# Simple Financial Analysis and Abnormal Stock Returns - Analysis of Piotroski's Investment Strategy

Hauke Rathjens* and Hendrik Schellhove†

Master Thesis in Accounting and Financial Management
at the Stockholm School of Economics

**Abstract**

We investigate (1) whether Piotroski's (2000) simple financial statement analysis can successfully be applied to the UK market in a more recent time period (1991-2008) and (2) whether the observed return patterns indicate abnormal returns. Piotroski shows that his strategy increases market-adjusted returns by 7.5 percentage points annually and that shorting expected losers and buying expected winners generates an average 23% annual return within a value stock portfolio in the US between 1976 and 1996. We find that the strategy is also successful when applied to the UK market as a whole. In the growth stock portfolio alone, shorting expecting losers and buying expected winners generates an average market-adjusted return of 13.8% and a 9.6 percentage points higher return compared to the entire growth stock portfolio. However, in contrast to Piotroski, we do not find that the strategy generates any significant returns in the value stock portfolio in the UK. In addition to his study, our study demonstrates that the results persist after adjusting returns with risk characteristic-matched portfolio returns, that the strategy explains future returns for the entire market after controlling for known risk variables, and that the risk-adjusted returns do not decrease over time. Overall, the findings suggest that an investor, using simple financial analysis, could have systematically earned abnormal returns not explicable by common risk factors.

**Keywords:** Simple Financial Statement Analysis, Abnormal Returns, Risk Adjustment, Piotroski, Market Efficiency

**Tutor:** Kenth Skogsvik[1]
**Date:** 2011-05-16

*40110@student.hhs.se
†40128@student.hhs.se
[1]The authors would like to thank Kenth Skogsvik for his valuable guidance and support throughout the research period.

# Contents

**Acronyms** **II**

**List of Tables** **III**

**1 Introduction** **1**

**2 Previous Research** **4**
2.1 The Efficient Market Hypothesis . . . . . 4
2.2 The Concept of Risk-Compensation . . . . . 5
2.3 Fundamental Investing . . . . . 7
2.4 Value Investing and the Book-to-Market Effect . . . . . 9

**3 Piotroski's Investment Strategy** **12**
3.1 The Investment Idea and the F_SCORE . . . . . 12
3.2 Sample Selection, Methodology, and Empirical Results . . . . . 13
3.3 Performed Tests . . . . . 15
3.4 Follow-Up Study . . . . . 18

**4 Limitations of Piotroski's Study** **20**

**5 Applying Piotroski to the UK Market** **22**
5.1 Sample Selection and Methodology . . . . . 22
5.1.1 Sample Selection . . . . . 22
5.1.2 B/M Computation . . . . . 23
5.1.3 F_SCORE Computation . . . . . 24
5.1.4 Return Computation . . . . . 26
5.2 Empirical Results . . . . . 27
5.2.1 Descriptive Statistics . . . . . 29
5.2.2 Returns Conditioned on B/M . . . . . 29
5.2.3 Returns Conditioned on Size . . . . . 34
5.3 Analysis of Empirical Results . . . . . 34
5.3.1 Value Stock Portfolio . . . . . 36
5.3.2 Growth Stock Portfolio . . . . . 38

**6 Tests for Abnormal Returns** **40**
6.1 Asset Pricing Models . . . . . 40
6.2 Characteristic-Matched Returns . . . . . 42
6.2.1 Methodology . . . . . 43
6.2.2 Empirical Results . . . . . 44
6.3 Regression Analysis . . . . . 46
6.3.1 Methodology . . . . . 47
6.3.2 Empirical Results . . . . . 49
6.4 Returns Over Time . . . . . 51

**7 Conclusion** **54**

**References** **56**

## Acronyms

**Amex** American Stock Exchange

**ARCA** Archipelago Exchange

**B/M** Book-to-Market

**CAPM** Capital Asset Pricing Model

**CB** Characteristic-Balanced

**CFO** Cash Flow from Operations

**CRSP** Center for Research in Security Prices

**DS** Thomson Reuters Datastream

**FB** Factor-Balanced

**FTSE** Financial Times Stock Exchange

**HML** High-Minus-Low

**IPO** Initial Public Offering

**LSE** London Stock Exchange

**NYSE** New York Stock Exchange

**RI** 'Total Return Index'

**ROE** Return on Equity

**SEO** Seasoned Equity Offering

**SMB** Small-Minus-Big

**WS** Worldscope

## List of Tables

1 Definitions of F_SCORE Variables . . . . . . . . . . . . . . . . . . . . . . . 25
2 Descriptive Statistics of Sample Firms . . . . . . . . . . . . . . . . . . . . . 28
3 Buy-and-Hold Raw Returns Across B/M Quintiles . . . . . . . . . . . . . . 30
4 Buy-and-Hold Market-Adjusted Returns Across B/M Quintiles . . . . . . . 31
5 Buy-and-Hold Raw Return Distribution . . . . . . . . . . . . . . . . . . . . 33
6 Buy-and-Hold Raw Returns Across Size Terciles . . . . . . . . . . . . . . . 35
7 Correlation Analysis of Returns and F_SCORE Indicator Variables . . . . . 36
8 Buy-and-Hold Characteristic-Matched Returns Across B/M Quintiles . . . 45
9 Regression Analysis of Individual Firm-Year Observations . . . . . . . . . . 50
10 Buy-and-Hold Characteristic-Matched Returns Over Three Time Periods . 52
11 Buy-and-Hold Characteristic-Matched Returns Across Time . . . . . . . . . 53

## List of Figures

1 Overview of Piotroski's Investment Strategy . . . . . . . . . . . . . . . . . . . 14

# 1 Introduction

In an efficient stock market in the semi-strong form, as defined in Fama (1970), the stock price of any listed firm would timely adjust to the publication of new value-relevant information, such as annual financial statements or press releases, and have all the information from historical stock prices already incorporated. In this world there are no unexploited profit opportunities and, assuming that investors want to be compensated for taking risks, investors always have to make riskier investments in order to achieve higher returns (e.g. Sharpe, 1964).

However, several market anomalies have been observed by researchers that are not in line with the market efficiency hypothesis, such as the post-earnings announcement drift (Bernard & Thomas, 1989). These anomalies indicate that stock prices might at least temporarily divert from their true, fundamental value. Consequently, practitioners and researchers have tried to figure out investment strategies that would earn investors superior returns without merely increasing their risk exposure.

One well-studied strategy is fundamental investing which builds upon Ball and Brown's (1968) revelation that accounting information is value-relevant. Fundamental investors try to predict future earnings or returns of a firm based on the analysis of its available accounting information and estimate its fundamental value. If the current market price of the stock is higher (lower) than its fundamental value, a short (long) position in the stock is taken. The fundamental strategies show superior returns (e.g. Ou & Penman, 1989), but since these strategies often use very complex, time-consuming, and hence costly statistical methods, they are criticised for not being applicable in practice (e.g. Ball, 1992). Others argue that they create new information previously not available to the markets and that the observed higher returns compensate for gaining proprietary knowledge (e.g. Foster, 1979).

Another well-known strategy is value investing which aims at identifying stocks temporarily undervalued. Value investments are often identified by the ratio of the book value to the market value of equity, the Book-to-Market (B/M) ratio. Value investors buy stocks with a high B/M ratio (value stocks) and sell low B/M shares (growth stocks). The value strategy builds upon the assumption that the fundamental values of stocks are measurable and that some market prices currently deviate from their fundamental values. Research has shown that value stocks have historically outperformed growth stocks in numerous stock markets worldwide over long periods of time (e.g. Haugen, 2009). The existence of this so-called B/M effect is not disputed in academics. Nonetheless, the proponents of market efficiency attribute this effect to higher risk, assuming that a high B/M ratio proxies for higher bankruptcy probability and lower liquidity of value stocks (e.g. Fama & French, 1992). In contrast, opponents of the market efficiency view believe that it results from overpessimism or short-term disinterest for value stocks (e.g. Lakonishok, Shleifer, & Vishny, 1994).

Piotroski (2000) combines both strategies, fundamental and value investing, in his study. He formulates a simple fundamental strategy based on easily observable accounting numbers in order to avoid practical implementation problems associated with too sophisticated analyses. Then, he applies this fundamental strategy to the value stock portfolio, consisting of all firms in the highest B/M quintile. In this portfolio he identifies those firms that have the strongest financial position and should therefore have the lowest risk of all value stocks. Surprisingly, investing only in the healthiest, presumably low-risk value stocks leads to high market-adjusted returns. Thus, he claims that this strategy can increase returns without increasing the risk exposure.

Piotroski's (2000) claim implies that markets are inefficient and that investors can earn a return exceeding an appropriate risk compensation (abnormal return) with fairly simple financial analyses. However, this assertion is not supported by a more detailed investigation whether the realised returns are abnormal. Furthermore, he tests his strategy only on a sample of the US market and limits it to the value stock portfolio. Since there is solely one observable pattern in historical stock price information, the strategy may have been merely a fortuitous observation. The conclusions about abnormal returns and market efficiency are therefore limited.

In this study we address the above mentioned limitations and investigate whether *applying Piotroski's (2000) simple financial analysis generates abnormal stock returns.* We replicate his investment strategy on the whole UK market between 1991 and 2008 and evaluate whether the observed returns are abnormal and continuous. Thereby we address the limitations of his study and contribute (1) to previous research by providing further evidence on the question if markets are efficient and (2) to investment practice by investigating if practitioners can really generate higher returns without incurring higher risks.

We find that his strategy is successful when applied to the UK market as a whole. In the growth stock portfolio alone, shorting expecting losers and buying expected winners generates an average market-adjusted return of 13.8% and a 9.6 percentage points higher return compared to the entire growth stock portfolio. However, in contrast to Piotroski, we do not find that the strategy generates any significant returns in the value stock portfolio. In addition to his study, our study demonstrates that the results persist after adjusting the returns with risk characteristic-matched portfolio returns. In the entire market expected winners still outperform expected losers by 9.5 percentage points. Moreover, the strategy explains future returns after controlling for known risk variables and the adjusted returns do not decrease over time.

This study proceeds as follows. Section two summarises the relevant previous literature followed by a description of Piotroski's (2000) study in section three. We explain the limitations of his study in section four and present the results from our replication on the UK market in section five. In section six we investigate whether observed returns are abnormal and continuous, and conclude in section seven.

# 2 Previous Research

The overview of the previous research is organised in four parts. At first, the efficient market hypothesis is presented and the concept of risk compensation is elaborated on. These aspects address Piotroski's (2000) claim that markets are inefficient and that investors can earn returns exceeding an appropriate risk compensation. Next, the development of the fundamental investment strategies is portrayed and the idea of value investing and the B/M effect are explained.

## 2.1 The Efficient Market Hypothesis

Building upon Fama (1965), Fama et al. (1969, p.1) define an efficient market as "a market that adjusts rapidly to new information". Thus, whenever new value-relevant information becomes available, it is immediately incorporated into the stock's price. Fama (1970) extends this definition by describing three different forms of market efficiency. The forms differ in their timely adjustment to different information subsets. First, in the weak form all information from past price histories is considered in market pricing. Second, in the semi-strong form also all other publicly available value-relevant information is incorporated in a timely manner. Thus, especially financial information from the firm's published annual reports and other corporate publications are utilised in the price determination. Third, in the strong form, in addition to all publicly, also all privately available value-relevant information is timely incorporated. Hereby privately available information is e.g. only accessible to institutional investors by exclusive access to management.

In most research articles, stock markets are assumed to be efficient in the semi-strong form. In this view investors cannot generate abnormal returns based on publicly available information, since all the available information has already been incorporated. Piotroski (2000) bases his investment strategy on simple analysis of publicly available financial statements. He claims that he generates returns above an appropriate risk compensation and consequently opposes the view of market efficiency in the semi-strong form.

In general, accounting-based investment strategies require that information from financial statements is value-relevant and that stock markets are temporarily inefficient. First, only if accounting information is value-relevant, it can be useful for investment decisions. Ball and Brown (1968) assess the value relevance of accounting income numbers and test whether their informational content is timely incorporated into stock prices. They find that accounting income numbers are value-relevant and that accounting statements do therefore constitute an important source of information. They also find that about 80% of the information content is already included in the stock prices on the announcement date of the annual report. Therefore, they conclude that the information is most likely earlier disseminated to the markets by other means, such as interim reports. This supports the view that markets react timely to new value-relevant information. Essentially, their study is the basis for all accounting-based investment strategies.

Second, if markets were inefficient in the long-run, investing in fundamentally under- or overvalued stocks would not yield superior returns as they would not revert back to their true values. Ball and Brown (1968) are also the first researchers observing temporary inefficiencies. They find that the stock prices drift in a foreseeable direction after the earnings announcement date. Bernard and Thomas (1989) examine this post-earnings announcement drift on US stock exchanges in more detail. They confirm the observation that there is a significant drift after the announcement of new earnings and cannot find a risk-based explanation for the drift. Hence, they conclude that stock markets seem to be temporarily inefficient after the earnings announcement, since the market prices only partially reflect the publicly available earnings information and deviate from their fundamental values. Bartov, Lindahl, and Ricks (1998) find another announcement drift in stock prices. They observe that it takes up to two years for the market to fully incorporate the information from asset write-offs into the stock price. Additionally, Chan, Jegadeesh, and Lakonishok (1996) show that the announcement drift to new information is more pronounced if the firm's past performance is contrary to the news.

To sum up, Piotroski (2000) opposes the view of market efficiency in the semi-strong form. Previous research has shown that financial statement information is value-relevant and that market prices might temporarily divert from fundamental values. Both observations are crucial for fundamental investment strategies, such as Piotroski's (2000).

### 2.2 The Concept of Risk-Compensation

To investigate whether fundamental investing can generate abnormal returns the concept of risk compensation is introduced. If observed stock returns exceed the expected or required returns, the returns are abnormal. In an efficient market with risk-averse investors the expected stock return is a direct function of the risk inherent in investing in the stock and abnormal returns are not possible. The investor can only increase returns by incurring additional risks (e.g. Sharpe, 1964). Ideally, the risk would be measured to estimate the expected and abnormal returns. However, since risk cannot be measured directly, different models approximating expected returns have been introduced over time. These models assume that markets are efficient and that the stock returns equal the expected returns on average.

The first model is the Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964) and Lintner (1965) building upon the work of portfolio formation by Markowitz (1952). The CAPM assumes that there are two types of risks, idiosyncratic and systematic. The idiosyncratic risk can be eliminated by investing in a large, diversified portfolio. An investor is therefore not compensated for taking this risk. The systematic risk is inherent to all stocks in a market and cannot be eliminated by diversification. Thus, in the CAPM the investor is compensated for an investment in a stock by the theoretical risk-free rate $r_f$ and by the additional market risk premium $r_m - r_f$ for taking the systematic risk. Depending on the stock return's sensitivity to changes in the returns of the whole market (co-variance risk), the investor expects to receive a higher or lower proportion of the risk premium, generally represented by the market beta $\beta$. The expected return *E(r)* for stock $i$ at time $t$ based on the available information $\Phi$ is:

$$\mathrm{E}(r_{i,t+1}|\Phi_t) = r_{f,t+1} + [\mathrm{E}(r_{m,t+1}|\Phi_t) - r_{f,t+1}]\,\frac{\mathrm{cov}(r_{i,t+1}, r_{m,t+1}|\Phi_t)}{\mathrm{var}(r_{m,t+1}|\Phi_t)} \tag{1}$$

$$= r_{f,t+1} + [\mathrm{E}(r_{m,t+1}|\Phi_t) - r_{f,t+1}]\,\beta_{i,t+1} \tag{2}$$

Later studies, especially by Fama and French (1992), demonstrate the flaws of the CAPM and present evidence that risk is multidimensional. They show that the combination of the Book-to-Market (B/M) ratio and the market capitalisation of a firm are stronger in explaining variations in historical stock returns than the market beta, market capitalisation, debt/equity ratio, or earnings/price ratio alone. This result builds upon two findings. First, Banz (1981) finds that in terms of market capitalisation small (large) firms have on average a too high (low) return given their market beta estimates. The size effect is assumed to be related to the higher risks that accompany investing in smaller firms, such as higher liquidity risk (Stoll & Whaley, 1983) or different underlying systematic risks faced by small firms (Chan & Chen, 1991). Second, Rosenberg, Reid, and Lanstein (1985) observe that firms' B/M ratios are on average positively related to their stock return. High B/M firms earn on average higher risk-adjusted returns than low B/M firms (B/M effect). The B/M ratio is assumed to be a proxy for distress risks not captured in the other factors (Fama & French, 1992).

Fama and French (1993) add two additional risk factors related to size and B/M to the CAPM in order to incorporate these findings. The first factor, Small-Minus-Big (SMB), measures the size return premium, and the second, High-Minus-Low (HML), measures the value premium provided to investors for investing in companies with high B/M values. The factors are estimated with historical data by dividing the complete market into 25 different portfolios based on the intersections of five size and five B/M quintiles. The expected return according to an adopted Fama and French's (1993) three-factor model is:

$$\mathrm{E}(r_{i,t+1}|\Phi_t) = r_{f,t+1} + \beta_3 \left(\mathrm{E}(r_{m,t+1}|\Phi_t) - r_{f,t+1}\right) + b_s \times SMB_t + b_v \times HML_t \quad (3)$$

The $\beta_3$ is different from the CAPM $\beta$ due to the inclusion of the two new factors. Fama and French (1993) prove that their three-factor model has a higher explanatory power of the cross-sectional return variations in the US market than the CAPM. Still, it does not capture all of the return variations. Thus, some researchers propose to include a forth factor that captures the momentum effect into the model (Carhart, 1997). The momentum effect describes the observation that stock prices which experienced high (low) returns over the past months will yield similar returns in the near future (Jegadeesh & Titman, 1993). Momentum may explain future returns but it is likely not an additional risk. Instead, irrational investor behaviour may explain the observed anomaly (e.g. Barberis, Shleifer, & Vishny, 1998).

## 2.3 Fundamental Investing

Following Ball and Brown's (1968) research that accounting information is value-relevant, fundamental investors try to generate abnormal returns by analysing a firm's fundamentals. Fundamentals encompass all qualitative and quantitative information of a firm that contributes to the firm's valuation. The firm's fundamental value incorporates all its fundamentals. Fundamental investors aim at identifying mispriced stocks, buy (sell) stocks if the market price is lower (higher) than their fundamental value, and attempt to generate abnormal returns as stock prices subsequently gravitate back to fundamental values. Thus, they assume that there is a systematic, temporary bias which violates the efficient market hypothesis. In general, researchers indirectly use fundamentals to predict future stock returns by forecasting future profitability based on a firm's fundamentals (e.g. Penman, 1991; Sloan, 1996). A similar line of research investigates the direct relationship between financial statement information and future stock returns (e.g. Ou & Penman, 1989).

First, focusing on the indirect relationship, Penman (1991) finds that current Return on Equity (ROE) is related to future profitability (future ROE) of the firm. This shows that past fundamentals can predict future accounting profitability. In this respect, since there is a relation between profitability and stock prices (Ball & Brown, 1968), past fundamentals can also predict future stock returns. However, (Penman, 1991) states that ROE alone is not a good indicator for distinguishing the future profitability of firms and should therefore not be used as a single measure in financial statement analysis. While Penman shows that past accounting information is related to future profitability and returns, he cannot confirm that an investor can successfully trade solely based on this information.

On the other hand, Sloan (1996) concludes that accruals predict future earnings and, more importantly, that this information is not fully incorporated into stock prices despite its value relevance. However, this does not necessarily mean that markets are inefficient in the semi-strong form and that investors can capitalise on unexploited profit opportunities. He states that implementing his strategy entails information acquisition and processing costs. Hence, the observed returns may be line with the efficient market hypothesis, since they are merely a compensation for costs associated with the investment strategy. Using UK data, Soares and Stark (2009) find that investors can most likely not capitalise on the the accruals anomaly. They argue that implementing the strategy requires the ability to short significant proportions of small firms' stocks and entails costs associated with trading.

Second, focusing on the direct relationship, Ou and Penman (1989) investigate whether information from financial accounting statements is useful to predict stock prices. Applying extensive financial analysis, they test a large number of accounting ratios. Next, they select the most relevant ratios to predict future stock returns with statistical analysis. This provides them with a summary measure of 16-18 ratios that identifies mispriced firms. Consequently, they conclude that their fundamental analysis systematically predicts abnormal returns. However, Greig (1992) tests these findings by controlling the returns for market beta and size. He finds no incremental predictive power of Ou and Penman's (1989) summary measure. Thus, he concludes that the summary measure only predicts expected and not abnormal returns. Other critics argue that observed returns of the trading strategy are not feasible, since the strategy entails considerable information processing costs for the extensive statistical analysis (e.g. Ball, 1992).

To address the latter critique, a simplification of the models was sought. Relying on analysts' practice instead of extensive statistical search, Lev and Thiagarajan (1993) identify financial (e.g. gross margin, provisions) and operational (e.g. order backlog, sales/employee) variables that are useful in security valuation. Next, they regress future excess returns on the change in earnings and test whether including the identified fundamental variables in the regression increases the explanatory power of the model. They find that the inclusion of most of the variables adds about 70% to the explanation of excess stock returns, compared to only using earnings. Following up, Abarbanell and Bushee (1998) examine whether the information described in Lev and Thiagarajan (1993) is immediately impounded into the share prices. They rank firms based on the previously identified ratios and assign them to portfolios. Their results indicate that investors can earn size-adjusted abnormal returns over a one year holding period using zero-investment portfolios.

To sum up, prior research has shown that past fundamentals can predict future profitability and stock returns. However, the realisation of returns is typically associated with high information acquisition and processing costs. Consequently, researchers developed more simple models to make fundamental investment strategies feasible in practice.

### 2.4 Value Investing and the Book-to-Market Effect

In addition to fundamental investing, value investing is a common investment strategy. This strategy capitalises on the B/M effect and takes long (short) positions in high (low) B/M firms (e.g. DeBondt & Thaler, 1985). Rosenberg et al. (1985) are among the first to find that firms' B/M ratios are on average positively related to their stock returns (the B/M effect). They show that high B/M firms have historically outperformed the market. Since the discovery of the B/M effect, researchers have tried to find compelling reasons for the observed returns and two main research streams have developed, the risk-based and the mispricing view.

According to the risk-based view, different risk levels across firms with different B/M ratios cause the B/M effect. In this view high B/M firms are riskier investments than low B/M firms are. Therefore, the corresponding higher returns for value stocks are only an appropriate risk compensation. This view is most prominently advocated by Fama and French (1992) (see section 2.2). Based on their empirical observation that size and B/M explain returns, they suggest that the B/M ratio of a firm proxies for its distress risk. Chen and Zhang (1998) examine the characteristics of high B/M firms in several countries. They find that other distress risk proxies can explain the B/M effect within each country, e.g. market leverage, dividend reduction, and standard deviation of prior earnings. Thus, they conclude that the average high B/M firm is financially distressed and argue that the higher observed returns are a risk compensation.

According to the mispricing view, irrational investor behaviour explains the B/M effect. Researchers argue that investors tend to naïvely extrapolate a firm's historical performance into the future (e.g. Lakonishok et al., 1994). In this view, high (low) B/M firms are undervalued (overvalued), since investors extrapolate their poor (good) performance too far in the future. Lakonishok et al. (1994) examine if investors extrapolate a firm's past performance too far into the future or if high B/M firms have higher risk. They find that value stocks have historically outperformed the market in the US and that these results are mainly based on a wrong extrapolation of past results. Additionally, they do not find compelling evidence that value strategies are riskier than investing in growth stocks. La Porta (1996) reinforces the extrapolation argument by examining investment strategies based on analysts' earnings growth forecasts. He finds that a portfolio of firms with low expected earnings growth significantly outperforms a portfolio of firms with a high expected earnings growth in absence of a significantly different risk. Hence, he concludes that the market, represented by analysts, is too optimistic (pessimistic) about the earnings growth trajectory. Moreover, DeBondt and Thaler (1985) try to figure out if the market overreacts to unexpected new information that contradicts the past performance of the firm. They find that long-term past losers outperform long-term past winners over the next three to five years after the publication date. They conclude that investors do not fully incorporate the implications of the new information about future profitability into the stock prices. Also, La Porta, Lakonishok, Shleifer, and Vishny (1997) investigate whether the B/M effect is caused by expectational errors made by investors. The study finds evidence that positive earnings surprises on announcement day are larger for value than growth stocks and persist long after portfolio formation. This is inconsistent with a risk-based explanation of the B/M effect.

To sum up, there is compelling evidence for both a risk-based and a mispricing explanation of the B/M effect. Although investing in value stocks, Piotroski (2000) does advocate neither the risk-based nor the mispricing view.

# 3 Piotroski's Investment Strategy

Piotroski (2000) combines the two research streams of fundamental and value investing and couples a simple financial statement analysis with an examination of the B/M effect. In this section we present Piotroski's basic idea, his sample, methodology and empirical results, his additional tests, as well as existing follow-up research.

## 3.1 The Investment Idea and the F_SCORE

Piotroski's (2000) investment idea is based on the observation that the success of investing in value stocks relies on the strong performance of relatively few firms (winners or outperformers), while tolerating the poor performance of many other companies (losers or underperformers). Arguing that accounting information is especially suitable for analysing high B/M firms, he computes the so-called F_SCORE, an aggregation of nine simple binary accounting-based proxies. This score is designed to capture the firm's financial position. The decision to purchase a firm's stock is then based on the strength of this signal. If the F_SCORE is value-relevant and if the market has not already incorporated its information, the signal should assist in identifying the potential winning firms and in improving the observed returns. The nine binary variables aggregated in the F_SCORE capture three areas of firms' financial condition: profitability, financial liquidity/leverage, and operating efficiency.

To capture profitability he uses four variables: $ROA$, $CFO$, $\Delta ROA$, and $ACCRUAL$. $ROA$ and $CFO$ are defined as net income before extraordinary items and cash flow from operations, respectively, divided by total assets at the beginning of the year. If $ROA$ ($CFO$) is positive, the corresponding binary indicator variable $F\_ROA$ ($F\_CFO$) is equal to one, and equal to zero in all other cases. $\Delta ROA$ is defined as the current year's $ROA$ less the prior year's $ROA$. $F\_\Delta ROA$ is equal to one, if the firm improves its $ROA$, i.e. if $\Delta ROA > 0$. The $ACCRUAL$ variable incorporates Sloan's (1996) findings that accrual information is value-relevant (see section 2.3). It is defined as $ROA$ less $CFO$ and its indicator variable $F\_ACCRUAL$ is equal to one if the firm's cash flow is higher than its earnings, i.e. $CFO > ROA$.

To assess financial liquidity and leverage he defines three signals: $\Delta LEVER$, $\Delta LIQUID$, and $EQ\_OFFER$. The variable $\Delta LEVER$ measures the historical change in the ratio of total long-term debt to average total assets. Assuming that an increase in leverage is bad for a distressed firm, $F\_\Delta LEVER$ equals zero (one) if its financial leverage increases (decreases). $\Delta LIQUID$ is defined as the change in the firm's liquidity ratio (assets over liabilities at fiscal year end less assets over liabilities at year start). An improvement in liquidity is seen as a good signal and hence $F\_\Delta LIQUID$ equals one if $\Delta LIQUID > 0$, and zero otherwise. Whether a firms issues seasoned equity is measured by the variable $EQ\_OFFER$. Assuming that issuing additional equity by a distressed firm is a bad sign $F\_EQ\_OFFER$ equals zero if the firm issued equity, and one otherwise.

The third financial condition, operating efficiency, is assessed by the two variables $\Delta MARGIN$ and $\Delta TURN$. He defines $\Delta MARGIN$ as the firm's current gross margin (current gross profit divided by current sales) less the firm's prior year's gross margin. $F\_\Delta MARGIN$ equals one if the margin improves, zero otherwise. $\Delta TURN$ is defined as the firm's current asset turnover (current sales over the assets at the beginning of the year[2]) less its prior year's asset turnover. An improvement in turnover is seen as positive and hence $F\_\Delta TURN$ equals one if $\Delta TURN > 0$, zero otherwise.

An overview of all F_SCORE variables is presented in table 1 on p. 25. Selecting these variables, Piotroski does not aim to find a single best set of accounting ratios to predict future stock returns. He relies on practice and qualitative arguments instead of statistical search. Therefore, his strategy is easy to implement as it does not require complex, costly statistical models.

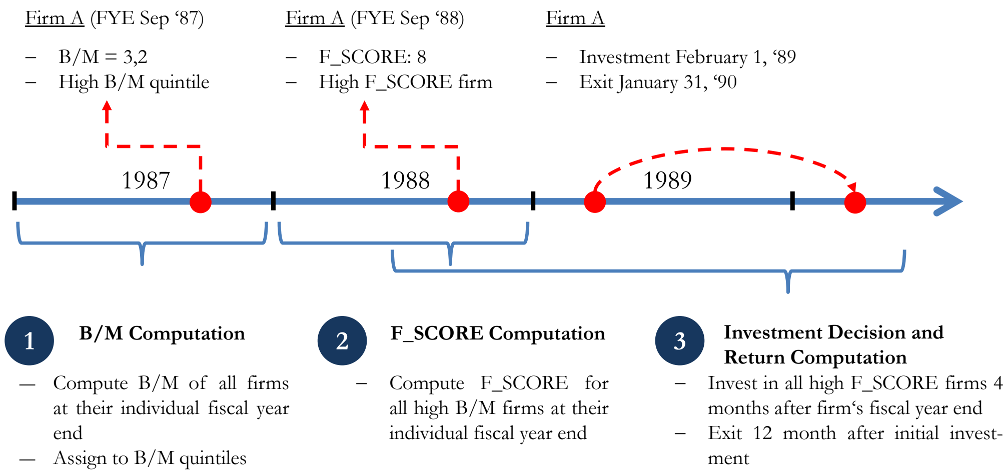
## 3.2 Sample Selection, Methodology, and Empirical Results

In each year between 1976 and 1996, Piotroski (2000) identifies firms with sufficient stock price and book value of equity data on Compustat, a database which includes all firms with a primary listing on the New York Stock Exchange (NYSE), American Stock Exchange (Amex), NASDAQ, or Archipelago Exchange (ARCA), and therefore mainly US-based companies. Then, he proceeds in three steps explained below.

First, to make an investment decision, he computes the firm's B/M ratio at fiscal year end for all sample firms. Next, he assigns each firm to a B/M quintile for each calendar year. For example, for all firms, whose fiscal year ends during 1987, the firm's B/M ratio

[2] Piotroski's (2000) definition of the variable changes in his article. We use the definition on page 9 and not the one from the footnotes of table 1 on page 14 of his article.

**Figure 1** – Overview of Piotroski's Investment Strategy




is calculated on their individual fiscal year end (e.g. 30/6/1987, 30/9/1987) providing the investor with the five quintiles for 1987.

Second, he computes the F_SCORE for each firm at each fiscal year end. In each year he classifies firms with an F_SCORE equal to eight or nine as high F_SCORE firms and firms with an F_SCORE equal to zero or one as low F_SCORE firms. While high F_SCORE firms are expected to outperform the market, low F_SCORE firms are expected to underperform. Firms with an F_SCORE between three and seven, inclusive, are not considered.

Third, for the investment decision, he considers all firms in the highest B/M quintile (value stocks) in year $t-1$ and each firm's F_SCORE from year $t$. Of all value stocks as of $t-1$ he invests only in the high F_SCORE firms as of year $t$. The investment is made four months after the end of each firm's fiscal year in order to ensure that the financial statements are publicly available at that time. For example, as shown in figure 1, if firm A belongs to the highest B/M quintile in 1987 and has an F_SCORE of eight based on its fiscal year end on 30/9/1988, an investment in the stock is made on 1/2/1989. The firm specific returns are measured as the one year buy-and-hold return earned from the beginning of the investment until one year later.[3] For example, the one year stock return for firm A's stock bought on 1/2/1989 is measured as of 31/1/1990. If a firm's stock is delisted during the holding period, the return is assumed to be zero. The returns of all

[3]Piotroski (2000) partly shows results for a two year holding period, which are very similar to the ones for the one year holding period.

firms with a high and low F_SCORE in year $t$ are assigned to the same year $t$, although the investment is later. Overall, the returns of the strategy are computed as the equally-weighted average of all return observations. This implies that an equal amount is invested in each stock.

Applying his investment strategy to the US market between 1976 and 1996, Piotroski (2000) finds that an investor could have increased the mean stock returns by 7.5 percentage points annually when selecting value stocks with a high F_SCORE compared to investing in the whole portfolio of value stocks. Furthermore, he shows that the entire return distribution is shifted to the right when investing in high F_SCORE firms within a value stock portfolio. He also demonstrates that shorting low F_SCORE firms (expected losers) and buying high F_SCORE firms (expected winners) generates average annual returns of 23% over the twenty year period. Overall, these findings seem to indicate that investors can generate abnormal returns when applying a simple accounting-based investment strategy. This means that investors do not immediately incorporate available financial information.

### 3.3 Performed Tests

To further analyse whether the strategy can really generate abnormal returns and to address other potential criticism, Piotroski (2000) pursues several additional tests and shows various return partitions. More specifically, he tests (1) if his investment strategy contributes to predict future returns beyond previously known anomalies, (2) if the returns are feasible for an investor, (3) if the F_SCORE is barely an ad hoc generated metric, (4) if a risk-based explanation of the above-market returns is likely, and (5) if the market only slowly incorporates publicly available information.

First, Piotroski tests if rather three previously known effects than his fundamental analysis explain the future return generation. Prior research has shown that historical levels of accruals (Sloan, 1996), equity offerings (Loughran & Ritter, 1995; Spiess & Affleck-Graves, 1995), and momentum strategies (Chan et al., 1996) predict future returns. Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) find that firms issuing equity in an Initial Public Offering (IPO) or Seasoned Equity Offering (SEO) have lower stock returns over the subsequent years than firms that do not. Chan et al. (1996) find evidence that strategies trading on the momentum effect can be successful and are not a statistical fluke. Piotroski's F_SCORE embeds the accrual and equity offering effect by including the *ACCRUAL* and *EQ_OFFER* variables in its computation. According to Piotroski, the F_SCORE's underlying success is based on the underreaction to historical information and financial events. The same effect is supposed to drive the momentum effect. Hence, these previously known effects correlate with the F_SCORE. Thus, the performance metric may only aggregate these effects, but does not itself contribute to the prediction of future returns. However, Piotroski finds that the inclusion of variables designed to capture these effects to a regression of market-adjusted returns has no impact on the robustness of the F_SCORE to predict future returns.

Second, Piotroski tests if the return improvements are feasible for an investor. If the returns were limited to stocks with low trading volumes, low share prices, and small market capitalisations, it would be unrealistic to assume that the returns are feasible for an investor. Then, an actual meaningful investment in the stocks could have significantly influenced the historical price determination. To address this, Piotroski shows three additional partitions of the returns: returns conditioned on terciles of size, trading volume, and share price. The mean market-adjusted return difference for small (medium) sized firms is 27.0 (17.3) percentage points and highly significant. In contrast, the difference is not significant for large firms. The significance for medium sized firms suggests that the strategy is feasible for investors. Furthermore, he finds that the high returns do not disappear when controlling for a low share price effect or low trading volumes.

Third, Piotroski shows that two other accepted measures of firm health can differentiate winners from losers. This way he undermines criticism that the F_SCORE is a specifically designed, ad hoc score to make investment decisions. He uses Altman's (1968) Z-score and the historical change in profitability, measured by the change in *ROA*, as other indicators for financial health. He divides the whole sample into terciles with low, medium, and high risk of financial distress based on Altman's Z-score. He finds that firms with high returns have a low risk of financial distress. In addition, he divides the whole sample into terciles with low, medium, and high historical change in profitability and demonstrates that firms with high levels of historical profitability have high future returns. To sum up, other financial statement-based indicators for financial health can also differentiate between winners and losers. Since other common indicators can also indicate future returns in the same data set, it is unlikely that F_SCORE is ad hoc designed.

Fourth, Piotroski states that a risk-based explanation is unlikely for three reasons.

First, high F_SCORE firms show the strongest subsequent returns, but have the smallest amount of ex-ante operating and financial risk as measured by the historical performance signals, i.e. the respective ratios to compute the F_SCORE. Second, small differences in size and B/M ratios are unlikely to account for a 23 percentage points differential in market-adjusted returns and the strategy generates positive returns in 18 out of 21 years.[4] Third, Piotroski computes $ROA_{t+1}$ and presents subsequent business failures, measured by performance-related delistings, for the various F_SCOREs. This demonstrates that the performance metric identifies firms with high levels of future profitability and low future failure risk. These findings contradict Fama and French's (1992) suggestion that the B/M effect is related to financial distress risk, since healthy firms within the high B/M portfolio yield higher returns and have stronger subsequent financial performance. In summary, Piotroski concludes that these findings contradict a risk-based explanation.

Fifth, Piotroski shows that the market slowly incorporates past performance and that fundamental analysis is most effective when investing in companies with limited available information. He presents the mean stock returns conditioned on F_SCORE over the subsequent four quarterly announcement periods following portfolio formation. Returns are measured as the buy-and-hold returns over a three day window surrounding the announcement date. According to his results winners experience a stronger earnings announcement surprise than losers and earnings announcement differences are stronger for small firms with low share turnover and without analyst coverage. This supports the argument that fundamental investing is most effective for companies with limited available information.

To sum up, Piotroski (2000) finds that returns are not explicable by known anomalies, that the returns are feasible for an investor, that the F_SCORE is not ad hoc generated, and that a risk-based explanation of the observed returns is unlikely. Furthermore, he shows that the market only slowly incorporates historical financial information and that his investment strategy is most effective for companies with limited available information. His findings support the view that the investment strategy does not only increase the investor's risk exposure, but indeed helps to identify over- and undervalued stocks.

[4] To show that the strategy generates positive returns over time, Piotroski (2000) adjusts the strategy and shorts value stocks with F_SCOREs of 4 or less and buys stocks with F_SCOREs of 5 or higher.

### 3.4 Follow-Up Study

There is only one major published follow-up study on developed markets. Using an adjusted performance metric, Mohanram (2005) investigates if one can achieve similar results in a growth stock portfolio in the US market between 1978 and 2001. Additionally, he re-tests the F_SCORE in the value stock portfolio and also applies it to a growth stock portfolio. In this section we present the methodology, the empirical results, and the main criticism of Mohanram's (2005) study.

Mohanram argues that growth firms have different characteristics than value firms due to generally higher investor and analyst following, more sources of information available other than the financial statements, and higher growth rendering fundamental accounting data less important in their valuation. Therefore, he adjusts the F_SCORE so that it is more suitable for analysing growth stocks and calls his performance metric GSCORE. The GSCORE methodology differs in four important aspects from the F_SCORE. First, the financial information used to construct the GSCORE is generally compared to the industry median in that year. All information is therefore considered relative to an assumed industry average. Second, the used ratios are different. The indicators measure profitability (ROA, cash flows, and net income), but also the results of naïve extrapolation (earnings and sales growth variability) and the effects of accounting conservatism (research and development spendings, capital expenditures, and advertising intensity in earnings). Third, in contrast to the F_SCORE, which requires the availability of all financial information to construct it, the GSCORE requires firms only to have earnings and cash flow information available. Firms with insufficient information available to calculate all binary signals are thus only able to achieve a lower GSCORE, but are not dropped from the sample. Fourth, all investments are done simultaneously on May 1 in all investment years. Based on a firm's GSCORE in a particular year, he builds portfolios of firms with high or low GSCOREs. Then, he differentiates between winners and losers in the lowest B/M quintile of the US stock market.

Mohanram achieves significantly higher size-adjusted returns for high than for low GSCORE firms in the growth stock portfolio. In addition, Mohanram shows that the F_SCORE strategy also works in the growth stock portfolio, but yields weaker results than the GSCORE strategy. This supports the view that contextual financial analysis matters, since the GSCORE (F_SCORE) is specifically designed for growth (value) stocks. He finds that the effectiveness of fundamental analysis for the growth stock portfolio is driven by high information availability, whereas the success of fundamental analysis for the value stock portfolio is driven by the neglection of stocks and low information availability. The GSCORE strategy's returns in the growth stock portfolio are positively related to analyst coverage and to firm size, both proxy for information availability. These findings document that mispricing in the two extreme B/M portfolios is of different nature.

Mohanram's study is criticised, because the main part of the observed returns to the GSCORE strategy originate from shorting underperforming firms. Thus, the ability to buy shorting instruments in practice is crucial. This is problematic since the sample starts in the late 1970s, when the availability of these instruments was limited. Hence, it is questionable whether the observed returns could have been realised in practice.

To conclude, according to Mohanram (2005) the F_SCORE is also applicable within the growth stock portfolio and the nature of mispricing seems to be different across B/M portfolios. Within a growth stock portfolio, financial analysis identifies mainly underperforming firms and works better for firms with a high degree of information availability.

# 4 Limitations of Piotroski's Study

Piotroski (2000) concludes that markets are inefficient and that investors can earn abnormal returns with fairly simple financial analyses. However, these conclusions are limited. His assertions are not supported by a more detailed investigation whether the realised returns are limited to a single market, are abnormal, and are continuous.

Piotroski only tests his strategy on the US market in the value stock portfolio. He provides evidence that the strategy is not specifically designed for his data by showing that other performance metrics can also identify out- and underperformers in the same data set. However, he applies the strategy neither in other portfolios nor in other markets. Mohanram (2005) tests Piotroski's strategy also in the growth stock portfolio, but still uses US data (see section 3.4). In addition, both studies confine their research to the extreme B/M quintiles. Extending the sample to the whole market would supply evidence whether Piotroski's metric is more generally applicable.

Furthermore, Piotroski does not appropriately evaluate whether the realised returns are abnormal. He does neither use an asset-pricing model nor adjusts for the three major risk factors market beta, size, and B/M simultaneously. Instead he uses three main arguments to undermine a risk-based explanation for the high (low) returns of high (low) F_SCORE firms and thus attempts to demonstrate that the observed returns are abnormal. First, he argues that the high F_SCORE firms show the strongest subsequent returns, but have the smallest amount of ex ante operating and financial risk as measured by the historical performance signals. This argumentation implies that the F_SCORE estimates ex ante operating and financial risk. However, the F_SCORE is not indicating ex ante risk per se. Instead, it indicates the one year historical *change* in ex ante risk, since most of its binary variables consider one year historical changes in financial or operational condition. In an efficient market with risk-averse investors an unexpected decrease in risk should lower the cost of capital and cause an increase in the firm's valuation. Hence, a historical risk decrease should have an immediate effect on stock prices. This reaction to historical changes should have occurred prior to an investment with the F_SCORE strategy. In contrary, the firm's average returns in the long run are not based on the historical change in risk (e.g. change in leverage), but on the actual level of ex ante risk at investment (e.g. current leverage).

Second, Piotroski argues that the small differences in size and B/M ratios among value stocks are unlikely to account for a 23 percentage points differential in market-adjusted returns. In addition, he demonstrates that the a hedge portfolio based on the strategy generates positive returns in 18 out of 21 years. These arguments imply that investors can short stocks without incurring additional costs. However, short selling constraints, insufficient liquidity, and costs associated with holding short positions over a long time period could lead to lower or negative realisable returns.

Third, Piotroski demonstrates that high F_SCORE firms have high levels of future profitability and low future failure risk. These findings contradict the idea that high failure risk drives the expected returns as argued by Fama and French (1992). However, high F_SCORE firms could be associated with other risks, e.g. liquidity risk or high market betas.

In addition, Piotroski does not investigate how returns develop over time. He shows the returns for each year between 1976 and 1996 but does not analyse if returns differences have changed over time. Also, it is not shown whether the return differences are significant per year. One could expect that the effectiveness of F_SCORE decreases in more recent years when developing and executing fundamental investment strategies became considerably cheaper. Testing the investment strategy over time provides evidence if the results are only a short-term anomaly or if they persist.

In conclusion, investigating the limitations of Piotroski's (2000) study shows that further out-of-sample evidence is necessary. His study has been neither tested in other developed markets nor over the entire market. Additionally, he has not analysed with more sophisticated methods if the returns are abnormal and if abnormal returns persist over time. We seek to address these limitations by applying Piotroski's study to the entire UK market and by evaluating whether the observed returns are abnormal and continuous.

# 5 Applying Piotroski to the UK Market

In this section our application of Piotroski's (2000) investment strategy to the UK market and eventual modifications are described. Generally, we aim at an exact replication of his investment strategy in order to make our results comparable to Piotroski's. Therefore, modifications are only done if different market characteristics or data availability necessitate adjustments to keep the investment strategy realistic in practice. In this section we describe our sample selection and methodology first. Then, the empirical results are presented and finally analysed.

## 5.1 Sample Selection and Methodology

We describe mainly aspects in the sample selection and methodology in case they differ from Piotroski's (2000) as explained in section 3.2. First, our differences in the selection process of the sample is described, followed by the B/M, F_SCORE, and return computation.

### 5.1.1 Sample Selection

The sample selection differs in the used database, time period, stock market, scope of the market, and treatment of multiple stock classes. For the years 1990-2007, we identify firms on the London Stock Exchange (LSE) having sufficient stock price and book value of equity data in the Thomson Reuters Datastream (DS) database. All selected firms must have their primary listing on the LSE. DS is used instead of Compustat, which is used by Piotroski (2000), as the latter only recently started to cover non-US markets.

Although DS includes observations prior to 1990, we exclude them from the sample for two reasons. On the one hand, DS's total return index is computed differently before 1988. On the other hand and most importantly, information on small companies is only available since the 1990s in Worldscope (WS)/DS (Thomson Financial, 2007). Thus, an earlier sample would not have been representative of the whole UK stock market. Moreover, our selected time period is more recent. This is important since anomalies tend to disappear once they have been discovered (Dimson & Marsh, 1999). In addition, they may disappear over time since the costs for developing and performing investment strategies (i.e. computing costs and database availability) have constantly decreased.

The UK market is selected for three reasons. First, successfully testing the strategy on the UK market could provide out-of-sample evidence that Piotroski's observed returns are not country-specific and more generally applicable. Second, the LSE is Europe's largest stock exchange in terms of market capitalisation (World Federation of Exchanges, 2010). A large sample stock exchange increases the power of the empirical tests, since a large number of observations over a relative long period of time is required. Third, financial statement analysis requires a market with well functioning accounting systems to ensure that our results are not distorted by unreliable accounting figures. UK accounting standards show historically only few value-relevant differences compared to US standards (Weetman & Gray, 1990).

The decision to extend the sample scope to the whole stock market is mainly motivated by Mohanram (2005). He shows that the success of this strategy is not confined to one B/M quintile. Thus, we explore whether the success differs when applying the strategy to different quintiles and to the whole market. This tests for a broader applicability of the investment strategy and shows whether the strategy works irrespective of the B/M ratios.

Selecting the sample firms we notice that some companies have multiple classes of stocks. However, we include all classes in our sample since Piotroski (2000) does not explicitly state how he handles firms with multiple ones. Besides, only 2% of all firms with primary listings on the LSE have multiple classes of stocks. Ideally an investor would like to invest only in the most liquid stock of all classes. Yet, it is difficult to set up such a rule in historical data without implying foresight bias.

### 5.1.2 B/M Computation

Like Piotroski (2000) we create B/M quintiles for every year between 1990 and 2007 first. We exclude observations with negative B/M values, i.e. firms with a negative book value of equity. Piotroski does not explicitly state if he in- or excludes negative B/M firms, but excluding them is a common practice in previous B/M research (e.g. Fama & French, 1995). In total we obtain 28703 positive firm-year B/M observations directly from the DS database.

#### 5.1.3 F_SCORE Computation

If a firm has a B/M value in year $t-1$, all accounting information from the firm's fiscal year end reporting in year $t$, $t-1$, and $t-2$ necessary for the F_SCORE calculation is downloaded from the WS/DS database. This is the same logic as described in section 3.2. However, we do not restrict the sample firms to the highest B/M quintile, but download the financial information for all firms regardless of their B/M quintile membership. Overall, financial statement data is obtained for the fiscal year ends from 1989-2008. As in Piotroski (2000), we proceed to the investment step only with those companies for which all financial information necessary to compute all F_SCORE variables is available (see table 1 for an overview of the required items). As gross profits or margins are typically not available for a broad range of financial companies, such as banks and investment companies, these financial service firms are indirectly excluded in this investment strategy. In the end, we have 18878 complete firm-year observations over the sample period with an average of 1049 firms per year. The number of firms per year remains fairly constant over the whole sample period with a minimum number of 948 in year 1998 and a maximum of 1159 in year 2007.

Whereas most of the accounting items are similarly organised in DS as in Compustat, two F_SCORE variables differ. The first one is $CFO$. In DS the Cash Flow from Operations (CFO), as disclosed in the UK cash flow statements, is only available since 1995 for some companies. Thus, relying on this DS item would lead to the exclusion of too many sample firms. Piotroski (2000) uses the CFO as disclosed in US GAAP statements to construct the $CFO$ variable. The US GAAP CFO deducts, for example, full interest and tax expenses and does not include investments in fixed assets. Thus, it is comparable to earnings. This comparability is important as the $ACCURAL$ variable compares the $CFO$ with the $ROA$ variable. We approximate the US GAAP CFO with the DS item 'Funds from Operations'. This item equals net income plus non-cash expenses, such as depreciation or expenses for provisions. Next, we deduct increases in working capital from this figure and use the result for the $CFO$ variable.

Second, in contrast to the Center for Research in Security Prices (CRSP) database used by Piotroski (2000), DS has no variable indicating whether a company has issued common equity in a seasoned offering during a fiscal year. Therefore, a reasonable approximation

**Table 1**

DEFINITIONS OF F_SCORE VARIABLES

| Variable name | Description |
|---|---|
| F_SCORE | $\text{F\_SCORE}_t = \text{F\_}ROA_t + \text{F\_}CFO_t + \text{F\_}\Delta ROA_t + \text{F\_}ACCRUAL_t + \text{F\_}\Delta LEVER_t + \text{F\_}\Delta LIQUID_t + \text{F\_}EQ\_OFFER_t + \text{F\_}\Delta MARGIN_t + \text{F\_}\Delta TURNOVER_t$ |
| *Profitability* | |
| Return on Assets | $ROA_t = (\text{Net Income Before Extraordinary Items}_t)/(\text{Assets}_{t-1})$<br>$\text{F\_}ROA_t = 1$ *if* $ROA_t > 0$, *else* 0 |
| Operational Cash Flow | $CFO_t = (\text{Cash Flow from Operations})/(\text{Assets}_{t-1})$<br>$\text{F\_}CFO_t = 1$ *if* $CFO_t > 0$, *else* 0 |
| Change in ROA | $\Delta ROA_t = ROA_t - ROA_{t-1}$<br>$\text{F\_}\Delta ROA_t = 1$ *if* $\Delta ROA_t > 0$, *else* 0 |
| Accruals | $ACCRUAL_t = CFO_t - ROA_t$<br>$\text{F\_}ACCRUAL_t = 1$ if $ACCRUAL_t > 0$, *else* 0 |
| *Financial Liquidity/Leverage* | |
| Change Leverage | $\Delta LEVER_t = (\text{Long-term Debt}_t)/(\frac{1}{2}\text{Assets}_t + \frac{1}{2}\text{Assets}_{t-1}) - (\text{LT Debt}_{t-1})/(\frac{1}{2}\text{Assets}_{t-1} + \frac{1}{2}\text{Assets}_{t-2})$<br>$\text{F\_}\Delta LEVER_t = 1$ *if* $\Delta LEVER_t < 0$, *else* 0 |
| Change Liquidity | $\Delta LIQUID_t = (\text{Current Assets}_t)/(\text{Current Liabilities}_t) - (\text{Current Assets}_{t-1})/(\text{Current Liabilities}_{t-1})$<br>$\text{F\_}\Delta LIQUID_t = 1$ *if* $\Delta LIQUID_t > 0$, *else* 0 |
| Equity Offer | $EQ\_OFFER_t = \textit{Issue New Equity}_t$<br>$\text{F\_}EQ\_OFFER_t = 0$ *if* $\Delta\text{Total common shares}_t > 0$ *and* Proceeds from Equity $\text{Issue}_t > 0$, *else* 1 |
| *Operating Efficiency* | |
| Change Margin | $\Delta MARGIN_t = (\text{Sales}_t - \text{COGS}_t)/\text{Sales}_t - (\text{Sales}_{t-1} - \text{COGS}_{t-1})/\text{Sales}_{t-1}$<br>$\text{F\_}\Delta MARGIN_t = 1$ *if* $\Delta MARGIN_t > 0$, *else* 0 |
| Change Turnover | $\Delta TURNOVER_t = \text{Sales}_t/\text{Assets}_{t-1} - \text{Sales}_{t-1}/\text{Assets}_{t-2}$<br>$\text{F\_}\Delta TURNOVER_t = 1$ *if* $\Delta TURNOVER_t > 0$, *else* 0 |

has to be found for the $EQ\_OFFER$ variable. We assume that a company issues seasoned common equity if the total number of common shares increases between two fiscal year ends and if the company simultaneously has a positive entry in the DS item 'net proceeds from issued equity'. Total common shares are the sum of number of common shares outstanding and treasury shares held by the firm at fiscal year end. If the change in total common shares is positive for a firm and if the firm receives proceeds from issued equity, we assume the firm to issue seasoned equity. These items are used in combination to exclude share issues for other purposes, such as management compensation, share splits, or share dividends. If the information about total common shares is available for a company, but the item 'net proceeds from issued equity' is not filled or amounted to zero, we assume that the company does not issue any equity. In case no common shares' data entry is available, $EQ\_OFFER$ is not computed. Every time a company is included for the first time in DS, the change in total common shares cannot be computed, because only the current year's number of common shares is available. Since Piotroski (2000) considers only SEOs, in other words new equity issues by an already publicly-traded company, we assume the change in common shares for first-time data entries to equal zero.

In contrast to Piotroski (2000) we define a low F_SCORE as a score of not only 0 or 1, but also 2, since we have hardly any observations with an F_SCORE of 0. Additionally, looking at the high B/M quintile, our sample (3190) is considerably smaller than Piotroski's (14043 firm-year observations). Therefore, we introduce the broader F_SCORE categories weak (0-3), medium (4-6), and strong (7-9).

#### 5.1.4 Return Computation

To compute the holding returns we use the 'Total Return Index' (RI) item provided by DS. This data type adjusts not only for stock price changes but also for dividends and capital actions, such as stock splits. Hence, it reflects the total return for a shareholder under the assumption that dividends are immediately reinvested. Using the RI bears the risk that the return computation might be inaccurate, since DS provides only two decimals. This is more pronounced if the RI is priced very low for a firm at the investment time. For example, if at investment the RI was priced at 0.01 and the stock increased by 51%, the two decimal RI would increase to 0.02 observing a return of 100%. However, these effects should average out. Excluding all return computations for which the total return index is below 0.1 at investment would not change the overall results.

Contrary to Piotroski (2000) we invest at the beginning of the seventh (not fifth) month after each firm's individual fiscal year end, because companies are required to publish their audited financial statements within a period of six month in the UK (FSA, 2010). Investing earlier after the firms' fiscal year ends could imply a foresight bias and hinder a replication in practice.[5] Consequently, each firm's return is calculated as the

[5] In this respect, we want to highlight that we invest always at the beginning of the seventh months after the firm's *individual* fiscal year end. In other, non-peer-reviewed or non-published replications of Piotroski's strategy this is commonly simplified and all investments take place four or six month after December 31, the last available fiscal year end in any calendar year (see e.g. Duong, Pescetto, & Santamaria, 2010; Lovric & Rados, 2010).

one year holding return from an investment in the stock at the beginning of the seventh month after the firm's fiscal year end. For an overview of the investment logic see section 3.2. If the stock delisted during the holding period, a return of zero is assumed in line with Piotroski (2000). Otherwise, the raw return RAWRET and market-adjusted return MARET for firm $i$ and calendar year $t$ are computed as:

$$\mathrm{RAWRET}_{i,t} = \frac{\mathrm{RI}_{i,t_{exit,i}}}{\mathrm{RI}_{i,t_{invest,i}}} - 1 \tag{4}$$

$$\mathrm{MKTRET}_{i,t} = \frac{\mathrm{RI}_{\mathrm{FTSE},t_{exit,i}}}{\mathrm{RI}_{\mathrm{FTSE},t_{invest,i}}} - 1 \tag{5}$$

$$\mathrm{MARET}_{i,t} = \mathrm{RAWRET}_{i,t} - \mathrm{MKTRET}_{i,t} \tag{6}$$

The RAWRET for firm $i$ in period $t$ is market-adjusted by subtracting the return of the Financial Times Stock Exchange (FTSE) All-Share Index (MKTRET) over the corresponding investment period, i.e. the same *invest* and *exit* dates. The FTSE All-Share Index, obtained through DS, includes all companies listed on the LSE that surpass a liquidity test (FTSE, 2010). This test ensures that the market returns are achievable and appropriate for adjusting the raw returns.

Last, as in Piotroski (2000) the one year buy-and-hold returns of all firms from all time periods are pooled and their equally-weighted average is compared. Since the F_SCORE strategy is based on investing in high (strong) and shorting low (weak) F_SCORE firms, the return difference between the high and low (strong and weak) F_SCORE firms is calculated. Additionally, we calculate the return differences between the high (strong) F_SCORE firms and all the firms in their respective sample segment groups. This way we assess whether the high (strong) F_SCORE firms performed better than their peer groups.

## 5.2 Empirical Results

The empirical results are presented as the equally-weighted average one year buy-and-hold returns of the pooled firm-year observations in the respective sample segmentations. The presentation is divided into three parts. First, we present the descriptive statistics of our sample, then the returns conditioned on B/M, and finally the returns conditioned on size.

**Table 2**

Descriptive Statistics of Sample Firms

**Financial Characteristics**

| **Variable** | Mean | Median | Standard Deviation | Proportion with Positive Signal |
|---|---|---|---|---|
| ROA | 0.004 | 0.048 | 0.238 | 73.47% |
| CFO | 0.057 | 0.081 | 0.748 | 79.99% |
| ΔROA | 1.965 | -0.002 | 317.1 | 48.44% |
| ACCRUAL | -0.053 | -0.043 | 0.707 | 30.69% |
| ΔLEVER | 0.005 | 0.000 | 0.096 | 47.83% |
| ΔLIQUID | -0.270 | -0.013 | 19.515 | 47.44% |
| EQ_OFFER | NA | NA | NA | 36.46% |
| ΔMARGIN | 0.198 | 0.001 | 18.493 | 51.12% |
| ΔTURN | -0.723 | -0.005 | 63.319 | 48.65% |
| Market Capitalisation | 865.471 | 47.569 | 5752.0 | NA |

**One Year Buy-and-Hold Returns and B/M Quintiles**

| *B/M Quintiles* | *low* 1 | 2 | 3 | 4 | *high* 5 | All |
|---|---|---|---|---|---|---|
| Raw Returns | 7.43% | 9.20% | 11.73% | 11.51% | 14.48% | 10.70% |
| Market-Adjusted Returns | -1.24% | 0.12% | 2.80% | 2.32% | 5.76% | 1.78% |

**One Year Buy-and-Hold Returns and Size Terciles**

| *Size Terciles* | *low* 1 | 2 | *high* 3 | ALL |
|---|---|---|---|---|
| Raw Returns | 13.79% | 8.80% | 9.70% | 10.70% |
| Market-Adjusted Returns | 4.53% | -0.13% | 1.09% | 1.78% |

The table shows the mean, median, and standard deviation of the sample firms' value on the nine F_SCORE variables as defined in table 1. Next the proportion of positive binary F_SCORE indicator signals is reported. The total sample consists of 18878 firm-year observations between 1991-2008. NA means that the value is not available. The raw and market-adjusted returns, B/M quintiles, and size tericles are calculated as described in tables 3, 4 and 6.

### 5.2.1 Descriptive Statistics

Table 2 provides descriptive statistics about the financial characteristics, especially F_SCORE statistics, as well as return statistics for all firm-year observations. Contrary to Piotroski (2000), the mean and median $ROA$, $CFO$, and $MARGIN$ of our sample firms are positive. Yet, this is reasonable, because he investigates only high B/M firms that generally perform poorly. Using histogram analysis, the comparably high $\Delta ROA$ and low $\Delta TURN$ display many outliers. However, in this respect outliers do not matter for our analysis since the F_SCORE consists of binary, not continuous variables.

Consistent with previous research, high B/M firms outperform low B/M firms both in terms of raw and market-adjusted returns (B/M effect). Also the size effect is visible. The bottom third of market capitalisation outperforms the top third. However, the middle tercile shows slightly stronger returns than the largest tercile. Market-adjusted returns of all sample firms are close but not equal to zero. This suggests that the FTSE All-Share index is a suitable but not perfect index for our sample, since it does not encompass all sample firms.

### 5.2.2 Returns Conditioned on B/M

Tables 3 and 4 show the raw and market-adjusted returns, respectively, as well as the corresponding observations conditioned on B/M quintiles and the F_SCORE. The number of observations differ across B/M quintiles, because we consider only those firms for which we have both the required data to calculate the F_SCORE in year $t$ as well as the information for B/M in year $t-1$ available. Since the B/M quintiles are calculated one year earlier, some data required for the F_SCORE calculation might be missing in the subsequent year. Then, in line with the strategy these firms drop out of the sample and the number of observations of the B/M quintiles differ. The high B/M quintile is dominated by small firms, which are more likely not to publish all required accounting data. Therefore, it has fewer observations than the low B/M quintile. Moreover, the growth quintile has relatively more low and weak F_SCORE firms than the value stocks as growth firms they are more likely to issue equity.[6] First, we present the returns and then the

[6] Analysing the proportions with a positive signal of the binary variable $EQ\_OFFER$ for the low and high B/M firms, we find that in the low (high) B/M portfolio 78% (35%) have issued seasoned common equity. Compared to this 54 percentage points difference, all other ratios are fairly alike.

## Table 3

### Buy-and-Hold Raw Returns Across B/M Quintiles

***Raw Returns***

| F_SCORE | *low* 1 | B/M Quintiles 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | -47.68% | 58.62% | -19.51% | | | -14.06% |
| 1 | -12.17% | -6.75% | -20.59% | 8.45% | 1.31% | -7.03% |
| 2 | 4.00% | -4.39% | 4.60% | 4.67% | 10.49% | 3.39% |
| 3 | -2.14% | 0.78% | 10.45% | 3.59% | 10.62% | 3.88% |
| 4 | 2.64% | 12.58% | 10.42% | 11.83% | 11.70% | 9.60% |
| 5 | 9.01% | 6.11% | 12.67% | 10.39% | 15.47% | 10.37% |
| 6 | 10.65% | 10.56% | 11.58% | 14.87% | 17.78% | 12.82% |
| 7 | 15.35% | 12.67% | 14.12% | 12.17% | 15.25% | 13.80% |
| 8 | 13.42% | 19.93% | 16.80% | 12.14% | 12.61% | 14.74% |
| 9 | 42.09% | 25.18% | -3.19% | 27.82% | 23.58% | 22.26% |
| **ALL** | **7.43%** | **9.20%** | **11.73%** | **11.51%** | **14.48%** | **10.70%** |
| High | 17.24% | 20.66% | 14.32% | 14.57% | 14.09% | 15.77% |
| Low | 0.45% | -4.30% | 0.68% | 5.33% | 9.06% | 1.58% |
| High - Low | 16.79% | 24.96% | 13.64% | 9.24% | 5.03% | 14.19% |
| *t*-Statistic | 1.98** | 3.965*** | 2.504*** | 1.019* | 0.804 | 4.199*** |
| High - All | 9.81% | 11.46% | 2.59% | 3.06% | -0.38% | 5.07% |
| *t*-Statistic | 2.483*** | 3.166*** | 0.778 | 0.995 | 0.158 | 3.592*** |
| Weak | -1.31% | -0.61% | 7.59% | 4.09% | 10.20% | 3.21% |
| Medium | 7.62% | 9.45% | 11.67% | 12.42% | 15.19% | 10.99% |
| Strong | 15.88% | 14.90% | 14.18% | 12.95% | 14.77% | 14.45% |
| Strong - Weak | 17.19% | 15.52% | 6.59% | 8.87% | 4.57% | 11.23% |
| *t*-Statistic | 4.465*** | 4.97*** | 2.014** | 2.14** | 0.969 | 6.175*** |

***Number of Observations (n)***

| F_SCORE | *low* 1 | B/M Quintiles 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | 2 | 1 | 1 | 0 | 0 | 4 |
| 1 | 41 | 21 | 21 | 23 | 16 | 122 |
| 2 | 173 | 129 | 119 | 108 | 87 | 616 |
| 3 | 460 | 398 | 342 | 326 | 276 | 1802 |
| 4 | 780 | 787 | 720 | 598 | 549 | 3434 |
| 5 | 1034 | 1047 | 989 | 837 | 705 | 4612 |
| 6 | 807 | 940 | 890 | 842 | 659 | 4138 |
| 7 | 463 | 595 | 593 | 613 | 528 | 2792 |
| 8 | 156 | 198 | 247 | 250 | 320 | 1171 |
| 9 | 24 | 32 | 35 | 46 | 50 | 187 |
| **ALL** | **3940** | **4148** | **3957** | **3643** | **3190** | **18878** |
| High | 180 | 230 | 282 | 296 | 370 | 1358 |
| Low | 216 | 151 | 141 | 131 | 103 | 742 |
| Weak | 676 | 549 | 483 | 457 | 379 | 2544 |
| Medium | 2621 | 2774 | 2599 | 2277 | 1913 | 12184 |
| Strong | 643 | 825 | 875 | 909 | 898 | 4150 |

This table represents one year buy-and-hold raw returns across B/M quintiles. F_SCORE is equal to the sum of the nine binary variables described in the table 1, where each binary equals one (zero) if the variable indicates a good (bad) future firm performance. An F_SCORE of nine (zero) indicates that the firm has the most (least) favourable fundamental signals. The high (low) F_SCORE portfolio consists of firms with an aggregate score of 8 or 9 (0, 1, or 2). The weak, medium and strong F_SCORE portfolios compromise firms with an aggregate F_SCORE of 0 to 3, 4 to 6, and 7 to 9, respectively. B/M quintiles are based on the preceding year's B/M ratios measured at individual firm's fiscal year end. *t*-Statistics for mean returns are from a one-tailed two sample test with unequal variances, where *, **, and *** indicate that the mean returns are significantly higher at the 10%, 5%, and 1% level. Raw returns are calculated as the 12-month buy-and-hold return of the firm starting at the first day of trading at the beginning of the seventh month after fiscal year end. Compounding ends one year later at the last day of trading at the end of the month, preceding the investment month.

## Table 4

### Buy-and-Hold Market-Adjusted Returns Across B/M Quintiles

***Market-adjusted Returns***

| F_SCORE | *low* 1 | B/M Quintiles 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | -23.57% | 48.44% | -36.26% | | | -8.74% |
| 1 | -20.80% | -17.77% | -22.98% | 1.95% | -16.50% | -15.80% |
| 2 | -1.64% | -10.76% | -3.97% | -2.58% | 0.48% | -3.87% |
| 3 | -8.80% | -6.31% | 2.43% | -4.89% | 1.13% | -3.89% |
| 4 | -4.32% | 5.19% | 2.79% | 2.78% | 6.11% | 2.26% |
| 5 | -0.39% | -3.22% | 3.94% | 0.76% | 6.78% | 1.20% |
| 6 | 0.98% | 0.83% | 2.06% | 5.58% | 7.51% | 3.15% |
| 7 | 3.89% | 1.94% | 3.62% | 3.40% | 6.43% | 3.79% |
| 8 | 4.78% | 8.34% | 7.46% | 1.58% | 3.55% | 4.93% |
| 9 | 31.48% | 13.70% | -13.55% | 17.81% | 13.18% | 11.76% |
| **ALL** | **-1.24%** | **0.12%** | **2.80%** | **2.32%** | **5.76%** | **1.78%** |
| High | 8.34% | 9.09% | 4.86% | 4.10% | 4.85% | 5.87% |
| Low | -5.48% | -11.34% | -7.03% | -1.79% | -2.16% | -5.85% |
| High - Low | 13.82% | 20.43% | 11.89% | 5.89% | 7.01% | 11.72% |
| *t*-Statistic | 1.685** | 3.435*** | 2.339*** | 0.676 | 1.172 | 3.615*** |
| High - All | 9.58% | 8.96% | 2.05% | 1.78% | -0.91% | 4.09% |
| *t*-Statistic | 2.497*** | 2.604*** | 0.641 | 0.594 | 0.378 | 2.989*** |
| Weak | -7.74% | -7.70% | -0.33% | -4.00% | 0.23% | -4.46% |
| Medium | -1.13% | 0.54% | 2.98% | 3.07% | 6.84% | 2.16% |
| Strong | 5.13% | 3.93% | 4.02% | 3.63% | 5.78% | 4.47% |
| Strong - Weak | 12.87% | 11.63% | 4.35% | 7.63% | 5.54% | 8.93% |
| *t*-Statistic | 3.498*** | 3.909*** | 3.909*** | 1.928** | 1.206 | 5.571*** |

***Number of Observations (n)***

| F_SCORE | *low* 1 | B/M Quintiles 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | 2 | 1 | 1 | 0 | 0 | 4 |
| 1 | 41 | 21 | 21 | 23 | 16 | 122 |
| 2 | 173 | 129 | 119 | 108 | 87 | 616 |
| 3 | 460 | 398 | 342 | 326 | 276 | 1802 |
| 4 | 780 | 787 | 720 | 598 | 549 | 3434 |
| 5 | 1034 | 1047 | 989 | 837 | 705 | 4612 |
| 6 | 807 | 940 | 890 | 842 | 659 | 4138 |
| 7 | 463 | 595 | 593 | 613 | 528 | 2792 |
| 8 | 156 | 198 | 247 | 250 | 320 | 1171 |
| 9 | 24 | 32 | 35 | 46 | 50 | 187 |
| **ALL** | **3940** | **4148** | **3957** | **3643** | **3190** | **18878** |
| High | 180 | 230 | 282 | 296 | 370 | 1358 |
| Low | 216 | 151 | 141 | 131 | 103 | 742 |
| Weak | 676 | 549 | 483 | 457 | 379 | 2544 |
| Medium | 2621 | 2774 | 2599 | 2277 | 1913 | 12184 |
| Strong | 643 | 825 | 875 | 909 | 898 | 4150 |

The table shows the one year buy-and-hold market-adjusted returns. Market-adjusted returns are the raw returns for each firm-year observation minus the returns to the FTSE All Share index over the same investment horizon as for that firm. All other variables as described in table 3.

return distribution for the high and the low B/M portfolio as well as for the entire sample.

Whereas Piotroski (2000) finds that high F_SCORE firms outperform the high B/M portfolio as well as low F_SCORE firms within the high B/M portfolio, we find that his strategy is not working within the high B/M portfolio. Although in terms of raw returns high and strong F_SCORE firms outperform low and weak F_SCORE firms by about 5 percentage points, respectively, the return differences are not significant. Also in terms of market-adjusted returns there are no significant return differences in the high B/M portfolio.

On the other hand, the results indicate that the F_SCORE strategy works well within the lower two B/M quintiles, in which return differences are economically and statistically significant. Within the lowest (second lowest) B/M portfolio high F_SCORE firms outperform low F_SCORE firms in terms of raw returns by 16.8 (25.0) and all firms in their respective quintile by 9.8 (11.5) percentage points. The results are fairly similar for market-adjusted returns, but most return differences are somewhat smaller. Apart from the return difference high minus low in the lowest B/M quintile, all return differences in the lowest and second lowest B/M quintile are significant at the 1% level (table 4). In addition, the growth stocks' returns show a perfect monotonic relationship with the F_SCORE except for firms with an aggregated score of 2.

Irrespective of B/M segmentation, the F_SCORE differentiates well between expected losers and winners over all sample firms. High F_SCORE firms outperform low F_SCORE firms in terms of raw (market-adjusted) returns by 14.2 (11.7) percentage points. All these return differences are highly significant.

Analysing the raw return distributions presented in table 5, our results indicate that the strategy is not shifting the *entire* return distribution within the value stock portfolio in the UK market. While F_SCORE clearly improves the median return and works well within the lower percentiles, the score does not differentiate between losers and winners within the stocks with higher returns. The 90th percentile of low F_SCORE is about 4.6 percentage points higher than the 90th percentile of high F_SCORE firms in terms of raw returns.

In contrast, it can be observed that investing in high F_SCORE firms within the low B/M portfolio shifts the entire return distribution to the right. The 10th percentile of high (low) F_SCORE firms is -38.8% (-75.5%) and the 90th percentile is 75.6% (70.0%).

Table 5

BUY-AND-HOLD RAW RETURN DISTRIBUTION

***Raw Returns in High B/M Portfolio***

| F_SCORE | Mean | 10th %ile | 25th %tile | Median | 75th %tile | 90th %ile | % positive | $n$ |
|---|---|---|---|---|---|---|---|---|
| **ALL** | **14.48%** | **-40.00%** | **-14.98%** | **2.83%** | **33.20%** | **69.58%** | **52.26%** | **3190** |
| High | 14.09% | -32.73% | -7.46% | 9.79% | 32.12% | 59.92% | 59.46% | 370 |
| Low | 9.06% | -45.63% | -26.83% | 0.00% | 27.51% | 64.55% | 42.72% | 103 |
| High - Low | 5.03% | 12.90% | 19.37% | 9.79% | 4.61% | -4.63% | 16.74% | |

***Raw Returns in Low B/M Portfolio***

| F_SCORE | Mean | 10th %ile | 25th %ile | Median | 75th %ile | 90th %ile | % positive | $n$ |
|---|---|---|---|---|---|---|---|---|
| **ALL** | **7.43%** | **-57.55%** | **-28.47%** | **0.00%** | **28.22%** | **68.29%** | **47.18%** | **3940** |
| High | 17.24% | -38.84% | -12.76% | 9.28% | 38.07% | 75.76% | 57.78% | 180 |
| Low | 0.45% | -75.54% | -47.95% | -14.45% | 21.93% | 70.08% | 36.57% | 216 |
| High - Low | 16.79% | 36.69% | 35.19% | 23.73% | 16.14% | 5.68% | 21.20% | |

***Raw Returns in All Sample Firms***

| F_SCORE | Mean | 10th %ile | 25th %tile | Median | 75th %tile | 90th %ile | % positive | $n$ |
|---|---|---|---|---|---|---|---|---|
| **ALL** | **10.70%** | **-48.16%** | **-20.04%** | **1.10%** | **29.88%** | **65.21%** | **50.50%** | **18878** |
| High | 15.77% | -36.94% | -10.96% | 7.47% | 35.72% | 73.28% | 57.73% | 1358 |
| Low | 1.58% | -69.81% | -41.33% | -1.76% | 22.24% | 64.72% | 39.62% | 742 |
| High - Low | 14.19% | 32.87% | 30.37% | 9.23% | 13.49% | 8.57% | 18.11% | |

This table shows the distribution of raw returns in the high and low B/M portfolio, and in the whole sample. All percentiles are calculated beginning with the lowest observation. All variables are as described in table 3.

Moreover, within the low B/M portfolio around 57.8% (36.6%) of the high (low) F_SCORE firms have positive stock returns and the median return increases monotonically with an increasing F_SCORE. Over all sample firms, the return distribution is shifted to the right. For all shown percentiles the high minus low difference is positive.

To sum up, the F_SCORE strategy works in the whole UK market. It differentiates successfully between out- and underperforming firms, especially in the growth stock portfolio. However, it does not provide significant results in the value stock portfolio. Moreover, the return differences between high and low F_SCORE firms are also substantially lower for the value stock portfolio than for the growth stock portfolio. In addition, the strategy does not shift the entire return distribution for the value stock portfolio, while it does for the growth stock portfolio and the entire market. These findings lead to the preliminary conclusion that the F_SCORE works well among growth but not among value stocks in the UK.

### 5.2.3 Returns Conditioned on Size

In table 6 the raw returns conditioned on size terciles and the F_SCORE are presented for the high B/M quintile (Panel A), the low B/M quintile (Panel B), and all sample firms (Panel C). As in Piotroski (2000) the size terciles for investments in year $t$ are based on all firms' market values of common equity at their fiscal year end in calendar year $t-1$. Differences in the number of observations per tercile occur for the same reason as for B/M quintiles. Also, conditioned on size the F_SCORE does not result in significant return differences within the entire value stock portfolio in our sample, but it differentiates between out- and underperformers (high-low F_SCORE firms) within the bottom third of market capitalisation at the 10% significance level. On the contrary, within the growth stock portfolio the F_SCORE works in the top two-thirds of market capitalisation with return differences being economically and statistically significant. Our results support the findings by Mohanram (2005) who shows that within a growth stock portfolio the F_SCORE and his similar performance metric GSCORE work best for larger companies. Across all sample firms the F_SCORE works best in the medium sized tercile, but generates also significant high minus low return differences in the other terciles. Return differences between high and all firms are not significant for large companies.

## 5.3 Analysis of Empirical Results

The results are surprising. Although the F_SCORE has been specifically designed to work for value stocks, it works best among growth stocks and worst among value stocks. Whereas Piotroski (2000) finds significant return differences between high and low (all) F_SCORE firms of 23.5 (7.4) percentage points within the value stock portfolio, we find no significant differences within this portfolio. Instead we find a significant return difference between high and low (all) F_SCORE firms of 17.2 (9.8) percentage points within the growth stock portfolio. In section 5.3.1 we analyse why the F_SCORE is potentially not working for value stocks and in section 5.3.2 we evaluate why the F_SCORE may work for growth stocks.

## Table 6

### Buy-and-Hold Raw Returns Across Size Terciles

#### Panel A: High B/M Portfolio

*Raw Returns*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **14.64%** | **11.69%** | **20.33%** | **14.48%** |
| High | 12.67% | 15.63% | 25.83% | 14.09% |
| Low | 2.39% | 23.58% | 16.19% | 9.06% |
| High - Low | 10.28% | -7.95% | 9.63% | 5.03% |
| $t$-Statistic | 1.329* | 0.496 | 0.755 | 0.804 |
| High - All | -1.96% | 3.94% | 5.50% | -0.38% |
| $t$-Statistic | 0.637 | 0.852 | 0.781 | 0.158 |
| Weak | 5.59% | 12.50% | 23.95% | 10.20% |
| Medium | 16.79% | 10.27% | 18.53% | 15.19% |
| Strong | 13.85% | 14.39% | 22.97% | 14.77% |
| Strong - Weak | 8.26% | 1.89% | -0.99% | 4.57% |
| $t$-Statistic | 1.17 | 0.284 | 0.108 | 0.969 |

*Number of Observations (n)*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **1946** | **878** | **366** | **3190** |
| High | 258 | 93 | 19 | 370 |
| Low | 65 | 22 | 16 | 103 |
| Weak | 224 | 96 | 59 | 379 |
| Medium | 1150 | 532 | 231 | 1913 |
| Strong | 572 | 250 | 76 | 898 |

#### Panel B: Low B/M Portfolio

*Raw Returns*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **11.51%** | **4.36%** | **7.88%** | **7.43%** |
| High | 20.70% | 19.17% | 15.18% | 17.24% |
| Low | 35.45% | -16.86% | -12.31% | 0.45% |
| High - Low | -14.75% | 36.03% | 27.49% | 16.79% |
| $t$-Statistic | 0.579 | 3.282*** | 3.978*** | 1.98** |
| High - All | 9.19% | 14.81% | 7.30% | 9.81% |
| $t$-Statistic | 0.727 | 1.631* | 1.883** | 2.483*** |
| Weak | 7.59% | -7.54% | -1.02% | -1.31% |
| Medium | 10.17% | 5.42% | 8.15% | 7.62% |
| Strong | 21.77% | 17.42% | 12.74% | 15.88% |
| Strong - Weak | 14.18% | 24.96% | 13.76% | 17.19% |
| $t$-Statistic | 1.243 | 4.246*** | 3.106*** | 4.465*** |

*Number of Observations (n)*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **792** | **1314** | **1834** | **3940** |
| High | 37 | 42 | 101 | 180 |
| Low | 66 | 87 | 63 | 216 |
| Weak | 183 | 272 | 221 | 676 |
| Medium | 477 | 865 | 1279 | 2621 |
| Strong | 132 | 177 | 334 | 643 |

#### Panel C: All Sample Firms

*Raw Returns*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **13.79%** | **8.80%** | **9.70%** | **10.70%** |
| High | 17.19% | 17.10% | 11.29% | 15.77% |
| Low | 5.88% | -4.25% | 3.47% | 1.58% |
| High - Low | 11.31% | 21.35% | 7.82% | 14.19% |
| $t$-Statistic | 1.602* | 4.754*** | 1.517* | 4.199*** |
| High - All | 3.41% | 8.29% | 1.59% | 5.07% |
| $t$-Statistic | 1.417* | 3.166*** | 0.789 | 3.592*** |
| Weak | 2.32% | 1.24% | 6.36% | 3.21% |
| Medium | 15.33% | 8.92% | 9.42% | 10.99% |
| Strong | 16.48% | 13.46% | 12.90% | 14.45% |
| Strong - Weak | 14.15% | 12.23% | 6.54% | 11.23% |
| $t$-Statistic | 4.056*** | 4.718*** | 2.656*** | 6.715*** |

*Number of Observations(n)*

| F_SCORE | *small* 1 | Size Terciles 2 | *large* 3 | ALL |
|---|---|---|---|---|
| **ALL** | **6017** | **6376** | **6485** | **18878** |
| High | 641 | 397 | 320 | 1358 |
| Low | 275 | 267 | 200 | 742 |
| Weak | 855 | 888 | 801 | 2544 |
| Medium | 3576 | 4146 | 4462 | 12184 |
| Strong | 1586 | 1342 | 1222 | 4150 |

Size terciles are determined for each calendar year by dividing the whole sample into three terciles based on the firm's market value of equity on their previous fiscal year end. All other variables are described in table 3.

**Table 7**

Correlation Analysis of Returns and F_SCORE Indicator Variables

| | *Profitability* | | | | *Financial Performance* | | | *Operating Efficiency* | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ROA | CFO | ΔROA | ACCRUAL | ΔLEVER | ΔLIQUID | EQ_OFFER | ΔMARGIN | ΔTURN | F_SCORE |
| *Low B/M Firms* | | | | | | | | | | |
| Raw Return | 4.27% | 5.19% | 3.37% | 1.82% | 2.60% | 1.06% | 4.32% | 1.22% | 2.86% | 8.07% |
| Market-adjusted Return | 2.04% | 3.56% | 2.61% | 1.70% | 2.52% | 0.28% | 4.74% | 1.07% | 2.26% | 6.29% |
| *High B/M Firms* | | | | | | | | | | |
| Raw Return | 0.08% | 2.38% | -2.21% | 1.28% | 0.90% | 0.42% | 3.78% | -0.53% | 3.66% | 2.75% |
| Market-adjusted Return | -0.30% | 1.61% | -3.12% | 0.29% | 2.05% | 0.59% | 4.58% | -0.96% | 2.84% | 2.12% |
| *Difference Low B/M - High B/M* | | | | | | | | | | |
| Raw Return | 4.19% | 2.81% | 5.58% | 0.54% | 1.70% | 0.64% | 0.54% | 1.74% | -0.81% | 5.32% |
| z-Value | 1.758** | 1.178 | 2.342* | 0.228 | 0.715 | 0.267 | 0.226 | 0.732 | -0.339 | 2.234** |
| Market-adjusted Return | 2.34% | 1.94% | 5.73% | 1.41% | 0.47% | -0.31% | 0.16% | 2.03% | -0.57% | 4.17% |
| z-Value | 1.758** | 1.178 | 2.342* | 0.228 | 0.715 | 0.267 | 0.226 | 0.732 | -0.339 | 2.234** |

The table shows the correlation between the F_SCORE indicator variables and returns in the two extreme B/M portfolios and their difference. The F_SCORE indicator variable prefix "F_" is omitted for succinctness. All variables are defined in table 3. z-Values for the difference in correlation are from the Fisher r-to-z transformation, where *, **, *** indicate that the correlation in the low B/M portfolio is significantly higher at the 10%, 5% and 1% (one-tailed p value) level than in the high B/M portfolio.

### 5.3.1 Value Stock Portfolio

A nearby conclusion is that the fundamental analysis strategy is not working for value stocks, because investors have already incorporated the information in a timely manner. This would imply that markets are efficient in the value stock portfolio.

The correlation table 7 as well as the table 5 showing the return distribution provide some further insights why the F_SCORE does not work in the value stock portfolio. Looking at the differences in correlations between the raw returns as well as the market-adjusted returns and the binary performance metrics, it can be seen that the correlation for $ROA$ and $\Delta ROA$ is significantly higher for growth than for value stocks. In fact, in the value stock portfolio the correlation of $ROA$ and $\Delta ROA$ with market-adjusted returns is negative, indicating that these profitability measures do not proxy future returns in the portfolio. The general weak correlations of the F_SCORE and its binary variables with future returns could imply that most of the available public information has already been incorporated in the stock prices. Furthermore, the return distribution for the value stocks presented in table 5 shows that the F_SCORE narrows the return distribution, but does not shift it to the right. This implies that the performance metric fails to identify the outperforming stocks within the high B/M portfolio.

However, so far we have seen that the F_SCORE works for all firms in our sample and within the lower B/M quintiles. In this respect, it would be surprising if the markets were efficient for value but not for growth stocks, since inefficiencies have usually been attributed to value stocks caused by low analyst following and neglection by investors (Lakonishok et al., 1994; Piotroski, 2000). Thus, apart from market efficiency, there could be two other main reasons why the performance metric does not work within the high B/M portfolio.

First, since most of the companies which delist during an investment are within the high B/M portfolio, Piotroski's assumption of zero returns in case of a delisting could significantly impact our observed returns for the high B/M quintile for two causes. On the one hand, his assumption is simplified, because stocks which stop trading have different returns depending on the delisting reason. Some of the companies may delist for performance-related reasons, whereas others delist due to a merger or an acquisition. While low F_SCORE firms are more likely to delist due to performance-related reasons, high F_SCORE firms are more likely to delist due to other reasons, such as an acquisition. Hence, when accounting more carefully for delisting returns, low (high) F_SCORE firms are likely to have lower (higher) returns. Implementing this could potentially show that the performance metric works also in the value stock portfolio. However, we do not apply different delisting returns, since we want to replicate Piotroski's study as closely as possible. On the other hand, assuming zero delisting returns my be inappropriate. As argued by Kaiser (1996) and Agarwal and Taffler (2008) it is common that shareholders do not receive any consideration in the case of performance-related delistings in the UK. However, adjusting all delisting returns regardless of the cause for delisting would not have any significant impact on our overall results. Nevertheless, more appropriate delisting returns, in other words distinguishing performance and non-performance-related delistings, may have altered our results. Due to a lack of access to the necessary data this cannot be tested.

Second, potential DS issues can influence the success of the F_SCORE in our sample. Especially among small companies DS could have a selection bias, meaning that it does not include firms which have quickly delisted again. Comparing DS to the CRSP database Ince and Porter (2006) report DS coverage and data integrity issues mostly among smaller firms. In our sample, this would especially matter for the high B/M quintile, because 61.0% of its firms are in the small size tercile. In contrast, the low B/M quintile consists of 20.1% small firms (table 6).

In summary, the investigation why the performance metric does not work in a value stock portfolio remains inconclusive. It may be due to market efficiency, simplified assumptions about delisting returns, or DS issues for small firms.

### 5.3.2 Growth Stock Portfolio

On the other hand, within the growth stock portfolio the F_SCORE does work. This observation is in line with our previous remarks about delisting returns and DS quality issues, as these should not affect the growth stock portfolio that strongly. In addition, although the performance metric has been originally designed for value stocks, it is not surprising that it works for growth stocks for two reasons outlined in the following paragrahs.

First, most of the binary F_SCORE variables proxy financial health and can be considered as positive signals not only for potentially distressed value stocks, but also for growth stocks. This applies to all binary variables, except for the financial performance signals $\Delta LEVER$ and $\Delta LIQUID$, since an increase in leverage or liquidity cannot necessarily be viewed as a bad signal for healthy, non-distressed companies. Furthermore, whereas $EQ\_OFFER$ most likely works for value stocks, it also indicates returns for growth stocks. Among value stocks, firms signal distress when issuing equity while their stock prices are depressed. Additionally, they also signal their inability to internally generate sufficient funds (Piotroski, 2000). Among growth stocks in turn, firms signal that their stock might be overvalued (pecking order theory as described in Myers & Majluf, 1984). Loughran and Ritter (1995) confirm this view by finding that firms issuing seasoned equity have lower subsequent stock returns.

Second, Mohanram's (2005) test results when applying the F_SCORE to the growth stock portfolio in the US market are similar to our findings. He shows that, whereas the F_SCORE works especially well among the bottom third of market capitalisation for value stocks, it works better for firms in the top two-thirds of market capitalisation within the growth stock portfolio. In addition, he shows that the effectiveness of financial analysis within the growth stock portfolio increases with analyst coverage. This is also in line with our findings, since size correlates with analyst coverage (Bhushan, 1989) and the performance metric works only in the top two-thirds of the growth stock portfolio.

Nevertheless, our findings for the growth stock portfolio are not fully in line with prior findings. According to Piotroski (2005) Mohanram's findings indicate that investors have overweighted the past growth-related performance attributes of growth stock firms, implying that a part of the growth stock portfolio is overvalued. He views this as the reason why Mohanram mainly identifies under- but not outperforming firms. However, since our findings show that the raw and market-adjusted returns of high F_SCORE firms within the growth stock portfolio are positive and considerably contribute to the hedge return, we cannot support this finding. In this respect, the mispricing nature seems to be different from the US market. Within the growth stocks, the strategy identifies overvalued firms, i.e. stocks with negative returns, and undervalued firms, i.e. stocks with positive returns.

To conclude, unsurprisingly the performance metric seems to work well for large and medium firms in the high B/M quintile. Contrary to prior research, we find that the F_SCORE cannot only identify under- but also outperforming firms.

# 6 Tests for Abnormal Returns

As shown in section 5.2, Piotroski's (2000) simple financial analysis strategy differentiates between out- and underperforming firms across all sample firms and within the higher B/M quintiles. However, additional tests are necessary to fully evaluate the success of the F_SCORE. These tests focus on whether the observed returns are abnormal, i.e. if the returns persist after adjusting for known risks. Such an additional analysis is necessary, since Piotroski's tests and argumentation are insufficient to evaluate the effectiveness of the strategy (see section 4). The F_SCORE is a summary metric of mostly accounting ratios. These ratios could vary across firms and cross-sectionally in a systematic manner as a function of the risk proxies B/M, size, and market beta. These risk proxies are in turn the determinants of expected returns as described in the three-factor model (see section 2.2). If the F_SCORE was indirectly related to risk, the observed returns of the strategy would solely be expected returns in line with risk compensation.

In general, there are three common approaches to investigate whether the observed returns are a compensation for risk or abnormal. One can specify an asset pricing model, match firm returns with risk characteristic-matched portfolio returns, or regress firm returns on firm characteristics known to capture the cross-section of returns as well as F_SCORE. In this section we test if the returns are abnormal using two out of the three approaches and analyse if abnormal returns persist over time.

## 6.1 Asset Pricing Models

First, one can specify an asset pricing model, based on historical data. In section 2.2 we have introduced the two common pricing models, the CAPM and the three-factor model. To test for abnormal returns one can regress the observed returns of an investment strategy on the risk factors, in other words the return premiums. Hereby, the intercept of the estimated regression represents abnormal returns. If the pricing model included all risk factors, regressing the returns of an investment strategy would not indicate abnormal returns in an efficient market. Hence, if one observes abnormal returns, the asset pricing model does either not include all risk factors or markets are not efficient.

For the UK, Hussain, Toms, and Diacon (2002) show that the three-factor model performs better than the CAPM. However, they also show that the three-factor model does insufficiently capture the risk. As outlined in the following paragraphs, this is confirmed by other researchers and therefore the model is inappropriate to explain stock returns in the UK stock market (e.g. Gregory, Harris, & Michou, 2001; Lee, Liu, & Strong, 2007; Michou, Mouselli, & Stark, 2007; Gregory, Tharyan, & Christidis, 2009).

Gregory et al. (2001) show that value portfolios substantially outperform growth portfolios in the UK after controlling for their loadings on the market, B/M, and size factors. They examine the UK market between 1975 and 1998 using accounting data from DS and stock prices from the London Share Price Database. Initially, they sort stocks into decile portfolios on the basis of the B/M ratio, earnings yield, cash flow yield, or past sales growth (one-way classification). Additionally, they sort stocks into three portfolios on the basis of their past sales performance first. Then, within each portfolio, they further sort stocks by B/M, earnings yield, or cash flow yield (two-way classification). They find that the return differences between value and growth portfolios can be explained by the three-factor model when using a one-way classification, but not when using a two-way classification. Regressing the returns of the extreme portfolios based on two-way classifications yields significant, positive intercepts. Thus, they reject the validity of the three-factor model for the UK market, since otherwise a simple two-way portfolio classification would easily yield abnormal returns.

Michou et al. (2007) argue that, while the estimation of the SMB and HML factors has become increasingly standardised in the USA, the estimation of the factors has so far been done in nine different ways in the UK. Hence, when analysing an investment strategy, the nine different ways may lead to different factor loadings and different conclusions about abnormal returns. In addition, Gregory et al. (2009) find significant intercepts (abnormal returns) for the UK market also when including momentum in the three-factor model. Therefore, the researchers recommend to use other available methods to test for abnormal returns instead.

In conclusion, regardless of the estimation method and also when including a momentum factor, the existing asset pricing models do not sufficiently explain the observed returns in the UK. Hence, we choose not to investigate the observed returns using an asset pricing model.

## 6.2 Characteristic-Matched Returns

Due to the lack of an appropriate asset pricing model for the UK market, other approaches to evaluate whether returns from investment strategies are abnormal have been studied. In a second approach characteristic-matched returns are calculated by adjusting the raw returns of a firm by the returns of a portfolio which encompasses firms with similar risk characteristics. If risk characteristics determined returns, selecting a firm and investing in a portfolio including firms with similar risk characteristics would yield on average the same expected returns in an efficient market. There has been an ongoing debate whether the co-variance used in asset pricing models or characteristics better capture risks (e.g. Daniel & Titman, 1997; Davis, Fama, & French, 1999). We have already argued that the common asset pricing models do not explain returns in the UK in section 6.1. However, it is still necessary to evaluate whether characteristics actually perform better than common asset pricing models.

Based on Daniel and Titman's (1997) work on the US market, Lee et al. (2007) evaluate whether characteristics or co-variance risk can better explain the size and value premiums in the UK market. They conclude that controlling for risk by matching a firm's returns with portfolio returns, which encompass stocks with similar B/M ratios and size, is more appropriate than using the three-factor model. They form two different types of zero-investment hedge portfolios: Characteristic-Balanced (CB) and Factor-Balanced (FB) portfolios. CB portfolios take long (short) positions in firms with high (low) ex-ante factor loadings and with similar size and B/M characteristics. On the contrary, FB portfolios go long (short) in stocks with characteristics related to higher (lower) returns (e.g. high B/M, small size), but with similar ex-ante factor loadings. If characteristics determined returns, the CB portfolio would have an average return of zero, while the FB portfolio would have significant positive expected average returns. Vice versa, if factor loadings determined returns, CB portfolios would yield a positive return and FB portfolios would yield zero returns. Lee et al. (2007, table 1) find that the CB portfolio has a positive but insignificant average monthly return (0.134%, $t = 1.081$) and that the FB portfolio has a positive and significant estimated average return of 0.226% ($t = -2.044$). Thus, they conclude that characteristics rather than co-variance risk explain the value premium in the UK stock market. Therefore, to measure abnormal returns, we use this approach and present first the methodology and then the empirical results.

### 6.2.1 Methodology

As previously defined, characteristic-matched returns are calculated by adjusting the raw returns of a firm by the returns of a portfolio which encompasses firms with similar risk characteristics.[7] These portfolio returns reflect the opportunity costs to investing in the firm, since both investments have similar risk. We form the benchmark portfolios based on the matching procedure as described in Hirshleifer et al. (2004). Each year we assign all firms to size quintiles based on the market value of equity at the end of their individual fiscal year. Next, we sort the firms further into B/M quintiles based on the financial information at fiscal year end, excluding firms with negative book values. This results in 25 portfolios. Contrary to Hirshleifer et al. (2004), we did not further re-divide our portfolios based on prior momentum for two reasons. First, we are more interested in investigating whether the returns are due to a risk factor and not a previously known anomaly, such as momentum. Second, we avoid too small benchmark portfolios by not re-dividing the portfolios. We have on average 1613 firms with sufficient B/M data per year, which are assigned to 25 portfolios with an average number of 64 firms. By dividing this portfolio in momentum quintiles, we would only have 12-13 firms left per portfolio. Furthermore, in order to have not only similar risk characteristics, but also similar transaction costs, we do not rebalance these portfolios during the one year holding period.

We assume that for each calendar year $t$, B/M and size quintiles are available on June 30, the date when the companies with the latest fiscal year end (December 31) must have published its audited financial statements. Hence, the B/M and size information in year $t$ is used for creating portfolio investments starting between July 1 in year $t$ and June 30 in year $t+1$. In each calendar year $t$ the firm $i$ is assigned to the risk characteristic-matched portfolio $w$.

Based on the 'Total Return Index' (RI) from DS, the portfolio return PORRET and

[7] In a variation of this approach the benchmark return is calculated as the return on a stock of a single firm with the most comparable risk characteristics. Using the return of such a control stock as a benchmark should lead to better specified test statistics (Barber & Lyon, 1997; Kothari & Warner, 1997). However, this is most relevant for 36 to 60 months holding periods. Since they also find that the reference portfolio method is generally more powerful in explaining abnormal returns, the control stock method is not used.

the corresponding characteristic-matched return CMRET for firm $i$ in year $t$ are:

$$\begin{aligned} \text{PORRET}_{w,t} &= \left[ \sum_{i_w=1}^{n} \frac{\text{RI}_{i_w,t_{exit,i}}}{\text{RI}_{i_w,t_{invest,i}}} - 1 \right] \times \frac{1}{n} && (7) \\ \text{CMRET}_{i_w,t} &= \text{RAWRET}_{i_w,t} - \text{PORRET}_{w,t} && (8) \end{aligned}$$

The characteristic-matched return for $i$ belonging to $w$ ($\text{CMRET}_{i_w,t}$) is computed as the difference between the firm's raw return and the equally-weighted return of all firms belonging to portfolio $w$ ($\text{PORRET}_{w,t}$) over the corresponding investment period, i.e. same *invest* and *exit* dates.

### 6.2.2 Empirical Results

If (1) markets are efficient, (2) characteristics determine returns, and (3) size and B/M are the only firm-specific risk characteristics, the adjusted return under the above outlined approach is zero. This should hold for any portfolios based on size, B/M, or F_SCORE. The difference of 0.02 percentage points between the observed characteristic-matched returns for the high and low B/M quintiles shows that this is the case for B/M partitions (see table 8). Since the difference is not significant, an investor cannot earn considerable abnormal returns based on a value strategy alone. Comparable return differences are observed for size tercile partitions (not tabulated).

If F_SCORE does not generate abnormal returns, one expects similar results for F_SCORE partitions. However, as shown in table 8, characteristic-matched returns deviate from zero and these returns still show a fairly positive relationship with the F_SCORE for all sample firms. The higher the performance indicator F_SCORE, the higher the characteristic-matched returns. The performance metric also successfully differentiates between out- and underperformers. High F_SCORE firms significantly outperform low F_SCORE firms as well as all other sample firms. Similarly, strong F_SCORE firms significantly outperform weak F_SCORE firms. However, compared to market-adjusted returns, we can also observe that the return differences are now mainly due to identifying underperformers. Strong and high F_SCORE firms show returns of only about 2%, while weak and low F_SCORE firms have negative returns of approximately -7%. This is important since capitalising on underperforming firms is often not possible for an investor, e.g. if no put options are available or if shorting stocks is associated with additional transaction

**Table 8**

BUY-AND-HOLD CHARACTERISTIC-MATCHED RETURNS ACROSS B/M QUINTILES

***Characteristic-Matched Returns***

| **F_SCORE** | *low* 1 | **B/M Quintiles** 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | -12.76% | 40.16% | -32.50% | | | -4.46% |
| 1 | -20.71% | -35.68% | -23.70% | 2.14% | -25.42% | -19.80% |
| 2 | -2.26% | -11.04% | -1.74% | -5.50% | -3.95% | -4.81% |
| 3 | -6.73% | -9.97% | -0.08% | -9.54% | -5.95% | -6.56% |
| 4 | -4.41% | 4.19% | -1.89% | -1.29% | -1.88% | -0.96% |
| 5 | -0.02% | -3.62% | 1.90% | -2.92% | -0.77% | -1.07% |
| 6 | 2.33% | 0.57% | 0.85% | 2.52% | 1.37% | 1.50% |
| 7 | 3.88% | 1.05% | 1.29% | 2.02% | 2.12% | 1.98% |
| 8 | 5.50% | 6.28% | 4.43% | -2.81% | -1.99% | 1.58% |
| 9 | 28.66% | -2.20% | -11.67% | 8.24% | 11.61% | 6.25% |
| **ALL** | **-0.62%** | **-0.99%** | **0.50%** | **-1.04%** | **-0.60%** | **-0.55%** |
| High | 8.63% | 5.10% | 2.43% | -1.09% | -0.15% | 2.22% |
| Low | -5.92% | -13.92% | -5.28% | -4.11% | -7.17% | -7.27% |
| High - Low | 14.55% | 19.02% | 7.71% | 3.02% | 7.02% | 9.50% |
| *t*-Statistic | 1.873** | 3.368*** | 1.578* | 0.374 | 1.227 | 3.12*** |
| High - All | 9.25% | 6.09% | 1.93% | -0.05% | 0.45% | 2.77% |
| *t*-Statistic | 2.493*** | 1.749** | 0.61 | 0.018 | 0.202 | 2.086** |
| Weak | -6.47% | -11.04% | -1.59% | -8.02% | -6.28% | -6.77% |
| Medium | -0.60% | 0.01% | 0.50% | -0.48% | -0.35% | -0.16% |
| Strong | 5.20% | 2.18% | 1.66% | 1.01% | 1.18% | 2.06% |
| Strong - Weak | 11.67% | 13.22% | 3.25% | 9.02% | 7.46% | 8.83% |
| *t*-Statistic | 3.329*** | 4.841*** | 1.123 | 2.424*** | 1.632* | 5.816*** |

***Number of Observations (n)***

| **F_SCORE** | *low* 1 | **B/M Quintiles** 2 | 3 | 4 | *high* 5 | ALL |
|---|---|---|---|---|---|---|
| 0 | 2 | 1 | 1 | 0 | 0 | 4 |
| 1 | 40 | 19 | 21 | 23 | 15 | 118 |
| 2 | 165 | 125 | 117 | 103 | 85 | 595 |
| 3 | 441 | 391 | 340 | 323 | 270 | 1765 |
| 4 | 765 | 778 | 713 | 596 | 544 | 3396 |
| 5 | 1013 | 1040 | 986 | 833 | 703 | 4575 |
| 6 | 797 | 936 | 887 | 840 | 658 | 4118 |
| 7 | 461 | 594 | 592 | 613 | 528 | 2788 |
| 8 | 154 | 198 | 246 | 249 | 320 | 1167 |
| 9 | 24 | 32 | 35 | 46 | 50 | 187 |
| **ALL** | **3862** | **4114** | **3938** | **3626** | **3173** | **18713** |
| High | 178 | 230 | 281 | 295 | 370 | 1354 |
| Low | 207 | 145 | 139 | 126 | 100 | 717 |
| Weak | 648 | 536 | 479 | 449 | 370 | 2482 |
| Medium | 2575 | 2754 | 2586 | 2269 | 1905 | 12089 |
| Strong | 639 | 824 | 873 | 908 | 898 | 4142 |

This table shows the one year buy-and-hold average characteristic-matched returns. Characteristic-matched returns are the raw returns minus the portfolio return, the firm belongs to, with the corresponding investment horizon of the firm. Compared to previous tables showing raw and market-adjusted returns, the sample size is slightly reduced, since some firms have no book value of equity or size information in year t, but B/M data in year t-1 and data to compute F_SCORE in year t. All other variables as described in table 3.

costs.

Within the value stock portfolio, the return differences are statistically insignificant, except for the difference between strong and weak firms, significant at the 10% level. However, the return differences for these firms are solely due to the identification of underperformers. On the other hand, within the growth stock portfolio high F_SCORE firms still outperform low F_SCORE firms and all differences are economical and highly statistically significant. In addition, these return differences are not due to the identification of underperforming firms. High (strong) growth F_SCORE firms have an average characteristic-matched return of 8.6% (5.2%).

In conclusion, also when adjusting returns with characteristic-matched portfolios the F_SCORE successfully differentiates between out- and underperformers over all sample firms and works especially well within the lowest B/M quintile. Compared to the previous analysis of market-adjusted returns (see table 4), the return differences over the entire sample are now mainly due to identifying underperforming firms.

### 6.3 Regression Analysis

In addition to the characteristic-matched return analysis, we continue to evaluate whether the observed returns are abnormal by regressing the individual firm returns (dependent variable) on firm characteristics associated with risk as well as on F_SCORE (independent variables). In an efficient market only risk is directly related to expected returns. As shown in section 5, the F_SCORE seems to be positively related to future returns since the higher the performance metric the higher the observed returns. In an efficient market this would mean that the F_SCORE is merely a proxy for risk. Since prior research has shown that size and B/M are the best proxies for risk and hence for expected returns (see section 2.2), the performance metric should not explain returns beyond these two risk proxies. We test for abnormal returns by estimating the regression with and without the F_SCORE variable. If markets are efficient and if the regression variables include already all risks, adding the F_SCORE as an independent variable should not significantly increase the explanatory power of the model. In addition, the F_SCORE itself should not get a positive, significant coefficient. Contrary results would indicate that the returns are abnormal. In the following we present the methodology for estimating the regression first and then our empirical results.

### 6.3.1 Methodology

To analyse whether the F_SCORE adds to explaining future returns beyond known risk factors, we estimate the following regressions, each with and without the F_SCORE as an independent variable.

$$\left.\begin{array}{l}\text{RAWRET}_{i,t} \textit{ or}\\ \text{MARET}_{i,t} \textit{ or}\\ \text{CAPMRET}_{i,t} \textit{ or}\\ \text{CMRET}_{i,t}\end{array}\right\} = \alpha + \beta_1 ln(\text{MVE}_{i,t}) + \beta_2 ln(\text{B/M}_{i,t}) + \beta_3 \text{F\_SCORE}_{i,t} \qquad (9)$$

MVE (B/M) is the market value of common equity (book value of equity scaled by MVE) at the fiscal year end in year $t$ of firm $i$ and CAPMRET is the CAPM-adjusted return, which is the raw return RAWRET adjusted by the expected return based on the CAPM described in section 2.2. CAPMRET is computed as follows:

$$Rf_{i,t} = \frac{RI_{Rf,t_{exit,i}}}{RI_{Rf,t_{invest,i}}} - 1 \qquad (10)$$

$$\text{CAPMRET}_{i,t} = \text{RAWRET}_{i,t} - \big(Rf_{i,t} + \beta_{i,t_{invest,i}} \times (\text{MKTRET}_{i,t} - Rf_{i,t})\big) \qquad (11)$$

$Rf$ is the risk free rate based on DS's RI for the UK one month Treasury Bill Tender and $\beta$ is estimated for firm $i$ at the day of investment $t_{invest}$ with DS data from a regression of monthly raw returns on the equal weighted market return index using up to 60 weeks preceding return data.

In the next paragraphs we present the methodology and rationale for estimating the regression by explaining first the choice of dependent variables, second the choice of independent variables, and third the general statistical methods.

The regression is done using individual firm-year returns, instead of portfolio returns, as dependent variables for two reasons. First, the set-up of the F_SCORE strategy does not permit building portfolios that invest and exit on the same date for all included stocks in a given year. Due to the differing fiscal year ends, the investment period differs between firms. In order to build suitable portfolios we would need to modify Piotroski's (2000) strategy so that every year the investment date would be the same for all stocks. This means that in each calendar year $t$ all stocks are invested on July 1 in year $t + 1$, six month after the last firm has published its financial statements. Considering that 32.7% of the sample firms have their fiscal year end before May, the investment strategy would be altered significantly. The accounting information used when investing may be outdated, since about one third of all firms' financial statements would have been publicly available for at least eight month already. Second, firm-year observations enable us to include individual size and B/M values directly in the regression as an independent variable. Using portfolios we would have to take an average value of size and B/M, which is less precise.

We use four different returns as the dependent variable. Next to the raw returns, market-adjusted returns are used in the pooled regression in order to control for the variations in returns between years.[8] If we did not control for the return variations over time, estimating a linear regression would be distorted. As before, the FTSE All-Share index returns are used for the adjustment. Using market-adjusted returns implicitly assumes that there is an equal co-variance risk, i.e. that all firms have a market beta equal to one. Since this is a simplified assumption, we also use the CAPM-adjusted returns. Regressing CAPM-adjusted returns on the firm's risk characteristics size and B/M, we include all three important risk variables in one regression model. The fourth return metric, the characteristic-matched return, is used to demonstrate on the one hand that the portfolios capture all the return variation due to size and B/M. On the other hand we test if the F_SCORE does have an incremental power in explaining characteristic-matched returns.

As independent variables we use the natural logarithm of the market value of equity and of the B/M ratio at fiscal year end for three reasons. First, continuous variables are more precise than percentile cut-offs, such as quint- or deciles. Using percentile cut-offs builds on the assumption that the size and B/M effect is constant and similar across all firms within each percentile portfolio. Second, cut-off points can only be arbitrarily determined. Third, we can confirm the results from previous studies that find that the logarithm of firm characteristics provides greater explanatory power for returns compared to using the absolute values (Amihud & Mendelson, 1986; Lamoureux & Sanger, 1989). The F_SCORE is not adjusted, since it can only amount to integers between 0 and 9.

[8] As an alternative, the market return could have been included as an independent variable in the regression of raw returns. The results (not shown) are similar to the one shown in table 9. The coefficients for market return and ln(B/M) are significant at the 1% level and for ln(MVE) at the 5% level before and after the inclusion of F_SCORE. F_SCORE has a 1% significant coefficient and adjusted $R^2$ significantly increases after its inclusion in the model.

Analysing the B/M ratios and returns for the 18878 observations for raw and market-adjusted returns as well as for the 18713 observations for characteristic-matched returns with scatter plots[9], we find a number of outliers. Based on the scattered plots, we exclude the most extreme outliers for B/M ratios and returns.[10] In addition, we exclude CAPM $\beta$ estimates below 0.2 and above 5.0.[11] Since the F_SCORE strategy builds on identifying exceptionally well performing firms, all outliers are neither removed nor down-weighted by using more sophisticated robust regression analysis. Since the Breusch-Pagan test and plotting the residuals against the returns show heteroskedasticity in our sample, we adjust the regression technique. As the sample size is reasonably large, we can use heteroskedasticity-robust *t*-statistics to correct for this as described in Wooldridge (2008).

### 6.3.2 Empirical Results

As explained previously, we estimate the regression with and without the F_SCORE variable. In efficient markets adding the F_SCORE as an independent variable to the regression should not significantly increase the explanatory power of the model. Also, the F_SCORE variable should not get a positive, significant coefficient.

The results of the regressions in table 9 show coefficients and the $R^2$ before and after including the F_SCORE as an independent variable in the model. Low $R^2$s are common in regressions when regressing return on risk characteristics instead of on risk premiums (three-factor model). It can be observed that the F_SCORE's coefficient is significant at the 1% level in the regression of raw returns. Furthermore, after including the F_SCORE the $R^2$ of the model increases. At the same time the significance of size decreases, but size is still significant at the 10% level. Controlling for return variations over time by regressing on market-adjusted returns, the F_SCORE still gets a highly significant coefficient. Also, the $R^2$ of the model increases, though being lower than in the raw return model. This

[9] As stated in table 8, the sample size for characteristic-matched returns is lower, since some sample firms have no B/M information in year $t$, but the data to compute F_SCORE in year $t$ and B/M in year $t-1$.

[10] We excluded 3 observations with B/M equal to 100, and 82 observations with annual returns larger 300%. Hence, the sample for raw and market-adjusted returns is reduced by 85. The sample for characteristic-matched returns is reduced by 84 observations. Not adjusting for outliers does not change our general conclusions regarding the effectiveness of F_SCORE. As an alternative, we adjust for outliers by winsorizing returns and B/M ratios at the 1% level. The choice of method does not affect our general results.

[11] This reduces our sample by 2465 firm-year observations, of which most of the observations are below 0.2 and negative. Not excluding these observations does not change our general conclusions.

**Table 9**

Regression Analysis of Individual Firm-Year Observations

| | Const. | ln(MVE) | ln(B/M) | F_SCORE | $R^2$ | n |
|---|---|---|---|---|---|---|
| ***Raw Returns*** | | | | | | |
| Coefficient | 0.078 | 0.004 | 0.049 | | 0.0077 | 18793 |
| *t*-Statistic | 3.82*** | 10.83*** | 2.05** | | | |
| Coefficient | -0.061 | 0.003 | 0.044 | 0.026 | 0.0144 | 18793 |
| *t*-Statistic | -2.51** | 1.88* | 9.79*** | 10.89*** | | |
| ***Market-Adjusted Returns*** | | | | | | |
| Coefficient | -0.029 | 0.004 | 0.034 | | 0.0041 | 18793 |
| *t*-Statistic | -1.51 | 2.60*** | 8.08*** | | | |
| Coefficient | -0.141 | 0.004 | 0.030 | 0.021 | 0.0090 | 18793 |
| *t*-Statistic | -6.15*** | 2.46** | 7.18*** | 9.26*** | | |
| ***CAPM-Adjusted Returns*** | | | | | | |
| Coefficient | -0.015 | 0.002 | 0.026 | | 0.0023 | 16328 |
| *t*-Statistic | -0.71 | 1.41 | 5.62*** | | | |
| Coefficient | -0.123 | 0.002 | 0.022 | 0.021 | 0.0068 | 16328 |
| *t*-Statistic | -4.82*** | 1.26 | 4.88*** | 8.26*** | | |
| ***Characteristic-Matched Returns*** | | | | | | |
| Coefficient | -0.107 | 0.008 | 0.010 | | 0.0012 | 18629 |
| *t*-Statistic | -5.59*** | 4.75*** | 2.25** | | | |
| Coefficient | -0.227 | 0.008 | 0.006 | 0.023 | 0.0070 | 18629 |
| *t*-Statistic | -10.16*** | 4.66*** | 1.32 | 10.20*** | | |

The table presents the results of regressing the individual firm-year return observation on size, B/M ratio and achieved F_SCORE. Size (B/M) is proxied by the natural logarithm of the market value of equity (B/M ratio) at the firm's financial year end prior to investing in the firm. All other variables as described in table 3. The equation of the regression is equation 9.

indicates that the previous model explains also some return variations over time.

The first two regressions only include the two risk characteristics size and B/M. To evaluate whether the F_SCORE proxies for market beta, we use CAPM-adjusted returns. Nevertheless, including F_SCORE in the regression of CAPM-adjusted returns, the performance metric still has a highly significant, positive coefficient and the explanatory power of the model increases.

When estimating the regression for characteristic-matched returns, size and B/M should have no significant coefficient, since we have already adjusted for B/M and size by substracting risk characteristic-matched portfolio returns from raw returns. However, we find that size and B/M still have significant positive coefficients. This suggests that the characteristic-matched returns do not fully capture the return variation due to B/M and size effects. This may result from using continuous, more precise size and B/M values when estimating the regression and less precise quintile cut-offs when calculating the risk characteristic-matched portfolio returns. Nevertheless, the F_SCORE's coefficient is highly significant at the 1% level and the explanatory power of the model increases.

Overall, F_SCORE's positive coefficients range between 2.1% (CAPM-adjusted and market-adjusted returns) and 2.6% (raw returns), suggesting that raw returns increase by 2.6% and risk-adjusted returns by 2.1% if the performance metric improves by one. Additionally, we observe that the $R^2$ always increases after adding the F_SCORE to the regression. F-tests comparing the fit of the models before and after including the F_SCORE indicate in all four cases that F_SCORE increases significantly the fit of models. Thus, it can be concluded that F_SCORE contains additional information not already captured by size, B/M, or market beta.

In summary, in all estimated regressions the F_SCORE's coefficient is significant and positive. Also, when including the performance metric, the fits of the models significantly increase. These findings show that the F_SCORE can predict returns even after controlling for size, B/M, and market beta, indicating that the returns are abnormal.

### 6.4 Returns Over Time

In the previous sections, we have shown that F_SCORE generates abnormal returns after controlling for known risk variables. However, abnormal returns tend to disappear once they have been discovered (Dimson & Marsh, 1999). In addition, they may be limited to earlier time periods, when they were merely a reward for the costs associated with cumbersome data gathering and complex computations. If this were the case, we would expect that the abnormal return differences between high and low F_SCORE firms disappear over time, latest after the publication of Piotroski's study (2000). In order to investigate this, we present characteristic-matched returns for the three six-year time periods 1991-1996, 1997-2002, and 2003-2008 (table 10) as well as for each year (table 11). Furthermore, we test if the success of F_SCORE is restriced to certain market conditions.

Table 10 shows that characteristic-matched return differences between strong and weak firms are significant over all three time periods. The return differences between high and low, as well as between high and all sample firms are only significant in the most recent time period (2003-2008). In general, characteristic-matched return differences rather increase over time. Between 2003 and 2008 high F_SCORE firms have outperformed low F_SCORE

**Table 10**

BUY-AND-HOLD CHARACTERISTIC-MATCHED RETURNS OVER THREE TIME PERIODS

| | ***Characteristic-Matched Returns*** | | | ***Number of Observations (n)*** | | |
|---|---|---|---|---|---|---|
| *Years* | *1991-1996* | *1997-2002* | *2003-2008* | *1991-1996* | *1997-2002* | *2003-2008* |
| **ALL** | **-1.20%** | **-0.43%** | **-0.02%** | **6281** | **6100** | **6497** |
| High | 0.46% | 1.97% | 4.22% | 460 | 436 | 462 |
| Low | -3.70% | -4.40% | -11.80% | 213 | 220 | 309 |
| High - Low | 4.16% | 6.37% | 16.02% | | | |
| $t$-Statistic | 1.029 | 0.79 | 4.332*** | | | |
| High - All | 1.66% | 2.40% | 4.24% | | | |
| $t$-Statistic | 0.776 | 0.933 | 1.93** | | | |
| Weak | -5.16% | -4.39% | -10.25% | 781 | 825 | 938 |
| Medium | -1.32% | -0.21% | 1.03% | 4037 | 4019 | 4128 |
| Strong | 1.23% | 1.40% | 3.50% | 1463 | 1256 | 1431 |
| Strong - Weak | 6.39% | 5.79% | 13.74% | | | |
| $t$-Statistic | 2.388*** | 1.837** | 6.574*** | | | |

This table shows the one year buy-and-hold characteristic-matched returns over the three time periods 1991-1996, 1997-2002, and 2003-2008. Each period includes firms which financial years ended during this period. All other variables as defined in table 3.

firms by 16.0 percentage points, an increase 11.9 percentage points compared to the period 1991-1996. Moreover, in the most recent time period high (strong) F_SCORE firms have 3.8 (2.3) percentage points higher returns compared to high (strong) F_SCORE firms between 1991 and 1996. Both differences are significant at the 10%-level ($t$=1.293 and 1.445, respectively).

However, the previous analysis is constrained by the fact that setting the time periods to three times six years is an arbitrary choice. Other more shorter time periods might show that return patterns have not increased over time. In this respect, we show characteristic-matched returns for high and low F_SCORE firms for each year $t$ in table 11. The returns for calendar year $t$ are not the returns an investor would have achieved in calendar year $t$, but the ones based on the F_SCORE computation in calendar year $t$.

Apart from the early nineties, high F_SCORE firms significantly outperform low F_SCORE firms fairly continuously. The investment strategy yields significant positive

**Table 11**

BUY-AND-HOLD CHARACTERISTIC-MATCHED RETURNS ACROSS TIME

| | High F_SCORE | | Low F_SCORE | | | |
|---|---|---|---|---|---|---|
| Year | Mean CMRET | n | Mean CMRET | n | Difference | *t*-Statistic |
| 1991 | 21.03% | 42 | -2.65% | 61 | 23.68% | 2.168** |
| 1992 | -3.98% | 80 | -6.45% | 33 | 2.47% | 0.24 |
| 1993 | 2.56% | 92 | -6.23% | 30 | 8.79% | 1.146 |
| 1994 | -3.84% | 105 | 2.33% | 22 | -6.17% | 0.595 |
| 1995 | -8.00% | 73 | -0.11% | 31 | -7.89% | 0.591 |
| 1996 | 6.03% | 66 | -7.96% | 33 | 13.98% | 1.592* |
| 1997 | 4.04% | 89 | -30.56% | 16 | 34.59% | 3.616*** |
| 1998 | -15.88% | 69 | 17.26% | 29 | -33.13% | 0.764 |
| 1999 | 1.83% | 69 | -16.01% | 26 | 17.84% | 1.679** |
| 2000 | 9.23% | 68 | -17.97% | 45 | 27.20% | 4.119*** |
| 2001 | 9.81% | 56 | -8.97% | 49 | 18.78% | 2.612*** |
| 2002 | 3.43% | 84 | 16.65% | 44 | -13.22% | 0.583 |
| 2003 | 1.12% | 107 | -19.97% | 41 | 21.08% | 2.145** |
| 2004 | 6.18% | 93 | -9.39% | 45 | 15.58% | 1.61* |
| 2005 | 9.41% | 72 | -25.63% | 45 | 35.04% | 3.27*** |
| 2006 | 1.17% | 72 | -10.52% | 52 | 11.68% | 1.678** |
| 2007 | 5.04% | 60 | -7.56% | 61 | 12.60% | 2.173** |
| 2008 | 3.29% | 57 | -2.11% | 54 | 5.41% | 0.471 |

The table shows the mean characteristic-matched returns and the corresponding number of observations as well as the difference between the high and low F_SCORE portfolios and its significance across all sample periods.

characteristic-matched returns in 14 out of 18 years. Furthermore, only investing in high F_SCORE firms also yields positive characteristic-matched returns in 14 out of 18 years, although their expected return equals zero.

Furthermore, in an additional analysis we test whether differing market conditions explain the variation of F_SCORE's success over time. The performance metric's success is measured in terms of the hedge returns from buying high and selling low F_SCORE firms over the entire sample. However, the correlation between F_SCORE's success and FTSE All-Share market returns is -0.30 but insignificant ($t$=-1.246). This means that we cannot find any support that F_SCORE works better under certain market conditions.

In summary, we do not find any indication for a decrease in abnormal returns over time. The differences between high and low F_SCORE firms remain significant and are not related to certain market conditions.

# 7 Conclusion

In this study we replicate Piotroski's (2000) simple financial analysis strategy and investigate whether it generates abnormal returns in the UK between 1991 and 2008. Using the so-called F_SCORE, which aggregates nine simple binary accounting-based proxies, the strategy identifies expected out- and underperforming firms. The decision to purchase a firm's stock is based on the strength of the F_SCORE signal, where a high (low) F_SCORE represents a buy (sell) recommendation. We show that this investment strategy generates abnormal returns in the whole UK stock market, especially in the growth stock portfolio. Contrary to Piotroski, we find that it does not work for value stocks alone.

Within the entire sample, high F_SCORE firms outperform low F_SCORE firms in terms of average annual market-adjusted returns by 11.7 and all other firms by 4.1 percentage points. Adjusting raw returns with risk characteristic-matched portfolio returns reduces the effectiveness of the strategy compared to only using raw or market-adjusted returns. Nevertheless, high F_SCORE firms still outperform low F_SCORE firms by 9.5 percentage points. Yet, the vast amount of the return difference is attributed to identifying firms with negative characteristic-matched returns now. This casts doubt on whether the full returns are actually realisable, since capitalising on underperforming stocks for long time periods is often associated with additional costs. On the other hand, only investing in high F_SCORE firms leads to a significant 2.8 percentage points return difference compared to investing in all other stocks.

Furthermore, we regress individual stock returns on the risk characteristics. By adding F_SCORE to the regression, we show that the performance metric explains returns beyond the risks captured by size, B/M, and market beta. This indicates that the returns to the strategy are abnormal. Additionally, we demonstrate that the success of the strategy does neither deteriorate over time nor is it linked to specific market cycles. In fact, the strategy generates positive risk-adjusted returns in 14 out of 18 years.

Focusing on the extreme B/M portfolios, within a growth stock portfolio the F_SCORE does successfully differentiate between out- and underperformers. High F_SCORE firms significantly outperform low F_SCORE firms in terms of market-adjusted returns by 13.8 and the entire growth stock portfolio by 9.6 percentage points. In addition, applying the strategy positively shifts the entire raw return distribution. In line with prior research, the F_SCORE works especially well among the top two-thirds of market capitalisation. When matching the firms' raw returns with risk characteristic-matched portfolio returns, the strategy still generates significant positive returns. High F_SCORE firms have a characteristic-matched return of 8.6% and outperform low F_SCORE firms by 14.6 percentage points. In contrast to prior research on the US market, these findings demonstrate that the possibility to short stocks is not required for an F_SCORE application within a growth stock portfolio.

On the other hand, within a value stock portfolio we do not find any evidence of a significant relation between F_SCORE and future returns. The strategy fails to identify firms which outperform all other value stocks. Also it does not successfully differentiate between out- and underperforming firms. These findings contrast Piotroski's (2000) results that the strategy works for value stocks in US market. In addition, while Piotroski claims that the strategy can shift the entire return distribution, we mainly observe a positive shift of the negative returns (left tail). However, using characteristic-matched returns we find that strong F_SCORE firms outperform weak ones by 7.4 percentage points. Yet, the difference is only significant at the 10% level and mainly attributable to short positions potentially entailing additional costs.

In conclusion, we find that in general Piotroski's strategy does successfully differentiate between out- and underperformers in the UK and that the observed returns cannot be explained by known risk factors. Thus, we conclude that the strategy based on simple financial analysis can generate abnormal returns in the UK.

Still, we cannot replicate the findings of the original study for the value stock portfolio. This impedes more general conclusions, so that further out-of-sample evidence from other large developed stock markets is needed. Besides, future studies could also explore whether the returns are caused by other, so far unknown risk factors. Simplified assumptions about delisting returns and potential data quality issues are limitations of this study. However, as argued in section 5.3.1, more precise delisting returns and data would most likely improve the results for the value stocks, but would not alter our overall conclusion for the whole market and the growth stock portfolio.

## References

Abarbanell, J. S., & Bushee, B. J. (1998). Abnormal returns to a fundamental analysis strategy. *The Accounting Review*, *73*(1), pp. 19-45.

Agarwal, V., & Taffler, R. (2008). Does financial distress risk drive the momentum anomaly? *Financial Management*, *37*(3), 461-484.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), pp. 589-609.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223 - 249.

Ball, R. (1992). The earnings-price anomaly. *Journal of Accounting and Economics*, *15*(2-3), 319 - 345.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, *6*(2), 159 - 178.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3 - 18.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, *43*(3), 341 - 372.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307 - 343.

Bartov, E., Lindahl, F. W., & Ricks, W. E. (1998). Stock price behavior around announcements of write-offs. *Review of Accounting Studies*, *3*(4), 327 - 346.

Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research*, *27*, pp. 1-36.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, *11*(2-3), 255 - 274.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), pp. 57-82.

Chan, L. K. C., & Chen, N.-F. (1991). Structural and return characteristics of small and large firms. *The Journal of Finance*, *46*(4), pp. 1467-1484.

Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, *51*(5), pp. 1681-1713.

Chen, N.-F., & Zhang, F. (1998). Risk and return of value stocks. *The Journal of Business*, *71*(4), pp. 501-535.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, *52*(1), pp. 1-33.

Davis, J. L., Fama, E. F., & French, K. R. (1999). Characteristics, Covariances, and Average Returns: 1929-1997 [Working Paper Series]. *SSRN eLibrary*.

DeBondt, W., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, *40*(3), pp. 793-805.

Dimson, E., & Marsh, P. (1999). Murphy's law and market anomalies. *Journal of Portfolio Management*, *25*(2), 53 - 69.

Duong, C. M., Pescetto, G. M., & Santamaria, D. (2010). How value-glamour investors use financial information: UK evidence of investor's confirmation bias [Working Paper Series]. *SSRN eLibrary*. Available from http://ssrn.com/abstract=1688255

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), pp. 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, *25*(2), 383 - 417.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, *10*(1), 1.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), pp. 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stock and bonds. *Journal of Financial Economics*, *33*(1), 3 - 56.

Fama, E. F., & French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The Journal of Finance*, *50*(1), pp. 131-155.

Foster, G. (1979). Briloff and the capital market. *Journal of Accounting Research*, *17*(1), pp. 262-274.

FSA. (2010). *Listing rules handbook.* Available from http://fsahandbook.info/FSA/html/handbook/LR

FTSE. (2010). *Factsheet - FTSE All-Share index.* Available from http://www.ftse.com/Indices/UK_Indices/Downloads/FTSE_All-Share_Index_Factsheet.pdf

Gregory, A., Harris, R. D., & Michou, M. (2001). An analysis of contrarian investment strategies in the UK. *Journal of Business Finance & Accounting*, *28*(9-10), 1192–1228.

Gregory, A., Tharyan, R., & Christidis, A. C. (2009). The Fama-French and momentum portfolios and factors in the UK [Working Paper Series]. *SSRN eLibrary*. Available from http://ssrn.com/abstract=1506960

Greig, A. C. (1992). Fundamental analysis and subsequent stock returns. *Journal of Accounting and Economics*, *15*(2-3), 413 - 442.

Haugen, R. A. (2009). *The new finance* (4th ed.). Prentice Hall.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, *38*, 297 - 331.

Hussain, S. I., Toms, S., & Diacon, S. (2002). Financial distress, market anomalies and single and multifactor asset pricing models: New evidence [Working Paper Series]. *SSRN eLibrary*. Available from http://ssrn.com/abstract=313001

Ince, O. S., & Porter, R. B. (2006). Individual equity return data from thomson datastream: Handle with care! *Journal of Financial Research*, *29*(4), 463–479.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), pp. 65-91.

Kaiser, K. M. J. (1996). European bankruptcy laws: Implications for corporations facing financial distress. *Financial Management*, *25*(3), pp. 67-85.

Kothari, S., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics*, *43*(3), 301 - 339.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, *49*(5), pp. 1541-1578.

Lamoureux, C. G., & Sanger, G. C. (1989). Firm size and turn-of-the-year effects in the OTC/NASDAQ market. *The Journal of Finance*, *44*(5), pp. 1219-1245.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *Journal of Finance*, *51*(5), 1715 - 1742.

La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *Journal of Finance*, *52*(2), 859 - 874.

Lee, E., Liu, W., & Strong, N. (2007). UK evidence on the characteristics versus covariance debate. *European Financial Management*, *13*(4), 742-756.

Lev, B., & Thiagarajan, S. R. (1993). Fundamental information analysis. *Journal of Accounting Research*, *31*(2), pp. 190-215.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), pp. 13-37.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *Journal of Finance*, *50*(1), 23 - 51.

Lovric, T., & Rados, D. (2010). I Piotroskis fotspar : Forslag pa forbattringar av Piotroskis hog book-to-market investeringsstrategier. *Uppsala University, Department of Business Studies*. Available from http://uu.diva-portal.org/smash/get/diva2:285865/FULLTEXT01

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), pp. 77-91.

Michou, M., Mouselli, S., & Stark, A. (2007). Estimating the Fama and French factors in the UK - An empirical review. *Manchester Business School Working Paper*. Available from http://www.mbs.ac.uk/cgi/apps/research/working-papers/view/paper.ashx?wId=112

Mohanram, P. S. (2005). Separating winners from losers among low book-to-market stocks using financial statement analysis. *Review of Accounting Studies*, *10*(2-3), 133 - 170.

Myers, S. C., & Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, *13*(2), 187 - 221.

Ou, J. A., & Penman, S. H. (1989). Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics*, *11*(4), 295 - 329.

Penman, S. H. (1991). An evaluation of accounting rate-of-return. *Journal of Accounting, Auditing and Finance*, *6*(2), 233 - 255.

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

Piotroski, J. D. (2005). Discussion of 'Separating winners from losers among low book-to-market stocks using financial statement analysis'. *Review of Accounting Studies*, *10*(2-3), 171 - 184.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *Journal of Portfolio Management*, *11*(3), 9 - 16.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), pp. 425-442.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, *71*(3), pp. 289-315.

Soares, N., & Stark, A. (2009). The accruals anomaly - Can implementable portfolio strategies be developed that are profitable net of transactions costs in the UK? *Accounting and Business Research*, *39*(4), 321–345.

Spiess, D. K., & Affleck-Graves, J. (1995). Underperformance in long-run stock returns following seasoned equity offerings. *Journal of Financial Economics*, *38*(3), 243 - 267.

Stoll, H. R., & Whaley, R. E. (1983). Transaction costs and the small firm effect. *Journal of Financial Economics*, *12*(1), 57 - 79.

Thomson Financial. (2007, April). *Worldscope database - datatypes definitions guide.*

Available from http://www.unisg.ch/de/Bibliothek/Recherche/Datenbanken/~/media/Internet/Content/Dateien/Bibliothek/Recherche/Datenbanken/ThomsonOneBanker/T1B_DatatypeDefinitionsGuide.ashx

Weetman, P., & Gray, S. J. (1990). International financial analysis and comparative corporate performance: The impact of UK versus US accounting principles on earnings. *Journal of International Financial Management & Accounting*, *2*(2-3), 111–130.

Wooldridge, J. (2008). *Introductory econometrics: A modern approach* (4th ed.). South-Western College Pub.

World Federation of Exchanges. (2010, January). *2010 WFE market highlights.* Available from http://www.world-exchanges.org/files/file/stats%20and%20charts/2010%20WFE%20Market%20Highlights.pdf