Accounting and Management Information Systems
Vol. 21, No. 3, pp. 355-372, 2022
DOI: http://dx.doi.org/10.24818/jamis.2022.03003

Can financial strength indicators form a profitable investment strategy? The case of F-Score in Europe

Andreas G. Koutoupis [a], Christos G. Kampouris [1,b] and Athanasia V. Sakellaridou [c]

[a] *Department of Accounting and Finance, University of Thessaly, Greece*
[b] *Department of Tourism Studies, University of Piraeus, Greece*
[c] *Department of Social Science-MBA, Hellanic Open University, Greece*

Abstract

Research Question: Can the F-Score predict the stock market returns in the cross section of international stock markets?

Motivation: The majority of the literature, in the area of the F-Score metric, has examined whether it can be used to predict future financial profitability, the relationship of F-Score with book-to-value metrics and the momentum premium and whether it can be used as a successful investment strategy tool. There only three studies that examine the relationship between the F-Score and future stock returns, without the use of complementary variables, and in other countries except Europe. This paper seeks to fill this gap.

Data: The dataset of the present research consists of listed European companies from 21 countries (in random order: Finland, United Kingdom, Switzerland, Turkey, Hungary, Portugal, Spain, Poland, Norway, Luxembourg, Italy, Netherlands, Ireland, Greece, Belgium Germany, Denmark, France Czech Republic, Sweden, Austria), from 1989 to 2016. We collect firm-level accounting information as provided by Worldscope, as well as the monthly total returns for common stocks from Datastream.

Tools: With the use of a dataset consisting of European companies from 21 countries, portfolio analysis and time series regressions are performed using abnormal monthly returns (monthly returns minus risk-free interest rates).

Findings: We find that the F-Score is a statistically significant predictor as well as an economically meaningful index. Its performance forecasting ability is visible in developed Europe, both in small and large companies, and remains stable after controlling for

[1] *Corresponding author:* Department of Tourism Studies, University of Piraeus, Piraeus, Greece, email: chriskamp@unipi.gr

established cross-sectional determinants (such as book market, investment, and company size).

Contribution: This study seeks to fill the gap in the stock return and F-Score relationship in a European setting controlling for the other financial variables. Our empirical models are tested across a number of different economic and stock market backgrounds and the implications of our results are of particular interest for academics, for investors (retail and institutional) and for policy makers.

Keywords: F-Score, Stock Returns, Value Investing, Portfolio Analysis.

JEL codes: G11, G12, G15, M41.

1. Introduction

Over the last decades, academic research has focused on examining the structural reasons for equity under or over performance. Focused research has been performed on specific indexes that could be applied in selecting stocks with the best value. Among the most recent explanatory indicators that gained ground in predicting firms' future performance and returns, seems to be the F-Score (Piotroski, 2000), which has reflected strong ability in identifying stocks that outperform the market.

Piotroski' F-Score, a method that measures a firm's fundamental strength, is a robust predictor of future profitability and subsequent stock returns. According to Mohr (2019), high F-Score firms outperform in emerging markets low F-Score firms by about 10% on a yearly basis. Furthermore, F-Score preserves its power to determine future returns across all size segments after examining established cross-sectional return variables, such as book-to-market, investment, operating profitability, and firm size. The findings indicate that investors teen to incorporate substantial information only into stock prices gradually.

As Piotroski (2000) observed, value stocks show abnormal returns: the average return adjusted to the market for value stocks is negative, while the average adjusted market return for value stocks is significantly positive. Therefore, the average returns for value stocks are below the market. However, several value stocks present very significant returns. Therefore, Piotroski generates a new index that aims to separate the few "winning" value shares from the majority of the value shares that seem to lose ground due to their distressed nature (Fama & French, 1993).

Grantham (2010) arrived at similar results in his work that explains the fundamental returns pattern of value stocks' performance. He illustrates an overall long-term return on value stocks that can nevertheless be severely affected in times of severe economic downturn. Grantham (2010) highlights that value stocks' performance was significantly lower than the market during important historical events such the Great Recession and the global financial crisis. This result is explained by the stable financial strength required for companies to overcome extreme situation and survive over time - a characteristic value that stocks rarely possess (Mohr, 2019).

The main objective of the present research is to study the return-predictive ability of the F-Score in the cross section of international stock markets with the use of data from 21 European countries. We base our methodology on the work of Walkshäusl (2020) and the main hypothesis that we examine is whether low F-Score firms produce higher future stock returns than high F-Score firms. In doing so, we extend the scope of Piotroski's research (2000) by observing the F-Score application for the first time across a wide sample of European firms.

Our purpose is to enlighten the international economic significance of the F-Score as a return predictor. Finally, as far as the sample period is concerned, the present study is performed in a quite unique time frame, 1989-2016, since it covers both periods of "bull" and "bear" equity markets and periods of economic expansion (1989-2002, 2004-2008, 2010-2011) and economic downturn (2002-2004, 2008-2011). In particular, the second period of recession was quite severe following the burst of the real estate bubble in the US which triggered the global financial crisis. Thus, our empirical models are tested across a number of different economic and stock market backgrounds and the implications of our results may be of particular interest not only for academics, but also for investors (retail and institutional) or policy makers.

The rest of the article is organised as follows: Section 2 presents the literature review; section 3 details the methodology and data analysis of the study; section 4 presents the empirical results, and section 5 concludes the article.

2. Literature review

Although F-Score stands as an individual performance forecasting tool, it has been investigated in the past extensively among corporate sub-samples (value shares) as well as in combination with other financial indicators, such as the book -to- market.

However, only three studies have explicitly examined the net F-Score return relationship lacking complementary variables. These studies examine firms in USA (Turtle & Wang, 2017), in Australia (Hyde, 2018), and in five individual Asian markets (Ng & Shen, 2019). Apart from the latter whose focal point is on the post 2000 period, the two other studies examine a more extended period.

Piotroski, through his seminal work, develops a complex accounting metric of a company's financial strength, namely the F-Score, employing historical financial statement information, aiming to create a reliable tool that would identify fundamentally strong and weak firms (Mohr, 2019). The results of his research in the US market indicated a significantly positive F-Score return relationship between companies that incorporated a high book-to-market ratio, that is robust to the standard controls established in that period. Since then, in the academic US literature, we evidence that the F-Score has become a particularly popular stock screening tool among US investors (Novy-Marx, 2014).

Specifically, in the work of Fama and French (2006) it has been utilized to predict firms' future profitability. Sias and Choi (2012) reports that it has been applied for institutional investor demand as well as an instrument indicator in further examining the way public important information is reflected into prices (Turtle & Wang, 2017). Towards that direction, Piotroski and So (2012) along with Safdar and Ahmed (2018) both present that investors' expectation errors concerning the entity's fundamental strength, as reported by F-Score, cause the US momentum premiums and consequently assist in explaining the particular anomalies. Moreover, a recent growing body of literature also documents the sustainability of F-Score among various applications outside the US.

Ng and Shen (2016) results are consistent with Piotroski and So (2012), revealing that F-Score provides a significant assistance in identifying, timely or aforehand, the subsequent winners, among Asian developed firms. Moreover, Ng and Shen (2016) and Walkshäusl (2017) in their studies implemented in seven Asia-Pacific markets, Europe, and the US, provide evidence that there is a strong performance relationship between F-Score and the full set of book to market ratios (i.e. stocks of value and growth). They suggest that positive value-growth returns are concentrated between high F-Score value stocks and low F-Score growth stocks but are absent between low F-Score value stocks and high F-Score growth stocks. Therefore, and according to an explanation based on inaccurate prices, their results reveal that the premium value of price corrections derives from the reversal of investor expectation errors. However, this seems to be applicable in companies where their market performance expectations are implied by the book-to-market ratio and are not compatible with the actual firm's fundamental performance, as this is calculated by the F-Score (Walkshäusl, 2020).

Similarly, Walkshäusl (2017, 2019) and Ahmed and Safdar (2018) provide supportive evidence that F-Score is useful in explicating the momentum premium by detecting strong interactions between the firms' past price performance and F-Score.

Considering that investors tend to react to changes in core fundamentals, they find that positive win-loss returns are concentrated among those firms where the performance of previous prices is in line with firm fundamentals but is absent among firms in which the previous price performance is incongruent with the fundamental strength of business (Walkshäusl, 2020).

The work of Tikkanen and Äijö (2018) proposes that the performance of European long-only value investing strategies that employ valuation indexes apart from the book to market ratio can be significantly improved by consolidating the data contained in F-Score. In the same study we find evidence that the consolidation of the data contained in F-Score may improve the performance of various long only value investing strategies in Europe, provided the set are formed on a variety of valuation indexes, that does not include the book to market ratio.

Summarizing the literature review, that previous work in the area of the F-Score metric examines whether it can be used to predict future financial profitability, the relationship of F-Score with book-to-value metrics and the momentum premium and whether it can be used as a successful investment strategy tool. There only three studies that examine the relationship between the F-Score and future stock returns, without the use of complementary variables, and in other countries except Europe.

Thus, from the review of the literature it is evident that there is a gap in the stock return and F-Score relationship in a European setting controlling for the other financial variables.

3. Methodology and data analysis

3.1 Data description

The dataset of the present research consists of listed European companies from 21 countries (in random order: Finland, United Kingdom, Switzerland, Turkey, Hungary, Portugal, Spain, Poland, Norway, Luxembourg, Italy, Netherlands, Ireland, Greece, Belgium Germany, Denmark, France Czech Republic, Sweden, Austria), from 1989 to 2016.

We start by collecting firm-level accounting information as provided by Worldscope, as well as the monthly total returns for common stocks from Datastream. One problem that should be considered in the collection of the data is survivorship, which is a type of sample selection bias that occurs when a data set only considers “surviving” or existing companies and fails to consider companies that already ceased to exist. Generally speaking, survivorship bias tends to create conclusions that are overly optimistic, and that may not be representative of real-life environments. The bias occurs because the “surviving” observations often tend to have survived due to their stronger-than-average resilience to difficult conditions and leaves out other observations that have ceased to exist as a result of such conditions.

To deal with the survivorship bias issue, the sample of the present research consists of active and inactive firms. Inactive firms are the ones that previously existed in the stock exchange and for several reasons, such as bankruptcy or delisting or having been acquired do not exist at the last year of the research. Upon data availability the starting year for each country included in the sample is possible to vary. We also choose not to include financial firms as the distinction between financing and operating activities is not explicit in the specific industry.

Stocks presenting price returns less than 50% or above 300% that are reversed within one month are excluded from our sample so to identify suspicious returns, (Ince and Porter, 2006). All company-level variables as well as returns are calculated in US dollars. All the variables among company-level are winsorized at the 1% and 99% level aiming to mitigate the impact of extreme prices. Also, according to Titman *et al.* (2013) the prerequisite for each country to be included in the sample is to have at least 30 stocks in each sampling year to ensure that the number of companies in the portfolio and cross-sectional regression testing is reasonable and can continue. Our last step for conducting a solid sample to examine was to restrict it to firm-year observations without missing data to calculate the primary variables of interest (for example size, book-to market values and F-Score).

The Book-to-Market index (BM) is defined as the ratio of equity at the end of the financial year (Worldscope item 03501) over the market capitalization (Fama & French, 1992; 1993). The company size (SZ) is the market equity, at the end of June yearly (Fama & French, 1992; 1993). It should be noted that firms presenting a negative book value of equity are excluded in the final sample.

The stock returns are calculated using the yield ratio provided by Datastream (element RI), which is defined as the theoretical increase in the value of one unit of shares at the closing price valid on the dividend date. The raw return on equity return for a company at month j is depicted as: r_j=RI_(j+1) / RI_j -1. Raw stock returns calculation is a procedure that will initiate six months after the end of the financial year.

For the calculation of size adjusted returns, the size benchmark portfolio is formed on a yearly basis allocating stocks into equally weighted portfolios, based upon market equity. The size-adjusted return of a company is the difference between its monthly total return and the corresponding monthly return of the benchmark portfolio of that particular company. The stock price performance of the portfolio designed according to the F-Score is calculated using size adjusted forward looking returns and monthly raw returns. Time series regressions are performed using abnormal monthly returns (monthly returns minus risk-free interest rates). For reciprocating cross-sections, we also calculate annual returns of one-year raw returns.

Table 1 below presents the basic summary statistics of the sample, categorized by country. Obs. is the number of firm-month observations, Ln(BM) is the natural logarithm of book-to-market ratio, Ln(SZ) is the is the natural logarithm of market equity measured as of June of each year, Mean ln(BM) and Mean ln(SZ) are the time-series average of the annual means over the sample period, and Time Period is the beginning and ending year of participation in the sample.

Table 1. Summary statistics on firm-level variables across countries

Country	Obs	Time Period	Mean ln(SZ)	Mean ln(BM)
Austria	12,660	1989-2016	12.548	-0.473
Belgium	17,508	1989-2016	12.616	-0.501
Chech Republic	432	2003-2013	12.205	0.554
Denmark	19,212	1989-2016	11.886	-0.316
Finland	20,628	1989-2016	12.549	-0.525
France	132,468	1989-2016	12.078	-0.648
Germany	110,700	1989-2016	12.108	-0.556
Greece	31,668	1989-2016	11.080	-0.118
Hungary	384	2001-2013	11.463	-0.299
Ireland	8,304	1989-2016	12.217	-0.515
Italy	39,348	1989-2016	12.235	-0.579
Luxemburg	528	1994-2016	12.569	-0.380
Netherlands	22,680	1989-2016	12.972	-0.685
Norway	22,056	1989-2016	12.216	-0.361
Poland	648	2004-2016	12.445	-0.258
Portugal	9,336	1989-2016	12.043	-0.216
Spain	22,908	1989-2016	13.054	-0.521
Sweden	48,564	1989-2016	11.564	-0.733
Switzerland	38,448	1989-2016	12.923	-0.520
Turkey	41,688	1989-2016	11.637	-0.492
United Kingdom	251,160	1989-2016	12.583	-0.815

Note: The above Appendix presents the basic statistics of ln(SZ) and ln(BM) variables by country. Obs. is the number of firm-month observations. Ln(BM) is the natural logarithm of book-to-market ratio. Ln(SZ) is the is the natural logarithm of market equity measured as of June of each year. Mean ln(BM), and Mean ln(SZ) are the time-series average of the annual means over the sample period. Time Period is the beginning and ending year of participation in the sample

3.2 Portfolio formation

The construction of F-Score differs from other indexes of multiple variable indicators. F-Score considers whether to meet the overall financial health and the development direction of the company's fundamentals. F-Score is a comprehensive measure of the company's basic strength. Based on the sum of nine binary variables, these variables can be divided into three dimensions of the company's health status and measure different aspects of the company's financial status: profitability, assets balance sheet health and operational efficiency. If the company's basic conditions are established, the indicator variable is equal to 1, otherwise it is not rated.

The F-Score is based on nine individual questions, each one indicating a different aspect of the firm's financial position. For every question the answer is "yes" one point is awarded on the Piotroski scale. A healthy score indicates that the stock trades at low prices compared to its fundamentals. The questions of the F-Score are the following:

1. The operating cash flow is positive.
2. The net profit before special items is positive.
3. The operating cash flow is greater than the net profit before special items.
4. The annual change in leverage (long-term debt divided by total assets) is negative.
5. Annual turnover change (sales divided by total lagged assets) is positive.
6. The company didn't issue shares.
7. Annual return on assets (net income before non-special items divided by total lagged assets) is positive.
8. Liquidity Annual change (current assets divided by current liabilities) is positive.
9. Annual change of gross profit margin (sales minus cost of sales divided by sales) is positive.

A score ranging between 7-9 would be classified as high, 4-6 would be average while a score between 0 and 3 is considered low. Low values of F-Score imply weak fundamentals while high values on F-Score imply the existence of strong fundamentals.

The key component on which our research design is based, involves the investigation of the implicit performance expectations of F-Score, under the assumption that the index may be used as a leading indicator of future firm performance. The sample firms are allocated into three different portfolios, depending on their F-Score. A firm is classified in a) the low portfolio when its F-Score is between zero and three, b) the medium portfolio when its F-Score is between four and six, and c) the high portfolio when its F-Score is between seven and nine respectively. The portfolios are rebalanced each year while the monthly size-adjusted returns are computed for the subsequent 12 months. 'High–Low' reflects the performance of the spread in terms of returns between high- and low-F-Score firms. We also report the t-statistic significance calculated for the average monthly return.

To calculate stock size adjusted returns, we form size reference portfolios on a yearly basis. The company's adjusted return is defined as the difference between its monthly total return and the monthly return corresponding to the company's benchmark size portfolio. In the second quarter of each year end t, stocks are classified into portfolios according to their F-Score, and the following 12 months (from July in year t to June in year t + 1) returns are calculated for each share. The same procedure is performed for the next year as well, resulting in annual rebalancing of the portfolios. When a company's F-Score falls into the top-ranked quintile of the portfolio, a company's ranking is high. When your F-Score falls in the lowest ranked quintile, the company's ranking is low.

4. Empirical results

Table 2 below reports the results from the F-Score portfolios in terms of returns. The returns were computed for the following twelve months, from the 7th month of year t to the 6th month of the year t+1. For the next year the same process is followed resulting in annual rebalancing of the portfolios. In this formation, firms are ranked as high when their F-Score is between 7 to 9, medium when the value is between 4 to 6, while when the value falls between 0 to 3 firms are considered as low.

Table 2. Returns of three F-Score portfolios

	Size adjusted returns		
	Mean	**SZ**	**BM**
Low F-Score (0-3)	1.30 %***	245.35	3.51
Medium F-Score (4-6)	0.80%***	251.96	3.86
High F-Score (7-9)	0.30%***	194.16	1.68
High – Low (Portfolio A)	-1.00%***		

Note: This table shows the size-adjusted average monthly return of the investment portfolio formed based on the F-Score ranking. In June of each year, based on the F-Score characteristics generated at the fiscal year ending, all sample companies are assigned to three investment portfolios. When the F-Score is between 0 and 3, companies are classified in the lower portfolio, in the medium portfolio we have companies with F-Score between 4 and 6, and in the high portfolio we have companies with F-Score between 7 and 9. Monthly adjusted returns are calculated with equal weights for the investment portfolio for the next 12 months, and the investment portfolio are rebalanced yearly. The `High-Low` metric provides the return spread between high and low F-Score companies.

The average returns as illustrated in Table 2 are 1.30% in regards with the low F-Score firms, 0.80% in medium rated F-Score firms and 0.30% for low F-Score firms, indicating that the lower the F-Score the higher are the average returns. The findings from the univariate sorts (Portfolio A) are consistent with a statistically significant F-Score effect, which has been supported significantly in the existing literature. High F-Score firms are displayed to underperform low F-Score firms by -1.00% on a monthly raw return basis. The results from the univariate analysis reflect that returns of European stocks vary significantly according to F-Score classification.

From the results it has been verified that the return difference between high and low F-Score firms depends on their F-Score classification. Investment strategies involving low F-Score company stocks will achieve a large and statistically significant positive average total return of 1.30%. In contrast, an investment strategy that selects stocks from high F-Score companies has an average return of 0.30%. The specific results are statistically insignificant.

Moving beyond and analyzing further the rest portfolio characteristics, we observe that firms incorporating high F-Score are smaller, in terms of size, than low F-Score firms (Fama and French, 1992). Thus, the negative returns reflected to the High – Low strategy can be attributed to a wider spread in the F-Score features. Firms characterized as low illustrate higher returns, whereas firms characterized as high employ lower returns, leading in a raw return difference of -0.010 or -0.97% respectively, that is statistically significant. Specifically, the firms allocated in low and medium F-Score portfolios present average returns of 0.013 (1.27%) and 0.008 (0.79%), while firms with high F-Score have returns of 0.003 (0.30%). We observe significantly positive return differences between low F-Score firms and high F-Score firms. We also observe significantly negative return differences between the so-called strong firms and weak firms (High-Low).

The next step in the analysis involves the segmentation of the sample companies into four portfolios depending upon their F-Score. Table 3 below presents the results from this part of the analysis. We have allocated Low F-Score firms between 0-3, high includes firms with F-Score between 7-9, while the medium portfolio is now consisted of two parts: the lower part which includes firms ranked between 3-5 and the upper part including firms ranked between 5-7 points. The table illustrates the monthly average returns (size-adjusted) for the firms in the examined sample according to their F-Score allocation.

Table 3. Returns of four F-Score portfolios

	Size adjusted returns		
	Mean	**SZ**	**BM**
Low F-Score (0-3)	1.20%***	267.64	3.78
Lower Medium F-Score (3-5)	0.95%***	261.72	2.24
Upper Medium F-Score (5-7)	0.83%***	246.38	3.90
High F-Score (7-9)	0.69%***	245.91	7.30
High – Low ((Portfolio B)	-0.51%***		

Note: This table shows the size-adjusted average monthly return of the investment portfolio formed based on the F-Score ranking. In June of each year, based on the F-Score characteristics generated at the fiscal year ending, all sample companies are assigned to three investment portfolios. When the F-Score is between 0 and 3, companies are classified in the lower portfolio, in the medium portfolio we have companies with F-Score between 4 and 6, and in the high portfolio we have companies with F-Score between 7 and 9. Monthly adjusted returns are calculated with equal weights for the investment portfolio for the next 12 months, and the investment portfolio are rebalanced yearly. The `High-Low` metric provides the return spread between high and low F-Score companies.

As it can be seen in Table 3 above the returns are 1.20%, 0.95%, 0.83%, 0.69% for portfolios with companies of low, lower medium, upper medium and high F-Score accordingly. We observe that the firms allocated in low and lower medium F-Score portfolios illustrate stronger returns 1.20% and 0.95%, while firms included in upper medium and high F-Score portfolios have 0.83% and 0.69% returns respectively. The table is illustrating that the lower the F-Score the higher are the average returns.

The findings from the univariate sorts of four portfolios are consistent with a statistically significant F-Score effect supported significantly in the existing literature. High F-Score firms allocated in the highest F-Score portfolio underperform against those that are allocated in the lowest F-Score portfolio by -0.51% monthly. The results from the analysis reflect that returns of European stocks vary significantly according to F-Score classification and is verified that the return difference between firms with high F-Score to firms with low F-Score diverse conditionally upon their classification. An investment strategy involving low F-Score stocks is awarded with a large and highly statistically significant positive average return of 1.20%. In contrast, an investment strategy based upon high F-Score stocks presents only 0.69% average returns.

Analyzing the portfolio characteristics, we observe that firms with high F-Score are smaller, in terms of size, than low F-Score firms (Fama & French, 1992). Thus, the negative returns reflected to the High–Low strategy can be attributed to a wider spread in the F-Score features.

The next step of the analysis involves the segmentation of the sample stocks into five equally weighted portfolios according to their F-Score and the calculation of size adjusted returns for the following twelve months, of year t + 1. The process is repeated for the following year resulting in the annual balancing of the portfolios, and the results are presented into Table 4 below.

Table 4 illustrates average returns on a monthly basis (size-adjusted) for the univariate portfolios that are sorted based on F-Score. The portfolios are formed in quantiles, meaning five equally weighted portfolios based on F-Score and the returns were computed for the following twelve months, from the 7^{th} month of year t to the 6^{th} month of the year t+1. For the next year the same process is followed resulting in annual rebalancing of the portfolios. In this formation we may observe that a firm is ranked as high when its total F-Score falls into the highest-ranked quintile portfolio while is ranked as low when its total F-Score falls into the lowest-ranked quintile portfolio. The three portfolios in between represent the medium ranked firms, those with F-Scores that cannot been included either in the highest or the lowest quintile portfolio.

Table 4. Returns of five F-Score portfolios

	Size adjusted returns		
	Mean	**SZ**	**BM**
Low F-Score	1.20%***	267.64	3.79
Lower Medium F-Score	0.95%***	261.72	2.25
Medium F-Score	0.80%***	246.38	3.91
Upper Medium F-Score	0.69%***	245.91	7.31
High F-Score	0.49%***	223.21	1.36
High – Low (Portfolio C)	-0.71%***		

Note: This table shows the size-adjusted average monthly return of the investment portfolio formed based on the F-Score ranking. In June of each year, based on the F-Score characteristics generated at the fiscal year ending, all sample companies are assigned to quantiles, five equally weighted portfolios based on F-Score. A firm is ranked as high when its F-Score falls into the highest-ranked quintile portfolio, its ranked as low when its F-Score falls into the lowest-ranked quintile portfolio. The three portfolios in between represent the medium ranked firms, these with F-Score that cannot been included either in the highest or the lowest quintile portfolio Monthly adjusted returns are calculated with equal weights for the investment portfolio for the next 12 months, and the investment portfolio balances per year. \`High-Low\` metric provides the return of the spread between high and low F-Score companies.

As shown, the returns presented are from low to high F-Score, have a descending order, illustrating that the lower the F-Score the highest the return. The findings from the portfolio analysis are consistent with a statistically significant F-Score effect, which has been supported significantly in the existing literature. High F-Score firms underperform low F-Score firms by -0.71% on a monthly basis.

Our results are in contrast with the evidence provided by of Fama and French (2008), as the portfolios imply that there is a robust low F-Score to high F-Score return difference among small and large European companies. Our findings are consistent with those of Cooper *et al.* (2008), who accordingly provided sufficient evidence to prove that all market segments have significant F-Score effects. Moreover, we estimate that new equity financing activities have generated huge and very significant size adjustment jedgine returns across all size segments. (Fama French, 2008).

Generally, our initial findings are sound and valid across all firm segments of firm size. F-Score high and low F-Score hedging strategy can be further enhanced in the European stock market, combined with the information captured by equity financing activities, suitable for small and large companies.

To examine the way that average returns behave across the variable under investigation we employed portfolios imitation formed through bivariate shorts. However, this aggregation procedure might not identify individual formations per stock. Following this assumption, we try to explore the predictive power regarding returns of F-Score at a panel level employing the OLS regression with grouped standard errors to elaborate the residual dependence created by the time effect and the firm effect.

Our estimation applies in yearly based data regression, using OLS with grouped standard errors, of yearly raw returns, F-Score, book-to-market ratio (BM) and firm size (SZ), as common control variables. The independent variables in the regressions are updated annually at the end of each June to predict yearly stock returns from July of the current year to June of the subsequent year (forward-looking returns). Table 5 documents average coefficient estimations. The reported regression is an estimation regarding the full sample.

Table 5. Panel regressions using OLS with clustered s.e.

	All	**T-statistic**
F-Score	-0.16	***
Ln (SZ)	0.00	***
Ln (BM)	0.00	***

Note: The table presents average coefficient estimates derived from panel analysis using OLS regressions with clustered s.e., along with the relevant t-statistics, which are given in parentheses, of yearly raw returns on F-Score, size (SZ), and book-to-market (BM). The sample consists of 851,328 firm-month observations covering firms (except financial firms) without missing data to compute the primary variables of interest. The time period of the sample covers years from 1989 to 2016. Note that the sample consists of different countries with different time periods. *, **, *** denotes statistical significance at the 10%, 5%, and 1% level respectively.

As it is evident in Table 5, F-Score caries a significant negative coefficient while it is statistically significant in the full sample. Hence, since F-Score's explanatory ability is independent from the considered sample, our findings seem to be more consistent with a risk-based explanation. As the results at the table 5 suggest, in general when there is any incongruence in the valuation signals provided by the two indicators, F-Score provides a stronger effect than size and book to market.

We would expect that in the three High-Low portfolios the average stock returns should illustrate congruence in the valuation signals in terms of the total F-Score ratio and with respect to high returns when F-Score is high and low returns when F-Score is low. For example, a company that is classified as a high F-Score, due to its F-Score ratio, expressing its sound financial performance, is expected to stipulate accordingly high returns. According to the mispricing-based explanation (Bali *et al.*, 2010), the forecasting ability in terms of return of the F-Score should be statistically significant only in the case that the examined sample is consisted of congruent companies, whereas not significant when the examined sample is consisted of incongruent companies, as noisier valuation signals are observed. Furthermore, according to the risk-based explanation, the forecasting ability in terms of returns for F-Score related firms should be independent from the examined sample, as the higher returns to low F-Score companies are generally a compensation of risk.

Table 6 shows the average monthly returns for all three spread portfolios that were generated from the three different shorting formations used previously based on F-Score. In all three of them the column high -low illustrates the return difference between firms of high F-Score vs firms of low F-Score.

Table 6. Portfolio hedges

Portfolio A Hedge (H-L)	**Portfolio B Hedge (H-L)**	**Portfolio C Hedge (H-L)**
-0.97%***	-0.51%***	-0.71%***

Our findings are not consistent with the hypothesis that the highest the F-Score a firm has the highest average returns it may accomplish, hence the fundamental strength that illustrates is robust enough.

In the first formation where there are 3 available portfolios based on F-Score the deviation of High-Low is illustrated to be -0.97% indicating that low F-Score firms outperforms on average the high F-Score firms. Similarly, in the second formation where 4 available portfolios were created, the deviation captured was lower at -0.51% while in the third formation upon quantiles method so five F-Score portfolios were available the deviation was -0.71%, explicitily declaring the outperformance of low F-Score firms vs high. The return difference is evaluated economically and statistically significant.

Piotroski and So (2012) in their research support with U.S. evidence that when we observe market expectation errors, i.e. high F-Score firms that illustrate a weak financial condition vs low F-Score firms that present to have a robust financial condition, we assume that the value premium is significantly higher. Nevertheless, when we observe that the market expectation errors do not exist, i.e. high F-Score firms that illustrate a strong financial condition and low F-Score firms that present a weak financial condition, then the value premium does not exist. Before further investigating the explanation based on erroneous prices, we need to check whether the Piotroski and So (2012) market expectation error approach also applies to the anomaly of asset growth in European stock markets.

As mentioned, we employ an approach so to realize whether the findings of our research could be attributed to market errors. As we have based our research in forming shorts upon Piotroski's (2000) F-Score, we focus on whether the information provided by F-Score is affected by market expectation errors as well. Towards that direction, F-Score is considered an index of market expectations regarding the firm's future performance as provides an independent overview of the firm's strength in its fundamentals.

In a rational base, firms acquiring a low F-Score index signal weak fundamentals and pessimistic future performance are expected, while firms with high F-Score, express a strong financial condition (in terms of profitability, financial leverage/liquidity, and operational efficiency) and would expect optimistic future performance. According to the specific theoretical framework, high F-Score companies illustrate strong fundamentals while low F-Score exhibit weak financial strength.

5. Conclusions

In this research, we examined Piotroski's F-Score and the stock returns it produced among a wide range of European firms, aiming to clarify its actual predictive ability in a global financial environment. We observe that the F-Score is a statistically significant predictor as well as an economically meaningful index as tests upon a wide section of European stock returns imply. Its performance forecasting ability is visible in developed Europe, both in small and large companies, and remains stable after controlling for established cross-sectional determinants (such as book market, investment, and company size).

Although our results are not aligned with the original evaluation of Walkshaul, we still support the significance of F-Score, that is established as a fairly global financial indicator gaining supporters throughout the world. Furthermore, it seems unreasonable to assume that firms with strong fundamentals, as these are indicated by F-Score index, may be considered riskier than firms with weak fundamentals. However, the results derived from our research remain consistent with the Piotroski's point of view almost 20 years ago when he insisted that basic information is only gradually incorporated in the stock price by the investors. Walkshaul indicated in his seminar paper that over the long term, the financial behavior of companies based on their F-Score value is consistent. Two prominent explanations can be proposed to explain F-Score abnormal returns. The first is more rational, while the second indicates some form of incorrect pricing. Regarding the explanation for the incorrect pricing or mispricing's abnormal performance that was caused by naive investors who failed in successfully including the complete F-Score information into the company's share price. In a frictionless and rational asset pricing framework, higher average returns for low F-Score companies reflect compensation for higher systemic risks.

In a rational and frictionless asset pricing framework, the higher average returns coming from firms with low F-Score index reflect compensation for higher systematic risks. We should pay attention to the fact that that what really causes the effects of asset changes on stock prices is still debatable in the literature. This question prompts what we are trying to do in a particular article. Using a complete sample of the European stock market from 21 countries / regions during 1989 to 2016 we directly assess whether the performance of the F-Score - based portfolio can be attributed to risk or incorrect pricing. We believe that investigating the potential sources of anomalies on specific topics outside of the United States, especially in Europe, can provide additional insights to understand whether this important asset pricing law poses a significant challenge to market efficiency.

Watanabe *et al.* (2013) and Titman *et al.* (2013) in their prominent work argue that the impact of asset growth on equity returns is widespread outside the United States, especially in countries where equity markets are more efficient. Although their arguments may lean toward a risk-based explanation, we must consider that his empirical research design is not focused on distinguishing between mispricing and reasonable explanations of abnormal asset growth, but rather on determining what may be wrong.

Following the consensus that investors misinterpret changes in the company's future business prospects implicit in the expansion or contraction of assets, leading to systematic pricing errors, we found that the use of variables believed to reveal signs either of overvaluation and undervaluation should help in identifying timely mispriced stocks. According to Bali *et al.* (2010) and Walkshäusl (2015), the existing literature refers the use of net capital financing activities as an indicator variable that is an important predictor of the actions of the profitability in a wide range of the cross-section.

Extrapolating the argumentation in Bali *et al.* (2010), High F-Score Returns - Low F-Score returns show the consistency of the valuation signals. Therefore, the conditional probability that each valuation signal is caused by incorrect prices rather than noise is high. On the other hand, the high F-Score yield leading to the low F-Score yield shows inconsistent valuation signals. In this case, the ability of each signal to identify price errors will be reduced, because the inconsistency between the signals means that the signal is more likely to be due to noise.

In addition, in order to support mispricing, we hope to find at the individual level of our empirical analysis (ie, cross-sectional regression) that there is a negative and significant relationship between total F-Score and subsequent returns only in sub-samples, where F-Score points in the same direction as the hedging performance indicator (concurrent signal). In the case of conflicting signals, that is, the sub-sample with the loudest signal, we expect that the impact of these two indicators on stock returns will be weakened, or even negligible.

As far as suggestions for further research it would be of interest to examine whether the results of the present study are similar between developed and less developed or developing countries, thus, examining the effect of qualitative characteristics on the F-Score phenomenon. Furthermore, one could examine if the predictive power of the F-Score changes across time, industries or countries.

References

Ahmed, A.S. & Safdar, I. (2018*)* “Disserting stock price momentum using financial statement analysis”, *Journal of Accounting and Finance*, vol.58, no.1: 3-43.

Bali, T.G., Demirtas, K.O. & Hovakimian, A. (2010*)* “Corporate financing activities and contrarian investment”, *Review of Finance*, vol.14, no.3: 543-584.

Cooper, M. J., Gulen, H. & Schill, M. J. (2008*)* “Asset growth and the cross-section of stock returns”, *Journal of Finance*, vol.63, no.4: 1609-1651.

Hyde, C.E. (2018*)* “The Piotroski F-score: evidence from Australia”, *Journal of Accounting and Finance*, Vol.58, no.2: 423-444.

Fama, E. F. & French, K. R. (1992*)* “The cross-section of expected stock returns”, *Journal of Finance*, vol.47, no.2: 427-465.

Fama, E.F. & French, K. R. (1993*)* “Common risk factors in the returns on stocks and bonds”, *Journal of Financial Economics,* vol.33, no.1: 3-56.

Fama, E. F. & French, K.R. (2006*)* “Profitability, investment and average returns”, *Journal of Financial Economics*, vol.82, no.3: 491-518.

Fama, E.F. & French, K. R. (2008*)* “Dissecting anomalies”, *Journal of Finance*, vol.63, no.4: 1653-1678.

Grantham, J. (2010*)* On the Potential Disadvantages of Graham and Dodd-Type Investing, *GMO Letters to the Investment Committee 16*.

Ince, O. & Porter, B. (2006*)* “Individual equity return data from Thomson Datastream: handle with care”, *Journal of Financial Research*, vol.29, no.4: 463-479.

Mohr, J. (2019*)* “Utility of Piotroski F-Score for predicting growth-stock returns”, *MFIE Capital Working Paper*, Corpus ID: 203657082

Ng, C.C.A. & Shen, J. (2016*)* “Screen winners from losers using simple fundamental analysis in the Pacific-Basin stock markets”, *Pacific-Basin Finance Journal*, vol.39(C): 159-177.

Ng, C.C.A. & Shen, J. (2019) "Quality investing in Asian stock markets", *Accounting & Finance*, vol.60, no.3, https://doi.org/10.1111/acfi.12446.

Novy-Marx, R. (2014) "Predicting anomaly performance with politics, the weather, global warming, sunspots, and the stars", *Journal of Financial Economics*, vol.112, no.2: 137-146.

Piotroski, J.D. (2000) "Value investing: the use of historical financial statement information to separate winners from losers", *Journal of Accounting Research*, vol.38: 1-41.

Piotroski, J.D. & So, E.C. (2012) "Identifying expectation errors in value/glamour strategies: a fundamental analysis approach", *Review of Financial Studies*, vol.25, no.9: 2841-2875.

Sias, R.W. & Choi, N.Y. (2012) "Why does financial strength forecast stock returns? Evidence from subsequent demand by institutional investors", *The Review of Financial Studies*, vol.25, no.5: 1550-1587.

Tikkanen, J. & Äijö, J. (2018) "Does the F-score improve the performance of different value investment strategies in Europe?", *Journal of Asset Management*, vol.19, no.7: 495-506.

Titman, S., Wei, K.C.J. & Xie, F. (2013) "Market development and the asset growth effect: international evidence", *Journal of Financial and Quantitative Analysis*, vol.48, no.5: 1405-1432.

Turtle, H.J., & Wang, K. (2017) "The value in fundamental accounting information", *Journal of Financial Research*, vol.40, no.1: 113-140.

Walkshäusl, C. (2015) "Equity financing activities and European value-growth returns", *Journal of Banking and Finance*, vol.57: 27-40.

Walkshäusl, C. (2017) "Expectation errors in European value-growth strategies", *Review of Finance*, vol.21, no.2: 845-870.

Walkshäusl, C. (2019) "The fundamentals of momentum investing: European evidence on understanding momentum through fundamentals", *Accounting & Finance*, vol.59, no.1: 831-857.

Walkshäusl, C. (2020) "Piotroski's F-SCORE: international evidence", *Journal of Asset Management*, vol.21, no.2: 106-118.

Watanabe, A., Xu, Y., Yao, T. & Yu, T. (2013) "The asset growth effect and market efficiency: insights from international stock markets", *Journal of Financial Economic*, vol.108, no.2: 529-563.