HOW VALUE–GLAMOUR INVESTORS USE FINANCIAL INFORMATION: UK EVIDENCE OF INVESTOR'S CONFIRMATION BIAS

Chau Minh Duong, Gioia Pescetto, Daniel Santamaria *

Abstract: The paper investigates investor's behaviour in the context of value–glamour investing and fundamental analysis, and provides a direct test of the confirmation bias by bringing together the evidence from several strands of literature into a well-defined framework of investor behaviour. The empirical evidence presented is in line with a model of investor's asymmetric reaction to good and bad news due to confirmation bias. Pessimistic value investors typically under-react to good financial information, but they process bad information rationally or over-confidently. On the contrary, glamour investors are often too optimistic to timely update prices following bad financial information, but they are likely to fairly price or even over-react when receiving good information.

Key words: *Value–glamour investing, financial statement analysis, contextual fundamental analysis, market efficiency, behavioural finance, confirmation bias*

---

* The authors are from Canterbury Christ Church University, North Holmes Road, Canterbury, Kent, CT1 1QU, United Kingdom. The corresponding author is Chau Duong, email: c.m.duong41@canterbury.ac.uk, from whom all the unreported results of the paper can be requested. We gratefully acknowledge comments from the participants of the Financial Management Association European Conference, the British Accounting Association Annual Conference, the Joint Conference of the British Accounting Association Northern Region Accounting Group and the Interdisciplinary Perspectives Special Interest Group. We especially acknowledge the comments from William Rees and Viet Nga Cao. All remaining errors are our own.

I. INTRODUCTION

The value-glamour effect is one of the most striking findings in the empirical finance literature. It is consistently documented that value stocks, which are those with high fundamental-to-price ratios, such as book-to-market ratio (BM hereafter), earning-price ratio and cash flows yields etc., outperform their lower ratios counterparts, known as glamour stocks (Fama and French, 1992, 1998; Lakonishok et al., 1994; Gregory et al., 2001). The strong evidence from the academic literature has inspired the widespread style-oriented investment strategies implemented by the investment community in the last few decades. However, despite the strong (weak) performance of value (glamour) portfolios, many value (glamour) stocks do not over-perform (under-perform) (e.g. Piotroski, 2000; Bird and Casavecchia, 2007). Thus, it is important to identify investment strategies that can recognise the better performing stocks in value and glamour portfolios as implied by simple value measures, such as the fundamental-to-price ratios.

In response to this need, one recent strand of literature employs fundamental analysis models to try improving the performance of value – glamour investing. The evidence is striking. Piotroski (2000) measures the overall firm's strength with a composite score, namely the F Score, that cumulates nine individual fundamental signals covering many aspects of the firm's past performance, including profitability, liquidity, capital structure, sources of funding and operating efficiency. Piotroski finds that, among high BM value stocks in the US during the period between 1976 and 1996, those with higher F Score tend to outperform their lower F Score counterparts by as much as 23% per year, and such outperformance persists for at least two years after the financial statements are released. In a similar vein, Mohanram (2005) designs a G Score from eight fundamental signals and finds that this measure of financial strength can differentiate winners from losers among low BM glamour stocks in the US stock market during the 1979-1999 period. Subsequently, Bird and Casavecchia (2007) use a dynamic model based on 24 accounting variables to predict the probability of a stock having an improved earnings-per-share performance, and they employ that probability as a measure of financial strength. Bird and Casavecchia's evidence is consistent with Piotroski's and Mohanram's conclusions that fundamental analysis can help differentiate "good" from "bad" value and glamour stocks. Needless to say, such evidence has important practical implications for investors, especially those using style-oriented investment strategies.

The evidence accumulated in the context of value-glamour investing and fundamental analysis, however, suffers from a major weakness related to the lack of understanding about the drivers of the observed return patterns. Most of the research in this area (Piotroski, 2000; Mohanram, 2005; Bird and Casavecchia, 2007) defines excess returns against a benchmark, such as the market portfolio or the corresponding size deciles. Bird and Casavecchia (2007) also find evidence that the Fama and French three-factor model cannot explain the observed returns. Although the debate is far from being settled, the cumulative evidence seems to reject risk-based explanations and suggests instead that behavioural biases may be responsible for the observed return pattern. In this context, the paper aims at making inroads into further explaining the source of the observed excess returns.

Following Ball (1992), who argues that the rejection of the efficient market hypothesis should emerge from the inability to reject a plausible inefficiency hypothesis, the paper makes a significant contribution to the understanding of the issues discussed above by offering a well-defined framework of investor's biased behaviour as an alternative hypothesis. In the extant literature, when the risk-based explanation is rejected, the value premium is often attributed to investor's over-reaction, while the abnormal returns from trading strategies based on fundamental analysis are claimed to be the result of investor's under-reaction. The question of how over-reaction can be reconciled to under-reaction in the context of value–glamour investing and fundamental analysis cannot be ignored. Thus the paper aims at contributing to the literature by addressing the following question: why do investors over-react to the future prospects of value and glamour firms and, at the same time, under-react to financial information about these same firms?

The paper hypothesises that the observed returns on portfolios formed on the bases of both value measures and fundamental analysis may be explained by behavioural biases. From a psychological viewpoint, value investors could potentially behave differently from glamour investors because value investors are typically pessimistic about the firm's future prospects, whereas glamour investors are on average optimistic. The literature on psychological biases in economic behaviour presents strong evidence of confirmation bias (e.g. Rabin and Schrag, 1999) and argues that people tend to misinterpret new information to support their prior beliefs. In this context thus, pessimistic value investors can be expected to be more biased when processing good information, and optimistic glamour investors to be more biased when processing bad information. The literature on value–glamour investing and fundamental analysis reviewed above provides further motivation for a test of investor's confirmation bias. Piotroski (2000) shows that, among value stocks, a hedge portfolio with a long position in financially strong stocks and short in weak stocks earns on average returns of 23% per annum, and these strong returns result mainly from the long position in financially strong stocks (13.4%). On the contrary, among glamour stocks, Mohanram (2005) shows that the contribution of a short position in financially weak stocks to the annual returns on the hedge portfolio dominates (17.9% out of 21.2%). Those observations suggest that, among value (glamour) stocks, those that are financially stronger (weaker) are more severely mispriced. A similar observation also derives from Bird and Casavecchia (2007). Overall, the cumulative intuition is that the pattern may follow a "rule", which is consistent with investor's behaviour suffering from a confirmation bias.

The paper provides a direct test of the confirmation bias by bringing together the evidence from several strands of literature into a well-defined framework of investor behaviour. More specifically, the paper tests the empirical predictions of the confirmation bias by applying Piotroski's (2000) and Moharam's (2005) models across sub-samples of value and glamour stocks, and by analysing whether value investors are most severely biased when they receive good financial information and glamour investors when bad information arrive. Piotroski's and Mohanram's models are originally tailored for and applied only among either value or glamour stocks. Both models concentrate on recent financial information and are constructed in a similar way by transforming financial signals into binary scores. Thus, applying them outside their designated contexts, simultaneously to both glamour and value stocks, can enhance the strength and robustness of the analysis and generate additional insights.

The paper contributes to the literature in at least three important ways. First and foremost, it is the first study to provide evidence of investor's confirmation bias in relation to the use of financial information in pricing value and glamour stocks. Not only such evidence has important practical implications for investors, but it also contributes a well-defined behavioural framework to the existing literature on fundamental analysis models applied to value–glamour investing. Second, the paper provides additional evidence on the determinants of abnormal returns from value-glamour investment strategies. Third, the evidence that both Piotroski's F Score and Mohanram's G Score perform indistinguishably well outside their designated contexts provides an additional reason to criticise the reliance on stock's styles when developing contextual fundamental analysis models. The strength of the evidence presented in the paper arises mainly from two sources. First, the paper argues that the methodology applied here is more appropriate than the methodologies used by the extant literature, given the well-documented problems associated with statistical tests on long-term buy-and-hold abnormal returns (Kothari and Warner, 1997; and Barber and Lyon, 1997). Specifically, the paper adopts Barber and Lyon's (1997) approach and uses control firms matched on book-to-market ratio and size to adjust returns for risk. This approach has been shown to be more appropriate for examining long-term abnormal returns. In addition, bootstrap procedures are also employed to reinforce the evidence obtained from the traditional parametric tests. Second, because existing studies on contextual fundamental analysis (Piotroski, 2000; Mohanram, 2005) restrict their investigation to samples of either value or glamour stocks, a test of the models' performance across both value and glamour contexts provides a better platform to explain the patterns of returns.

Consistent with much of the extant literature, the paper presents strong evidence that, among listed stocks in the UK during 1991-2007, value stocks earned higher returns and were more financially distressed compared to glamour stocks. The paper also finds evidence that financially stronger stocks earn higher returns than weaker stocks. The main, and more interesting, evidence rests with the application of fundamental analysis models across value–glamour contexts. First, the paper finds evidence that both Piotroski's F Score and Mohanram's G Score, despite being carefully tailored for value and glamour stock respectively, work equally well across different value–glamour contexts. This evidence implies that in practice the F Score and G Score can be employed in any context, and it challenges the value of using stock styles to define the framework for developing contextual fundamental analysis models. Second, the findings suggest that value (glamour) investors tend to under-react to good (bad) financial information, and either fairly react or overreact to bad (good) financial information, in line with the behavioural model of asymmetric reaction to information due to confirmation bias (Rabin and Schrag, 1999). Third, the evidence also suggests that the outperformance of financially strong stocks and underperformance of financially weak stocks are due to investor's behaviour rather than to different loadings on risk factors. Lastly, the efficiency of fundamental analysis strategies is found to be fairly consistent across time, even in bad states of the world[1] where the marginal utility of consumption is high. Such consistency strengthens the practical implications of the findings and gives confidence to those investors who want to apply the strategies in practice,

[1] Within the scope of the paper, periods of general negative performance of the stock market are defined as "bad". Please refer to section V.4. for more details.

especially in the light of the current financial crisis. Moreover, the consistency also lends very little support, if any, for the rational view of a risk-based explanation.

The paper proceeds as follows. Section II motivates and develops testable hypotheses, followed by the explanation of the sample selection procedures and the discussion of the descriptive statistics in section III. The main methodologies are explained in section IV. The results are presented and discussed in section V, while section VI concludes the paper.

II. HYPOTHESIS DEVELOPMENT

The main claim of Piotroski's (2000) model is that it performs particularly well when applied to US value stocks, while Mohanram's (2005) model is specifically designed for US glamour stocks. First, whether these models work in the UK market is still an open question. Second, there is a lack of evidence on the efficiency of the models when applied outside the universe of value and glamour stocks. If Piotroski's and Mohanram's claims hold, one would expect to observe a decline in the efficiency of the models when applied outside the contexts for which they were originally designed. The first hypothesis is tested using our UK sample to supplement these evidence:

H1a: Stocks that are financially stronger outperform those that are financially weaker.

H1b: The outperformance of high F Score over low F Score stocks is greatest among value stocks and decreases as stock style improves, while the outperformance of high G Score over low G Score stocks is greatest among glamour stocks and decreases as stock style deteriorates.

One of the most controversial aspects in the extant literature is the lack of an alternative theory to explain the stock return behaviour in the contexts of value and glamour stocks. When empirical evidence departs from the efficient market hypothesis, the most frequently cited alternative explanation is biases in investor behaviour. When the risk-based explanation is rejected, value premiums are often attributed to investor over-reaction (Lakonishok et al., 1994). The over-reaction model relates investor's behaviour to evidence about consistently biased human judgement errors, as documented in the psychology literature. It argues that investors tend to over-react to firm's past performance, and this results in lower (higher) than rationally justified expectations about the future performance of value (glamour) stocks, which typically have weaker (stronger) past performance. The lower (higher) expectations are then responsible for the under-pricing (over-pricing) of value (glamour) stocks, which causes a value premium to occur. Moreover, the success of fundamental analysis models is generally attributed to investor's under-reaction to published information due to psychological evidence about human conservatism bias (Edwards, 1968). This model predicts that investors fail to update their expectation timely when new information arrives. As a result, when good (bad) financial information are published, the price will not adjust instantaneously, but only slowly move up (down), creating an arbitrage opportunity and making it possible for fundamental analysis models to lead to abnormal returns.

Given the contradictory predictions of the over-reaction and under-reaction models, it is important to study the behaviour of investors who may rely on fundamental analysis to inform their value–glamour investing strategies. How can the over-reaction behaviour of value-glamour investors be reconciled with the under-reaction behaviour of investors who use fundamental analysis? In other words, why does the market over-react in a sample of value or glamour stocks, but then under-react to fundamental information relating to these same stocks? The paper proposes an investor's behaviour model that is built upon the psychological evidence on confirmation bias. This behavioural bias suggests that people tend to misinterpret information in a way that supports their current beliefs (Rabin and Schrag, 1999). Because value stocks are generally more financially distressed than glamour stocks (Fama and French, 1995), value investors typically hold more pessimistic believes than glamour investors. Therefore, a confirmation bias behaviour is consistent with value investors interpreting financial information in an overly pessimistic way and thus under-reacting to good information, in line with their pessimistic beliefs. Similarly, glamour investors who are affected by confirmation bias would under-react to bad information, since they hold optimistic beliefs. On the contrary, when new information is in support of investor's current beliefs, i.e. bad information in value context and good information in glamour context, investors process such information rationally. However, as argued by Rabin and Schrag (1999), confirmation bias could also lead to overconfidence. Therefore, when new information is in support of investor's current beliefs, some investors could be overconfident and thus over-react to such information. The paper investigates whether a confirmation bias in investor's behaviour can explain the observed pattern of stock returns by testing the following hypotheses:

> *H2a: Among value (glamour) stocks, those which are financially stronger (weaker) earn abnormally positive (negative) returns.*
>
> *H2b: Among value (glamour) stocks, those which are financially weaker (stronger) earn either no abnormal returns or abnormally positive (negative) returns.*

III. DATA

III.1 Sample Selection

The paper uses stocks listed on the UK stock market during the period 1991-2007. All data are collected from Datastream and Worldscope. The 1991-2007 period has been chosen to ensure data availability and consistency. In order to avoid survivorship bias, stocks that are delisted during the period are also included. Some stocks are excluded from the analysis according to the following criteria. First, financial firms are excluded. Second, companies with more than one type of ordinary share are also excluded. Third, being aware of the problems associated with Datastream's return data (Ince and Porter, 2006), the following items are also excluded: (i) items which are not companies or shares (e.g. ADR, Index, etc.); (ii) non-equity securities; (iii) companies incorporated outside the UK; (iv) shares not traded on the main UK stock exchange; and (v) non ordinary shares. Fourth, for stock to be included in the sample, all data needed to calculate all variables as described in the Appendix must be available. Moreover, observations with a negative BM are also excluded because negative BM could not be explained in terms of future expectations. Fifth, because calculating returns from the Datastream's Return Index suffers from errors when the level of Return Index and stock price is very small due to the discreteness of the Return Index and price data, stocks whose Return Index, prices and market capitalisation at the time of portfolio formation are very small are also excluded. This exclusion also frees the sample from very small firms, which are usually very illiquid. The thresholds for the exclusion of stocks with low Return Index, low price and small capitalisation are set at 0.10, £0.25 and £1 million, respectively[2]. Sixth, due to the intra-industry scoring scheme for some signals of the G Score, if the number of firms in an industry (as defined by Datastream level 2 industry classification) in any given year is below five, all firms in that industry are excluded. The above exclusion process produces a sample of 13,761 firm-year observations[3]. On a year-by-year basis, the number of observations is fairly stable, with a minimum of 679 observations in 1991 and a maximum of 937 observations in 1999. This stability mitigates concerns that the analysis results could be biased by a few influential years.

[2] Although the thresholds are chosen somewhat arbitrarily, thresholds of 0.5 for the Return Index, 5, 10, 20 pence for stock price, and £0.5 million for market value were also used and the unreported results show that the main findings are qualitatively the same.

[3] To mitigate concerns regarding the reliability of Datastream's returns data, the correlation between the annual returns on the value-weighted portfolio that holds all sample stocks and returns on the FTSE All Shares Index are examined (Ince and Porter, 2006). The two return series are found to be highly correlated with a correlation coefficient is 0.95, which is significant at 1% level.

III.2 Descriptive Statistics

Columns 2 to 6 of Table 1 provide some main descriptive statistics of the sample. The definitions of the variables are given in the Appendix. The statistics suggest the existence of some observations with very large DROA and DMargin. However, the magnitude of these variables does not matter because these figures are transformed into binary signals before being used in the main analysis. It is also notable that the means of F_ROA, F_CFO and F_Accrual are very large, which implies that most of the observations have good signals regarding profitability, cash flows and accruals (82.01%, 85.28% and 74.35%, respectively). This makes the distribution of observations across F Score heavily tilted toward larger values[4], which is very similar to the US evidence presented in Piotroski (2000). Moreover, the mean of the size is remarkably larger than the median (£817 million as compared to £75 million). This suggests the existence of some observations of very large size, which could dominate the returns on value-weighted portfolios. This characteristic of the sample motivates the use in the paper of an equally-weighted approach in portfolio's formation[5].

[Insert Table 1 here]

IV. METHODOLOGY

IV.1 Portfolio's Formation

In each year, stocks are ranked by F Score and G Score. Table 2 reports the distributions of firm-year observations across the F Score and G Score. The distribution

[4] See Table 2 and the related discussion in section IV.1.

[5] See section IV.1. for a more detailed discussion on the choice of the portfolio's formation approach.

across the F Score is very heavily tilted toward higher values, with very few observations in the lower values (only 4 and 98 observations with an F Score of 0 and 1, respectively). While the distribution of the G Score is more symmetric, there are fewer observations in the extreme values. Thus, rather than using only observations in the extremes to create portfolios of financially strong and weak stocks, as in Piotroski (2000) and Mohanram (2005), stocks with an F Score from 0 to 3 (G Score from 0 to 2) are grouped into the low F Score (G Score) portfolios; and stocks with an F Score from 7 to 9 (G Score from 5 to 7) are assigned to the high F Score (G Score) portfolios. In addition, stocks are also ranked yearly by BM, which is independent to the ranking by F Score and G Score. Stocks in the highest BM tercile are referred to as value stocks, those in the lowest BM tercile are glamour stocks and the rest are neutral stocks[6].

[Insert Table 2 here]

Three types of portfolios are investigated, namely: (i) the long portfolios, with long positions in high F Score and high G Score stocks; (ii) the short portfolios, with short positions in low F Score and low G Score stocks; and (iii) the hedge portfolios, with long positions in high F Score and high G Score stocks, and short positions in low F Score and low G Score stocks. Each year, portfolios are formed at the end of June to avoid any look-ahead bias. All positions are closed after one year from the date when the portfolio is

[6] Because the main interest lies in the sub-groups of financially strong and weak value stocks and financially strong and weak glamour stocks, sorting the sample into deciles or quintiles would make the sub-samples so small that the reliability of the statistical tests could be reduced. As shown in section V.I., sorting into terciles could reasonably differentiate value from glamour stocks in terms of returns and financial characteristics.

formed[7]. In the remainder of the paper, F Score and G Score strategies refer to trading strategies that hold one of the above portfolios.

The following illustrates in more detail the time convention used in the paper. Portfolios are formed at the beginning of July each year during the seventeen-year period 1991-2007. At the portfolio forming date, i.e. on the 1st of July in each year *t*, the book-to-market ratios are calculated using book value at the fiscal year-end of year $t – 1$ and market value at the end of December of year $t – 1$. The fundamental signals are calculated using the financial statements for the fiscal year ending in any month in year $t – 1$. Firm's size is measured as market value at the portfolio forming date. Returns are measured for the twelve-month period starting from the beginning of July of year *t*. Two issues concerning this approach need further discussion. First, the approach results in portfolios that comprise of firms with different year-end dates. The concern is that there could be a gap between the date of the financial statement release and the portfolio forming date (especially for those firms with early-in-the-year fiscal year-end dates) and ignoring returns during such period could be problematic (Bernard et al., 1997). The alternative is to measure returns from, say, six months after the (varying) fiscal year-ends. However, the chosen methodology in the paper represents a more practicable trading strategy because the portfolio forming date is fixed. Second, market values at the end of December are used to calculate BM ratios, which makes the numerator and the denominator of the ratio misaligned in terms of time. Fama and French (1992) argue that although this is not a perfect approach, switching to using market values at fiscal year-ends would not be preferable as it would suffer from the bias from market-wide variation in BM ratio during the year.

Fama (1998) emphasizes the importance of the issues associated with choosing an equally-weighted or a value-weighted approach in portfolio's formation. He shows that in

[7] In unreported results, when the two-year returns are measured, most of the main findings of the paper are not altered.

relation to many anomalies, changing the way of calculating portfolio's returns between an equally-weighted and a value-weighted approach can affect significantly the magnitudes of the anomalous returns, or even in some cases eliminate the anomalies. Although the equally-weighted approach is commonly used in the long-term anomaly literature, it suffers from the influence of very small stocks. Fama and French (2008) highlight the influential roles of very small stocks by showing that 60% of stocks listed on the NYSE, AMEX and NASDAQ can be classified as very small stocks, while they account for only 3% of the whole market capitalization. However, the value-weighted approach can also be problematic, as it can be influenced by very large stocks. Therefore, since there are some stocks of very large size in the sample[8], the paper employs an equally-weighted approach in forming portfolios[9].

IV.2 Measurement of Returns

The existing literature often emphasizes the importance of treating delisting returns, especially in studies on market anomalies (Shumway, 1997; Beaver et al., 2007). Shumway (1997) finds that delisting returns are commonly missed by the CRSP database for stocks that are delisted due to performance-related reasons. This observation reveals the delisting bias of CRSP and highlights the importance of a proper treatment of missing delisting returns. Many studies simply use a single value, usually -100% or 0, to replace missing delisting returns (e.g. Piotroski, 2000), regardless of the reasons. However, such approach is obviously inappropriate because the consequences of a performance-related delisting are clearly different from a delisting due to merger or acquisition. Shumway (1997) suggests that for the US market, if a delisting return is missing, rather than excluding the stock, a replacement value of -30% should be used if the delisting is performance-related, or zero

[8] See the discussion on descriptive statistics in section III.2.

[9] With reference to the concerns discussed above about the equally-weighted approach, unreported results show that most of the main results do not change significantly when the sample is free from micro stocks, defined as those stocks with market capitalizations below the 20th percentile of the sample.

otherwise. To illustrate the important sensitivity of tests of market efficiency to the treatment of delisting returns, Shumway and Warther (1999) find that, among NASDAQ stocks, the size effect disappears if performance-related delisting returns are properly treated. Rather than using a single replacement value, Beaver et al. (2007) adopt an approach that estimates different replacement values for different delisting reasons. They find that the premiums in the book-to-market, earning-to-price and cash-flows-to-price effects increase, and those in the accruals anomaly decrease, when delisting returns are properly handled. They also argue that the treatment of delisting returns can affect the estimate of market returns.

In the UK, however, although Datastream is among the most commonly used databases in accounting and finance research, there is little evidence of the existence of a delisting bias in the Datastream database and of the effects of different treatments of delisting returns on findings. A common approach is to adopt a single replacement value, where a -100% value is used as replacement if the delisting of a stock is performance related, and zero otherwise (Liu et al., 1999; Chi-Hsiou Hung et al., 2004). However, as noted by Beaver et al. (2007), this approach disregards the partial returns when the delisting date is not exactly at the beginning of the period of compounding returns. To be more specific, it is problematic to use a zero value as replacement in cases when the stock is delisted because of non-performance, since it does not take into account the returns from the portfolio formation date to the delisting date.

In the paper, raw returns are measured as the arithmetic growth of the Datastream's Return Index from the beginning of July of year *t* to the end of June of year *t + 1*. Based on the discussion above, if a stock is delisted within one year of portfolio formation, the raw return is calculated as:

$$RR_{t,t+1} = \left(1 + RR_{t,d}\right) \times (1 + DR) - 1 \tag{1}$$

where: $RR_{t,t+1}$ is the one-year raw return; $RR_{t,d}$ is the return from the portfolio forming date to the date when the stock is delisted[10]; DR is the delisting returns, which is -100% if the firm is delisted due to performance-related reasons and zero otherwise[11]. This procedure is equivalent to assuming that: (i) investors will lose all of their investment if the delisting is performance related; and (ii) if the delisting is not performance related, investors will receive the partial returns from the portfolio forming date to the date of delisting. While there is no obvious way to fully validate these assumptions, the first assumption is justified by the fact that it is common practice in the UK for stockholders not to receive any return after the stock is delisted (Kaiser, 1996; Agarwal and Taffler, 2008). With respect to the second assumption, it must be emphasised that the main focus of the paper is the analysis and comparison of groups of value and glamour stocks, and sub-groups of financially strong and weak stocks. Since there is no clear motive why a stock delisted because of non-performance related reasons should belong to a certain group of stocks, the findings of the paper should not be affected by this assumption. In other words, there is no reason to believe that this assumption would create a systematic bias in the results.

It is very common in the literature to measure abnormal returns by using a benchmark portfolio. For example, Piotroski (2000) uses the value-weighted market portfolio and Mohanram (2005) uses size deciles as benchmark. However, it is well documented that using referencing portfolios can make the test statistics severely misspecified, and that the null hypothesis is rejected more often than the theoretical rejection rates would suggest (Barber and Lyon, 1997; Kothari and Warner, 1997). Because the issue of whether the outperformance of financially stronger stocks over financially weaker stocks is due to risk is crucial in this area of study, this can significantly distort the results. Barber and Lyon (1997)

[10] It is assumed that this return is wholly reinvested.

[11] Delisting reasons are identified through scanning for key words in the Datastream's footnotes to company status.

compare different approaches to calculate abnormal returns, including those using reference portfolios, using the Fama and French three-factor model, and using the control firm matching approach. They document that the control firm matching approach is the best performing model in relation to alleviating most of the sources of bias, and it is therefore likely to provide the best well-specified test statistics. This evidence motivates the use in the paper of the control firm matching approach. Specifically, abnormal returns are measured by subtracting from the raw returns of each sample firm the corresponding raw returns of a control firm having similar BM and size. Following Barber and Lyon (1997), the market capitalization of the control stock falls between 70% and 130% that of the sample stock, and the BM is the closest to that of the sample stock. Nonetheless, to enable comparison between the paper's findings and the findings of prior studies, market-adjusted returns are also reported in the portfolio analysis[12]. However, in the regression analysis, only Barber and Lyon's (1997) control-firm-adjusted returns are discussed.

IV.3 Significance Testing

In the portfolio analysis, two parametric tests are performed. First, the one-sample t-test is employed to investigate whether portfolio's abnormal returns are significantly different from zero. This is applicable to all measures of abnormal returns, but not to raw returns. Second, the independent sample's t-test compares the mean returns between the long and short portfolios (i.e. it tests whether the returns on the hedge portfolios are significantly different from zero) and is applicable to all three measures of returns. However, in the context of long-run buy-and-hold abnormal returns, these parametric tests suffer from the departure of the abnormal returns from basic assumptions, including normality and stationary. Even under the Barber and Lyon's (1997) control-firm-matching approach, the problems could be pervasive as there is no evidence that such approach is effective outside

[12] Unreported results also show that the main evidence presented in the paper is very similar when returns are size-adjusted.

the contexts investigated by Barber and Lyon (Kothari and Warner, 1997). In such case, the bootstrap technique represents a reasonable alternative that can supplement the statistical inferences from the traditional parametric tests. Therefore, the paper also employs the bootstrap approach to further support the evidence from the parametric testing. The following illustrates the bootstrap procedure for testing the null hypothesis that market-adjusted return on the high-BM high-F-Score portfolio is zero (the same procedure is applied to all other return metrics and all other portfolios). First, observations with BM ratios in the highest tercile are drawn randomly and assigned to a pseudo portfolio. The random selection continues until the number of stocks in the pseudo portfolio is equal to the number of stocks in the actual high-BM high-F-Score portfolio. Second, the equally-weighted market-adjusted return on the pseudo portfolio is calculated, and this represents one observation of the empirical distribution of abnormal returns under the null hypothesis. Third, the process is repeated 1,000 times to create an empirical distribution of abnormal returns consisting of 1,000 observations. Finally, the two-tailed bootstrap p-value is estimated as:

$$p - value_b = 2 \times \min\{P(\bar{R}_i^* > \bar{R}), P(\bar{R}_i^* < \bar{R})\} \tag{2}$$

where: $p - value_b$ is the bootstrap p-value; $\bar{R}_i^*$ are the bootstrap mean returns of the pseudo portfolios (i = 1, 2 ... 1000); $\bar{R}$ is the mean return of the actual portfolio; P(.) is the probability function.

IV.4 Cross Section Regressions

The regression approach provides additional evidence of the power of the F Score and the G Score to predict future cross-sectional stock returns in value–glamour contexts, and whether such power is subsumed by other known risk factors, such as size and BM. Besides, a regression approach can provide a useful interpretation of how the F Score and the G Score are related to future returns. In the regression analysis, the control-firm-adjusted returns are used as the dependent variable. Firstly, the relation between abnormal returns and the F Score, and the G Score is examined using the following regression:

$$CFAR1 = \alpha + \beta_1 \log(BM) + \beta_2 \log(Size) + \beta_3 \text{Score} \quad (3)$$

where, CFAR1 is the control-firm-adjusted returns; BM is the book-to-market ratio; Size is the market value; and Score is either the F Score or the G Score[13]. This specification allows to test whether there is a relationship between financial conditions and abnormal returns.

In the second specification, the F Score and G Score are decomposed into two components to enable the investigation of the marginal effect of strong versus weak financial conditions on future returns. More specifically, the following regression is estimated:

$$CFAR1 = \alpha + \beta_1 \log(BM) + \beta_2 \log(Size) + \beta_3 \text{High\_Score} + \beta_4 \text{Low\_Score} \quad (4)$$

where, High_Score is either (i) the dummy variables that take the value of 1 if the stock is in the high F Score group, and zero otherwise; or (ii) the dummy variables that take the value of 1 if the stock is in the high G Score group, and zero otherwise; Low_Score is either (i) the dummy variables that take the value of 1 if the stock is in the low F Score group, and zero otherwise; or (ii) the dummy variables that take the value of 1 if the stock is in the low G Score group, and zero otherwise. This specification allows for the investigation of whether the relation between financial conditions and returns are driven by both/either financially strong stocks and/or financially weak stocks, and provides evidence about the confirmation bias.

Only in the regression analysis, all continuous variables are winsorized at the 1th and 99th percentiles to mitigate concerns regarding outliers. Following Fama and MacBeth (1973), regressions are run in relation to every year from 1991 to 2007, and the coefficients are the averages across 17 annual regressions. To gain statistical inference on the significance of

[13] Although control-firm-adjusted returns are already adjusted for size and BM as part of the matching process, we still introduce size and BM as control variables in the regressions to guard our findings against any remained size and BM effects that are not fully captured by the matching process.

the coefficients, standard errors are corrected for autocorrelation using the Newey and West (1987) methodology. Regressions are run in samples of value, neutral and glamour stocks[14]. The results are reported in Table 5 and discussed in section V.3.

V. RESULTS

V.1 Value Premiums and Fundamental Characteristics of Value–Glamour Stocks

Since the paper is premised on the existence of value premiums, it is important to firstly confirm this observation in the light of the sample adopted by the paper. Moreover, the main test of investor's confirmation bias is also based on the basic premise that value investors are more pessimistic and glamour investors are more optimistic. Therefore, whether or not value stocks are more financially distressed than glamour stocks in the sample used is very important, not only because it could provide justification for the contextual fundamental analysis approach using the value–glamour separation, but also because such observation could validate the basic premises underlying the main test of the paper. To provide such validation, the last four columns of Table 1 compare raw returns and fundamental characteristics of value stocks with those of glamour stocks - i.e., stocks in the highest versus those in the lowest BM tercile. The findings suggest that the RR1 of high BM (value) stocks are higher than those of low BM (glamour) stocks by 6.36% (statistically significant at 1% level), which confirms the well-documented value-glamour effect[15]. It is also shown that value stocks are more financially distressed, once again in line with earlier evidence (e.g. Fama and French, 1995). Specifically, value stocks tend to be less profitable with significantly lower ROA. With respect to the components of earnings, value stocks also

[14] Unreported results also show that pooled regressions would make no material change to the main conclusions in this section.

[15] See, for example, Fama and French (1992) and Lakonishok et al. (1994) for US evidence, and Gregory et al. (2001) for evidence on the UK market.

appear to have lower cash flows and larger accruals components. Moreover, value stocks are associated with lower capital expenditure and research and development expenses, which again signal more negative financial prospects. When being translated into binary signals, value stocks tend once again to be less profitable with lower F_ROA, F_DROA, F_Margin, G_ROA and G_CFO. Value stocks also have more volatile profits and sales growth, in addition to less capital expenditure and research and development expenses, and thus a worse prospect for future sales and earnings.

Although the individual binary signals generally suggest that value stocks are more financially distressed than glamour stocks, as discussed above, the overall firm's strength reveals a rather unpredictable picture. While value stocks have significantly lower G Scores than glamour stocks as predicted, surprisingly value and glamour stocks have indistinguishable F Scores. However, it can be noted that EQ_Offer is the main cause of higher F Scores for value stocks (the difference is 0.2831, which is significant at 1% level). While not issuing equity may signal good financial prospects, as proposed by the pecking order hypothesis, it can also be simply due to the constraints faced by distressed firms. Thus, it is hard to conclude that higher EQ_Offer makes value stocks less distressed than glamour stocks. Unreported results show that, as predicted, excluding EQ_Offer gives an average F Score for value stocks that is significantly lower than that of glamour stocks. To provide further evidence on whether value stocks are more financially distressed, Table 3 presents the correlations between BM, F Score and G Score. As predicted, the F Score and the G Score exhibit a strong and significant positive relationship as they are both designed to measure firm's financial strength. The evidence also shows that both F Score and G Score are negatively related with BM, which tends to suggest that value stocks are more financially distressed.

[Insert Table 3 here]

In summary, the evidence presented in this section: (i) confirms the existence of value premium in the sample; and (ii) suggests that value stocks are more financially distressed than glamour stocks. It suggests that the efficiency of fundamental analysis models could be enhanced by contextualising them to address the unique properties of value and glamour stocks. Piotroski's (2000) and Mohanram's (2005) models were originally designed to respond to this evidence. However, it is argued in the next section that these models are equally efficient when applied outside the contexts for which they were designed.

V.2 Portfolio Analysis

Table 4 reports returns on F Score and G Score strategies across sub-samples of value, neutral and glamour stocks, as well as a general sample containing all available stocks, together with the corresponding results from the significance tests. It can be observed across value–glamour contexts and return metrics that returns on the hedge portfolios are consistently positive. Under the parametric t-tests, the only occasions when the F Score premium (i.e. the outperformance of high F Score over low F Score stocks) is insignificant are in relation to glamour stocks when returns are market-adjusted, and value stocks when returns are control-firm-adjusted (t-statistics of 1.545 and 1.425, respectively). Similarly, the G Score premium is statistically significant in most cases, except for the VMAR1 among glamour stocks and CFAR1 among neutral stocks (t-statistics of 1.570 and 1.585 respectively). However, the bootstrap tests give very strong and consistent evidence of the statistical significance of returns on the hedge portfolios. In general thus, the evidence suggests that financially stronger stocks outperform their weaker counterparts, even after returns are adjusted for risks. Moreover, both F Score and G Score premiums appear to change unsystematically when the value measure, BM, increases. Furthermore, the unconditional application of an F Score and G Score strategy on BM (i.e. among all available stocks) also yields significant positive premiums. These observations are consistent across all return metrics. Thus, it is hard to conclude that the F Score is more efficient among value stocks and the G Score is more efficient among glamour stocks. In summary, the evidence thus far lends strong support to hypothesis *H1a* and tends to reject hypothesis *H1b*.

[Insert Table 4 here]

The evidence presented here is an interesting supplement to Piotroski (2000) and Mohanram (2005), suggesting that the F Score (G Score), although being aggregated from a range of fundamental signals that address the unique characteristics of value (glamour) stocks, works equally well outside the context of value (glamour) stocks. This finding also strengthens the US evidence documented by Fama and French (2006) that the F Score is effective in a general sample of stocks as a predictor of future cross-section stock returns. To the extent that Piotroski's F Score and Mohanram's G Score represent careful attempts to tailor fundamental analysis models to value and glamour stocks, the evidence in the paper challenges the appeal of separating stocks based on styles to build contextual fundamental analysis models.

Table 4 also shows evidence that strongly supports hypothesis *H2a*. Specifically, value stocks with a high F Score earn positive abnormal returns (VMAR1 is 0.0706 and CFAR1 is 0.0282), and glamour stocks with a low F Score earn negative abnormal returns (VMAR1 is -0.0746 and CFAR1 is -0.1111). Similarly, value stocks with a high G Score also earn positive abnormal returns (VMAR1 is 0.0825 and CFAR1 is 0.0784), and glamour stocks with a low G Score earn negative abnormal returns (VMAR1 is -0.051 and CFAR1 is -0.0578). All of these abnormal returns are statistically significant as inferred by both the parametric t-tests and the bootstrap p-values, suggesting that a typical value (glamour) investor would under-react to good (bad) financial information.

Moreover, Table 4 also lends strong supports to hypothesis *H2b*, although the picture is not entirely similar under the two abnormal return metrics (i.e. VMAR1 and CFAR1) and measures of overall firm's strength (i.e. F Score and G Score). Considering F Score strategies first, value stocks with low F Scores earn negative abnormal returns, but such abnormal returns seem not statistically significant. VMAR1 is -0.0419 with a relatively low t-statistic of 1.841 and high bootstrap p-value of 0.088 (i.e. the null cannot be rejected at 5% level), while CFAR1 is -0.0238, but it is clearly insignificant under both the parametric and bootstrap tests. On the contrary, for the strong glamour stocks, while the positive CFAR1 of 0.0329 seems insignificant with a relatively low t-statistic of 1.870 and a high bootstrap p-value of 0.104, the market-adjusted return seems to suggest a market overreaction (VMAR1 is -0.0044, which is significant at least under the bootstrap test with p-value of 0.068). Second, when firm’s strength is measured by the G Score, the overall evidence is also very much in line with hypothesis *H2b*. When returns are market-adjusted, there is evidence of value investor’s overreaction to bad financial information (VMAR1 on value stocks with low G Score is 0.0283, which is statistically significant under both the parametric and bootstrap tests) and glamour investor’s overreaction to good financial information (VMAR1 on glamour stocks with high G Score is -0.0071, which is significant under the bootstrap test with p-value of 0.000). Under CFAR1, hypothesis *H2b* is also supported, at least among value stocks with low G Score where CFAR1 is negative but insignificant (t-statistics is -0.217 and bootstrap p-value is 0.808). Generally, the evidence seems to suggest that value (glamour) investors’ reaction to bad (good) financial information is somewhere between a rational response and overreaction.

The support for hypotheses *H2a* and *H2b* can also be strengthened by looking at the contribution to the hedge portfolios of long positions in strong stocks relatively to that of short positions in weak stocks. The evidence suggests that the relative contribution of long positions in strong stocks to the hedge portfolios is greatest among value stocks. When firm’s strength is measured by the F Score, it is 63% if returns are measured net of market returns, or 54% as measured by CFAR1; the corresponding percentages are 152% or 95%, respectively, in relation to the G Score. It also systematically decreases while the stock style improves (i.e. when the cells go from left to right in Table 4), leading to the dominant contribution of short positions among glamour stocks (106% under VMAR1, or 77% under CFAR1 when F Score is used, and 116% or 61% if G Score is used). Such evidence further suggests that a typical investor would be most severely biased when new information arrives that is not in line with expectations, i.e. when good financial information reaches a value investor, or bad information is received by a glamour investor.

Overall, the evidence presented in this section lends strong supports for the hypotheses *H1a, H2a, H2b* and tends to reject hypothesis *H1b*. First, high F Score and G Score stocks outperform low F Score and G Score stocks across value, neutral and glamour contexts, and also in the general sample of stocks. Second, F Score and G Score strategies are not most effective when applied among value stocks and glamour stocks, respectively. Third, the evidence provides very strong support for the confirmation bias of value and glamour investors. This implies that the main benefit of financial statement analysis in a value context comes mainly from the identification of financially strong stocks, while the main benefit of analyzing glamour stock's financial statements is to avoid the poorly-performing firms. Moreover, unreported results also show that the evidence documented in this section is robust when: (i) sales-to-price ratio is used as the value measure[16], thus showing that the main findings are not unique to the use of BM as the value measure; and (ii) the sample is free from micro stocks (defined as stocks whose market capitalizations are below the 20th percentile of the sample), thus mitigating the concerns about employing an equally-weighted portfolio approach.

V.3 Regression Analysis

Table 5 reports the results of the regression analysis. The evidence from estimating equation (3) strongly suggests that there is a positive relationship between F Score, G Score

[16] Sales-to-price ratio is chosen for robustness reasons because other value measures are correlated with some components of the composite F Score and G Score and thus could introduce some bias into the analysis.

and abnormal returns, as the coefficients of the F Score and the G Score are all positive and statistically significant. This observation suggests that the predictive powers of the F Score and the G Score go beyond the well-documented size and BM effects, since both size and BM are included in the regressions. Furthermore, there is no clear pattern to suggest that these effects are greater or smaller when applied in a particular context, as the magnitude of the coefficients vary unsystematically when stock's style changes. Therefore, the evidence further supports the acceptance of hypothesis *H1a* and the rejection of hypothesis *H1b*, in line with the portfolio analysis in the previous section.

The evidence of confirmation bias is provided by estimating equation (4). If the confirmation bias exists, it is expected that: (i) among value stocks, the coefficients of High F Score and High G Score are significantly positive, and the coefficients of Low F Score and Low G Score are insignificant; (ii) among glamour stocks, the relation is driven mainly by financially weaker stocks, i.e. the coefficients of High F Score and High G Score are insignificant, and the coefficients of Low F Score and Low G Score are significantly negative; and (iii) the magnitude of the coefficients of High F Score and High G Score decreases systematically as stock style improves (i.e., when going from left to right in Table 5), while the magnitude of the coefficients of Low F Score and Low G Score decreases systematically as stock style deteriorates (i.e., when going from right to left in Table 5). Most of these empirical predictions of the confirmation bias are supported by the findings presented in Table 5. The coefficient of High G Score among value stocks is significantly positive, and the magnitude of this coefficient decreases as stock style improves, from 0.0649 among value stocks to 0.0233 among glamour stocks. Meanwhile, the coefficient of Low G Score among value stock is insignificant. Taken together with the significantly positive relationship between G Score and returns as discussed earlier, the evidence suggests that the relationship between G Score and returns is driven mainly by financially stronger stocks. On the contrary, the coefficient of Low G Score among glamour stocks is -0.0870 and it is significant at 1% level, and its magnitude decreases to as low as -0.0240 among value stocks and it becomes insignificant. Similarly, the coefficient of Low F Score is significantly negative among glamour stocks (-0.1092 and significant at 1% level) and it decreases as stock style deteriorates. Meanwhile, the coefficients of High F Score and High G Score among glamour stocks are both positive and statistically insignificant, suggesting that among glamour stocks the positive relation between strong financial conditions and returns is subsumed by the stronger negative relation between weak financial conditions and returns. In short, aside from the only exception of the statistically insignificant coefficient of High F Score among value stocks, the general evidence as reported in Table 5 provides striking evidence in support of the confirmation bias and hypotheses *H2a* and *H2b*.

[Insert Table 5 here]

Overall, the regression analysis provides evidence which is fully in line with the findings documented using portfolio tests in the previous section. The consistency of the evidence across different methodologies leads to our conclusion that the confirmation bias does exist.

V.4 F Score and G Score Strategies Across Time

This section tests the efficiency of F Score and G Score strategies across time. The purpose of this analysis is twofold: (i) to give confidence to investors who are interested in applying these strategies, the consistency across time is important, especially in the context of the current financial crisis; and (ii) the test across time could shed light on the debate central to the paper between risk-based versus behavioural motives. Although the evidence from both portfolio and regression analyses seem to be inconsistent with a risk-based explanation, it could be argued that the observed returns could be due to some omitted risk factors, in addition to size and BM. Although this possibility cannot be conclusively excluded within the scope of this paper, the findings presented in this section should mitigate the concerns. Lakonishok et al. (1994) argue that for value stocks to be riskier than glamour stocks, glamour stocks should outperform value stocks in some states of the world, and these states should on average be ‘bad’ states. In ‘bad’ states of the world, a risk-averse investor should prefer glamour stocks to value stocks, if value stocks are indeed riskier. Following this line of argument, this section further investigates whether high F Score and G Score stocks are riskier than low F Score and G Score stocks.

Table 6 presents the raw returns on the high and low F Score and G Score stocks, together with the returns on the hedge portfolios that long (short) stocks with high (low) F Score and G Score. The findings show that over the sample period, high F Score stocks outperform low F Score stocks in 16 out of 17 years, the return difference being statistically significant in 11 out of the 16 years, except for 1999 where the hedge return is negative but insignificant. High G Score stocks also outperform low G Score stocks in 12 years, and the return difference being statistically significant in 9 out of the 12 years. The second column in Table 6 classifies the state of the stock market into either ‘good’ or ‘bad’. There are different views and definitions of a ‘good’ or ‘bad’ market condition. The paper identifies a ‘good’ market condition with a stock market in an upward trend, and vice versa a ‘bad’ state when the market is falling (Chi-Hsiou Hung et al., 2004). Accordingly, within the scope of the paper, a ‘good’ (‘bad’) market is identified when the contemporary return on the FTSE All Shares Index is positive (negative). Using this classification, the years 2000, 2001, 2002 and 2007 are classified as ‘bad’ years, and in all these ‘bad’ years high F Score and high G Score stocks consistently outperforms stocks with weaker financial performance.

[Insert Table 6 here]

In summary, the analysis in this section provides further challenges to the view that investors base their decisions on risk-based motives. Moreover, the evidence on the consistent performance of the F Score and G Score strategies across time could lend more confidence to investors who are keen to apply these strategies in practice.

VI. CONCLUSIONS

The paper investigates the performance of fundamental analysis strategies across value and glamour stocks. Piotroski's (2000) F Score and Mohanram's (2005) G Score are the two fundamental analysis models employed, as they are originally designed to suit value and glamour stocks, respectively. Using a sample of UK listed stocks during the period 1991-2007, the paper firstly compares returns and fundamental characteristics of value and glamour stocks. Earlier evidence (Fama and French, 1992, 1995; and Gregory et al., 2001) is confirmed by the findings that the value premium exists and that value stocks are indeed more financially distressed.

The main focus of the investigation then turns to the question of how fundamental analysis strategies perform in different value and glamour contexts. While there is strong evidence that firms with stronger financial position outperform weaker stocks, there is very little evidence that such effect is limited to a specific group of stocks with similar value or glamour characteristics. Thus, the F Score and the G Score could be used as a fundamental analysis model for any type of stocks. While Piotroski (2000) and Mohanram (2005) include careful and reasonable procedures to tailor the models for either value or glamour stocks, the evidence in the paper challenges the feasibility of tailoring fundamental analysis models to value and glamour stocks. There is also little evidence in support of a risk-based explanation for the outperformance of financially strong stocks.

More interestingly, the evidence documented in the paper is consistent with a behavioural investment model, where due to confirmation bias (1) value investors under-react to good information while they fairly react or even overreact to bad information, and (2) glamour investors under-react to bad information but good information is processed quite efficiently or even overconfidently. This investment behaviour implies that in the context of value stocks, the main benefit of analyzing financial statements is to identify the financially strong stocks, while fundamental analysis among glamour firms is mainly aimed at avoiding the poor performing firms. The framework fits well with the practice of value and glamour investing. While restricting their attention to “cheap” stocks, which are typically financially distressed, value investors are seeking the ‘dusty gems’, i.e. those distressed stocks that are actually worth more. In contrast, glamour investors, whose money are invested in safe places with those well-known growth stocks, need to ensure that they exclude ‘falling stars’, that is those traditionally good companies which are facing recent financial problems. The fundamental analysis models investigated in the paper can help value and glamour investors analyse recent financial statements and identify the ‘dusty gems’ and the ‘falling stars’. In addition, the findings presented here also have useful practical implications, especially for value – glamour investors. Our evidence suggests that in value–glamour contexts investor’s behaviour is asymmetric and dependent upon the recent financial information released. In a value context, a typical value investor updates good news only slowly, while bad news is often reflected fairly or even too quickly. In contrast, in a glamour context, the release of good news is expected, thus it is often reflected instantly or even too quickly into stock prices, while bad news travel too slowly. Therefore, a value or glamour trading strategy that exploits this sup-optimal behaviour could increase profitability.

Finally, the evidence presented in the paper of investor’s confirmation bias creates some interesting research opportunities. For example, researchers may ask how managers respond to a market that is more willing to accept information that confirms its priors. If the confirmation bias hypothesis holds, and if managers understand this behaviour, it could be expected that at times manager may try to report financial information that is in line with market expectations through, for example, the mechanism of accruals management, or even real operations management. Evidence along these lines may further support the existence of the market confirmation bias and provide interesting knowledge about manager's behaviour in respond to the behavioural bias of the market.

APPENDIX

Definitions of Variables

RR1 is the buy-and-hold raw returns, measured by the arithmetic growth of Datastream's Return Index from July of year *t* to June of year *t + 1*. If a stock is delisted in one year after portfolio formation, raw returns are calculated as: $RR_{t,t+1} = (1 + RR_{t,d}) (1 + DR) – 1$ (where: $RR_{t,t+1}$ is one-year raw return, $RR_{t,d}$ is return from portfolio forming date to the date when the stock is delisted, DR is delisting returns, which is -100% if the firm is delisted due to performance-related reasons and 0 otherwise).

VMAR1 is the one-year market-adjusted returns calculated as the difference between raw returns and returns on the value-weighted market portfolio.

CFAR1 is the control-firm-adjusted returns calculated as the difference between raw returns and the corresponding returns on the control stock. Following Barber and Lyon (1997), for each sample stock, the control stock is the stock whose market capitalization is between 70% and 130% that of the sample stock, and whose BM is closest to that of the sample stock.

Size is the market value of the firms at the end of June.

BM is the book-to-market ratio, calculated as book value of common equity at the end of the fiscal year divided by the market value at the end of December.

ROA is the returns-on-assets ratio, calculated as the net income, before extraordinary income and after preferred dividend, scaled by initial total assets.

DROA is the change in ROA since the last fiscal year.

CFO is the cash flow from operations[17] scaled by initial total assets.

Accrual is the total accrual component of earning, calculated as ROA less CFO.

DLever is the change in debt ratio from the previous year, where debt ratio is the ratio of long-term debt, including the portion of long-term debt classified as short-term debt, to average total assets.

[17] If a firm has a Cash Flow Statement, net cash flow from operations is used. If a firm does not have a Cash Flow Statement for a year, cash flow from operations is estimated by: CF = OpFund – (DCA – DCash) + DCL (where: CF is cash flow from operations, OpFund is funds from operations, which include net income plus all non-cash charges, such as depreciation; DCA is change in current assets from last fiscal year; DCash is change in cash and cash equivalence from last fiscal year; DCL is change in current liabilities from last fiscal year).

DLiquid is the change in current ratio from the previous year, where the current ratio is the ratio of current assets to current liabilities.

DMargin is the change in gross margin ratio from the previous year, where the gross margin ratio is the ratio of gross income (which equals to net sales less cost of goods sold[18]) to net sales.

DTurn is the change in asset turnover ratio from the previous year, where the asset turnover ratio is the ratio of net sales to average total assets.

SALESVAR and ***ROAVAR*** is the variance of a firm's sales and ROA during the last five years[19].

CAPEX is the capital expenditure on assets ratio, calculated as capital expenditure scaled by initial total assets.

RD is the research and development expenditure on assets ratio, calculated as research and development expenditure scaled by initial total assets.

If a firm has positive *ROA, DROA, CFO, Dliquid, DMargin, DTurn* or negative *Accruals* and *DLever,* ***F_ROA, F_DROA, F_CFO, F_Dliquid, F_DMargin, F_DTurn, F_Accrual, F_DLever*** are one, and zero otherwise. ***EQ_Offer*** is one if there is no seasoned common equity offering during the year prior to portfolio formation, and zero otherwise[20].

[18] When cost of goods sold is unavailable from Datastream, it is assumed that the corresponding data equals to zero. The validity of this assumption is *ad hoc* checked as follows. First, a list of observations is compiled, which satisfy all other sample selection requirements but do not have data for cost of goods sold. It is noticed that many of the missing information relates to either the services industry, or to firms in the very early stage (within one year) of incorporation (152 out of 263 observations), which could potentially mean the corresponding firms simply have zero cost of goods sold in that year. Second, the actual annual reports of a random sub-sample of the compiled list are checked and this shows that the unavailability of observations on Datastream is more likely to be due to a zero value rather than a real missing data. Therefore, unreported data are treated as having a value of zero. The same treatment is applied to other data that are likely to be zero, including long-term debt, capital expenditure, research and development expenditure, and net proceeds from issuing stock.

[19] If there is not enough data for the last five years, data for at least three most recent years are used.

[20] A seasoned equity offer is identified when: (i) the firm has positive net proceeds from issuing common or preferred stock in the last fiscal year; and (ii) the firm's number of outstanding common shares has increased by at least 0.1% since the last fiscal year.

If a firm has *ROA, CFO, CAPEX, RD* larger than the industry median or *SALESVAR* and *ROAVAR*[21] smaller than the industry median, ***G_ROA, G_CFO, G_CAPEX, G_RD, G_SALESVAR, G_ROAVAR*** are one, and zero otherwise.

F Score = F_ROA + F_DROA + F_CFO + F_Accrual + F_DLever + F_DLiquid + F_DTurn + F_DMargin + EQOffer

G Score[22] = G_ROA + G_CFO + F_Accrual + G_ROAVAR + G_SALESVAR + G_CAPEX + G_RD

[21] If data for at least three most recent years are not available, *G_SALESVAR* and *G_ROAVAR* are assumed to be zero.

[22] The G Score in this study is not exactly defined as in Mohanram (2005), because one of Mohanram's signals relates to advertising expenses and this data is not available from Datastream.

References

Agarwal, V., and R. Taffler (2008), 'Does Financial Distress Risk Drive the Momentum Anomaly?', *Financial Management,* Vol. 37, No. 3, pp. 461-484.

Ball, R. (1992), 'The Earnings-Price Anomaly', *Journal of Accounting and Economics,* Vol. 15, No. 2&3, pp. 319-345.

Barber, B.M., and J.D. Lyon (1997), 'Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics', *Journal of Financial Economics,* Vol. 43, No. 3, pp. 341-372.

Beaver, W., M. McNichols, and R. Price (2007), 'Delisting Returns and Their Effect on Accounting-Based Market Anomalies', *Journal of Accounting & Economics,* Vol. 43, No. 2, pp. 341-368.

Bernard, V., J. Thomas, and J. Wahlen (1997), 'Accounting-Based Stock Price Anomalies: Separating Market Inefficiencies from Risk', *Contemporary Accounting Research,* Vol. 14, No. 2, pp. 89-136.

Bird, R., and L. Casavecchia (2007), 'Sentiment and Financial Health Indicators for Value and Growth Stocks: The European Experience', *European Journal of Finance,* Vol. 13, No. 8, pp. 769-793.

Chi-Hsiou Hung, D., M. Shackleton, and X. Xu (2004), 'Capm, Higher Co-Moment and Factor Models of Uk Stock Returns', *Journal of Business Finance & Accounting,* Vol. 31, No. 1&2, pp. 87-112.

Edwards, W. (1968), *Conservatism in Human Information Processing* (New York: Wiley).

Fama, E.F. (1998), 'Market Efficiency, Long-Term Returns, and Behavioral Finance', *Journal of Financial Economics,* Vol. 49, No. 3, pp. 283-306.

Fama, E.F., and K.R. French (1992), 'The Cross-Section of Expected Stock Returns', *Journal of Finance,* Vol. 47, No. 2, pp. 427-465.

Fama, E.F., and K.R. French (1995), 'Size and Book-to-Market Factors in Earnings and Returns', *Journal of Finance,* Vol. 50, No. 1, pp. 131-155.

Fama, E.F., and K.R. French (1998), 'Value Versus Growth: The International Evidence', *Journal of Finance,* Vol. 53, No. 6, pp. 1975-1999.

Fama, E.F., and K.R. French (2006), 'Profitability, Investment and Average Returns', *Journal of Financial Economics,* Vol. 82, No. 3, pp. 491-518.

Fama, E.F., and K.R. French (2008), 'Dissecting Anomalies', *Journal of Finance,* Vol. 63, No. 4, pp. 1653-1678.

Fama, E.F., and J. MacBeth (1973), 'Risk, Return and Equilibrium: Empirical Tests', *Journal of Political Economy,* Vol. 81, No. 3, pp. 607-636.

Gregory, A., R.D.F. Harris, and M. Michou (2001), 'An Analysis of Contrarian Investment Strategies in the Uk', *Journal of Business Finance & Accounting,* Vol. 28, No. 9&10, pp. 1193-1228.

Ince, O., and R. Porter (2006), 'Individual Equity Return Data from Thomson Datastream: Handle with Care!', *Journal of Financial Research* Vol. 29, No. 4, pp. 463-479.

Kaiser, K. (1996), 'European Bankruptcy Laws: Implications for Corporations Facing Financial Distress', *Financial Management,* Vol. 25, No. 3, pp. 67-85.

Kothari, S.P., and J.B. Warner (1997), 'Measuring Long-Horizon Security Price Performance', *Journal of Financial Economics,* Vol. 43, No. 3, pp. 301-339.

Lakonishok, J., A. Shleifer, and R.W. Vishny (1994), 'Contrarian Investment, Extrapolation, and Risk', *Journal of Finance,* Vol. 49, No. 5, pp. 1541-1578.

Liu, W., N. Strong, and X. Xu (1999), 'The Profitability of Momentum Investing', *Journal of Business Finance and Accounting,* Vol. 26, No. 9, pp. 1043-1091.

Mohanram, S.P. (2005), 'Separating Winners from Losers among Low Book-to-Market Stocks Using Financial Statement', *Review of Accounting Studies,* Vol. 10, No. 2&3, pp. 133-170.

Newey, W., and K. West (1987), 'A Simple, Positive Semi-Definite, Heteroscedastic and Autocorrelation Consistent Covariance Matrix', *Econometrica,* Vol. 55, No. 3, pp. 703-708.

Piotroski, J.D. (2000), 'Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers', *Journal of Accounting Research,* Vol. 38, No. 3, pp. 1-41.

Rabin, M., and J.L. Schrag (1999), 'First Impressions Matter: A Model of Confirmatory Bias', *Quarterly Journal of Economics,* Vol. 114, No. 1, pp. 37-82.

Shumway, T. (1997), 'The Delisting Bias in Crsp Data', *Journal of Finance,* Vol. 52, No. 1, pp. 327-340.

Shumway, T., and V. Warther (1999), 'The Delisting Bias in Crsp's Nasdaq Data and Its Implications for the Size Effect', *Journal of Finance,* Vol. 54, No. 6, pp. 2361-2389.

Table 1

Descriptive statistics for the whole sample and comparison of mean returns and fundamental characteristics of value and glamour stocks

Variables	Whole sample (13,761 firm-year observations)					Value vs. Glamour			
	Mean	Median	Maximum	Minimum	Std. Dev.	Value	Glamour	Diff.	t-stat
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
RR1	0.0967	0.0652	17.7276	-1	0.5632	0.1277	0.0641	0.0636	5.164***
Size	817	75	164,306	1	4,935	349	1259	-910	-8.885***
BM	0.6939	0.5038	29.8726	0.0004	0.7729	1.3379	0.2167	1.1212	72.458***
ROA	0.0391	0.0598	3.1256	-7.4569	0.2118	0.0161	0.0574	-0.0413	-8.678***
DROA	0.0228	-0.0021	200.6431	-9.2948	1.781	0.0045	0.023	-0.0185	-1.529
CFO	0.0883	0.0924	80.9861	-5.8704	0.7179	0.0536	0.1059	-0.0523	-11.760***
ACCRUAL	-0.0492	-0.0408	2.9193	-80.9592	0.706	-0.0375	-0.0485	0.0109	3.441***
DMARGIN	0.0431	0.0002	476.0476	-30.9205	4.2571	-0.0077	0.1451	-0.1528	-1.406
DLIQUID	0.0019	-0.0013	50.9735	-57.9082	1.9695	0.063	-0.0576	0.1206	2.637***
DLEVER	0.0015	0	0.9944	-3.8576	0.1069	-0.0011	0.0015	-0.0026	-1.121
DTURN	-0.0288	-0.0059	7.1668	-12.9227	0.3406	-0.0315	-0.0128	-0.0187	-2.688***
CAPEX	0.0784	0.0511	28.3369	0	0.2609	0.0627	0.0865	-0.0238	-11.239***
RD	0.0198	0	1.7843	0	0.0719	0.0085	0.0347	-0.0261	-16.532***
F_ROA	0.8201	1	1	0	0.3841	0.7538	0.8434	-0.0896	-10.769***
F_DROA	0.4785	0	1	0	0.4996	0.4303	0.5245	-0.0942	-9.071***
F_CFO	0.8528	1	1	0	0.3543	0.8284	0.8571	-0.0288	-3.786***
F_ACCRUAL	0.7435	1	1	0	0.4367	0.7455	0.7314	0.014	1.531
F_DMARGIN	0.5258	1	1	0	0.4994	0.496	0.5509	-0.0549	-5.272***
F_DLIQUID	0.4968	0	1	0	0.5	0.4975	0.4927	0.0048	0.459
F_DLEVER	0.6234	1	1	0	0.4846	0.6251	0.6267	-0.0016	-0.156
F_DTURN	0.4777	0	1	0	0.4995	0.4659	0.4973	-0.0314	-3.012***
EQOFFER	0.4357	0	1	0	0.4959	0.6028	0.3198	0.2831	28.362***
G_ROA	0.497	0	1	0	0.5	0.2517	0.6935	-0.4418	-47.264***
G_CFO	0.497	0	1	0	0.5	0.3115	0.6532	-0.3418	-34.856***
G_ROAVAR	0.4599	0	1	0	0.4984	0.4569	0.4298	0.0272	2.620***
G_SALESVAR	0.4595	0	1	0	0.4984	0.4554	0.4417	0.0137	1.316
G_CAPEX	0.497	0	1	0	0.5	0.4113	0.5611	-0.1498	-14.513***
G_RD	0.3011	0	1	0	0.4588	0.2246	0.3633	-0.1387	-14.750***
F SCORE	5.4542	5	9	0	1.5729	5.4451	5.4437	0.0015	0.044
G SCORE	3.4549	3	7	0	1.6664	2.8569	3.8741	-1.0172	-30.842***

Notes:

Columns 2 to 6 present summary statistics using the whole sample. The last 4 columns compare mean raw returns and fundamental characteristics of value stocks with those of glamour stocks. Stocks in the highest BM tercile are value stocks and those in the lowest BM tercile are glamour stocks. Definitions of the variables are provided in the Appendix. *, **, *** denotes significance at 10%, 5% and 1% levels, respectively, using a 2-tailed independent two-sample t-test.

Table 2

Distribution of observations across F Score and G Score

Panel A: Distribution of observations across F Score

F Score	n	F Score Group	N
0	4		
1	98	Low F Score	1,513
2	351		
3	1,060		
4	2,156		
5	3,215	Medium F Score	8,630
6	3,259		
7	2,346		
8	1,048	High F Score	3,618
9	224		

Panel B: Distribution of observations across G Score

G Score	n	G Score Group	N
0	305		
1	1,463	Low G Score	4,349
2	2,581		
3	2,777	Medium G Score	5,520
4	2,743		
5	2,133		
6	1,363	High G Score	3,892
7	396		

Notes:

This table reports the number of firm-year observations grouped by F Score and G Score. Stocks with F Score from 0 to 3 (G Score from 0 to 2) are grouped into the low F Score (G Score) portfolio. Stocks with F Score from 4 to 6 (G Score from 3 to 4) are assigned to the medium F Score (G Score) portfolio. Stocks with F Score from 7 to 9 (G Score from 5 to 7) are assigned to the high F Score (G Score) portfolio. Definitions of F Score and G Score are provided in the Appendix.

Table 3

Correlations between BM and the measures of firm's financial strength

	F SCORE	G SCORE	BM
F SCORE	1.0000		
G SCORE	0.2886***	1.0000	
BM	-0.0313***	-0.2000***	1.0000

Notes:
This table reports the Pearson's correlation coefficients between BM and the two measures of firm's financial strengths (F Score and G Score). Definitions of the variables are provided in the Appendix. *** denotes the coefficient is significant at 1% level.

Table 4

Returns on the F Score and G Score strategies across value–glamour contexts

	F Score				G Score			
	Value	Neutral	Glamour	All	Value	Neutral	Glamour	All
Panel A: RR1								
Strong	0.1598	0.137	0.0987	0.1324	0.1596	0.1107	0.0789	0.1051
Weak	0.0386	0.007	-0.021	0.0088	0.1125	0.0578	0.0244	0.0753
Hedge	0.1212	0.1301	0.1198	0.1236	0.0471	0.0529	0.0545	0.0297
t-stat	4.552***	4.058***	2.588***	7.210***	2.035**	2.699***	1.909*	2.390**
Bootstrap p-value	(0.010)	(0.000)	(0.006)	(0.000)	(0.000)	(0.006)	(0.024)	(0.052)
Panel B: VMAR1								
Strong	0.0706	0.0474	-0.0044	0.0385	0.0825	0.0275	-0.0071	0.0217
t-stat	5.353***	3.715***	-0.348	5.153***	4.591***	2.832***	-0.763	3.382***
Bootstrap p-value	(0.006)	(0.148)	(0.068)	(0.604)	(0.000)	(0.768)	(0.000)	(0.056)
Weak	-0.0419	-0.056	-0.0746	-0.0572	0.0283	-0.0239	-0.051	-0.006
t-stat	-1.841*	-1.937*	-1.711*	-3.016***	1.984**	-1.439	-1.937**	-0.58
Bootstrap p-value	(0.088)	(0.008)	(0.000)	(0.000)	(0.052)	(0.002)	(0.000)	(0.000)
Hedge	0.1125	0.1034	0.0701	0.0958	0.0542	0.0514	0.0439	0.0277
t-stat	4.276***	3.273***	1.545	4.695***	2.362**	2.670***	1.57	2.267**
Bootstrap p-value	(0.004)	(0.004)	(0.076)	(0.000)	(0.000)	(0.004)	(0.106)	(0.094)
Strong/Weak contribution (%) +	63/37	46/54	-6/106		152/-52	54/46	-16/116	
Panel C: CFAR1								
Strong	0.0282	0.0696	0.0329	0.0433	0.0784	0.0246	0.0375	0.0398
t-stat	1.429	3.961***	1.870*	4.081***	3.294***	1.804**	2.782***	4.443***
Bootstrap p-value	(0.062)	(0.000)	(0.104)	(0.000)	(0.000)	(0.19)	(0.038)	(0.000)
Weak	-0.0238	-0.0713	-0.1111	-0.0678	-0.0039	-0.0153	-0.0578	-0.0202
t-stat	-0.775	-2.046**	-1.717*	-2.543**	-0.217	-0.723	-1.568	-1.473
Bootstrap p-value	(0.51)	(0.004)	(0.000)	(0.000)	(0.808)	(0.234)	(0.002)	(0.026)
Hedge	0.0519	0.1409	0.144	0.1111	0.0824	0.04	0.0953	0.0601
t-stat	1.425	3.610***	2.147**	3.8714***	2.756***	1.585	2.427**	3.663***
Bootstrap p-value	(0.013)	(0.000)	(0.000)	(0.000)	(0.000)	(0.074)	(0.000)	(0.000)
Strong/Weak contribution (%) +	54/46	49/51	23/77		95/5	62/38	39/61	54/46

Notes:

This table reports the one-year raw returns, market-adjusted returns and control-firm-adjusted returns on the strong, weak and hedge portfolios across value, neutral and glamour samples, and for the whole sample. Stocks with F Score from 0 to 3 (G Score from 0 to 2) are grouped into the weak portfolio and stocks with F Score from 7 to 9 (G Score from 5 to 7) are assigned to the strong portfolio. The hedge portfolios are those take long in strong stocks and short in weak stocks. Stocks in the highest BM tercile are value stocks, those in the lowest BM tercile are glamour stocks, and the rest are neutral. Definitions of the variables are provided in the Appendix. *, **, *** means significant at 10%, 5%, 1% levels, respectively. The t-statistics are from the 2-tailed t-tests under the null hypothesis that the portfolio returns are zero. The bootstrap p-values are estimated from 1,000 iterations.

+ The contribution of the long positions relatively to the contributions of the short positions to the returns of the hedge portfolios.

Table 5

Cross-sectional regressions of abnormal returns on F Score and G Score

	Value				Neutral				Glamour			
	Eq. (3)		Eq. (4)		Eq. (3)		Eq. (4)		Eq. (3)		Eq. (4)	
Indep. variables	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**
Panel A: F Score												
C	-0.0492	-1.31	0.0308	1.496	-0.1293	-2.402**	0.0232	0.649	-0.0367	-0.989	0.0666	2.566**
log(BM)	0.0034	0.178	0.0094	0.396	-0.0105	-0.423	-0.0125	-0.486	0.0113	0.889	0.0182	1.3
Log(Size)	-0.0007	-0.27	-0.0019	-0.666	-0.0033	-1.068	-0.0030	-0.849	-0.0037	-1.719	-0.0040	-1.352
F Score	0.0128	2.265**			0.0295	3.67***			0.0155	2.202**		
High F Score			0.0001	0.005			0.0607	4.958***			0.0204	1.247
Low F Score			-0.0599	-1.743			-0.1080	-1.968*			-0.1092	-3.046***
Panel B: G Score												
C	-0.0414	-1.483	0.0267	1.157	-0.0017	-0.057	0.0568	1.669	-0.0408	-1.523	0.0823	2.672**
log(BM)	0.022	1.078	0.0213	0.917	-0.0011	-0.037	-0.0073	-0.26	0.0027	0.191	0.0159	0.996
Log(Size)	-0.0028	-1.015	-0.0027	-0.948	-0.0059	-1.704	-0.0046	-1.092	-0.0097	-3.358***	-0.0079	-1.629
G Score	0.0232	3.419***			0.0151	3.034***			0.026	5.127***		
High G Score			0.0649	2.02*			-0.0054	-0.38			0.0233	1.336
Low G Score			-0.0240	-0.818			-0.0500	-2.479**			-0.0870	-3.107***

Notes:

The table reports results from the Fama–MacBeth (1973) regressions. The regressions are estimated each year from 1991 to 2007 and the coefficients are the averages across 17 annual regressions. The t-statistics are calculated using Newey-West corrected standard errors. The regressions are run among value, neutral and glamour stocks. Stocks in the highest BM tercile are value stocks, those in the lowest BM tercile are glamour stocks, and the rest are neutral stocks. The table reports results from two regressions:

Equation (3): $CFAR1 = \alpha + \beta_1 \log(BM) + \beta_2 \log(Size) + \beta_3 \text{Score}$

Equation (4): $CFAR1 = \alpha + \beta_1 \log(BM) + \beta_2 \log(Size) + \beta_3 \text{High\_Score} + \beta_4 \text{Low\_Score}$

where: CFAR1 are the one-year control-firm-adjusted returns; BM is the book-to-market ratio; Size is the market capitalization; Score is replaced by either F Score (in panel A) or G Score (in Panel B); High_Score is replaced by either High F Score, which is a dummy variable that takes the value of 1 if the stock is in the high F Score group and zero otherwise (in panel A) or High G Score, which is a dummy variable that takes the value of 1 if the stock is in the high G Score group and zero otherwise (in panel B); Low_Score is replaced by either Low F Score, which is a dummy variable that takes the value of 1 if the stock is in the low F Score group and zero otherwise (in panel A) or Low G Score, which is a dummy variable that takes the value of 1 if the stock is in the low G Score group and zero otherwise (in panel B).

Stocks with F Score from 0 to 3 (G Score from 0 to 2) are in the low F Score (G Score) group and stocks with F Score from 7 to 9 (G Score from 5 to 7) are in the high F Score (G Score) group. All continuous variables are winsorized at the 1 and 99 percentiles. *, **, *** means significant at 10%, 5%, 1% levels, respectively. Definitions of the variables are provided in the Appendix.

Table 6

Raw returns for F Score and G Score strategies across years

Year	Market condition	G Score Strong	G Score Weak	G Score Strong-Weak	G Score t-stat	F Score Strong	F Score Weak	F Score Strong-Weak	F Score t-stat
1991	Good	0.2243	0.0397	0.1846	4.153***	0.1915	-0.0143	0.2058	2.956***
1992	Good	0.2095	0.2035	0.0060	0.114	0.2529	0.0590	0.1938	2.962***
1993	Good	0.0866	0.2419	-0.1553	-4.394***	0.2375	0.1303	0.1072	1.779*
1994	Good	0.1090	-0.0312	0.1402	4.837***	0.0744	-0.0470	0.1214	2.789***
1995	Good	0.2148	0.2878	-0.0729	-1.694*	0.2779	0.2463	0.0315	0.443
1996	Good	0.0042	0.0273	-0.0231	-0.746	0.0406	0.0088	0.0318	0.685
1997	Good	0.2443	0.1062	0.1381	3.143***	0.2316	0.0707	0.1609	2.574**
1998	Good	-0.0006	-0.0156	0.0151	0.440	0.0190	-0.0496	0.0686	1.027
1999	Good	0.1220	0.3297	-0.2076	-2.342***	0.1922	0.4964	-0.3042	-1.582
2000	Bad	0.0545	-0.0256	0.0802	1.979**	0.0269	-0.1512	0.1780	2.882***
2001	Bad	-0.0829	-0.1742	0.0913	2.383**	0.1073	-0.2767	0.3840	7.897***
2002	Bad	0.0137	-0.0879	0.1016	2.252**	-0.0052	-0.1414	0.1363	2.841***
2003	Good	0.3220	0.3509	-0.0289	-0.559	0.3294	0.2874	0.0421	0.716
2004	Good	0.2117	0.1331	0.0786	1.838*	0.1788	0.0171	0.1617	2.414**
2005	Good	0.1660	0.0681	0.0979	2.024**	0.1833	0.0233	0.1601	2.451**
2006	Good	0.1949	0.1776	0.0172	0.278	0.1994	0.0917	0.1077	1.615
2007	Bad	-0.2320	-0.3595	0.1275	3.564***	-0.2583	-0.3649	0.1066	2.351**

Notes:

This table reports the one-year raw returns on the strong and weak portfolios. Stocks with F Score from 0 to 3 (G Score from 0 to 2) are grouped into the weak portfolio and stocks with F Score from 7 to 9 (G Score from 5 to 7) are assigned to the strong portfolio. A "good" ("bad") market is identified if the contemporary return on the FTSE All Shares Index is positive (negative). Definitions of the variables are provided in the Appendix. *, **, *** means significant at 10%, 5%, 1% levels, respectively. The t-statistics are from the 2-tailed independent two-sample t-test.