Does the F-score improve the performance of different value investment strategies in Europe?

Authors: Tikkanen, Jarno; Äijö, Janne

Title: Does the F-score improve the performance of different value investment strategies in Europe?

Year: 2018

Version: Accepted manuscript

Please cite the original version:

Tikkanen, J. & Äijö, J., (2018). Does the F-score improve the performance of different value investment strategies in Europe? *Journal of asset management* 19:7, 495–506. https://doi.org/10.1057/s41260-018-0098-3

Metadata of the article that will be visualized in OnlineFirst

ArticleTitle	Does the F-score improve the performance of different value investment strategies in Europe?	
Article Sub-Title		
Article CopyRight	Springer Nature Limited (This will be the copyright line in the final PDF)	
Journal Name	Journal of Asset Management	

Corresponding Author	Family Name	**Äijö**
	Particle	
	Given Name	**Janne**
	Suffix	
	Division	
	Organization	University of Vaasa
	Address	Vaasa, Finland
	Phone	
	Fax	
	Email	janne.aijo@uva.fi
	URL	
	ORCID	
Author	Family Name	**Tikkanen**
	Particle	
	Given Name	**Jarno**
	Suffix	
	Division	
	Organization	University of Vaasa
	Address	Vaasa, Finland
	Phone	
	Fax	
	Email	
	URL	
	ORCID	

	Received	
Schedule	Revised	4 October 2018
	Accepted	

Abstract

This study examines whether the performance of different value investment strategies can be improved with Piotroski's (J Account Res 38:1–41, 2000) F-score screening method for the European stock markets. Our aim is to investigate the ability of the screening method to distinguish between winners and losers among several value investment strategies that use different financial ratios to form portfolios, such as B/M, E/M, D/M, and EBITDA/EV ratios. The results of the study provide compelling evidence that the F-score screening method significantly improves the performance of all investigated investment strategies. The results regarding the superior performance of the high F-score portfolios are robust across investment strategies, various performance measures and risk-adjustment methods. The results are useful for individual investors and professional portfolio managers.

Keywords (separated by '-') F-score - Anomalies - Value investment strategy - Portfolio management - Portfolio performance

JEL Classification (separated by '-')	G11 - G14 - G15
Footnote Information	

J Asset Manag
https://doi.org/10.1057/s41260-018-0098-3

ORIGINAL ARTICLE

Does the F-score improve the performance of different value investment strategies in Europe?

Jarno Tikkanen[1] · **Janne Äijö**[1]

Revised: 4 October 2018

Abstract This study examines whether the performance of different value investment strategies can be improved with Piotroski's (J Account Res 38:1–41, 2000) F-score screening method for the European stock markets. Our aim is to investigate the ability of the screening method to distinguish between winners and losers among several value investment strategies that use different financial ratios to form portfolios, such as B/M, E/M, D/M, and EBITDA/EV ratios. The results of the study provide compelling evidence that the F-score screening method significantly improves the performance of all investigated investment strategies. The results regarding the superior performance of the high F-score portfolios are robust across investment strategies, various performance measures and risk-adjustment methods. The results are useful for individual investors and professional portfolio managers.

Keywords F-score · Anomalies · Value investment strategy · Portfolio management · Portfolio performance

JEL Classification G11 · G14 · G15

Introduction

Value investing is one of the most popular investment strategies and has received considerable attention among practitioners and academics. There is extensive evidence to show that portfolios formed on the basis of high Earnings to Market (E/M), Book-to-Market (B/M), Dividends to Market (D/M) and Earnings Before Income and Taxes to Enterprise Value (EBIT/EV) ratios, for example, provide abnormal returns world-wide (see, e.g., Fama and French 2012; Gray and Vogel 2012; Asness et al. 2013). Higher returns of value stocks have been explained by greater fundamental risk[1] (see, e.g., Fama and French 1992, 1996; Griffin and Lemmon 2002; Vassalou and Xing 2004; Kapadia 2011; Cakici and Tan 2014) and by mispricing-based explanations, where the value premium is caused by investor behavior[2] (see, e.g., Lakonishok et al. 1994; La Porta 1996; Chan and Lakonishok 2004; Campbell et al. 2008).

> "*The success of the value investment strategy relies on the strong performance of a few companies, while tolerating the poor performance of many deteriorating companies*" Piotroski (2000)

Indeed, Piotroski (2000) reports that 44% of all high B/M stocks yield negative market adjusted returns in a 2-year holding period after portfolio formation. In attempt to exclude these poor performing stocks from the value

✉ Janne Äijö
janne.aijo@uva.fi

1 University of Vaasa, Vaasa, Finland

[1] These studies broadly indicate that high B/M companies tend to be less profitable, more leveraged, have more distress risk, respond more negatively to economic shocks and carry a greater probability of default. Because of the greater risks, these stocks should be priced with a higher risk premium.

[2] This strand of literature broadly suggests that investors are overly pessimistic (optimistic) about the companies that have done poorly (well) in the past, which causes value (glamor) investments to be undervalued (overvalued). Much of this type of mispricing is related to earnings announcement periods and mispricing is especially strong among small companies with low analyst coverage.

portfolio, Piotroski (2000) developed a screening method, called the F-score, which is based on nine accounting variables. They measure a company's profitability, leverage, liquidity and source of funds as well as the operating efficiency of the company. Using these accounting variables as binary signals to screen stocks, Piotroski (2000) aims to separate winners (high F-score stocks) from losers (low F-score stocks) among the universe of high B/M stocks, and to shift the distribution of returns of the traditional value portfolio. Indeed, by selecting *ex ante* apparently financially strong companies among the high B/M stocks, Piotroski (2000) finds that returns to the traditional high book-to-market portfolio can be increased by as much as 7.5%. Furthermore, a zero-cost strategy that buys (sells) financially strong (weak) value stocks yields an astonishing 23.0% return annually. A considerable proportion of these returns is related to small and medium-sized companies that have low share turnover and analyst coverage, whereas the applicability of the F-score screening method is rather limited among large company stocks.

The follow-up study by Piotroski and So (2012) focused on examining the value-growth anomaly using the F-score as a criterion for separating financially strong/weak companies among value and growth stocks. They find that the return spread between value and growth stocks is 22% (0%) when value stocks with strong (weak) fundamentals are compared with growth stocks with weak (strong) fundamentals, implying that investors' expectation errors explain the value-growth anomaly. Choi and Sias (2012), in turn, find that financial strength, as measured by the F-score, is useful not only in predicting future stock returns, but also the demand of institutional investors. First, they find that the difference in returns between high and low F-score portfolios is 25.7%, which is in a similar region to the findings of Piotroski (2000). Moreover, they conclude that because of the financial strength of some companies, the demand among institutional investors is high for these stocks, driving the spread between the performance of high and low F-score stocks. Recently, Walkshäusl (2017) focused on the ability of the F-score also to explain the value-growth anomaly in the European stock markets. Consistent with the empirical evidence on the US stock markets, they find that investors' expectation errors are greatest between value stocks with strong fundamentals and growth stocks with weak fundamentals, which explains the value-growth anomaly. Moreover, they document that prior external financing activities, such as equity issuance and stock repurchases, affect the anomalous stock returns.[3]

Inspired by these results, the research question we aim to answer is simple: Does the F-score improve the performance of different value investment strategies? The research question is timely and is also motivated by the wide practical acceptance of the F-score among professional money managers as a stocks screener (see Novy-Marx 2014). It is nowadays also easily available on internet for individual investors to select stocks. The related literature has examined the usefulness of the F-score in improving the returns of value portfolios consisting of high B/M stocks (see Piotroski 2000), momentum portfolios (Turtle and Wang 2017) or they have focused in studying the value-growth anomaly (Piotroski and So 2012; Walkshäusl 2017). In addition, most of these studies have focused on the performance of long-short strategies. However, several studies have revealed a number of short-sale restrictions, such as institutional constraints, arbitrage risk, behavioral biases of traders and trading costs that may prevent individual, professional and institutional investors implementing the long-short strategy.[4] Therefore this paper focuses on the usefulness of the F-score screening method among various long-only value investment strategies. Given that value investment strategies are one of the most popular investment strategies among practitioners this is an important task.

The purpose of the study is thus to exclude *ex ante* the poorly performing companies from the value portfolios with Piotroski's (2000) composite score analysis using the European stock markets as a laboratory. Besides examining the applicability of Piotroski's (2000) screening method among high B/M stocks, we investigate the mechanism's effectiveness among other investment strategies that use different valuation ratios, such as E/M, D/M, and EBIT/EV ratios to form portfolios. Moreover, to ascertain that the returns are not driven by various measures of risk, the paper examines different performance ratios of high and

[3] Similar to these studies, Fama and French (2006) show that F-score is a leading indicator of future profitability implying that it is a good proxy for the strength of companies' fundamentals, while Broussard et al. (2016) show that the F-score is useful in differentiating which companies migrate from value portfolios to neutral or growth portfolios and therefore also serves to explain the returns of different investment styles. Finally, Turtle and Wang (2017) show that the F-score can be used to improve the performance of momentum strategies, while Novy-Marx (2014) uses F-score as a stand-alone criterion to screen stocks and finds that the F-score screening method, without sorting stocks first based on B/M ratio, offers one of the best performance among several quality investment strategies investigated in the study.

[4] For example, while individual investors rarely take short positions, many institutional investors are simply not allowed to take short positions. For references and for more detailed discussion of the literature, please see Stambaugh et al. (2012).

low F-score portfolios, such as various drawdowns, Sortino and Sharpe ratios, and the Fama and French (2015) five-factor adjusted abnormal returns. Finally, we are also interested in examining the factor loadings of the portfolios on the Fama and French (2015) five-factor model as they provide more detail on the risk characteristics of the portfolios.

By so doing, the study differs from earlier studies in two main ways. First, the study extends the empirical findings of Piotroski (2000) using the B/M ratio to form value portfolios by examining F-score screening method with several different value investment strategies. This is important, as shown by Gray and Vogel (2012) for example, who demonstrate that there exist significant differences between the performances of various value investment strategies. In addition, it is important to investigate whether the F-score screening works for all strategies similarly in terms of performance improvement and whether the screening strategy identifies the returns of low and high score portfolios better for certain investment strategies. The value investment strategies investigated in our study are based on high B/M, E/M and D/M stocks as well as EBIT/EV and EBITDA/EV stocks (see Loughran and Wellman 2011; Gray and Vogel 2012). Following Gray and Vogel (2012), we also examine the recently proposed investment strategy that uses the profitability ratio in stock selection, as it has been shown to be one of the best performing investment strategies (see Novy-Marx 2013, 2014). Given that it is already a more quality-like[5] investment strategy even before conditioning with the F-score screening, one could anticipate that the F-score screening would not necessarily work equally well for the strategy. However, it is interesting to examine whether the F-score screening also contains useful additional information for the strategy, i.e., whether the screening can also improve its performance.

Secondly, in addition to extensive analysis of various risk-adjusted performance measures, such as examining for the first time whether the Fama and French (2015) five-factor model can explain the anomalous returns, one novel aspect of the study is that we focus on the risk characteristics of the high and low F-score portfolios as measured by the loadings of the portfolios for the Fama and French (2015) five-factor model. This is particularly important as it reveals the sensitivities of the portfolio returns to these risk factors and serves to explain the return patterns.

The results of the study demonstrate that Piotroski's (2000) screening method works for all examined strategies in the European stock markets. The best performance is achieved by applying the F-score screening for the EBITDA/EV strategy with compound-annual growth rates (CAGR) of 19.62%, while the greatest improvement over the traditional portfolio is obtained for the D/M and B/M strategies, suggesting that the F-score screening is particularly useful for these strategies. The main results regarding the improvement in the performance of all investment strategies are consistent with those of various performance measures and risk-adjustment methods. The results also show that high and low F-score portfolios have different risk characteristics in terms of loadings for the Fama and French (2015) five-factor model. The results have important implications for professional money managers and individual investors alike. The paper is organized as follows. The next section provides the empirical framework including the presentation of the data and methodology. "Result" section presents the empirical results, while the final section concludes the research.

Data and methodology

The sample consists of public European companies examined during the period 1992–2014. Following Fama and French (2012), the countries included in the sample are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the UK. The data were obtained from the Thomson Reuters Datastream database.[6] Financial companies are excluded from the sample owing to the differing interpretation of their financial statements (see, e.g., Fama et al. 1992; Piotroski 2000; Asness et al. 2013). Due to possible illiquidity issues AQ4 the smallest 10% of companies are excluded from the sample (see, e.g., Gray and Vogel 2012). We also required that all the variables necessary to calculate the F-score were available; if they were not, the company was excluded from the sample.

Following Walkshäusl (2017), we use equally weighted portfolios in our empirical analysis. Therefore, the STOXX Europe 600 equal weighted gross return index is used as the market portfolio (see Plyakha et al. 2012). Since the data for the gross return index is not available prior to 2001, the STOXX Europe 600 equal weighted net return index is used between 1992 and 2000. This naturally leads to underestimating the market return by the amount of

[5] According to Novy-Marx (2014), the strategy selects stocks that are of high quality (i.e., high profitability), while value strategies tend to select stocks that are of low quality (e.g., low profitability, high debt).

[6] It is acknowledged that Datastream may have problems with the small stock market data. Following Ince and Porter (2006), we manually screened our data sample to ensure that there were no unusually high monthly returns or dividend yields.

Journal : Large 41260 | Dispatch : 21-11-2018 | Pages : 12
Article No. : 98 | □ LE | □ TYPESET
MS Code : JASE-D-18-00030 | ✔ CP | ✔ DISK

withholding taxes on paid dividends over that period.[7] Finally, following Fama and French (2012), the 1-month US T-bill rate is used as the risk-free rate and all returns are reported in US dollars. The results remain qualitatively the same if local currencies are utilized.

The F-score and portfolio formation

Following Piotroski (2000), the composite score of a company is based on nine individual binary signals. Therefore, the F-score of a single company may vary from 0 (low) to 9 (high). These signals can be classified into three different categories. The first category is based on a company's profitability. The second category is designed to measure the company's leverage, liquidity and source of funds. Finally, the third category measures the operating efficiency of the company. These individual signals are defined as follows:

1. If the company's return on assets (ROA) is positive, the indicator variable F_ROA is equal to one, and otherwise it is equal to zero. The ROA is defined as net income before extraordinary items divided by the beginning of year total assets.
2. F_ΔROA is defined as a current year's ROA less the previous year's ROA. If the difference is greater than zero, the variable F_ ΔROA equals one, otherwise it will be zero.
3. F_CFO is the cash flow from operations (CFO) of the company. If it is positive, the variable equals one, otherwise it equals zero. CFO is defined as the cash flow from operations divided by the beginning of year total assets.
4. F_ACCRUAL is defined as the difference of the current year's net income before extraordinary items and cash flow from operations divided by the beginning of year total assets. If the CFO is greater than the ROA, F_ACCRUAL takes a value of one, otherwise zero.
5. F_ΔMARGIN is defined as the company's current gross margin ratio (gross margin divided by total sales) minus the previous year's gross margin ratio. If the current gross margin ratio is greater than the previous year's gross margin ratio, the variable F_ ΔMARGIN is equal to one. Otherwise, it will be equal to zero.
6. F_ΔTURN is the difference between the current year asset turnover ratio (total sales divided by beginning of the year total assets) and the previous year's asset turnover ratio. The F_ΔTURN indicator variable is one if the difference is positive, and zero otherwise.
7. The indicator variable F_ΔLEVER is defined as the change (compared to the previous year's ratio) in the ratio of total long-term debt to average total assets. The variable is given as equal to one if the company's leverage ratio is lower than in the previous year, and otherwise as zero.
8. F_ΔLIQUID measures the change of the company's current ratio between the current and the previous year. The current ratio is defined as the ratio of current assets to current liabilities. If the difference of the current and previous year's current ratio is positive, the indicator variable F_ΔLIQUID equals one, and otherwise zero.
9. EQ_OFFER equals one if the company did not issue new shares in the previous year, otherwise it will be stated as zero.

Next, book-to-market (B/M), earnings before interest and taxes to enterprise value (EBIT/EV), earnings before interest, taxes, depreciation, and amortization to enterprise value (EBITDA/EV), earnings to market capitalization (E/M), dividends to market capitalization (D/M), and Novy-Marx's profitability ratio (Novy-Marx 2013) are calculated for the purposes of the various investment strategies examined. Enterprise value is defined as market capitalization + preferred stock + minority interest + total debt − cash. Following Novy-Marx (2013), the profitability ratio is calculated as total revenue minus cost of goods sold scaled down by total assets.

Following the traditional literature (see, e.g., Piotroski, 2000; Fama et al. 1992, 2012), the portfolios are formed every year on the last trading day of June. Therefore, the portfolio formation is based on the accounting variables as per the end of the previous year. This approach reduces the risk of the results being influenced by the look-ahead bias that may potentially be a problem in studies of this kind. We require a 5-month minimum difference between the fiscal year end and the portfolio formation date.

Following Piotroski (2000), the top 20% of stocks are included in the value investment portfolios based on each financial ratio (e.g., the B/M portfolio contains the top 20% of high B/M stocks, the EBIT/EV portfolio contains the top 20% of high EBIT/EV stocks). This yields a final sample of 12,272 high B/M company-year observations, 12,279 high EBIT/EV observations, 12,255 high EBITDA/EV observations, 12,276 high E/M observations, 11,977 high D/M observations, and 12,273 observations based on the Novy-Marx profitability ratio. Overall, the number of companies in these portfolios varies from 405 to 697 on an annual basis. Table 1 shows the number of companies that are included for each value portfolio from each country and

[7] Given that the main purpose of the study is to investigate whether the Piotroski screening method improves the performance of various value investment strategies, the potential underestimation of the performance of the market portfolio prior to 2001 should not affect the main conclusions of the paper.

Journal : Large 41260 | Dispatch : 21-11-2018 | Pages : 12
Article No. : 98 | ☐ LE | ☐ TYPESET
MS Code : JASE-D-18-00030 | ☑ CP | ☑ DISK

Table 1 Summary statistics

	B/M	E/M	D/M	EBIT/EV	EBITDA/EV	Novy-Marx
Panel A: sample countries						
Austria	227	250	208	228	304	75
Belgium	279	321	275	289	357	116
Denmark	427	371	198	357	373	333
Finland	353	354	611	270	272	344
France	1859	1866	1256	2162	2358	642
Germany	1285	1358	1298	1854	2250	1985
Greece	1015	498	568	450	349	223
Ireland	127	181	111	147	106	123
Italy	932	532	627	657	701	511
Norway	450	469	351	365	365	238
Portugal	309	173	182	96	153	14
Spain	421	399	399	332	291	131
Sweden	450	644	658	597	508	540
Switzerland	644	508	284	485	526	521
The Netherlands	330	648	487	520	517	578
UK	3164	3704	4464	3470	2825	5899
Panel B: variables						
Mean	1.736	0.164	0.077	0.384	0.486	0.688
25th	1.129	0.103	0.044	0.148	0.228	0.513
Median	1.464	0.123	0.055	0.178	0.272	0.600
75th	1.944	0.161	0.075	0.237	0.359	0.749

Panel A shows the total number of company-year observations per country in the final sample during the sample period. The company-year observations are calculated for the top 20% of stocks for each financial variable. Panel B shows the mean, 25th percentile, median and 75th percentile of the variables

also the mean, 25th percentile, median and 75th percentile of the financial variables used in the portfolio formation.

Next, for each value portfolio that consists of the top 20% of stocks based on each financial ratio, we form the high and low F-score portfolios on the basis of the company's aggregate F-score. In a method similar to that used by Piotroski and So (2012), companies with the lowest aggregate score are classified as low F-score companies (F-score 1–3). Conversely, companies with the highest aggregate score are classified as high F-score companies (F-score 7–9). The number of companies with an F-score of 0 is very small (less than 15 company-year observations), and therefore, such companies are excluded from the sample.

Company-specific return is measured as the 1-year buy-and-hold return with dividends reinvested, earned from the last trading day of June to the last trading day of June of the following year. The returns are measured using compound-annual growth rates (CAGR) as in Gray and Vogel (2012). Following Piotroski (2000), if the stock was delisted during the holding period, the delisting return is assumed to be zero.[8] In addition, the weight changes of the stocks stemming from the variability of constituent stock returns during the holding periods are taken into account in the calculation of time-series returns for portfolios.

Finally, the companies are also classified by size into *small*, *medium*, or *large* categories. Specifically, companies are ranked based on their previous year's market capitalization in descending order. Large stocks are defined as those that account for 90% of the coverage of the total market capitalization of the sample. Beyond that segment, the stocks that, when added, take the total market capitalization to 98%, are defined as the medium stocks and the remainder of the stocks are defined as small stocks. This method is similar to that of Asness et al. (2013) and the method the MSCI uses to define its global size references. The result is subsamples based on size that contain 1837–2708 large; 3422–4493 medium; and 5071–7014 small company-year observations.

[8] This approach could potentially cause some bias in our results. However, according to Beaver et al. (2007), this has the most significant impact on the returns of growth stocks, which are not examined in the study.

Journal : Large 41260 Dispatch : 21-11-2018 Pages : 12
Article No. : 98 □ LE □ TYPESET
MS Code : JASE-D-18-00030 ✔ CP ✔ DISK

Risk-adjusted performance measures

The risk-adjusted performance of the portfolios is measured using the Sharpe (1966) ratio that is commonly used in value anomaly literature. The statistical significances of the differences between Sharpe ratios are tested using the Ledoit and Wolf (2008) test, which is based on the circular block bootstrap method. In order to avoid problems stemming from negative excess returns used to calculate the Sharpe ratios, we use the following refinement for the denominator of the Sharpe ratio as suggested by Israelsen (2005):

$$S_p = \frac{R_p - R_f}{\sigma_p^{(\mathrm{ER}/|\mathrm{ER}|)}} \tag{1}$$

where ER = Excess return that is equal to $R_p - R_f$, σ_p = Standard deviation of monthly excess returns of portfolio p.

However, the Sharpe ratio has been criticized for penalizing very high positive returns as they also increase the standard deviation (Goetzmann et al. 2007). Consequently, we also calculate the Sortino ratio, which uses only negative returns to measure risk (Sortino and van der Meer 1991). The Sortino ratio uses the root-mean-square deviation below the minimum acceptable return (i.e., downside deviation). The risk-free rate is used in this study as a minimum acceptable return. More formally, the Sortino ratio can be represented by the following equation:

$$SR_p = \frac{R_p - \mathrm{MAR}}{\sqrt{\frac{1}{n}\sum_{R_p<\mathrm{MAR}}\left(R_p - \mathrm{MAR}\right)^2}} \tag{2}$$

where MAR = Minimum acceptable return, R_p = Return of portfolio p.

In addition to these portfolio performance measures, we use the Fama and French (2015) five-factor model to measure abnormal returns in order to determine whether the returns of the portfolios can be explained by the market, size, value, profitability and investment factors.[9] Different loadings for the risk factors will also reveal the differences in the risk characteristics of the high and low F-score portfolios. In order to avoid autocorrelation and heteroscedasticity, Newey and West's (1987) standard errors are used in the regressions. Abnormal returns are estimated using the following equation from the Fama and French (2015) five-factor model:

$$R_{it} - R_{Ft} = \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i\mathrm{SMB}_t + h_i\mathrm{HML}_t + r_i\mathrm{RMW}_t + c_i\mathrm{CMA}_t + \varepsilon_{it} \tag{3}$$

where R_i and R_{Mt} are the return on the portfolio i and the market, respectively, R_{Ft} is the risk-free rate, SMB_t is the difference between the returns on portfolios of small stocks and big stocks, HML_t is the return on a portfolio of high B/M stocks minus the return on a portfolio of low B/M stocks, RMW_t is the return on a portfolio of stocks with robust profitability minus the return on a portfolio of stocks with weak profitability, CMA_t is the return on a portfolio of companies with low investment levels (conservative companies) minus the return on a portfolio of companies with high investment levels (aggressive companies) and α_i measures the abnormal return for a portfolio i.

The statistical significance of the differences in the alphas (i.e., abnormal returns) between portfolios is tested with the following alpha spread test:

$$t = \frac{\alpha_i - \alpha_j}{\sqrt{\mathrm{SE}^2_{\alpha i} + \mathrm{SE}^2_{\alpha j}}} \tag{4}$$

where α_i = alpha of a portfolio i, α_j = alpha of a portfolio j, SE^2_{α} = standard error of portfolio's alpha.

Results

Returns of the portfolios

Table 2 presents the compound-annual growth rates (CAGR) for the high and low F-score portfolios and also for the market over the sample period 1992–2014. It also shows the annual returns of each traditional investment strategy that consists of all the top 20% of stocks based on each ratio (e.g., high B/M strategy). Panel A of Table 2 provides several interesting results when applying the F-score screening method to the various value investment strategies in the European stock markets. First, consistent with Loughran and Wellman (2011) and Gray and Vogel (2012) from the US markets, our results show that without applying the F-score screening, the EBIT/EV strategy yields the highest returns (17.42%) followed by the EBITDA/EV strategy (17.40%), while, interestingly, the Novy-Marx profitability strategy yields the lowest returns (12.24%). Second, and more importantly, applying the F-score screening method improves the returns for all strategies investigated. Consistent with the traditional portfolios, the best performance of the high F-score portfolios is obtained for the EBIT/EV and EBITDA/EV strategies with returns of 19.28% and 19.62%, respectively. Even though the F-score also works for the Novy-Marx's profitability strategy, suggesting that the screening also adds value to the quality-like investment strategy, the high F-score portfolio again yields the lowest returns (14.66%). The greatest difference between the traditional value

[9] The factors for the European financial markets are publicly available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Journal : Large 41260 | Dispatch : 21-11-2018 | Pages : 12
Article No. : 98 | ☐ LE | ☐ TYPESET
MS Code : JASE-D-18-00030 | ✔ CP | ✔ DISK

Table 2 Returns of portfolios

	B/M	E/M	D/M	EBIT/EV	EBITDA/EV	Novy-Marx
Panel A: annual returns						
Hi	18.51%	18.94%	18.35%	19.28%	19.62%	14.66%
Lo	6.50%	1.61%	4.92%	8.00%	11.62%	6.44%
All	14.72%	17.14%	14.44%	17.42%	17.40%	12.24%
Market	9.12%	9.12%	9.12%	9.12%	9.12%	9.12%
Hi-Lo	12.01%	17.33%	13.43%	11.28%	8.00%	8.22%
Hi-All	3.79%	1.80%	3.91%	1.86%	2.22%	2.42%
Hi-Market	9.39%	9.82%	9.23%	10.16%	10.49%	5.54%
Panel B: number of companies						
Hi	3767	5935	3393	6103	5945	4751
Lo	1662	227	1070	272	290	840
All	12,272	12,276	11,977	12,279	12,255	12,273
Panel C: Fama–French 5F Alphas						
Hi	6.22%***	6.16%***	6.22%***	6.78%***	7.44%***	5.47%***
t-stat	(3.757)	(4.269)	(4.019)	(4.529)	(4.944)	(4.154)
Lo	− 4.65%**	− 10.50%**	− 6.68%***	− 2.92%	− 2.04%	1.29%
t-stat	(− 2.007)	(− 2.654)	(− 2.815)	(− 0.788)	(− 0.575)	(0.318)
All	2.74%*	4.42%***	2.20%	5.17%***	5.34%***	3.89%**
t-stat	(1.938)	(3.050)	(1.533)	(3.727)	(3.879)	(2.740)
Panel D: alpha spreads						
Hi-Lo	10.87%***	16.65%***	12.90%***	9.69%***	9.48%***	4.18%
t-stat	(3.820)	(3.956)	(4.552)	(2.426)	(2.460)	(0.977)
Hi-All	3.48%	1.74%	4.01%*	1.61%	2.10%	1.58%
t-stat	(1.599)	(0.852)	(1.902)	(0.788)	(1.029)	(0.816)

Panel A in the table reports compound-annual growth rates of each investigated portfolio. *Hi (lo)* refers to a high (low) F-score portfolio, which consists of stocks that receive a score between 7 and 9 (1 and 3) at the time of portfolio formation. *All* refers to the portfolio without applying the F-score screening (i.e., it includes the top 20% of stocks based on each ratio). Panel B reports company-year observations of different portfolios between June 1992 and June 2014. Panel C of the table presents the intercepts (abnormal returns) and their t-statistics (in parentheses) of the Fama and French (2015) five-factor model. Monthly return series are used in the regression calculations but reported numbers are annualized. Finally, alpha spreads are presented in panel D. Test-statistics for alpha spreads are calculated as presented in Eq. 4 of this paper. Returns are reported in US dollars. The portfolios are formed on the last trading day of June each year and each portfolio is held for 1 year

*, ** and ***Indicate statistical significance at the 10%, 5% and 1% levels, respectively

portfolio (14.44%) and the high F-score portfolio (18.35%) is found for the D/M strategy, implying that the F-score screening method seems to be particularly useful for that strategy. A similar level of benefit is found for the B/M strategy (3.79%). Finally, the greatest differences between the high and low F-score portfolios are found for E/M and D/M strategies with return spreads of 17.33% and 13.43%, respectively, implying that the screening method seems to have the best ability to identify the returns for these high and low F-score portfolios.

Next, we examine whether the returns can be explained through the market, size, value, profitability and investment factors of the Fama and French (2015) five-factor model, which has not been previously examined. Panel C of Table 2 reports the abnormal returns for different portfolios; and it is clear that the abnormal returns of all high F-score portfolios are positive and statistically highly significant, demonstrating that the five-factor model is not fully able to explain the high returns. The greatest abnormal return can be achieved with the F-score screening applied with the EBITDA/EV strategy followed by the EBIT/EV, B/M, and D/M strategies. For example, the abnormal return of the high (low) F-score EBITDA/EV portfolio is 7.44% (− 2.04%), while the abnormal return of the portfolio containing the top 20% of EBITDA/EV stocks is 5.34%. The results are somewhat similar in the case of other strategies. Similar to the results in Panel A, the greatest differences between the alphas of the high score and traditional portfolios can be found for the B/M and D/M strategies, as the alpha spreads are 3.48% and 4.01%,

Journal : Large 41260 Dispatch : 21-11-2018 Pages : 12
Article No. : 98 LE TYPESET
MS Code : JASE-D-18-00030 CP DISK

Table 3 Loadings of the Fama and French (2015) five-factor model

	B/M	E/M	D/M	EBIT/EV	EBITDA/EV	Novy-Marx
Panel A: high F-score portfolios						
Alpha (annual)	6.22%***	6.16%***	6.22%***	6.78%***	7.44%***	5.47%***
	(3.757)	(4.269)	(4.019)	(4.529)	(4.944)	(4.154)
Market	0.725***	0.725***	0.693***	0.722***	0.703***	0.756***
	(29.882)	(24.642)	(22.839)	(23.631)	(26.029)	(33.815)
SMB	0.682***	0.646***	0.600***	0.628***	0.614***	0.622***
	(12.808)	(12.722)	(10.034)	(12.122)	(12.460)	(14.250)
HML	0.345***	0.354***	0.275***	0.279***	0.286***	− 0.072
	(4.013)	(5.275)	(3.793)	(3.910)	(3.817)	(− 1.119)
RMW	0.268***	0.409***	0.353***	0.413***	0.358***	0.211***
	(3.378)	(4.954)	(4.523)	(4.973)	(4.390)	(2.678)
CMA	0.093	0.037	0.155*	0.071	0.042	0.024
	(0.995)	(0.469)	(1.878)	(0.891)	(0.497)	(0.308)
Adj. R-sqr.	84.44%	86.20%	80.80%	84.38%	84.91%	88.03%
Panel B: low F-score portfolios						
Alpha (annual)	− 4.65%**	− 10.50%***	− 6.68%***	− 2.92%	− 2.04%	1.29%
	(− 2.007)	(− 2.654)	(− 2.815)	(− 0.788)	(− 0.575)	(0.318)
Market	0.936***	0.881***	0.950***	0.900***	0.996***	0.929***
	(24.560)	(11.788)	(21.742)	(14.286)	(12.620)	(12.285)
SMB	0.779***	0.860***	0.716***	0.623***	0.636***	0.835***
	(8.637)	(6.707)	(8.001)	(3.881)	(5.539)	(5.489)
HML	0.237**	0.187	0.093	0.087	0.154	− 0.578**
	(1.968)	(0.984)	(0.807)	(0.641)	(1.015)	(− 2.305)
RMW	− 0.022	0.409*	0.176	0.032	0.426	− 0.195
	(− 0.147)	(1.758)	(1.423)	(0.171)	(1.493)	(− 0.733)
CMA	0.311**	0.310	0.501***	0.537***	0.502**	0.003
	(2.019)	(1.387)	(3.214)	(2.649)	(2.370)	(0.012)
Adj. R-sqr.	77.73%	55.27%	78.80%	56.34%	53.75%	61.47%

The table reports the annualized abnormal returns and the factor loadings of the Fama and French (2015) five-factor model for each investigated portfolio. Panel A (B) reports the results for high (low) F-score portfolios that consist of stocks that received the scores between seven and nine (one and three) at the time of the portfolio formation. Returns are reported in US dollars

*, ** and ***Indicate statistical significance at the 10, 5, and 1% levels, respectively

respectively, while the greatest differences between the alphas of high- and low-score portfolios can be found for the E/M and D/M strategies, as the alpha spreads are 16.65% and 12.90%, respectively. Overall, these results show that the high returns of the high F-score portfolios cannot be explained by the Fama and French (2015) factors.

The results presented in Table 3 further show that the loadings of the risk factors vary between high- and low-score portfolios, implying that the portfolios have different risk characteristics. For example, the low F-score portfolios have a higher loading for market risk than do the high F-score portfolios. The loadings for the market risk vary between 0.881 and 0.996 (0.693 and 0.756) for the low (high) F-score portfolios. Similarly, the low-score portfolios seem to have higher loadings for the size factor than high score portfolios. The high F-score portfolios have a statistically significant and positive loading for the profitability (RMW) factor, suggesting that the high F-score serves as an indicator of good future profitability, which is similar to the results reported in Fama and French (2006). The low F-score portfolios, on the other hand, have a mostly insignificant or negative loading for the RMW factor. In addition, low-score portfolios have a positive loading for the investment (CMA) factor, broadly suggesting that financially unhealthy companies make conservative investments. This loading also explains the insignificance of the HML factor loading implying that HML is redundant to CMA, which is consistent with Fama and French (2015). Finally, with the exception of the

Journal : Large 41260 Dispatch : 21-11-2018 Pages : 12
Article No. : 98 ☐ LE ☐ TYPESET
MS Code : JASE-D-18-00030 ✔ CP ✔ DISK

Table 4 Sharpe and Sortino ratios

	B/M	E/M	D/M	EBIT/EV	EBITDA/EV	Novy-Marx
Panel A: sortino ratios						
Hi	1.473	1.501	1.573	1.581	1.665	1.087
Lo	0.235	− 0.067	0.139	0.310	0.533	0.218
All	0.982	1.266	1.019	1.344	1.365	0.789
Market	0.440	0.440	0.440	0.440	0.440	0.440
Hi-Lo	1.238	1.568	1.433	1.271	1.132	0.869
Hi-All	0.491	0.235	0.553	0.237	0.300	0.298
Hi-Market	1.033	1.061	1.132	1.141	1.225	0.647
Panel B: sharpe ratios						
Hi	0.951	1.004	1.002	1.041	1.082	0.750
Lo	0.169	− 0.050	0.103	0.224	0.347	0.152
All	0.659	0.861	0.688	0.903	0.909	0.557
Market	0.322	0.322	0.322	0.322	0.322	0.322
Hi-Lo	0.782***	1.055***	0.899***	0.817***	0.735**	0.598**
p value	(0.000)	(0.003)	(0.000)	(0.008)	(0.024)	(0.031)
Hi-All	0.292***	0.144***	0.314***	0.138***	0.173***	0.193***
p value	(0.000)	(0.002)	(0.001)	(0.001)	(0.000)	(0.003)
Hi-Market	0.629***	0.682***	0.680***	0.719***	0.760***	0.428***
p value	(0.008)	(0.008)	(0.007)	(0.004)	(0.002)	(0.007)

Panels A and B in the Table 3 present the annualized Sortino and Sharpe ratios of each portfolio investigated. Sortino and Sharpe ratios are calculated using monthly returns over the full sample period (June 1992 and June 2014). *Hi (lo)* refers to a high (low) F-score portfolio, which consists of stocks that receive a score between 7 and 9 (1 and 3) at the time of portfolio formation. *All* refers to the portfolio without applying the F-score screening (i.e. it includes the top 20% of stocks based on each ratio). The table also shows the differences in the Sharpe ratios and *p* values (in parentheses). The *p* values are obtained using Ledoit and Wolf (2008) test. This test is based on the circular block bootstrap method. See "Risk-adjusted performance measures" section for further information

*, ** and ***Indicate statistical significance at 10%, 5% and 1% levels, respectively

Novy-Marx profitability strategy, all other portfolios have positive loadings for the value factor varying between 0.271 and 0.371. The negative loading of the profitability strategy for the value factor is also consistent with Novy-Marx (2014). In summary, we can conclude that the F-score screening can be used to modify the risk characteristics of the traditional portfolios.

Next, we report the Sharpe and Sortino ratios for each portfolio. Panel A in Table 4 shows that the Sortino ratios for the high F-score portfolios are consistently higher than those for the market, portfolios without screening and the low F-score portfolios. For example, the Sortino ratio for the high F-score EBITDA/EV portfolio is 1.665, while it is, respectively, 0.533 and 1.365 for the low F-score portfolio and the EBITDA/EV portfolio without applying screening. The greatest differences between the high and traditional (high and low score) portfolios can be found for the D/M (E/M) strategy. Similar to the results reported in Table 2, the high F-score portfolio of the EBITDA/EV strategy outperforms all other strategies, also when measured by the Sortino ratio.

Panel B in Table 4 shows the Sharpe ratios for different portfolios formed. Again, the outperformance of the high F-score portfolios is apparent, providing further evidence that the F-score screening is a useful tool for investors. In absolute terms, the high F-score portfolio of the EBITDA/EV strategy has the highest mean Sharpe ratio of 1.082, while the high F-score portfolio of the Novy-Marx strategy has the lowest Sharpe ratio. Overall, using the F-score screening methods provides the greatest improvement for the traditional D/M and B/M portfolios as measured by the Sharpe ratio.

We also test whether the differences in the Sharpe ratios are statistically significant using the Ledoit and Wolf (2008) test, which is based on the circular block bootstrap method. Panel B in Table 4 shows that all the differences in the Sharpe ratios are statistically significant at the 5% of level. These results suggest that the high F-score portfolios offer a higher excess return per unit of total risk than other portfolios.

Finally, we report various worst-case scenarios (i.e., various drawdowns) for each portfolio formed, as these could also be considered different forms of risk. Panel A in

Journal : Large 41260 | Dispatch : 21-11-2018 | Pages : 12
Article No. : 98 | □ LE | □ TYPESET
MS Code : JASE-D-18-00030 | ✔ CP | ✔ DISK

Table 5 Drawdowns

	B/M (%)	E/M (%)	D/M (%)	EBIT/EV (%)	EBITDA/EV (%)	Novy-Marx (%)
Panel A: worst monthly						
Hi	− 20.81	− 22.97	− 20.96	− 22.04	− 21.00	− 20.69
Lo	− 25.84	− 31.45	− 25.55	− 27.75	− 29.35	− 20.14
All	− 23.86	− 23.95	− 23.34	− 22.69	− 22.58	− 21.30
Market	− 24.24	− 24.24	− 24.24	− 24.24	− 24.24	− 24.24
Panel B: worst 12 Month						
Hi	− 32.36	− 32.71	− 30.45	− 29.82	− 28.24	− 26.06
Lo	− 49.38	− 36.31	− 42.28	− 34.52	− 33.78	− 31.48
All	− 36.06	− 33.83	− 32.86	− 30.40	− 29.57	− 28.54
Market	− 28.05	− 28.05	− 28.05	− 28.05	− 28.05	− 28.05
Panel C: worst drawdown						
Hi	− 55.53	− 55.26	− 51.51	− 52.56	− 50.29	− 54.14
Lo	− 73.34	− 60.73	− 67.86	− 63.38	− 70.38	− 69.90
All	− 60.66	− 56.84	− 57.10	− 54.42	− 53.53	− 57.35
Market	− 62.37	− 62.37	− 62.37	− 62.37	− 62.37	− 62.37

Panel A in the table presents the worst 1-month returns of each portfolio investigated. *Hi (lo)* refers to a high (low) F-score portfolio, which consists of stocks that receive a score between 7 and 9 (1 and 3) at the time of portfolio formation. *All* refers to the portfolio without applying the F-score screening (i.e., it includes the top 20% of stocks based on each ratio). Panel B presents the worst annual returns (1-year returns) of different portfolios. Finally, Panel C in the table shows the maximum drawdown of each portfolio. Returns are reported in US dollars

Table 5 presents the worst monthly return of each portfolio. It is clear that, with the exception of the Novy-Marx profitability strategy, all high F-score portfolios manage to outperform the other portfolios. The results are very similar when the worst annual returns and maximum drawdowns are compared. If the worst-case scenario (maximum drawdown) is interpreted as a form of risk, the results indicate that the high F-score portfolios have been less risky than other portfolios.[10]

Performance of high and low F-score portfolios of different sizes

An important issue is whether the high returns of high F-score portfolios are concentrated mostly on small companies, as suggested by Piotroski (2000), or if it is a useful screening method for companies of all sizes. Therefore, the study also investigates the returns of different-sized portfolios. For the classification of stocks based on size, please see "The F-score and portfolio formation" section.

Table 6 presents the returns of different-sized portfolios. Panel A of Table 6 shows compound-annual growth rates of portfolios containing large stocks. Clearly, the high F-score portfolios outperform the low F-score portfolios and those portfolios without F-score screening demonstrating that it also works for large stocks. Interestingly, the best annual returns are achieved by applying the F-score screening method to the B/M strategy, while the high F-score portfolio for the EBITDA/EV strategy loses its superiority among large stocks. A high F-score portfolio with the B/M strategy applied produces a return of 19.90%, while the return of the low F-score (traditional) portfolio is 7.35% (16.17%). However, as the sample size for the low F-score portfolios is quite small, the results of the low F-score portfolios should be interpreted with caution.

Panel B of Table 6 presents the results of portfolios containing medium-sized companies. High F-score portfolios outperform all other portfolios across all investment strategies studied. The highest returns are achieved with the high F-score portfolio of the EBITDA/EV strategy (19.86%), while the best improvement compared to the traditional portfolio is found for the B/M and D/M strategies. Next, panel C of Table 6 shows the returns of small stock portfolios. The high F-score portfolio of the EBIT/EV strategy yields the highest returns (20.65%), while the best improvement of 4.48% is achieved for the D/M

[10] Overall, the results discussed in "Returns of the portfolios" section remain qualitatively the same in the face of alternative screening approaches, where high (low) F-score portfolios consist of stocks that receive a score between 8 and 9 (1 and 2) at the time of portfolio formation. In our further robustness analysis we show that the results regarding the outperformance of high F-score portfolios over other portfolios also remain the same in different market states, i.e., bull and bear markets as defined by Lunde and Timmermann (2004). For the sake of brevity, we do not report these results but they are available from the authors on request.

Journal : Large 41260 | Dispatch : 21-11-2018 | Pages : 12
Article No. : 98 | ☐ LE | ☐ TYPESET
MS Code : JASE-D-18-00030 | ☑ CP | ☑ DISK

Table 6 Returns of different size portfolios

	B/M	E/M	D/M	EBIT/EV	EBITDA/EV	Novy-Marx
Panel A: large						
Hi	19.90%	17.27%	15.58%	16.28%	16.82%	14.29%
Lo	7.35%	0.34%	5.08%	6.25%	9.27%	4.14%
All	16.17%	16.13%	14.39%	15.40%	15.37%	10.42%
Market	9.42%	9.42%	9.42%	9.42%	9.42%	9.42%
# of companies						
Hi	822	1018	590	1034	974	789
Lo	228	41	143	31	39	45
All	2708	2450	2213	2331	2171	1837
Panel B: medium						
Hi	17.51%	18.04%	17.57%	18.22%	19.86%	13.31%
Lo	3.79%	5.35%	4.01%	10.39%	7.63%	3.79%
All	13.69%	17.32%	13.87%	17.28%	18.33%	11.09%
Market	9.42%	9.42%	9.42%	9.42%	9.42%	9.42%
# of companies						
Hi	1401	1672	949	1795	1717	1353
Lo	536	71	235	59	64	118
All	4493	3650	3468	3668	3599	3422
Panel C: small						
Hi	17.58%	19.52%	19.38%	20.65%	19.98%	15.13%
Lo	6.03%	− 0.36%	3.46%	6.31%	8.54%	7.35%
All	14.32%	17.28%	14.90%	18.56%	17.77%	13.16%
Market	9.42%	9.42%	9.42%	9.42%	9.42%	9.42%
# of companies						
Hi	1544	3245	1854	3274	3254	2609
Lo	898	115	692	182	187	677
All	5071	6176	6296	6280	6485	7014

Panel A in the table reports the compound-annual growth rates of large stock portfolios. *Hi (lo)* refers to a high (low) F-score portfolio, which consists of stocks that receive a score between 7 and 9 (1 and 3) at the time of portfolio formation. *All* refers to the portfolio without applying the F-score screening (i.e., it includes the top 20% of stocks based on each ratio). Panels B and C present these results for medium and small stocks. Returns are reported in US dollars

strategy. Overall, the returns for small stocks are higher than for medium-sized and large stocks. These findings also suggest that although the high F-score screening method is very useful for large companies, it is more profitable for small companies, which concurs with Piotroski (2000).

Conclusions

This study investigates the suitability of Piotroski's (2000) composite score analysis for use in separating the winners from the losers among several value investment strategies that use different valuation ratios to form portfolios, such as E/M, D/M and EBITDA/EV ratios. To verify that the returns of the formed portfolios are not achieved by taking more risk, we use an extensive set risk and performance measures, such as various drawdowns, Sortino and Sharpe ratios, as well as the Fama and French (2015) five-factor adjusted abnormal returns. We also perform various additional tests, such as controlling for the size of the companies, to confirm the robustness of the results.

The results reported in the current study show that Piotroski's (2000) screening method improves all investigated value investment strategies in the European stock markets. The best performance is achieved when applying the F-score screening method to the EBITDA/EV strategy, while the best improvement compared to the traditional investment strategy is achieved B/M and D/M strategies, suggesting that it is particularly useful for these strategies. In addition to the enhanced performance of all strategies investigated, the F-score screening can also be used to modify the risk characteristics of the traditional value portfolios. Overall, the results of the study demonstrate the

F-score provides useful information for professional portfolio managers and individual investors alike. Future research could potentially examine the time-varying performance of the high and low F-score portfolios and their responses to macroeconomic and other shocks (see, e.g., Cakici and Tan 2014).

References

Asness, C., T. Moskowitz, and L. Pedersen. 2013. Value and Momentum Everywhere. *The Journal of Finance* 68 (3): 929–985.

Beaver, W., M. McNichols, and R. Price. 2007. Delisting Returns and Their Effect on Accounting-Based Market Anomalies. *Journal of Accounting and Economics* 43 (2–3): 341–368.

Broussard, J., J. Mikkonen, and V. Puttonen. 2016. Style Migration in Europe. *European Financial Management* 22: 797–816.

Cakici, N., and S. Tan. 2014. Size, Value, and Momentum in Developed Country Equity Returns: Macroeconomic and Liquidity Exposures. *Journal of International Money and Finance* 44: 179–209.

Campbell, J., J. Hilscher, and J. Szilagyi. 2008. In Search of Distress Risk. *The Journal of Finance* 63 (6): 2899–2939.

Chan, L., and J. Lakonishok. 2004. Value and Growth Investing: Review and Update. *Financial Analysts Journal* 60 (1): 71–86.

Choi, N., and R. Sias. 2012. Why Does Financial Strength Forecast Stock Returns? Evidence from Subsequent Demand by Institutional Investors. *Review of Financial Studies* 25: 1550–1587.

Fama, E., and K. French. 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance* 47 (2): 427–465.

Fama, E., and K. French. 1996. Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance* 51: 55–84.

Fama, E., and K. French. 2006. Profitability, Investment and Average Returns. *Journal of Financial Economics* 82: 491–518.

Fama, E., and K. French. 2012. Size, Value, and Momentum in International Stock Returns. *Journal of Financial Economics* 105: 457–472.

Fama, E., and K. French. 2015. A Five-Factor Asset Pricing Model. *Journal of Financial Economics* 116: 1–22.

Goetzmann, W., J. Ingersoll, and M. Spiegel. 2007. Portfolio Performance Manipulation and Manipulation-proof Performance Measures. *The Review of Financial Studies* 20 (5): 1503–1546.

Gray, W., and J. Vogel. 2012. Analyzing Valuation Measures: A Performance Horse Race over the Past 40 Years. *Journal of Portfolio Management* 39 (1): 112–121.

Griffin, J., and M. Lemmon. 2002. Book-to-Market Equity, Distress Risk and Stock Returns. *Journal of Finance* 57: 2317–2336.

Ince, O., and R. Porter. 2006. Individual Equity Return Data from Thomson Datastream: Handle with Care! *Journal of Financial Research* 29 (4): 463–479.

Israelsen, C.L. 2005. A Refinement to the Sharpe Ratio and Information Ratio. *Journal of Asset Management* 5: 423–427.

Kapadia, N. 2011. Tracking Down Distress Risk. *Journal of Financial Economics* 102 (1): 167–182.

La Porta, R. 1996. Expectations and the Cross-section of Stock Returns. *The Journal of Finance* 51 (5): 1715–1742.

Lakonishok, J., A. Shleifer, and R. Vishny. 1994. Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance* 49 (5): 1541–1578.

Ledoit, O., and M. Wolf. 2008. Robust Performance Hypothesis Testing with the Sharpe Ratio. *Journal of Empirical Finance* 15 (5): 850–859.

Loughran, T., and J. Wellman. 2011. New Evidence on the Relation Between the Enterprise Multiple and Average Stock Returns. *Journal of Financial and Quantitative Analysis* 46: 1629–1650.

Lunde, A., and A. Timmermann. 2004. Duration Dependence in Stock Prices. *Journal of Business and Economic Statistics* 22 (3): 253–273.

Newey, W., and K. West. 1987. A Simple Semidefinite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55 (3): 703–708.

Novy-Marx, R. 2013. The Other Side of Value: The Gross Profitability Premium. *Journal of Financial Economics* 108 (1): 1–28.

Novy-Marx, R. 2014. Quality Investing. *Working Paper*.

Piotroski, J. 2000. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. *Journal of Accounting Research* 38: 1–41.

Piotroski, J., and E. So. 2012. Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach. *Review of Financial Studies* 25: 2841–2875.

Plyakha, Y., R. Uppal, and G. Vilkov. 2012. Why Does an Equal-Weighted Portfolio Outperform Value- and Price-Weighted Portfolios. *Working Paper*, Goethe University Frankfurt.

Sharpe, W. 1966. Mutual Fund Performance. *The Journal of Business* 39 (1): 119–138.

Sortino, F., and L. Price. 1994. Performance Measurement in a Downside Risk Framework. *The Journal of Investing* 3 (3): 59–64.

Sortino, F., and R. van der Meer. 1991. Downside risk. *The Journal of Portfolio Management* 17 (4): 27–31.

Turtle, H., and K. Wang. 2017. The Value in Fundamental Accounting Information. *The Journal of Financial Research* 40 (1): 113–140.

Vassalou, M., and Y. Xing. 2004. Default Risk in Equity Returns. *Journal of Finance* 59: 831–868.

Walkshäusl, C. 2017. Expectation Errors in European Value-Growth Strategies. *Review of Finance* 21 (2): 845–870.

Jarno Tikkanen is a portfolio analyst at the OP Wealth Management, Finland. He holds a Master's Degree in Finance and has worked for several years as a portfolio analyst. In addition, he has published a research article on portfolio performance and value strategies.

Janne Äijö is a Professor of Accounting and Finance at the University of Vaasa, Finland. His research interests include financial markets research, such as investment strategies, financial market anomalies, option-implied information and currency markets. His articles have appeared in the Journal of Futures Markets, International Review of Financial Analysis, Quantitative Finance, Research in International Business and Finance, Global Finance Journal and Finance Research Letters.

Journal : **41260**
Article : **98**

Author Proof

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details Required	Author's Response
AQ1	Please check and confirm that the authors and their respective affiliations have been correctly identified and amend if necessary.	
AQ2	Please check and confirm the corresponding author is correctly identified and amend if necessary.	
AQ3	Please check and confirm the organisation name, city and country name are correctly identified and amend if necessary.	
AQ4	References Fama et al. (1992, 2012) and Stambaugh et al. (2012) are cited in text but not provided in the reference list. Please provide references in the list or delete these citations.	
AQ5	Reference Sortino and Price (1994) is given in list but not cited in text. Please cite in text or delete from list.	