Essays on value, momentum and the preference for skewness

Yiqing Dai

Business School, University of Adelaide

January 2017

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Thesis Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree. I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Yiqing Dai, Candidate Date

Context Statement

This dissertation attempts to explain the cross-section of expected returns based on the valuation equation of Miller and Modigliani (1961) and the cumulative prospect theory of Tversky and Kahneman (1992). It consists of three essays on empirical asset pricing.

The first essay proposes an alternative metric for value investing: the dividend-to-market ratio (DM). The book-to-market ratio (BM) which is currently used in academia and industry, is a noisy measure for value investing, because book value is a weak indicator of intrinsic value. Motivated by the valuation equation of Miller and Modigliani (1961), this paper suggests that DM is much more efficient in identifying undervalued stocks than BM, due to the strong link between expected dividends and intrinsic value. DM also provides a better estimation of expected stock returns compared to the linear combination of BM, profitability and investment used in the five-factor model of Fama and French (2015a), because it allows for non-linearities between expected returns and these variables. Results of cross-sectional regressions at the firm level and time-series regressions at the portfolio level consistently show that DM has a far stronger link with expected returns than BM, and it also outperforms a linear combination of BM, profitability and investment.

The second essay examines the prediction of cumulative prospect theory whereby investors prefer lottery-like (or positively skewed) payoffs, resulting in overpricing and low expected returns to such assets. Given that earnings surprises are associated with lottery-like payoffs, investors should be willing to pay more for stocks with a high probability of generating positive earnings surprises. Empirical tests in this study consistently suggest that there is a strong negative correlation between the predicted profitability shocks (PPS) and expected stock returns. This essay contributes to the literature in asset pricing by revealing the link between skewness preference and prominent anomalies such as BM, profitability and price momentum.

The explanatory power of BM and operating profitability disappears after controlling for PPS, which indicates that both could be noisy proxies for PPS in predicting average returns. Further, the price momentum effect cannot be driven out by earnings momentum once PPS is taken into account, which demonstrates that price momentum has incremental explanatory power for stock returns over that provided by earnings momentum.

The third essay (co-authored with Takeshi Yamada and Tariq Haque) examines if crash-risk is systematically priced in momentum portfolio returns. A recent paper by Daniel and Moskowitz (2016) documents that crashes occur in the momentum strategy, and investors may take many years to recover from the resulting losses. Thus, this essay asks, if crash risk exists in momentum portfolios, is such risk priced in the market? To this end, we develop a measure, tail coskewness, that focuses exclusively on how tail events (low-probability events leading to large gains or losses) contribute to the systematic skewness of momentum portfolios. The results show that tail coskewness not only subsumes the risk premium associated with coskewness which may be a determinant of cross-sectional returns as shown by Harvey and Siddique (2000), but also that associated with firm size. The paper uses US data from 1927 as well as international data, where robust results are found across different time-periods and markets.

Acknowledgement

I would like to express my deep gratitude to my thesis supervisors, Professor Takeshi Yamada and Dr. Tariq Haque for their invaluable advice and guidance. I am incredibly grateful to Professor Takeshi whose insightful comments and encouragement have much improved the quality of this dissertation. In addition to his wisdom, great knowledge and rigorous attitude, his kind mind and selflessness also make him a great lifelong mentor. I am also hugely indebted to Dr. Tariq Haque for his generous help, enormous patience, and immense knowledge that have helped me in too many ways to enumerate throughout my PhD journey. I will always remember his great support and encouragement as a mentor and as a friend that have guided me through difficult times of research and writing of this thesis.

I am also grateful to the financial support provided by the Adelaide Graduate Centre and the Faculty of Accounting and Finance. Were it not for their generosity, my PhD journey would not have been completed.

I would like to thank the members of Faculty of Accounting and Finance. The faculty, staff, and students really made my stay in the University of Adelaide one I'll always remember.

My wife, Yanping, deserves special thanks for giving enormous support to my research. Her sacrifice for the family has made this dissertation possible.

Table of Contents

List of Figures

List of Tables

Chapter 1
Value investing with dividend-to-market ratio

1.1. Abstract

The book-to-market ratio (*BM*) is a noisy metric for value investing because book value is a weak indicator of intrinsic value. Using the dividend discount model of Miller and Modigliani (1961), this paper proposes an alternative metric for value investing: the dividend-to-market ratio (*DM*), where dividend is measured as profitability minus investment. Test results show that *DM* effectively distinguishes between undervalued stocks and overvalued ones, and substantially outperforms *BM*. Further, *DM* is a parsimonious, more efficient measure to estimate expected returns than a linear model consisting of *BM*, profitability and investment. An investor can increase a portfolio's Sharpe ratio by adding just the *DM* factor than by adding all the *BM*, profitability and investment factors.

1.2. Introduction

A great deal of academic research has been published on value investing, which suggests to buy undervalued stocks and avoid overvalued ones with respect to their intrinsic value. Academic evidence on value investing is overwhelmingly dominated by the book-to-market ratio (*BM*), pioneered by the findings of Rosenberg, Reid, and Lanstein (1985) and Fama and French (1992). However, because book value is a weak indicator of intrinsic value, BM is a noisy metric for value investing[1]. Specifically, BM does not allow for a differentiation between a low-priced stock with a high intrinsic value, and one that its low price is consistent with its low intrinsic value (low expectation of future cash flows). Since the second scenario is more likely in a highly competitive market, BM is rather inefficient in identifying the best value opportunities. A high BM portfolio is heavily populated by stocks that are not undervalued by the market and therefore is sub-optimal for value investing.

Using the dividend discount model of Miller and Modigliani (1961), this paper proposes an alternative metric for value investing: the dividend-to-market ratio (*DM*),

$$DM = Dividend/Market\ Value = (Profitability - Investment)/Market\ Value$$

where dividend is defined as the maximum payable dividend (profitability minus investment), following the notation of Fama and French (2006). Miller and Modigliani (1961) claim that a firm's value is justified by its expected dividends — the difference between the earning power of the firm's assets and the reinvestment of earnings required to generate future cash flows. Given estimates of expected dividends and current market value, we can solve the market discount rate on expected dividends (i.e., long-term average expected returns) (see Fama and

[1] In fact, Graham and Dodd (1934) strongly criticize the view of equating intrinsic value to book value because neither the average earnings nor the average market price evinced any tendency to be governed by book value.

French 2006 and 2015a). Hence, it is a straightforward choice to use the ratio of expected dividends to market value to estimate the cross-section of market discount rate.

Because expected dividends are a strong indicator of intrinsic value, the dividend-to-market ratio effectively distinguishes between undervalued stocks and overvalued ones. A high (low) *DM* indicates the firm's expected future cash flows are currently discounted at a high (low) rate, and hence its stocks are in the value (growth) category. If two firms are identical in market valuation but different in dividends, the firm with larger dividends must have a higher market discount rate. Likewise, if two firms are identical in dividends but different in market valuation, the firm with higher market valuation should have a lower market discount rate. Value investors could thus maximize their economic gains per dollar of investment by constructing a high *DM* portfolio, holding stocks with strong fundamentals at moderate prices, as well as stocks with average fundamentals at discount prices.

Using portfolios formed by double sorts (3×3) on *DM* and *BM*, I show that 30% of high *BM* stocks have low *DM* value, indicating that they are the low-priced ones with low intrinsic value. These high *BM*, low *DM* stocks substantially underperform the market, which directly evidences that *BM* is a noisy metric for value investing. Consistent with the prediction of the dividend discount model, value investing with *DM* leads to substantial economic gains in the sample period July 1963 to December 2013. For zero-cost mimicking factors formed by double sorts (2×3) on size and *DM*, a $1 factor exposure delivers a cumulative profit of $28.84 for the *DM* value factor, while the cumulative profit for the *BM* value factor is only $4.35. The Sharpe ratio improves from 0.39 for the *BM* value factor to 0.81 for the *DM* value factor. Thus, *DM* is superior to *BM* for value investing from both theoretical justification and empirical regularity.

This paper adds to a growing literature using the dividend discount model of Miller and Modigliani (1961) to enhance estimates of expected stock returns (see Fama and French 2006, 2015a and 2015c, Aharoni et al. 2013 and Novy-Marx 2013). The focus of those papers is to decompose the valuation equation of the dividend discount model into three component variables (BM, profitability and investment), and then combine them linearly to estimate expected returns. My work differs from those papers in one important way in that I take an integrated approach by using *DM* alone to estimate expected return. The dividend discount model indicates that *DM* provides a closed end solution to expected returns as an interaction term between expected dividends and market value($\rho = Dividend/Market\ Value$). That is, the marginal effect of expected dividends depends on the level of market value, while the marginal effect of market value depends on the level of expected dividends. Therefore, *DM* provides a better estimate of expected stock returns than a linear combination of *BM*, profitability and investment because it automatically accounts for the economic non-linearity between expected returns and these variables.

Throughout this paper, I conduct tests to assess whether *DM* can outperform a linear model consisting of *BM*, profitability and investment measures in predicting average stock returns. In Fama-Macbeth (FM) cross-sectional regressions of stock returns on firm characteristics, *DM* simultaneously subsumes the explanatory powers of *BM*, profitability and investment statistically and economically. In the extreme deciles predicted to have high (low) returns by the FM regression jointly controlling for *BM*, profitability and investment, 27% (47%) stocks are not associated with extreme *DM* value, showing no extreme returns. In time series regressions using 2×3 sorts to form mimicking factors, the *DM* value factor generates significant alpha relative to the five-factor model (FF5) of Fama and French (2015a) that includes the market, size, *BM*, profitability, and investment factors. In contrast, the *BM*, profitability and investment factors are fully explained by a parsimonious model that includes only the market, size and *DM* factors. Furthermore, in GRS tests to explain a set of prominent anomalies that are not related to the dividend discount model, the *DM* factor does a better job in explaining stock returns than a linear combination of the *BM*, profitability, and investment factors.

Since profitability is the source of dividends, I also conduct a horse race between *DM* and several alternative profitability measures in predicting average stock returns. These alternative profitability measures are the earnings-to-price ratio of Ball (1978), the cash flow-to-price ratio of Lakonishok et al. (1994), the gross profitability of Novy-Marx (2013), the operating profitability of Ball et al. (2015) and the return-to-equity ratio of Hou et al. (2014). In time series spanning tests based on mimicking factors, the *DM* factor dramatically outperforms the set of alternative profitability factors. In particular, none of the alternative profitability factors exhibits statistically reliable alpha after controlling for the *DM* factor. In contrast, the *DM* factor consistently produces large, highly significant alpha after controlling for alternative profitability factors. These results show that the *DM* factor is much closer to the efficient frontier than alternative profitability factors.

The paper proceeds as follows. Section 1.3 provides a simple theoretical framework for the dividend-to-market ratio. Section 1.4 presents the *DM* measure and data used in this study. Section 1.5 compares *DM* with a linear combination of *BM*, profitability and investment using firm-level FM regressions. Section 1.6 presents the performance of mimicking portfolios. Section 1.7 conducts a horse race among the competing models to explain several prominent return anomalies. Section 1.8 implements robustness tests with several alternative measures of profitability. Section 1.9 concludes.

1.3. Dividend discount model

Based on the dividend discount model of Miller and Modigliani (1961), Fama and French (2006, 2008, 2015a and 2015c) and Aharoni, Grundy, and Zeng (2013) show that the market value of a firm is the present value of its expected dividends:

$$M_t = \sum_{\tau=0}^{\infty} \frac{E(D_{t+\tau})}{(1+\rho)^{\tau+1}} \tag{1}$$

where M_t is the market value of the firm at the start of period t, $E(D_t)$ is the expected dividends (expected earnings minus expected additional investment required to generate future earnings) for period t, and ρ is the market discount rate on expected dividends or the long-term average expected return (these two terms are used interchangeably).

From the perspective of market discount rate, its relation with expected dividends and current market value can be brought out with a bit of manipulation on equation (1):

$$1 = \sum_{\tau=0}^{\infty} \frac{E(D_{t+\tau})/M_t}{(1+\rho)^{\tau+1}} \tag{2}$$

where equation (2) reveals that the expected dividend to market ratio, $E(D_{t+\tau})/M_t$, provides a closed form solution for the discount factor $(1/(1+\rho)^{\tau+1})$. If dividends are considered in perpetuity at D_0, by setting $t = 0$, we can algebraically simplify equation (2) into a much more compact equation:

$$1 = \sum_{\tau=0}^{\infty} \frac{D_0/M_0}{(1+\rho)^{\tau+1}} = \frac{D_0/M_0}{\rho}, \text{ or equivalently } \rho = D/M \tag{3}$$

where subscripts are dropped without leading to ambiguity in the present context. In this case, the dividend to market provides a closed-end solution to the market discount rate ρ. That is, given estimates of future dividends and market value, the market discount rate on dividends is uncovered to investors.

In equation (2), the expected dividend to market ratio, $E(D_{t+\tau})/M_t$, can be decomposed into three component variables: *BM*, profitability and investment, when the expected dividends are expressed as expected earnings minus expected reinvestment of earnings:

$$1 = \sum_{\tau=0}^{\infty} \frac{[E(Y_{t+\tau}) - E(\triangle B_{t+\tau})]/M_t}{(1+\rho)^{\tau+1}} = \sum_{\tau=0}^{\infty} \frac{\left[\frac{E(Y_{t+\tau})}{B_t} - \frac{E(\triangle B_{t+\tau})}{B_t}\right] * \frac{B_t}{M_t}}{(1+\rho)^{\tau+1}} \tag{4}$$

where $E(Y_t)$ is the expected earnings, B_t is book equity, and $E(\triangle B_t)$ is the expected change in book equity. Each of *BM*, expected earnings-to-book equity ratio and expected growth in book equity alone acts as an incomplete measure of expected returns, because expected returns also vary with the other two variables[2]. To improve estimates of expected returns, Fama and French (2006, 2015a) linearly combine *BM*, profitability and investment to explain the cross section of average stock returns. Equations (4), however, indicates that expected returns is linearly related to the term $[E(Y_{t+\tau}) - E(\triangle B_{t+\tau})]/M_t$, but nonlinearly related to the terms $E(Y_{t+\tau})/B_t$, $E(\triangle B_{t+\tau})/B_t$ and B_t/M_t . Therefore, a linear combination of *BM*, *OP* and *INV* in asset pricing models might be misspecified. In contrast, the term $E(D_{t+\tau})/M_t$ provides a better description of expected returns by automatically accounting for the economic non-linear relation.

1.4. Measure of dividend-to-market ratio and data

One challenging task for measuring *DM* is to identify a reliable proxy for expected dividends. Graham and Dodd (1934) point out that the past financial record affords at least a rough guide to the future. Earlier studies find that simple proxies for expected profitability and investment provided by the most recent record are powerful forecasting variables for average returns. For equity earnings, Novy-Marx (2013) finds that gross profitability (revenue minus cost of goods sold, $REVT - COGS$) has great power in predicting the cross section of average returns, and interprets this as a clean accounting measure of true economic profitability. Ball et al. (2015) show that operating profitability (revenue less cost of goods sold less selling, general &

[2] Three conditional hypotheses are generated immediately by equation (5): Holding two other component variables fixed, a higher *BM* implies a higher expected return; a higher expected earning-to-book equity ratio implies a higher expected return; and a higher expected growth in book equity implies a lower expected return. These hypotheses are supported by enormous evidence in literature on the explanatory powers of *BM*, profitability and investment in expected returns. For example, Rosenberg et al. (1985) and Fama and French (1992) show that higher *BM* predicts higher average returns; Novy-Marx (2013) and Ball et al. (2015) document that profitability is positively correlated with average returns; Titman, Wei, and Xie (2004) and Aharoni, Grundy, and Zeng (2013) show that investment is negatively correlated with average returns.

administrative expenses excluding expenditures on research & development, $REVT - COGS - XSGA + XRD$) can further improve the predictive power of profitability.[3] For additional investment, Aharoni, Grundy, and Zeng (2013) and Fama and French (2015a) find that the growth of total assets ($\triangle AT/AT$) is negatively related to average returns.[4] Together with equations (2) to (4), I measure the dividend-to-market ratio for each firm at the end of each June as

$$DM = [\underbrace{(REVT - COGS - XSGA + XRD - XINT)}_{Profitability} - \underbrace{B * \triangle AT/AT}_{Investment}]/\underbrace{M}_{Market\ Cap} \quad (5)$$

Where revenue (*REVT*), cost of goods sold (*COGS*), selling, general & administrative expenses (*XSGA*), research & development (*XRD*), interest expense (*XINT*) and book equity (*B*) are measured with accounting data for the fiscal year ending in year t-1; $\triangle AT/AT$ is the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-1 total assets; *M* is the market capitalization at the end of December of year t-1, adjusted for changes in shares outstanding between the measurement date for *B* and the end of December.

I use monthly stock returns data from the Centre for Research in Security Prices (CRSP) and annual accounting data from Compustat. The asset pricing tests cover July 1963 through December 2013. I exclude financial firms and very small firms with total assets of less than \$25 million or book equity of less than \$12.5 million. Table 1.1 reports summary statistics for three sets of portfolios formed by double sorts (3×3) on *DM* and one second sort variable [*BM,* operating profitability (*OP*)*,* and investment (*INV*)]. At the end of each June, stocks are

3 Ball et al. (2015) argue that, since the allocation of COGS and XSGA is not determined by Generally Accepted Accounting Principles, operating profitability that deducts both expenses is an even cleaner measure of profitability.

4 Fama and French (2015a) find that the lagged growth of assets has greater power in predicting the cross section of average returns than the lagged growth in book equity, and argue that the lagged growth of assets is a better proxy for expected future growth in book equity than the lagged growth in book equity. We argue that one possible explanation is that the lagged growth in book equity also contains information on the change in financial leverage, which disturbs the relation between average returns and book equity growth.

independently assigned to three *DM* groups, and three *BM, OP* and *INV* groups using the NYSE 30th and 70th percentiles as breakpoints. The intersections of the two classifications produce three sets of portfolios: *DM-BM*, *DM-OP* and *DM-INV*. Panel A reports portfolio market and size adjusted returns, which are the intercepts form time-series regression of portfolio value-weighted returns on market value-weighted index and size factor (large minus small). Panel B, C and D reports number of firms, times series average *DM* and second sort variable. Measures of *BM*, operating profitability (*OP)*, investment (*INV)* are constructed in the same way as Fama and French (2015a) in order to facilitate a direct comparison.

For *DM-BM* portfolios, portfolio with high *BM* and low *DM* substantially underperform most of other portfolios by producing an average adjusted return of -0.10%, despite it has the second highest averaged *BM* (0.42) in panel C. These high *BM* stocks are not undervalued by the market, as the low *DM* value indicates that their low prices are associated with low cashflows. Note that stocks with high *BM* and low *DM* have an average number of firms of 315 in Panel D, accounting for 30% of high *BM* stocks. On the other hand, stocks with low *BM* and high *DM* produce an impressive average adjusted return of 0.19%, the third highest among *BM-DM* portfolios, despite that they have a very negative averaged *BM* of -1.24 in Panel C. These low *BM* and high *DM* stocks are not overvalued by the market, accounting for 7.55% of low *BM* stocks.

Most importantly, holding *DM* fixed, stocks with high *BM* do not significantly outperform stocks with low *BM*. The column of H-L shows that the spread of size and market adjusted returns between low *BM* and high *BM* stocks is only 0.10% (t = 063), 0.07% (t = 0.62) and 0.20% (t = 1.25) for the group of low, medium and high *DM* stocks, respectively. In contrast, the row of H-L shows that the high *DM* portfolio consistently outperforms the low *DM* stock by 0.39% (t = 2.46), 0.44% (t = 4.64) and 0.49% (t = 4.51) for the group of low, medium and high *BM* stocks, respectively.

Similarly, for *DM-OP* portfolios and *DM-INV* portfolios, controlling for *DM* invalids the predictive power of profitability and investment, except for the group of low *DM* stocks among *DM-OP* portfolios. In contrast, the predictive power of *DM* persists after controlling for profitability and investment. In a nutshell, these three sets of portfolios tell a consistency story that *DM* subsumes the predictive power of *BM*, profitability and investment in returns.

1.5. FM regression and economic significance comparison

This section studies the cross-sectional relation between individual stock returns and explanatory variables using FM regressions, and then compares the economic importance of competing variables.

1.5.1. Firm-level cross-sectional regression

Table 1.2 presents the time-series average slopes from FM regressions of stock monthly excess returns on *BM*, *OP*, *INV* and *DM,* and the corresponding Newey-West (1987) adjusted *t*-statistics. In this section, I do not control other variables not implied by valuation theory, except for size (*ME)*. Independent variables are trimmed at the 1% and 99% levels on a table-by-table basis to ensure different regressions within each table panel are based on the same observations. Following Fama and French (2008a), except for all stocks, I also run separate FM regressions for All but Micro stocks (ABM) and Microcap stocks (below the 20th percentile of NYSE market cap) to isolate the influence of microcap stocks in testing results.

I first examine the explanatory powers of *BM*, *OP* and *INV* separately. Consistent with prior studies, regressions 1-4 show that all three measures help to explain the cross-section of stock returns.[5] The average slopes for these measures are all more than 2.2 standard errors from zero

[5] Regressions 1 and 4 correspond to the FF3, FF5 models with market beta absent in the cross-sectional regressions. We do not include the market beta in our cross-sectional regressions, following Fama and French (2008) who argue that there is little reason to expect the market beta to be correlated with anomaly variables.

for All stocks. The positive slopes for *BM* and *OP* say that value stocks tend to have higher average returns than growth stocks, and stocks of profitable firms tend to outperform stocks of unprofitable firms. The negative slope for *INV* says that higher investment is related to lower expected returns. These results are not unduly influenced by microcaps, as the slopes for ABM stocks are roughly the same as for All stocks. For microcaps, however, it is surprising to find that the slopes for *OP* are not significant (t = 0.56 in regression 2, and t = 1.26 in regression 4).

Our primary interest is regression 5, which uses *DM* to explain expected returns, with the assistance of *ME*. Regression 5 shows that *DM* has a strong role in explaining the cross-section of average returns. For All stocks, the *DM* slope (2.409) has an impressive t-value of 6.99, much larger than that of *BM*, *OP* and *INV* in previous regressions. Note also that the slopes on *DM* are close in absolute value with large t-statistics for ABM (2.152, t = 5.24) and Microcaps (2.527, t = 7.12), indicating that the effect of *DM* is pervasive across the full spectrum of stocks. Regression 6 shows that, except for microcaps, the positive relation between *DM* and expected return persists after controlling for *BM*, *OP* and *INV*. The average slope for *DM* is 1.706 (t = 4.47) for All stocks and 1.285 (t = 2.98) for ABM. In contrast, controlling for *DM* seems to absorb the roles of *BM*, *OP* and *INV* in average returns, as their t-values drop dramatically to non-significant levels in all sample sets. These results are in line with the prediction of the dividend discount model that *DM* has a far strong link with expected return than *BM*, profitability and investment.

I caution that including all component variables in an interaction model increases multicollinearity, such that regressions may not give accurate results about any individual parameter or about which parameters are redundant with respect to others. For robustness, I use return residuals from regression 4 as a dependent variable to test the incremental explanatory power of *DM* relative to *BM*, *OP* and *INV*. Those residuals are by definition orthogonal to *BM*, *OP* and *INV*. Regression 7 in Panel B shows that *DM* still captures significant variation in residual returns. For All stocks, the positive *DM* slope remains 2.37 standard errors from zero. The slopes for *DM* in ABM and microcaps are less impressive (0.68, t = 1.71 for A*BM* stocks and 0.677, t = 1.70 for microcaps), which might be due to the small sample size. Next, regression 8 reverses the process by regressing return residuals from regression 5 on *BM*, *OP* and *INV*. The fact that all three variables lose significance in All Stocks, ABM stocks and Microcaps confirms that *BM*, *OP* and *INV* have no unique effects in average returns relative to *DM*. Thus, these results reconfirm the capability of *DM* to fully subsume the explanatory powers of *BM*, *OP* and *INV* in average returns.

1.5.2. Economic significance

Fama and French (2015b) show that a variable's importance can be judged by its incremental contribution to the average return spread for portfolios sorted by fitted values from a multivariate cross-sectional regression. Thus, I compute the average return spreads between portfolios of stocks forecasted to have high versus low return based on FM regression results. In particular, at the end of June each year, stocks are formed into portfolios according to their predicted returns, which are the fitted values of the FM regression estimated over the 50-year sample period. These fitted values are the average regression slopes multiplied by the value of explanatory variables at the end of each June. Equal-weighted portfolio monthly returns are then calculated from July through June of the following year. In Table 1.3, results are presented for the average return spreads between low and high tertiles, low and high quintiles, and low and high deciles in sequence.

Table 1.3 focuses on the return spreads forecasted by regression 1 (controlling for *ME* and *BM*), regression 4 (controlling for *ME*, *BM, OP* and *INV*) and regression 5 (controlling for *ME* and *DM*) from Table 1.2. The differences in return spread between regression 1 and regression

4 show that adding *OP* and *INV* into the three-factor model of Fama and French (FF3) delivers substantial economic gains. Of more significance is that regression 5 generates generally larger return spreads than regression 4. For All stocks, the return spreads predicted by regression 5 and regression 4 are 0.99% versus 0.83% for tertiles, 1.17% versus 1.04% for quintiles, and 1.42% versus 1.30% for deciles. For ABM, the return spreads predicted by regression 5 and regression 4 are very close in magnitude, 0.75% versus 0.69% for tertiles, 0.78% versus 0.73% for quintiles, and 1.06% versus 1.10% for deciles.

A more impressive gain for the return spreads predicted by regression 5 is a substantial reduction in standard deviation for all partitions used. For example, in regression 5 using All Stocks, the standard deviation of return spreads provided by regression 4 falls from 3.60% to 2.74% for tertile portfolios, from 4.33% to 3.11% for quintile portfolios, and from 5.28% to 3.41% for decile portfolios. As a result, regression 5 dramatically increases the *t*-values and Sharpe ratios for return spreads relative to regression 4. The Sharpe ratios for tertile, quintile and decile return spreads predicted by regression 5 are 1.25, 1.30 and 1.44 respectively, while the corresponding values for regression 4 are only 0.80, 0.83 and 0.85, a difference of more than 50% in Sharpe ratio. For both ABM and Micro stocks, the difference in Sharpe ratio for return spreads remains large (about 40%) between regression 4 and regression 5. Figure 1.1 shows the probability density function for the time-series return spreads predicted by regression 5 (solid line) and regression 4 (dot line) using tertile portfolios for All stocks. Controlling for *DM* is associated with higher peaks and shorter tails for return spreads compared to that controlling for *BM, OP* and *INV*. Overall, these results suggest that *DM* is more efficient in estimating expected returns compared to a linear combination of *BM*, *OP* and *INV*.

1.5.3. Return spread on subset portfolios

Table 1.4 reports the return spreads produced by different subsets of stocks in the extreme expected return deciles predicted by regression 4 (controlling for *ME*, *BM, OP* and *INV*) and regression 5 (controlling for *ME* and *DM*) from Table 1.2. Subset 1 includes stocks that are listed in the extreme decile predicted by regression 4 but not in the extreme decile predicted by regression 5. Subset 2 includes stocks that are commonly listed in the extreme deciles predicted by both regression 4 and regression 5. Subset 3 includes stocks that are listed in the extreme decile predicted by regression 5 but not in the extreme decile predicted by regression 4.

Subset 1 shows that, in the low (high) extreme return decile predicted by *BM, OP* and *INV,* there are 47% (27%) of stocks that are not associated with extreme *DM* value. Consequently, among all three subsets, subset 1 produces the most unattractive average return spreads between stocks listed into the low and high deciles. The average monthly return spreads for subsets 1, 2 and 3 are 0.98% (t = 3.9), 1.51% (t = 7.27), and 1.29% (t = 7.46), respectively. The Sharpe ratios for subset 1 are 0.55, which are about half of that for subset 2 (1.02). The underperformance of subset 1 shows that extreme value for *BM, OP* and *INV* does not imply extreme expected returns once it is not related to extreme value for *DM*. In contrast, the Sharpe ratios on subset 3 (1.05) are nearly identical to subset 2. This says that, in the extreme deciles predicted by *DM*, portfolio performances are largely the same for stocks with or without showing extreme value for *BM*, *OP* and *INV*. In other words, stocks with extreme expected returns need not have extreme value for *BM, OP* and *INV.* These results reiterate that it is *DM* predicting the cross-section of expected returns, and a linear combination of *BM*, *OP* and *INV* misses out the critical interaction effect between expected dividends and market value, ending up as a noisy approach to estimate expected returns.

1.6. Comparison of mimicking portfolios

Since investors are concerned with whether this opportunity set produced by the *DM* is actually exploitable, this section compares the performance of mimicking portfolios for *DM, BM*, *OP* and *INV*. To evaluate the pervasiveness of average return patterns, I use three sets of mimicking portfolios formed by the 2×3 sorts, 2×5 sorts and 2×10 sorts, where the latter two focus on more extreme characteristic values. Following Fama and French (1993), the 2×3 sorts for the value factor are constructed at the end of June of each year. I use the median NYSE size to split NYSE, Amex, and NASDAQ stocks into small and big stocks. Independently, I break stocks into three *BM* groups using the NYSE breakpoints for the lowest 30%, middle 40%, and highest 30% of *BM* values for the fiscal year ending in calendar year $t - 1$. *HML* (high minus low *BM*) is the average return on the two value (high *BM*) portfolios minus the average return on the two growth (low *BM*) portfolios. The same construction process also applies to *OP, INV* and *DM,* where *RMW* (robust minus weak profitability) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios; *CMA* (conservative minus aggressive investment) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios; DM^r (high minus low *DM*) is the average return on the two high *DM* portfolios minus the average return on the two low *DM* portfolios. Factor mimicking portfolios in the 2×5 sorts and the 2×10 sorts are formed in the same way as in the 2×3 sorts and report the average spread between the top and bottom portfolios, except that stocks are assigned independently to quintile and decile portfolios for the second sort.

Figure 1.2 shows that the cumulative returns to $\mathrm{DM^r}$, *HML, RMW* and *CMA* constructed by 2×3 sorts. For a \$1 factor exposure over the sample period July 1963 to December 2013, the cumulative profit is \$28.84 for DM^r, while it is only \$4.35 for *HML*, \$3.30 for *RMW*, and \$4.72 for *CMA*. Table 1.5 reports summary statistics for factor portfolios, including the average of time series returns, standard deviations, *t*-statistics and Sharpe ratios. In our observation period,

DM^r has the highest average return and most significant *t*-statistics among competing factors. For instance, in the 2×3 sorts of Panel A, DM^r has a monthly average return of 0.59% with a *t*-statistic of 5.77. By contrast, the average returns for *HML, RMW* and *CMA* are only 0.32%, 0.28% and 0.31% per month (*t* = 2.74, 2.55 and 2.28), respectively. This outperformance is not accompanied by a large volatility, as we see that the standard deviation of DM^r is moderate, 2.53%, compared with 2.86% for *HML*, 2.67% for *RMW*, and 2.28% for *CMA*. Consequently, DM^r has a much higher Sharpe ratio of 0.81, compared to *HML, RMW* and *CMA* (SR=0.39, 0.36 and 0.48). Moreover, the outperformance of DM^r holds for extreme portfolios, as it also has much higher Sharpe ratios than *HML, RMW* and *CMA* in the 2×5 sorts (0.78 versus 0.38, 0.37 and 0.52) and in the 2×10 sorts (0.95 versus 0.52, 0.26 and 0.60).

I also present subperiod results where I split the overall sample before and after July 1988. Panels B and C show that regardless of the observation period, DM^r consistently outperforms *HML*, *RMW* and *CMA*. For instance, in the 2×3 sorts, the Sharpe ratios on DM^r for the two subperiods (SR=0.94 and 0.73) are close to that for the overall period (SR=0.81). The Sharpe ratios on *HML, RMW* and *CMA* for the two subperiods (SR=0.46 and 0.32 for *HML*, 0.36 and 0.38 for *RMW,* 0.53 and 0.44 for *CMA*) are also close to that for the overall period (SR=0.39, 0.36 and 0.48). The 2×5 factors and the 2×10 factors offer similar results with the 2×3 factors for the subperiods. The results thus confirm that DM^r is closest to the efficient frontier among the mimicking portfolios considered here.

For portfolio management, it is critical to know whether these patterns in average returns show up reliably for large stocks that account for more than 90% of total market capitalization, or rely mostly on small stocks that are much less liquid. Table 1.6 shows separate results for small and big stocks. We find that the effect of *DM* among big stocks, DM_B^r, is impressively strong, although it is weaker than that among small stocks, DM_S^r. In the 2×3, 2×5 and 2×10 sorts, the average returns of DM_B^r are 0.44%, 0.50% and 0.74% per month (*t* = 3.54, 3.40 and 4.55), while the average returns of DM_S^r are 0.75%, 0.82% and 1.04% per month (t = 7.17, 6.81 and 7.46). In contrast, HML, RMW and CMA for big stocks are lack consistent statistical power, although they are highly significant for small stocks. In the 2×3, 2×5 and 2×10 sorts, the t-values are 1.19, 1.05 and 2.50 for HML_B, 1.93, 2.21 and 1.15 for RMW_B, and 1.14, 1.43 and 1.94 for CMA_B, although the t-values are 3.81, 3.83 and 3.95 for HML_S, 2.64, 2.48 and 2.16 for RMW_S, and 5.41, 5.53 and 5.84 for CMA_S. The results confirm the evidence in table 1.4 of Fama and French (2015a) that the value, profitability and investment premiums do not show consistent significance for big stocks, but do for small stocks. In short, for big stocks, DM^r is much more reliable and exploitable for investors than *HML, RMW* and *CMA.*

Figure 1.3 shows the trailing ten-year Sharpe ratios of the 2×3 sorted DM^r, *HML, RMW* and *CMA* factors. The Sharpe ratios for *HML, RMW* and *CMA* fluctuate dramatically over time. For instance, *RMW* fares poorly from the late 1970s to the early 1980s, but recovers sharply from the late 1980s to the early 1990s, which is followed by a big loss in the late 1990s. In contrast, the Sharpe ratio for DM^r remains relatively stable around its mean of 0.94, with the only two significant falls occurring during the nifty fifty boom in the early 1970s and the dot-com boom in the late 1990s. For most of the sample, the Sharpe ratios for DM^r dominate the Sharpe ratios for *HML, RMW* and *CMA*.

In addition, Figure 1.3 also shows that the Sharpe ratio for DM^r is generally strongly positively correlated with those for *HML, RMW* and *CMA*, except for its relation with *RMW* in 1970s. The common fall in Sharpe ratio for *HML*, *RMW* and *CMA* in the 1990s is accompanied by a drop for DM^r, and the common rise in Sharpe ratio for *HML*, *RMW* and *CMA* in the 2000s is followed by an increase for DM^r. These high correlations provide evidence that the effects of *DM*, *BM*, profitability and investment are commonly derived from the dividend discount model.

Table 1.7 presents further details on correlations between the mimicking factors. DM^r has strong positive correlations with *HML* and *CMA* (ρ = 0.85 and 0.78) for the overall period. The correlations of DM^r with *HML* and *CMA* show up consistently for both the first subperiod (ρ = 0.81 and 0.75) and the second subperiods (ρ = 0.89 and 0.80). Unlike its relations with *HML* and *CMA*, DM^r is only moderately correlated with *RMW* (ρ = 0.40) for the full sample period. The correlation between DM^r and *RMW* is moderate negative (ρ = -0.32) for the first subperiod, during which *RMW* exhibits exceptionally poor performance during the 1970s as shown in Figure 1.3. However, their correlation turns out to be strongly positive (ρ = 0.71) for the second subperiod.

1.7. Explanatory powers of the FF5 and DM models

This section compares the FF5 model with the DM model in explaining average stock returns using time-series regressions, where the FF5 model includes the market, size, *BM*, profitability, and investment factors, while the DM model includes the market, size and *DM* factors.

1.7.1. Explaining mimicking factors

Table 1.8 analyzes the performance of mimicking factors relative to the FF5 model or the DM model. In the regressions of DM^r on the FF5 model, the intercepts (0.21, 0.18 and 0.39 for the 2x3, 2x5 and 2×10 sorts respectively) are still roughly one-third of the original monthly return of DM^r in table 1.5, with large *t*-statistics (*t* = 5.13, 3.73 and 5.00 respectively). Although DM^r loads heavily on *HML, RMW* and *CMA*, such loadings only explain about two-thirds of DM^r. The finding that DM^r is not fully explained by the FF5 model is in line with my earlier findings in FM regressions that the effect of *DM* cannot be fully driven out by *BM*, *OP* and *INV*.

In contrast, when the DM model is used to explain the *HML, RMW* and *CMA* in sequence, the intercepts are either negative or close to zero. For example, in the *HML* regressions, the intercepts are -0.21 (t = -3.29), -0.20 (t = -2.50) and -0.13 (t = -1.04) for the 2x3, 2x5 and 2×10 sorts respectively, as a result of heavy loadings on DM^r. The evidence suggests that a pure value strategy does not add abnormal returns for investors after accounting for DM^r. For the regressions of *RMW* and *CMA*, the intercepts are all close to zero and statistically insignificant. The result that the large returns of *HML, RMW* and *CMA* are completely absorbed by their exposure to DM^r reconfirms my earlier findings that controlling for *DM* drives out the effects of *BM*, *OP* and *INV* in average returns.

1.7.2. Explaining return anomalies

One level playing field to see which model provides a better description of average returns is using these models to analyse prominent anomalies that are not directly associated with valuation theory. Following Fama and French (2015b), the set of anomalies scrutinized in this study includes market beta,[6] net stock issues, volatility, accruals and momentum. Six sets of value-weighted anomaly portfolios are constructed: 25 *Size-Beta* (β) portfolios, formed at the end of each June, from independent 5×5 sorts of stocks on size and market β using NYSE breakpoints, where β is estimated using the most recent five years of past monthly returns (at least 24 past monthly observations); 25 *Size-Net stock issue* (*NI*) portfolios, formed in the same way as 25 *Size-Beta* portfolios, where the second sort variable *NI* is the change in the natural log of split-adjusted shares outstanding from the fiscal year-end in *t*-2 to the fiscal year-end in *t*-1; 25 *Size-Variance* (*Var*) portfolios, formed using monthly independent 5×5 sorts on size and the variance of daily returns in month *t*-1; 25 *Size-Residual variance* (*RVar*) portfolios, formed in the same way as 25 *Size-Var* portfolios, where the second sort variable *RVar* is the

[6] The relationship between market beta and average return is much flatter than implied by the Sharpe-Lintner CAPM.

variance of daily residuals in month *t*-1 from the FF3 model; 25 *Size-Accruals* (*AC*) portfolios, formed at the end of each June, from independent 5×5 sorts of stocks on size and accruals which are the change in operating working capital from the fiscal year-end in *t*-2 to *t*-1 divided by book equity in *t*-1; 25 *Size-Momentum* portfolios, formed using monthly independent 5×5 sorts on size and cumulative returns from month *t*-12 to *t*-2.

I assess the performance of the FF3 (the three-factor model of Fama and French 1993), FF5 and DM models by running the Gibbons, Ross, and Shanken (1989) test based on time-series regressions. The *GRS* test jointly tests whether the intercepts are different from zero. In other words, the GRS test asks whether the highest Sharpe ratio one can construct using both the left hand side portfolio (LHS) and the right hand side factors (RHS) is reliably higher than using RHS factors only [see Fama and French (2015b)].

Table 1.9 provides *GRS* statistics, *p*-values, the average absolute intercepts ($A|\alpha|$), the average standard errors of the intercept (*SE*) and the Sharpe ratios of the intercept (*SR*) for various models. Results on models 1-2 are generally in line with Fama and French (2015c). Except for the *Size-AC* sorts, the FF5 model performs at least as well as and generally better than the FF3 model in the *GRS* tests on different size-anomaly portfolios. This evidence suggests the FF5 model that includes profitability and investment factors improves the description of average returns provided by the FF3 model. However, the results on the *Size-AC* portfolios indicate that FF5 is likely to fare poorly when applied to portfolios with strong accrual tilts.

Of primary interest to us, the results from GRS tests for model 3 show that the DM model consistently outperforms the FF5 and FF3 models in its ability to provide better descriptions of average excess returns. For each panel, the DM model delivers lower GRS-statistics than the FF5 and FF3 models. Panel G show that, taking the average of the 6 anomaly portfolio sets, the *GRS*-statistics for the DM*,* FF5 and FF3 models are 3.33, 3.63 and 3.77, respectively. For model 4, where all relevant factors in the FF5 and DM models are included, the *GRS*-statistics hardly change relative to the DM model (*GRS* = 3.30 versus 3.33). Thus, it seems that *HML, RMW* and *CMA* are redundant when DM^r is included in the model.

Models 5 and 6 test the augmented versions of the FF5 and DM model by adding a momentum factor. *WML* (winner minus loser) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios constructed monthly using the 2×3 sorts on size and prior (month *t*-12 to *t*-2) returns. For each panel, the *GRS*-statistic for the augmented DM model (model 6) is lower than that for the augmented FF5 model (model 5). In panel G, the augmented FF5 and DM models have average *GRS*-statistics of 3.11 and 2.81, respectively.

Interestingly, in Panel E of the 25 *size-AC* portfolios, the DM model does not suffer the same problem that the FF5 model has — being less efficient than the FF3 model in explaining average returns. The *GRS*-statistic for the DM model (2.23) is slightly lower than the FF3 model (2.36) and is much lower than the FF5 model (2.93). The same outperformance also applies to the augmented models for the 25 *size-AC* portfolios. Here the GRS-statistics are 1.74 and 2.30 for the augmented DM and FF5 models, respectively.

1.8. Alternative profitability measures

Since profitability is the source of expected dividend, the dividend to market ratio can also be classified as a profitability measure. This section compares the performance of *DM* with alternative profitability measures head-to-head. These alternative profitability measures are the earnings-to-price ratio (E/P) of Ball (1978), the cash flow-to-price ratio (C/P) of Lakonishok et al.(1994), the gross profitability-to-asset ratio (*GP/AT*) of Novy-Marx (2013), the operating profitability-to-asset ratio (*OP/AT*) of Ball et al. (2015) and the quarterly earnings-to-equity ratio (ROE) of Hou et al. (2014).

Table 1.10 reports results of time-series regressions of alternative profitability factor using the *DM* factors and other control variables (Panel A), and time-series regressions of the *DM* factors using alternative profitability factor and other control variables (Panel B). All factors are constructed by double sorts (2×3) on size and respective measure. The regression intercepts reveals which of the profitability factor generate significant alpha relative to others. For robustness and facilitating comparison with prior studies, all regressions control for factors on market, size, *BM*, past return in month *t*-1 (*REV*), cumulative returns for the 11 months from *t*-12 to *t*-2 (*WML*) and standardized unexpected earnings (*SUE*). Regressions in this table cover January 1975 through December 2013, determined by the quarterly data requirements for constructing ROE and SUE measures.

Panel A shows all of alternative profitability factors do not exhibit significant alpha (intercept) over the sample, with the exception of operation profitability, which retains a weak significant alpha of 0.18% (t = 1.78). These results indicate that investors trading the *DM* factor largely cannot enhance their performance by incorporating alternative profitability factors. These alternative profitability factors all exhibit large loadings (from 0.20 to 0.51) on DM^r with high t - statistics (from 2.75 to 9.16), which indicates that these alternative profitability factors are substantially attenuated by the *DM* factor.

The results in panel B show that controlling for alternative profitability factors does not drive out the *DM* premium, as all regressions produce large and highly significant alphas. In first 5 regressions controlling for the earnings-price ratio, gross profitability, operating profitability and return on equity factors in sequence, intercepts are very similar, from 0.36 to 0.39 with *t*-statistics from 5.46 to 6.02. In the final regression that control for all alternative profitability factors together, the DM factor is associated with a large intercept of 0.32 (*t* = 5.20). This evidence shows that *DM* plays a much stronger role in explaining the cross-section of average returns than alternative profitability measures.

1.9. Conclusion

For value investing, *BM* is a noisy measure to identify undervalued opportunities, because book value is a weak indicator of intrinsic value. Thus, based on the dividend discount model of Miller and Modigliani (1961), this paper proposes an alternative metric for value investing, the dividend-to-market ratio. Since expected dividend has a strong link with intrinsic value, *DM* is much more efficient in identifying undervalued stocks than *BM*. Furthermore, *DM* provides a better estimation of expected stock returns than a linear combination of *BM*, profitability and investment, because the latter one omits the economic non-linear relation. Consistent with the prediction of the dividend discount model, my test results show that *DM* has a far stronger link with expected returns than *BM,* and it also outperforms a linear combination of *BM*, profitability and investment. These results persist in FM regressions using firm characteristics to explain stock returns, and in time series regressions using mimicking factors. Note also that, using *DM* instead of a combination of profitability, investment and *BM* to explain average stock return is also in line with the principle of parsimony, which prefers a model with fewer variables whenever it yields the same descriptive accuracy as the larger more complex model.

1.10. Appendix A

A.1. Composite trading strategies

If *DM* determines the cross-section of expected returns, and a linear combination of *BM*, profitability and investment can only work as a rough approximation for *DM,* then the *DM* factor should be able to outperform composite trading strategies that combine the effects of *BM*, profitability and investment. This section thus compares DM^r with two well-known composite trading strategies. The first strategy I implement is an equal combination of value, profitability and investment strategies. For each sorting algorithm (2×3, 2×5 and 2×10 sorts), the (*HML*+*RMW*)/2 strategy is the average return on *HML* and *RMW,* while the (*HML*+*CMA*)/2, (*RMW*+*CMA*)/2 and (*HML*+*RMW*+*CMA*)/3 strategies are formed similarly but based on different combination. The second strategy is a combined ranking strategy constructed from the 2×3, 2×5 and 2×10 sorts on size and an average ranking of particular multiple characteristics including *BM* (low to high), profitability (weak to robust) and investment (aggressive to conservative). Specifically, the approach is the same as for constructing factor mimicking portfolios in section 2.1, except that the second sort is based on the average ranking of characteristics: {*BM* & *OP*}, {*BM* & *INV*}, {*OP* & *INV*} and {*BM* & *OP* & *INV*}.

In Panel A of Table A1, the equal combination trading strategies show substantial reductions in standard deviations without enhancing economic rewards. For example, the strategy of (*HML*+*RMW*)/2 in the 2×3 sorts reduces the standard deviation to 2.11% from an average of 2.77% for *RMW* and *CMA* (2.86% and 2.67% in table 1.7). This results in an improved Sharpe ratio of 0.49 from an average of 0.38 for *RMW* and *CMA* (0.39 and 0.36 in table 1.7). However, such improvement is dwarfed by DM^r, as the ratio of these Sharpe ratios to the Sharpe ratio for DM^r, denoted by *SR/SR**, is only 60%. The synergistic effect between *CMA* and *HML* is limited as *CMA* fares poorly in the first subperiod as seen in table 1.7, and then strongly and positively correlates with *HML* (ρ = 0.60) in the second subperiod as seen in table 1.9. Similarly, due to correlations among factors, other strategies in the 2×3 sorts do not deliver much improvement in performance, with *SR/SR** ratios of 0.56, 0.70 and 0.67 for the (*HML*+*CMA*)/2, (*RMW*+*CMA*)/2 and (*HML*+*RMW*+*CMA*)/3 strategies. The results from the 2×5 and 2×10 sorts show that the underperformance of these equal combination trading strategies relative to DM^r persists.

In Panel B of Table 1.10, the combined ranking strategies slightly enhance mean returns but also push up standard deviations. For example, the {*BM* & OP} strategy in the 2×3 sorts produces a mean return of 0.42%, which is higher than the average of 0.30% for *RMW* and *CMA* (0.32% and 0.28% in table 1.7). However, its standard deviation also increases substantially to 3.22% from an average of 2.77% for *RMW* and *CMA* (2.67% and 2.28% in table 1.7), resulting in an *SR/SR** ratio of 0.55. The combined ranking strategies of {*BM* & *INV*}, {*OP* & *INV*} and {*BM* & *OP* & *INV*} in the 2×3 sorts also do not fare well as their Sharpe ratios are merely 50% to 58% of the Sharpe ratio for DM^r. The combined ranking strategies in the 2×5 and 2×10 sorts offer similarly poor results.

A.2. Profitability-to-market ratio and investment-to-market ratio

With a bit of manipulation, *DM* can be decomposed into a profitability-to-market ratio (*PM*) and an investment-to-market (*IM*) ratio, whereby equation (4) becomes

$$1 = \sum_{\tau=0}^{\infty} \frac{E(Y_{t+\tau})/M_t - E(\Delta B_{t+\tau})/M_t}{(1+\rho)^{\tau+1}} \tag{A}$$

Equation (A) says that the cross-section of expected returns could be explained by a linear combination of *PM* and *IM*. Stocks with high ratios of *PM* tend have higher expected returns, while stocks with high ratios of *IM* tend to have lower expected returns. Following equation (6), *PM* is constructed as revenues minus cost of goods sold, minus selling, general, and administrative expenses, plus expenditures on research & development, minus interest expense, all divided by market equity [$(REVT - COGS - XSGA + XRD - XINT)/M$]. *IM* is constructed as the product of book equity and the percentage change in total assets, divided by market equity [$(B * \Delta\, AT/AT)/M)$].

Table A2 tests the predictive powers of *PM* and *IM* in cross-sectional returns using FM regressions. In panel A, regression 1 shows the predictive power of *DM* for observations used in this table. Regressions 2 and 3 show that there is a strong positive relation between average returns and *PM* (t = 5.80), and a strong negative relation between average returns and *IM* (t = -3.16). These relations remain strong when *PM* and *IM* are simultaneously included in regression 4 (t = 6.17 for *PM* and t = -3.69 for *IM*). In panel B, regression 5 shows that *PM* and *IM* have incremental explanatory power (t = 4.17 and -3.69) after controlling for *ME* and *BM*. Regression 6 shows that, after controlling for *ME, BM, OP* and *INV*, the incremental power of *IM* becomes insignificant (t = -0.44) but the incremental power of *PM* remains significant (t = 2.03). In regression 7, controlling for *ME* and *DM* drives out the incremental explanatory power of *PM* (t = 1.29) and the slope on *IM* becomes positive (t = 2.4), which indicates that the predictive powers of *PM* and *IM* are not incremental relative to *DM*. In short, these results implies that both *PM* and *IM* have strong predictive power for average returns, and the profitability measure *PM* is the primary predictor variable for average returns as demonstrated by its higher significance levels in FM regressions.

1.11. References

Aharoni, G., B. Grundy, and Q. Zeng, 2013, Stock Returns and the Miller Modigliani Valuation Formula: Revisiting the Fama French Analysis, *Journal of Financial Economics* 110, 347-357.

Asness, C., A. Frazzini, R. Israel, and T. Moskowitz, 2015, Fact, Fiction, and Value Investing, *Journal of Portfolio Management* 42, 34-52.

Ball, R., 1978, Anomalies in Relationships between Securities' Yields and Yield-Surrogates, *Journal of Financial Economics* 6, 103-126.

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. V. Nikolaev, 2015, Deflating profitability, *Journal of Financial Economics* 117, 225-248.

Graham, Benjamin, 1973. The Intelligent Investor (4th Rev. ed.). New York: Haper.

Graham, B., and D. Dodd, 1934. Security Analysis. New Yor: Mcgraw Hill.

Fama, E. F., and K. R. French, 1992, The cross-section of expected returns, *Journal of Finance* 47, 427-465.

———, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

———, 2006, Profitability, Investment and Average Returns, *Journal of Financial Economics* 82, 491-518.

———, 2008, Average Returns, B/M, and Share Issues, *Journal of Finance* 63, 2971-2995.

———, 2015a, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

———, 2015b, Incremental variables and the investment opportunity set, *Journal of Financial Economics*.

———, 2015c, Dissecting Anomalies with a Five-Factor Model, *Journal of Financial studies*.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting Anomalies: An Investment Approach. *The Review of Financial Studies,28*(3), 650-705.

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

Lewellen, J., 2014, The Cross Section of Expected Stock Returns, working paper.

Miller, M. H., and F. Modigliani, 1961, Dividend Policy, Growth, and the Valuation of Shares, *Journal of Business* 34, 411-433.

Novy-Marx, R., 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1-28.

———, 2014, Quality Investing, Unpublished working paper. University of Rochester.

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

Rosenberg, B., K. Reid, and R. Lanstein, 1985, Persuasive Evidence of Market Inefficiency, *Journal of Portfolio Management* 11, 9-17.

Titman, S., K. C. J. Wei, and F. Xie, 2004, Capital Investments and Stock Returns, *Journal of Financial and Quantitative Analysis* 39, 677-700.

Williams, John, 1938. The theory of investment value. Amsterdam: North-Holland Pub.

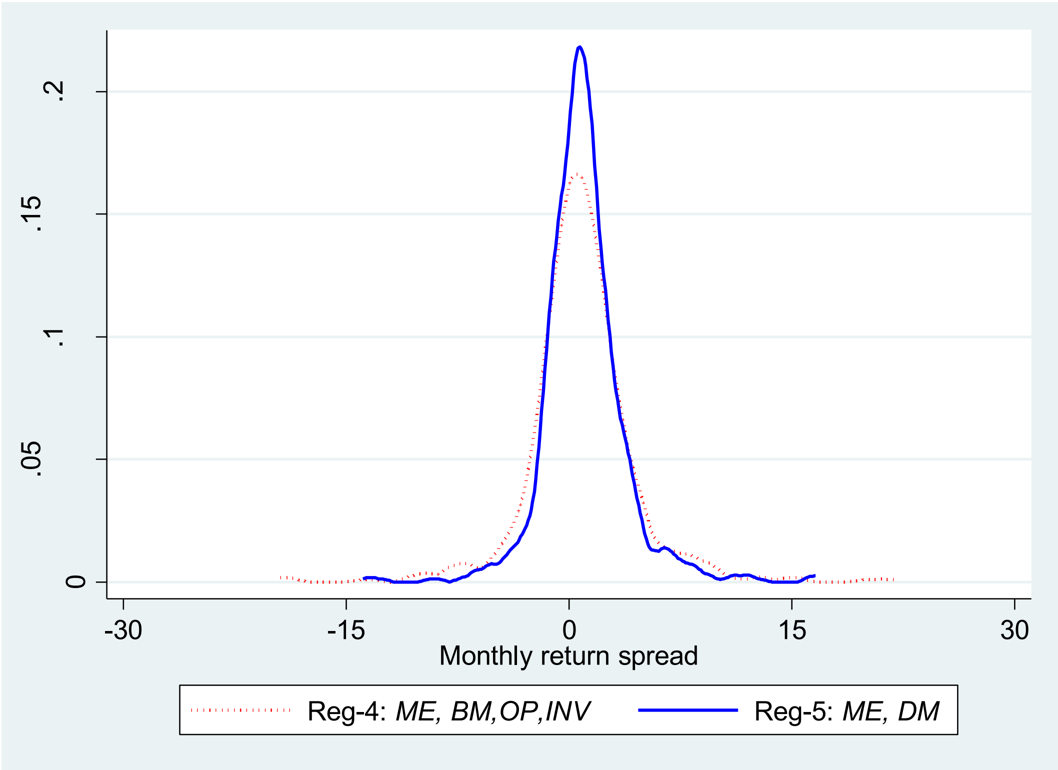

Figure 1.1 The density function for the time-series return spreads on tertile portfolios (All stocks)

For each of regression 1, 4 and 5 in Table 1.2, stocks are allocated to tertile portfolios using regression predicted values. The predicted values for month *t* are computed using the explanatory variables for month *t* and the average slopes from our Fama-Macbeth (FM) regressions. The density function of the spreads between the time-series returns on the top and bottom portfolios is constructed for All stocks. Reg-4 is regression 4 from Table 1.2, that uses the natural log of market cap (*ME*) and the natural log of the book to market ratio (*BM*), operating profitability (*OP*) and investment (*INV*) as independent variables. Reg-5 is regression 5 from Table 1.2, that uses *ME* and the dividend-to-market ratio (*DM*) as independent variables.

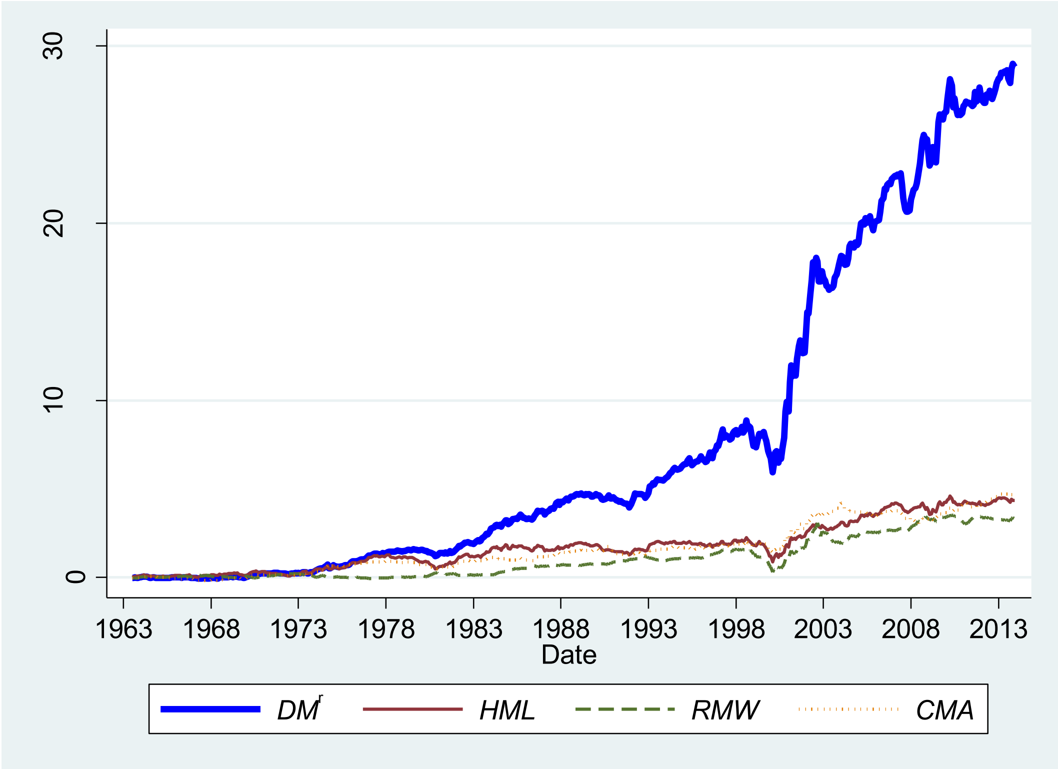

Figure 1.2 The cumulative returns to *DM*^{r}, *HML, RMW* and *CMA*

The plot shows the cumulative returns to *DM*^{r}*, HML, RMW* and *CMA*, which represent mimicking portfolios for the dividend-to-market ratio (*DM)*, book-to-market ratio (*BM*), operating profitability (*OP*) and investment (*INV*) factors, and are constructed using 2×3 sorts covering July 1963 to December 2013. At the end of each June, stocks are assigned to two size groups using the NYSE median market cap as breakpoint. Stocks are also assigned independently to tertile portfolios using the NYSE 30th and 70th percentiles of the ranked *DM* values. The intersections of the two sorts produce six value-weight *Size-DM* portfolios. *DM*^{r} is the average return on the two high *DM* portfolios minus the average return on the two low *DM* portfolios. In the sort for June of year *t*, *DM* is measured as revenues minus cost of goods sold, minus selling, general, and administrative expenses, plus expenditures on research & development, minus interest expense, minus the product of book equity and the percentage change in total assets for the last fiscal year ending in *t*-1, all divided by market cap in December of *t*-1. *HML* (high minus low), *RMW* (robust minus weak) and *CMA* (conservative minus aggressive) are formed in the same way as *DM*^{r}, except the second sort variable is *BM*, *OP* or *INV* respectively. In the sort for June of year *t*, *BM* is the ratio of book equity for the last fiscal year ending in *t*-1 divided by market cap in December of *t*-1; *OP* is revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity for the last fiscal year ending in *t*-1; *INV* is the change in total assets from the fiscal year ending in year *t*-2 to the fiscal year ending in *t*-1, divided by *t*-1 total assets.

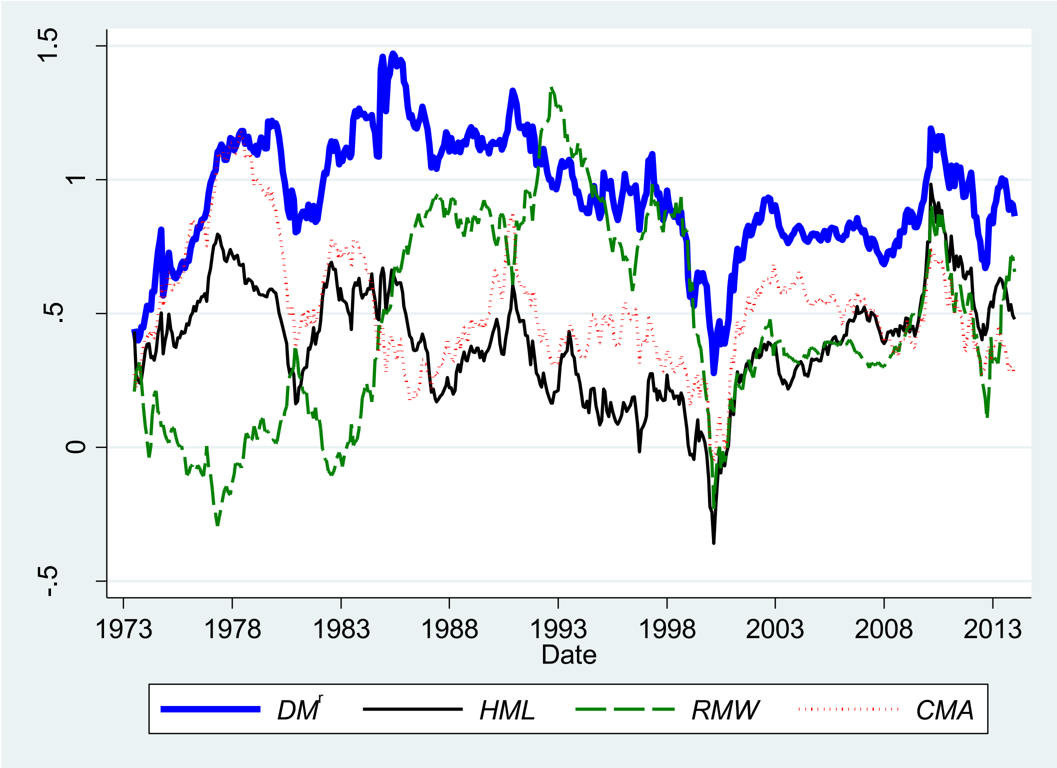

Figure 1.3 The trailing ten-year Sharpe ratios for DM^r, *HML*, *RMW* and *CMA*

The plot shows the trailing ten-year Sharpe ratios for DM^r, *HML*, *RMW* and *CMA,* which represent mimicking portfolios for the dividend-to-market ratio (*DM)*, book-to-market ratio (*BM*), operating profitability (*OP*) and investment (*INV*) factors, and are constructed using 2×3 sorts over the period from July 1963 to December 2013. At the end of each June, stocks are assigned to two size groups using the NYSE median market cap as breakpoint. Stocks are also assigned independently to tertile portfolios using the NYSE 30th and 70th percentiles of the ranked *DM* values. The intersections of the two sorts produce six value-weight *Size-DM* portfolios. DM^r is the average return on the two high *DM* portfolios minus the average return on the two low *DM* portfolios. In the sort for June of year *t*, *DM* is measured as revenues minus cost of goods sold, minus selling, general, and administrative expenses, plus expenditures on research & development, minus interest expense, minus the product of book equity and the percentage change in total assets for the last fiscal year ending in *t*-1, all divided by market cap in December of *t*-1. *HML* (high minus low), *RMW* (robust minus weak) and *CMA* (conservative minus aggressive) are formed in the same way as DM^r, except that the second sort variable is *BM*, *OP* or *INV* respectively. In the sort for June of year *t*, *BM* is the ratio of book equity for the last fiscal year ending in *t*-1 divided by market cap in December of *t*-1; *OP* is revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity for the last fiscal year ending in *t*-1; *INV* is the change in total assets from the fiscal year ending in year *t*-2 to the fiscal year ending in *t*-1, divided by *t*-1 total assets.

Table 1.1 Summary statistics; July 1963 to December 2013

This table reports summary statistics for portfolios formed by double sorts (3×3) on *DM* and one second-sort variable (*BM, OP* and *INV*). At the end of each June, stocks are independently assigned to three *BM* groups, and three Stocks are assigned independently to three groups based on one second-sort variable. The NYSE 30th and 70th percentile are used as sorting breakpoints. The intersections of the two sorts produce nine portfolios. Dividend-to-market ratio (*DM*) is revenues minus cost of goods sold, minus selling, general, and administrative expenses, plus expenditures on research & development, minus interest expense, minus the product of book equity and the percentage change in total assets for the last fiscal year ending in t-1, all divided by market cap in December of t-1; book-to-market ratio (*BM*) is the ratio of book equity for the last fiscal year ending in t-1 divided by market cap in December of t-1; operating profitability (*OP*) is revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity for the last fiscal year ending in t-1; investment (*INV*) is the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-1 total assets. The sample excludes financial firms and very small firms with total assets of less than $25 million or book equity of less than $12.5 million. All variables are trimmed at the 1% and 99% level. Adjusted returns are market and size factors adjusted portfolio returns, which are the intercepts form time-series regression of portfolio value-weighted returns on market value-weighted index and size factor (large minus small). *t*-stat is the test statistic.

	BM					*OP*					*INV*				
	Low	Med	High	H-L	*t*-stat	Low	Med	High	H-L	*t*-stat	Low	Med	High	H-L	*t*-stat
Panel A: Adjusted Returns (market and size factors adjusted)															
DM-Low	-0.20	-0.14	-0.10	0.10	0.63	-0.41	-0.23	-0.04	0.37	2.74	-0.13	-0.12	-0.23	-0.10	-0.58
DM-Med	0.10	0.11	0.17	0.07	0.52	0.04	0.07	0.16	0.12	1.13	0.16	0.10	0.15	-0.01	-0.12
DM-High	0.19	0.31	0.39	0.20	1.25	0.24	0.32	0.32	0.08	0.61	0.35	0.31	0.25	-0.10	-0.49
H-L	0.39	0.44	0.49			0.66	0.55	0.36			0.48	0.43	0.48		
t-stat	2.46	4.64	4.51			5.46	4.55	2.52			2.66	3.62	2.60		
Panel B: Average *DM*															
DM-Low	-0.01	-0.05	-0.08			-0.17	-0.01	0.03			-0.19	-0.04	-0.09		
DM-Med	0.12	0.14	0.13			0.12	0.13	0.13			0.14	0.13	0.12		
DM-High	0.27	0.31	0.56			0.54	0.38	0.44			0.50	0.36	0.48		
Panel C: Average Second sort variable															
		BM					*OP*					*INV*			
DM-Low	-1.47	-0.51	0.42			-0.05	0.26	0.45			-0.12	0.08	0.30		
DM-Med	-1.19	-0.44	0.15			0.08	0.27	0.48			-0.05	0.07	0.19		
DM-High	-1.24	-0.36	0.49			0.04	0.27	0.52			-0.12	0.06	0.21		
Panel D: Number of firms															
DM-Low	454	321	315			633	300	147			155	232	686		
DM-Med	317	454	234			281	437	286			233	525	244		
DM-High	63	280	491			278	316	241			492	269	69		

Table 1.2 Fama and Macbeth regressions; July 1963 to December 2013

The table shows the average slopes and their *t*-statistics (in parentheses) from cross-sectional regressions that predict monthly returns, for July 1963 to December 2013. Separate regressions are run for All stocks, All but micro stocks and Microcap stocks which are below the 20th percentile of NYSE market cap as at the end of June of each year. In Panel A, the dependent variable is monthly excess returns for individual stocks. In Panel B, the dependent variable is characteristic-adjusted returns (residual returns) for individual stocks from either regression 4 or 5.

Panel A: Regressions of stock excess returns

	All					All but Micro					Micro				
	BM	*OP*	*INV*	*DM*	*ME*	*BM*	*OP*	*INV*	*DM*	*ME*	*BM*	*OP*	*INV*	*DM*	*ME*
Reg-1	0.311				-0.034	0.238				-0.048	0.536				0.106
	(3.41)				(-0.74)	(2.55)				(-1.10)	(5.42)				(0.75)
Reg-2		0.610			-0.103		0.538			-0.088		0.197			-0.110
		(2.22)			(-2.58)		(1.92)			(-2.12)		(0.56)			(-0.79)
Reg-3			-1.524		-0.068			-1.465		-0.073			-1.810		-0.023
			(-4.89)		(-1.72)			(-4.18)		(-1.86)			(-6.00)		(-0.18)
Reg-4	0.287	1.200	-1.014		-0.064	0.207	1.028	-0.940		-0.081	0.409	0.589	-1.250		0.224
	(3.11)	(4.08)	(-4.23)		(-1.67)	(2.12)	(3.56)	(-3.67)		(-2.08)	(3.64)	(1.26)	(-3.70)		(1.58)
Reg-5				2.409	-0.081				2.152	-0.086				2.527	0.051
				(6.99)	(-2.08)				(5.24)	(-2.17)				(7.12)	(0.39)
Reg-6	0.155	0.440	-0.054	1.706	-0.060	0.101	0.462	-0.261	1.285	-0.079	0.934	2.272	-3.246	-2.271	0.292
	(1.51)	(1.27)	(-0.14)	(4.47)	(-1.57)	(0.93)	(1.35)	(-0.63)	(2.98)	(-2.03)	(0.39)	(0.27)	(-0.34)	(-0.18)	(2.05)

Panel B: Regression of characteristic-adjusted returns (residual returns)

Regression	Dependent variable	All				All but Micro				Micro			
		BM	*OP*	*INV*	*DM*	*BM*	*OP*	*INV*	*DM*	*BM*	*OP*	*INV*	*DM*
Reg-7	Char-adj returns from reg-4: *ME, BM,OP* and *INV*				0.790				0.680				0.677
					(2.37)				(1.71)				(1.70)
Reg-8	Char-adj returns from reg-5: *ME* and *DM*	0.098	0.228	0.411		0.060	0.327	0.291		0.151	-0.797	0.414	
		(1.09)	(0.72)	(1.44)		(0.61)	(1.02)	(0.95)		(1.36)	(-1.56)	(1.28)	

Table 1.3 Average return spreads between the low and high expected return portfolios formed on regression fitted values

At the end of June each year, stocks are allocated to equal-weight tertile, quintile or decile portfolios using regression predicted values, which are the fitted values of regressions 1, 4 and 5 in Table 1.2. These fitted values are the average regression slopes multiplied by the value of explanatory variables at the end of each June. The average spread (Mean) between monthly returns to the top and bottom portfolios, and standard deviations (STD), *t*-statistics and annualized Sharpe ratios (SR) for those spreads are reported for All stocks, All but micro stocks and Microcap stocks which are below the 20th percentile of NYSE market cap.

Regression # from Table 1.2	Tertile spread				Quintile spread				Decile spread			
	Mean	STD	t-stat	SR	Mean	STD	t-stat	SR	Mean	STD	t-stat	SR
All												
Reg-1: *ME, BM*	0.64	3.04	5.15	0.72	0.72	3.56	5.00	0.70	0.84	4.27	4.86	0.68
Reg-4: *ME, BM, OP, INV*	0.83	3.60	5.70	0.80	1.04	4.33	5.91	0.83	1.30	5.28	6.07	0.85
Reg-5: *ME, DM*	0.99	2.74	8.90	1.25	1.17	3.11	9.22	1.30	1.42	3.41	10.25	1.44
All but Micro												
Reg-1: *ME, BM*	0.51	3.02	4.16	0.59	0.44	3.69	2.93	0.41	0.70	4.14	4.14	0.58
Reg-4: *ME, BM, OP, INV*	0.69	3.61	4.71	0.66	0.73	4.28	4.20	0.59	1.10	5.27	5.14	0.72
Reg-5: *ME, DM*	0.75	2.79	6.62	0.93	0.78	3.27	5.89	0.83	1.06	3.52	7.39	1.04
Micro												
Reg-1: *ME, BM*	0.75	3.57	5.13	0.72	0.85	4.35	4.82	0.68	1.10	5.54	4.88	0.69
Reg-4: *ME, BM, OP, INV*	0.99	3.79	6.41	0.90	1.13	4.59	6.04	0.85	1.49	5.64	6.43	0.91
Reg-5: *ME, DM*	1.08	3.23	8.25	1.16	1.37	3.71	9.07	1.28	1.63	4.18	9.52	1.35

Table 1.4 Return spreads on subsets of extreme deciles formed on regression predicted values (All stocks)

For each of regressions 1, 4 and 5 in Table 1.2, stocks are allocated to equal-weight decile portfolios using regression predicted values. Subset 1 includes group of stocks that are listed in the extreme deciles (low and high) by regression 4 controlling for *ME*, *BM, OP* and *INV,* but not in the extreme deciles by regression 5 controlling for *ME* and *DM*. Subset 2 includes group of stocks that are listed in the extreme deciles by both regression 4 and regression 5. Subset 3 includes group of firms that are listed in the extreme deciles by regression 5, but not in the extreme deciles by regression 4. Subset N/decile N is the percentage of stocks in a subset relative to total stocks in the extreme decile. The average spreads (Mean) between the returns to the top and bottom deciles, and corresponding standard deviations (STD), t-statistics and annualized Sharpe ratios (SR) are constructed for each subset.

Subset	Subset N/decile N		Mean	STD	t-stat	SR
	Low decile	High decile				
1.(Reg-4: *ME, BM,OP,INV*)\(Reg-5: *ME, DM*)	47%	27%	0.98	6.17	3.9	0.55
2.(Reg-4: *ME, BM,OP,INV*)∩(Reg-5: *ME, DM*)	53%	73%	1.51	5.13	7.24	1.02
3.(Reg-5: *ME, DM*)\(Reg-4: *ME, BM,OP,INV*)	47%	27%	1.29	4.24	7.46	1.05

Table 1.5 Summary statistics for mimicking portfolio monthly returns; July 1963 to December 2013

The table reports summary statistics for DM^r, *HML, RMW* and *CMA*, which represent mimicking portfolios for the dividend-to-market ratio (*DM)*, book-to-market ratio (*BM*), operating profitability (*OP*) and investment (*INV*) factors. At the end of each June, stocks are assigned to two Size groups using the NYSE median market cap as breakpoint. Using 2×3 sorts, stocks are also independently assigned to tertile portfolios using the NYSE 30th and 70th percentiles of ranked values for the dividend-to-market ratio (*DM*). The intersections of the two sorts produce six value-weight *Size-DM* portfolios. The *DM* factor, DM^r, is the average return on the two high *DM* portfolios minus the average return on the two low *DM* portfolios. *HML* (high minus low), *RMW* (robust minus weak) and *CMA* (conservative minus aggressive) are formed in the same way as DM^r, except the second sort variable is *BM*, *OP* and *INV* respectively. Factor mimicking portfolios in the 2×5 sorts and the 2×10 sorts are formed in the same way as for the 2×3 sorts and report the average spread between the top and bottom portfolios, except that stocks are assigned independently to quintile and decile portfolios for the second sort. Mean is the time-series mean of monthly returns, STD is its time-series standard deviation, *t*-stat is the test statistic, and SR is the annualized Sharpe ratio.

	2×3 factors				2×5 factors				2×10 factors			
	Mean	STD	*t*-stat	SR	Mean	STD	*t*-stat	SR	Mean	STD	*t*-stat	SR
Panel A: July 1963 to December 2013 (606 Monthly Observations.)												
DM^r	0.59	2.53	5.77	0.81	0.66	2.95	5.51	0.78	0.89	3.22	6.78	0.95
HML	0.32	2.86	2.74	0.39	0.37	3.30	2.73	0.38	0.61	4.00	3.72	0.52
RMW	0.28	2.67	2.55	0.36	0.34	3.14	2.65	0.37	0.31	4.06	1.86	0.26
CMA	0.31	2.28	3.39	0.48	0.39	2.61	3.68	0.52	0.55	3.20	4.25	0.60
Panel B: July 1963 to June 1988 (300 Monthly Observations.)												
DM^r	0.58	2.15	4.70	0.94	0.63	2.51	4.33	0.87	0.82	2.90	4.91	0.98
HML	0.38	2.86	2.28	0.46	0.40	3.26	2.15	0.43	0.64	4.01	2.77	0.55
RMW	0.19	1.87	1.80	0.36	0.19	2.12	1.57	0.31	0.08	2.86	0.50	0.10
CMA	0.31	2.07	2.63	0.53	0.33	2.47	2.32	0.46	0.51	2.94	2.98	0.60
Panel C: July 1988 to December 2013 (306 Monthly Observations.)												
DM^r	0.60	2.86	3.68	0.73	0.69	3.32	3.64	0.72	0.95	3.52	4.74	0.94
HML	0.26	2.86	1.60	0.32	0.33	3.34	1.72	0.34	0.57	3.99	2.49	0.49
RMW	0.36	3.27	1.91	0.38	0.48	3.88	2.16	0.43	0.53	4.95	1.86	0.37
CMA	0.31	2.47	2.22	0.44	0.45	2.74	2.86	0.57	0.60	3.44	3.04	0.60

Table 1.6 Mimicking portfolio monthly returns for small and big stocks; July 1963 to December 2013

The table reports summary statistics for DM^r, *HML*, *RMW* and *CMA* for small and big stocks. DM_S^r is the difference in average return for the small high *DM* portfolio and the small low *DM* portfolio. DM_B^r is similarly defined for portfolios of big stocks, and DM_{S-B}^r is the difference between DM_S^r and DM_B^r. $HML_S^r, HML_B^r, HML_{S-B}^r$ and $RMW_S^r, RMW_B^r, RMW_{S-B}^r$ and $CMA_S^r, CMA_B^r, CMA_{S-B}^r$ are formed in the same way, but for portfolios sorted by book-to-market ratio (*BM*), operating profitability (*OP*) and investment (*INV*). Mean is the time-series mean of monthly returns, STD is its time-series standard deviation, *t*-stat is the test statistic, and SR is the annualized Sharpe ratio.

	2×3 factors				2×5 factors				2×10 factors			
	Mean	STD	t-stat	SR	Mean	STD	t-stat	SR	Mean	STD	t-stat	SR
DM_S^r	0.75	2.58	7.17	1.01	0.82	2.97	6.81	0.96	1.04	3.44	7.46	1.05
DM_B^r	0.44	3.03	3.54	0.50	0.50	3.60	3.40	0.48	0.74	3.98	4.55	0.64
DM_{S-B}^r	0.32	2.46	3.18	0.45	0.33	2.98	2.69	0.38	0.31	3.70	2.04	0.29
HML_S^r	0.49	3.14	3.81	0.54	0.58	3.72	3.83	0.54	0.75	4.67	3.95	0.56
HML_B^r	0.15	3.09	1.19	0.17	0.15	3.58	1.05	0.15	0.46	4.53	2.50	0.35
HML_{S-B}^r	0.34	2.48	3.35	0.47	0.43	3.13	3.35	0.47	0.29	4.56	1.56	0.22
RMW_S^r	0.31	2.92	2.64	0.37	0.35	3.47	2.48	0.35	0.39	4.39	2.16	0.30
RMW_B^r	0.24	3.06	1.93	0.27	0.33	3.63	2.21	0.31	0.23	4.89	1.15	0.16
RMW_{S-B}^r	0.07	2.69	0.66	0.09	0.02	3.33	0.17	0.02	0.16	4.52	0.85	0.12
CMA_S^r	0.49	2.22	5.41	0.76	0.58	2.60	5.53	0.78	0.76	3.21	5.85	0.82
CMA_B^r	0.14	3.01	1.14	0.16	0.20	3.38	1.43	0.20	0.34	4.34	1.94	0.27
CMA_{S-B}^r	0.35	2.67	3.22	0.45	0.39	3.03	3.13	0.44	0.42	4.16	2.49	0.35

Table 1.7 Correlations between mimicking portfolios; July 1963 to December 2013

The table reports Pearson correlations for DM^r, *HML, RMW* and *CMA*, which represent mimicking portfolios for the dividend-to-market ratio (*DM)*, book-to-market ratio (*BM*), operating profitability (*OP*) and investment (*INV*) factors.

	2×3 factors				2×5 factors				2×10 factors			
	DM^r	*HML*	*RMW*	*CMA*	DM^r	*HML*	*RMW*	*CMA*	DM^r	*HML*	*RMW*	*CMA*
Panel A: July 1963 to December 2013 (606 Monthly Observations.)												
DM^r	1.00				1.00				1.00			
HML	0.85	1.00			0.82	1.00			0.69	1.00		
RMW	0.40	0.16	1.00		0.42	0.17	1.00		0.41	0.15	1.00	
CMA	0.78	0.74	0.06	1.00	0.76	0.70	0.02	1.00	0.65	0.62	-0.01	1.00
Panel B: July 1963 to June 1988 (300 Monthly Observations.)												
DM^r	1.00				1.00				1.00			
HML	0.81	1.00			0.79	1.00			0.62	1.00		
RMW	-0.32	-0.61	1.00		-0.25	-0.60	1.00		-0.05	-0.45	1.00	
CMA	0.75	0.76	-0.63	1.00	0.74	0.76	-0.59	1.00	0.64	0.68	-0.43	1.00
Panel C: July 1988 to December 2013 (306 Monthly Observations.)												
DM^r	1.00				1.00				1.00			
HML	0.89	1.00			0.85	1.00			0.76	1.00		
RMW	0.71	0.60	1.00		0.69	0.59	1.00		0.63	0.50	1.00	
CMA	0.80	0.73	0.39	1.00	0.78	0.66	0.32	1.00	0.66	0.57	0.20	1.00

Table 1.8 Time-series regressions of mimicking portfolios using the DM and FF5 models; July 1963 to December 2013

The table shows alpha (percent per month) from monthly time-series regressions using the DM and FF5 models, where the DM model includes the market, size and DM factors, while the FF5 model includes the market, size, BM, profitability, and investment factors. *Mkt* is the value-weight return on the market portfolio of all stocks, minus the one month Treasury bill rate; the size factor (*SMB*: small minus big) is the value-weight return on the small-cap portfolio minus the value-weight return to the large-cap portfolio, which are formed at the end of each June using the NYSE median market cap as breakpoint. DM^r, *HML, RMW* and *CMA* are constructed as described in Table 1.7 using 2x3, 2x5 and 2x10 sorts. The DM model includes *Mkt, SMB* and DM^r, while the FF5 model includes *Mkt, SMB*, *HML, RMW* and *CMA* as explanatory variables.

	Int	*Mkt*	*SMB*	*HML*	*RMW*	*CMA*	DM^r	*R2*
Panel A: 2×3 factors								
DM^r	0.21	0.01	0.08	0.46	0.30	0.43		0.86
	(5.13)	(0.54)	(4.93)	(22.07)	(19.49)	(16.02)		
HML	-0.21	-0.04	-0.01				0.93	0.72
	(-3.29)	(-2.31)	(-0.35)				(35.02)	
RMW	0.07	0.01	-0.20				0.41	0.20
	(0.70)	(0.55)	(-5.38)				(9.81)	
CMA	-0.03	-0.07	-0.03				0.65	0.62
	(-0.47)	(-4.94)	(-1.28)				(26.43)	
Panel B: 2×5 factors								
DM^r	0.18	0.01	0.08	0.41	0.33	0.50		0.84
	(3.73)	(0.44)	(4.27)	(19.43)	(20.01)	(18.73)		
HML	-0.20	-0.05	0.01				0.89	0.67
	(-2.50)	(-2.32)	(0.42)				(31.44)	
RMW	0.12	0.00	-0.27				0.42	0.23
	(1.01)	(0.13)	(-6.20)				(10.12)	
CMA	0.01	-0.07	0.01				0.63	0.59
	(0.07)	(-4.05)	(0.26)				(25.35)	
Panel C: 2×10 factors								
DM^r	0.39	-0.02	0.06	0.30	0.29	0.42		0.68
	(5.00)	(-1.21)	(1.92)	(12.51)	(14.43)	(13.64)		
HML	-0.13	-0.05	0.07				0.84	0.48
	(-1.04)	(-1.84)	(1.63)				(21.66)	
RMW	-0.00	-0.05	-0.33				0.46	0.22
	(-0.03)	(-1.46)	(-5.79)				(9.55)	
CMA	0.05	-0.09	0.04				0.61	0.44
	(0.47)	(-3.60)	(0.98)				(18.74)	

Table 1.9 Tests of FF3, FF5 and DM models using return anomalies: July 1963 to December 2013

The table provides GRS statistics (*GRS*), *p*-values, the average absolute intercepts (A|α|), the average standard error of the intercepts (SE) and the Sharpe ratio of the intercepts (SR) for competing models and augmented versions of these models that include the momentum factor. The six sets of anomaly portfolios are: 25 *Size-Beta* (β) portfolios, formed at the end of each June, from independent 5×5 sorts of stocks on size and market β using NYSE breakpoints, where β is estimated using the most recent five years of past monthly returns (at least 24 past monthly observations); 25 *Size-Net stock issue* (*NI*) portfolios, formed in the same way as 25 *Size-Beta* portfolios, where the second sort variable *NI* is the change in the natural log of split-adjusted shares outstanding from the fiscal year-end in *t*-2 to the fiscal year-end in *t*-1; 25 *Size-Variance* (*Var*) portfolios, formed using monthly independent 5×5 sorts on size and the variance of daily returns in month *t*-1; 25 *Size-Residual variance* (*RVar*) portfolios, formed in the same way as 25 *Size-Var* portfolios, where the second sort variable *RVar* is the variance of daily residuals in month *t*-1 from the FF3 model; 25 *Size-Accruals* (*AC*) portfolios, formed at the end of each June, from independent 5×5 sorts of stocks on Size and accruals which are the change in operating working capital from the fiscal year-end in *t*-2 to *t*-1 divided by book equity in *t*-1; 25 *Size-Momentum* portfolios, formed using monthly independent 5×5 sorts on size and cumulative returns from month *t*-12 to *t*-2. Panel G show average statistics for the 6 anomaly portfolio sets. *Mkt, SMB, HML, RMW, CMA* and DM^r are constructed using 2×3 sorts as in Tables 6 and 8. The momentum mimicking factor (*WML*: winner minus loser) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios, formed using monthly independent 2×3 sorts on size and prior (2-12) returns.

	Model factors	*GRS*	*p*	*A*|*a*|	*SE*	*SR*
	Panel A: 25 Size-Beta portfolios					
(1)	*Mkt, SMB, HML*	1.66	0.02	0.11	0.09	0.27
(2)	*Mkt, SMB, HML, RMW, CMA*	1.68	0.02	0.08	0.08	0.28
(3)	*Mkt, SMB,* DM^r	1.43	0.08	0.09	0.09	0.26
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	1.37	0.11	0.08	0.09	0.26
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	1.31	0.15	0.07	0.08	0.25
(6)	*Mkt, SMB,* DM^r*, WML*	1.09	0.35	0.07	0.09	0.23
	Panel B: 25 Size-NI portfolios					
(1)	*Mkt, SMB, HML*	3.12	0.00	0.13	0.08	0.38
(2)	*Mkt, SMB, HML, RMW, CMA*	2.50	0.00	0.09	0.08	0.34
(3)	*Mkt, SMB,* DM^r	2.34	0.00	0.09	0.08	0.34
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	2.40	0.00	0.09	0.08	0.34
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	2.01	0.00	0.08	0.08	0.31
(6)	*Mkt, SMB,* DM^r*, WML*	1.83	0.01	0.08	0.08	0.30
	Panel C: 25 Size-Var portfolios					
(1)	*Mkt, SMB, HML*	5.15	0.00	0.17	0.09	0.48
(2)	*Mkt, SMB, HML, RMW, CMA*	4.87	0.00	0.12	0.08	0.48
(3)	*Mkt, SMB,* DM^r	4.58	0.00	0.13	0.09	0.47
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	4.45	0.00	0.12	0.08	0.47
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	4.55	0.00	0.12	0.08	0.47
(6)	*Mkt, SMB,* DM^r*, WML*	4.17	0.00	0.11	0.09	0.46
	Panel D: 25 Size-RVar portfolios					
(1)	*Mkt, SMB, HML*	5.63	0.00	0.17	0.08	0.51
(2)	*Mkt, SMB, HML, RMW, CMA*	5.20	0.00	0.11	0.08	0.49
(3)	*Mkt, SMB,* DM^r	5.08	0.00	0.11	0.08	0.49
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	5.13	0.00	0.11	0.08	0.50
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	4.87	0.00	0.11	0.08	0.49
(6)	*Mkt, SMB,* DM^r*, WML*	4.68	0.00	0.10	0.08	0.48
	Panel E: 25 Size-AC portfolios					
(1)	*Mkt, SMB, HML*	2.36	0.00	0.09	0.08	0.33
(2)	*Mkt, SMB, HML, RMW, CMA*	2.93	0.00	0.11	0.08	0.37
(3)	*Mkt, SMB,* DM^r	2.23	0.00	0.10	0.08	0.33

	Model factors	*GRS*	*p*	$A\|a\|$	*SE*	*SR*
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	2.45	0.00	0.10	0.08	0.35
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	2.30	0.00	0.10	0.08	0.34
(6)	*Mkt, SMB,* DM^r*, WML*	1.74	0.01	0.09	0.08	0.30
	Panel F: 25 Size-Momentum portfolios					
(1)	*Mkt, SMB, HML*	4.72	0.00	0.29	0.10	0.46
(2)	*Mkt, SMB, HML, RMW, CMA*	4.60	0.00	0.30	0.10	0.47
(3)	*Mkt, SMB,* DM^r	4.29	0.00	0.31	0.10	0.45
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	3.99	0.00	0.26	0.10	0.44
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	3.64	0.00	0.12	0.08	0.42
(6)	*Mkt, SMB,* DM^r*, WML*	3.33	0.00	0.12	0.08	0.41
	Panel G: Average					
(1)	*Mkt, SMB, HML*	3.77	0.00	0.16	0.09	0.40
(2)	*Mkt, SMB, HML, RMW, CMA*	3.63	0.00	0.13	0.08	0.41
(3)	*Mkt, SMB,* DM^r	3.33	0.01	0.14	0.09	0.39
(4)	*Mkt, SMB, HML, RMW, CMA,* DM^r	3.30	0.02	0.13	0.08	0.39
(5)	*Mkt, SMB, HML, RMW, CMA, WML*	3.11	0.02	0.10	0.08	0.38
(6)	*Mkt, SMB,* DM^r*, WML*	2.81	0.06	0.09	0.08	0.36

Table 1.10 Spanning tests

The table shows alphas (percent per month) from monthly time-series regressions of alternative profitability factor using the *DM* factor and other control variables (Panel A), or from time-series regressions of the *DM* factors using alternative profitability factor and other control variables (Panel B) for January 1975 to December 2013. E/P^r, C/P^r, GP/AT^r, OP/AT^r, REV^r and SUE^r are the mimicking factors for the earnings-price ratio (*E/P*), the cash flow to price ratio (*C/P*), the gross profitability-to-asset ratio (*GP/AT*), the operating profitability-to-asset ratio *(OP/AT)*, short-term reversal and standardized unexpected earnings (*SUE*). These mimicking factors are formed in the same 2×3 sorts as DM^r in Table 1.7, except that the second sort variable is the respective variable. Cash flow in *C/P* is net income plus amortization and depreciation minus changes in working capital and capital expenditures. *SUE* is measured as the most recent year-over-year change in income before extraordinary items, deflated by the standard deviation of the innovations of income before extraordinary items over the last eight announcements (at least six to be included).

Panel A: Alternative profitability factors as dependant variables

	Int	DM^r	*Mkt*	*HML*	*SMB*	REV^r	*WML*	SUE^r	*R2*
E/P^r	*-0.03*	*0.50*	*-0.13*	*0.31*	*-0.23*	*0.08*	*0.03*	*0.02*	*0.69*
	(-0.28)	*(7.97)*	*(-6.13)*	*(5.30)*	*(-6.52)*	*(3.12)*	*(1.36)*	*(0.42)*	
C/P^r	-0.19	0.63	-0.10	-0.02	-0.20	0.11	-0.06	0.30	0.53
	(-1.79)	(9.13)	(-4.30)	(-0.27)	(-5.21)	(3.69)	(-2.30)	(5.18)	
GP/AT^r	0.07	0.2	0.01	-0.55	0.16	0.04	-0.07	0.24	0.32
	(0.67)	(2.75)	(0.30)	(-8.18)	(3.97)	(1.18)	(-2.58)	(4.04)	
OP/AT^r	0.18	0.27	-0.03	-0.62	0.05	-0.03	0.01	0.15	0.35
	(1.78)	(4.15)	(-1.43)	(-10.06)	(1.30)	(-1.17)	(0.38)	(2.80)	
ROE^r	-0.11	0.51	-0.09	-0.32	-0.14	0.04	0.03	0.89	0.69
	(-1.32)	(9.16)	(-4.61)	(-6.11)	(-4.69)	(1.56)	(1.47)	(19.46)	

Panel B:The *DM* factor as dependant variable

Int	E/P^r	C/P^r	GP/AT^r	OP/AT^r	ROE^r	*Mkt*	*HML*	*SMB*	REV^r	*WML*	SUE^r	*R2*
0.36	0.24					0.01	0.61	0.06	-0.04	0.04	0.01	0.78
(5.68)	(7.97)					(0.40)	(19.46)	(2.31)	(-2.30)	(2.33)	(0.38)	
0.39		0.24				-0.00	0.66	0.05	-0.05	0.06	-0.05	0.79
(6.22)		(9.13)				(-0.00)	(25.78)	(2.12)	(-2.62)	(3.67)	(-1.47)	
0.39			0.08			-0.03	0.81	-0.01	-0.03	0.06	0.00	0.75
(5.82)			(2.75)			(-1.86)	(30.15)	(-0.36)	(-1.44)	(3.30)	(0.04)	
0.37				0.13		-0.02	0.83	-0.00	-0.02	0.05	0.00	0.76
(5.46)				(4.15)		(-1.53)	(30.58)	(-0.11)	(-1.06)	(2.88)	(0.02)	
0.38					0.30	0.00	0.75	0.05	-0.03	0.04	-0.25	0.79
(6.02)					(9.16)	(0.12)	(33.47)	(1.93)	(-1.81)	(2.17)	(-5.44)	
0.31	0.05	0.25	-0.05	0.26	0.04	0.02	0.70	0.07	-0.04	0.05	-0.12	0.82
(5.16)	(1.17)	(5.84)	(-1.09)	(4.74)	(0.89)	(1.63)	(22.73)	(2.83)	(-2.39)	(2.87)	(-2.55)	

Table A1 Performance of composite trading strategies

Panel A reports summary statistics for equal combinations of the value, profitability and investment strategies. For each sorting algorithm (2×3, 2×5 and 2×10 sorts), the (*HML*+*RMW*)/2 strategy is the average of the returns to *HML* and *RMW*. Returns to the (*HML*+*CMA*)/2, (*RMW*+*CMA*)/2 and (*HML*+*RMW*+*CMA*)/3 strategies are formed similarly but based on different factors. Panel B reports summary statistics for combined ranking strategies constructed from the 2×3, 2×5 and 2×10 sorts on size and the average ranking of particular multiple characteristics including *BM* (low to high), profitability (weak to robust) and investment (aggressive to conservative). For example, the {*BM* & *OP*} strategy is formed in the same way as a mimicking portfolio, except that the second sort variable is the average of the *BM* & *OP* rankings for individual stocks. The {*BM* & *INV*}, {*OP* & *INV*} and {*BM* & *OP* & *INV*} strategies are formed in the same way as the {*BM* & *OP*} strategy but based on different characteristic combinations. Mean is the time-series mean of monthly returns, STD is its time-series standard deviation, *t*-stat is the mean divided by its time-series standard error, and SR is the annualized Sharpe ratio. SR/SR* is the ratio of the Sharpe ratio on the composite trading strategy relative to the Sharpe ratio for DM^r.

	2×3 factors					2×5 factors					2×10 factors				
	Mean	STD	t-stat	SR	SR/SR*	Mean	STD	t-stat	SR	SR/SR*	Mean	STD	t-stat	SR	SR/SR*
Panel A: Equal combination															
(HML+RMW)/2	0.30	2.11	3.47	0.49	0.60	0.35	2.47	3.51	0.49	0.63	0.46	3.06	3.67	0.52	0.54
(HML+CMA)/2	0.32	2.40	3.24	0.46	0.56	0.38	2.73	3.41	0.48	0.62	0.58	3.24	4.39	0.62	0.65
(RMW+CMA)/2	0.30	1.81	4.01	0.57	0.70	0.36	2.06	4.34	0.61	0.78	0.43	2.58	4.10	0.58	0.61
(HML+RMW+CMA)/3	0.30	1.93	3.87	0.54	0.67	0.36	2.20	4.08	0.57	0.73	0.49	2.65	4.53	0.64	0.67
Panel B: Combined ranking															
BM & OP	0.42	3.22	3.17	0.45	0.55	0.51	3.61	3.47	0.49	0.63	0.64	4.40	3.57	0.50	0.53
BM & INV	0.33	2.77	2.90	0.41	0.50	0.46	3.27	3.45	0.49	0.62	0.62	3.99	3.80	0.54	0.56
OP & INV	0.36	2.70	3.27	0.46	0.57	0.45	3.04	3.60	0.51	0.65	0.63	3.64	4.24	0.60	0.63
BM & OP & INV	0.43	3.18	3.32	0.47	0.58	0.50	3.72	3.29	0.46	0.59	0.69	4.48	3.81	0.54	0.56

Table A2 Fama and Macbeth regressions of returns using profitability-to-price and investment-to-price ratios

The table shows the average slopes and their *t*-statistics (in parentheses) from cross-sectional regressions for July 1963 to December 2013. In Panel A, the dependent variable is stock monthly excess returns. In Panel B, the dependent variable is characteristic-adjusted returns (residual returns) from cross-sectional regressions using residual returns from either regression 1, 4 or 5 from Table 1.2. The independent variables are defined as follows: the profitability-to-market ratio (*PM*) is revenues minus cost of goods sold, minus selling, general, and administrative expenses, plus expenditures on research & development, minus interest expense for the last fiscal year ending in *t*-1, all divided by market cap in December of year *t*-1; the investment-to-market ratio (*IM*) is the product of book equity with the percentage change in total assets for the last fiscal year ending in *t*-1, divided by market cap in December of year *t*-1, where the percentage change in total assets is the change in total assets from the fiscal year ending in year *t*-2 to the fiscal year ending in *t*-1, divided by *t*-1 total assets. The measures of dividend-to-market ratio (*DM*), book-to-market equity (*BM*) and market cap (*ME*) are defined as in Table 1.1.

Panel A: Regression of excess returns

	DM	*PM*	*IM*	*BM*	*ME*
(1)	2.109			0.076	-0.063
	(7.40)			(0.88)	(-1.53)
(2)		2.663		0.018	-0.068
		(5.80)		(0.21)	(-1.67)
(3)			-1.278	0.243	-0.046
			(-3.16)	(2.62)	(-1.11)
(4)		2.752	-1.439	0.024	-0.066
		(6.17)	(-3.69)	(0.28)	(-1.65)

Panel B: Regression of characteristic-adjusted returns

	Dependent variable	*PM*	*IM*
(5)	Char-adj returns using reg-4: ME and BM	2.152	-1.494
		(4.17)	(-3.82)
(6)	Char-adj returns using reg-4: ME, BM,OP and INV	1.013	-0.170
		(2.03)	(-0.44)
(7)	Char-adj returns using reg-5: ME and *DM*	0.672	0.936
		(1.29)	(2.40)

Chapter 2
Forecasting profitability shocks and the preference for skewness

2.1. Abstract

The strong stock price reactions to positive earnings surprises are associated with lottery-like payoffs. If investors have a strong skewness preference, as suggested by the cumulative prospect theory, they should be willing to pay more for stocks with a high probability of generating positive earnings surprises, leading to low subsequent returns. I find a negative and significant relation between predicted profitability shocks (PPS) and stock returns. The four-factor alpha on the negative-minus-positive decile portfolio long negative PPS stocks and short positive PPS stocks is 0.925% (t = 5.63) per month. Further, controlling for PPS greatly enhances the predicting power of price momentum for stock returns, which can no longer be fully explained by earnings momentum as suggested by Novy-Marx (2015). In contrast, controlling for PPS largely subsumes the predicting power of the book-to-market ratio and operating profitability for stock returns.

2.2. Introduction

A great deal of finance research documents that the preference for positively skewed payoffs has important pricing implications for financial assets. Following the cumulative prospect theory of Tversky and Kahneman (1992), Barberis and Huang (2007) model that a security with positive skewness can be overpriced and earn a negative average excess return because investors with a skewness preference opt to overweight tail events. In a similar attempt, Brunnermeier, Gollier, and Parker (2007) show investors with optimal expectations choose to be optimistic about the states associated with the most skewed assets, and that their upward biased beliefs cause the overinvestment and low returns of such assets. Empirically, there is rapidly growing evidence supporting the preference for skewness, as researchers employ the theory to explain the low returns on assets with lottery-like payoffs, e.g., IPO stocks (Jay Ritter 1991), over-the-counter stocks (Eraker and Ready 2015), private equity (Moskowitz and Vissing-Jorgensen 2002), financial distressed stocks (Campbell, Hilscher and Szilagyi 2008) and high idiosyncratic volatility stocks (Bali, Cakici and Whitelaw 2011).

A natural laboratory to understand the implications of skewness preference is the swift and substantial stock price adjustments during the days surrounding quarterly earnings announcements. Because earnings numbers convey new information to markets, stock prices react dramatically once firms release earnings surprises (Nichols and Wahlen 2004). It is not uncommon to see stock prices rise/drop more than 10% in a few working days (even a few hours[7]) due to an earnings hit/miss, and such lottery-like payoffs are eye-catching events dominating financial press headlines during the earnings season. The strong reactions to earnings surprises provide strong incentives for investors to use all sources of available

[7] A most recent example is Netflix, whose stock price soared as much as 20% after a few hours on 18 October 2018 as the company posted third-quarter earnings per share of 12 cents compared to 7 cents a share in the year-earlier period.

information to forecast earnings changes, and investors with a strong preference for lottery-like payoffs should be willing to pay more for stocks with a high probability of generating positive earnings surprises[8].

In this paper, I investigate the relation between predicted profitability shocks (PPS) for the coming quarter and stock returns. I focus on the change in operating profitability (OP) instead of bottom-line income, motivated by the finding of Ball, Gerakos, Linnainmaa and Nikolaev (2015) that operating profitability has a stronger link with expected returns than bottom-line income. I define profitability shock as the innovation of a firm's quarterly operating profitability that is uncorrelated with past profitability changes, i.e., the residual from a cross-sectional regression of SUEs (standardized unexpected earnings computed with firm-level quarterly operating profitability) on lagged SUEs. It is well known that the measure of SUE is serially correlated over four lags (Ball and Bartov 1996). Without controlling appropriately for this autocorrelation, the pricing effect of PPS is likely to conflate with the post-earnings announcement drift, i.e. stock prices continue to drift in the direction of an earnings change for several months following an earnings announcement.

My analysis proceeds in two stages: I firstly develop a parsimonious model to predict future profitability shocks, and then investigate how PPS explain the cross-section of stock returns. I combine various accounting and equity market variables in My benchmark model to predict future profitability shocks. There are three variables exhibiting substantial power in predicting the cross-section of future profitability shocks: the level of operating profitability (OP), book-to-market (BM) ratio and past returns. First, Firms with a high level of profitability tend to have negative profitability shocks, which is consistent with Fama and French’s (2000; 2006)

[8] Other timely sources of firm earnings-relevant information, like new product development, major contract wins, management forecasts and peer earnings announcements might attenuate stock price reactions on earnings announcements. However, such earnings-relevant information might also trigger strong price reactions with lottery-like payoffs by themselves, and thus do not undermine the overall price consequences of earnings changes.

finding that profitability is mean reverting in a competitive environment. Second, after controlling for OP, there is a strong negative relation between BM and future profitability shocks. Intuitively, since the market value for a firm is the present value of its future cash flows, a low market value (high BM) must indicate a negative shift of future profitability for a given level of current profitability. My results provide novel evidence for the presumption of Fama and French (1992) that value stocks have poorer prospects than growth stocks. Third, firms with high past returns (winners) are more likely to generate positive profitability shocks than firms with low past returns (losers), which is in line with earlier literature on stock returns having strong predictive power with respect to future earnings changes (Beaver, Lambert and Morst 1980; Kothari and Sloan 1992).

Using predicted profitability shocks (PPS) from out-of-sample predictive tests, I show there is a strong and negative relation between PPS and expected stock returns, as predicted by the cumulative prospect theory. A negative-minus-positive decile portfolio long stocks in the lowest PPS decile and short stocks in the highest PPS decile generates a four-factor alpha of 0.925% per month with a t-statistic of 5.63. In bivariate portfolio analysis, the predicting power of PPS for stock returns is robust after controlling for various cross-sectional effects. In the Fama and MacBeth (1973) regression, where I control for multiple variables simultaneously, the average slope on the PPS is -4.934%, with a t-statistic of -5.03. Of particular interest, I show that the PPS effect is more pronounced among large stocks than among small stocks, which differs substantially from some earlier findings that showed skewness-related anomalies mainly manifest themselves in stocks with small market capitalization (Campbell, Hilscher and Szilagyi 2008; Bali, Cakici and Whitelaw 2011).

After documenting the pricing effect of PPS, I explore the impact of PPS on three prominent anomalies: price momentum, operating profitability and book-to-market ratios. First, I show that controlling for PPS greatly enhances the power of price momentum in predicting stock returns. Most recently, Chordia and Shivakumar (2006) and Novy-Marx (2015) argue that price momentum is merely a manifestation of the earnings momentum and show that post-earnings announcement drift fully subsumes the explanatory power of past returns in the cross-section of stock returns in their studies. I find that controlling for the PPS effect resurrects the predicting power of price momentum in stock returns, which cannot be driven out by post-earnings announcement drift. Since a high past return is a strong leading indicator of positive future profitability shock, a negative-minus-positive portfolio based on PPS is essentially a contrarian strategy long past losers and short past winners. The highly negative correlation (-0.62) between the PPS strategy and the price momentum strategy allows a combination of the two to perform much better than either factor alone. My evidence thus indicates that price momentum has separate explanatory power for stock returns compared to earnings momentum.

Second, I show PPS largely subsumes the explanatory power of operating profitability for stock returns, and thus provides a possible explanation for the pricing effect of operating profitability. It is a well-known and widely accepted economic assumption that profitability is mean reverting in a competitive environment (Fama and French 2000; 2006). My PPS measure captures the tendency that profitable firms are more likely to have negative profitability shocks, and non-profitable firms are more likely to have positive profitability shocks. Driven by the strong preference for lottery-like payoffs associated with profitability shocks, investors might thus overprice stocks of non-profitable firms and underprice stocks of profitable firms, resulting in the pricing effect of profitability. My empirical results show that after controlling for PPS, the predicting power of operating profitability largely disappears, with the exception of stocks in the smallest size quintile that merely account for 3.5% of market capitalization.

Third, I show that PPS fully absorbs the roles of BM in predicting stock return, and thus the value effect may simply be a proxy for PPS. By fixing operating profitability, I observe that BM is a strong predictor of future profitability shock, and that stocks with high (low) BM are more likely to have negative (positive) PPS. The preference for skewness associated with profitability shocks might entice investors to overprice growth stocks with high positive PPS, and underprice value stocks with high negative PPS.

The paper proceeds as follows. Section 2.3 provides a prediction model of profitability shocks. Section 2.4 presents the relation between PPS and average returns. Section 2.5 analyses the impact of PPS on the predicting power of price momentum, operating profitability and BM in stock returns. Section 2.6 compares the performance of PPS with other skewness anomalies. Section 2.7 concludes the paper.

2.3. A prediction model of profitability shocks

In this section, I show how profitability shocks lead to swift price adjustments with lottery-like payoffs. I then develop a parsimonious model in which various accounting and equity market measures are used to predict future profitability shocks at firm-level.

2.3.1. Profitability shocks and lottery-like payoffs

Using a seasonal random walk model with trend, I construct SUE (standardized unexpected earnings) based on firm-level quarterly operating profitability (see Appendix A for details). Ball and Bartov (1996) argue that the term "unexpected" in SUE is misleading because it ignores that firms' SUEs are serially correlated over four lags. Therefore, I define profitability shock as the innovation of firm quarterly operating profitability that is uncorrelated to lagged SUEs. Following the procedure of Ball and Bartov (1996), the estimator for profitability shock for firm i in quarter *t* is the regression residual, $\varepsilon_{i,t}$, from cross-sectional regressions of current SUE on lagged SUEs:

$$SUE_{i,t} = a_0 + \sum_{j=1}^{4} a_j SUE_{i,t-j} + \varepsilon_{i,t} \qquad (1)$$

where $\varepsilon_{i,t}$ is the current profitability shock that is supposed to be orthogonal to lagged SUEs. Following Ball and Bartov (1996), all values of SUEs in the regressions are replaced by their decile rankings and then scaled so that they range from 0 (for the lowest decile) to 1 (for the highest decile).

The impact of profitability shock on stock prices is considerable. Table 2.1 reports distributional statistics on cumulative seven-day abnormal returns (CAR7) surrounding quarterly earnings announcements to decile portfolios sorted by profitability shocks, over the period 1975 to 2015. The CAR7 for each firm is the sum of daily abnormal returns in the 7 days (t = -5 to t = 2) relative to the earnings announcement date (t = 0). Daily abnormal returns are residuals from quarterly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model. The first column of Table 2.1 shows the average CAR7 rises monotonically across profitability shock decile portfolios, as the lowest profitability shock portfolio suffers an average loss of -1.388%, whereas the highest profitability shock decile portfolio yields an average gain of 1.954%. During the short window of 7 days, profitability shocks trigger a 3.342% difference in CAR7 between extreme profitability shock decile portfolios.

The swift and significant price adjustments in the earnings announcement season are associated with lottery-like payoffs — a small chance of very substantial capital gains over a very short trading window. The skew column in Table 2.1 shows a monotonic relation between skewness and profitability shocks; CAR7 is skewed to the left (-0.204) for the lowest profitability shock portfolio and is highly skewed to the right (0.409) for the highest profitability shock portfolio. Figure 2.1 shows intuitively how profitability shocks result in lottery-like payoffs. This figure focuses on large gains/losses in CAR7 that are higher than 15% or lower than -15%, which is equivalent to a $495 million gain/loss for an average-sized firm. The highest (positive) profitability shock decile has a relatively higher probability of enjoying large gains than suffering large losses, i.e. 4.84% versus 1.11%. For the lowest (negative) profitability shock decile, the profitability of suffering large losses is much higher than that of enjoying large gains, i.e. 4.42% versus 1.67%. This result demonstrates that stocks with positive profitability shocks are associated with a higher probability of lottery-like payoffs compared to stocks with negative profitability shocks.

2.3.2. A model to predict profitability shocks

In the prior section, I show that profitability shocks cause swift and significant price adjustments associated with lottery-like payoffs. The stock price consequences of profitability shocks provide enormous incentives for investors to forecast future profitability shocks. I develop a parsimonious model in which various accounting and equity market variables are used to predict future profitability shocks, as follow:

$$\varepsilon_{i,t+1} = \beta + \lambda' A_{i,t} + \mu' M_{i,t} + \omega_{i,t} \tag{2}$$

where $\varepsilon_{i,t+1}$ is profitability shocks for the next earnings season, $A_{i,t}$ is a vector of accounting variables, and $M_{i,t}$ is a vector of market-based variables.

The vector of accounting variables include quarterly operating profitability (OPQ), quarter-end book-to-market ratio (BMQ) and quarterly sale growth (SGQ). The inclusion of operating profitability is motivated by the standard economic argument that profitability is mean reverting in a competitive environment, such that changes in profitability are predictable (Fama and French 2000). The choice of operating profitability to measure an individual firm's earning power is motivated by Ball, Gerakos, Linnainmaa and Nikolaev (2015), who find that operating profitability better explains the cross section of average returns than gross profitability and net income. The use of the book-to-market ratio is motivated by Fama and French (1992), who extrapolate that low prices (high book-to-market ratio) is a signal of poor prospects, leading to high expected stock returns for value stocks as they are penalized with higher costs of capital.

I also include sales growth as a predictor variable since firms enjoying strengthened competitive positions in the market are more likely to achieve strong profits growth.

My first market-based variable is the past monthly excess return over the market index (EXRET), because stock returns have significant predictive power for future earnings changes (Beaver, Lambert and Morst 1980; Kothari and Sloan 1992). The second market-based variable is the maximum daily return (MAX). Bali, Cakici and Whitelaw (2011) show that stocks with extreme positive returns in a given month have a great probability to exhibit extreme positive returns in the following month. The rest of the market-based variables are similar to those employed by Chen, Hong, and Stein (2001) and Boyer, Mitton and Vorkink (2010) in predicting idiosyncratic skewness: idiosyncratic volatility (IVOL) and idiosyncratic skewness (ISKEW) over the past six months t-5 to t, turnover (TURN), a NASDAQ dummy, small and medium-size dummies and industry dummies (49 industries as defined on Ken French's website). The Appendix provides more details on the construction of these variables.

The vast majority of U.S. firms release their quarterly earnings in the earnings season, which is four times a year: from early January to late February, from early April to late May, from early July to late August, and from early October to late November. I align market-based variables measured at the end of each earnings season (i.e., at the end of February, May, August, and November) with accounting variables computed with the latest quarterly accounting data. This alignment ensures all information is available for investors to predict the next profitability shocks at the end of the current earnings season.

Although the latest quarterly accounting data provides a timely source of information, we might also need lagged information to account for seasonality and transitory factors in order to get a clean measure of material economic events. Following the approach of Campbell, Hilscher and

Szilagyi (2008), I impose geometrically declining weights to lagged information for measures of OPQ, BMQ and SGQ, as follows:

$$\mathrm{AVG(X)_t} = \frac{1-\emptyset^3}{1-\emptyset^{12}}(\mathrm{X_t} + \cdots + \emptyset^{11}\mathrm{X_{t-11}}) \quad (3)$$

where X is the variable of interest, and $\emptyset = 2^{-1/3}$ implies that the weight is halved each quarter[9]. Similarly, I also implement this procedure to adjust the market-based measures of EXRET and TURN, where the weight is halved each month. In the exploratory regressions, I find lagged monthly returns and turnover data also provide incremental information in predictive tests.

2.3.3. Results from full-sample predictive tests of profitability shocks

Table 2.2 reports the time-series averages of the slopes from the full-sample predictive tests for profitability shocks in the coming earnings season using quarter-by-quarter Fama-MacBeth (FM) regressions, over the period 1975 to 2015. I exclude very small firms with total assets of less than \$25 million or book equity of less than \$12.5 million. Following the treatment of SUEs in equation (1), all predictor variables in the regressions are replaced by their decile rankings and then scaled so that they range from 0 (for the lowest decile) to 1 (for the highest decile).

Specifications (1) to (5) show that all three accounting variables help to explain the cross-section of predicted profitability shocks. Among these three accounting variables, lagged profitability, AVG(OPQ), has the most impressive predicting power, with an average slope of -0.085 (t = -13.26) in the univariate setting of specification (1). The result that profitable firms tend to have negative profitability shocks is in line with standard economic arguments that profitability is mean reverting in a competitive environment. For the book-to-market ratio,

[9] To avoid losing data, we replace missed lagged observations with their cross-sectional means.

AVG(BMQ), its predicting power does not show up clearly when it is used alone in specification (2). However, after controlling for operating profitability, it becomes highly significant in specification (3), with an average slope of -0.057 (t = -11.63). Therefore, value (high BM) stocks are associated with negative profitability shocks, supporting the conjecture of Fama and French (1992) that value stocks might have poorer prospects than growth stocks. As expected, specification (4) shows that sales growth is associated with positive profitability shocks, as the average slope on AVG(SGQ) 0.020 shows a t-statistic of 5.45. When I control for all three accounting variables in specification (5), operating profitability displays the strongest power in predicting profitability shocks, followed by the book-to-market ratio and sales growth.

Specification (6) confirms that stock return is a leading indicator of future earnings changes. The positive and highly significant slope (0.076, t = 17.73) on AVG(EXRET) shows that positive profitability shocks are more likely to occur among past winners than past losers. Specification (7) indicates a positive and significant relation between maximum daily return and profitability shocks, as the average slope on MAX alone is 0.012 with a t-statistic of 2.95. Specification (8) is similar to the model used by Boyer, Mitton and Vorkink (2010) in predicting idiosyncratic skewness. I observe that idiosyncratic skewness (ISKEW), past returns (AVG(EXRET)) and turnover (AVG(TURN) show significant power in predicting profitability shocks. Thus, stocks with high idiosyncratic skewness and high past returns are associated with positive profitability shocks, whereas stocks with high turnover are related to negative profitability shocks.

Specification (9) is the benchmark model that I focus on in the paper, combining accounting variables with market-based variables to predict profitability shocks. In economic terms indicated by average slopes, operating profitability (-0.145, t = -18.01) has the strongest predicting power for profitability shocks, followed by past returns (0.076, t = 19.01) and book-to-market ratio (-0.053, t = -14.30). Maximum daily return, idiosyncratic volatility and skewness, and sales growth also display considerable roles in predicting profitability shocks, with average slopes more than 5 standard errors from zero. Note that the average slope on IVOL (-0.034, t = -8.49) in specification (9) indicates that a high level of volatility is associated with negative future profitability shocks after accounting for other variables. Moreover, turnover and the NASDAQ dummy also helps predict profitability shocks, but their effects are trivial in economic terms. Lastly, specification (9) produces a relatively low adjusted R-square of 0.111, which reflects the fact that forecasting profitability shock is a difficult task.

2.4. Predicted profitability shocks and average returns

In this section, I examine the asset pricing implications of my profitability shock model based on out-of-sample predictive tests. I first assess how average returns vary across portfolios sorted by predicted profitability shocks using univariate and bivariate portfolio analysis, and then estimate the power of predicted profitability shocks in explaining the cross-section of average returns using Fama and Macbeth (FM) regressions.

2.4.1. Portfolios sorted on predicted profitability shocks

To alleviate the concern of look-ahead bias, predicted profitability shocks are estimated using only historically available data known at the end of earnings season t. At the end of each earnings season (i.e., at the end of every February, May, August, and November) from November of 1984 to November of 2015, the model of specification (9) in Table 2.2 is re-estimated using data on an expanding window, starting from February of 1975 to November of 1984[10]. Predicted profitability shocks are fitted values of the predictive model, i.e., the average

[10] The dependent variable, profitability shock ($\varepsilon_{i,t}$), is also re-estimated by equation (1) at the end of each earnings season t, using only historically available data.

regression slopes multiplied by values of explanatory variable at the end of earnings season t, as follows:

$$E_t(\varepsilon_{i,t+1}) = \beta + \lambda' A_{i,t} + \mu' M_{i,t} \tag{3}$$

We then sort stocks into decile portfolios based on predicted profitability shocks using NYSE breakpoints, and rebalance the portfolios quarterly at the end of each earnings season.

Figure 2.2 graphically summarises the time-series average of realized profitability shocks for each decile portfolio in portfolio holding periods in the sample, from January 1985 to December 2015. Portfolio 1 (negative) includes stocks with the lowest predicted profitability shocks, and portfolio 10 (positive) includes stocks with the highest predicted profitability shocks. I find that the average value of realized profitability shocks rises monotonically with predicted profitability shocks across deciles. The negative decile exhibits an average realized profitability shock of -0.092 while the positive decile exhibits an average realized profitability shock of 0.07, a difference of 0.162 (significant at the $p < 0.001$ level). This result demonstrates that the model, using only historically available data, works well in predicting future profitability shocks.

Table 2.3 presents the value-weighted (VW) average monthly excess returns and alphas relative to the four-factor (Fama-French-Carhart) model for decile portfolios, sorted by predicted profitability using NYSE breakpoints. The four-factor model includes a market factor (MKT), a size factor (SMB), a value factor (HML) and a momentum factor (WML), which are constructed with the same methodology employed by the Fama and French library. I observe that the negative-minus-positive decile portfolio earns a mean excess return of 0.469% per month with a t-statistic of 2.17. Returns on the negative-minus-positive portfolio are even higher after controlling for risk. The four-factor alpha is 0.925% per month with a t-statistic of 5.63, which is economically and statistically significant at all conventional levels. Similarly, the negative-minus-positive quintile portfolio generates a mean excess return of 0.350% per month (t = 1.94) and a four-factor alpha of 0.788% per month (t = 6.07).

Of particular interest, the negative-minus-positive decile portfolio has an economically and statistically large negative exposure to the momentum factor WML, with an average slope of -0.653 and a t-statistic of -17.55. Given that positive profitability shocks are more likely to occur among past winners than past losers, as shown in Table 2.2, a negative-minus-positive portfolio based on predicted profitability shock is essentially a contrarian strategy long on past losers and short on past winners. Figure 2.3 shows the trailing 5-year Sharpe ratios of an NMP (negative-minus-positive predicted profitability shock) decile portfolio and a WML (winner-minus-loser) decile portfolio. It is impressive to see that there is a strong negative correlation (-0.62) between these two strategies. More specifically, a rise of one strategy tends to be closely followed by a fall of the other, and a peak of one strategy typically appears at the same time with a trough of the other. Due to this strong negative correlation and the positive returns on its own, controlling for the WML factor substantially enhances the return spread across portfolios, sorted by predicted profitability shock.

To provide a detailed picture of the composition of each portfolio, I calculate the time-series averages of a variety of stock characteristics for each portfolio and report the results in Table 2.4. From negative to positive deciles ranked by predicted profitability shocks, operating profitability AVG(OPG) falls monotonically, from 0.057 for the negative decile to 0.015 for the positive decile. In contrast, the average values of AVG(BMQ) across deciles are approximately the same. The result on AVG(BMQ) is not a surprise, as I have shown that AVG(BMQ) does not help predict future profitability shock in univariate FM regressions in Table 2.2. I thus construct a measure of residual BMQ by running a regression of BMQ on AVG(OPG). As predicted profitability shocks increase across deciles, the residual BMQ decreases dramatically, which demonstrates that high book-to-market ratios are associated with low future profitability shocks after controlling for operating profitability. Sales growth AVG(SGQ) also exhibits pronounced patterns across predicted decile portfolios, which indicates that stocks with fast sales growth tend to have positive future profitability shocks. For the group of equity-based variables, I find that high past returns (EXRET) and high idiosyncratic skewness (ISKEW) are associated with positive predicted profitability shocks. In contrast, MAX, idiosyncratic volatility (IVOL) and turnover (TURN) display convex patterns across portfolios. The column of Table 2.4 for market capitalization (CAP) shows that small stocks are more likely to be associated with negative profitability shocks relative to other stocks. The last two columns of Table 2.4 present the mean and skewness of realized CAR7 for each portfolio in the holding periods. A portfolio with high predicted profitability shocks tends to yield higher CAR7 than a portfolio with low predicted profitability shocks. Moreover, skewness rises significantly from low decile (-0.003) to high decile (0.392).

2.4.2. Bivariate portfolio analysis

To examine the robustness of the results, Table 2.5 reports the four-factor alphas of the quintile portfolios sorted by predicted profitability shocks after controlling for various cross-sectional effects. These control variables include firm characteristics (size, book-to-market, operating profitability), measures of price movement (price momentum, Max, idiosyncratic volatility, idiosyncratic skewness), measures of earnings momentum (standardized unexpected profitability and cumulative seven-day abnormal returns) and measures of institutional holding and analyst coverage. I first sort stocks based on one control variable using NYSE quintile breakpoints, and then within each quintile portfolio I sort stocks into quintile portfolios based on predicted profitability shocks. For each quintile of the control variable, I report value-weighted four-factor alphas to portfolio 1 (negative), portfolio 5 (positive), negative-minus-positive (1-5) portfolio and their t-statistics. The second column "Ave." refers to the average alphas across five quintile portfolios with different levels of the control variables but with a similar level of predicted profitability shocks.

Controlling for firm characteristics (size, book-to-market, operating profitability)

For portfolio management, it is critical to know whether the pattern in average returns appears reliably for large stocks that account for the vast majority of total market capitalization, or rely mostly on small stocks that are much less liquid. I first examine the relation between stock returns and predicted profitability shocks after controlling for the size effect. Panel A of Table 2.5 shows that the negative PPS quintile outperforms the positive PPS quintile, and the outperformance is most pronounced among the largest stocks. The average alpha, "Ave.", on the negative-minus-positive portfolios across five size quintiles is 0.60% per month with a t-statistic of 5.95, which indicates that market capitalization does not explain the negative relation between predicted profitability shocks and stock returns. The largest quintile has the largest four-factor alpha for the negative-minus-positive portfolio, 0.84% per month (t = 6.03), which is more than three times larger than that of the smallest quintile, 0.24% per month (t = 1.65).

Panel B of Table 2.5 shows the results after controlling for the book-to-market ratio. The four-factor alphas of the negative-minus-positive portfolios between growth (low BM) stocks and value (high BM) stocks are largely the same in magnitude. The average of the alphas of the negative-minus-positive portfolios across five BM quintiles, 0.70% per month (t = 5.80), indicates that stocks with negative predicted profitability shocks significantly outperform stocks with positive predicted profitability shocks after controlling for the value effect.

When I control for operating profitability in Panel C, the effect of predicted profitability shocks on average stock returns persists, with an average alpha of 0.53% per month and a t-statistic of 4.41 for negative-minus-positive portfolios. However, it is worth noting that the effect does not show up among stocks with the lowest level of profitability. In particular, the quintile 1 portfolio with low operating profitability and low predicted profitability shocks has a negative alpha of -0.30 (t = 1.24), as opposed to positive alphas on other quintile 1 portfolios. This underperformance of non-profitable firms with negative profitability outlook is in line with the distress risk puzzle of Campbell, Hilscher and Szilagyi (2008), which states that financially distressed stocks are associated with anomalously low returns.

Controlling for measures of price movement (price momentum, Max, idiosyncratic volatility and idiosyncratic skewness)

In Panel D of Table 2.5, I examine whether price momentum can account for the effect of predicted profitability shocks. The average of the alphas for negative-minus-positive portfolios remains large at 0.73% per month (t = 7.98). Thus, controlling for momentum does not correct the pattern of returns across stocks with different levels of predicted profitability shocks. Moreover, it is a surprise to see that the loser portfolio with low predicted profitability shocks earns an impressively large alpha of 0.97% per month (t = 5.33). This could be the result of investors' excessive pessimism about the prospect of this type of stocks, leading to oversold prices which subsequently rebound. Consequently, the alpha on the negative-minus-positive portfolio is more than three times greater for loser stocks (1.40% per month, t=6.37) than for winner stocks (0.43% per month, t=2.33).

We also control for MAX, idiosyncratic volatility and idiosyncratic skewness in Panels E, F and G, and find similar results whereby the effect of predicted profitability shock is preserved. The average alpha of negative-minus-positive portfolios is economically and statistically significant, at 0.72% (t=5.86) per month, 0.64% (t = 4.88) per month and 0.76% (t = 6.46) per month after controlling for max, idiosyncratic volatility and idiosyncratic skewness, respectively. Moreover, there is relatively little difference in alphas for negative-minus-positive portfolios between low and high controlling quintiles.

Controlling for measures of earnings momentum (standardized unexpected profitability and cumulative seven-day abnormal returns)

Measures of SUE and CAR7 capture the tendency of firms reporting strong earnings to subsequently outperform firms reporting weak earnings. When I control for SUE, negative-minus-positive portfolios have an average alpha of 0.84%, with a robust t-statistic of 7.07. Controlling for CAR7 generates similar results, the average alpha of negative-minus-positive portfolios is still economically and statistically significant. Hence, earnings momentum cannot be an explanation for the effect of predicted profitability shocks on stock returns.

Controlling for measures of institutional holding and analyst coverage

My last set of control variables includes residual institutional holding and residual analyst coverage. Following Campbell, Hilscher and Szilagyi (2008), the residual institutional holding is the residual from a regression of institutional holding (13-F filings) on firm size with time-fixed effects. Similarly, the residual analyst coverage is the residual from a regression of the log of one plus the number of analysts covering each firm (IBES) on firm size with time-fixed effects.

In Panel J, controlling for residual institutional holding, I find that the effect of predicted profitability shocks is much stronger for stocks with low residual institutional holding compared to stocks with high residual institutional holding; the alpha of negative-minus-positive portfolios for stocks with low institutional holding is 0.96% (t=4.03) per month, versus 0.36% (t=1.96) per month for stocks with high institutional holding. The result supports the conclusions of Kumar (2009), who finds that individual investors have a strong preference for stocks with lottery-like payoffs compared to institutional investors. A similar but less pronounced pattern is also found in Panel K when controlling for residual analyst coverage, which shows that stocks with low analyst coverage generate a larger alpha for negative-minus-positive portfolios than stocks with high analyst coverage (1.01% and t=5.25 per month versus 0.69 and t=3.26 per month).

2.4.3. Fama-MacBeth regressions

So far my analysis has been based on portfolio sorts, which makes it difficult to determine which anomaly variable holds unique information on average returns when we need to control for multiple variables simultaneously. This subsection attempts to address this issue by directly estimating the marginal effect of predicted profitability shocks using Fama and Macbeth regressions. Table 2.6 presents the time-series average slopes and their *t*-statistics from FM regressions of stock monthly excess returns on predicted profitability shocks and other control variables. The control variables include the log of market capitalizations (Size), the log of book-to-market ratio (BM), operating profitability (OP), investment (INV), standardized unexpected profitability (SUE), cumulative seven-day abnormal returns (CAR7), monthly excess returns (EXRET), the cumulative returns over months t-12 through t-2 (MOM), the maximum daily return over the past one month (MAX) and idiosyncratic volatility (IVOL) over the past six months. Following Fama and French (2015) and Ball, Gerakos, Linnainmaa and Nikolaev (2015), Size, BM, OP and INV are measured with annual accounting data for the fiscal year t-1, and updated at the end of June for the calendar year t. Independent variables are trimmed at the 1% and 99% levels on a table-by-table basis to ensure different regressions within each table panel are based on the same observations. Except for the full example from January 1985 to December 2015, separate FM regressions for two subsamples using the end of 1999 as the breakpoint are also run.

After a background review on the control variables in specification (1), specification (2) shows that predicted profitability shock (PPS) is economically and statistically significant in explaining average stock returns when considered alone, with an average slope of -3.081% (t = -2.50). The pattern is more pronounced after controlling for other variables in specification (3), where the average slope on predicted profitability shock is -4.685% with a t-statistic of -5.04. The results indicate a negative relation between predicted profitability shock and average stock returns after controlling for multiple variables, which is in line with Barberis and Huang's (2008) finding that positively skewed securities can become overpriced and earn negative average returns. Results from the subsample analysis reveal that the predicting power of predicted profitability shocks persists over time, as the average slope on PPS for subsamples remains highly significant: -4.423 (t = -3.77) for the first subsample and -4.931 (t = -3.45) for the second subsample.

2.5. Momentum, value and operating profitability effects

The FM regressions in Table 2.6 provides interesting insight into the relation between predicted profitability shock and momentum, value and operating profitability effects. It appears that adding predicted profitability shocks to regressions drives out the significance of BM and OP, but drives up the significance of MOM. Specification (1) shows that when I omit the control for PPS, the average slopes of BM and OP are statistically significant, at 0.124% per month (t = 2.02) and 1.925% per month (t = 4.92), respectively. However, after controlling for PPS in specification (3), the average slopes of BM and OP shrink dramatically and become statistically insignificant at 0.074% per month (t = 1.22) and 0.636% per month (t = 1.42), respectively. This result suggests both BM and OP could be noisy proxies for predicted profitability shocks in predicting average returns. Meanwhile, adding PPS to the regression causes the average slope of MOM to improve from 0.031% per month in specification (1) to 0.222% per month in specification (3). This indicates that controlling for predicted profitability shocks could enhance the effect of price momentum. Note that the statistically insignificant average slope of MOM could be caused by the non-linear relation between average stock returns and MOM, where return difference is apparent only for stocks with extreme past returns. Such a non-linear relation might weaken the return effect of price momentum in multivariate FM regressions, which imposes a functional form on the relation between explanatory variables and returns.

We now turn to time-series spanning tests based on mimicking factors, which alleviate the concern of non-linear relations between anomaly variables and average returns. Specifically, I regress the BM, OP and MOM factors on explanatory variables including the PPS factor, and regress the PPS factor on explanatory variables including the BM, OP and MOM factors. A significant abnormal intercept would suggest the test factor contributes to the description of average returns provided by the explanatory variables.

The mimicking factors are constructed using independent 2×3 sorts as described by the Fama and French library. Taking the predicted profitability shock factor as an example, the 2×3 sorts uses the median NYSE market cap to split NYSE, Amex, and NASDAQ stocks into two groups, and also independently uses the 30^{th} and 70^{th} NYSE percentiles of PPS values to break stocks into three PPS groups. Thus six value-weight intersection portfolios are constructed quarterly at the end of each earnings season, and the NMP (negative minus positive predicted profitability shock) is the average return on two high PPS portfolios minus the average return on two low PPS portfolios. Similarly, I also construct the operating profitability factor (RMW, robust minus weak profitability), the investment factor (CMA, conservative minus aggressive investment), and the SUE factor (positive minus negative SUE), based on measures of OP, INV and SUE for the second sort and rebalance once the respective measure is updated.

Panel A shows the results of the spanning tests on WML. WML earns a significant average return of 0.548% (t = 2.35) per month in specification (1). This large return is fully subsumed by other variables in specification (2), as we see WML yields an insignificant alpha of -125% (t = -0.65) per month, resulting from its heavy loading on the SUE factor (1.643%, t = -14.64). However, after controlling for NMP in specification (3), WML retains its economically and statistical significance with an alpha of 0.466% (t = 3.23). The negative loading of WML on NMP (-1.120%, t = -17.99) suggests that controlling for NMP significantly contributes to the performance of WML. The result that price momentum cannot be fully explained by SUE lends new evidence to the conclusion of Jegadeesh and Lakonishok (1996) that price momentum has distinct marginal explanatory power on average returns relative to earnings momentum. This result runs contrary to the findings of Chordia and Shivakumar (2006) and Novy-Marx (2015), i.e., that the price momentum anomaly is merely a weak expression of the earnings momentum — a conclusion which was drawn without controlling for predicted profitability shocks.

Panel B of Table 2.7 shows that although HML does not earn a significant average return in specification (1), it earns a significant alpha (0.255% per month, t = 2.72) relative to explanatory variables without controlling for NMP in specification (2). After controlling for NMP in specification (3), the alpha shrinks to 0.126% per month and becomes statistically insignificant (t = 1.32), resulting from a substantial loading on NMP (0.243% per month, t = 4.44). This result confirms the prior finding using the FM regressions that predicted profitability shock helps explain the BM effect in average returns.

The results of the spanning tests on RMW are presented in Panel C. The operating profitability factor earns an average return of 0.277% (t = 2.41) per month in specification (1), and a highly reliable alpha of 0.391% (t = 4.33) per month relative to explanatory variables without controls of NMP in specification (2). In specification (3), resulting from its heavy loading on NMP (0.406% per month, t = 8.69), RMW fails to generate a significant alpha relative to explanatory variables (alpha=0.131% per month, t = 1.49). This result reiterates that predicted profitability shock subsumes operating profitability in predicting average stock returns.

Panel D reports the results of the spanning tests on NMP, the predicted profitability factor. In specification (1), the average return of NMP is not impressive at 0.206% (t = 1.50) per month. However, after controlling for other factors in specification (2) and specification (3), NMP earns a highly significant alpha of 0.528% (t = 4.48) and 0.475% (t = 5.53) per month, respectively. NMP loads heavily on WML and the SUE factor, suggesting the price momentum and SUE factors greatly enhance the performance of the predicted profitability shock factor. The time series correlation between NMP and WML is -74%, while for NMP and SUE, the correlation is -57%. The greatly enhanced performance of NMP is a direct result of integrating negatively correlated assets with positive returns with each other.

However, spanning tests on mimicking portfolios do not indicate whether the result is pervasive across size groups or is merely dominated by a particular size group. To address this concern, Table 2.8 provides results of the same spanning tests within each size quintile. Specifically, I report the alphas of value-weighted quintile portfolios sorted by BM, OP, MOM and predicted profitability shock portfolios within size quintile using NYSE breakpoints. The set of explanatory variables for specifications (2) and (3) are defined in Table 2.7. For example, for portfolios sorted by BM, the alphas of specification (2) are intercepts from the spanning tests of quintile portfolios' excess returns on the set of explanatory variables used in specification (2) in panel A of Table 2.7, including RMW, WML, MKT, SMB, CMA and the SUE factors.

For quintiles sorted by price momentum, alphas on the winner-minus-loser portfolios in specification (2) show there is a momentum effect for the smallest size group, but there is also a reversal effect for the largest size group. In specification (3), controlling for NMP brings up the alpha of the winner-minus-loser portfolios substantially. For the two-smallest size quintiles, the momentum premium shows up positively and significantly. For the three-largest size quintiles, alphas of the winner-minus-loser portfolios turn into positive numbers, though they are statistically insignificant. Thus, the price momentum effect is greatly improved with the help of the NMP control, particularly among small and medium stocks.

For quintiles sorted by BM, results on specification (2) show that the value effect is concentrated on size quintiles 2 to 4. After controlling for NMP in specification (3), we do not find significant value effect for all size groups, as all value-minus-growth portfolios fail to generate significant abnormal returns. For quintiles sorted by OP, the effect of operating profitability is initially pervasive across all size groups, as shown by alphas of specification (2). From alphas of high-minus-low portfolios in specification (3), I see an additional control of NMP in the spanning tests removes the significance of the OP effect for all size groups, except for the smallest size group which merely accounts for 3.5% of market capitalization.

Lastly, for quintiles sorted by predicted profitability shocks, the alphas of specification (2), without controlling for WML, and specification (3), controlling for WML, both show that the effect of predicted profitability shocks is pervasive across size groups. Note also that the negative-minus-positive portfolios earn substantially larger premiums for the largest quintile than for the smallest quintile, namely, 0.68% ($t = 4.22$) versus 0.31% ($t = 1.97$) per month in specification (2) and 0.62% ($t = 4.56$) versus 0.28% ($t = 1.86$) per month in specification (3).

In summary, cross-sectional regressions and spanning tests provide strong corroborating evidence that controlling for predicted profitability shocks greatly improves the predicting power of price momentum in average returns, whereas it largely subsumes the effect of BM and OP size effects in average returns. The impact of predicted profitability shocks on the predicting power of BM, OP and price momentum are largely robust across all size groups.

2.6. MAX and expected idiosyncratic skewness

My final empirical exercise is to investigate how predicted profitability shocks relate to other well-known measures of idiosyncratic skewness in explaining cross-sectional variation in average stock returns. Specifically, I assess whether the pricing effects of expected idiosyncratic skewness (Boyer, Mitton, and Vorkink 2010) and maximum daily return (Bali, Cakici, and Whitelaw 2011) can be captured by predicted profitability shocks, or the other way around.

Table 2.9 reports the results of the spanning tests based on mimicking portfolios constructed by independent 2×3 sorts. The MAX factor is a strategy of long stocks with the lowest maximum daily return and short stocks with the highest maximum daily return, and the E(IKEW) factor is a strategy of long stocks with the lowest expected idiosyncratic skewness and short stocks with the highest expected idiosyncratic skewness. My estimates of expected idiosyncratic skewness are based on the same methodology that I use to estimate predicted profitability shocks in prior sections, whereby the dependent variable is the idiosyncratic volatility over the past three months. Accordingly, the control variable of past idiosyncratic volatility is also measured over a period of three months.

Panel A of Table 2.9 shows that predicted profitability shocks substantially attenuates the pricing effect of MAX. Before controlling for NMP, the MAX factor generates an average abnormal return of 0.479% (t = 3.29) per month in specification (2). An inclusion of NMP in the control panel shrinks the abnormal return by a third to 0.311% (t = 2.09), resulting from its heavy loading on NMP (0.353%, t = 4.06). Similarly, I also find that predicted profitability shocks play a critical role in explaining the relation between expected idiosyncratic skewness and average returns. The average abnormal return of the E(IKEW) factor before and after controlling for NMP is 0.461% (t = 3.11) versus 0.368% (t = 2.39). In contrast, results in Panel

C suggest the pricing effect of predicted profitability shock is essentially unaffected by additional controls for MAX and E(ISKEW). For specification (2) that does not control for the MAX and E(ISKEW) factors, NMP is associated with an abnormal return of 0.475% (t = 5.53). The abnormal return is only slightly reduced to 0.441% (t = 5.14) after additional controls for both the MAX and E(ISKEW) factors in specification (3).

Table 2.10 provides further details on the same spanning tests within each size quintile. For quintiles sorted by MAX, alphas of specification (2) show that the pricing effect of MAX is most pronounced among small stocks, and is decreasing monotonically across size quintiles. After controlling for NMP in specification (3), the MAX factor only exhibits significant abnormal returns for quintile 1 and 2, which indicates the pricing effect of MAX is limited to small stocks after controlling for predicted profitability shocks.

For quintiles sorted by expected idiosyncratic skewness, the alphas of specification (2) show that the effect of E(ISKEW) is concentrated on medium stocks (quintiles 2 to 4). The average abnormal returns to the low-minus-high strategy on quintiles 2 to 4 is 0.54% (t = 2.73), 0.58% (t = 3.03) and 0.51% (t = 2.64), respectively. The additional control of NMP brings down the average abnormal returns on quintiles 2 to 4 to 0.37% (t = 1.83), 0.55% (t = 2.77) and 0.41% (t = 2.03), respectively. Hence, it is medium stocks that drive the effect of expected idiosyncratic skewness, which constitutes about 24.5% of market capitalization.

For quintiles sorted by predicted profitability shocks, the alphas on the negative-minus-positive portfolios are highly significant for medium and large stocks, and is largely the same with and without controlling for the MAX and E(ISKEW) factors. For quintiles 2 to 5 of specification (2), stocks with low predicted profitability shocks outperform stocks with high predicted profitability shocks by 0.65% (t = 4.48), 0.52% (t = 3.58) , 0.56% (t = 3.91) and 0.62% (t =

4.56) per month, respectively, while for specification (3), the results were 0.62% (t = 4.21), 0.48% (t = 3.24) , 0.56% (t = 3.84) and 0.62% (t = 4.46), respectively.

In short, these results in the spanning tests show that predicted profitability shocks play a substantial role in explaining the pricing effects of MAX and expected idiosyncratic skewness, but not vice versa. Moreover, each predicted profitability shock, MAX and expected idiosyncratic skewness has unique predicting power in average returns. The pricing effect of predicted profitability shocks is also more pronounced among large and medium stocks, whereas MAX is more specific to small stocks, and expected idiosyncratic skewness is concentred on medium stocks. Thus, from a practical perspective on investable assets, the pricing effect of predicted profitability shocks appears to have the upper hand over that of MAX and expected idiosyncratic skewness.

2.7. Conclusion

The strong stock price reaction to earnings surprises provides strong incentives for investors to use all sources of available information to forecast earnings changes and allocate assets accordingly. Following the cumulative prospect theory, investors with a strong preference for lottery-like payoffs should be willing to pay more for stocks with a high probability of generating positive earnings surprises, causing overinvestment and low returns. This paper documents that there is a negative and significant relation between predicted profitability shocks (PPS) and stock returns, such that stocks with low PPS outperform stocks with high PPS. To do this, I firstly developed a parsimonious model to predict future profitability shocks, in which operating profitability, book-to-market ratio and past returns displayed significant power in predicting the cross-section of future profitability shocks. I then investigated how predicted profitability shocks explain the cross-section of stock returns, using portfolio-level analyses and firm-level FM regressions. I find that a negative-minus-positive decile portfolio long negative PPS stocks and short positive PPS stocks generate a four-factor alpha of 0.925% (t = 5.63) per month. My results also show that the pricing effect of PPS is more pronounced among large stocks than among small stocks.

After documenting the pricing effect of PPS, I explored the impact of PPS on three prominent anomalies, including price momentum, operating profitability and the book-to-market ratio. I find that controlling for PPS greatly enhances the predicting power of price momentum for stock returns, which can no longer be fully explained by earnings momentum. In contrast, controlling for PPS largely subsumes the predicting power of the book-to-market ratio and operating profitability for stock returns. Therefore, investors with a strong preference for lottery-like payoffs may simply use the book-to-market ratio and operating profitability as noisy proxies for PPS.

2.8. Appendix

AVG(.): Variables coined with AVG(.) are adjusted with lagged information by imposing geometrically declining weights, as follows:

$$AVG(X)_t = \frac{1-\emptyset^3}{1-\emptyset^{12}}(X_t + \cdots + \emptyset^{11}X_{t-11})$$

where X is the variable of interest, and Ø=2^(-1/3) implies that the weight is halved each quarter for quarterly observations or each month for monthly observations.

Book-to-market ratio (BM): BM is measured at the end of June of year t is the ratio of book equity for the last fiscal year ending in t-1 divided by the market cap in December of t-1.

Quarterly book-to-market ratio (BMQ): BMQ is measured at the end of earnings season t is the ratio of book equity from the quarter-end accounting data released on earnings season t, divided by market cap measured at the end of the earnings release month.

Cumulative seven-day abnormal returns (CAR7): CAR7 for each firm is the sum of daily abnormal returns in the 7 days (t = -5 to t = 2) relative to the earnings announcement date (t = 0). Daily abnormal returns are residuals from monthly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model.

Market capitalization (CAP): CAP is the market capitalization ($\$10^9$) measured at the end of the earnings release month.

Monthly excess return over market index (EXRET): EXRET is the monthly excess return for individual stocks relative to the value-weighted market index.

Industry dummies: firms are allocated into 49 industries as defined on Ken French's website.

Investment (INV): INV measured at the end of June of year t is the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-1 total assets.

Idiosyncratic volatility (IVOL): IVOL is the standard deviation of daily residuals over the six months finishing at the end of earnings season t, where daily residuals are from the monthly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model.

Idiosyncratic skewness (ISKEW): ISKEW is the skewness of daily residuals over the six months finishing at the end of earnings season t, where daily residuals are from monthly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model.

Maximum daily return (MAX): Max is the maximum daily return for the last month of earnings season t.

Momentum (MOM): MOM is the cumulative return over months t-12 through t-2.

Operating profitability (OP): Following Ball, Gerakos, Linnainmaa and Nikolaev (2015), OP is measured at the end of June of year t, based on annual accounting data for the fiscal year ending in year t-1, and is revenue less cost of goods sold less selling, general and administrative expenses (excluding expenditure on research & development), divided by total assets.

Quarterly operating profitability (OPQ): OPQ is measured at the end of earnings season t, based on quarterly accounting data released on earnings season t, with the same equation as OP.

Predicted profitability shock (PPS): PPS is estimated using the model of specification (9) in Table 2.2, based on historically available data known at the end of earnings season t.

Profitability shock: profitability shock is the innovation of firm quarterly operating profitability that is uncorrelated to lagged SUEs. The estimator for profitability shock for firm i in quarter t is the regression residual, $\varepsilon_{i,t}$, from cross-sectional regressions of current SUE on lagged SUEs:

$$\mathrm{SUE}_{i,t} = a_0 + \sum_{j=1}^{4} a_j \mathrm{SUE}_{i,t-j} + \varepsilon_{i,t}$$

Size: size is the log of market capitalization measured at the end of June of year t.

Size dummy: firms are allocated into small, medium and large groups using 20^{th} and 50^{th} percentiles of NYSE market cap.

Quarterly sale growth (SGQ): SGQ is measured at the end of earnings season t and is the change of quarterly revenue from the quarter released on earnings season t-1 to the quarter released on earnings season t, divided by t-1 quarterly revenue.

Standardized unexpected earnings (SUE): SUE is calculated with a seasonal random walk model with trend (Jegadeesh and Livnat 2006) based on firms' quarterly operating profitability (OPQ):

$$\mathrm{SUE}_{i,t} = \frac{OPQ_{i,t} - OPQ_{i,t-4} - D_{i,t}}{\sigma_{i,t}}$$

where the drift, $D_{i,t}$, and the standard error, $\sigma_{i,t}$, are estimated as follow:

$$D_{i,t} = \frac{\sum_{j=1}^{8} (OPQ_{i,t} - OPQ_{i,t-4})}{8}$$

and

$$\sigma_{i,t} = \frac{1}{7}\sqrt{\sum_{j=1}^{8} \left(OPQ_{i,t-j} - OPQ_{i,t-j-4} - D_{i,t}\right)^2}$$

Turnover (TURN): TURN is the sum of trading volumes during that month divided by the number of shares outstanding.

2.9. References

Bali, T. G., N. Cakici, and R. F. Whitelaw, 2011, Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns, *Journal of Financial Economics* 99, 427-446.

Ball, R., and E. Bartov, 1996, How naive is the stock market's use of earnings information?, *Journal of Accounting and Economics* 21, 319-337.

Ball, R., and P. Brown, 1968, An Empirical Evaluation of Accounting Income Numbers, *Journal of Accounting Research* 6, 159-178.

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. V. Nikolaev, 2015, Deflating Profitability, *Journal of Financial Economics* 117, 225-248.

Barberis, N., and M. Huang, 2008, Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, *American Economic Review* 98, 2066-2100.

Beaver, W., R. Lambert, and D. Morse, 1980, The information content of security prices, *Journal of Accounting and Economics* 2, 3-28.

Bernard, V. L., and J. K. Thomas, 1989, Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?, *Journal of Accounting Research* 27, 1-36.

Brunnermeier, M. K., C. Gollier, and J. A. Parker, 2007, Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns, *American Economic Review* 97, 159-165.

Campbell, J. Y., J. Hilscher, and J. Szilagyi, 2008, In Search of Distress Risk, *Journal of Finance* 63, 2899-2939.

Chordia, T., and L. Shivakumar, 2006, Earnings and price momentum, *Journal of Financial Economics* 80, 627-656.

Eraker, B., and M. Ready, 2015, Do Investors Overpay for Stocks with Lottery-Like Payoffs? An Examination of the Returns of OTC Stocks, *Journal of Financial Economics* 115, 486-504.

Fama, E. F., and K. R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427-465.

———, 2000, Forecasting Profitability and Earnings, *Journal of Business* 73, 161.

———, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491-518.

Jegadeesh, N., and J. Livnat, 2006, Revenue surprises and stock returns, *Journal of Accounting and Economics* 41, 147-171.

Kothari, S. P., and R. G. Sloan, 1992, Information in prices about future earnings, *Journal of Accounting and Economics* 15, 143-171.

Kumar, A., 2009, Who Gambles in the Stock Market?, *Journal of Finance* 64, 1889-1933.

Livnat, J., and R. R. Mendenhall, 2006, Comparing the Post--Earnings Announcement Drift for Surprises Calculated from Analyst and Time Series Forecasts, *Journal of Accounting Research* 44, 177-205.

Moskowitz, T. J., and A. Vissing-Jorgensen, 2002, The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?, *American Economic Review* 92, 745-778.

Nichols, D. C., and J. M. Wahlen, 2004, How Do Earnings Numbers Relate to Stock Returns? A Review of Classic Accounting Research with Updated Evidence, *Accounting Horizons* 18, 263-286.

Novy-Marx, R., 2015, *Fundamentally, Momentum is Fundamental Momentum* Working paper).

Ritter, J. R., 1991, The Long-run Performance of Initial Public Offerings, *Journal of Finance* 46, 3-27.

Tversky, A., and D. Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty* 5, 297-323.

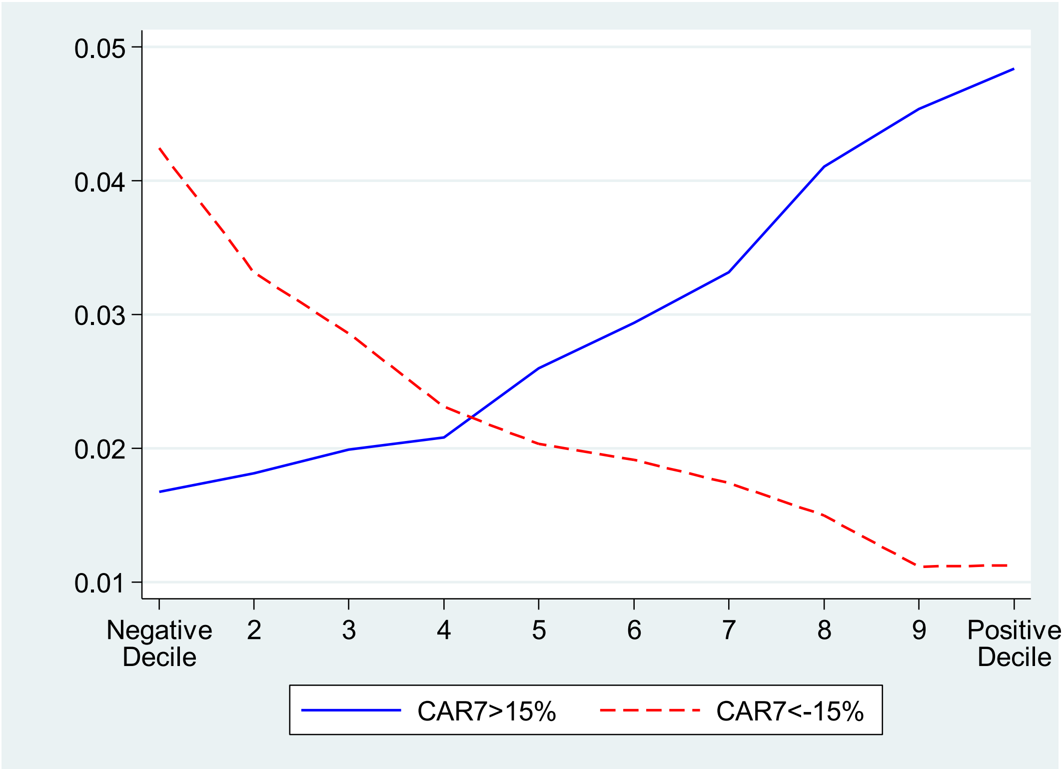

Figure 2.1 Probability of large gain or loss in CAR7

This plot shows the probability of a large gain or loss in CAR7 for individual stocks in decile portfolios sorted by profitability shocks. Stocks are allocated to decile portfolios each quarter based on profitability shocks. Large gains/losses are defined as CAR7 higher than 15% or lower than -15%. CAR7 for each firm is the sum of daily abnormal returns in the 7 days (t = -5 to t = 2) relative to the earnings announcement date (t = 0). Daily abnormal returns are residuals from quarterly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model.

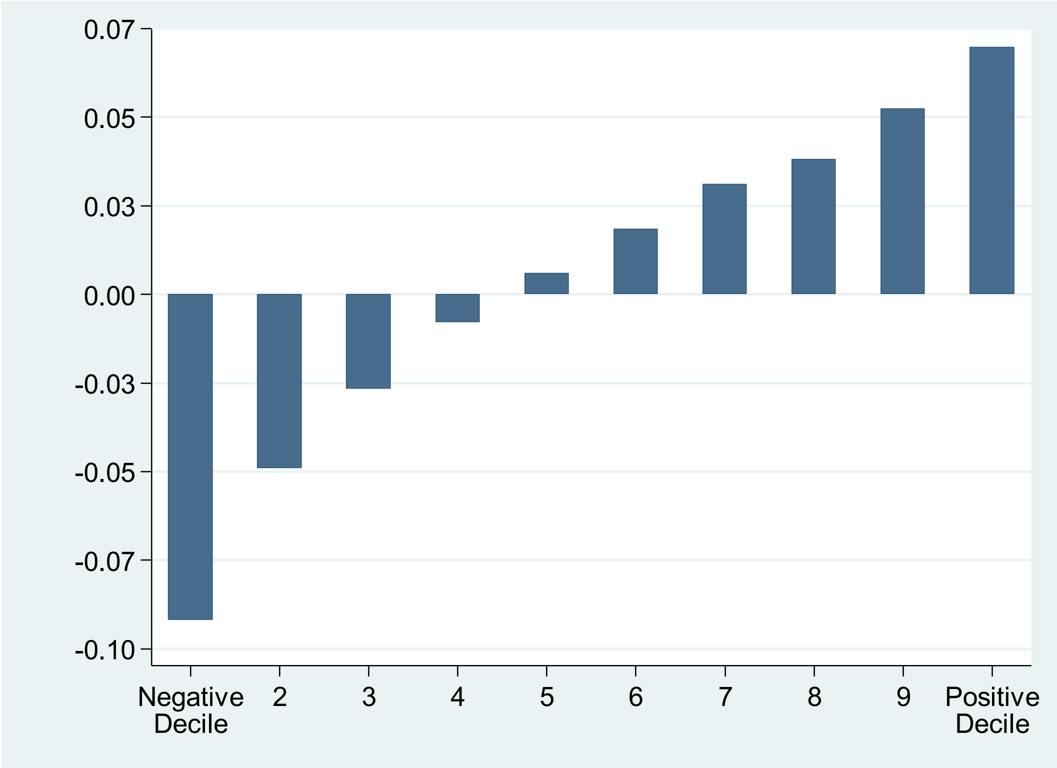

Figure 2.2 Realized profitability shocks

The plot shows the time-series average of holding-period realized profitability shocks for individual stocks in each decile portfolio sorted by predicted profitability shocks. The negative decile includes stocks with the lowest predicted profitability shocks, and the positive decile includes stocks with the highest predicted profitability shocks.

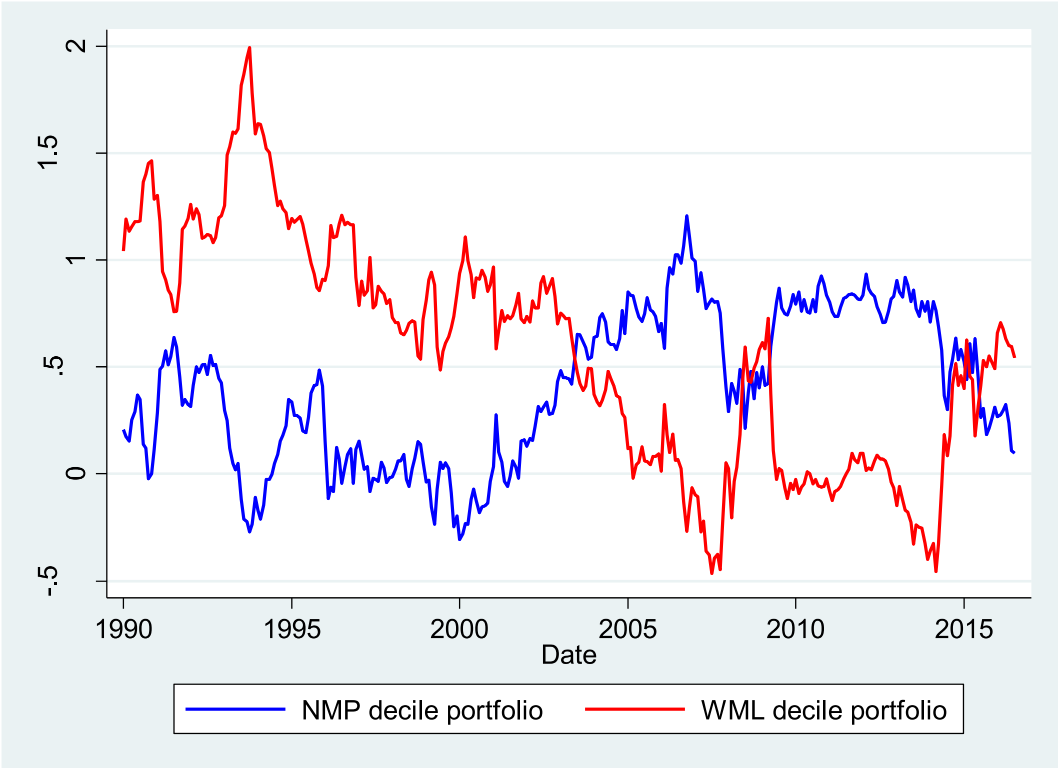

Figure 2.3 Trailing 5-year Sharpe ratios

The plot shows the trailing 5-year Sharpe ratios of an NMP (negative-minus-positive predicted profitability shock) decile portfolio and a WML (winner-minus-loser) decile portfolio over the sample period of January 1985 to December 2015.

Table 2.1 Distribution of cumulative seven-day abnormal returns surrounding earnings announcements to portfolios sorted by profitability shocks

This table reports the distribution of cumulative seven-day abnormal returns (CAR7) surrounding earnings announcements to decile portfolios sorted by profitability shocks, over the period 1975 to 2015. The CAR7 for each firm is the sum of daily abnormal returns in the 7 days (t = -5 to t = 2) relative to the earnings announcement date (t = 0). Daily abnormal returns are residuals from quarterly time-series regressions of individual stocks' daily returns on the Carhart-four-factor model. Profitability shock is defined as the innovation of a firm's quarterly operating profitability that is uncorrelated to lagged SUEs, as the error term in equation (1). I exclude very small firms with total assets of less than $25 million or book equity of less than $12.5 million. The sample has 279,228 firm-quarter observations on firms from the CRSP and Compustat quarterly databases, with accounting and market information necessary to calculate the explanatory variables for in-sample predictive FM tests for profitability shocks.

Decile	Mean	Median	Std.	Skew	p1	p5	p10	p25	p50	p75	p90	p95	p99
Low	-1.388	-1.040	7.391	-0.204	-22.966	-13.878	-9.975	-4.953	-1.040	2.386	6.651	10.060	18.781
2	-1.069	-0.921	7.251	-0.014	-21.674	-13.070	-9.482	-4.690	-0.921	2.644	6.919	10.381	19.263
3	-0.631	-0.557	7.102	0.046	-20.267	-12.356	-8.748	-4.236	-0.557	2.984	7.247	10.807	19.551
4	-0.265	-0.271	6.978	0.017	-19.786	-11.494	-8.164	-3.810	-0.271	3.224	7.657	11.095	19.493
5	0.150	0.030	6.983	0.177	-19.163	-10.971	-7.611	-3.399	0.030	3.551	8.007	11.619	20.648
6	0.504	0.282	7.135	0.239	-18.841	-10.674	-7.441	-3.185	0.282	3.970	8.697	12.403	21.460
7	0.943	0.637	7.154	0.251	-18.302	-10.251	-6.921	-2.699	0.637	4.443	9.208	12.993	21.759
8	1.306	0.888	7.293	0.286	-17.730	-9.904	-6.634	-2.525	0.888	4.837	9.914	13.909	23.036
9	1.797	1.269	7.214	0.387	-16.423	-9.094	-6.025	-2.132	1.269	5.296	10.588	14.512	23.095
High	1.954	1.337	7.226	0.409	-16.516	-8.719	-5.616	-1.985	1.337	5.453	10.643	14.646	23.674

Table 2.2 Firm-level full-sample predictors of profitability shocks

The table shows the average slopes and their *t*-statistics (in parentheses) from full-sample predictive tests for profitability shocks using quarter-by-quarter Fama-MacBeth (FM) regressions, over the period 1975 to 2015. I exclude very small firms with total assets of less than $25 million or book equity of less than $12.5 million. The sample has 279,228 firm-quarter observations on firms from the CRSP and Compustat quarterly databases. with accounting and market information necessary to calculate explanatory variables. Predictor variables include quarterly operating profitability (OPQ), quarter-end book-to-market ratio (BMQ), quarterly sales growth (SGQ), past monthly excess returns over the market index (EXRET), maximum daily returns over the past one month (MAX), idiosyncratic volatility (IVOL) and idiosyncratic skewness (ISKEW) over the past six months, and average daily turnover over the past month (TURN). Variables coined with AVG(.) are adjusted with lagged information by imposing geometrically declining weights, as defined in equation (3). All predictor variables in regressions are replaced by their decile rankings and then scaled so that they range from 0 (for the lowest decile) to 1 (for the highest decile).

Model	AVG(OPQ)	AVG(BMQ)	AVG(SGQ)	AVG(EXRET)	MAX	IVOL	ISKEW	AVG(TURN)	NASDAQ dummy	Small dummies	Medium dummy	Industry dummies	A(R^2)
(1)	-0.085											No	0.016
	(-13.26)												
(2)		0.000										No	0.003
		(0.07)											
(3)	-0.119	-0.057										No	0.022
	(-16.36)	(-11.63)											
(4)			0.020									No	0.004
			(5.45)										
(5)	-0.121	-0.053	0.023									No	0.025
	(-16.81)	(-12.64)	(6.64)										
(6)				0.076								No	0.012
				(17.73)									
(7)					0.012							No	0.003
					(2.95)								
(8)				0.066		0.000	0.025	-0.008	-0.001	-0.000	-0.003	yes	0.088
				(15.31)		(0.04)	(11.41)	(-3.88)	(-0.51)	(-0.18)	(-1.85)		
(9)	-0.145	-0.053	0.014	0.076	0.021	-0.034	0.016	-0.007	0.004	-0.004	-0.003	yes	0.111
	(-18.01)	(-14.30)	(5.16)	(19.01)	(7.05)	(-8.49)	(7.56)	(-3.32)	(2.71)	(-1.48)	(-1.67)		

Table 2.3 Excess return and alpha of portfolios sorted by out-of-sample predicted profitability shocks

The table shows the value-weighted (VW) average monthly excess returns and alphas relative to the four-factor Fama-French-Carhart model for decile portfolios sorted by predicted profitability using NYSE breakpoints. The sample starts in January 1985 and ends in December 2015.

Decile	Mean excess return	Four-factor model					
		a	β_{MKT}	β_{SMB}	β_{HML}	β_{WML}	$A(R^2)$
Negative	0.948	0.511	1.022	0.145	-0.154	-0.475	0.855
2	0.861	0.437	0.959	-0.021	-0.130	-0.388	0.850
3	0.767	0.315	0.899	-0.106	-0.065	-0.262	0.839
4	0.733	0.194	0.885	-0.101	0.032	-0.110	0.811
5	0.793	0.154	0.946	-0.099	0.058	-0.013	0.866
6	0.670	0.003	0.948	-0.070	0.076	0.026	0.856
7	0.581	-0.126	0.978	0.003	-0.031	0.073	0.871
8	0.834	0.065	1.028	0.010	-0.086	0.136	0.885
9	0.585	-0.244	1.083	0.048	0.072	0.128	0.900
Positive	0.478	-0.414	1.134	-0.031	0.145	0.178	0.897
Neg. - Pos.	0.469	0.925	-0.112	0.176	-0.299	-0.653	0.463
(deciles)	(2.17)	(5.63)	(-2.95)	(3.28)	(-4.81)	(-17.55)	
Neg. - Pos.	0.350	0.788	-0.130	0.0606	-0.279	-0.579	0.515
(quintiles)	(1.94)	(6.07)	(-4.31)	(1.43)	(-5.68)	(-19.71)	

Table 2.4 Summary statistics of portfolios sorted by out-of-sample predicted profitability shocks

The table shows the time-series averages of a variety of stock characteristics for stocks in decile portfolios sorted by predicted profitability using NYSE breakpoints. Reported characteristics include quarterly operating profitability (OPQ), quarter-end book-to-market ratio (BMQ), residual BMQ (the residuals from a regression of BMQ on AVG(OPG)), quarterly sales growth (SGQ), maximum daily returns over the past one month (MAX), idiosyncratic volatility (IVOL) and idiosyncratic skewness (ISKEW) over the past six months, past monthly excess returns over the market index (EXRET), turnover (TURN), market capitalization (CAP), mean and skewness of realized CAR7 in the portfolio holding period. Variables coined with AVG(.) are adjusted with lagged information by imposing geometrically declining weights, as defined in equation (3). The sample starts in January 1985 and ends in December 2015.

Decile	AVG(OPQ)	AVG(BMQ)	Residual BMQ	AVG(SGQ)	AVG(EXRET)	MAX	IVOL	ISKEW	AVG(TURN)	CAP($\$10^9$)	Realized CAR7 Mean	Realized CAR7 Skew
Negative	0.057	-0.593	0.331	0.037	-0.037	6.401	5.432	-1.171	0.064	1.592	0.097	-0.003
2	0.048	-0.609	0.138	0.041	-0.022	6.067	4.599	0.344	0.019	3.066	0.249	0.053
3	0.044	-0.600	0.048	0.044	-0.012	5.774	4.225	0.857	0.006	3.447	0.363	0.126
4	0.041	-0.586	-0.010	0.046	-0.005	5.614	4.057	1.294	0.002	3.969	0.262	0.035
5	0.039	-0.602	-0.080	0.050	0.002	5.552	3.902	1.593	0.000	4.054	0.365	0.155
6	0.037	-0.620	-0.154	0.054	0.008	5.531	3.862	1.809	0.004	4.515	0.403	0.086
7	0.034	-0.626	-0.217	0.059	0.012	5.624	3.834	2.040	0.003	4.627	0.414	0.149
8	0.030	-0.607	-0.275	0.063	0.016	5.662	3.858	2.277	0.000	4.407	0.484	0.235
9	0.025	-0.604	-0.359	0.072	0.020	5.854	4.021	2.555	0.001	4.130	0.458	0.253
Positive	0.015	-0.776	-0.729	0.092	0.030	6.600	4.809	3.693	0.033	3.742	0.510	0.392

Table 2.5 Alphas of portfolios sorted by predicted profitability shocks after controlling for various effects

This table reports the four-factor alphas to quintile portfolios sorted by predicted profitability shocks after controlling for various cross-sectional effects. These control variables include firm characteristics (size, book-to-market, operating profitability), measures of price movement (price momentum, Max, idiosyncratic volatility, idiosyncratic skewness), measures of earnings momentum (standardized unexpected profitability and cumulative seven-day abnormal returns) and measures of institutional holding and analyst coverage. I first sort stocks based on one control variable using NYSE quintile breakpoints, and then within each quintile portfolio I sort stocks into quintile portfolios based on predicted profitability shocks. For each quintile of the control variable, I report value-weighted four-factor alphas of portfolio 1 (negative), portfolio 5 (positive), negative-minus-positive (1-5) portfolio and their t-statistics. The second column "Ave." refers to the average alphas across five quintile portfolios with different levels of the control variables but with a similar level of predicted profitability shocks. The sample starts in January 1985 and ends in December 2015.

	Four-factor alpha						t-statistic					
Quintile	Ave.	Small	2	3	4	Large	Ave.	Small	2	3	4	Large
Panel A: Controlling for Size												
1 Negative	0.35	0.22	0.32	0.31	0.35	0.53	5.08	2.20	3.31	2.88	3.04	5.41
5 Positive	-0.26	-0.02	-0.32	-0.35	-0.28	-0.32	-4.00	-0.17	-3.08	-3.12	-2.65	-4.29
Neg. - Pos.	0.60	0.24	0.65	0.66	0.62	0.84	5.95	1.65	4.51	4.36	4.42	6.03

	Four-factor alpha						t-statistic					
Quintile	Ave.	Low	2	3	4	High	Ave.	Low	2	3	4	High
Panel B: Controlling for Book-to-Market												
1 Negative	0.44	0.56	0.55	0.74	0.14	0.19	4.71	4.24	3.68	4.65	0.82	0.97
5 Positive	-0.27	-0.07	-0.30	-0.28	-0.25	-0.43	-4.06	-0.72	-2.92	-2.28	-2.01	-3.04
Neg. - Pos.	0.70	0.63	0.86	1.02	0.39	0.62	5.80	3.49	4.84	5.04	1.82	2.49
Panel C: Controlling for Operating Profitability												
1 Negative	0.30	-0.30	0.33	0.46	0.38	0.63	3.21	-1.24	2.05	2.88	2.37	4.23
5 Positive	-0.23	-0.49	-0.23	-0.38	-0.10	0.05	-3.55	-3.53	-1.89	-3.13	-0.90	0.49
Neg. - Pos.	0.53	0.19	0.56	0.84	0.49	0.58	4.41	0.68	2.59	4.35	2.40	3.05
Panel D: Controlling for Price Momentum												
1 Negative	0.46	0.97	0.48	0.64	0.15	0.06	5.33	5.26	3.30	4.26	1.12	0.44
5 Positive	-0.27	-0.44	-0.16	-0.25	-0.14	-0.37	-3.94	-2.91	-1.27	-2.14	-1.26	-2.68
Neg. - Pos.	0.73	1.40	0.64	0.88	0.29	0.43	6.37	5.68	3.28	4.81	1.77	2.33

Panel E: Controlling for Max												
	Four-factor alpha						t-statistic					
Quintile	Ave.	Low	2	3	4	High	Ave.	Low	2	3	4	High
1 Negative	0.42	0.42	0.61	0.42	0.35	0.33	4.86	2.84	4.10	2.67	2.01	1.77
5 Positive	-0.30	-0.13	-0.20	-0.23	-0.54	-0.40	-4.17	-1.16	-1.91	-2.03	-3.57	-2.19
Neg. - Pos.	0.72	0.54	0.81	0.65	0.89	0.72	5.86	3.19	4.39	3.39	4.08	2.94
Panel F: Controlling for Idiosyncratic Volatility												
1 Negative	0.35	0.31	0.47	0.40	0.35	0.22	3.87	2.49	3.23	2.29	2.02	1.08
5 Positive	-0.28	-0.04	-0.18	-0.47	-0.40	-0.34	-3.56	-0.38	-1.52	-3.87	-2.54	-1.70
Neg. - Pos.	0.64	0.35	0.65	0.87	0.76	0.55	4.88	2.19	3.48	4.02	3.37	2.14
Panel G: Controlling for Idiosyncratic Skewness												
1 Negative	0.47	0.51	0.69	0.45	0.42	0.28	5.43	3.41	4.96	3.02	2.54	1.58
5 Positive	-0.29	-0.06	-0.12	-0.25	-0.48	-0.54	-4.33	-0.50	-1.14	-2.21	-3.95	-3.70
Neg. - Pos.	0.76	0.57	0.81	0.70	0.90	0.82	6.46	3.11	4.49	3.60	4.44	3.80
Panel H: Controlling for SUE												
1 Negative	0.51	0.29	0.62	0.61	0.62	0.43	6.08	1.66	4.21	4.26	3.96	2.77
5 Positive	-0.33	-0.54	-0.25	-0.25	-0.27	-0.31	-5.22	-4.92	-2.01	-2.14	-2.42	-2.38
Neg. - Pos.	0.84	0.83	0.87	0.86	0.89	0.74	7.07	3.86	4.32	4.48	4.52	3.44
Panel I: Controlling for CAR7												
1 Negative	0.51	0.44	0.53	0.59	0.62	0.39	5.99	2.58	3.77	4.59	4.36	2.43
5 Positive	-0.32	-0.38	-0.46	-0.24	-0.25	-0.26	-4.89	-2.87	-4.19	-2.26	-2.38	-1.94
Neg. - Pos.	0.83	0.82	0.99	0.83	0.87	0.66	6.92	3.86	5.71	4.93	4.58	3.08
Panel J: Controlling for Institutional Holding												
1 Negative	0.28	0.78	0.61	0.60	0.03	-0.60	3.29	4.84	4.07	4.06	0.22	-3.93
5 Positive	-0.35	-0.18	-0.15	-0.25	-0.20	-0.96	-4.95	-1.22	-1.32	-2.06	-1.55	-6.87
Neg. - Pos.	0.63	0.96	0.77	0.85	0.23	0.36	5.43	4.03	3.74	4.30	1.45	1.96
Panel K: Controlling for Analyst Coverage												
1 Negative	0.61	1.08	0.78	0.38	0.40	0.41	7.27	7.95	5.79	2.68	2.63	2.52
5 Positive	-0.17	0.07	-0.20	-0.15	-0.29	-0.28	-2.71	0.49	-1.60	-1.33	-2.68	-2.42
Neg. - Pos.	0.78	1.01	0.98	0.54	0.69	0.69	7.20	5.25	5.33	3.02	3.55	3.26

Table 2.6 Firm-level predictors of monthly stock returns

The table shows the average slopes and their t-statistics (in parentheses) from cross-sectional regressions that predict monthly returns for January 1985 to December 2015. Independent variables include predicted profitability shock (PPS), the log of market capitalizations (Size), the log of book-to-market ratio (BM), operating profitability (OP), investment (INV), standardized unexpected profitability (SUE), cumulative seven-day abnormal returns (CAR7), monthly excess returns (EXRET), cumulative returns over months t-12 through t-2 (MOM), the maximum daily return over the past one month (MAX) and idiosyncratic volatility (IVOL) over the past six months. Independent variables are trimmed at the 1% and 99% levels on a table-by-table basis to ensure different regressions within each table panel are based on the same observations. Except for the full example from January 1985 to December 2015, the table also presents results of separate FM regressions for two subsamples using the end of 1999 as the breakpoint.

	Panel A: Full sample			Panel B: 01/1985 to 12/1999			Panel C: 01/2000 to 12/2015		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
PPS		-3.081	-4.685		-1.915	-4.423		-4.174	-4.931
		(-2.50)	(-5.04)		(-1.37)	(-3.77)		(-2.09)	(-3.45)
SIZE	-0.071		-0.059	-0.009		0.006	-0.129		-0.119
	(-2.00)		(-1.64)	(-0.19)		(0.11)	(-2.51)		(-2.29)
BM	0.124		0.074	0.160		0.119	0.091		0.032
	(2.02)		(1.22)	(2.21)		(1.69)	(0.93)		(0.33)
OP	1.925		0.636	2.424		1.127	1.458		0.177
	(4.92)		(1.42)	(3.95)		(1.75)	(2.96)		(0.28)
INV	-0.722		-0.724	-0.832		-0.838	-0.619		-0.616
	(-4.04)		(-4.07)	(-3.31)		(-3.36)	(-2.43)		(-2.43)
SUE	0.187		0.192	0.191		0.196	0.183		0.189
	(11.84)		(12.18)	(8.61)		(8.83)	(8.14)		(8.40)
CAR7	0.019		0.020	0.026		0.027	0.012		0.014
	(6.64)		(7.23)	(5.64)		(5.94)	(3.64)		(4.20)
MOM	0.031		0.222	0.598		0.780	-0.501		-0.301
	(0.14)		(1.05)	(2.60)		(3.55)	(-1.36)		(-0.85)
EXRET	-3.972		-4.033	-4.756		-4.823	-3.237		-3.293
	(-7.63)		(-7.77)	(-7.20)		(-7.31)	(-4.08)		(-4.17)
MAX	-0.028		-0.025	-0.022		-0.021	-0.033		-0.029
	(-2.39)		(-2.20)	(-1.37)		(-1.29)	(-1.99)		(-1.81)
IVOL	0.004		0.002	-0.013		-0.016	0.020		0.019

	Panel A: Full sample			Panel B: 01/1985 to 12/1999			Panel C: 01/2000 to 12/2015		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
	(0.16)		(0.07)	(-0.41)		(-0.50)	(0.55)		(0.50)
$A(R^2)$	0.057	0.005	0.059	0.053	0.005	0.055	0.061	0.005	0.063

Table 2.7 Spanning tests for the value, operating profitability, momentum and predicted profitability shock factors

This table reports results of time-series regressions to explain the BM factor HML (Panel A), the OP factor RMW (Panel B), the price momentum factor WML (Panel C) and the predicted profitability shock factor NMP (Panel D). The explanatory factors are the returns to various combinations of the market factor MKT, the size factor SMB, the investment factor CMA, the standardized unexpected profitability (SUE) factor, the book-to-market factor HML, the operating profitability factor RMW, the price momentum factor WML and the predicted profitability shock factor NMP. The sample period is from January 1985 to December 2015.

	Dependent variable											
	Panel A: WML			Panel B: HML			Panel C: RMW			Panel D: NMP		
Specification	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
a	0.548	-0.125	0.466	0.141	0.255	0.126	0.277	0.391	0.131	0.206	0.528	0.475
	(2.35)	(-0.65)	(3.23)	(1.00)	(2.72)	(1.32)	(2.41)	(4.33)	(1.49)	(1.50)	(4.48)	(5.53)
β_{HML}		-0.069	0.200					-0.569	-0.557		0.240	0.211
		(-0.65)	(2.53)					(-13.70)	(-14.71)		(3.68)	(4.44)
β_{RMW}		0.226	0.593		-0.597	-0.669					0.328	0.423
		(2.08)	(7.27)		(-13.70)	(-14.71)					(4.94)	(8.69)
β_{WML}					-0.017	0.086		0.052	0.214			-0.420
					(-0.65)	(2.53)		(2.08)	(7.27)			(-17.99)
β_{MKT}		-0.146	-0.155		-0.026	-0.008		-0.074	-0.033		-0.008	-0.069
		(-3.26)	(-4.74)		(-1.15)	(-0.34)		(-3.45)	(-1.65)		(-0.29)	(-3.41)
β_{SMB}		0.027	-0.106		0.057	0.080		0.051	0.086		-0.119	-0.108
		(0.45)	(-2.43)		(1.96)	(2.79)		(1.81)	(3.30)		(-3.28)	(-4.07)
β_{CMA}		0.169	-0.053		0.596	0.596		0.045	0.089		-0.198	-0.127
		(1.35)	(-0.57)		(11.23)	(11.52)		(0.74)	(1.61)		(-2.59)	(-2.27)
β_{SUE}		1.643	0.619		-0.213	-0.148		-0.074	0.030		-0.915	-0.224
		(14.64)	(6.21)		(-3.11)	(-2.16)		(-1.08)	(0.48)		(-13.30)	(-3.55)
β_{NMP}			-1.120			0.243			0.406			
			(-17.99)			(4.44)			(8.69)			
$A(R^2)$	0.000	0.440	0.704	0.000	0.624	0.643	0.000	0.465	0.557	0.000	0.393	0.679

Table 2.8 Spanning tests for quintile portfolios sorted by momentum, book-to-market, operating profitability and predicted profitability shocks within size quintiles

This table reports results of spanning tests for value-weight quintile portfolios sorted by momentum, the book-to-market ratio, operating profitability and predicted profitability shocks constructed within size quintiles using NYSE breakpoints. The alphas of specifications (2) and (3) are intercepts of the spanning tests using a set of explanatory variables specified in specifications (2) and (3) of Table 2.7. The sample period is from January 1985 to December 2015.

	Four-factor alpha					t-statistic				
	Small	2	3	4	Large	Small	2	3	4	Large
Sorting by price momentum										
Specification (2)										
1 Loser	-0.38	-0.09	0.18	0.24	0.40	-1.63	-0.51	1.00	1.26	2.52
5 Winner	0.33	-0.07	-0.13	-0.23	-0.12	2.61	-0.67	-1.11	-1.74	-1.05
Win. - Los.	0.72	0.03	-0.31	-0.47	-0.53	2.40	0.10	-1.22	-1.72	-2.12
Specification (3)										
1 Loser	-0.85	-0.53	-0.25	-0.20	-0.04	-3.87	-3.32	-1.65	-1.17	-0.30
5 Winner	0.55	0.14	0.11	0.06	0.13	4.54	1.50	0.95	0.51	1.29
Win. - Los.	1.40	0.66	0.36	0.26	0.17	5.29	3.10	1.69	1.15	0.87
Sorting by BM										
Specification (2)										
1 Growth	-0.09	-0.23	-0.18	-0.09	0.05	-0.87	-2.98	-1.98	-0.98	0.63
5 Value	0.11	0.07	0.19	0.16	0.06	0.79	0.65	1.58	1.46	0.60
Val.- Gro.	0.20	0.30	0.38	0.25	0.01	1.24	1.94	2.18	1.73	0.08
Specification (3)										
1 Growth	-0.03	-0.16	-0.07	0.03	0.07	-0.25	-1.93	-0.68	0.34	0.92
5 Value	0.13	-0.01	0.12	0.06	-0.03	0.85	-0.06	0.91	0.55	-0.31
Val.- Gro.	0.15	0.15	0.18	0.03	-0.10	0.91	0.94	1.02	0.20	-0.69
Sorting by OP										
Specification (2)										
1 Low	-0.27	-0.29	-0.17	-0.22	-0.21	-2.48	-3.23	-1.76	-2.27	-1.93
5 High	0.34	0.24	0.22	0.27	0.18	3.75	3.02	2.64	2.83	2.50
High - Low	0.61	0.53	0.39	0.49	0.39	4.99	3.83	2.78	3.47	2.65
Specification (3)										
1 Low	-0.20	-0.12	0.02	-0.10	-0.14	-1.68	-1.29	0.17	-1.02	-1.21
5 High	0.21	0.07	0.05	0.13	0.03	2.26	0.87	0.63	1.32	0.39
High - Low	0.41	0.19	0.04	0.23	0.17	3.23	1.36	0.27	1.62	1.10
Sorting by expected profitability shock										
Specification (2)										
1 Negative	0.18	0.22	0.18	0.27	0.42	1.49	1.84	1.45	1.98	3.79
5 Positive	-0.14	-0.50	-0.41	-0.36	-0.26	-1.17	-4.75	-3.53	-3.17	-3.19
Neg. - Pos.	0.31	0.72	0.59	0.63	0.68	1.97	4.10	3.33	3.45	4.22
Specification (3)										
1 Negative	0.14	0.17	0.13	0.22	0.38	1.33	1.80	1.30	1.95	4.04
5 Positive	-0.13	-0.48	-0.39	-0.34	-0.24	-1.13	-4.73	-3.49	-3.13	-3.15
Neg. - Pos.	0.28	0.65	0.52	0.56	0.62	1.86	4.48	3.58	3.91	4.56

Table 2.9 Spanning tests for the maximum daily return, expected idiosyncratic skewness and predicted profitability shock factors

This table reports the results of spanning tests for the maximum daily return MAX factor (Panel A), the expected idiosyncratic skewness E(IKEW) factor (Panel B) and the predicted profitability shock factor NMP (Panel C). The explanatory factors are the returns to various combinations of the market factor MKT, the size factor SMB, the investment factor CMA, the standardized unexpected profitability (SUE) factor, the book-to-market factor HML, the operating profitability factor RMW, the price momentum factor WML, the maximum daily return MAX factor, the expected idiosyncratic skewness E(IKEW) factor and the predicted profitability shock factor NMP. The sample period is from January 1985 to December 2015.

	Dependent variable								
	Panel A: MAX			Panel B: E(IKEW)			Panel C: NMP		
Specification	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
a	0.257	0.479	0.311	0.282	0.461	0.368	0.206	0.475	0.441
	(1.16)	(3.29)	(2.09)	(1.09)	(3.11)	(2.39)	(1.50)	(5.53)	(5.14)
β_{MKT}		-0.428	-0.403		-0.489	-0.476		-0.069	-0.025
		(-12.44)	(-11.79)		(-13.95)	(-13.42)		(-3.41)	(-1.00)
β_{SMB}		-0.600	-0.562		-0.816	-0.795		-0.108	-0.040
		(-13.41)	(-12.55)		(-17.87)	(-17.11)		(-4.07)	(-1.11)
β_{HML}		0.466	0.391		0.370	0.329		0.211	0.155
		(5.79)	(4.84)		(4.51)	(3.92)		(4.44)	(3.20)
β_{CMA}		0.122	0.167		0.025	0.050		-0.127	-0.167
		(1.29)	(1.79)		(0.26)	(0.52)		(-2.27)	(-2.99)
β_{RMW}		0.076	-0.073		0.035	-0.048		0.423	0.438
		(0.92)	(-0.83)		(0.42)	(-0.52)		(8.69)	(8.91)
β_{SUE}		-0.233	-0.154		-0.090	-0.046		-0.224	-0.184
		(-2.19)	(-1.45)		(-0.82)	(-0.42)		(-3.55)	(-2.96)
β_{WML}		0.218	0.366		0.403	0.485		-0.420	-0.428
		(5.51)	(6.88)		(9.99)	(8.79)		(-17.99)	(-16.36)
β_{NMP}			0.353			0.196			
			(4.06)			(2.17)			
$\beta_{E(IKEW)}$									-0.113
									(-2.06)
β_{MAX}									0.083
									(1.47)
β_{IVOL}									0.136
									(2.01)
$A(R^2)$	0.000	0.650	0.665	0.000	0.729	0.733	0.000	0.679	0.697

Table 2.10 Spanning tests for quintile portfolios sorted by maximum daily return, expected idiosyncratic skewness and predicted profitability shocks within size quintiles

This table reports the results of spanning tests for value-weight quintile portfolios sorted by maximum daily return, expected idiosyncratic skewness and predicted profitability shocks constructed within size quintiles using NYSE breakpoints. The set of explanatory variables used in specification (2) and (3) are specified in Table 9. The sample period is from January 1985 to December 2015.

	Four-factor alpha					t-statistic				
	Small	2	3	4	Large	Small	2	3	4	Large
Sorting by MAX										
Specification (2)										
1 Low	0.52	0.37	0.34	0.24	0.18	5.57	3.78	3.29	2.22	1.83
5 High	-0.40	-0.42	-0.15	-0.15	-0.13	-2.53	-3.90	-1.09	-1.09	-1.01
Low - High	0.92	0.80	0.49	0.39	0.31	4.40	4.30	2.45	1.86	1.53
Specification (3)										
1 Low	0.38	0.28	0.26	0.19	0.12	4.04	2.73	2.42	1.66	1.21
5 High	-0.31	-0.30	-0.04	0.01	-0.07	-1.88	-2.70	-0.28	0.08	-0.50
Low - High	0.68	0.58	0.30	0.18	0.19	3.21	3.05	1.45	0.81	0.90
Sorting by E(ISKEW)										
Specification (2)										
1 Low	0.21	0.19	0.24	0.20	0.02	2.04	1.86	2.23	1.81	0.19
5 High	-0.17	-0.35	-0.33	-0.31	-0.25	-0.82	-2.44	-2.49	-2.25	-1.98
Low - High	0.39	0.54	0.58	0.51	0.26	1.52	2.73	3.03	2.64	1.46
Specification (3)										
1 Low	0.10	0.16	0.25	0.21	0.06	0.91	1.47	2.24	1.77	0.64
5 High	-0.10	-0.22	-0.29	-0.20	-0.11	-0.44	-1.47	-2.12	-1.43	-0.86
Low - High	0.19	0.37	0.55	0.41	0.17	0.74	1.83	2.77	2.03	0.89
Sorting by Predicted profitability shock										
Specification (2)										
1 Negative	0.14	0.17	0.13	0.22	0.38	1.33	1.80	1.30	1.95	4.04
5 Positive	-0.13	-0.48	-0.39	-0.34	-0.24	-1.13	-4.73	-3.49	-3.13	-3.15
Neg. - Pos.	0.28	0.65	0.52	0.56	0.62	1.86	4.48	3.58	3.91	4.56
Specification (3)										
1 Negative	0.14	0.09	0.10	0.25	0.37	1.26	0.99	0.91	2.23	3.83
5 Positive	-0.07	-0.52	-0.38	-0.31	-0.25	-0.60	-5.07	-3.38	-2.82	-3.20
Neg. - Pos.	0.21	0.62	0.48	0.56	0.62	1.40	4.21	3.24	3.84	4.46

Chapter 3
Tail coskewness risk in momentum portfolios

3.1. ABSTRACT

Over long sample periods, momentum portfolios exhibit negative skewness caused by extreme left-tail events or momentum crashes. We use tail coskewness to capture momentum crash risk caused by the tendency for winners to crash following bull markets and for losers to rebound following bear markets. Allowing for tail coskewness, which captures momentum crash risk, causes the loadings on past returns to fall by as much as 65.8% in our cross-sectional regressions. Tail coskewness subsumes the well-known size effect in average returns to size and momentum portfolios.

3.2. Introduction

Early studies on price momentum well recognize that past stock returns and skewness are negatively related, as past winners have negative skewness while past losers have less negative or even slightly positive skewness (Harvey and Siddique, 2000; Chen, Hong and Stein: 2001). A number of recent studies (Daniel and Moskowitz, 2015; Barroso and Santa-Clara: 2015) highlight that a winners-minus-losers strategy, though generating high average monthly returns can also give rise, in some months, to huge negative returns. These momentum crashes can be very damaging to an investor and could take many years, or even decades, to recover from even in the presence of offsetting positive returns.[11] In this study, we develop an empirical methodology for estimating the risk premium associated with momentum crashes, or left-tail events. More generally, we investigate the role of tail events and skewness in explaining returns to momentum portfolios.

To this end, we introduce the concept of tail coskewness derived from coskewness by Kraus and Litzenberger (1976) and Harvey and Siddique (2000) and apply it in a momentum context. These authors showed that an asset's contribution to the skewness of the market portfolio, or its coskewness with respect to the market portfolio, was priced. Although we confirm that this coskewness pricing effect also exists in our momentum portfolios, we find that such effect is driven almost exclusively by tail events. Therefore, we develop a measure, tail coskewness, that focuses exclusively on how tail events (low-probability events leading to large gains or losses) contribute to the systematic skewness of momentum portfolios.[12] This contrasts with

[11] Barroso and Santa-Clara (2015) note that the long-short momentum strategy experienced a crash of -91.59% over two months in 1932 and a crash of -73.42% over three months in 2009.

[12] Here, we measure tail events as those returns that lie in the lowest x% and the highest x% of returns based on the full return distribution of individual assets where we vary x from 1% to 8% in our empirical tests.

the coskewness metric which measures, for a given asset, the contribution of tail and non-tail returns, towards portfolio skewness.

Our empirical results show that tail coskewness has strong explanatory power over the cross-section of size/momentum returns, even allowing for size and past returns. By our estimates, a two standard deviation increase in tail coskewness, based on exposure to tail events only, implies that average returns to size/momentum portfolios increases by 0.313% (t=-2.49) per month. This is nearly two-thirds of overall average market excess returns, which are 0.46% per month. Importantly, when the coskewness calculation omits tail events then its pricing effect usually becomes insignificant. Our tail coskewness measure also gives rise to a larger implied risk premium than that associated with the existing coskewness measure. Our results show that investors in momentum strategies care about skewness caused by extreme events and demand additional compensation for exposure to that risk.

The ability of past returns to explain momentum profits also reduces significantly once tail coskewness is included. Adding tail coskewness in our multivariate regressions causes the average slopes on past returns to drop by as much as 65.8% and these slopes sometimes become statistically insignificant. Our paper is the first to document estimates of momentum profits corresponding to systematic tail coskewness risk. Moreover, we show that the familiar size effect completely loses its explanatory power for average momentum returns after controlling for tail coskewness.

We also show that winner portfolios have more negative tail coskewness than losers following both bull and bear markets because of their reduced capacity to rebound following bear markets and their greater potential to crash following bull markets. In other words, the explanatory power of tail coskewness is attributable to the fact that it captures the two scenarios when momentum portfolios may crash: when the market portfolio rebounds following a bear market and when the market portfolio crashes following a bull market. Finally, we show that tail coskewness explains momentum returns better than asymmetric beta, i.e., the difference between upside and downside beta (DeBondt and Thaler, 1987; Rouwenhorst, 1998; Dobrynskaya: 2015).

Our empirical methodology follows Chan and Chen (1988) and Fama and French (1992) who investigated the role of the Sharpe-Lintner-Black beta in cross-sectional returns. Specifically, we use portfolios formed on firm size and past returns with frequent rebalancing, over the period July 1963 to June 2013 and compute tail coskewness for each portfolio over the full 50-year sample period. We then run a series of Fama-Macbeth (FM) regressions to examine the cross-sectional relationship between observed portfolio returns and the full-period tail coskewness and other variables such as size and past returns. This technique allows our portfolios to have the same relative exposure to tail coskewness risk over the full 50-year period, even if the true risk exposure of the portfolios is time-varying. We use the full sample period to estimate tail coskewness since tail events are very infrequent so that the precision of our tail coskewness estimates can be improved. The increased precision of the full sample estimates was the main reason for their use in Chan and Chen (1988) and Fama and French (1992). We also follow the approach of Fama and French (2006) and Boyer, Mitton and Vorkink (2010) to predict tail coskewness using a two-stage regression approach which alleviates concerns that future information is included when tail coskewness is computed using the full sample period. However, we find that the use of forecast tail coskewness instead of full period tail coskewness does not change our main results.

Our paper adds to a growing literature (Daniel and Moskowitz, 2015; Barroso and Santa-Clara, 2015) describing how skewness in momentum portfolios can cause momentum crashes. Though the focus of those papers is on forecasting momentum crashes and reducing exposure to the momentum strategy accordingly, our focus is on interpreting momentum crashes as the realization of tail coskewness risk and then quantifying the magnitude of that risk. We interpret the high average returns associated with the momentum strategy as compensation for tail coskewness risk and argue that investors in the winners-minus-losers strategy would be mostly concerned by the small probability of large crashes that could take years to recover from. Investors should then demand high average returns as compensation for those negative outcomes. We also expand on a growing recent literature that shows that tail risk is priced in the cross-section of stock returns even in the presence of coskewness (Kelly and Jiang, 2014; Weigert, 2015; Chabi-Yo, Ruenzi and F.Weigert, 2016) and in returns to well-diversified mutual funds (Xiong, Idzorek and R.Ibbotson, 2014), by showing that these tail effects are present in momentum returns.

The paper proceeds as follows. Section 3.3 discusses our empirical methodology including the formation of portfolios in our testing procedures. Section 3.4 discusses our estimates for tail coskewness. Section 3.5 analyses the explanatory power of tail coskewness in the cross-section of average excess returns to size/momentum portfolios. In Section 3.6, we check if our findings hold across sub-samples and internationally. Section 3.7 concludes.

3.3. Empirical Methodology and Data

This study follows Chan and Chen (1988) and Fama and French (1992) in forming portfolios based on instrumental variables that are highly correlated with the risk factor of interest which is tail coskewness.

The instrumental variables we use are past returns and firm size. Harvey and Siddique (2000) and Chen, Hong and Stein (2001) show that winner portfolios are more likely to experience a substantial negative shock compared to loser portfolios as winner portfolios have lower skewness than loser portfolios. Accordingly, past returns should be negatively related to future skewness, coskewness and tail coskewness with respect to the market portfolio. It is important to note that we avoid using the skewness of individual stocks to form test portfolios since the skewness of individual stock returns does not persist across different periods and is not a good predictor of future skewness (Singleton and Wingender, 1986; Chen, Hong and Stein, 2001; Boyer, Mitton and Vorkink, 2010).

We use firm size as the second instrumental variable together with past returns. Chan and Chen (1991) show that portfolios of small stocks may contain a large portion of less efficient and highly leveraged firms that are prone to crashing in poor economic conditions while Harvey and Siddique (2000) and Barone-Adesi, Gagliardini, and Urga (2004) show that small firms tend to be more negatively skewed than large firms.

We then investigate the cross-sectional relationship between tail coskewness estimates, computed from the full sample period, and observed returns to those portfolios after allowing for size and past returns. An advantage of this technique is that though the exposure of our portfolios to tail coskewness risk may vary through time, the relative strength of the exposures to tail coskewness risk may persist over a long period. Using more observations then leads to more precise estimates of the exposures to tail coskewness risk than those obtained using shorter time periods. It should be noted that using the same value for tail coskewness in each of our cross-sectional regressions13 does not mean that the exposure of our portfolios to tail coskewness risk is constant over time. Instead we assume that the exposure to tail coskewness risk follows a time-varying but stationary process. In addition, the constituent stocks comprising our portfolios change over time due to changes in size and past return performance relative to other stocks and because each constituent stock may not have constant exposure to tail coskewness risk.

[13] The value of tail coskewness in our cross-sectional regressions varies slightly from month to month because, when calculating tail coskewness, we omit the single data point corresponding to the month of interest. This is to avoid spurious correlation in our FM regressions.

We use four datasets in our empirical tests. The first dataset we use comprise the 25 value-weighted size/momentum portfolios from the Kenneth French Data Library, that incorporates all CRSP firms in the US and all NYSE, AMEX and NASDAQ firms with share codes 10 and 11. Our sample period is July 1963 through June 2013. The portfolios are constructed monthly based on the intersection of 5 portfolios formed on size and 5 portfolios formed on prior (2-12) returns. The monthly size and prior (2-12) return breakpoints are the NYSE quintile breakpoints. To be included in a portfolio for month t (formed at the end of month t-1), a stock must have had a valid return for month t-12 and for month t-2. Each included stock also must have had a valid firm size as at the end of month t-1.

The second dataset consists of 100 value-weighted size/momentum portfolios, comprising all CRSP firms included in the first dataset of 25 value-weighted size/momentum portfolios. We implement two pass sorts to construct the 100 size/momentum portfolios, because there is some panel imbalance (i.e., empty cells) under the double-independent-sort procedure used for the 25 size/momentum portfolios. That is, we first sort stocks based on market capitalization into deciles and then subdivide each size portfolio into 10 momentum portfolios based on prior returns. Again, our size and prior return breakpoints are the NYSE decile breakpoints.

For the momentum component of the 100 size/momentum portfolios, we use different combinations of the ranking period R and the holding period H. While we keep updating the size decile monthly based on firm size at formation month (t=-1), momentum deciles are constructed with different combinations of R=5 or 11 months and H=1, 3, 6, 9 or 12 months. Following Jegadeesh and Titman (1993), we construct overlapping portfolios to increase the power of our tests. In any given month t, the portfolio monthly return is the simple average of H overlapping portfolios that update their momentum deciles in the preceding t-H months as well as in month t-1. For example, for H=3, the monthly portfolio return for April 1970, t, is the simple average of three portfolios: these are based on past returns for the R months to February 1970 (t-2), January 1970 (t-3), and December 1969 (t-4).

The third dataset comprises the 25 value-weight size/momentum portfolios for developed markets (November 1990 to March 2014) from the Kenneth French Data Library. This dataset combines 23 developed markets into four regions: (i) North America, including the United States and Canada; (ii) Japan; (iii) Asia Pacific, including Australia, New Zealand, Hong Kong, and Singapore (but not Japan); and (iv) Europe, including Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. The size breakpoints for each region are the 3rd, 7th, 13th, and 25th percentiles of that region's aggregate market capitalization. For portfolios formed at the end of month t–1, the prior returns are a stock's cumulative returns from month t–12 to t–2. The momentum breakpoints for all size quintiles in a region are the 20th, 40th, 60th, and 80th percentiles based on prior returns for the top 90% of stocks, by market capitalization, for that region.

For the last dataset, taken from Datastream, we construct 25 value-weighted size/momentum portfolios, incorporating all United Kingdom stocks traded on the London Stock Exchange (LSE). We implement two pass sorts to construct this set of portfolios for July 1975 to June 2014. The size breakpoints are the market capitalization breakpoints for NYSE quintiles, to avoid sorts dominated by the plentiful but less important tiny stocks from the LSE. For portfolios formed at the end of month t–1, the prior returns for a constituent stock are that stock's cumulative returns for months t–12 to t–2. The momentum breakpoints are the 20th, 40th, 60th, and 80th percentiles based on prior returns to stocks within each size quintile.

3.4. Measures of Coskewness and Tail Coskewness

Coskewness reflects an asset's contribution to the skewness of a well-diversified portfolio. An asset with positive coskewness, if included in a portfolio, makes that portfolio more positively skewed and vice versa for an asset with negative coskewness. Since risk averse investors dislike negative skewness, coskewness and average returns should be negatively correlated. Thus, an asset with more negative coskewness should have a higher expected return and vice versa. Harvey and Siddique (2000) construct a measure of coskewness, standardized coskewness, as shown in Equation (1):

$$\widehat{CS}_{\mathrm{i}}^{S} = T^{-1}\sum_{t=1}^{T}\frac{\epsilon_{i,t}\left(r_{M,t}-\widehat{\mu_M}\right)^2}{\widehat{\sigma_{\epsilon_{\mathrm{i}}}}\widehat{\sigma_M^2}} \quad (1)$$

where T denotes the total number of observations being used to compute estimates; $\epsilon_{it} = r_{it} - a_i - \beta_i\left(r_{Mt} - r_{ft}\right)$, is the residual from the regression of asset i's excess returns on the contemporaneous market excess return; r_{Mt} and $\widehat{\mu_M}$ are the market return and its mean respectively; $\widehat{\sigma_{\epsilon_i}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2}$ and $\widehat{\sigma_M^2} = \frac{1}{T}\sum_{t=1}^{T}(r_{Mt} - \widehat{\mu_M})^2$.

Motivated by the finding that tail events matter most for the shape of skewness (Badrinath and Chatterjee, 1988; Peiro, 1999; Kim and White, 2004), we propose a new measure, tail coskewness, that focuses exclusively on the contribution of tail events. We define standardized tail coskewness as:

$$\widehat{TCS}_{\mathrm{i}}^{S} = T^{-1}\sum_{t=1}^{T}\frac{\omega_{i,t}\left(r_{M,t}-\widehat{\mu_M}\right)^2}{\widehat{\sigma_{\epsilon_{\mathrm{i}}}}\widehat{\sigma_M^2}}$$

$$\omega_{i,t} = \begin{cases}\epsilon_{it,} \text{ if } F(\epsilon_{it}) \leq \alpha/2 \text{ or } F(\epsilon_{it}) \geq 1 - \alpha/2 \\ 0, \text{ if } \alpha/2 < F(\epsilon_{it}) < 1 - \alpha/2\end{cases} \quad (2)$$

where T is kept in the numerator to allow for comparison between tail coskewness TCS_i^S and coskewness CS_i^S; ϵ_{it} is a tail event observation lying in the extreme left or right tail of the residual distribution. Specifically we define a two-tailed α (alpha), equally divided in both tails, such that α/2 is in each tail as given by the cumulative distribution function for residuals, denoted F. The tail coskewness of an asset represents an asset's marginal effect on the skewness of a well-diversified portfolio via its extreme residuals. Our empirical tests vary the two-tailed α from 0.02 to 0.16 to identify which value of α helps to fully capture the risk premium associated with tail coskewness.

It is worth emphasizing that the standardized coskewness will tend to be more negative (positive) when a portfolio has many negative (positive) CAPM residuals corresponding with when the market portfolio is either well above or well below its mean. In this study, the feature that is most prominent is the propensity for winner portfolios to crash heavily when the market portfolio crashes following a bull market and the propensity for losers to rebound very strongly when the market rebounds following a bear market. A more negative value for standardized coskewness compared to standardized tail coskewness is indicative of the fact that when the market portfolio crashes following a bull market then it is highly likely that the portfolio in question crashes even more heavily while a more positive value indicates that when the market portfolio rises following a bear market then it is highly likely that the portfolio rises even more significantly.

We also use two non-standardized measures of coskewness. Although the standardized coskewness measure makes the pricing effect of coskewness independent of the CAPM beta, it is well documented that the linear relationship between the CAPM beta and average returns does not necessarily hold in practice (Reinganum, 1981; Stambaugh, 1982; Lakonishok and Shapiro, 1986; Fama and French, 1992, 2004). Thus, the benefit of using the standardized coskewness to isolate the effect of beta on coskewness seems doubtful. More seriously, the residual distribution that is being used to define tail events may not resemble the original distribution of tail events, which results in biased estimates of coskewness. We therefore construct the non-standardized coskewness:

$$\widehat{CS_{\mathrm{i}}^{N}} = T^{-1} \sum_{t=1}^{T} \frac{(r_{i,t} - \widehat{\mu_{\mathrm{i}}})(r_{M,t} - \widehat{\mu_{M}})^2}{\widehat{\sigma_{\mathrm{i}}} \widehat{\sigma_{M}^{2}}} \tag{3}$$

where r_{it} is the excess return of an individual asset and $\widehat{\mu_{\mathrm{i}}}$ is its full sample mean. Similarly, the non-standardized tail coskewness is defined as:

$$\widehat{TCS_{\mathrm{i}}^{N}} = T^{-1} \sum_{t=1}^{T} \frac{\psi_{i,t}(r_{M,t} - \widehat{\mu_{M}})^2}{\widehat{\sigma_{\mathrm{i}}} \widehat{\sigma_{M}^{2}}}$$

$$\psi_{i,t} = \begin{cases} r_{it} - \widehat{\mu_{\mathrm{i}}}, \text{ if } F(r_{it}) \leq \alpha/2 \text{ or } F(r_{it}) \geq 1 - \alpha/2 \\ 0, \text{ if } \alpha/2 < F(r_{it}) < 1 - \alpha/2 \end{cases} \tag{4}$$

where r_{it} represents an individual asset's excess returns that lie in the tail areas of that asset's return distribution. When comparing the non-standardized measures of coskewness and tail coskewness, the tail coskewness is always less negative reflecting the fact that for both winner and loser portfolios the propensity for extreme negative returns outweighs the propensity for extreme positive returns but where the propensity for extreme negative returns is higher for winners relative to losers.

In Figure 3.1, we plot on a monthly basis the contribution of each month in the sample towards the calculation of full-sample non-standardized coskewness. We do this for an average loser portfolio and for an average winner portfolio14. Figure 3.1 shows that both winner and loser portfolios suffer more extreme downside events (allied with a contemporaneous extreme downside event to the market portfolio) than extreme upside events. The tail events for these two portfolios mainly occur during turbulent periods such as the 1970s Oil Crisis, the 1987

[14] This analysis is based on the 25 value-weighted size and momentum portfolios from the Kenneth R. French data library. Here the return to the average loser portfolio is defined as the average return to the loser portfolios within each of the five size quintiles. The return to the average winner portfolio is defined as the average return to the winner portfolios within each of the five size quintiles.

Black Monday Crash, the 1998 Russian Financial Crisis, the 2000 Dot Com Bubble crash and the 2007 Global Financial Crisis. These occasional crashes (particularly for winners) and sharp rebounds (particularly for losers) represent the spikes in the monthly components of the non-standardized coskewness calculation and they overwhelmingly dominate other less extreme observations. This provides further motivation for our emphasis on tail events in this study.

Earlier studies show that coskewness is relevant to market valuations and that the omission of coskewness can lead to a misspecification of asset pricing models (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000). Table 3.1 suggests that coskewness might provide a risk based explanation for the momentum effect since coskewness estimates vary systematically across the winner and loser portfolios within each size quintile. In the right hand columns of Panel A, Table 3.1, we observe a monotonic relationship between coskewness and portfolio returns; for each size group the winner portfolio has lower negative coskewness (for both standardized: CS_i^S and non-standardized: CS_i^N) than the loser portfolio. The high correlation between coskewness and portfolio mean excess return (-0.807 for CS_i^S and -0.837 for CS_i^N from Panel B) also suggests that coskewness is considerably correlated with variation in excess returns to size/momentum portfolios.

For tail coskewness, Panel A, Table 3.1 shows that, when the two tailed α is set to 10%, the patterns of tail coskewness across our portfolios closely resembles that of coskewness for both standardized and non-standardized measures. In the last column of Panel A, the values for non-standardized tail coskewness estimates are very close to the corresponding coskewness estimates. This suggests that coskewness estimates are mainly attributable to extreme tail events. Panel B shows that the correlation between coskewness and tail coskewness is extremely high, 0.964 for the standardized measure and 0.979 for the non-standardized measure. Tail coskewness estimates are also highly correlated with mean excess return (-0.796 for

TCS_i^S and -0.821 for TCS_i^N). Tail coskewness may therefore account for much of the cross-sectional variation in excess returns associated with coskewness.

Interestingly, in the second last column of Panel A, we observe that coskewness estimates for the non-standardized measure are all negative. The uniformly negative sign can be explained by the asymmetric correlation reported by Longin and Solnik (2001) and Ang and Chen (2002). They suggest that correlations between equity portfolios and the aggregate market are much greater for downside moves than for upside moves, especially for extreme returns. As coskewness measures the correlation between equity portfolios and the square of the market, the magnitude of downside correlations outweighs upside correlations for all 25 size/momentum portfolios. Thus, for portfolios such as winner portfolios that suffer more frequent extreme losses than gains, their tail coskewness is more negative. Conversely, for loser portfolios that have more frequent extreme gains than losses compared to winners, their tail coskewness is less negative.

Two other results in Panel B of Table 3.1 are noteworthy. First, CAPM betas have weak correlation with coskewness and tail coskewness, ranging from 0.149 to 0.289. This finding alleviates concerns that the pricing effect of coskewness and tail coskewness found in our subsequent multivariate tests may actually be due to an omitted CAPM beta. Second, coskewness and tail coskewness are both mildly correlated with size, ranging from 0.412 to 0.506, and with past returns, ranging from -0.524 to -0.572. Hence, the issue of multi-collinearity should not arise when we use all of size, past returns and tail coskewness (or coskewness) in our subsequent regressions.

3.5. Empirical Results

3.5.1. Tail Coskewness Versus Coskewness

Table 3.2 shows time-series averages of the slopes from Fama-Macbeth (FM) regressions of excess returns on size, prior return, and tail coskewness, for our 25 value-weighted size/momentum portfolios15. To help understand the economic magnitude of risk premia for coskewness and tail coskewness, we also report, where necessary, the change in risk premia corresponding to a two standard deviation change in coskewness and tail coskewness respectively. Following Fama and French (2008), who argue that there is little reason to expect the CAPM beta to correlate with anomaly variables, we do not include the CAPM beta in our cross-sectional regressions16.

Column 1 of Panel A shows that the size and momentum effects (prior 2 to 12 month cumulative returns) help to explain the cross-section of size/momentum returns. We observe that the average slope on size is -0.06 with a t-statistic of -1.67. The momentum effect is much stronger with an average slope of 0.951 on past returns and a t-statistic of 4.25. However, the average intercept of FM regression is 0.763% with a t-statistic of 2.02, which is anomalously high relative to the contemporaneous average market excess return of 0.46% per month. The large intercept indicates that size and past returns fail to fully capture the patterns in average returns to our size and momentum portfolios.

Column 2 of Panel A shows that when all returns are used to compute standardized coskewness then the pricing effect of coskewness is observed, albeit at a significance level of 10%. This result and the fact that the standardized coskewness is always more negative for winners compared to losers (Table 3.1), shows that investors require significantly higher average returns from momentum strategies to compensate for coskewness risk. Perhaps more

[15] As our data set covers 600 months from July 1963 to June 2013, the Newey-West autocorrelation consistent covariance estimator with 5 lags is used, guided by the common practice of using a lag length of $T^{0.25}$. To avoid spurious correlation in our FM regressions, when computing our tail coskewness estimates, we leave out the single data point corresponding to the test month and use all other observations to compute the numerators of those estimators.

[16] In unreported univariate and multivariate regression results, the CAPM beta coefficients have an insignificant negative sign.

significantly, when coskewness excludes left and right-tail observations then its statistical significance disappears and the coefficient on past returns increases. This shows that the higher average returns that investors require as compensation for coskewness risk in momentum strategies, is in fact compensation for the risk of extreme tail events.

Panels B and C of Table 3.2 use standardized tail coskewness and standardized tail coskewness together with standardized coskewness that omits left and right-tail observations respectively, to explain the cross-section of size/momentum returns. The results in Panel B show that the standardized tail coskewness provides a stronger description of average returns to size/momentum portfolios, as the average slopes on tail coskewness are always negative and statistically significant with t-statistics greater than 2. Column 3 of Panel B shows for example that, using α =0.02, tail coskewness has an average slope of -0.481 with a t-statistic of -2.24. When the two-tailed α is increased to 0.08 (Column 6), the t-statistic for tail coskewness peaks in magnitude at 2.65 with a corresponding risk premium of 0.249% per month. Further increasing the two-tailed α to 0.10 and 0.12 does not bring incremental economic gains since the corresponding risk premia on tail coskewness in Columns 7 and 8 (0.235% and 0.249% per month) are largely similar to that from Column 6. Panel C of Table 3.2 reports the same regressions as for Panel B but adds robust coskewness which is coskewness calculated by omitting observations from the left and right tail. When the two-tailed α is set at 0.08 in Column 6, the average slope on tail coskewness is most negative at -0.808 (t = -2.41) corresponding to a risk premium of 0.313% per month. This result is consistent with a number of studies (Kelly and Jiang, 2014; Weigert, 2015; Chabi-Yo, Ruenzi and F.Weigert, 2016) that show that tail risk is priced in the cross-section of stock returns even allowing for coskewness risk. It is also important to note that since winner portfolios have lower tail coskewness than loser portfolios (Table 3.1), our results imply that investors in momentum strategies require higher average returns for higher exposure to tail risk.

Moreover, compared with Column 1 of Panel A, controlling for tail coskewness in Panels B and C dramatically changes the slopes on lnME, PR and the intercept. Here, we use Column 6 of Panel B to illustrate the key findings. First, the relationship between firm size and excess returns seems to be completely absorbed by coskewness, since the slope on lnME is no longer significant (-0.020, t = -0.46). This disappearing size effect confirms the result of Barone-Adesi, Gagliardini and Urga (2004) that the explanatory power of size in the cross-section of returns is due to it being a proxy for omitted coskewness risk. Second, the economic significance of the momentum effect falls when tail coskewness is included. The slope on PR reduces from 0.951 (t = 4.25) in Column 1 (Panel A) to 0.738 (t = 3.89) in Column 6 Panel B, a reduction of 22.4%. Finally, controlling for tail coskewness now makes the regression intercept statistically insignificant (0.567, t = 1.30), which clearly demonstrates that tail coskewness is a critical explanatory variable for the cross-section of size/momentum portfolio returns.

Table 3.3 reports the explanatory power of non-standardized coskewness and tail coskewness. Column 1 of Panel A shows that the non-standardized coskewness has an average slope of -1.165 with a t-statistic of -1.80. Regressions 2 to 7 of Panel A then show that robust estimates of coskewness (trimmed estimates that discard observations in tail areas) cannot help to explain the cross-section of size/momentum returns. In Column 6, after omitting 5% of observations belonging to the left tail and 5% of observations in the right tail for a two tailed α of 0.10, we find that the average slope on coskewness is now very close to zero and insignificant (-0.249, t=-0.34). In Panel B, using non-standardized tail coskewness, our results again imply that the pricing effect of coskewness is mainly attributable to tail events. Importantly since non-standardized tail coskewness is more negative for our winner portfolios compared to loser portfolios that implies that investors in momentum strategies require higher expected returns for the net positive exposure to tail risk. Further, the risk premium associated with tail risk is maximized when we measure tail events to be the top and bottom 5% of the return distribution and is equal to 0.313% per month.

Finally, Panel C provides further strong evidence that the pricing effect of coskewness is attributable to tail events since the simultaneous presence of tail coskewness and coskewness measured without the extreme left and right tails, shows that tail coskewness is priced but not trimmed coskewness.

In summary, the evidence from Tables 2 and 3 suggests that investors are concerned with tail risk when implementing momentum strategies, requiring higher average returns as compensation for this risk. This risk premium is because of their fear of severe momentum crashes as captured by the tail coskewness measure and is present even in the presence of past returns and size. For any level of two tailed α, the non-standardized measure gives a higher risk premium than that implied by its standardized counterpart and also more significantly reduces the slope on past returns. Non-standardized tail coskewness therefore seems to have stronger ability to explain the cross-section of size/momentum returns than does standardized coskewness. It is important to recognize that tail risk and past returns are correlated in the sense that past returns do imply some probability of a subsequent tail event. By excluding tail risk, the implied risk premium associated with past returns mistakenly includes a tail risk premium. We demonstrate this by showing that the risk premium implied by past returns is reduced once tail risk is allowed for.

3.5.2. Tail Coskewness Versus the Momentum Effect

To fully examine the relationship between tail coskewness and the momentum effect, we construct 100 size/momentum portfolios based on different ranking and holding periods. In this section, we use only non-standardized tail coskewness with a two-tailed α=0.10 since, from earlier, this estimator gives the largest implied risk premium for tail coskewness. We implement the overlapping procedure introduced by Jegadeesh and Titman (1993) to reduce the estimation errors in tail coskewness. Specifically, we first compute tail coskewness for each of the H overlapping portfolios based on the full sample period and take the simple average of these H values as an explanatory variable in our FM regressions. If the errors in tail coskewness estimates for our H overlapping portfolios are less than perfectly positively correlated, this averaging approach can greatly improve estimation precision.

Table 3.4 shows the average slopes and intercepts for the FM regressions of portfolio excess returns on size, momentum and tail coskewness. In Panel A, the results of bivariate regressions of excess returns on size and past returns show that past returns have a strong role in explaining the cross-section of size/momentum returns. In Panel B, using multivariate regressions, adding tail coskewness significantly reduces the explanatory power of past returns in all portfolio sets. In particular, the average slope for past returns becomes statistically insignificant after holding periods extend to 9 or 12 months. For example, the average slope on past returns for the R=5/H=12 portfolio set declines from 2.5 (t=4.71) in Panel A to 0.855 (t=1.51) in Panel B, and the average slope on past returns for the R=11/H=12 portfolio set declines from 0.637 (t=2.05) in Panel A to 0.267 (t=0.96) in Panel B. In contrast, the corresponding average slopes on tail coskewness are highly significant, t=-4.43 for the R=5/H=12 portfolio set and t=-3.32 for the R=11/H=12 portfolio set. This is a strong result as it shows that excess returns to some commonly implemented momentum strategies are consistently and strongly explained by the possibility of future crashes as well as by past returns and in some cases the tail risk is strong enough to cause the statistical significance of past returns to disappear.

To illustrate the economic significance of tail coskewness, we further decompose momentum profits into that part explained by tail coskewness and that part that remains unexplained. The last row of Panel B, Table 3.4 shows that tail coskewness explains a larger and larger fraction of momentum profits as the length of the holding period H increases. Figure 3.2 shows that, although tail coskewness (orange region) accounts for only a relatively small fraction of momentum profits for short holding periods, it accounts for an increasing fraction of momentum profits as H increases. Eventually, 65.8% of momentum profits for the R=5/H=12 strategy, and 58.1% of momentum profits for the R=11/H=12 strategy, are explained by tail coskewness. In addition, if we exclude the R=5/H=1 and R=5/H=3 portfolios, then the last row of Panel B, Table 3.4, shows that a two standard deviation increase in tail coskewness causes the expected return to momentum strategies with R=5 to increase by between 0.240% and 0.383% per month. For R=11, the expected return to momentum strategies increase by between 0.257% per month and 0.368% per month. This is our estimate of the risk premium associated with tail coskewness, allowing for size and past returns, for longer horizon momentum strategies.

An intuitive explanation for the higher explanatory power of tail coskewness in momentum profits, as H increases, is that momentum portfolios are increasingly exposed to tail risk as in any given month, there are now H constituent portfolios that could suffer a tail event and not one only as in our earlier analysis. However, it could also be the result of improvements in measurement precision from the use of overlapping portfolios. In particular, measurement errors in the tail coskewness estimate for each overlapping portfolio tend to cancel each other out during the averaging process, resulting in greater precision.

In short, Table 3.4 and Figure 3.2 show that, after controlling for tail coskewness, past returns now often have significantly reduced explanatory power in the cross-section of size/momentum returns. This is strong evidence that the high average returns to momentum strategies may be compensation for exposure to tail coskewness risk, the risk of winners being exposed to extreme downside events and losers being exposed to extreme upside events. This result is also consistent with recent literature on the significance of tail risk in the cross-section of stock returns, even in the presence of coskewness risk.

3.5.3 Expected Tail Coskewness and Returns to Size/Momentum Portfolios

Using the full-period tail coskewness raises a concern about whether expected tail coskewness conditional on information at time t, as opposed to future information, has predictive power in average returns. Using a two-stage regression approach similar to Fama and French (2006) and Boyer, Mitton and Vorkink (2010), we first form estimates as at the end of each month t of expected tail coskewness for each portfolio for month t+1. We then test the cross-sectional relationship between average returns and expected tail coskewness, computed using information up to and including month t, valid for month t+1. We run the two-stage regression as follows. At the end of each month t from June 1963 to May 2013, we separately run the following first-stage cross-sectional regression to explain tail coskewness:

$$TCS_{p,T} = a_{0,T} + a_{1,T} lnME_{p,T} + a_{2,T} PR_{p,T} + a_{3,T} \sigma_{p,T} + \omega_{p,T} \qquad (5)$$

where $TCS_{p,T}$ is the tail coskewness estimated for portfolio p for window T, which is an initial sample of 210 observations from January 1945 to June 1963 that expands by one month each time and ends at month t; $lnME_{p,T}(PR_{p,T})$ is the time-series average of firm size (prior returns) for portfolio p in the expanding window T; $\sigma_{p,T}$ is the time-series average of volatility (the standard deviation of daily returns computed each month) to portfolio p in the expanding window T. The monthly returns to the 25 size/momentum portfolios and daily returns used in the calculation of monthly portfolio volatility, are from the Kenneth French library. The slopes obtained from the first stage regression using the most recent expanding window are then used in the estimation of expected tail coskewness:

$$E\left[TCS_{p,t+1}\right] = a_{0,T} + a_{1,T} lnME_{p,t} + a_{2,T} PR_{p,t} + a_{3,T} \sigma_{p,t} \qquad (6)$$

where $E[TCS_{p,t+1}]$ is the expected tail coskewness of portfolio p for month t+1; $lnME_{p,t}$, $PR_{p,t}$ and $\sigma_{p,t}$ are the average values of firm size, prior return and volatility over the most recent expanding window which ends at month t for portfolio p.

We then run the following second-stage FM cross-sectional regressions to explain excess returns to our 25 size/momentum portfolios, for July 1963 to June 2013:

$$r_{p,t+1} = \gamma_0 + \gamma_1 E[TCS_{p,t+1}] + \gamma_2 lnME_{p,t} + \gamma_3 PR_{p,t} + \varepsilon_{i,t+1} \quad (7)$$

where $r_{p,t+1}$ is the monthly return for portfolio p observed at the end of month t+1; $lnME_{p,t}$ is the natural logarithm of firm size for portfolio p at the end of month t and $PR_{p,t}$ is past returns to portfolio p from month t-12 to t-2. This approach provides feasible estimates of expected tail coskewness each month, which are time varying, and which assumes that the long-run relationship between expected tail coskewness and returns remains unchanged over time.

Panel A of Table 3.5 reports results from the first-stage regressions to predict tail coskewness. For the full sample result, there are 600 monthly cross-sectional regressions using expanding windows from {January 1945 - June 1963} to {January 1945 - May 2013} to estimate the dependent variable of tail coskewness based on our explanatory variables that include the time-series average of firm size, prior returns and volatility. Consistent with our hypothesis that prior returns and size work as proxies for future tail coskewness, the average slopes on prior return are positive and significant in all of the monthly regressions, and the average slopes on size are negative and significant in 69.9% of the monthly regressions. Volatility, $\sigma_{p,T}$, also has strong explanatory power for tail coskewness, since its average slopes are significant in 73.2% of the monthly regressions. The positive relationship between volatility and tail coskewness shows that higher volatility leads to higher tail coskewness. In addition, the average F-stat (33.309) shows that the model has statistically significant predictive capability for future tail coskewness.

We also report subsample results where each subsample adds 120 extra monthly observations relative to the previous subsample and where the first subsample uses an initial expanding window of {January 1945 - June 1963} that expands to {January 1945 - June 1973}. The subsample results show that the explanatory power of size and volatility only become highly reliable (100% and 80% of point estimates are statistically significant) in the third subsample, where the expanding window covers at least 480 monthly observations {January 1945 - June 1983}. The adjusted R^2 also increases from 0.448 in the first subsample to 0.84 in the third subsample. These results support our argument that increasing sample size enhances the precision of tail coskewness estimation for size/momentum portfolios.

Panel B of Table 3.5 presents the results of the second-stage cross-sectional regressions examining the relationship between expected tail coskewness and excess returns to size/momentum portfolios for July 1963 to June 2013. Regression 1 is a baseline regression controlling for size and past returns only. Regression 2 uses estimates of expected tail coskewness based on the results of the relevant first-stage cross-sectional regression, i.e. the cross-sectional regression implemented at the end of month t. The slope on $E[TCS_{p,t+1}]$ is highly significant (-5.623, t = -5.65), which indicates that expected tail coskewness has strong predictive power in the cross section of average returns to size/momentum portfolios. After controlling for size and past returns in regression 3, the predictive power of expected tail coskewness is still statistically significant (-5.173, t = -2.66). The reduced but still highly significant coefficient on expected tail coskewness reflects the fact that expected tail coskewness is correlated with size and past returns. In addition, adding expected tail coskewness to our regressions increases the adjusted R^2 from 0.472 in regression 1 to 0.565 in regression 3. In summary, the results of Table 3.5 indicate that expected tail coskewness helps to explain the cross-sectional variation in returns to size/momentum portfolios. The concern that our earlier use of the full sample period to compute tail coskewness uses future returns currently unknown to investors is therefore significantly reduced.

3.5.4 Tail Coskewness Versus Time-varying Beta of Momentum Portfolios

A well-documented result in the momentum literature is the time-varying beta of momentum strategies (Grundy and Martin, 2001). If the return on the market is low over the ranking period (bear market), the loser portfolio will tend to include high market beta assets, and the winner portfolio will tend to include low market beta assets. Thus the beta of the momentum strategy is likely to be negative following bear markets and positive following bull markets. Daniel and Moskowitz (2015) additionally show that, when the market recovers sharply after bear markets, losers experience much stronger gains than do winners. These infrequent but strong gains generate relatively high but still negative tail coskewness for the loser portfolio. It is therefore important to ask whether the pattern of tail coskewness for momentum portfolios observed in the previous sections is largely attributable only to bear market states or whether it also applies to bull markets. To address this issue, we examine the risk characteristics of our 25 value-weighted size/momentum portfolios following both bull and bear markets. We define the market as a bull market when the market cumulative prior return for months t-12 to t-2, is positive, and as a bear market otherwise.

Table 3.6 shows that losers have less negative tail coskewness exposure than winners following both bear markets (Panel A) and bull markets (Panel B). However, following bear markets, none of the loser portfolios displays significant tail coskewness at the 10 percent level for any size group. Following bull markets, the negative tail coskewness for all portfolios is much more significant than the corresponding values in bear markets and winners still have more negative tail coskewness than losers. The difference in tail coskewness between winners and losers is larger following bull markets than following bear markets. For example, tail coskewness in the small size group after bull markets has a spread of 0.133 (-0.328 for the losers less -0.461 for the winners), while the spread after bear markets is 0.082 (-0.091 for the losers less -0.173 for the winners). Our result therefore suggests that crash risk in momentum strategies, as proxied by the difference in TCS values for winners versus losers, is likely to be more severe following bull markets.

We further investigate the cause of the observed pattern of tail coskewness for momentum portfolios and momentum crash risk using the concept of upside beta and downside beta by Bawa and Lindenberg (1977). The upside beta, β^+ , is computed over periods when the excess market return is above its mean, and the downside beta β^- is computed over periods when the excess market return is below its mean:

$$\beta_i^+ = \frac{Cov(r_i, r_M | r_M > \mu_m)}{Var(r_M | r_M > \mu_m)} \tag{8}$$

$$\beta_i^- = \frac{Cov(r_i, r_M | r_M < \mu_m)}{Var(r_M | r_M < \mu_m)} \tag{9}$$

Where μ_m is the average market excess return. Figure 3.3 shows intuitively how the upside beta and downside beta are related to momentum returns. In this figure the analysis is based on our 25 size/past return portfolios where the winner (loser) portfolio is defined as the average of the 5 winner (loser) portfolios within each size quintile. This figure shows that following bear markets losers have greater upside beta than downside beta while winners have greater downside beta than upside beta. Following bull markets, the figure shows that winners, particularly, have greater downside beta than upside beta17. Both of these results suggest the

[17] In Figure 3, the upside beta and the downside beta: (β^+, β^-), for the loser portfolio following bear markets is 1.814 and 1.439 respectively. For the winner portfolio, following bear markets, (β^+, β^-) = (0.630, 1.050). For the loser portfolio, following bull markets, (β^+, β^-) = (1.079, 1.081). For the winner portfolio, following bull markets, (β^+, β^-) = (1.141, 1.439).

possibility of momentum crashes when loser stocks rebound strongly following bear markets and winner stocks crash heavily following bull markets.

Table 3.6 provides summary statistics on upside and downside betas for our 25 size/past return portfolios and confirms the intuition provided in Figure 3.3. In every size quintile, following bear markets, Panel A, β^+ is three times larger for losers compared to winners. This indicates that momentum crash risk is present in all size quintiles following a bear market, particularly most pronounced in large stocks where the difference in β^+ between losers and winners is 1.413 (2.028 for losers versus 0.615 for winners). Overall, Panel A shows that losers tend to rebound strongly when the market recovers from losses, giving rise to less negative tail coskewness.

Following bull markets, Panel B, for every size quintile, β^- is larger for winners compared to losers while β^+ is similar for winners and losers. Thus momentum crash risk is present in all size quintiles following a bull market and this crash risk is again most acute for the large size group where the difference in β^- between winners and losers is 0.465 (1.244 for winners versus 0.779 for losers). Taken together, the results in Panel B suggest that large negative tail coskewness is driven by large negative betas following bull markets, particularly for winner portfolios.

In sum, our tail coskewness measure captures the extreme tail events following both bull and bear markets that give rise to momentum crashes. Table 3.6 extends the results of Grundy and Martin (2001) by examining betas following both bear and bull markets, to enable a better understanding of the momentum crashes of Daniel and Moskowitz (2015) and Barroso and Santa-Clara (2015).

3.5.5 Tail Coskewness versus Asymmetric Beta

As described in the previous section, tail coskewness is closely related to asymmetric beta, which is the difference between upside beta and downside beta, $\beta_i^+ - \beta_i^-$. Ang, Chen and Xing (2006) show that stock returns are negatively related to asymmetric beta. More recently, Dobrynskaya (2015) demonstrates that asymmetric beta helps to explain the cross-section of momentum returns. Since asymmetric beta and (tail) coskewness reflect asymmetry in return distributions, it is possible that they are different measures reflecting the same underlying risk factor. Thus, it is necessary to investigate which of asymmetric beta, coskewness and tail coskewness best explains the momentum effect.

The data used for this test are the 25 value-weighted size/momentum portfolios from July 1963 to June 2013. Panel A, Table 3.7 shows that asymmetric beta falls monotonically from losers to winners, which is very similar to the patterns observed for both coskewness and tail coskewness with respect to size/momentum portfolios. Asymmetric beta is highly correlated with coskewness and tail coskewness with correlations of 0.94 and 0.92, respectively.

Panel B, Table 3.7 presents the results from FM regressions of excess returns on asymmetric beta, coskewness, tail coskewness, size and prior returns. We first examine the explanatory power of asymmetric beta, coskewness and tail coskewness in univariate regressions. Regressions 1-3 reveal that all three measures help to explain the cross-section of size/momentum returns with t-statistics on the average slopes all less than -4.5. In regressions 4-6, adding size and past returns as independent variables results in tail coskewness and coskewness remaining significant, but asymmetric beta is now insignificant. The average slope on asymmetric beta (-0.177, t=-0.42) indicates that the role of asymmetric beta in explaining size/momentum returns might be weaker than that of coskewness and tail coskewness. Finally, the results from regression 7 confirm that there is considerable collinearity between asymmetric beta and tail coskewness. After both are included together as explanatory variables, the average slope on asymmetric beta becomes positive (1.307, t=2.78) whereas theoretically the slope should be negative. This abnormal slope is a typical sign of collinearity when the correlations among independent variables are very strong. Our results here suggest that while asymmetric beta and tail coskewness capture similar risk, tail coskewness has stronger explanatory power. This may be because extreme momentum crashes tend to occur when market variance is high (Barroso and Santa-Clara, 2015; Daniel and Moskowitz, 2015) and coskewness becomes larger in magnitude when market variance is high, when compared with asymmetric beta18.

To examine how much asymmetric beta captures tail events, we compute the simple correlation between asymmetric beta and trimmed non-standardized coskewness, i.e. the non-standardized coskewness metric based on the middle 90% of return observations. We find a correlation of 0.55, whereas the correlation between asymmetric beta and non-standardized coskewness, which includes all returns, is significantly higher at 0.94. This result suggests that the explanatory power of asymmetric beta in explaining the cross-section of size/momentum returns, at least in a univariate sense, seems likely to be driven by the impact of tail events. This is consistent with the pricing impact of tail events in the cross-section of stock returns shown by Kelly and Jiang (2014), Weigert (2015) and Chabi-Yo, Ruenzi and F.Weigert (2016).

3.6. Robustness

3.6.1 Subperiod analysis

To investigate the robustness of tail coskewness, we examine the 25 size/momentum portfolios from the Kenneth French Data Library in three separate sub-periods (January 1927 to June 1963; July 1963 to June 1983; July 1983 to June 2013).

[18] Dobrynskaya (2015) also shows the explanatory power of asymmetric beta over the cross-section of size/momentum returns. However, she does not relate her analysis to momentum crashes or examine the impact of different market states.

Like our basic result for July 1963 to June 2013, our results for all three sub-periods confirm that tail coskewness is statistically and economically important. The first two panels of Table 3.8 show that the average slopes for tail coskewness are significantly negative for July 1963 – June 1983 and July 1983 – June 2013 (-0.944, t=-2.07 for the first sub-period; -1.140, t=-1.71 for the second sub-period) in our multivariate FM regressions. The use of tail coskewness also significantly reduces the slopes and t-statistics on size, past returns and the intercept. Table 3.8 also shows that the effect of coskewness is weak in the 1983-2013 sub-period, with a t-statistic of -1.36. Asymmetric beta is also insignificant in the 1983-2013 (t=-0.47) and the 1963-1983 (t=0.96) sub-periods. The sub-period results thus support the conclusion that tail coskewness has a stronger role than asymmetric beta in explaining average returns to size/momentum portfolios.

The last panel of Table 3.8 shows the results for the pre-1963 sub-period (1927-1963). In FM regressions controlling for size, momentum and tail coskewness, the average slope for tail coskewness is -0.482, with a t-statistic of -2.03. The effect of tail coskewness is thus also significant in the pre-1963 period. Note also that tail coskewness completely subsumes the explanatory power of past returns as its t-statistic falls to 1.00 compared with 1.88 in FM regressions controlling for size and past returns. In the regressions controlling for coskewness, the average slope on coskewness (-0.496, t=-2.12) is very close to that observed for tail coskewness in previous FM regressions. This result is consistent with our previous finding that the explanatory power of coskewness in average returns is driven mainly by tail events. Finally, the last regression in Table 3.8 shows that asymmetric beta has a positive average slope (0.540, t=1.96), which is contrary to our theoretical prediction. Overall, extending our tests to the pre-1963 sub-period provides further evidence that tail coskewness has consistent explanatory power over the cross-section of size/momentum returns.

3.6.2. Value, Profitability, Investment and Industry Momentum Effects

We have also calculated the coskewness and tail coskewness measures and implemented FM regressions for 25 size/book-to-market, 25 size/operating profitability, 25 size/investment and 25 size/industry-momentum portfolios in the spirit of Moskowitz and Grinblatt (1999). For brevity, the descriptive statistics for these value-weighted portfolios are not reported. They show no clear relationship between coskewness or tail coskewness and any of book-to-market, operating profitability and investment. There is, however, a relationship between coskewness and industrial momentum such that winner industries have more negative coskewness or tail coskewness than their loser counterparts.

Table 3.9 presents the results from our FM regressions. They show that tail coskewness has significant incremental explanatory power over the cross-section of size/book-to-market (lnBM), size/operating profitability (OP), size/investment (INV) and size/industry-momentum returns. However, tail coskewness does not significantly impact the point estimates of lnBM, OP and INV in the respective regressions, suggesting that tail coskewness is not significantly related to these risk factors. Our result suggests that inclusion of these additional factors in our earlier FM regressions would not significantly affect our estimates of the risk premium corresponding to momentum crash risk.

3.6.3 International Market

We now examine whether tail coskewness is important in explaining returns to size/momentum portfolios in international data. We first examine the 25 value-weighted size/momentum portfolios for developed markets (Global, North America, European, Japan and Asia Pacific) from the Kenneth French data library. Two limitations arise from using these data. The first is the lack of past returns to use as an explanatory variable in our FM regressions. The second is the relatively short sample period from November 1990 to March 2014. To mitigate these problems, we separately construct 25 value-weighted size/momentum portfolios for the United Kingdom for a longer period, from July 1975 to June 2014.

Table 3.10 demonstrates that the pricing effect of tail coskewness is pervasive across all developed markets except Japan. Excluding Japan, the average slopes on tail coskewness are always more than two standard errors from zero. For the global market, tail coskewness has an average slope of -5.158 (t=-5.30) showing that tail coskewness is important in explaining cross-sectional variation in size/momentum returns. An important question is why tail coskewness shows no explanatory power over Japanese size/momentum portfolios. It is worth noting that Japan is the only developed market where the momentum effect does not exist (Chui, Titman and Wei, 2010; Fama and French, 2012; Moskowitz and Pedersen, 2013)[19]. This simultaneous absence of the momentum anomaly and the pricing effect of tail coskewness might provide indirect evidence that tail coskewness is the genuine risk factor that underlies the momentum effect.

Table 3.11 shows that the UK equity market delivers similar results to the US equity market. Again, there exists a reward for exposure to tail coskewness and the momentum effect becomes subsumed by tail coskewness. In regressions 1-3, of Table 3.11, all of the average slopes on tail coskewness, coskewness and asymmetric beta are statistically significant with t-statistics exceeding four. Regression 4 says that the momentum effect shows up strongly in the UK equity market, since past returns have an average slope of 1.102 with a t-statistic of 2.71. However, in regression 5, adding tail coskewness reduces the average slope on past returns to 0.572 with a t-statistic of 1.26 whereas the average slope for tail coskewness is statistically

[19] Asness (2011) argues that momentum does exist in Japan in that the returns adjusted for the Fama and French (1993) three factor model are significantly positive. However the raw momentum returns are not significantly different from zero. We also calculate tail coskewness for 25 size/momentum portfolios in Japan, covering the period from 1990-2014, and find that the tail coskewness for all of these portfolios is positive. This is a striking result as the tail coskewness for almost all of the corresponding portfolios for the Global, North America, Europe and Asia-Pacific regions, is negative.

significant with a t-statistic of -4.58 that exceeds the t-statistics observed for coskewness and asymmetric beta in regressions 6 and 7. Thus average returns to the 25 UK size/momentum portfolios seem to be more explained by exposure to tail coskewness risk compared than by exposure to coskewness risk or asymmetric beta.

3.7. Conclusion

We provide robust evidence to support a risk-based explanation for the momentum effect. Using size/momentum portfolios from the US, UK and other major developed markets, we show that momentum returns are associated with tail coskewness and interpret this as a systematic risk factor. This is the risk of a momentum crash due to a more significant downside event to winners relative to losers following a bull market or due to a more significant upside event occurring to losers relative to winners following a bear market. We estimate the risk premium corresponding to tail coskewness or momentum crash risk to be 0.313% per month. The effective risk premium is even higher as the slope on past returns reduces significantly, by as much as 65.8%, in the presence of tail coskewness. We also show that the familiar size effect completely loses its explanatory power for size/momentum portfolios after controlling for tail coskewness. Our risk-based explanation for momentum and size effects sheds light on how risk-averse investors use portfolio characteristics, such as size and past returns as simple proxies to identify and price tail risk. This would otherwise be very difficult to estimate due to the infrequency of tail events. Finally, our results for momentum portfolios, are consistent with recent literature (Kelly and Jiang, 2014; Weigert, 2015 and Chabi-Yo, Ruenzi and F.Weigert, 2016) that shows that tail risk is priced in the cross-section of stock returns.

3.8. References

Albuquerque, R., 2012. Skewness in stock returns: Reconciling the evidence on firm versus aggregate stock returns. Review of Financial Studies 25, 1630-1673.

Amaya, D., Christofferson, P., Jacobs, K., Vasquez, A., 2015. Does realized skewness predict the cross-section of equity returns? Journal of Financial Economics 118, 135-167.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity returns. Journal of Financial Economics 63, 443–494.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Review of Financial Studies 19, 1191-1239.

Arditti, F., 1967. Risk and the required return on equity. Journal of Finance 22, 19-36.

Asness, C., 2011. Momentum in Japan: The exception that proves the rule. Journal of Portfolio Management 37, 67-75.

Asness, C., Moskowitz, T., Pedersen, L., 2013. Value and momentum everywhere. Journal of Finance 68, 929-985.

Badrinath, S., Chatterjee, S., 1988. On measuring skewness and elongation in common stock return distributions: The case of the market index. Journal of Business 61, 451-472.

Bali, T., Cakici, N., Whitelaw, R., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. Journal of Financial Economics 99, 427-446.

Barone-Adesi, G., Gagliardini, P., Urga, G., 2004. Testing asset pricing models with coskewness. Journal of Business and Economic Statistics 22, 474-485.

Barroso, P., Santa-Clara, P., 2015. Momentum has its moments. Journal of Financial Economics 116, 111-120.

Bawa, V., Lindenberg, E., 1977. Capital market equilibrium in a mean-lower partial moment framework, Journal title?? 12, 635-635.

Bowley, A., 1920. Elements of Statistics. Charles Scribner's Sons, New York.

Boyer, B., Mitton, T. Vorkink, K., 2010. Expected idiosyncratic skewness. Review of Financial Studies 23, 169-202.

Chabi-Yo, F., Ruenzi,S., Weigert,F., 2016, Crash Sensitivity and the Cross-Section of Expected Stock Returns, Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2011746

Chan, K., Chen, N., 1988. An unconditional asset-pricing test and the role of firm size as an instrumental variable for risk. Journal of Finance 43, 309-325.

———, 1991, Structural and return characteristics of small and large firms. Journal of Finance 46, 1467-1484.

Chang, B., Christofferson, P., Jacobs, K., 2013. Market skewness risk and the cross section of stock returns. Journal of Financial Economics 107, 46-68.

Chen, J., Hong, H., Stein, J.,2001. Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. Journal of Financial Economics 61, 345-381.

Chordia, T., Shivakumar, L., 2002. Momentum, business cycle, and time-varying expected returns. Journal of Finance 57, 985-1019.

Chui, A.. Titman, S. Wei, K., 2010. Individualism and momentum around the World. Journal of Finance 65, 361-392.

Chung, Y., Johnson, H. Schill, M., 2006. Asset pricing when returns are nonnormal: Fama-French factors versus higher-order systematic comoments. Journal of Business 79, 923-940.

Conrad, J., Dittmar, R., Ghysels, E., 2013. Ex ante skewness and expected stock returns. The Journal of Finance 68, 85-124.

Cooper, M., Gutierrez Jr, R., Hameed, A., 2004. Market states and momentum. The Journal of Finance 59, 1345-1365.

Daniel, K., Moskowitz, T., 2015. Momentum crashes, forthcoming. Journal of Financial Economics.

De Bondt, W., Thaler, R., 1987. Further evidence on investor overreaction and stock market seasonality. Journal of Finance 42, 557-581.

Dobrynskaya, Victoria, 2015. Upside and Downside Risks in Momentum Returns. Available at SSRN: http://ssrn.com/abstract=2399359 or http://dx.doi.org/10.2139/ssrn.2399359

Elton, E., Gruber, M., Brown, S., Goetzmann, W., 2010. Modern Portfolio Theory and Investment Analysis. 6th edition. Wiley, New York.

Fama, E., 1976. Foundations of Finance. Basic Books, Inc., New York.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427-465.

---------, 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

———, 2004. The capital asset pricing model: Theory and evidence. Journal of Economic Perspectives 18, 25-46.

———, 2006. Profitability, investment and average returns. Journal of Financial Economics 82, 491-518.

———, 2008. Dissecting anomalies. Journal of Finance 63, 1653-1678.

———, 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457-472.

Griffin, J., Ji, X., Martin, J., 2003. Momentum investing and business cycle risk: Evidence from pole to pole. Journal of Finance 58, 2515-2547.

Grundy, B., Martin, J., 2001, Understanding the nature of the risks and the source of the rewards to momentum investing. Review of Financial Studies 14, 29-78.

Harvey, C., Siddique, A., 1999. Autoregressive conditional skewness. Journal of Financial and Quantitative Analysis 34, 465-487.

———, 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263-1295.

Jacobs, H., Regele, T and M.Weber, 2016. Expected Skewness and Momentum. Available at SSRN: http://ssrn.com/abstract=2600014 or http://dx.doi.org/10.2139/ssrn.2600014

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65-91.

———, 2001. Profitability of momentum strategies: An evaluation of alternative explanations. Journal of Finance 56, 699-720.

Kelly, B., Jiang, H., 2014. Tail risk and asset prices. Review of Financial Studies 27, 2841-2871.

Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics. Charles Griffin and Co., London.

Kim, T., White, H., 2004. On more robust estimation of skewness and kurtosis. Finance Research Letters 1, 56-73.

Kraus, A., Litzenberger, R. 1976. Skewness preference and the valuation of risk assets. Journal of Finance 31, 1085-1100.

Longin, F., Solnik, B., 2001. Extreme correlation of international equity markets. Journal of Finance 56, 649–676.

Menezes, C., Geiss, C., Tressler, J., 1980. Increasing downside risk. American Economic Review 70, 921-932.

Merton, R., 1990. Capital market theory and the pricing of financial securities. In: Friedman, B., Woodford, M., (Eds.) Handbook of Monetary Economics. North-Holland, Amsterdam, pp. 497-581.

Miller Jr., R., 1997. Beyond ANOVA: Basics of Applied Statistics. Chapman & Hall, London.

Mitton, T., Vorkink, K., 2007, Equilibrium underdiversification and the preference for skewness. Review of Financial Studies 20, 1255-1288.

Peiro, A., 1999. Skewness in financial returns. Journal of Banking & Finance 23, 847-862.

Rouwenhorst, K., 1998. International momentum strategies. Journal of Finance 54, 1439-1464.

Reinganum, M., 1981. A new empirical perspective on the CAPM. Journal of Financial and Quantitative Analysis 16, 439-462.

Singleton, J., Wingender, J., 1986. Skewness persistence in common stock returns. Journal of Financial and Quantitative Analysis 21, 335-341.

Stambaugh, R., 1982. On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis. Journal of Financial Economics 10, 237-268.

Vanden, J., 2006. Option coskewness and capital asset pricing. Review of Financial Studies 19, 1279-1320.

Wang, J., Wu, Y., 2011. Risk adjustment and momentum sources. Journal of Banking and Finance 35, 1427-1435.

Weigert, F., 2015. Crash Aversion and the Cross-Section of Expected Stock Returns Worldwide. Review of Asset Pricing Studies 6, 135-178.

Xiong, J., Idzorek,T and R.Ibbotson, 2014. Volatility versus Tail Risk: Which One is Compensated in Equity Funds? Journal of Portfolio Management 40, 112-121.

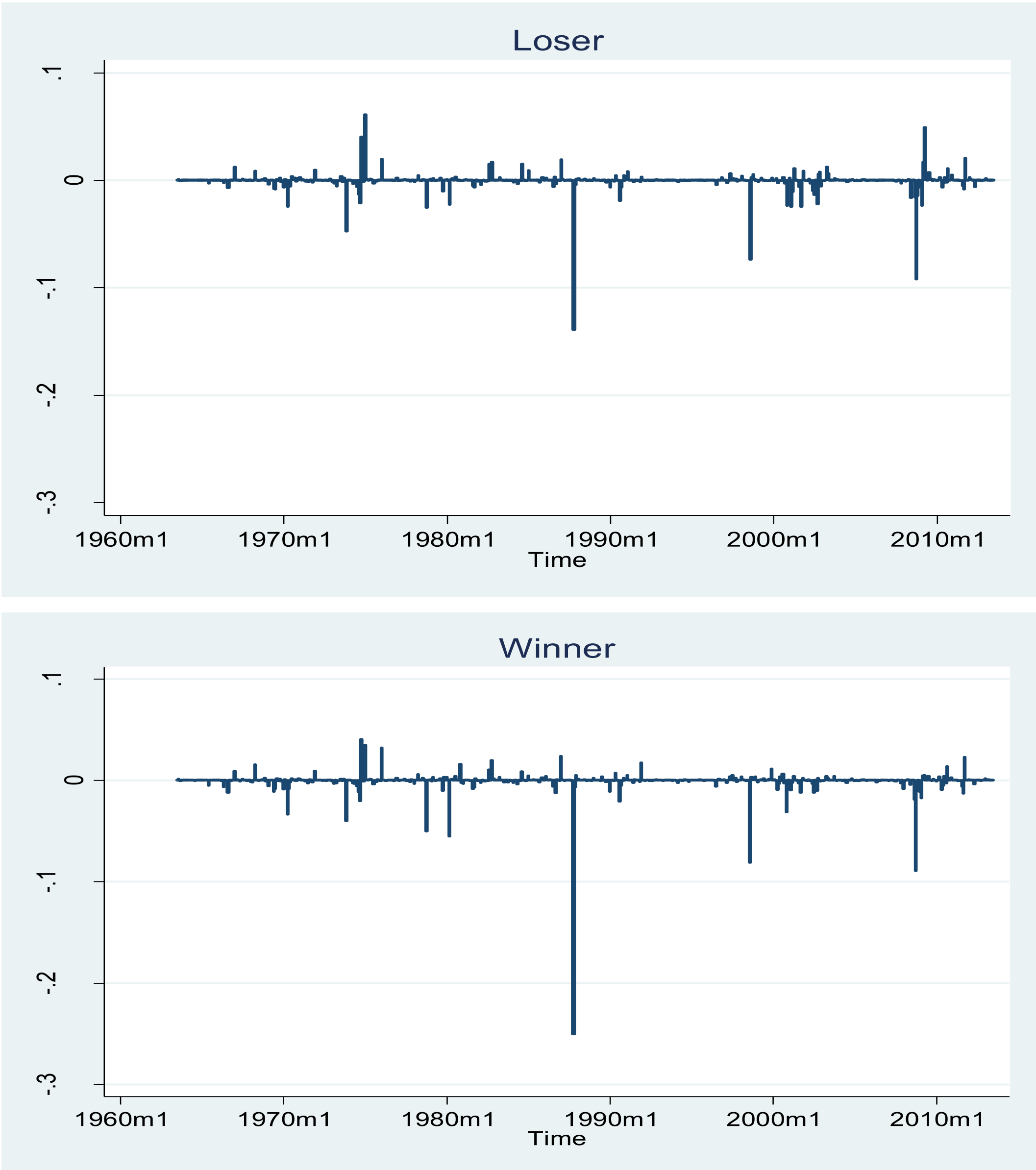

Figure 3.1 Time-series monthly components of non-standardized coskewness for two portfolios: average winners and average losers

This analysis is based on the 25 value-weighted size and momentum portfolios from the Kenneth R. French data library. The return to the average loser portfolio is the average return to the loser portfolios within each of the five size quintiles. The return to the average winner portfolio is the average return to the winner portfolios within each of the five size quintiles. The monthly component of non-standardized coskewness for portfolio i for each month t, from July 1963 to June 2013 is defined as $\frac{(r_{i,t}-\widehat{\mu_i})(r_{M,t}-\widehat{\mu_M})^2}{\widehat{\sigma}_i\widehat{\sigma_M^2}}$ where $r_{i,t}, r_{m,t}, \widehat{\mu}_i, \widehat{\mu_M}$ are the return to portfolio i, the return to the market portfolio in month t, the mean return to portfolio i and the mean return to the market portfolio based on T=600 monthly observations, and $\widehat{\sigma}_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_{i,t} - \widehat{\mu}_i\right)^2}$ and $\widehat{\sigma_M^2} = \frac{1}{T}\sum_{t=1}^{T}(r_{Mt} - \widehat{\mu_M})^2$ are the standard deviation and variance of returns to portfolio i and the market portfolio. The average of these monthly components from July 1963 to June 2013 is the value of the non-standardized coskewness defined as $\widehat{CS_i^N} = T^{-1}\sum_{t=1}^{T}\frac{(r_{i,t}-\widehat{\mu_i})(r_{M,t}-\widehat{\mu_M})^2}{\widehat{\sigma}_i\widehat{\sigma_M^2}}$.

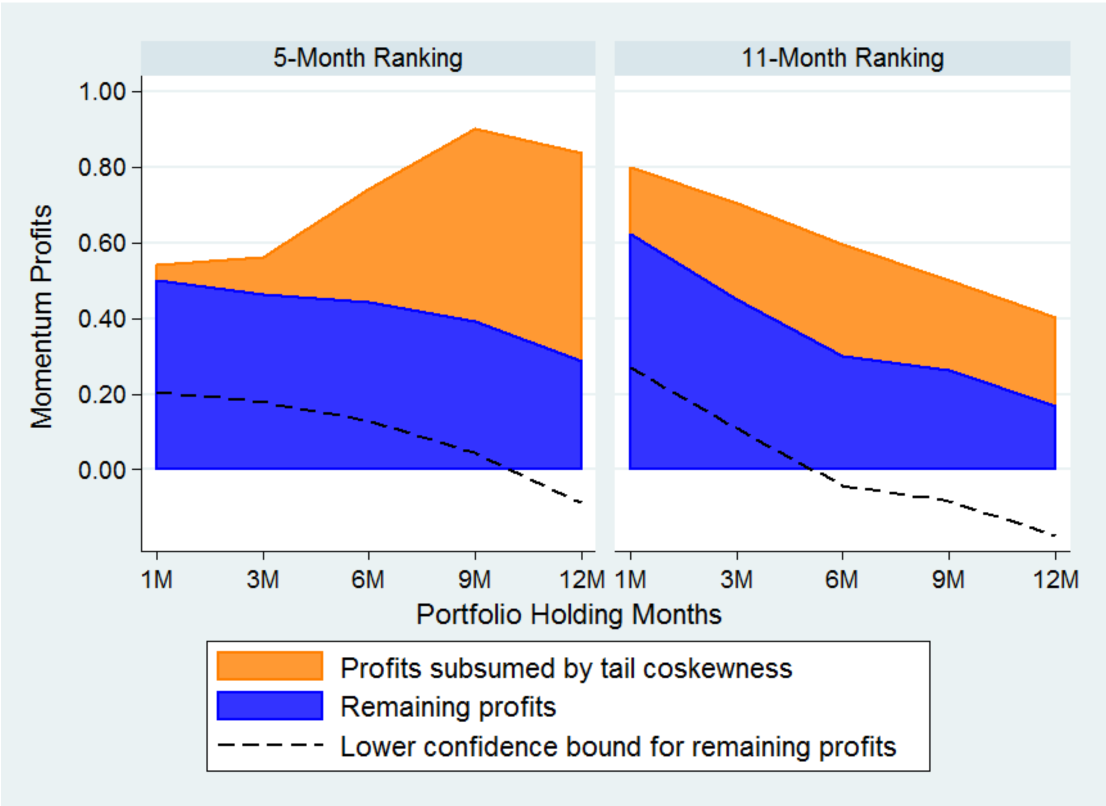

Figure 3.2 Decomposition of momentum profits for 100 size/momentum portfolios for different momentum strategies

In this analysis, momentum profits refers to the amount of excess returns explained by past returns in FM regressions of excess returns to 100 size/momentum portfolios, and is measured as the product of a two standard deviation change in past returns with the average regression slope on past returns. The height of the graph represents the amount of momentum profits explained by past returns in FM regressions controlling for size and past returns (see Panel A Table 3.4). The bottom (blue) region represents the amount of momentum profits explained by past returns in FM regressions with an additional control of tail coskewness (see Panel B Table 3.4). The top (orange) region represents the portion of the momentum profits that is subsumed by tail coskewness, measured as the difference of momentum profits between FM regressions with and without tail coskewness. The dashed line represents the lower confidence bound for the 95% confidence interval associated with remaining profits. The left (right) graph represents results for size/momentum portfolios with a ranking period of 5 months (11 months), and different holding periods of 1, 3, 6, 9 or 12 months for the momentum component.

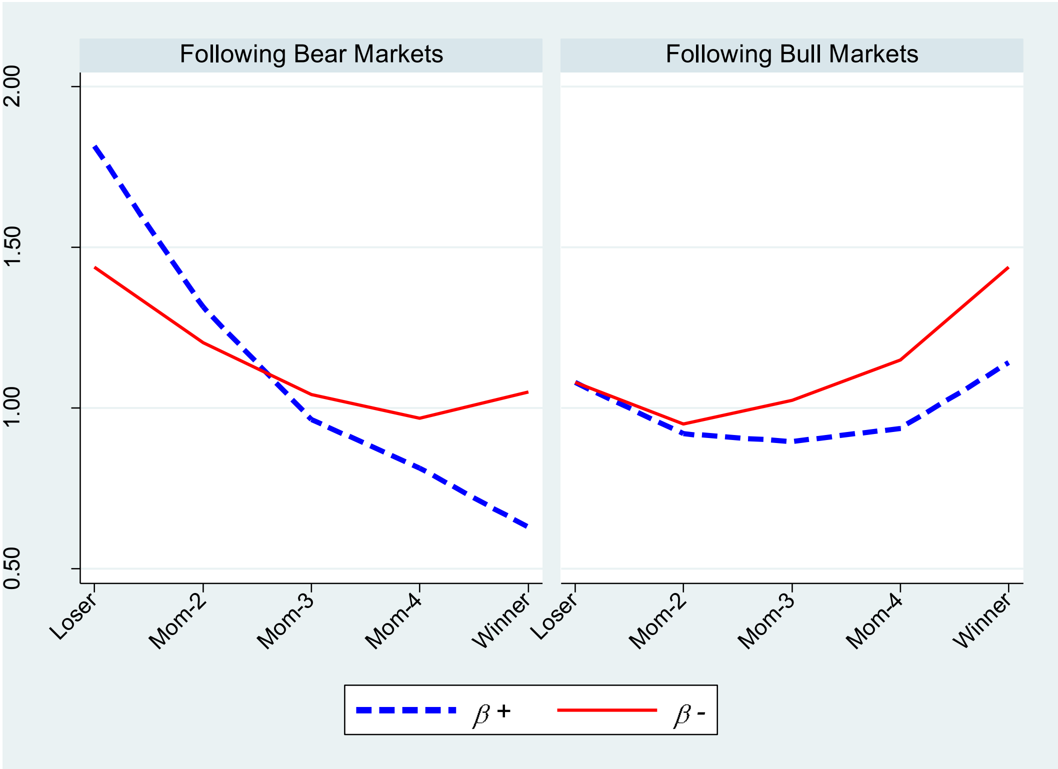

Figure 3.3 Time-varying beta for average winners and average losers

This figure shows the upside beta (β^{+}) and downside beta (β^{-}) for average winners and average losers for 25 value-weighted size/momentum portfolios following bull and bear markets, respectively. The return to the average loser portfolio is the average return to the loser portfolios within each of the five size quintiles. The return to the average winner portfolio is the average return to the winner portfolios within each of the five size quintiles. β^{+} is the upside beta in equation (8), and β^{-} is the downside beta in equation (9). A bull market occurs when the market cumulative prior return over months t-12 to t-2, is positive, and is a bear market otherwise.

Table 3.1 Summary statistics for 25 value-weighted size/momentum portfolios (July 1963 to June 2013)

This table summarizes the 25 value-weighted size/momentum portfolios, incorporating all CRSP firms in the US listed on the NYSE, AMEX, or NASDAQ. The sample period of July 1963 through June 2013 yields 600 observations. The 25 equal-weighted size/momentum portfolios are constructed monthly from the intersection of five size portfolios with five portfolios formed on prior (2-12) returns. The monthly size and prior (2-12) return breakpoints are the corresponding NYSE quintile values.
Mean excess return is the average monthly return of 25 size/momentum portfolios in excess of the one-month Treasury bill rate. CAPM betas are the average slopes from univariate regressions of the portfolio excess returns on the excess return to the value-weighted market index. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return. Standardized CS_i^S is the standardized coskewness defined in equation (1), computed with orthogonalized residuals from the CAPM model. Standardized $TCS_i^{S,\alpha=0.1}$ is the standardized tail coskewness defined in equation (2) with a two-tailed α=0.10. Non-standardized CS_i^S is the non-standardized coskewness defined in equation (3), computed with equity returns directly. Non-standardized $TCS_i^{N,\alpha=0.1}$ is the non-standardized tail coskewness defined in equation (4) with a two-tailed α=0.10. Coskewness is significant at the 10% level and at the 5% level if its absolute value exceeds 0.16 and 0.20 respectively based on a simulation process using a bivariate normal distribution with a correlation coefficient of 0.5.

Panel A: Time-series average characteristics

	Mean Excess return	lnME	PR	CAPM β	Standardized		Non-standardized	
					CS_i^S	$TCS_i^{S,\alpha=0.1}$	CS_i^N	$TCS_i^{N,\alpha=0.1}$
Size small								
Loser	-0.001	3.123	-0.299	1.365	-0.029	0.050	-0.415**	-0.379**
Mom-2	0.653	3.372	-0.041	1.052	-0.157	-0.053	-0.511**	-0.472**
Mom-3	0.890	3.430	0.093	0.987	-0.258**	-0.142	-0.574**	-0.525**
Mom-4	1.032	3.461	0.244	1.004	-0.327**	-0.272**	-0.615**	-0.574**
Winner	1.388	3.475	0.836	1.218	-0.307**	-0.172*	-0.602**	-0.595**
Size 2								
Loser	0.112	4.835	-0.263	1.454	0.050	0.045	-0.402**	-0.386**
Mom-2	0.636	4.856	-0.039	1.119	-0.057	-0.029	-0.473**	-0.458**
Mom-3	0.787	4.862	0.094	1.027	-0.180**	-0.097	-0.543**	-0.503**
Mom-4	0.991	4.865	0.244	1.056	-0.278**	-0.213**	-0.591**	-0.569**
Winner	1.213	4.862	0.787	1.286	-0.271**	-0.152	-0.586**	-0.526**
Size 3								
Loser	0.237	5.672	-0.245	1.367	0.173*	0.223**	-0.332**	-0.321**
Mom-2	0.585	5.692	-0.037	1.097	0.044	-0.003	-0.441**	-0.389**
Mom-3	0.684	5.698	0.094	1.012	-0.095	-0.023	-0.509**	-0.444**
Mom-4	0.767	5.697	0.243	1.005	-0.311**	-0.212**	-0.605**	-0.540**
Winner	1.187	5.683	0.730	1.225	-0.242**	-0.173**	-0.573**	-0.495**
Size 4								
Loser	0.161	6.567	-0.226	1.338	0.231**	0.260**	-0.294**	-0.260**
Mom-2	0.563	6.591	-0.036	1.105	0.131	0.134	-0.405**	-0.386**
Mom-3	0.624	6.594	0.095	0.998	0.002	0.032	-0.476**	-0.430**
Mom-4	0.771	6.592	0.243	0.996	-0.069	0.026	-0.510**	-0.464**
Winner	1.024	6.583	0.680	1.151	-0.261**	-0.086	-0.583**	-0.524**
Size large								
Loser	0.116	8.185	-0.199	1.240	0.268**	0.186*	-0.268**	-0.257**
Mom-2	0.432	8.386	-0.034	0.934	0.327**	0.227**	-0.272**	-0.244**
Mom-3	0.349	8.422	0.096	0.900	0.006	0.003	-0.477**	-0.440**
Mom-4	0.520	8.461	0.242	0.883	0.088	0.102	-0.441**	-0.385**
Winner	0.741	8.324	0.564	1.026	-0.085	-0.069	-0.496**	-0.398**

** and * denote t-statistics significant at the 5 percent and 10 percent levels, respectively.

Panel B: Cross-sectional correlation									
		Mean Excess return	lnME	PR	CAPM beta	Standardized		Non-standardized	
						CS_i^S	$TCS_i^{S,\alpha=0.1}$	CS_i^N	$TCS_i^{N,\alpha=0.1}$
Mean		1							
lnME		-0.263	1						
PR		0.719	0.006	1					
Beta		-0.277	-0.277	-0.139	1				
Standardized	CS_i^S	-0.807	0.506	-0.545	0.149	1			
	$TCS_i^{S,\alpha=0.1}$	-0.796	0.448	-0.526	0.232	0.964	1		
Non-standardized	CS_i^N	-0.837	0.412	-0.572	0.289	0.973	0.949	1	
	$TCS_i^{N,\alpha=0.1}$	-0.821	0.481	-0.524	0.232	0.961	0.933	0.979	1

Table 3.2 Fama-MacBeth regressions on measures of standardized coskewness and tail coskewness

The table shows the average slopes and their t-statistics (in parentheses) from monthly cross-sectional regressions to predict excess returns to 25 value-weighted size/momentum portfolios using measures of standardized coskewness and tail coskewness (defined in Equations [1] and [2]), for July 1963 to June 2013. Two-tailed α (alpha) represents the total proportion of observations from the left and right tails of the residual distribution that are excluded in calculating standardized coskewness and that are included in calculating standardized tail coskewness and ranges from 0.02 to 0.12. Residuals are from a regression of an asset's monthly excess returns on the contemporaneous market excess return. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return; _cons is the average intercept. $TCS^S(CS^S)$ risk premium is the change in excess return corresponding to a two standard deviation change in standardized tail coskewness and coskewness measures across 25 size/momentum portfolios, calculated as twice the product of the average slope on TCS^S (CS^S) from FM regressions with the average cross-sectional standard deviation of standardized TCS (CS) calculated over the 25 size/momentum portfolios.

Two-tailed α(alpha) Trimmed	(1) N/A	(2) α=0.00	(3) α=0.02	(4) α=0.04	(5) α=0.06	(6) α=0.08	(7) α=0.10	(8) α=0.12
Panel A: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 CS_i^S + e_i$								
lnME	-0.060*	-0.019	-0.044	-0.052	-0.053	-0.062*	-0.059	-0.059
	(-1.67)	(-0.42)	(-1.09)	(-1.33)	(-1.37)	(-1.65)	(-1.55)	(-1.59)
PR	0.951***	0.747***	0.835***	0.893***	0.885***	0.966***	0.947***	0.946***
	(4.25)	(4.05)	(4.32)	(4.44)	(4.35)	(4.52)	(4.44)	(4.49)
CS^S		-0.591*	-0.500	-0.265	-0.241	0.049	-0.055	-0.042
		(-1.89)	(-1.58)	(-1.11)	(-0.97)	(0.28)	(-0.27)	(-0.24)
_cons	0.763**	0.565	0.695*	0.745*	0.749*	0.793**	0.769*	0.773**
	(2.02)	(1.20)	(1.65)	(1.84)	(1.86)	(2.03)	(1.95)	(1.97)
CS^S Risk Premium		*0.225*	*0.122*	*0.065*	*0.056*	*-0.011*	*0.012*	*0.01*
Panel B: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 TCS_i^S + e_i$								
lnME			-0.044	-0.040	-0.028	-0.020	-0.024	-0.020
			(-1.16)	(-1.03)	(-0.66)	(-0.46)	(-0.56)	(-0.46)
PR			0.875***	0.848***	0.797***	0.738***	0.741***	0.726***
			(4.15)	(4.07)	(4.04)	(3.89)	(3.96)	(3.83)
TCS^S			-0.481**	-0.533**	-0.769**	-0.873***	-0.812**	-0.836**
			(-2.24)	(-2.41)	(-2.32)	(-2.65)	(-2.42)	(-2.47)
_cons			0.673*	0.650	0.605	0.567	0.603	0.580

	(1.68)	(1.60)	(1.41)	(1.30)	(1.38)	(1.31)
TCS^S Risk Premium	*0.116*	*0.144*	*0.204*	*0.249*	*0.235*	*0.249*
Panel C: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 CS_i^S + \lambda_4 TCS_i^S + e_i$						
lnME	-0.022	-0.023	-0.023	-0.023	-0.027	-0.023
	(-0.48)	(-0.51)	(-0.52)	(-0.52)	(-0.60)	(-0.50)
PR	0.736***	0.751***	0.755***	0.772***	0.766***	0.746***
	(4.06)	(4.18)	(4.16)	(4.22)	(4.18)	(4.10)
CS^S	-0.558	-0.396	-0.175	0.068	0.041	0.016
	(-1.53)	(-1.28)	(-0.70)	(0.38)	(0.21)	(0.09)
TCS^S	-0.531*	-0.581**	-0.708**	-0.808**	-0.752**	-0.773**
	(-1.96)	(-2.02)	(-2.12)	(-2.41)	(-2.27)	(-2.28)
_cons	0.598	0.617	0.621	0.622	0.639	0.616
	(1.26)	(1.32)	(1.35)	(1.36)	(1.40)	(1.34)

Table 3.3 Fama-MacBeth regressions on measures of non-standardized coskewness and tail coskewness

The table shows the average slopes and their t-statistics (in parentheses) from monthly cross-sectional regressions to predict excess returns to 25 value-weighted size/momentum portfolios using measures of non-standardized coskewness and tail coskewness (defined in equations [1] and [2]) for July 1963 to June 2013. Two-tailed α (alpha) trimmed represents the proportion of total observations from the left and right tails of the return distributions that have been discarded from the calculation of non-standardized coskewness and included in the calculation of non-standardized tail coskewness, and ranges from 0.02 to 0.12. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return. _cons is the average intercept. $TCS^N(CS^N)$ risk premium is the change in excess return corresponding to a two standard deviation change in standardized tail coskewness and coskewness measures across 25 size/momentum portfolios, calculated as twice the product of the average slope on TCS^N (CS^N) from FM regressions with the average cross-sectional standard deviation of non-standardized TCS (CS) calculated over the 25 size/momentum portfolios.

Two-tailed α(alpha) Trimmed	(1) a=0.00	(2) a=0.02	(3) a=0.04	(4) a=0.06	(5) a=0.08	(6) a=0.10	(7) a=0.12
			Panel A: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 CS_i^N + e_i$				
LnME	-0.023	-0.058	-0.061*	-0.068*	-0.065*	-0.062*	-0.061*
	(-0.52)	(-1.58)	(-1.69)	(-1.94)	(-1.86)	(-1.75)	(-1.69)
PR	0.709***	0.946***	0.894***	0.817***	0.870***	0.944***	0.949***
	(3.80)	(4.30)	(4.55)	(4.87)	(4.72)	(5.03)	(4.42)
CS^N	-1.165*	-0.774	-0.900*	-1.423*	-1.121	-0.249	0.121
	(-1.80)	(-1.35)	(-1.76)	(-1.69)	(-1.64)	(-0.34)	(0.19)
_cons	0.088	0.629	0.718*	0.775*	0.759*	0.811**	0.781**
	(0.13)	(1.49)	(1.77)	(1.87)	(1.88)	(2.04)	(2.03)
CS^N Risk Premium	*0.244*	*0.058*	*0.101*	*0.159*	*0.106*	*0.02*	*-0.005*
			Panel B: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 TCS_i^N + e_i$				
lnME		-0.039	-0.026	-0.003	-0.010	-0.008	-0.012
		(-0.96)	(-0.60)	(-0.07)	(-0.23)	(-0.16)	(-0.26)
PR		0.793***	0.777***	0.722***	0.712***	0.668***	0.664***
		(3.93)	(3.66)	(3.46)	(3.46)	(3.39)	(3.53)
TCS^N		-0.768*	-1.130*	-1.685**	-1.494**	-1.601**	-1.481**
		(-1.81)	(-1.88)	(-2.27)	(-2.39)	(-2.49)	(-2.25)
_cons		0.419	0.149	-0.201	-0.086	-0.164	-0.065

	(0.82)	(0.24)	(-0.28)	(-0.13)	(-0.24)	(-0.09)
TCS^N Risk Premium	*0.167*	*0.202*	*0.288*	*0.28*	*0.313*	*0.308*
Panel C: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 CS_i^N + \lambda_4 TCS_i^N + e_i$						
lnME	-0.021	-0.026	-0.031	-0.023	-0.011	-0.011
	(-0.47)	(-0.58)	(-0.73)	(-0.54)	(-0.24)	(-0.25)
PR	0.702***	0.698***	0.709***	0.704***	0.712***	0.678***
	(3.86)	(3.77)	(3.88)	(3.77)	(3.95)	(3.63)
CS^N	-1.650*	-1.068*	-1.004	-0.764	0.287	0.536
	(-1.79)	(-1.87)	(-1.36)	(-1.26)	(0.46)	(0.92)
TCS^N	-1.124*	-1.088	-1.025*	-1.191**	-1.512**	-1.484**
	(-1.81)	(-1.65)	(-1.66)	(-2.11)	(-2.49)	(-2.28)
_cons	0.002	0.112	0.164	0.077	-0.061	-0.050
	(0.00)	(0.16)	(0.25)	(0.12)	(-0.09)	(-0.07)

Table 3.4 Fama-MacBeth regressions for 100 value-weighted size/momentum portfolios (July 1963 to June 2013)

The table shows the average slopes and their t-statistics (in parentheses) from the monthly cross-sectional regressions to predict excess returns to 100 value-weighted size/momentum portfolios, for July 1963 to June 2013. For the momentum component of the 100 size/momentum portfolios, we use different combinations of the ranking period R and the holding period H. While we keep updating the size decile monthly based on firm size at formation month (t=-1), momentum deciles are constructed with different combinations of R=5 or 11 months and H=1, 3, 6, 9 or 12 months. Following Jegadeesh and Titman (1993), we construct overlapping portfolios to increase the power of our tests. In any given month t, the portfolio monthly return is the simple average of H overlapping portfolios that update their momentum deciles in the preceding t-H months as well as in month t-1. TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10. Fraction of momentum profit explained is the portion of the average slope on past returns that is reduced after controlling for tail coskewness in FM regressions. TCS risk premium is the average change in excess returns corresponding to a two standard deviation change in TCS across 100 size/momentum portfolios, computed as twice the product of the average slope on TCS in FM regressions with the average cross-sectional standard deviation of TCS across 100 size/momentum portfolios.

Portfolio Holding Month	5-month ranking (6-2)					11-month ranking (12-2)				
	1	3	6	9	12	1	3	6	9	12
Panel A: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + e_i$										
lnME	-0.04	-0.039	-0.045	-0.048	-0.048	-0.056	-0.056	-0.058	-0.061	-0.061
	(-1.10)	(-1.09)	(-1.24)	(-1.32)	(-1.31)	(-1.48)	(-1.50)	(-1.56)	(-1.63)	(-1.63)
PR	0.957***	1.164***	1.911***	2.575***	2.500***	0.869***	0.824***	0.772***	0.717**	0.637**
	(3.51)	(3.69)	(4.50)	(5.20)	(4.71)	(3.85)	(3.49)	(3.00)	(2.57)	(2.05)
_cons	0.649*	0.615	0.596	0.589	0.61	0.706*	0.711*	0.742*	0.772**	0.772**
	(1.65)	(1.58)	(1.52)	(1.49)	(1.53)	(1.79)	(1.81)	(1.89)	(1.97)	(1.98)
Panel B: $r_i = \lambda_0 + \lambda_1 \ln ME_i + \lambda_2 PR_i + \lambda_3 TCS_i + e_i$										
lnME	-0.039	-0.033	-0.021	0.003	0.020	-0.028	-0.017	-0.001	0.003	-0.001
	(-0.98)	(-0.83)	(-0.52)	(0.07)	(0.44)	(-0.63)	(-0.37)	(-0.01)	(0.05)	(-0.01)
PR	0.887***	0.955***	1.137***	1.122**	0.855	0.676***	0.526***	0.389*	0.375	0.267
	(3.34)	(3.21)	(2.79)	(2.22)	(1.51)	(3.47)	(2.60)	(1.72)	(1.50)	(0.96)
TCS	-0.160	-0.456	-1.386***	-2.669***	-3.390***	-1.356***	-1.913***	-2.796***	-3.074***	-2.916***
	(-0.49)	(-1.23)	(-3.29)	(-4.65)	(-4.43)	(-2.76)	(-3.06)	(-3.50)	(-3.60)	(-3.32)
_cons	0.583	0.416	-0.063	-0.731	-1.129	0.058	-0.215	-0.666	-0.802	-0.706
	(1.18)	(0.81)	(-0.11)	(-1.16)	(-1.54)	(0.10)	(-0.32)	(-0.88)	(-1.04)	(-0.90)
TCS Risk Premium	*0.016*	*0.051*	*0.110*	*0.153*	*0.157*	*0.129*	*0.153*	*0.162*	*0.146*	*0.126*
Fraction of Momentum Profit Explained										
	0.073	*0.179*	*0.405*	*0.565*	*0.658*	*0.223*	*0.362*	*0.496*	*0.478*	*0.581*

Table 3.5 The relationship between portfolio returns and expected tail coskewness

Panel A reports time-series average slopes and the percentage of slopes that are statistically significant at the 5% level (in parentheses), from monthly cross-sectional regressions to predict tail coskewness (a first-stage regression):

$$TCS_{p,T} = a_0 + a_1 lnME_{p,T} + a_2 PR_{p,T} + a_3 \sigma_{p,T} + \omega_{p,T}$$

where $TCS_{p,T}$ is the tail coskewness estimated for portfolio p based on an expanding window, *T*, with initial sample of January 1945-June 1963 that expands by one month at a time; $lnME_{p,T}(PR_{p,T})$ is the time-series average of firm size (prior return) for portfolio p in the expanding window *T*; $\sigma_{p,T}$ is the time-series average of the standard deviation of daily returns to portfolio p, computed monthly, in the expanding window *T*. The daily returns and monthly returns to 25 size/momentum portfolios are from the Kenneth French data library. The slopes obtained from the first stage regression using the most recent expanding window (ending at month *t*) are then used in the calculation of expected tail coskewness:

$$E[TCS_{p,t+1}] = a_0 + a_1 lnME_{p,T} + a_2 PR_{p,T} + a_3 \sigma_{p,T}$$

where $E[TCS_{p,t+1}]$ is the expected tail coskewness of portfolio p for month *t+1*; $lnME_{p,T}$, $PR_{p,T}$ and $\sigma_{p,T}$ are time-series averaged values based on the most recent expanding window (ending at month *t*) .
Panel B reports time-series average slopes and their t-statistics (in parentheses) from monthly FM cross-sectional regressions to predict excess returns to 25 size/momentum portfolios, for July 1963 to June 2013,

$$r_{p,t+1} = \gamma_0 + \gamma_1 E[TCS_{p,t+1}] + \gamma_2 lnME_{p,t} + \gamma_3 PR_{p,t} + \varepsilon_{i,t+1}$$

where $r_{p,t+1}$ is the monthly excess return for portfolio *p*; $lnME_{p,t}$ and $PR_{p,t}$ are the values of firm size and prior returns for portfolio p as at the end of month *t*.

Panel A: First-stage regression predictors of tail coskewness

$PR_{p,T}$		$lnME_{p,T}$		$\sigma_{p,T}$		F-stat	R-square
Full sample: T={January 1945 - June 1963} To T={January 1945 - May 2013}							
-0.232	(100%)	0.019	(69.9%)	0.297	(73.2%)	33.309	0.712
1st subsample: T={January 1945 - June 1963} To T={January 1945 - May 1973}							
-0.156	(100%)	0.006	(13.3%)	0.195	(51.7%)	8.659	0.448
2nd subsample: T={January 1945 - June 1973} To T={January 1945 - May 1983}							
-0.188	(100%)	0.008	(35.8%)	0.118	(34.2%)	15.981	0.560
3rd subsample: T={January 1945 - June 1983} To T={January 1945 - May 1993}							
-0.309	(100%)	0.028	(100%)	0.367	(80.0%)	43.534	0.840
4th subsample: T={January 1945 - June 1993} To T={January 1945 - May 2003}							
-0.303	(100%)	0.030	(100%)	0.480	(100%)	49.718	0.858
5th subsample: T={January 1945 - June 2003} To T={January 1945 - May 2013}							
-0.205	(100%)	0.024	(100%)	0.327	(100%)	48.528	0.854

Panel B: Second-stage regression predictors of returns to size/momentum portfolios

Reg.	$E[TCS_{p,t+1}]$		$PR_{p,t}$		$lnME_{p,t}$		_cons		R-square
1			0.951***	(4.24)	-0.060*	(-1.67)	0.763**	(2.02)	0.472
2	-5.623***	(-5.65)					-0.828***	(-1.95)	0.224
3	-5.173**	(-2.66)	0.252	(0.87)	0.003	(0.06)	-0.837***	(-0.82)	0.565

Table 3.6 Time-varying beta for 25 value-weighted size/momentum portfolios (July 1963 to June 2013)

This table reports the risk characteristics of the 25 value-weighted size/momentum portfolios following bear (Panel A) and bull (Panel B) markets, respectively. A bull market occurs when the market cumulative prior return over months t-12 to t-2, is positive, and is a bear market otherwise. TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10. β is the average slope from univariate regressions of the portfolio excess returns on the excess return to the value-weighted market index. β^+ is the upside beta in equation (5), and β^- is the downside beta in equation (6). Upside extreme event# (Downside extreme event#) is defined as the number of monthly excess return observations which are 3.158 or more standard deviations above (below) the mean, which occurs in one out of every 600 observations for a normal distribution.

		Panel A: Following Bear Markets						Panel B: Following Bull Markets					
		TCS	β	β^+	β^-	Upside Extreme Event#	Downside Extreme Event#	TCS	β	β^+	β^-	Upside Extreme Event#	Downside Extreme Event#
Small	Loser	-0.091	1.717**	1.662**	1.450**	5	1	-0.328**	1.172**	1.013**	1.338**	0	1
	Mom-2	-0.144	1.201**	1.122**	1.226**	2	3	-0.381**	0.969**	0.804**	1.111**	1	2
	Mom-3	-0.163*	1.012**	0.846**	1.124**	2	2	-0.429**	0.977**	0.799**	1.161**	1	4
	Mom-4	-0.163*	0.876**	0.643**	1.043**	1	3	-0.473**	1.082**	0.845**	1.285**	1	4
	Winner	-0.173*	0.956**	0.544**	1.173**	0	1	-0.461**	1.373**	1.094**	1.594**	2	4
Size 2	Loser	-0.095	1.840**	1.828**	1.499**	3	2	-0.321**	1.238**	1.099**	1.285**	0	2
	Mom-2	-0.135	1.304**	1.287**	1.279**	2	2	-0.369**	1.018**	0.966**	1.082**	0	2
	Mom-3	-0.141	1.024**	0.968**	1.048**	2	2	-0.435**	1.032**	0.937**	1.135**	1	4
	Mom-4	-0.158	0.953**	0.805**	1.055**	1	2	-0.466**	1.119**	0.959**	1.267**	1	4
	Winner	-0.184*	1.028**	0.658**	1.195**	0	2	-0.456**	1.437**	1.163**	1.527**	2	4
Size 3	Loser	-0.092	1.735**	1.696**	1.354**	5	1	-0.271**	1.161**	1.147**	1.085**	0	1
	Mom-2	-0.119	1.286**	1.314**	1.138**	2	2	-0.356**	0.991**	0.971**	1.033**	0	2
	Mom-3	-0.141	1.096**	1.058**	1.096**	2	2	-0.409**	0.968**	0.876**	1.020**	0	3
	Mom-4	-0.149	0.902**	0.827**	0.983**	1	1	-0.502**	1.068**	0.904**	1.206**	0	4
	Winner	-0.170*	0.948**	0.696**	1.060**	0	1	-0.472**	1.386**	1.188**	1.470**	1	4
Size 4	Loser	-0.115	1.825**	1.859**	1.447**	4	2	-0.214**	1.060**	1.052**	0.920**	0	1
	Mom-2	-0.112	1.361**	1.416**	1.243**	3	2	-0.333**	0.962**	0.987**	0.921**	0	2
	Mom-3	-0.146	1.079**	1.027**	1.023**	2	2	-0.382**	0.955**	0.952**	0.938**	0	2
	Mom-4	-0.131	0.937**	0.942**	0.970**	1	1	-0.440**	1.033**	0.984**	1.051**	1	4
	Winner	-0.198*	0.890**	0.637**	1.053**	0	1	-0.464**	1.301**	1.156**	1.359**	1	4
Large	Loser	-0.119	1.735**	2.028**	1.446**	5	4	-0.187*	0.952**	1.081**	0.779**	0	0
	Mom-2	-0.100	1.212**	1.434**	1.129**	2	3	-0.212**	0.777**	0.874**	0.610**	0	0
	Mom-3	-0.148	0.955**	0.929**	0.923**	1	2	-0.400**	0.870**	0.919**	0.872**	0	2
	Mom-4	-0.121	0.766**	0.847**	0.791**	1	1	-0.409**	0.950**	0.994**	0.940**	2	3
	Winner	-0.140	0.736**	0.615**	0.770**	0	0	-0.448**	1.193**	1.106**	1.244**	2	4

** and * denote t-statistics significant at the 5 percent and 10 percent levels, respectively.

Table 3.7 Time-varying beta for 25 value-weighted size/momentum portfolios (July 1963 to June 2013)

Panel A reports the asymmetric betas, $\beta^{+} - \beta^{-}$, which are the difference between upside and downside betas as defined in equations (5) and (6), for our 25 size/momentum portfolios. Panel B shows the average slopes and their t-statistics from the monthly cross-sectional regressions to explain excess returns to 25 size/momentum portfolios. CS is the non-standardized coskewness defined in equation (3). TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return; _cons is the average intercept from the monthly cross-sectional regressions.

Panel A: Asymmetric beta ($\beta^{+} - \beta^{-}$)

	Loser	2	3	4	Winner	All
Small	-0.020	-0.178	-0.282	-0.381	-0.502	-0.272
2	0.087	-0.004	-0.115	-0.257	-0.415	-0.141
3	0.232	0.073	-0.062	-0.225	-0.312	-0.059
4	0.277	0.164	0.045	-0.028	-0.304	0.031
Big	0.406	0.313	0.047	0.044	-0.171	0.128
All	0.196	0.073	-0.073	-0.169	-0.341	-0.063

Panel B: FM regressions

Regression	1	2	3	4	5	6	7
($\beta^{+} - \beta^{-}$)			-1.329***			-0.177	1.307***
			(-4.57)			(-0.42)	(2.78)
CS		-2.932***			-1.163*		
		(-4.62)			(-1.80)		
TCS	-3.144***			-1.601**			-2.771***
	(-4.87)			(-2.49)			(-3.45)
lnME				-0.008	-0.023	-0.045	-0.070
				(-0.16)	(-0.52)	(-0.87)	(-1.37)
PR				0.668***	0.709***	0.976***	1.197***
				(3.39)	(3.80)	(4.24)	(5.40)
_cons	-0.749**	-0.770*	0.575**	-0.164	0.088	0.811	-0.184
	(-2.02)	(-1.92)	(2.41)	(-0.24)	(0.13)	(1.55)	(-0.26)

Table 3.8 Sub-period and early period analysis using Fama-MacBeth regressions for 25 value-weighted size/momentum portfolios

The table shows the average slopes and their t-statistics from the monthly cross-section regressions to predict excess returns to 25 size/momentum portfolios, for two separate sub-periods (July 1963-June 1983 and July 1983 to June 2013) and one early sub-period (January 1927 to June 1963). The asymmetric beta $\beta^{+} - \beta^{-}$ is the difference between the upside beta β^{+} in equation (5) and the downside beta β^{-} in equation (6). CS is the non-standardized coskewness defined in equation (3). TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return; _cons is the average intercept from the cross-sectional regressions.

	July 1983 to June2013 (360 Months)				July 1963 to June1983 (240 Months)				January 1927 to June 1963 (438Months)			
$(\beta^{+} - \beta^{-})$				-0.252				0.337				0.540*
				(-0.47)				(0.96)				(1.96)
CS			-0.892				-0.974*				-0.496**	
			(-1.36)				(-1.96)				(-2.12)	
TCS		-1.140*				-0.944**				-0.482**		
		(-1.71)				(-2.07)				(-2.03)		
lnME	-0.002	0.030	0.027	0.020	-0.148**	-0.110	-0.115	-0.173**	-0.180**	-0.210***	-0.210***	-0.116**
	(-0.05)	(0.62)	(0.55)	(0.31)	(-2.25)	(-1.50)	(-1.63)	(-2.54)	(-2.46)	(-2.73)	(-2.75)	(-2.10)
PR	0.607*	0.430*	0.446*	0.612***	1.468***	1.251***	1.193***	1.732***	0.914*	0.616	0.612	1.493***
	(1.95)	(1.74)	(1.87)	(2.73)	(5.38)	(4.39)	(4.63)	(5.58)	(1.88)	(1.00)	(1.01)	(3.92)
_cons	0.489	-0.368	-0.223	0.468	1.173*	0.997	0.992	1.422**	1.256**	1.529**	1.543**	1.027**
	(1.07)	(-0.39)	(-0.23)	(0.66)	(1.80)	(1.42)	(1.43)	(2.11)	(2.22)	(2.48)	(2.50)	(2.16)

Table 3.9 Fama-MacBeth regressions for 25 size/book-to-market, 25 size/operating profitability, 25 size/investment and 25 size/industry momentum value-weighted portfolios

The table shows the average slopes and their t-statistics (in parentheses) from monthly cross-sectional regressions to predict excess returns to 25 size/book-to-market portfolios (regressions 1&2), 25 size/operating profitability portfolios (regressions 3&4), 25 size/investment portfolios (regressions 5&6), and 25 size/industry momentum portfolios (regressions 7&8). The first three sets of portfolios are value-weighted and taken from the Kenneth French Data Library, and the 25 size/industry momentum portfolios are the value-weighted intersection of five portfolios formed on size and five portfolios formed on prior industry returns (using a 6-month ranking period and a 6 month holding period following Moskowitz and Grinblatt: 1999). The sample period is July 1963 to June 2013. lnME is the logarithm of average firm size; lnBM is the logarithm of the book-to-market ratio; OP is operating profitability; INV is annual growth in total assets; PR is the 6-month cumulative industry return to which the stock belongs; TCS is the non-standardized tail coskewness defined in equation (4), using a two-tailed α=0.10.

	25 size/lnBM		25 size/profitability		25 size/investment		25 size/industry momentum	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lnME	-0.056	-0.020	-0.080**	-0.038	-0.065	-0.027	-0.051	-0.031
	(-1.40)	(-0.43)	(-2.02)	(-0.90)	(-1.49)	(-0.54)	(-1.51)	(-0.87)
lnBM	0.241***	0.216**	-0.002	0.007	0.144	0.125		
	(2.77)	(2.54)	(-0.02)	(0.06)	(0.89)	(0.76)		
OP			0.443**	0.440**	0.049	0.013		
			(2.07)	(2.01)	(0.16)	(0.04)		
INV			-1.387**	-1.445***	-0.504***	-0.540***		
			(-2.58)	(-2.64)	(-2.62)	(-2.81)		
PR							2.338***	1.877***
							(3.44)	(2.89)
TCS		-1.667***		-1.763**		-1.750***		-1.382**
		(-2.78)		(-2.58)		(-2.74)		(-2.15)
_cons	1.084***	0.060	1.239***	0.128	1.141***	0.069	0.697**	0.255
	(2.72)	(0.10)	(3.17)	(0.22)	(2.64)	(0.11)	(1.97)	(0.66)
r2	0.460	0.504	0.484	0.509	0.497	0.530	0.476	0.548

Table 3.10 Fama-MacBeth regressions for 25 international value-weighted size/momentum portfolios (November 1990 to March 2014)

The table shows the average slopes and their t-statistics from monthly cross-section regressions to predict excess returns to 25 international size/momentum portfolios. This dataset for developed markets, for November 1990 to March 2014, is taken from the Kenneth French data library. The 23 developed markets are combined into four regions: (i) North America, which includes the United States and Canada; (ii) Japan; (iii) Asia Pacific, including Australia, New Zealand, Hong Kong, and Singapore (but not Japan) and (iv) Europe, including Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. The size breakpoints for a region are the 3rd, 7th, 13th, and 25th percentiles of that region's aggregate market capitalization. For portfolios formed at the end of month t–1, the prior returns are a stock's cumulative return for months t–12 to t–2. The momentum breakpoints for all size quintiles in a region are the 20th, 40th, 60th, and 80th percentiles of the prior return for big (top 90% by market cap) stocks in that region. The US one month T-bill rate is used as the risk-free rate; lnME is the logarithm of average firm size; TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10; _cons is the average intercept from our cross-sectional regressions.

	Global		North America		Europe		Japan		Asia Pacific	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
lnME	-0.067**	0.002	-0.094**	-0.048	-0.011	0.054	-0.062	-0.055	-0.028	0.026
	(-2.21)	(0.05)	(-2.39)	(-1.04)	(-0.37)	(1.62)	(-1.00)	(-0.87)	(-0.57)	(0.51)
TCS		-4.972***		-2.446***		-4.875***		-0.757		-1.647**
		(-4.12)		(-2.68)		(-4.82)		(-0.52)		(-2.21)
_cons	1.093**	-2.322**	1.622***	-0.076	0.782	-2.372**	0.448	0.626	1.060	0.161
	(2.49)	(-2.09)	(3.06)	(-0.07)	(1.64)	(-2.51)	(0.65)	(0.94)	(1.60)	(0.20)

Table 3. 11 Fama-MacBeth regressions for 25 UK value-weighted size/momentum portfolios (July 1975 to June 2014)

The table shows the average slopes and their t-statistics from the monthly cross-section regressions to predict excess returns to 25 UK value-weighted size/momentum portfolios, taken from the Kenneth French data library. This dataset incorporates all stocks traded on the London Stock Exchange from July 1975 to June 2014. The size breakpoints are the market caps for NYSE quintiles. For portfolios formed at the end of month t–1, the lagged momentum return is a stock's cumulative return for months t–12 to t–2. The momentum breakpoints are the 20th, 40th, 60th, and 80th percentiles for lagged momentum based on all stocks. The UK three-month Treasury bill rate (from The Bank of England) is the risk-free rate. Asymmetric beta, $\beta^{+} - \beta^{-}$, is the difference between the upside beta, β^{+}, in equation (5) and the downside beta, β^{-}, in equation (6). CS is the non-standardized coskewness defined in equation (3). TCS is the non-standardized tail coskewness defined in equation (4), with a two-tailed α=0.10. lnME is the logarithm of average firm size. PR is the cumulative prior (2-12) return; _cons is the average intercept from our cross-sectional regressions.

Regression	1	2	3	4	5	6	7
$(\beta^{+} - \beta^{-})$			-2.669***				-0.929*
			(-4.49)				(-1.86)
CS		-3.619***				-2.399***	
		(-5.27)				(-4.20)	
TCS	-3.899***				-2.483***		
	(-5.93)				(-4.58)		
lnME				0.002	0.003	0.006	0.017
				(0.07)	(0.13)	(0.26)	(0.75)
PR				1.102***	0.572	0.559	0.742
				(2.71)	(1.26)	(1.23)	(1.47)
_cons	0.009	0.151	0.128	0.009	-0.113	-0.022	-0.130
	(0.04)	(0.66)	(0.56)	(0.04)	(-0.44)	(-0.08)	(-0.52)