa jednak w sposób idealny i niektóre ze spółek dokonują naruszenia jego zasad, czego najbardziej znanym przypadkiem był skandal Enronu.

Problem możliwie jak najwcześniejszego wykrywania manipulacji w sprawozdaniu finansowym jest przedmiotem badań wielu autorów stosujących różne podejścia: od prostej analizy wskaźnikowej, przez data mining, po modele analizy dyskryminacyjnej (Persons, 1995; Summers, Sweeney 1998; Spathis, 2002; Kotsiantis i in., 2006; Fich, Shivdasani, 2007; Ata, Seyrek, 2009; Johnson, Ryan, Tian, 2009; Skousen, Twedt, 2009; Pai, Hsu, Wang, 2011; Gupta, Gill, 2012; Kanapickienė, Grundienė, 2015). Syntetyczny przegląd badań w tym zakresie prezentuje M. Sylwestrzak (2017). Należy podkreślić, że wyniki badań nie są jednoznaczne – na przykład K. A. Kaminski, T. S. Wetzel i L. Guan (2004) kwestionują skuteczność wskaźników finansowych w wykrywaniu nieprawidłowości. Z kolei F. H. Glancy i S. B. Yadav (2011) oraz M. Zaki i B. Theodoulidis (2013) proponują połączenie wykorzystania wskaźników finansowych z analizą lingwistyczną sprawozdań.

W latach osiemdziesiątych XX wieku zaczęły powstawać narzędzia służące wykrywaniu manipulacji w sprawozdaniu finansowym. Jako przykłady można tu wymienić modele, których autorami są: L. DeAngelo (1986), P. M. Dechow i R. G. Sloan (1991), J. Jones (1991), S. H. Kang i K. Sivaramakrishnan (1995), P. M. Dechow, R. G. Sloan i A. P. Sweeney (1995), M. D. Beneish (1999), P. M. Dechow i I. D. Dichev (2002), M. F. McNichols (2002), P. M. Dechow, S. A. Richardson i I. Tuna (2003), S. P. Kothari, A. J. Leone i C. E. Wasley (2005), J. Ye (2006), S. R. Stubben (2010). Modele takie mogłyby stanowić prosty i tani sposób wstępnej oceny wiarygodności sprawozdania, jednak warunkiem ich stosowania jest potwierdzenie ich skuteczności w danym kraju. Wszystkie wymienione podejścia wykorzystują wskaźniki analizy finansowej do wyliczenia odpowiednio wartości,

na podstawie których następnie wyznaczane jest prawdopodobieństwo wystąpienia nieprawidłowości. Przedmiotem zainteresowania w niniejszym artykule jest model Beneisha w wersji ośmioelementowej. Jest to model probitowy, oszacowany przy wykorzystaniu ważonej metody największej wiarygodności (*Weighted Exogenous Sample Maximum Likelihood* – WESML), który zakłada, że na podstawie wybranych zmiennych pochodzących ze sprawozdań finansowych można ocenić, czy w przedsiębiorstwie doszło do manipulacji wynikami, rozumianej jako naruszenie przyjętych standardów rachunkowości (Beneish, 1999) Model oszacowano, wykorzystując próbę 50 amerykańskich spółek giełdowych dokonujących manipulacji oraz 1708 spółek, w stosunku do których nie pojawiły się tego typu podejrzenia. Dane finansowe wytypowanych podmiotów pochodziły z lat 1982–1989 i były testowane na próbie kontrolnej z lat 1989–1992. Elementy składowe modelu zostały zaprezentowane w tabeli 3, a przy ich doborze przyjęto założenie, że większe prawdopodobieństwo manipulacji występuje w przypadku spółek stających w obliczu pogarszających się perspektyw finansowych, w których występują znaczące różnice memoriałowe, oraz tych, w których obecne są mechanizmy skłaniające menedżerów do zniekształcania danych.

Tabela 3. Elementy składowe modelu Beneisha

Nazwa	Formuła	Uzasadnienie
DSRI – Days Receivable Index	$\frac{Należności_t / Przychody_t}{Należności_{t-1} / Przychody_{t-1}}$	Możliwość zawyżania przychodów
GMI – Gross Margin Index	$\frac{(Przychody_{t-1} - KWS_{t-1}) / Przychody_{t-1}}{(Przychody_t - KWS_t) / Przychody_t}$	Negatywny sygnał dotyczący perspektyw na przyszłość
AQI – Asset Quality Index	$\frac{[1 - (AO_t + RzAT_t + InwDT_t)] / Aktywa_t}{[1 - (AO_{t-1} + RzAT_{t-1} + InwDT_{t-1})] / Aktywa_{t-1}}$	Możliwość przesuwania kosztów na przyszłe okresy przez ich aktywowanie w bilansie
SGI – Sales Growth Index	$\frac{Przychody_t}{Przychody_{t-1}}$	Presja na utrzymanie tempa wzrostu
DEPI – Depreciation Index	$\frac{Amortyzacja_{t-1} / (RzAT_{t-1} + Amortyzacja_{t-1})}{Amortyzacja_t / (RzAT_t + Amortyzacja_t)}$	Praktyki mające na celu podwyższenie zysków
SGAI – SG&A Expense Index	$\frac{KSZ_t / Przychody_t}{KSZ_{t-1} / Przychody_{t-1}}$	Utrata kontroli nad kosztami administracyjnymi lub nadmierna aktywność wspierająca sprzedaż

Nazwa	Formuła	Uzasadnienie
TATA – Total Accruals to Total Assets	$\frac{WDK_t - \text{Przepływy operacyjne netto}_t}{\text{Aktywa}_t}$	Duże różnice między wynikiem a rzeczywistymi przepływami mogą sugerować działania mające na celu zmanipulowanie wyniku
LVGI – Leverage Growth Index	$\frac{(ZB_t + \text{Zadłużenie } DT_t) / \text{Aktywa}_t}{(ZB_{t-1} + \text{Zadłużenie } DT_{t-1}) / \text{Aktywa}_{t-1}}$	Obecność klauzul restrykcyjnych, które mogą skłaniać do manipulacji

RzAT – rzeczowe aktywa trwałe

KWS – koszt własny sprzedaży

AO – aktywa obrotowe

InwDT – inwestycje długoterminowe

KSZ – koszty sprzedaży i zarządu

ZB – zobowiązania bieżące

ZadłużenieDT – zadłużenie długoterminowe

WDK – wynik z działalności kontynuowanej

Źródło: opracowanie własne na podstawie Beneish, Lee, Nichols, 2013

Wymienione w tabeli 3 elementy składowe modelu służą do wyliczenia wartości *M-Score* według wzoru (1)².

$$\begin{aligned}
 M\text{-Score} = & -4,84 + 0,92 \cdot DSRI + 0,528 \cdot GMI + 0,404 \cdot AQI + \\
 & + 0,892 \cdot SGI + 0,115 \cdot DEPI - 0,172 \cdot SGAI + \\
 & + 4,4679 \cdot TATA - 0,327 \cdot LVGI.
 \end{aligned} \quad (1)$$

Otrzymaną wartość *M-Score* porównuje się następnie z wartością graniczną, która zależy od akceptowanego poziomu błędów pierwszego i drugiego typu, czyli odpowiednio: niewykazania manipulacji, w przypadku gdy rzeczywiście do niej doszło, lub wskazania manipulacji, kiedy nie miała ona miejsca. Najczęściej spotykanymi w literaturze punktami granicznymi są $-2,22$ lub $-1,78$. Przekroczenie tego progu, czyli wartość większa od granicznej, wskazuje, że według modelu spółkę należy zaklasyfikować do grupy określanej mianem manipulatorów (*manipulators*), wartości mniejsze niż graniczna powodują zaliczenie spółki do grupy niemanipulatorów (*non-manipulators*)³.

Wynik *M-Score* ma za zadanie informować użytkownika o prawdopodobieństwie wystąpienia manipulacji w danych finansowych, co nie jest jednoznaczne ze stwierdzeniem, że w badanych spółkach sklasyfikowanych jako manipulatorzy

2 Warto podkreślić, że w literaturze występują dwie wersje modelu Beneisha. Obok podstawowego modelu uwzględniającego osiem zmiennych, funkcjonuje również model pięcioczynnikowy, pomijający zmienne *SGAI*, *DEPI* oraz *LVGI*, z identycznymi wartościami współczynników dla pozostałych zmiennych.

3 Takie określenia będą stosowane w dalszej części opracowania.

z całą pewnością doszło do nieprawidłowości, a w podmiotach z drugiej grupy manipulacje nie miały miejsca. Jak wskazuje sam autor modelu, uzyskanie takiej pewności wymaga bardziej wnikliwej analizy, której jednym z elementów może być na przykład porównanie wartości poszczególnych zmiennych objaśniających z wartościami referencyjnymi (Beneish, Lee, Nichols, 2013).

4. Wykorzystanie i ocena skuteczności modelu Beneisha przez innych autorów

Dorobek publikacyjny odnoszący się do modelu Beneisha można podzielić na trzy grupy. Najbardziej popularnym podejściem wśród autorów jest wykorzystywanie modelu Beneisha do oszacowania skali zjawiska manipulacji wynikami w różnych krajach.

R. Kaur, K. Sharma i A. Khanna (2014) przeanalizowali dane 332 spółek indyjskich z różnych sektorów w latach 2012 i 2013. Wyniki ich badań sugerują, że największa skala manipulacji występowała w spółkach z sektora IT (32%) oraz handlu (31%). Ciekawym spostrzeżeniem był również fakt, że spółki z najgorszymi wynikami, niezależnie od sektora, w ramach którego funkcjonowały, zostały zaklasyfikowane przez model jako manipulatorzy. F. Paolone i C. Magazzino (2014) badali sprawozdania 1809 włoskich spółek akcyjnych w latach 2005–2012. Za punkt graniczny uznali oni wartość *M-Score* na poziomie $-1,78$ i stwierdzili, że aż w 51,4% analizowanych przypadków mogło dochodzić do manipulacji wynikami. E. Kara, M. Korpi i M. Ugurlu (2015) dokonali analizy sprawozdań 132 spółek tureckich notowanych na giełdzie w Istambule, zaliczanych do sektora przemysłowego, w okresie 2010–2012. W ich badaniu aż 66 podmiotów zostało zaklasyfikowanych do grupy manipulatorów.

S. Repousis (2016) w swoich analizach dotyczących lat 2011–2012 wykazał, że z 25,5 tysiąca greckich spółek poddanych analizie blisko 8,5 tysiąca (33%) osiągnęło wartość *M-Score* przekraczającą $-2,22$, co może sugerować relatywnie dużą skalę manipulacji. Różnice w poziomie wskaźników *DSRI*, *AQI*, *SGAI*, *LVGI*, *GMI* oraz *SGI* między grupą domniemyanych manipulatorów i niemanipulatorów okazały się statystycznie istotne.

Wspomniane publikacje obejmowały liczne grupy podmiotów, jednak dużym ograniczeniem tego typu badań jest fakt, że nie podjęto w nich próby oceny skuteczności tego narzędzia, czyli weryfikacji, czy spółki wskazane jako manipulatorzy rzeczywiście dokonywały manipulacji.

Wyniki badań N.H. Anh i N.H. Linh (2016) na próbie 229 spółek niefinansowych notowanych na wietnamskiej giełdzie wskazały, że problem manipulacji wynikami w latach 2013–2014 dotyczył aż 48,4 % podmiotów. Wskazania modelu Beneisha okazały się spójne z raportami audytorów.

Drugi typ analiz bazuje na studiach przypadków pojedynczych przedsiębiorstw, a więc nie może być przenoszony na szerszą grupę podmiotów i uogólniany. Przykładem może być wykorzystanie modelu Beneisha obok modelu Altmana w studium przypadku Enronu, opracowane przez M. Mahama (2015), malezyjskiej spółki dokonującej manipulacji (Omar i in., 2014) czy spółki słowackiej (Petřík, 2016).

Trzecią grupę stanowią publikacje, w których autorzy próbują ocenić skuteczność identyfikacji przedsiębiorstw jako manipulatorów lub niemanipulatorów. Badania przeprowadzone przez Tarjo i H. Herawati (2015) w latach 2001–2014 objęły 35 indonezyjskich spółek, w stosunku do których instytucje nadzorujące rynek kapitałowy przeprowadziły postępowania wskazujące na oszustwa w sprawozdaniach finansowych. Każdej ze spółek przypisano również parę – podmiot z tego samego sektora o zbliżonej wartości aktywów. Analizy wykazały, że przy wykorzystaniu ośmioczynnikowego modelu Beneisha udało się poprawnie zaklasyfikować 77% manipulatorów i 80% niemanipulatorów.

Niższą skuteczność wykazały badania P. Marinakisa (2011). Dokonał on analizy 185 spółek brytyjskich, w stosunku do których w okresie od 1994 do 2006 roku komitet Financial Reporting Council zgłosił zastrzeżenia dotyczące wiarygodności sprawozdań finansowych. Wynika z nich, że model Beneisha w próbie estymacyjnej poprawnie identyfikował od 52% do 62% manipulatorów przy błędnym zaklasyfikowaniu od 8,4% do 19,4% spółek niedokonujących manipulacji, w zależności od przyjętej wartości granicznej. Dla próby kontrolnej wartości te wyniosły odpowiednio 42–57% oraz 7–10%. Model wzbogacony o dodatkowe zmienne uwzględniające udział wynagrodzenia rewidenta w sumie bilansowej (*Audit Fees to Total Assets Index – AUDI*), efektywną stopę opodatkowania (*Effective Tax Rate Index – EFTAX*) oraz stosunek wynagrodzeń dyrektorów do przychodów ze sprzedaży (*Directors Remuneration to Sales Index – DIRSI*) pozwolił zwiększyć liczbę poprawnie zaklasyfikowanych podmiotów.

Badania przeprowadzone przez zespół w składzie M. E. M. Kamal, M. F. M. Saleh i A. Ahmad (2016) dla 17 spółek z malezyjskiego rynku kapitałowego, przeciwko którym w latach 1996–2014 zostało wszczęte postępowanie komisji papierów wartościowych i giełd, wykazały, że model Beneisha poprawnie identyfikował 82% manipulatorów.

Mimo przeprowadzenia szerokich studiów literaturowych autorce nie udało się znaleźć badań oceniających skuteczność modelu Beneisha w warunkach polskich.

5. Metodyczne aspekty badania

Ze względu na dostępność danych autorka skupiła się w badaniu na spółkach notowanych na Giełdzie Papierów Wartościowych w Warszawie i New Connect. Ocena skuteczności modelu Beneisha wymaga w pierwszej kolejności zdefiniowania kryterium uznania spółki za manipulatora i zidentyfikowania podmiotów, spełniających to kryterium. Dla obserwatora z zewnątrz, który nie posiada wielu niezbędnych informacji na temat spółki, nie jest możliwe do zweryfikowania, czy w przedsiębiorstwie dochodziło do manipulacji mających na celu świadome zniekształcenie wyników, ponieważ z oczywistych względów spółki nie przyznają się do tego. Jedynymi ewidentnymi dowodami, że dochodziło do manipulacji w sprawozdaniach finansowych, mogą być poważne zastrzeżenia biegłych rewidentów lub postępowania wszczynane przez różne organy nadzoru (np. komisje nadzoru finansowego, urzędy skarbowe) zakończone nałożeniem sankcji. Dla potrzeb badania przyjęto, że podstawą zakwalifikowania podmiotu jako manipulatora będzie negatywna opinia audytora o sprawozdaniu lub raporcie okresowym lub odmowa jej wydania w latach 2014–2017. Rok sprawozdania, w stosunku do którego audytor wyraził negatywną opinię bądź odmówił jej wydania, uznano za przedział, w którym doszło do manipulacji (rok t). W ten sposób zidentyfikowano 24 spółki, którym następnie przypisano parę (spółkę kontrolną) w postaci podmiotu z tego samego sektora, prowadzącego możliwie jak najbardziej zbliżony rodzaj działalności, w którym w badanym okresie nie występowały problemy ze sprawozdaniem, a także inne zdarzenia, jak na przykład restrukturyzacja, upadłość, czy brak walnego zgromadzenia w kodeksowym terminie. Największe trudności w połączeniu podmiotów w pary pojawiły się w przypadku spółek zaliczonych do sektora energetycznego i „innych dóbr konsumpcyjnych” ze względu na ich unikatowość. Szczegółowe informacje dotyczące klasyfikacji spółek przedstawiono w załączniku 1. W ten sposób wyodrębniono dwie grupy: pierwszą, zawierającą podmioty, w których wystąpił problem z wiarygodnością danych finansowych oraz drugą, uwzględniającą jednostki, w których problem u tego nie zidentyfikowano, stanowiącą grupę kontrolną.

Następnie zebrane zostały dane finansowe niezbędne do obliczenia wartości wskaźników będących składowymi modelu. Analizie w ramach każdej pary spółek poddawano rok manipulacji (rok t) oraz rok poprzedzający (rok $t - 1$), co wymagało zgromadzenia danych za trzy lata obrotowe (od $t - 2$ do t). Wyjątkami były spółki Qumak i Vistal, dla których odmowa wydania opinii przez audytora miała miejsce w 2017 roku i dotyczyła raportu półrocznego. W tych dwóch przypadkach, ze względu na brak sprawozdania za 2017 rok, za rok manipulacji przyjęto rok 2016. Podstawą obliczeń były informacje zawarte w sprawozdaniach skonsolidowanych, o ile spółka je przygotowywała. Dane pobrano z bazy Notoria, a w przy-

padku ich braku lub niekompletności były one szczegółowo analizowane i uzupełniane na podstawie sprawozdań badanych podmiotów.

W przypadkach gdy wskaźnika nie dało się obliczyć ze względu na występowanie wartości zerowych w mianowniku, podobnie jak to zrobił M. Beneish, a za nim także inni autorzy badający skuteczność jego modelu, obserwacja nie była eliminowana, a wskaźnikowi przypisywane były wartości „neutralne”, równe 1. Podobnie postępowano w przypadku braku możliwości obliczenia wskaźnika, na przykład ze względu na układ sprawozdania.

Elementem wymagającym komentarza jest również zastosowane podejście do wartości znacząco odbiegających od pozostałych w próbie. Powszechnym zabiegiem jest odcinanie wartości skrajnych. Autorka przeprowadziła analizę danych zarówno bez dodatkowej korekty, jak i przy winsoryzacji na poziomie 5% górnych i dolnych wartości. Wyniki obu analiz okazały się zbieżne, dlatego prezentacji i omówieniu podlegają jedynie wnioski z badania na podstawie danych bez korekty. Wybrane statystyki rozkładów wskaźników obliczonych dla grupy manipulatorów oraz niemanipulatorów zaprezentowano w tabeli 4.

Tabela 4. Parametry rozkładów oraz ich porównanie w grupie manipulatorów i niemanipulatorów

Wskaźnik	Manipulatorzy			Niemanipulatorzy			Test równości średnich U Manna-Whitneya		
	μ	σ	J-B	μ	σ	J-B	<i>U</i>	<i>Z</i>	<i>p</i>
<i>DSRI</i>	23,54	139,04	4400,29	1,00	0,26	1,40	936,0	1,58	0,11
<i>GMI</i>	1,06	0,24	102,88	0,99	0,05	382,38	1147,0	-0,03	0,97
<i>AQI</i>	0,91	3,57	657,30	1,02	0,44	19,18	1058,0	-0,69	0,49
<i>SIGI</i>	12,08*	79,27	4433,47	1,05*	0,18	5,98**	510,0	-4,70	0,00
<i>DEPI</i>	1,22	1,16	690,41	1,01	0,45	2062,44	1100,0	0,38	0,71
<i>SG&A</i>	1,84*	2,78	915,37	0,98*	0,18	5,33**	874,0	2,03	0,04
<i>LVI</i>	1,63*	1,67	625,08	1,10*	0,38	1103,25	636,0	3,78	0,00
<i>TATA</i>	-0,41*	0,84	1035,67	-0,07*	0,09	68,32	709,0	-3,24	0,00

*Wskaźniki, których średni poziom różnił się w grupie manipulatorów i niemanipulatorów w sposób statystycznie istotny ($p = 0,05$).

**Wskaźniki o rozkładzie normalnym – test rozkładu normalnego Jarque’a-Bera $\chi^2(2) = 5,99, \alpha = 0,05$.

Źródło: opracowanie własne

Z przedstawionego w tabeli 4 zestawienia wynika, że rozkłady wartości wskaźników jedynie w dwóch przypadkach – *SIGI* oraz *SG&A* dla grupy niemanipulatorów miały rozkład normalny. Kształt pozostałych rozkładów nie był zgodny z normalnym, co wskazuje na występowanie obserwacji znacznie różniących się od wartości średniej danego wskaźnika w obu badanych grupach spółek. Aby ocenić, czy różnice w poziomie średnich wartości wskaźników

w obu grupach były statystycznie istotne, wykorzystano test U Manna-Whitneya, który potwierdził występowanie różnic średnich dla wskaźników *SGI*, *SG&A*, *LVI* oraz *TATA*.

Podstawą oceny skuteczności modelu były powszechnie stosowane miary statystyczne wykorzystywane w celu wyznaczenia krzywej *ROC*, takie jak: czułość i swoistość oraz wartość predykcyjna dodatnia i ujemna, iloraz wiarygodności wyniku dodatniego i ujemnego oraz dokładność. Szczegółowy opis tych miar wraz z ich interpretacją przedstawiono w załączniku 1.

6. Ocena skuteczności modelu na przykładzie polskich spółek

W celu oceny skuteczności modelu Beneisha dla emitentów z rynku giełdowego i alternatywnego postawiono hipotezę badawczą, iż model Beneisha pozwoli przynajmniej w 75% przypadków poprawnie klasyfikować spółki do grupy manipulatorów i niemanipulatorów na podstawie danych z roku, którego dotyczyły nieprawidłowości lub roku poprzedzającego.

Pierwotnie przyjęto, że punktem odcięcia będzie wartość najczęściej wskazywana w badaniach innych autorów, tj. $-2,22$, dopuszczając możliwość jego przesunięcia, gdyby skuteczność modelu okazała się wyższa dla innej wartości różnicującej.

Wartości wskaźnika *M-Score* dla grup spółek manipulatorów i niemanipulatorów zaprezentowano w tabeli 5. Kolorem szarym oznaczono przekroczenie wartości $-2,22$ uznanej za punkt odcięcia, czyli klasyfikację przez model do grupy manipulatorów.

Z przedstawionych danych wynika, że w przypadku spółek manipulatorów model nie zawsze jednoznacznie identyfikuje problem w roku manipulacji. W roku t wartość *M-Score* tylko w 50% obserwacji przekroczyła zalecany poziom. W roku poprzedzającym dla 10 z 24 podmiotów wartość wskaźnika była powyżej wartości progowej. Należy jednak pamiętać, że baza Notoria zawiera dane, które mogły podlegać korekcie po decyzji audytora. W jednej trzeciej podmiotów zakwalifikowanych do grupy manipulatorów w żadnym z analizowanych okresów poziom *M-Score* nie wykazał problemów.

W przypadku grupy niemanipulatorów 75% spółek zostało poprawnie sklasyfikowanych. W roku t model wskazywał na manipulację wynikiem w 4 podmiotach, zaś w roku $t - 1$ w trzech podmiotach. Warto zauważyć, że w roku $t - 1$ wartości *M-Score* były bliskie wartości progowej.

Tylko dla jednej spółki wskaźnik w obu badanych okresach nieznacznie przekroczył zalecany poziom. Skuteczność klasyfikacji przedstawiono w tabeli 6.

Tabela 5. Wartości *M-Score* dla badanych grup spółek

Manipulatorzy	<i>t</i> - 1	<i>t</i>	<i>t</i> - 1	<i>t</i>	Niemanipulatorzy
Admiral Boats	3,51	-4,03	-2,51	-1,69	Robinson Europe
Alma Market	-4,09	-7,52	-2,21	-2,54	Emperia Holding
Aton-HT	17,46	6,55	-2,86	-2,86	Krynicky Recykling
BGE	-7,28	-7,21	-2,76	-2,69	Energa
Calatrava Capital	-1,40	-4,66	-3,82	-2,87	Skyline Investments
Drewex	-5,28	-2,08	-2,75	-2,65	KPD
Drop	-2,82	-6,50	-2,35	-2,63	Orzeł Biały
Euroinvest/Vision	882,76	480,77	-3,36	-2,36	ATM Grupa
Fiten	-0,69	-1,93	-3,46	-2,87	Tauron
Fota	-4,91	-2,44	-2,19	-2,05	Intercars
Genesis Energy	-7,51	7,28	-3,16	-2,44	Kogeneracja
Ideon	-4,47	-1,60	-3,54	-2,55	GTC
Imagis	-9,81	-2,96	-1,99	-2,73	Neptis
Mediatel	-3,11	13,36	-2,79	-3,91	MNI
Novavis	12,1	-2,37	-2,46	-4,99	ZE PAK
PBG	-1,90	-2,44	-2,60	-1,86	Polimex Mostostal
PC Guard	4,02	-1,31	-2,64	-2,71	Asseco BS
Petrolinvest	-3,55	-1,66	-2,81	-3,21	Lotos
Platynowe Inwestycje	-2,96	-13,66	-3,04	-2,32	Capital Park
Qumak	-9,97	-3,23	-2,53	-2,49	Comarch
Redwood Holding	-3,36	-2,18	-2,50	-2,05	Rafamet
Regnon	-1,57	-2,03	-2,49	-2,94	Komputronik
Stoppoint	-3,00	-31,52	-3,10	-2,58	Black Point
Vistal Gdynia	-2,24	-1,7	-3,06	-2,33	Budimex
Min./Maks.	-31,52/882,76		-4,99/-1,69		Min/Max

Źródło: opracowanie własne

Tabela 6. Skuteczność klasyfikacji badanych podmiotów dla *M-Score* = -2,22

Model	Rzeczywistość	Manipulatorzy	Niemanipulatorzy	Suma
Manipulatorzy	Wskazania prawdziwie pozytywne (TP) 16	Błąd I typu wskazania fałszywie dodatnie (FP) 6		22
Niemanipulatorzy	Błąd II typu wskazania fałszywie negatywne (FN) 8	Wskazania prawdziwie negatywne (TN) 18		26
Suma	24	24		n = 48

Źródło: opracowanie własne

Tak przyjęte kryterium klasyfikacji pozwoliło właściwie zidentyfikować 67% manipulatorów i 75% niemanipulatorów, czyli łącznie poprawnie sklasyfikowane zostało 71% badanych podmiotów. Największą skuteczność model uzyskał

dla punktu granicznego z przedziału od $-2,04$ do $-1,98$ włącznie, gdzie błędnie zidentyfikowanych zostało 42% manipulatorów i 8% niemanipulatorów, a łączna dokładność klasyfikacji wyniosła 75%. Szczegółową ocenę właściwości diagnostycznych modelu dla progów $-2,22$ oraz $-1,98$ przedstawiono zbiorczo w załączniku 1.

Analiza uzyskanych wyników pozwoliła na zaobserwowanie dwóch ciekawych zjawisk, które mogą stanowić skuteczniejszą podstawę różnicowania spółek manipulatorów i niemanipulatorów. Po pierwsze, obserwacja górnych i dolnych zakresów realizowanych wartości *M-Score* prowadzi do wniosku, że rozpiętość wyników jest znacznie większa w przypadku manipulatorów. Wśród niemanipulatorów nie obserwowano wartości mniejszych od -5 ani większych niż $-1,68$. Gdyby zatem uznać, że wątpliwości dotyczące jakości sprawozdania powinny budzić nie tylko wartości *M-Score* powyżej $-1,98$, ale również te poniżej -5 , to takie podejście umożliwiłoby poprawne zidentyfikowanie po 92% podmiotów (22 spółki) z każdej grupy.

Drugie spostrzeżenie dotyczy zmian wartości wskaźnika w roku manipulacji w stosunku do roku poprzedzającego. W przypadku manipulatorów zakres zmian wynosił od 1 do blisko 950%, ze średnią zmianą na poziomie 53%, podczas gdy dla niemanipulatorów zakres zmian mieścił się w przedziale od 0,3 do 103%, ze średnią wynoszącą 7,5%. Szczegółowe wyniki analizy przedstawia tabela 7, w której kolorem szarym oznaczono przekroczenie poziomu 34%.

Spółki, w których wartość wskaźnika *M-Score* zmieniła się o ponad 34%, w większości okazywały się manipulatorami. Skuteczność klasyfikacji na podstawie tego kryterium była zdecydowanie wyższa niż w przypadku wykorzystania jedynie wartości *M-Score*, co zaprezentowano w załączniku 1. Zastosowane podejście pozwoliło na zwiększenie dokładności i poprawne rozpoznanie 85% analizowanych podmiotów.

Tabela 7. Wartości zmian wskaźnika *M-Score* w roku t w stosunku do roku $t - 1$

Manipulatorzy	Wartość bezwzględna zmiany <i>M-Score</i>		Niemanipulatorzy
Admiral Boats	214,9%	32,6%	Robinson Europe
Alma Market	83,9%	15,0%	Emperia Holding
Aton-HT	62,5%	0,3%	Krynicki Recykling
BGE	1,0%	2,7%	Energa
Calatrava Capital	233,5%	24,8%	Skyline Investments
Drewex	60,7%	3,5%	KPD
Drop	130,2%	12,0%	Orzeł Biały
Euroinvestment/Vision	45,5%	29,8%	ATM Grupa
Fiten	179,5%	17,3%	Tauron
Fota	50,3%	6,4%	Intercars

Manipulatorzy	Wartość bezwzględna zmiany <i>M-Score</i>		Niemanipulatorzy
Genesis Energy	197,0%	23,0%	Kogeneracja
Ideon	64,2%	27,8%	GTC
Imagis	69,9%	37,2%	Neptis
Mediatel	529,7%	40,2%	MNI
Novavis	119,6%	102,7%	ZE PAK
PBG	28,6%	28,5%	Polimex Mostostal
PC Guard	132,5%	2,8%	Asseco BS
Petrolinvest	53,1%	14,5%	Lotos
Platynowe Inwestycje	361,4%	23,6%	Capital Park
Qumak	67,6%	1,5%	Comarch
Redwood Holding	35,1%	18,0%	Rafamet
Regnon	29,4%	17,8%	Komputronik
Stoppoint	949,7%	16,8%	Black Point
Vistal Gdynia	24,4%	24,0%	Budimex
Min	1,0%	0,3%	Min
Max	949,7%	102,7%	Max
Średnia	53,2%	7,5%	Średnia

Źródło: opracowanie własne

7. Wnioski i ograniczenia

Porównując otrzymane wyniki z rezultatami badań innych autorów, stwierdzić można, że poprawność klasyfikacji polskich spółek, w których wystąpiły problemy z wiarygodnością sprawozdań finansowych, w oparciu o graniczny poziom *M-Score* równy $-2,22$, była wyższa niż w przypadku badań prowadzonych dla rynku brytyjskiego, lecz niższa niż dla rynku indonezyjskiego czy malezyjskiego. W zakresie poprawnej klasyfikacji podmiotów do grupy niemanipulatorów na polskim rynku model Beneisha sprawdził się najgorzej, gdyż jego 75-procentowa poprawność klasyfikacji w tym obszarze była niższa niż u pozostałych autorów, gdzie każdorazowo przekraczała 80%. Dokładność modelu w warunkach polskich wyniosła 71% lub 75% dla punktów granicznych wynoszących odpowiednio $-2,22$ oraz $-1,98$ i była niższa niż w badaniu Tarjo i N. Herawati (2015), w którym poprawnie sklasyfikowane zostało blisko 79% wszystkich spółek.

Z przeprowadzonych analiz wynika, że model Beneisha może być skutecznym narzędziem do wstępnej identyfikacji manipulacji danymi finansowymi, jednak interpretując wyniki badania, należy mieć na względzie kilka istotnych ograniczeń. Analizie poddano relatywnie niewielką grupę podmiotów, ponieważ niełatwo dotrzeć do informacji na temat nieprawidłowości w sprawozdaniach finansowych. Tak mała liczba obserwacji nie pozwoliła na wydzielenie grupy kontrolnej,

na której można by potwierdzić skuteczność przyjętych kryteriów klasyfikacji. Będzie to możliwe w następnym kroku, w miarę pojawiania się kolejnych informacji, które pozwolą zwiększyć liczbę spółek klasyfikowanych jako manipulatorów.

Drugim ograniczeniem, na które warto zwrócić uwagę, jest fakt, że przypisując podmiotom z grupy manipulatorów parę, nie zawsze udaje się dobrać spółkę na tyle podobną, aby odwzorować wszystkie cechy mogące wpływać na zakres manipulacji. Przykładem mogą tu być spółki z sektora energetycznego zaliczone do grupy manipulatorów, które były podmiotami relatywnie małymi, notowanymi na rynku alternatywnym i prowadzącymi działalność głównie w zakresie odnawialnych źródeł energii i nie było możliwości dobrania do pary spółek z pozytywną opinią rewidenta posiadających taką charakterystykę.

Ponadto nawet w przypadku wydania pozytywnej opinii przez audytora należy to raczej utożsamiać z brakiem podstaw do uznania spółki za manipulatora niż z gwarancją, że wszystkie liczby w sprawozdaniu zostały przedstawione w sposób obiektywny.

Ponieważ analiza obejmuje podmioty z różnych sektorów i na przestrzeni różnych okresów, potencjalnym problemem mogą być różnice w sprawozdaniach, będące wynikiem specyfiki prowadzonej działalności, oraz zmiany zachodzące w czasie (np. w otoczeniu rynkowym). Podobnie jak w przypadku powszechnie znanego modelu Altmanna, wykorzystywanego do prognozowania bankructwa, konieczne może okazać się okresowe dopasowanie współczynników modelu.

Mimo wspomnianych ograniczeń badanie skuteczności modelu Beneisha na przykładzie polskich spółek kapitałowych może stanowić kolejny głos w globalnej dyskusji na temat metod wykrywania manipulacji danymi finansowymi.

Bibliografia

- Anh N.H., Linh N.H. (2016), *Using the M-score Model in Detecting Earnings Management: Evidence from Non-Financial Vietnamese Listed Companies VNU*, „Journal of Science: Economics and Business”, t. 32, nr 2, s. 14–23.
- Ata H., Seyrek I. (2009), *The Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms*, „The Journal of Faculty of Economics and Administrative Sciences”, nr 14(2), s. 157–170.
- Beneish M.D. (1999), *The detection of earnings manipulation*, „Financial Analysts Journal”, t. 55, nr 5, s. 24–36.
- Beneish M.D., Lee C.M.C., Nichols D.C. (2013), *Earnings Manipulation and Expected Returns*, „Financial Analysts Journal”, t. 69, nr 2, s. 57–82.
- DeAngelo L. (1986), *Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders*, „The Accounting Review”, nr 61, s. 400–420.
- Dechow P. M., Dichev I.D. (2002), *The quality of accruals and earnings: The role of accrual estimation errors*, „The Accounting Review”, nr 77, s. 35–59.
- Dechow P. M., Richardson S.A., Tuna I. (2003), *Why are earnings kinky? An examination of the earnings management explanation*, „Review of Accounting Studies”, nr 8, s. 355–384.

- Dechow P. M., Sloan R.G. (1991), *Executive incentives and the horizon problem: An empirical investigation*, „Journal of Accounting and Economics”, nr 14, s. 51–89.
- Dechow P. M., Sloan R.G., Sweeney A.P. (1995), *Detecting earnings management*, „The Accounting Review”, nr 70, s. 193–193.
- El Diri M. (2018), *Introduction to earning management*, Springer International Publishing, Cham.
- Fich E.M., Shivdasani A. (2007), *Financial Fraud, Director Reputation, and Shareholder Wealth*, „Journal of Financial Economics”, nr 86(2), s. 306–333.
- Glancy F.H., Yadav S.B. (2011), *A computational model for financial reporting fraud detection*, „Decision Support Systems”, t. 50, cz. 3, s. 595–601.
- Gupta R., Gill N. (2012), *Prevention and Detection of Financial Statement Fraud – An Implementation of Data Mining Framework*, „Editorial Preface”, nr 3(8), s. 150–160.
- Hashim H.A., Salleh Z., Ariff A.M. (2013), *The Underlying Motives for Earnings Management: Directors, Perspective*, „International Journal of Trade, Economics and Finance”, t. 4, nr 5, s. 296–299.
- Johnson S., Ryan H., Tian Y. (2009), *Managerial Incentives and Corporate Fraud: The Sources of Incentives Matter*, „Review of Finance”, nr 13(1), s. 115–145.
- Jones J. (1991), *Earnings management during import relief investigations*, „Journal of Accounting Research”, nr 29(2), s. 193–228.
- Kamal M.E.M., Salleh M.F.M., Ahmad A. (2016), *Detecting financial statement fraud by Malaysian public listed companies: The reliability of the Beneish M-Score model*, „Journal Pengu-rusan”, nr 46, s. 23–32.
- Kaminski K.A., Wetzel T.S., Guan L. (2004), *Can financial ratios detect fraudulent financial reporting?*, „Managerial Auditing Journal”, t. 19, cz. 1, s. 15–28.
- Kanapickienė R., Grundienė Ž. (2015), *The Model of Fraud Detection in Financial Statements by Means of Financial Ratios*, „Procedia – Social and Behavioral Sciences”, nr 213, s. 321–327.
- Kang S.H., Sivaramakrishnan K. (1995), *Issues in testing earnings management and an instrumental variable approach*, „Journal of Accounting Research”, nr 33, s. 353–367.
- Kara E., Korpi M., Ugurlu M. (2015), *Using Beneish model in identifying accounting manipulation: an empirical study in BIST manufacturing industry sector*, „Journal of Accounting, Finance and Auditing Studies”, nr 1(1), s. 21–39.
- Kaur R., Sharma K., Khanna A. (2014), *Detecting Earnings Management in India – A sector-wise Study on European*, „Journal of Business and Management”, t. 6, nr 11, s. 11–18.
- Kothari S.P., Leone A.J., Wasley C.E. (2005), *Performance matched discretionary accrual measures*, „Journal of Accounting and Economics”, nr 39, s. 163–197.
- Kotsiantis S., Koumanakos E., Tzelepis D., Tampakas V. (2006), *Forecasting Fraudulent Financial Statements Using Data Mining*, „International Journal of Computational Intelligence”, nr 3(2), s. 104–110.
- Mahama M. (2015), *Detecting corporate fraud and financial distress using the Altman and Beneish models*, „International Journal of Economics, Commerce and Management”, nr 3(1), s. 1–18.
- Marinakos P. (2011), *An investigation of earnings management and earnings manipulation in the UK*, praca doktorska, Nottingham University.
- McNichols M.F. (2002), *Discussion of: The quality of accruals and earnings – The role of accrual estimation errors*, „The Accounting Review”, t. 77, nr s–1, s. 61–69.
- Omar N., Koya R.K., Sanusi Z.M., Shafie N.A. (2014), *Financial statement fraud: A Case examination using beneish model and ratio analysis*, „International Journal of Trade, Economics and Finance”, t. 5, nr 2, s. 184–186.
- Pai P., Hsu M., Wang M. (2011), *A Support Vector Machine-Based Model for Detecting Top Management Fraud*, „Knowledge-Based Systems”, nr 24(2), s. 314–321.
- Paolone F., Magazzino C. (2014), *Earnings manipulation among the main industrial sectors: Evidence from Italy*, „Economia Aziendale”, nr 5, s. 253–261.

- Persons O. (1995), *Using Financial Statement Data to Identify Factors Associated with Fraudulent Financial Reporting*, „Journal of Applied Business Research”, nr 11(3), s. 38–46.
- Petrík V. (2016), *Application of Beneish M-Score on Selected Financial Statements*, Conference: Bezpečné Slovensko a Európska Únia at: Košice, Slovakia – The University of Security Management in Košice, t. 1, <https://www.researchgate.net/publication/311733912> [dostęp: 2.02.2018].
- Repousis S. (2016), *Using Beneish model to detect corporate financial statement fraud in Greece*, „Journal of Financial Crime”, t. 23 cz. 4, s. 1063–1073, <https://doi.org/10.1108/JFC-11-2014-0055>.
- Schilit H., Perler J. (2010), *Financial Shenanigans: How to Detect Accounting Gimmicks & Fraud in Financial Reports*, 3rd edition, McGraw-Hill, New York.
- Skousen Ch.J., Twedt B.J. (2009), *Fraud score analysis in emerging markets*, „Cross Cultural Management: An International Journal”, t. 16, cz. 3, s. 301–316.
- Spathis C. (2002), *Detecting False Financial Statements Using Published Data: Some Evidence From Greece*, „Managerial Auditing Journal”, nr 17(4), s. 179–191.
- Stubben S.R. (2010), *Discretionary revenues as a measure of earnings management*, „The Accounting Review”, t. 85, nr 2, s. 695–717.
- Summers S., Sweeney J. (1998), *Fraudulently Misstated Financial Statements and Insider Trading: An Empirical Analysis*, „Accounting Review”, nr 73(1), s. 131–146.
- Sylwestrzak M. (2017), *Wykorzystanie modelu CART-Logit do analizy fałszerstw sprawozdań finansowych*, „Finanse, Rynki Finansowe, Ubezpieczenia”, nr 4 (88/1), s. 403–412, <http://dx.doi.org/10.18276/frfu.2017.88/1-39> [dostęp: 1.02.2018].
- Tarjo, Herawati N. (2015), *Application of Beneish M-Score Models and Data Mining to Detect Financial Fraud*, „Procedia – Social and Behavioral Sciences”, nr 211, s. 924–930.
- Ye J., (2007), *Accounting Accruals and Tests of Earnings Management*, <https://ssrn.com/abstract=1003101> [dostęp: 1.02.2018].
- Zaki M., Theodoulidis B. (2013), *Analyzing Financial Fraud Cases Using a Linguistics-Based Text Mining Approach*, <https://ssrn.com/abstract=2353834> or <http://dx.doi.org/10.2139/ssrn.2353834> [dostęp: 1.02.2018].

Załącznik 1. Opis miar wykorzystanych do oceny modelu Beneisha wraz z uzyskanymi wynikami

Miara (opis)	Formuła	Kryterium*			Interpretacja miar dla $\Delta M > 34\%$
		$M > -2,22$	$M > -1,98$	$\Delta M > 34\%$	
Czułość Opisuje zdolność testu do wykrywania jednostek rzeczywicie posiadających daną cechę	$\text{czułość} = \frac{TP}{(TP + FN)}$	67%	58%	83%	Z 83-procentowym prawdopodobieństwem, manipulator zostanie poprawnie sklasyfikowany przez model.
Swoistość Opisuje zdolność testu do wykrywania jednostek rzeczywicie nieposiadających danej cechy	$\text{swoistość} = \frac{TN}{(FP + TN)}$	75%	92%	88%	Z 88-procentowym prawdopodobieństwem, niemanipulator zostanie poprawnie sklasyfikowany przez model.
Wartość predykcyjna dodatnia Prawdopodobieństwo, że jednostka posiada cechę, mając dodatni wynik testu	$PPV = \frac{TP}{(FP + TP)}$	73%	88%	87%	Z 87-procentowym prawdopodobieństwem, pozytywne wskazanie przez model wskazuje rzeczywistego manipulatora.
Wartość predykcyjna ujemna Prawdopodobieństwo, że dana jednostka nie ma danej cechy, mając negatywny wynik testu	$NPV = \frac{TN}{(FN + TN)}$	69%	69%	84%	Z 84-procentowym prawdopodobieństwem negatywne wskazanie przez model rzeczywicie wskazuje spółkę, która nie manipuluje wynikami.
Iloraz wiarygodności wyniku dodatniego Iloraz dwóch prawdopodobieństw: tego, że pozytywny wynik testu otrzyma jednostka posiadająca cechę, do tego, że pozytywny wynik otrzyma jednostka bez danej cechy	$LR_+ = \frac{\text{czułość}}{(1 - \text{swoistość})}$	2.67	7.00	6.67	Prawdopodobieństwo, że spółka wskazana przez model jako manipulator rzeczywicie jest nim, jest blisko siedmiokrotnie większe niż ryzyko pomyłki.

<p>Iloraz wiarygodności wyniku ujemnego Iloraz dwóch prawdopodobieństw: tego, że negatywny wynik testu otrzyma jednostka posiadająca cechę, do tego, że negatywny wynik otrzyma jednostka nieposiadająca cechy</p>	$LR_- = \frac{1 - \text{czułość}}{\text{swoistość}}$	0,44	0,45	0,19	<p>Prawdopodobieństwo, że spółka wskazana przez model jako niemanipulator w rzeczywistości jest nim, jest ponad pięciokrotnie niższe niż szansa poprawnej diagnozy.</p>
<p>Dokładność Zdolność do stawiania prawidłowej diagnozy przy wykorzystaniu testu</p>	$ACC = \frac{TP + FN}{n}$	71%	75%	85%	<p>Prawdopodobieństwo, że uzyskany wynik klasyfikacji na manipulatorów i niemanipulatorów jest w 85% poprawny.</p>

* Przyjęte kryteria oznaczają:

$M > -2,22$ miary skuteczności modelu dla wartości M -Score powyżej $-2,22$,

$M > -1,98$ miary skuteczności modelu dla wartości M -Score powyżej $-1,98$,

$\Delta M > 34\%$ miary skuteczności modelu dla zmian wartości M -Score przekraczających 34% w ujęciu rok do roku.



Źródło: opracowanie własne

Effectiveness of the Beneish Model in Detecting Financial Statement Manipulations

Abstract: The aim of this study is to verify whether Beneish *M-Score* model can be useful in detecting Polish companies involved in earning management practices that lead to adverse or disclaimer of auditors' opinion. The sample covers 24 pairs of firms listed on Warsaw Stock Exchange or New Connect (alternative market). The findings generally indicate that with -2.22 point cut-off the model was able to identify 67% of manipulators and 75% non-manipulators correctly. The accuracy of the model improved from 71% to 75% after shifting the cut-off point to -1.98. Another observation was that high changes in *M-Score* values turned out to be better indicator of manipulation and the classification based on 35% change in year-to-year values reached 85% accuracy.

Keywords: Beneish Model, *M-Score*, financial statement manipulation, Poland, listed companies

JEL: M42, G30

	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p>
	<p>This journal adheres to the COPE's Core Practices https://publicationethics.org/core-practices</p>