ISSN 2392-1641
e-ISSN 2450-0097

Economics and Business Review

Volume 9 (4) 2023

CONTENTS

Poznań University of Economics and Business Press

Economics and Business Review

Volume 9 (4) 2023

CONTENTS

sciendo
Economics and Business Review
https://www.ebr.edu.pl
Vol. 9 (4), 2023: 121–152
https://doi.org/10.18559/ebr.2023.4.1075

Is value investing based on scoring models effective? The verification of *F*-Score-based strategy in the Polish stock market

Bartłomiej Pilch[1]

Abstract

The aim of the paper is to analyse the effectiveness of *F*-Score-like models using the example of the Polish stock market. *F*-Score is a scoring model based on a high B/M investing strategy, which uses fundamental signals to assess the economic condition of an entity. So far, its effectiveness has been generally proven in numerous stock markets worldwide. However, no comprehensive study focusing on the Polish market has been conducted. Therefore, *F*-Score and similar models (*FS*-Score and PiotroskiTrfm) were analysed in this regard. It was shown that companies with higher scores generated positive both raw and market-adjusted returns on average. However, they were lower than the mean returns of low-score companies (for *FS*-Score) or total high B/M portfolio (regarding *F*-Score and PiotroskiTrfm). The results of the study show that *F*-Score, *FS*-Score and PiotroskiTrfm are generally effective investing tools. However, it might be more advisable for value investors to choose a total high B/M portfolio instead of shares of high-score entities according to *F*-Score or PiotroskiTrfm.

Keywords

- *F*-Score
- high B/M
- investment strategy
- value investing

JEL codes: G11, M20, M41

Article received 13 November 2022, accepted 30 October 2023.

Suggested citation: Pilch, B. (2023). Is value investing based on scoring models effective? The verification of *F*-Score-based strategy in the Polish stock market. *Economics and Business Review*, *9*(4), 121–152. https://doi.org/10.18559/ebr.2023.4.1075

[1] Cracow University of Economics, Institute of Finance, ul. Rakowicka 27, 31-510 Kraków, Poland, bpilch98@gmail.com, https://orcid.org/0000-0002-0081-4050.

Introduction

The investing process should focus on identifying entities whose intrinsic value exceeds the market price at a given moment (Graham & Dodd, 1934). This was the conclusion reached by these authors, who are widely recognised as among the most significant figures connected with fundamental analysis. It characterises the value investing approach. In line with this conclusion is the idea of investing in shares in undervalued entities, which are often measured by using a B/M ratio (book value to market value; entities whose B/M is below 1 are considered undervalued).

A high B/M strategy is the foundation of the *F*-Score model developed by Piotroski (2000). He proposed a model that consists of nine fundamental signs, which is used to select entities with strong economic foundations, based on scoring, and build an investment portfolio from shares in them. Such a portfolio should generate returns that outperform the market. After the publication of the *F*-Score model, a few modifications of this construct were also made by other authors. Their aim was to improve the initial model or build a new model based on the example of other stock markets. So far, many analyses of *F*-Score's effectiveness have been conducted using examples from European countries and other emerging markets. However, as yet there is no comprehensive analysis based on the specificity of the Polish stock market, apart from research conducted on small samples of the largest listed entities.

The aim of the paper is to analyse the effectiveness of the *F*-Score and similar models using the example of the Polish stock market. The main research hypothesis is that the *F*-Score model is effective and companies with higher scoring outperform both low-score entities and all high B/M companies. The supportive hypothesis is analogous for *FS*-Score and PiotroskiTrfm.

The structure of the paper is as follows: Section 1 includes the overview of the high B/M investing strategy, *F*-Score model and its modifications (*G*-Score, *FS*-Score, PiotroskiTrfm), while Section 2 presents a literature review of the research focused on the assessment of the *F*-Score effectiveness. Section 3 includes a description of the methods used, Section 4—empirical research, and last Section concludes.

1. *F*-Score model and its modifications

1.1. High B/M strategy

A high B/M strategy fits into the framework of value investing. According to Chan et al. (1991), stock returns generated from the portfolio of entities with high B/M values outperformed other portfolios. This was also supported by Fama and French (1992), in whose view it is prudent to invest in a portfolio consisting of undervalued shares. The authors stated that the companies with high B/M might be generally treated as financially distressed, which causes a lack of interest in their shares among investors. Lakonishok et al. (1994) argued that B/M values relate to the behavioural aspect of investors. It refers to their tendency to be over-pessimistic when evaluating entities affected by temporary financial problems. Hence, investors are not willing to invest in shares of high B/M entities. As a result, these become undervalued. On the other hand, financial surprises, quite common among companies in poor financial condition, are most likely to be avoided by high B/M entities (La Porta et al., 1997).

Although the positive association between B/M and rates of return was generally accepted, the relationships between these variables were continuously verified. The positive correlation between the B/M factor and future stock returns was empirically confirmed, e.g., by Auret and Sinclaire (2006), Hasan et al. (2015), da Cunha Araújo and Veras Machado (2018), and Fahreza and Rizkianto (2021).

An analysis conducted by Auret and Sinclaire (2006) used the example of entities listed on the Johannesburg Stock Exchange. The results showed that the next month's return on the shares was positively correlated with the B/M factor. However, this correlation was noticeably weaker after the inclusion of other explanatory variables (cash flow to price, dividend yield, and price-to--net asset value).

Hasan et al. (2015) conducted their analysis using the data of selected companies listed on the Karachi Stock Exchange. They concluded that B/M was the most significantly connected with stock returns out of the all variables examined (the authors also included debt-to-equity, firm size, and sales-to--price as exogenous variables). However, the degree of correlation between B/M and stock returns was moderate.

The study by da Cunha Araújo and Veras Machado (2018) was conducted using the example of Brazilian listed companies. B/M combined with relative earnings (measured by return on equity) was found to be positively linked with future stock returns. Such an association was also maintained after the inclusion of controlling variables regarding firm size and liquidity.

Fahreza and Rizkianto (2021) focused on the companies listed on the Indonesia Stock Exchange. Their results showed that high B/M values were positively connected with higher future stock returns. Such an association applied to both value-weighted and equally weighted portfolios built by the authors.

Generally, the above-mentioned research proved the effectiveness of a high B/M investing strategy. These studies concerned emerging stock markets, like the Polish one, and based on them, it could be stated that it is sensible to invest in undervalued companies. With this regard, the key issue arises, namely, how to select the entities with good financial condition within this group?

1.2. *F*-Score

The main concept of the research procedure proposed by Piotroski was to use accounting-based variables to measure the future financial condition of high B/M companies (Piotroski, 2000). Such entities were usually recognised as financially distressed ones (Fama & French, 1995). Therefore, in Piotroski's view, the best way to provide insight into the future economic situation was to use accounting indicators (Piotroski, 2000). The author argues that: (1) high B/M companies with high profitability present an ability to generate financial surpluses, (2) increasing financial leverage or/and decreasing liquidity might be treated as a sign of a growing risk to a company, (3) operating effectiveness indicators relate to the changes in two factors affecting total profitability—sales volume and relative margin (Piotroski, 2000). Hence, nine variables were chosen to assess the entity three times, taking into account its profitability, financial leverage, liquidity and funding sources, and operating effectiveness. Variables that constitute *F*-Score are presented in Table 1.

The form of the *F*-Score model is a sum of binary values (0 or 1) for a given variable. ΔACCRUAL and ΔLEVER are destimulants (their negative values are recognised as 1 point), while the other variables are stimulants. Entities with a score of 8 or 9 points were selected for the investment portfolio as high-score entities. For low-score companies, Piotroski (2000) postulated short selling.

The construction of the investing portfolio was in line with the buy-and-hold strategy, which led to the generation of significantly different rates of return between high- and low-*F*-Score entities. The mean annual raw yields generated by the investment portfolio (high B/M entities with 8-9 points scoring) amounted to 31.3%, with 7.8% for low *F*-Score companies. However, it is worth noting that absolute rates of return are not sufficiently objective: in periods of a bull market, generating positive yields is more likely than during a slump. Therefore, more meaningful results are provided by comparing the returns generated with market returns or calculating market-adjusted returns. Such adjusted returns were negative in the case of low *F*-Score entities

Table 1. Variables constituting *F*-Score

Variable	Area	Variable type*
$\text{ROA} = \dfrac{\text{Net income}_t - \text{extraordinary operations balance}_t}{\text{Total assets at the beginning of a period}_t}$	profit-ability	S
$\text{CFO} = \dfrac{\text{Net operating cash flows}_t}{\text{Total assets at the beginning of a period}_t}$		S
$\Delta\text{ROA} = \text{ROA}_t - \text{ROA}_{t-1}$		S
$\Delta\text{ACCRUAL} = (\text{ROA} - \text{CFO})_t - (\text{ROA} - \text{CFO})_{t-1}$		D
$\Delta\text{LEVER} = \left(\dfrac{\text{Total liabilities}}{\text{Average total assets}}\right)_t - \left(\dfrac{\text{Total liabilities}}{\text{Average total assets}}\right)_{t-1}$	financial leverage, liquidity, source of funds	D
$\Delta\text{LIQUID} = \left(\dfrac{\text{Current assets}}{\text{Current liabilities}}\right)_t - \left(\dfrac{\text{Current assets}}{\text{Current liabilities}}\right)_{t-1}$		S
$\text{EQ\_OFFER} = \begin{cases} 1 \text{ for no issue of ordinary shares in a given period} \\ 0 \text{ otherwise} \end{cases}$		S
$\Delta\text{MARGIN} = \left(\dfrac{\text{Gross margin on sales}}{\text{Sales revenues}}\right)_t - \left(\dfrac{\text{Gross margin on sales}}{\text{Sales revenues}}\right)_{t-1}$	operating effective-ness	S
$\Delta\text{TURN} = \left(\dfrac{\text{Sales revenues}}{\text{Total assets at the beginning of a period}}\right)_t - \left(\dfrac{\text{Sales revenues}}{\text{Total assets at the beginning of a period}}\right)_{t-1}$		S

* In the case of stimulants, the positive value of a given variable is recognised as 1 point.

Notes: S – stimulant, D – destimulant.

Source: based on (Piotroski, 2000).

(–9.6%) and significantly positive regarding high B/M companies—at the level of 13.4% (Piotroski, 2000). These results explicitly indicated the benefits of using the *F*-Score model.

Piotroski (2000) provided empirical confirmation of his model, which was also based on the two subsamples – entities with a score below 5 points (‘Weak *F*-Score’) and 5 or more points (‘Strong *F*-Score’). Only in two years (1976 and 1994) out of 21 analysed did mean yields generated by weak *F*-Score entities outperform the rates of return of strong *F*-Score companies. The average difference between the two groups of entities analysed was 9.7 p.p. (arithmetic mean) or 9.3 p.p. (the average weighted by the number of observations in a given year) in favour of strong *F*-Score companies. These results could be treated as an initial empirical confirmation of the strategy proposed by Piotroski.

Despite the potential usefulness of using the *F*-Score strategy, it was not a popular model for several years after its development. However, its recognition significantly increased during the financial crisis that started in 2007. This model led to the generation of an average 32.6% yield. This was the best result from the strategies analysed by the American Association of Individual Investors (Comparic, 2017) and affected the model's popularity in subsequent years. As a result, several modifications to the *F*-Score were developed. These are models that took into account different sets of variables and markets developed in recent years and include G-Score, *FS*-Score and PiotroskiTrfm.

1.3. *G*-Score

G-Score is a model developed by Mohanram (2005). Like Piotroski, Mohanram divided the indicators used to construct the model into three subgroups. These were signals referring to earnings and cash flow profitability, naive extrapolation, and accounting conservatism (Mohanram partially included behavioural factors in the model as well). The variables related to financial streams are among the main measures of the economic effectiveness of business management. However, the inclusion of variables related to the two areas listed next may come as a surprise. The motivation for including such variables was that stock markets make a naive extrapolation of the current fundamental values of growth companies (whose business specificity is, after all, focused on maintaining a significant positive growth rate) (La Porta, 1996). Moreover, the valuation of this type of entity should take into account variables that are not subject to reporting under conservative accounting. According to Trueman et al. (2000), these are:

a) non-financial factors such as the number of users (especially in the case of Internet companies),
b) public interest in the entity,
c) the effectiveness of its marketing activities.

It is worth noting that Mohanram did not include the absolute values of the variables in his model. He relates them to the medians of observable values in a sample of companies from the same industry. This approach differs from the one adopted by Piotroski. In addition, the G-Score model was developed to construct portfolios of entities with low B/M, which is opposite to the original strategy that the *F*-Score was based on (Mohanram, 2005). The variables of the *G*-Score model are presented in Table 2.

Table 2. *G*-Score exogenous variables

Variable	Area	Variable type
$G1 = ROA_{i_t} - ROA_t^*$	profitability	stimulant
$G2 = CFO_{i_t} - CFO_t^*$		stimulant
$G3 = CFO_{i_t} - ROA_{i_t}$		stimulant
$G4 = ERN\_VAR_{i_t} - ERN\_VAR_t^*$	naive extrapolation	destimulant
$G5 = SAL\_GR_{i_t} - SAL\_GR_t^*$		destimulant
$G6 = \left(\frac{R\&D}{A}\right)_{i_t} - \left(\frac{R\&D}{A}\right)_t^*$	accounting conservatism	stimulant
$G7 = \left(\frac{CAPEX}{A}\right)_{i_t} - \left(\frac{CAPEX}{A}\right)_t^*$		stimulant
$G8 = \left(\frac{ADV}{A}\right)_{i_t} - \left(\frac{ADV}{A}\right)_t^*$		stimulant

Notes: *i* – variable value for a given and entity, * – median value for entities from one industry, ERN_VAR – earnings variability, SAL_GR – sales growth variability, R&D – research and development expenses, CAPEX – capital expenditures, ADV – advertising intensity.

Source: based on (Mohanram, 2005).

Low *G*-Score values were set as 0–1, while high values as 6–8 points. Based on back-testing, the construct was found to be more effective than the market. Adjusted annual yields accounted for 2.4% among a set of high-scoring entities and –16.4% for low *G*-Score companies. According to the author, the overall effectiveness of the model was particularly due to an accurate assessment of which companies to avoid. As a result, these entities should not be included in the investment portfolio. Short selling might be also applied to them (Mohanram, 2005).

1.4. *FS*-Score

The *FS*-Score model, developed by Gray, presents greater similarity to the original *F*-Score model than Mohanram's construct. Like the previously mentioned models, the *FS*-Score contains indexes assigned to three groups. In this case, they are current profitability, financial stability, and recent operational improvements. The variables from which the *FS*-Score is built are presented in Table 3.

Table 3. *FS*-Score exogenous variables

Variable	Area	Variable type
$\text{ROA} = \dfrac{\text{Net income}_t - \text{extraordinary operations balance}_t}{\text{Total assets at the beginning of a period}_t}$	current profitability	stimulant
$\text{FCFTA} = \dfrac{\text{Free cash flow}_t}{\text{Total assets at the beginning of a period}_t}$		stimulant
$\text{ACCRUAL} = \text{ROA}_t - \text{FCFTA}_t$		destimulant
$\Delta\text{LEVER} = \Delta\left(\dfrac{\text{Long} - \text{term liabilities}}{\text{Total assets}}\right)$	financial stability	destimulant
$\Delta\text{LIQUID} = \Delta\left(\dfrac{\text{Current assets}}{\text{Current liabilities}}\right)$		stimulant
NEQUISS = Number of own shares repurchased$_t$ – number of shares issued$_t$		stimulant
$\Delta\text{ROA} = \text{ROA}_t - \text{ROA}_{t-1}$	operational improvements	stimulant
$\Delta\text{FCFTA} = \text{FCFTA}_t - \text{FCFTA}_{t-1}$		stimulant
$\Delta\text{MARGIN} = \Delta\left(\dfrac{\text{Gross margin on sales}}{\text{Sales revenues}}\right)$		stimulant
$\Delta\text{TURN} = \Delta\left(\dfrac{\text{Sales revenues}}{\text{Total assets at the beginning of a period}}\right)$		stimulant

Notes: Δ – difference between the value of a given variable in the t period and its value in the t–1 period.

Source: based on (Gray, 2015).

The main indices modified by Gray compared to the *F*-Score model are variables relating to the cash flow (which also caused the different construction of the ACCRUAL variable) and stock issues. This author proposed also the inclusion of an index characterising the difference in the value of free cash flow (ΔFCFTA), while there was no similar variable (i.e. ΔCFO) in the *F*-Score model. As a result, there were three variables concerning current profitability, three related to financial stability, and four regarding operational improvements.

The comparison of the effectiveness of the *F*-Score and *FS*-Score models prepared by Gray was based on a sample from 1974–2014. The verification of the effectiveness of both models showed their superiority over the benchmark, which was the S&P500 index. Although both constructs analysed allowed higher returns than the market index, the use of the *FS*-Score led to better results. However, it is worth noting that, in general, the observable differences between returns were not particularly large. The average annual return on the S&P500 index during the period under review was 11.2%, on the investment portfolio built in accordance with the *F*-Score and *FS*-Score strategy—12.6% and 13.3%, respectively (Gray, 2015). Nevertheless, the use of both models seemed more expedient than investing passively in the S&P500 index.

In line with the previous findings were those of Mehta et al. (2019). According to their research, the *FS*-Score as well as *F*-Score models proved to be more effective than the market. Since the sample period (2006-2015) included the global financial crisis, the performance of the strategies tested could be considered robust.

1.5. PiotroskiTrfm

The analyses conducted by Piotroski (2000), Mohanram (2005), and Gray (2015) focused on the U.S. stock market. However, the empirical verification of the effectiveness of the *F*-Score was also prepared based on the example of the South African stock exchange (different from the U.S. in both geographical and stock market development dimensions). Such research was carried out by Nast, who examined the effectiveness of the various variables that make up Piotroski's model. According to his analysis, only 6 of the 9 indices proved to be statistically significant in discriminating companies with above-market returns from entities that generated unsatisfactory results. Surprisingly, in the case of two of these—ΔLEVER, ΔTURN—the author pointed out the advisability of using reverse scoring to the one proposed by Piotroski (Nast, 2017).

Nast verified the effectiveness of individual *F*-Score variables and constructed a model consisting of its selected components. Finally, he included six original variables (using reverse scoring for two of them: ΔLEVER and ΔTURN), discarding the others (ΔROA, ΔLIQUID, ΔMARGIN). The model built in this way (PiotroskiTrfm) was quite effective. Investing in shares of high-score companies brought average annual returns of 22%. On the other hand, the shares of entities with weaker economic situations generated on average a capital gain of less than 6% (Nast, 2017).

Based on Nast's research, it could be stated that the specificity of a given stock market plays an important role in the construction of *F*-Score-type models. A set of variables that accurately characterises the economic situation of companies in one market will not necessarily prove effective for another. This is another argument for conducting research focused on the Polish stock market.

2. Assessment of *F*-Score model effectiveness—literature review

Taking into account the aim of the research, a literature review focused on the effectiveness of the *F*-Score model was prepared. This focused on various markets in different time frames to ensure data comparability. The results are presented in Table 4.

Table 4. Examples of research focused on assessing *F*-Score effectiveness

Source	Market	Analysed period	Results/conclusions
Almas and Duque (2008, p. 25–26)	Belgium, French, Netherlands, Ireland, Portugal	1993–2003	among the three investment strategies analysed, only *F*-Score proved to be effective. Returns generated with its use outperformed the market by an average of 9.2 p.p.
Noma (2010)	Japan	1986–2001	investing in line with the *F*-Score strategy led to the construction of a portfolio that generated a return of 7.8 p.p. above the market average, generating an annual return of 17.6%
Rathjens and Shellhove (2011, p. 26, 58–59)	United Kingdom	1991–2008	entities with high *F*-Score ratings generated returns higher than those of low-rated entities (by an average of 11.7 p.p.), and from other companies (by an average of 4 p.p.)
Mohr (2012)	United States of America	1976–1996	a strategy based on buying the shares of companies with high *F*-Score and selling the securities of entities low-rated by the model yielded returns higher than the market
Hyde (2013)	25 countries (emerging markets)	2000–2011	a study based on the example of emerging markets empirically confirmed the effectiveness of using an *F*-Score-based investment strategy
Singh and Kaur (2015)	India	1996–2010	the returns generated in accordance with the *F*-Score strategy were 18.4 p.p. higher than the values achievable by investing in equity securities of companies with high B/M values

Source	Market	Analysed period	Results/conclusions
Krauss et al. (2015)	United States of America	1976–1996	investment strategies based on the *F*-Score, both on a monthly and weekly basis, led to the generation of excess returns
Safdar (2016)	United States of America	1973–2015	the *F*-Score-based strategy is generally effective, but the degree of its effectiveness varies between industries
Hyde (2016)	Australia	1992–2013	the investment strategy based on the *F*-Score has proven to be effective only for the segment of smaller listed companies. Significant sensitivity of its effectiveness due to the selection of a given sample and time frame was also found
Oyebode (2016)	South Africa	2005–2014	the investment portfolio built from the shares of companies with low *F*-Score ratings generated returns below the market by 6.6 p.p. on average. Thus, the short-selling strategy based on the values of the model analysed proved to be purposeful
Banerjee and Deb (2017)	India	2003–2013	investment portfolios composed of stocks of companies with high *F*-Score ratings generated significantly higher returns than collections of low-scoring shares
Tripathy and Pani (2017)	India	2009–2015	higher-rated companies yielded rates of return significantly higher than those achieved by entities rated low by the *F*-Score.
Sareewiwatthana and Janin (2017)	Thailand	2002–2016	all of the investment strategies analysed (including the *F*-Score) generated returns significantly exceeding the market yields
Turtle and Wang (2017)	United States of America	1972–2012	entities with high *F*-Score values generated significantly higher returns than companies with low ratings
Bülow (2017)	United States of America	2003–2015	buying stocks with a high *F*-Score rating allowed the generation of a return of more than 6 p.p. above the market. A portfolio built from equities with low *F*-Score values yielded a return of 4%

Source	Market	Analysed period	Results/conclusions
Eremenko (2017)	Brazil, China, India, Germany, Russia, United Kingdom	2013–2015	most of the high-rated entities generated returns significantly above the market values (by 8.2 p.p. on average)
Pätäri et al. (2018)	Germany	2000–2015	the inclusion of the *F*-Score in the investment strategy used results in a significant increase in the rate of return achieved
Lalwani and Chakraborty (2018)	India	2001–2015	the strategy based on Piotroski's concept led to an average annual return exceeding the market in the 2001–2005 and 2011–2015 periods. The 2006–2010 period, however, yielded returns slightly lower than the benchmark
Tikkanen and Äijö (2018)	16 European countries	1992–2014	the selection of companies using the *F*-Score model positively affected the returns achieved using multiplier-based investment strategies
Walkshäusl (2020)	35 countries, including 20 developed and 15 emerging markets	2000–2018	returns earned by companies with high *F*-Score ratings exceeded the values generated from securities of entities deemed by the model to be in poor financial health by ca. 10 p.p. per year
Pilch (2021)	Poland	2017–2019	high-scored entities generated higher returns than other public companies from the IT and video game industries
Brindelid and Nilsson (2021)	Denmark, Finland, Norway, Sweden	2012–2021	a comparison of investment strategies based on the *F*-Score and „The Magic Formula" showed an advantage for Piotroski's strategy
Kusowska (2021)	Poland	2014–2020	the research on a sample of the largest Polish entities empirically confirmed the *F*-Score effectiveness

Source: based on the literature review.

It is quite astounding that the results of most of the research presented in Table 4 are generally uniform. They lead to the findings regarding the confirmation of the model's effectiveness. On the other hand, there were also some analyses that pointed out one of the weaknesses of the *F*-Score model—its sensitivity to the time frame adopted (Hyde, 2016; Lalwani & Chakraborty, 2018). Therefore, the results that were obtained from the period 2000–2011 will not necessarily be reflected on the basis of the more recent analysis.

Regarding the Polish stock market, two analyses conducted in recent years were identified in the literature (Kusowska, 2021; Pilch, 2021). However, these concern only a few dozen listed companies. This research provides a more comprehensive empirical analysis based on the sample (total of 225 entities with the highest B/M out of 672 entities analysed, as indicated in section 3) and period (2012–2022) taken into account as well as the use of other *F*-Score-type models.

3. Methods description

3.1. Research design

The empirical part of this paper focuses on assessing whether the models analysed are effective and whether entities with higher scoring outperform companies assessed low (and the market) in terms of returns generated. *F*-Score, *FS*-Score, and PiotroskiTrfm were analysed in this regard. G-Score was excluded from the research due to the lack of sufficient data (especially regarding the estimation of G6 and G8 variables). Based on the financial data, scoring for each entity was computed and returns generated by these companies (in the following period, as described below) were analysed.

Entities that achieved 0 or 1 point (for each model) were considered as having low scores, while a high score was 8–9 points for the *F*-Score, 6 points for PiotroskiTrfm, and 9–10 points for the *FS*-Score. In the following part of the paper, "low" is understood as low-score entities and "high" as high-score companies. Descriptive statistics for variables constituting the *F*-Score and *FS*-Score were analysed, as well as returns generated by entities with each scoring. Statistical tests for the significance of differences regarding means were also employed. Finally, investment portfolios comprising shares in high- and low-score entities (separately) were built (under the assumption of simple diversification).

Among the returns, raw and market-adjusted ones were analysed. The formulas for calculating them are as follows:

$$RR = \frac{P_t}{P_{t-1}} - 1 \tag{1}$$

$$MAR = RR - \frac{P\_WIG_t}{P\_WIG_{t-1}} - 1 \tag{2}$$

where:

RR – raw return,

MAR – market-adjusted return,

P – share price (at closing),

t – last working day in June,

t–1 – first working day in July the previous year,

P_WIG – level of WIG (Warszawski Indeks Giełdowy) index (at closing).

The financial data for a given year were used to explain returns from a year-long period starting six months after the closing date. For instance, scoring was based on the financial data from 2020, while respective returns were for the period 1.07.2021–30.06.2022. The aim of such a time shift was to ensure that the financial data were already known by investors. Hence, the period under analysis is 2011–2020 regarding the financial data used. Returns concerned the period from 2.07.2012 (the first working day of this month) to 30.06.2022.[2]

The financial data as well as share prices were obtained from the Orbis BvD Info database. Due to the data limitations, the following simplifications were implemented:

- Extraordinary operations balance was equal to extraordinary and other profit/loss (regarding ROA calculation),
- Free cash flow was estimated as operating income multiplied by 0.81 (as the Polish base CIT rate is 19%) plus depreciation and amortisation less change in net working capital (measured as inventory plus trade receivables less trade payables) and additions to fixed assets as CAPEX-related variable (regarding FCFTA calculation).

The outliers (the highest 2% and the lowest 2% values of each continuous explanatory variable analysed) were also removed from the sample.

3.2. Sample

The initial sample covered by the study comprised companies from several industries. Entities with core business activity connected with financial services (K – Financial and insurance activities under NACE Rev. 2 classification) were excluded, especially due to the different layouts of their financial statements. It is worth noting that the final sample (which is the focus of the following part of the study) refers to the set of 20% of entities with the highest B/M. They were considered undervalued, in line with Piotroski's approach.

[2] On the following pages, '2022' regarding stock returns means the period between 1.07.2021 and 30.06.2022 (analogous for other years).

The breakdown of entities constituting both the initial and final sample by industry is shown in Table 5.

Table 5. Sample – breakdown by industry

Industry (according to NACE Rev. 2 classification)	Initial sample	Final sample
C – Manufacturing	232	64
G – Wholesale and retail trade; repair of motor vehicles and motorcycles	100	35
J – Information and communication	86	21
M – Professional, scientific and technical activities	69	22
F – Construction	43	17
L – Real estate activities	24	16
D – Electricity, gas, steam and air conditioning supply	20	12
N – Administrative and support service activities	18	7
Q – Human health and social work activities	16	2
H – Transportation and storage	13	5
O – Public administration and defence; compulsory social security	12	5
A – Agriculture, forestry and fishing	10	5
S – Other service activities	10	6
Other industries	19	8
Total	**672**	**225**

Source: own work.

Generally, the share of companies from the top 5 industries amounted to ca. 79% of the total entities analysed. Moreover, manufacturing companies accounted for 34.5% of the initial sample. The shares of companies from the C, G, J, M and F industries in the total number of entities analysed did not significantly (based on the *t*-tests) differ between the initial and final samples. Therefore, it seems that sectoral differentiation did not play a significant role in this research.

3.3. Book-to-market values

Entities constituting the final sample were chosen based on the B/M values. Companies with the 20% highest values of this variable were selected, in line with Piotroski's (2000) approach. The number of companies and entities with the highest B/M values is presented in Figure 1.

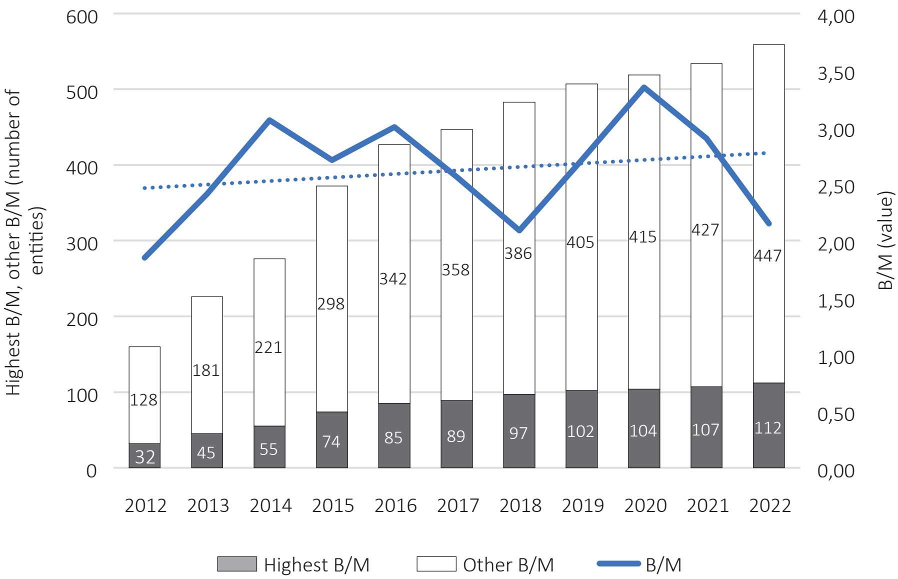

Figure 1. Number of entities with the highest B/M

Source: own work.

As can be indicated on the basis of Figure 1, the number of entities included in the research grew substantially year on year over the period analysed. The value of B/M that differentiated high and other B/M companies was between 1.85 and 3.35. Changes in this value might be explained to some extent by the post-crisis recovery, the period of shortening of the monetary policy and increasing globalisation, as well as the COVID pandemic. Please note that there are a total of 225 companies in the final sample and 902 items of firm-observation data, e.g. based in the Figure 1. These numbers differ, as some of the companies were considered undervalued several times in different years.

4. Verification of the effectiveness of *F*-Score-like model

4.1. Distribution of models scoring

The maximum scoring of models analysed is different. The distribution of the number of entities with each scoring is also significantly differentiated between *F*-Score, PiotroskiTrfm and *FS*-Score. Histograms for these models are presented in Figures 2–4.

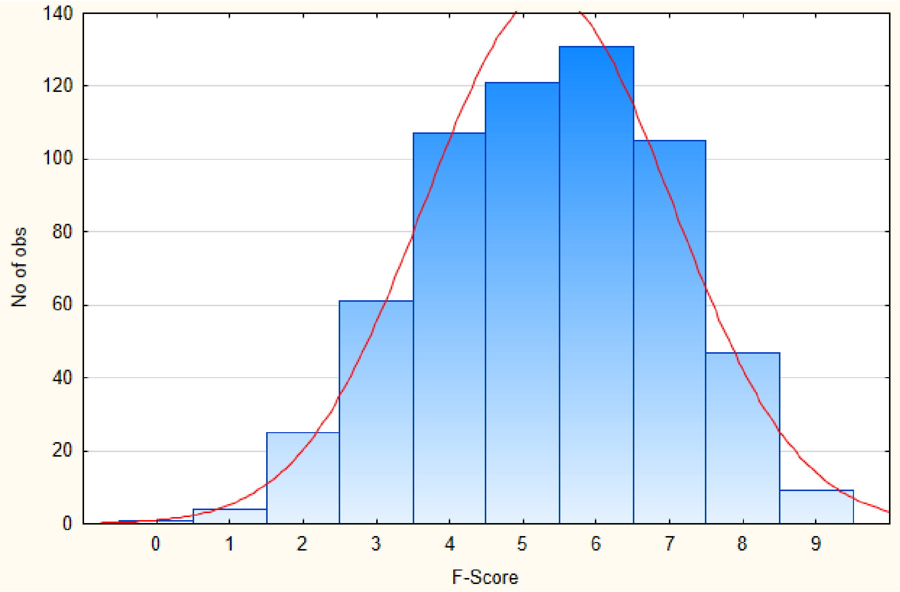

Figure 2. ***F*-Score histogram**

Source: own work using Statistica.

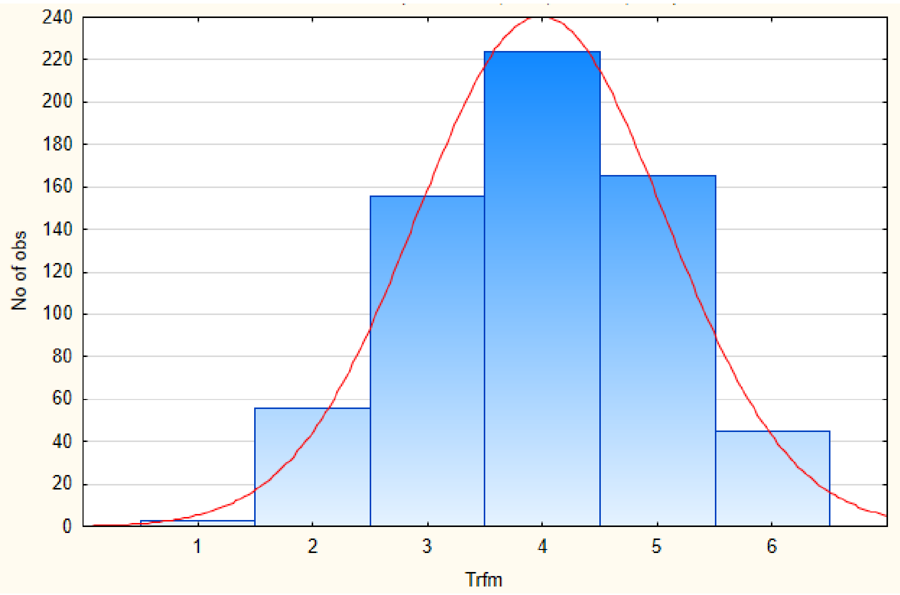

Figure 3. PiotroskiTrfm histogram

Source: own work using Statistica.

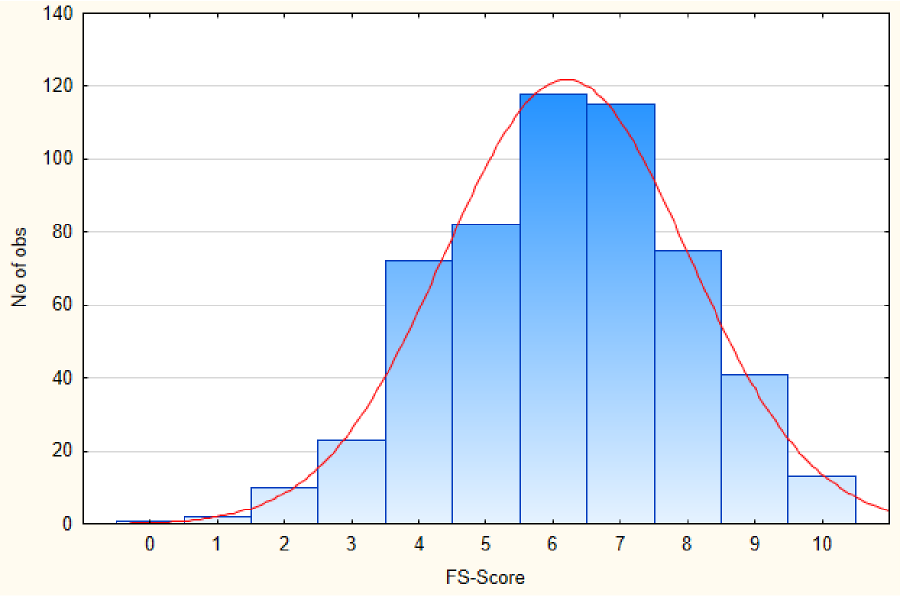

Figure 4. ***FS*-Score histogram**

Source: own work using Statistica.

As can be found based on Figures 2–4, the less frequent values were generally the lowest and the highest, similar to the Gauss distribution. The shares of companies with the two highest possible points under *F*-Score and *FS*-Score and with 6 points for PiotroskiTrfm (entities that were chosen to the investment portfolios) were 9.2%, 9.8%, 6.9%, respectively. For the companies with the lowest scoring (0–1 points), it was 0.8%, 0.5%, 0.5%, respectively. Based on these values, it seems that PiotroskiTrfm is significantly more restrictive in the choice of company. However, *F*-Score and *FS*-Score might benefit from greater portfolio diversification.

4.2. Descriptive statistics

Descriptive statistics for individual variables were further analysed. They concerned exogenous variables (ratios constituting *F*-Score, PiotroskiTrfm, and *FS*-Score) as well as returns. The results of such analysis are presented in Table 6 and Table 7.

No fewer than 25% of companies generated negative values—this applies to all variables except CFO and FCFTA. Most of the high B/M companies (over ¾) generated positive operating cash flows. For ROA, medians and means were positive—these are generally signs of positive average performance

Table 6. Descriptive statistics for exogenous variables included in the research

Variable	Mean (%)	1st quartile (%)	Median (%)	3rd quartile (%)	Standard deviation (%)	Percentage of positive signs (%)
ROA	1.09	–1.12	1.73	4.47	7.77	69.44
ΔROA	–0.40	–3.25	–0.23	2.23	8.44	46.46
CFO	5.08	0.25	4.71	9.66	7.18	76.59
ΔACCRUAL	–0.37	–5.47	–0.40	4.78	11.52	47.41
ΔLEVER	0.29	–2.70	0.31	3.32	6.44	54.00
ΔLIQUID	–4.56	–22.85	–1.05	23.19	148.08	48.13
ΔMARGIN	0.34	–2.82	–0.02	3.22	11.27	49.80
ΔTURN	–3.62	–10.67	–0.29	4.89	22.97	47.93
FCFTA	10.14	2.36	10.14	17.21	12.13	83.40
ΔFCFTA	–1.36	–8.44	–0.76	6.11	14.95	46.77
ACCRUAL	–8.03	–15.11	–8.34	–0.55	13.26	23.12

Source: own work.

Table 7. Descriptive statistics for raw and market-adjusted returns

Returns	Mean (%)	10th percentile (%)	25th percentile (%)	Median (%)	75th percentile (%)	90th percentile (%)	Percentage of positive signs (%)
RR	15.62	–41.57	–21.74	–1.12	28.41	71.88	45.36
MAR	12.40	–48.56	–26.56	–1.73	25.13	76.39	47.54

Source: own work.

of analysed entities. However, ΔROA and ΔFCFTA were mostly negative—on average, ROA and FCFTA declined over time. On the other hand, ACCRUAL was negative for ca. 77% of observations, which is a sign of high conversion of profits into operating cash flows. Such changes are also supported by the ΔACCRUAL variable—its values were negative in most cases, which is a positive sign in Piotroski's (2000) view.

Most of the entities (ca. 54%) increased their indebtedness and both the mean and median for ΔLEVER were positive. Moreover, the absolute value of the 3rd quartile was significantly higher than for the 1st quartile – the scale of debt increase was on average higher for 25% of entities than its decline for the opposite ones. In terms of liquidity, most observations generated a negative change. What is more, there were many entities with a significant drop in liquidity, as the mean for ΔLIQUID is noticeably lower than the median for this variable. The differentiation of changes in liquidity was highest among all the variables that were analysed.

The changes in operating effectiveness, measured by ΔMARGIN and ΔTURN, were mainly negative. Regarding ΔMARGIN, there were similar shares of entities that increased and declined the value of this variable in a given year (± 0.4 p.p.). However, the mean was positive—the scale of the improved gross margin on sales among the entities analysed was higher than the decreasing margin. On the other hand, for ΔTURN, the mean was significantly negative—most of the entities generated a noticeable drop in sales productivity.

Piotroski (2000) stated that high B/M entities are poorly performing on average. The results obtained partially support these findings—profitability ratios were mostly positive. However, they were also declining on average. An average drop in liquidity, increase in debt level, and inconclusive results regarding operating effectiveness were also noted.

The share prices of most of the entities analysed decreased, especially after the inclusion of the market rate of return (from WIG) as a benchmark. However, the means for RR and MAR were at the level of 12%–16%—positive returns generated by the minority of companies were on average higher than negative changes in the share price of other entities (in absolute values). Particularly, there were 10% of companies that achieved raw returns (market-adjusted ones) exceeding 71.8% (76.3%), while for 10% of the companies with the weakest performance, raw returns were lower than 41.5% (48.5%).

A correlation analysis was made for an initial insight into the relationship of the variables analysed to the returns and scoring of the models. It is presented in Table 8. Correlations between exogenous variables are shown in the Appendix.

Most of the correlations between exogenous variables and returns were insignificant. In the case of significant ones, the negative association between ROA and RR is surprising. On the other hand, the expected negative interdependence between ΔLIQUID and MAR was observed. The correlations between FCFTA, ΔFCFTA, ACCRUAL, and *F*-Score as well as PiotroskiTrfm are also with expected signs. It is not an obvious conclusion, as these variables appear only in *FS*-Score out of the three models analysed. A similar situation concerned the CFO and *FS*-Score.

Generally, there are no particularly strong correlations between exogenous variables and returns. Regarding ΔLIQUID, the results obtained support the findings of Nast (2017), but they are opposite to Piotroski's (2000). For ROA, the results are opposite to both authors' conclusions – asset profitability has proven to be negatively associated with future returns from the example of the Polish stock market. It seems that these results might be treated as preli-

Table 8. Correlations between exogenous variables, returns, and scoring of models analysed*

Variable	RR	MAR	*F*-Score	Trfm	*FS*-Score
ROA	–0.0876	–0.0628	–0.0060	**0.2604**	0.0021
ΔROA	–0.0698	–0.0244	–0.0272	0.0167	–0.0220
CFO	–0.0669	–0.0721	**0.1026**	**0.1895**	**0.0855**
ΔACCRUAL	–0.0807	–0.0336	–0.0393	0.0410	–0.0807
ΔLEVER	–0.0009	–0.0109	–0.0102	0.0081	0.0749
ΔLIQUID	–0.0786	**–0.0984**	0.0284	0.0263	–0.0249
ΔMARGIN	–0.0047	–0.0109	–0.0717	–0.0151	–0.0603
ΔTURN	–0.0669	–0.0433	0.0147	**0.1123**	0.0324
FCFTA	0.0267	0.0252	**0.2738**	**0.2595**	**0.4993**
ΔFCFTA	0.0341	0.0289	**0.1595**	0.0214	**0.4554**
ACCRUAL	–0.0773	–0.0614	**–0.2725**	**–0.1031**	**–0.4894**
EQ_OFFER	0.0560	0.0686	**0.1361**	**0.2410**	**0.0928**
NEQUISS	0.0535	0.0674	**0.1514**	**0.2561**	**0.1043**

* Bold values are significant at $p < 0.05$.

Source: own work.

minary evidence of the models' ineffectiveness. However, the analysis of returns by scoring was conducted in the following part. It should provide more credible results.

4.3. Returns by scoring

Assuming that the scoring approach adopted in *F*-Score, *FS*-Score and PiotroskiTrfm is effective, the returns of high-score companies should outperform the rates of return of low-score entities and the total (final) sample. Descriptive statistics for entities with each scoring were calculated. Tables 9–14 present raw and market-adjusted returns with this regard.

There was a minor share of entities with low scores (11–18 times lower than the number of high-score companies for individual models) among the companies analysed. Median market-adjusted returns were non-positive for entities with 0–1, 3, 5–6, and 8 scoring according to *F*-Score. Regarding raw returns, there was also a negative median for entities with scores of 2 and 7. A significantly different situation concerned the means—they were positive for the companies with 2–9 scoring for both raw and market-adjusted returns. The differences between returns from high-score entities and low-score companies were in favour of the first group for all statistics analysed. However, compared to the total (final) sample, high-score entities were less effective,

Table 9. Raw returns by scoring: *F*-Score

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	13.9	–36.7	–20.8	–1.3	27.8	69.3	46.6	611
0	–74.6	–74.6	–74.6	–74.6	–74.6	–74.6	0.0	1
1	–7.9	–45.9	–28.2	–9.7	10.6	31.7	25.0	4
2	5.9	–56.2	–45.7	0.0	33.1	83.7	48.0	25
3	10.3	–38.9	–26.2	0.0	34.0	56.5	49.2	61
4	18.9	–24.9	–15.2	4.0	33.2	83.2	55.1	107
5	21.9	–42.3	–20.5	–3.3	27.1	84.4	47.1	121
6	10.3	–36.7	–23.9	–5.8	19.4	57.7	39.7	131
7	11.7	–28.4	–19.4	–1.9	18.8	89.0	43.8	105
8	8.6	–29.8	–18.2	–0.3	24.7	45.2	46.8	47
9	19.7	–22.9	–15.2	14.3	24.5	52.1	66.7	9
Low score	–21.2	–67.8	–57.7	–18.4	–1.1	27.0	20.0	5
High score	10.4	–29.9	–17.8	0.3	25.2	45.8	50.0	56
High-all	–3.6	6.8	2.9	1.5	–2.6	–23.6	3.4	–
High-low	31.6	38.0	39.8	18.6	26.3	18.8	30.0	–

Notes: Perc. – percentile, % positive – percentage of positive signs, *n* – number of entities, high-all – the difference between high-score companies and total final sample, high-low – the difference between high-score and low-score companies.

Source: own work.

Table 10. Market-adjusted returns by scoring: *F*-Score

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	11.5	–44.4	–23.7	–1.5	25.1	77.6	47.6	611
0	–82.8	–82.8	–82.8	–82.8	–82.8	–82.8	0.0	1
1	–15.3	–48.8	–20.1	–3.5	1.4	8.9	25.0	4
2	11.9	–50.4	–29.6	7.6	28.4	80.6	52.0	25
3	8.3	–42.0	–21.8	–5.7	26.4	52.5	45.9	61
4	14.4	–37.8	–14.2	2.8	26.8	70.1	54.2	107
5	18.7	–51.0	–27.3	–2.4	27.7	84.0	46.3	121
6	5.8	–44.9	–25.4	–6.4	11.4	48.1	38.9	131
7	11.5	–39.8	–20.6	5.4	33.7	73.5	54.3	105
8	6.8	–36.9	–20.5	–4.9	27.3	60.7	44.7	47
9	28.3	–13.9	–0.6	15.1	30.5	67.7	66.7	9
Low score	–28.8	–76.9	–67.9	–4.2	–2.8	7.2	20.0	5
High score	10.2	–35.8	–18.2	–1.2	28.2	61.2	48.2	56
High-all	–1.2	8.6	5.5	0.3	3.0	–16.4	0.6	–
High-low	39.0	41.1	49.7	3.0	31.0	54.0	28.2	–

Source: own work.

Table 11. Raw returns by scoring: PiotroskiTrfm

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	18.4	–38.6	–20.5	–0.5	30.4	76.9	46.7	649
0	–	–	–	–	–	–	–	0
1	–11.2	–16.4	–14.1	–10.2	–7.9	–6.5	0.0	3
2	34.7	–42.4	–21.2	–4.8	43.0	147.7	42.9	56
3	14.3	–42.6	–24.6	–7.1	27.1	72.9	44.9	156
4	20.3	–37.3	–19.9	0.0	33.4	89.1	49.6	224
5	16.8	–32.2	–20.4	–1.3	29.0	59.8	44.2	165
6	11.4	–25.1	–15.3	3.4	20.6	64.0	55.6	45
Low score	–11.2	–16.4	–14.1	–10.2	–7.9	–6.5	0.0	3
High score	11.4	–25.1	–15.3	3.4	20.6	64.0	55.6	45
High-all	–7.0	13.5	5.2	3.9	–9.8	–12.9	8.9	–
High-low	22.6	–8.7	–1.3	13.7	28.5	70.5	55.6	–

Source: own work.

Table 12. Market-adjusted returns by scoring: PiotroskiTrfm

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	15.7	–44.5	–23.6	–0.8	26.4	80.6	48.5	649
0	–	–	–	–	–	–	–	0
1	–1.4	–19.4	–9.3	7.6	10.9	12.9	66.7	3
2	32.3	–54.3	–25.6	0.2	32.1	165.6	50.0	56
3	11.5	–45.0	–26.7	–5.9	24.4	78.6	44.2	156
4	19.1	–40.8	–20.3	0.8	29.7	86.0	51.8	224
5	12.6	–42.9	–20.8	–1.5	22.4	47.7	47.3	165
6	5.5	–47.5	–19.6	–1.1	19.3	53.3	48.9	45
Low score	–1.4	–19.4	–9.3	7.6	10.9	12.9	66.7	3
High score	5.5	–47.5	–19.6	–1.1	19.3	53.3	48.9	45
High-all	–10.2	–3.0	4.0	–0.3	–7.1	–27.3	0.4	–
High-low	6.9	–28.1	–10.3	–8.7	8.4	40.3	–17.8	–

Source: own work.

Table 13. Raw returns by scoring: *FS*-Score

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	11.6	–34.6	–20.6	–1.7	26.9	64.6	46.4	552
0	–74.6	–74.6	–74.6	–74.6	–74.6	–74.6	0.0	1
1	82.2	53.0	64.0	82.2	100.4	111.4	100.0	2
2	–21.3	–58.7	–55.2	–22.5	2.8	14.2	30.0	10
3	6.1	–45.7	–28.8	0.6	40.5	57.6	52.2	23
4	7.0	–37.5	–23.4	–0.1	20.7	50.4	45.8	72
5	26.4	–31.2	–16.6	6.2	45.8	90.4	56.1	82
6	20.2	–34.3	–19.6	–4.9	25.8	103.4	45.8	118
7	–0.2	–34.9	–24.2	–6.3	11.8	43.1	37.4	115
8	7.3	–30.6	–17.6	–4.9	13.8	56.9	41.3	75
9	18.3	–30.0	–14.6	14.3	35.0	91.9	61.0	41
9	4.1	–22.0	–15.5	0.5	16.8	39.7	53.8	13
Low score	29.9	–50.5	–14.4	45.7	82.2	104.1	66.7	3
High score	14.9	–29.5	–15.1	10.5	32.3	78.8	59.3	54
High-all	3.3	5.0	5.5	12.1	5.4	14.2	12.9	–
High-low	31.6	38.0	39.8	18.6	26.3	18.8	30.0	–

Source: own work.

Table 14. Market-adjusted returns by scoring: *FS*-Score

Scoring	Mean (%)	10th perc.	25th perc.	Median (%)	75th perc.	90th perc.	% positive	*n*
All entities	8.5	–44.4	–23.9	–1.9	23.3	65.1	47.5	552
0	–82.8	–82.8	–82.8	–82.8	–82.8	–82.8	0.0	1
1	64.7	24.1	39.3	64.7	90.2	105.4	100.0	2
2	–25.2	–68.3	–54.6	–20.1	–0.3	9.2	30.0	10
3	1.5	–54.1	–28.9	2.8	27.7	56.0	52.2	23
4	2.7	–50.4	–23.5	–1.4	17.7	39.7	47.2	72
5	17.4	–46.7	–20.9	2.8	28.2	79.2	52.4	82
6	18.9	–39.5	–22.1	–3.4	28.7	90.4	44.1	118
7	–0.7	–46.8	–24.2	–5.3	17.7	37.3	44.3	115
8	4.9	–38.7	–25.7	–4.9	19.8	66.4	44.0	75
9	16.1	–37.6	–13.9	4.0	36.6	71.6	58.5	41
9	4.7	–29.8	–19.6	4.6	30.5	36.1	61.5	13
Low score	15.6	–63.5	–34.5	13.9	64.7	95.3	66.7	3
High score	13.4	–33.7	–16.6	4.2	34.7	62.8	59.3	54
High-all	4.9	10.7	7.2	6.1	11.5	–2.3	11.8	–
High-low	–2.2	29.8	17.8	–9.7	–30.0	–32.4	–7.4	–

Source: own work.

taking into account means and 90th percentiles (for raw returns, it was observable also for the 75th percentile).

In the case of PiotroskiTrfm, median raw returns were negative or at 0 for entities with 0–5 scoring. On the other hand, the positive median in market-adjusted returns concerned companies with scores of 1–2 and 4, while for entities with a score of 6, the median was negative. Mean returns (both raw and market-adjusted) were positive for high-score companies, but they were outperformed by the results for the total sample. On the other hand, low-score entities performed worst.

Regarding *FS*-Score, medians differed significantly between entities with different scoring. They were positive for high-score entities. However, medians for low-score companies were also positive and even higher than in the case of high-score entities. A similar situation concerned means: the low-score entities generated higher values than high-score companies (however, the difference regarding market-adjusted returns was noticeably lower than for raw returns). On the other hand, high-score entities outperformed the total sample.

Based on the above results, it seems that the scoring approach did not lead to the generation of better returns than for a total portfolio of high B/M companies; this applies to *F*-Score and PiotroskiTrfm. Regarding *FS*-Score, though high-scored entities outperformed the total sample, they were worse than low-score companies. To test mean differences between high and low, as well as high and all, statistical tests were conducted. The results of these are presented in Table 15.

Table 15. Results of the *t*-test for mean differences

Groups	Statistic	*F*-Score		PiotroskiTrfm		*FS*-Score	
		RR	MAR	RR	MAR	RR	MAR
High-all	*t*-statistic	–0.3769	–0.1322	–0.5236	–0.7696	0.3852	0.5768
	p-value	0.7064	0.8949	0.6008	0.4418	0.7022	0.5643
High-low	*t*-statistic	1.4038	1.6271	0.8075	0.2499	–0.5459	–0.0812
	p-value	0.1656	0.1090	0.4236	0.8038	0.5873	0.9356

Source: own work.

No statistically significant results were obtained. In the case of high-low for *F*-Score and PiotroskiTrfm, it could be explained by a very low number of low-score companies (5 and 3, respectively). Therefore, despite the quite high absolute differences between the means for these groups, they were found to be statistically insignificant. Generally, the above results partially confirmed the effectiveness of the models analysed. However, high-score entities performed worse than either low-score companies (it applies to *FS*-Score) or the total sample (*F*-Score, PiotroskiTrfm).

4.4. Building investment portfolios

Since the purpose of the paper is connected with the empirical analysis of investment strategies' performance, separate portfolios consisting of shares of entities with high and low scores were built. The assumption was of equal shares of each entity in the portfolios, i.e. simple diversification. The returns generated from these portfolios are shown in Table 16.

Table 16. Mean returns from investment portfolios built

Model	***F*-Score**			**PiotroskiTrfm**			***FS*-Score**		
Scoring	**High (8–9)**			**High (6)**			**High (9–10)**		
Item	**RR (%)**	**MAR (%)**	***n***	**RR (%)**	**MAR (%)**	***n***	**RR (%)**	**MAR (%)**	***n***
2013	–	–	0	14.54	4.28	1	28.39	18.14	3
2014	6.13	–9.88	2	–7.25	–23.26	4	–	–	0
2015	19.64	16.58	4	5.72	2.66	3	31.99	28.93	5
2016	12.48	28.04	5	25.62	41.17	4	–1.73	13.82	7
2017	38.72	2.38	6	76.06	39.72	5	16.69	–19.66	6
2018	–18.83	–9.97	4	–30.28	–21.42	4	–29.57	–20.71	3
2019	–10.83	–19.03	8	–13.61	–21.81	5	–18.03	–26.23	4
2020	51.62	69.45	9	23.44	41.27	5	21.70	39.53	11
2021	1.21	–30.65	6	13.52	–13.82	9	55.09	23.23	8
2022	–9.59	10.25	12	–4.35	15.49	5	–6.99	12.86	7
2013–2022	*7.12*	*2.66*	*56*	*7.20*	*3.42*	*45*	*7.01*	*4.69*	*54*
Scoring	**Low (0–1)**			**Low (0–1)**			**Low (0–1)**		
Item	**RR (%)**	**MAR (%)**	***n***	**RR (%)**	**MAR (%)**	***n***	**RR (%)**	**MAR (%)**	***n***
2013	–57.68	–67.94	1	–	–	0	–	–	0
2014	–	–	0	–	–	0	–	–	0
2015	–1.12	–4.18	1	–	–	0	118.67	115.61	1
2016	–18.37	–2.81	1	–	–	0	–	–	0
2017	–	–	0	–	–	0	–	–	0
2018	–	–	0	–	–	0	–	–	0
2019	–74.61	–82.81	1	–17.93	–26.13	1	–74.61	–82.81	1
2020	–	–	0	–10.22	7.60	1	–	–	0
2021	45.74	13.88	1	–	–	0	45.74	13.88	1
2022	–	–	0	–5.56	14.28	1	–	–	0
2013–2022	*–18.69*	*–24.72*	*5*	*–3.56*	*–0.96*	*3*	*–2.10*	*–8.27*	*3*

Note: 2013–2022 – geometric mean return for 2013–2022 (including years with no entities chosen by the models).

Source: own work.

A high *F*-Score portfolio led to the generation of positive returns in 5 (for market-adjusted returns) or 6 (regarding raw returns) out of the 10 years analysed, similar to *FS*-Score. Regarding PiotroskiTrfm, it was 6 years for both RR and MAR. There were no high-score entities identified in 2013 based on *F*-Score nor in 2014 according to *FS*-Score.

The average annual raw return for high-score portfolios was between 7.0% and 7.2% for all the models analysed. However, market-adjusted returns differed significantly, but on average was highest for *FS*-Score. Generally, all high-score portfolios generated positive market-adjusted returns.

Low-score portfolios comprised only a few companies – up to 5 in total. Hence, years with no entities selected for the low-score portfolio were quite frequent (5 out of 10 according to *F*-Score and 7 according to *FS*-Score and PiotroskiTrfm). Such portfolios generated significantly negative returns on average (both raw and market-adjusted). However, they were noticeably lower for *F*-Score than for the other two models.

Based on the high-score portfolio construction, all models were found to be effective – they generated positive raw and market-adjusted returns. This supports the findings resulting from the backtesting by Piotroski (2000), Gray (2015), and Nast (2017). Piotroski (2000) and Oyebode (2016) stated that it is worth taking short positions in low-scored entities, which was also empirically confirmed, as the low-score portfolios generated negative returns. However, high-score companies did not outperform the benchmarks (either low-score portfolio or total sample of high B/M companies). This is not consistent with the statistical tests performed by Piotroski (2000). Generally, both the main hypothesis and the supporting one might be considered only partially confirmed.

Conclusions

High B/M investing is a foundation of investment strategies based on *F*-Score and similar models. So far, both high B/M investing and the effectiveness of the *F*-Score model have been empirically confirmed by examples from different markets worldwide. However, regarding the Polish stock market, the only analyses conducted so far focused on relatively small samples and timeframes. This research provides an empirical insight into *F*-Score-based strategy based on the example of a comprehensive sample of Polish listed companies. Other similar models inspired by *F*-Score (i.e. *FS*-Score and PiotroskiTrfm) were also taken into account.

To verify the usefulness of the models analysed, returns by scoring were analysed. Generally, the returns for the entities with higher scoring were positive. However, mean returns for high-score entities were lower than the rates of return for the low-score portfolio (*FS*-Score) or for a total sample of high B/M companies (*F*-Score, PiotroskiTrfm). Nevertheless, statistical tests did not confirm the significance of mean differences.

Investment portfolios consisting of high- and low-score entities were built (separately), with an assumption of simple diversification. The results showed that using the models analysed here to build a high-score portfolio generated positive raw and market-adjusted returns. The low-score-based portfolio generated negative yields. It is also worth noting that low-score portfolios consisted of only a few shares. Therefore, they are not diversified in any way, especially as there were only single shares chosen for the portfolio in individual years.

Overall, the above findings are partially in line with the conclusions produced by Piotroski (2000) and most of the other authors whose results are presented in Table 3, namely, Gray (2015) and Nast (2017), regarding the effectiveness of *F*-Score, *FS*-Score, and PiotroskiTrfm, respectively. Moreover, the findings of Kusowska (2021) and Pilch (2021) regarding the effectiveness of the *F*-Score-based strategy on the Polish stock market were also empirically confirmed. On the other hand, the advantage of high-score companies over the total sample of high B/M entities postulated by Piotroski (2000) was not supported.

The implications of the research mainly concern value investors. The results pointed out the effectiveness of *FS*-Score and *F*-Score-based strategies, but also their weaknesses. Therefore, it seems that it might be sensible not only to invest in line with the models analysed, but also to invest in a total portfolio of high B/M companies. The advantage of this approach is also a greater portfolio diversification than with using *F*-Score-like models.

Among the limitations of the study might be the assumption regarding the simple diversification applied, which was also pointed out by Mehta et. al (2019). Transaction costs were also not included in the research. Another issue relates to the way of verifying the models' effectiveness, i.e. backtesting—it is commonly used but, as it fully focuses on historical financial data, without the inclusion of forward-looking measures. Moreover, the research focused on a relatively short period (mostly resulting from the data availability). On the other hand, it includes periods of both favourable economic situations and economic slumps, and this should be considered an advantage.

The directions for future research are strongly associated with the limitations indicated. They mostly refer to the building of investment portfolios with different shares for individual entities. Moreover, the inclusion of transaction costs might be useful. Investment portfolios consisting of a larger sample of high B/M companies, without reliance on *F*-Score-type scoring, should also be analysed using the example of the Polish stock market.

Appendix

Table A1. Correlations between exogenous variables

Variable	ROA	ROA	CFO	ΔACCRUAL	ΔLEVER	ΔLIQUID	ΔMARGIN
ROA	1.0000	**0.4127**	**0.3882**	**0.2527**	**–0.2502**	0.0646	**0.1295**
ROA		1.0000	0.0307	**0.5971**	**–0.2786**	0.0474	**0.2904**
CFO			1.0000	**–0.3882**	–0.0448	0.0239	0.0106
ΔACCRUAL				1.0000	**–0.1829**	**0.1092**	**0.1794**
ΔLEVER					1.0000	**–0.2953**	–0.0625
ΔLIQUID						1.0000	0.0457
ΔMARGIN							1.0000

Variable	ΔTURN	FCFTA	ΔFCFTA	ACCRUAL	EQ_OFFER	NEQUISS
ROA	0.0687	**0.2687**	**–0.1784**	**0.3119**	0.0511	0.0725
ROA	**0.2886**	0.0531	–0.0608	**0.1857**	0.0466	0.0445
CFO	0.0195	0.3309	–0.1048	**–0.1024**	–0.0051	0.0122
ΔACCRUAL	**0.0975**	0.0187	–0.0098	**0.1274**	0.0684	0.0750
ΔLEVER	**0.1946**	**0.0883**	**0.1764**	**–0.2315**	–0.0110	–0.0190
ΔLIQUID	–0.0813	–0.0568	**–0.1212**	**0.0933**	0.0102	0.0088
ΔMARGIN	–0.0761	0.0825	–0.0029	–0.0067	0.0301	0.0282
ΔTURN	1.0000	0.0296	0.0575	0.0105	0.0331	0.0313
FCFTA		1.0000	**0.5614**	**–0.8313**	0.0267	0.0444
ΔFCFTA			1.0000	**–0.6567**	0.0251	0.0233
ACCRUAL				1.0000	0.0031	–0.0019
EQ_OFFER					1.0000	**0.9803**
NEQUISS						1.0000

Source: own work.

References

Almas, D., & Duque, J. (2008). *Value investing: The book-to-market effect, accounting information, and stock returns*. ADVANCE – Centro de Investigação Avançada em Gestão Working Papers, 1. https://www.repository.utl.pt/handle/10400.5/2558?locale=en

Auret, C. J., & Sinclaire, R. A. (2006). Book-to-market ratio and returns on the JSE. *Investment Analysts Journal*, *35*(63), 31–38. https://doi.org/10.1080/10293523.2006.11082476

Banerjee, P., & Deb, S. (2017). Abnormal returns using accounting information within a value portfolio. *Accounting Research Journal, 30*(1), 73–88. https://doi.org/10.1108/ARJ-01-2015-0003

Brindelid, L., & Nilsson, T. (2021). *Portfolio performance in Nordic countries: A quantitative comparison study of investment strategies in Denmark, Finland, Norway and Sweden.* Postgraduate thesis. Karlstad Business School.

Bülow, S. (2017). *The effectiveness of fundamental analysis on value stocks—an analysis of Piotroski's F-Score*. Bachelor thesis. Lund University.

Chan, L. K. C., Hamao, Y., & Lakonishok, J. (1991). Fundamentals and stock returns in Japan. *The Journal of Finance*, *46*(5), 1739–1764. https://doi.org/10.1111/j.1540-6261.1991.tb04642.x

Comparic. (2017). *Jak z zyskiem przejść krach, czyli strategia F-Score*. from https://comparic.pl/zyskiem-przejsc-krach-czyli-strategia-f-score/

Da Cunha Araújo, R. C., & Veras Machado, M. A. (2018). Book-to-market ratio, return on equity and Brazilian stock returns. *RAUSP Management Journal*, *53*(3), 324–344. https://doi.org/10.1108/RAUSP-04-2018-001

Eremenko, E. (2017). Quantitative fundamentals. Application of Piotroski F-Score on non-US markets*. SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3262154

Fahreza, M., & Rizkianto, E. (2021). Retained earnings factor analysis on return and excess return: A comparison study of retained earnings-to-market equity and book-to-market equity from 2008–2018. *Advances in Economics, Business and Management Research*, *177*, 225–231. https://doi.org/10.2991/aebmr.k.210522.030

Fama, E., & French, K. (1992). The cross-section of expected returns. *The Journal of Finance*, *47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E., & French, K. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, *50*(1), 131–155. https://doi.org/10.1111/j.1540-6261.1995.tb05169.x

Graham, B., & Dodd, D. (1934). *Security analysis*. McGraw Hill.

Gray, W. (2015). *Simple methods to improve the Piotroski F-Score*. Alpha Architect. https://alphaarchitect.com/2015/05/value-investing-research-simple-methods-to-improve-the-piotroski-f-score/

Hasan, S., Farooq, S., & Mudassir, M. (2015). Stock returns indicators: Debt to equity, book to market, firm size and sales to price. *Journal of Poverty, Investment and Development*, *16*, 25–32.

Hyde, C. (2013). An emerging markets analysis of the Piotroski F-Score. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.2274516

Hyde, C. (2016). The Piotroski F-Score: Evidence from Australia. *Accounting & Finance, 58*(2), 423–444. https://doi.org/10.1111/acfi.12216

Krauss, C., Krüger, T., & Beerstecher, D. (2015). *The Piotroski F-Score: A fundamental value strategy revisited from an investor's perspective*. Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung Discussion Papers, 13. https://www.econstor.eu/bitstream/10419/121238/1/837249988.pdf

Kusowska, M. (2021). Assessment of effectiveness of Piotroski F-Score strategy in the Warsaw Stock Exchange. *Research on Enterprise in Modern Economy – Theory and Practice*, *1*(32), 47–59. https://doi.org/10.19253/reme.2021.01.004

La Porta, R. (1996). Expectations and the cross section of stock returns. *The Journal of Finance*, *51*(5), 1715–1742. https://doi.org/10.1111/j.1540-6261.1996.tb05223.x

La Porta, R., Lakonishok, J., Schleifer, A., & Vishny, R. (1997). Good news for value stocks: Further evidence on market effectiveness. *Journal of Finance*, *52*(2), 859–874. https://doi.org/10.2307/2329502

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation and risk. *The Journal of Finance*, *49*(5), 1541–1578. https://doi.org/10.1111/j.1540-6261.1994.tb04772.x

Lalwani, V., & Chakraborty, M. (2018). Quality investing in the Indian stock market. *Managerial Finance, 44*(2), 127–141. https://doi.org/10.1108/MF-07-2017-0248

Mehta, N., Pothula, V., & Bhattacharyya, R. (2019). *A value investment strategy that combines security selection and market timing signals*. WorldQuant University MScFE Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3451859

Mohanram, P. (2005). Separating winners from losers among low book-to-market stocks using financial statement analysis, *Review of Accounting Studies*, *10*, 133–170. https:/doi.org/10.1007/s11142-005-1526-4

Mohr, J. H. M. (2012). *Utility of Piotroski F-Score for predicting growth-stock returns*. Working Paper, MFIE Capital. https://www.quant-investing.com/media/tztosoes/backtest_utility_of_piotroski_f-score_for_predicting_growth-stock_returns_14.pdf

Nast, T. K. (2017). *Transforming Piotroski's (binary) F-score into a real one*. Postgraduate thesis. University of Pretoria.

Noma, M. (2010). Value investing and financial statement analysis. *Hitotsubashi Journal of Commerce and Management*, *44*(1), 29–46. https://doi.org/10.15057/18701

Oyebode, A. (2016). *Application of the Altman Z-EM-Score and Piotroski F-Score to the Johannesburg securities exchange as short selling instrument*. Postgraduate thesis. University of Pretoria.

Pätäri, E.J., Leivo, T. H., Hulkkonen, L., & Honkapuro, J. V. S. (2018). Enhancement of value investing strategies based on financial statement variables: The German evidence. *Review of Quantitative Finance and Accounting, 51*(3), 813–845. https://doi.org/10.1007/s11156-017-0689-y

Pilch, B. (2021). Analiza wskaźnika F-Score dla spółek giełdowych z branż IT i gier wideo. *Annales Universitatis Mariae Curie-Skłodowska. Sectio H – Oeconomia, 55*(1), 41–50. https://doi.org/10.17951/h.2021.55.1.41-50

Piotroski, J. (2000). Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research*, *38*, 1–41. https://doi.org/10.2307/2672906

Rathjens, H., & Schellhove, H. (2011). *Simple financial analysis and abnormal stock returns—analysis of Piotroski's investment strategy*. Master thesis. Stockholm School of Economics.

Safdar, I. (2016). Industry competition and fundamental analysis. *Journal of Accounting Literature, 37*(1), 36–54. https://doi.org/10.1016/j.acclit.2016.09.001

Sareewiwatthana, P. & Janin, P. (2017). Tests of quantitative investing strategies of famous investors: Case of Thailand. *Investment Management and Financial Innovations, 14*(3), 218–226. https://doi.org/10.21511/imfi.14(3-1).2017.06

Singh, J., & Kaur, K. (2015). Adding value to value stocks in Indian stock market: An empirical analysis. *International Journal of Law and Management, 57*(6), 621–636. https://doi.org/10.1108/IJLMA-09-2014-0055

Tikkanen, J., & Äijö, J. (2018). Does the F-score improve the performance of different value investment strategies in Europe? *Journal of Asset Management, 19*(3), 495–506. https://doi.org/10.1057/s41260-018-0098-3

Tripathy, B., & Pani, B. (2017). Effect of F-Score on stock performance: Evidence from Indian equity market. *International Journal of Economics and Finance, 9*(2), 89–99. https://doi.org/10.5539/ijef.v9n2p89

Trueman, B., Wong, F., & Zhang, X. (2000). The eyeballs have it: Searching for the value in Internet. *Journal of Accounting Research*, *38*, 137–162. https://doi.org/10.2307/2672912

Turtle, H. J., & Wang, K. (2017). The value in fundamental accounting information. *The Journal of Financial Research, 40*(1), 113–140. https://doi.org/10.1111/jfir.12119

Walkshäusl, C. (2020). Piotroski's F-Score: International evidence, *Journal of Asset Management, 21*(2), 106–118. https://doi.org/10.1057/s41260-020-00157-2