Downside Risk and the Momentum Effect*

Andrew Ang
Columbia University and NBER

Joseph Chen
University of Southern California

Yuhang Xing
Columbia University

First Version: 31 Aug, 2001
This Version: 14 Feb, 2002

JEL Classification: C12, C15, C32, G12
Keywords: asymmetric risk, cross-sectional asset pricing, downside correlation, downside risk, momentum effect

*The authors would like to thank Brad Barber, Alon Brav, Geert Bekaert, John Cochrane, Randy Cohen, Kent Daniel, Bob Dittmar, Rob Engle, Cam Harvey, David Hirschleiffer, Qing Li, Terence Lim, Bob Stambaugh, Akhtar Siddique and Zhenyu Wang. We especially thank Bob Hodrick for detailed comments. We thank seminar participants at Columbia University, NYU and USC, and at the Five Star Conference, for helpful comments. This paper is funded by a Q-Group research grant. Andrew Ang: aa610@columbia.edu; Joe Chen: joe.chen@marshall.usc.edu; Yuhang Xing: yx35@columbia.edu.

Abstract

Stocks with greater downside risk, which is measured by higher correlations conditional on downside moves of the market, have higher returns. After controlling for the market beta, the size effect and the book-to-market effect, the average rate of return on stocks with the greatest downside risk exceeds the average rate of return on stocks with the least downside risk by 6.55% per annum. Downside risk is important for explaining the cross-section of expected returns. In particular, we find that some of the profitability of investing in momentum strategies can be explained as compensation for bearing high exposure to downside risk.

1 Introduction

We define “downside risk” to be the risk that an asset’s return is highly correlated with the market when the market is declining. In this article, we show that there are systematic variations in the cross-section of stock returns that are linked to downside risk. Stocks with higher downside risk have higher expected returns, which cannot be explained by the market beta, the size effect or the book-to-market effect. In particular, we find that high returns associated with the momentum strategies (Jegadeesh and Titman, 1993) are sensitive to the fluctuations in downside risk.

Markowitz (1959) raises the possibility that agents care about downside risk, rather than about the market risk. He advises constructing portfolios based on semi-variances, rather than on variances, since semi-variances weight upside risk (gains) and downside risk (losses) differently. In Kahneman and Tversky (1979)’s loss aversion and Gul (1991)’s first-order risk aversion utility, losses are weighted more heavily than gains in an investor’s utility function. If investors dislike downside risk, then an asset with greater downside risk is not as desirable as, and should have a higher expected return than, an asset with lower downside risk. We find that stocks with highly correlated movements on the downside have higher expected returns. The portfolio of greatest downside risk stocks outperforms the portfolio of lowest downside risk stocks by 4.91% per annum. After controlling for the market beta, the size effect and the book-to-market effect, the greatest downside risk portfolio outperforms the lowest downside risk portfolio by 6.55% per annum.

It is not surprising that higher-order moments play a role in explaining the cross-sectional variation of returns. However, which higher-order moments are important for cross-sectional pricing is still a subject of debate. Unlike traditional measures of centered higher-order moments, our downside risk measure emphasizes the asymmetric effect of risk across upside and downside movements (Ang and Chen, 2001). We find little discernable pattern in the expected returns of stocks ranked by third-order moments (Rubinstein, 1973; Kraus and Litzenberger, 1976; Harvey and Siddique, 2000), by fourth-order moments (Dittmar, 2001) by downside betas, or by upside betas (Bawa and Lindenberg, 1977).

We also find that the profitability of the momentum strategies is related to downside risk. While Fama and French (1996) and Grundy and Martin (2001) find that controlling for the market, the size effect, and the book-to-market effect increases the profitability of momentum strategies, rather than explaining it, the momentum portfolios load positively on a factor that reflects downside risk. A linear two-factor model with the market and this downside risk factor explains some of the cross-sectional return variations among momentum portfolios. The downside risk factor commands a significantly positive risk premium in both Fama-MacBeth

(1973) and Generalized Method of Moments (GMM) estimations and retains its statistical significance in the presence of the Fama-French factors. Although our linear factor models with downside risk are rejected using the Hansen-Jagannathan (1997) distance metric, our results suggest that some portion of momentum profits can be attributed as compensation for exposures to downside risk. Past winner stocks have high returns, in part, because during periods when the market experiences downside moves, winner stocks move down more with the market than past loser stocks.

Existing explanations of the momentum effect are largely behavioral in nature and use models with imperfect formation and updating of investors' expectations in response to new information (Barberis, Shleifer and Vishny, 1998; Daniel, Hirshleifer and Subrahmanyam, 1998; Hong and Stein, 1999). These explanations rely on the assumption that arbitrage is limited, so that arbitrageurs cannot eliminate the apparent profitability of momentum strategies. Mispricing may persist because arbitrageurs need to bear some undiversifiable factor risk, and risk-averse arbitrageurs demand compensation for accepting such risk (Hirshleifer, 2001). In particular, Jegadeesh and Titman (2001) show that the momentum effect has persisted since its discovery. We show that momentum strategies have high exposures to a systematic downside risk factor.

Our findings are closely related to Harvey and Siddique (2000), who argue that skewness is priced, and show that momentum strategies are negatively skewed. In our data sample, we fail to find any pattern relating past skewness to expected returns. Our findings are also related to DeBondt and Thaler (1987) who find that past winner stocks have greater downside betas than upside betas. Though the profitability of momentum strategies is related to asymmetries in risk, we find little systematic effect in the cross-section of expected returns relating to downside betas. Instead, we find that it is downside correlation which is priced.

While Chordia and Shivakumar (2000) try to account for momentum with a factor model, where the factor betas vary over time as a linear function of instrumental variables, they do not estimate this model using cross-sectional methods. Ahn, Conrad and Dittmar (2001) find that imposing these constraints reduces the profitability of momentum strategies. Ghysels (1998) also argues against time-varying beta models, showing that linear factor models with constant risk premia, like the models we estimate, perform better in small samples. Hodrick and Zhang (2001) also find that models that allow betas to be a function of business cycle instruments perform poorly, and they find substantial instabilities in such models.[1]

Our research design follows the custom of constructing and adding factors to explain devi-

[1] An alternative non-behavioral explanation for momentum is proposed by Conrad and Kaul (1998), who argue that the momentum effect is due to cross-sectional variations in (constant) expected returns. Jegadeesh and Titman (2001) reject this explanation.

ations from the Capital Asset Pricing Model (CAPM). However, this approach does not speak to the source of factor risk premia. Although we design our factor to measure an economically meaningful concept of downside risk, our goal is not to present a theoretical model that explains how downside risk arises in equilibrium. Our goal is to test whether a part of the factor structure in stock returns is attributable to downside risk. Other authors use factors which reflect the size and the book-to-market effects (Fama and French, 1993 and 1996), macroeconomic factors (Chen, Roll and Ross, 1986), production factors (Cochrane, 1996), labor income (Jagannathan and Wang, 1996), market microstructure factors like volume (Gervais, Kaniel and Mingelgrin, 2001) or liquidity (Pástor and Stambaugh, 2001), and factors motivated from corporate finance theory (Lamont, Polk and Saá-Requejo, 2001). Momentum strategies do not load very positively on any of these factors, nor do any these approaches use a factor which reflects downside risk.

The rest of this paper is organized as follows. Section 2 investigates the relationship between past higher-order moments and expected returns. We show that portfolios sorted by increasing downside correlations have increasing expected returns. On the other hand, portfolios sorted by other higher moments do not display any discernable pattern in their expected returns. Section 3 details the construction of our downside risk factor, shows that it commands an economically significant risk premium, and shows that it is not subsumed by the Fama and French (1993) factors. We apply the downside risk factor to price the momentum portfolios in Section 4 and find that the downside risk factor is significantly priced by the momentum portfolios. Section 5 studies the relation between downside risk and liquidity risk, and explores if the downside risk factor reflects information about future macroeconomic conditions. Section 6 concludes.

2 Higher-Order Moments and Expected Returns

Economic theory predicts that the expected return of an asset is linked to higher-order moments of the asset's return through the preferences of a marginal investor. The standard Euler equation in an arbitrage-free economy is:

$$\mathrm{E}_t[m_{t+1} r_{i,t+1}] = 0, \tag{1}$$

where m_{t+1} is the pricing kernel or the stochastic discount factor, and $r_{i,t+1}$ is the excess return on asset i. If we assume that consumption and wealth are equivalent then the pricing kernel is the marginal rate of substitution for the marginal investor: $m_{t+1} = U'(W_{t+1})/U'(W_t)$. By taking a Taylor expansion of the marginal investor's utility function, U, we can write:

$$m_{t+1} = 1 + \frac{W_t U''}{U'} MKT_{t+1} + \frac{W_t^2 U'''}{2U'} MKT_{t+1}^2 + \ldots, \tag{2}$$

where MKT_{t+1} is the rate of return on the market portfolio, in excess of the risk-free rate.

The coefficient on MKT_{t+1} in equation (2), $W_t U''/U'$, corresponds to the relative risk aversion of the marginal investor. The coefficient on MKT_{t+1}^2 is studied by Kraus and Litzenberger (1976) and motivates Harvey and Siddique (2000)'s coskewness measure, where risk-averse investors prefer positively skewed assets to negatively skewed assets. Dittmar (2001) examines the cokurtosis coefficient on MKT_{t+1}^3 and argues that investors with decreasing absolute prudence dislike cokurtosis. Empirical research rejects the standard specifications for U, such as power utility, and leaves unanswered what the most appropriate representation for U is.

However, economic theory does not restrict the utility function U to be smooth. Both Kahneman and Tversky (1979)'s loss aversion utility and Gul (1991)'s first-order risk aversion utility function have a kink at the reference point to which an investor compares gains and losses. These asymmetric, kinked utility functions suggest that polynomial expansions of U, such as the expansion used by Bansal, Hsieh and Viswanathan (1993), may not be a good global approximations of U. In particular, standard polynomial expansions may miss asymmetric risk.

We show in Section 2.1 that there is a positive relation between downside risk and expected returns. Stocks with high downside conditional correlations, which condition on moves of the market below its mean, have higher returns than stocks with low downside conditional correlations. However, there is no reward nor cost for bearing risk on the upside. In Section 2.2 we show that stocks sorted by other higher-order moments have no discernable patterns in their expected returns. We also show that stocks sorted by conditional downside or upside betas have little discernable patterns in Section 2.3. We provide an interpretation of our results in Section 2.4.

2.1 Downside and Upside Correlations

In Table (1), we show that stocks with high downside risk with the market have higher expected returns than stocks with low downside risk. We measure downside risk and upside risk by downside conditional correlations, ρ^-, and upside conditional correlations, ρ^+, respectively. We define these conditional correlations as:

$$\begin{aligned} & \rho^- && = \text{corr}\left(r_{i,t}, MKT_t | MKT_t < \overline{MKT}_t\right) \\ \text{and} \quad & \rho^+ && = \text{corr}\left(r_{i,t}, MKT_t | MKT_t > \overline{MKT}_t\right), \end{aligned} \tag{3}$$

where $r_{i,t}$ is the excess stock return, MKT_t is the excess market return, and $\overline{MKT}_t$ is the mean excess market return.

To ensure that we do not capture the endogenous influence of contemporaneously high returns on higher-order moments, we form portfolios sorted by past return characteristics and examine portfolio returns over a future period. To sort stocks based on downside and upside correlations at a point in time, we calculate ρ^- and ρ^+ using daily continuously compounded excess returns over the previous year. We first rank stocks into deciles, and then we calculate the holding period return over the next month of the value-weighted portfolio of stocks in each decile. We rebalance these portfolios each month. Appendix A provides further details on portfolio construction.

Panels A and B of Table (1) list monthly summary statistics of the portfolios sorted by ρ^- and ρ^+, respectively. We first examine the ρ^- portfolios in Panel A. The first column lists the mean monthly holding period returns of each decile portfolio. Stocks with the highest past downside correlations have the highest returns. In contrast, stocks with the lowest past downside correlations have the lowest returns. Going from portfolio 1, which is the portfolio of lowest downside correlations, to portfolio 10 which is the portfolio of highest downside correlations, the average return almost monotonically increases. The return differential between the portfolios of the highest decile ρ^- stocks and the lowest decile ρ^- stocks is 4.91% per annum (0.40% per month). This difference is statistically significant at the 5% level (t-stat = 2.26), using Newey-West (1987) standard errors with 3 lags.

The remaining columns list other characteristics of the ρ^- portfolios. The portfolio of stocks with the highest past downside correlations have the lowest autocorrelations, but they also have the highest betas. Since the CAPM predicts that high beta stocks should have high expected returns, we investigate in Section 3 if the high returns of high ρ^- stocks are attributable to high post-formation period betas. However, high returns of high ρ^- stocks do not appear to be due to the size effect or the book-to-market effect. The columns labeled "Size" and "B/M" show that high ρ^- stocks tend to be large stocks and growth stocks. Size and book-to-market effects would predict high ρ^- stocks to have low returns rather than high returns.

The second to last column calculates the post-formation conditional downside correlation of each decile portfolio, over the whole sample. These post-formation period ρ^- are monotonically increasing, which indicates that the top decile portfolio, formed by taking stocks with the highest conditional downside correlation over the past year, is the portfolio with the highest downside correlation over the whole sample. This implies that using past ρ^- is a good predictor of future ρ^- and that downside correlations are persistent.

The last column lists the downside betas, β^-, of each decile portfolio. We define downside beta, β^-, and its upside counterpart, β^+ as:

$$\beta^- = \frac{\text{cov}(r_{i,t}, MKT_t | MKT_t < \overline{MKT}_t)}{\text{var}(MKT_t | MKT_t < \overline{MKT}_t)}$$
$$\text{and} \quad \beta^+ = \frac{\text{cov}(r_{i,t}, MKT_t | MKT_t > \overline{MKT}_t)}{\text{var}(MKT_t | MKT_t > \overline{MKT}_t)}. \tag{4}$$

The β^- column shows that the ρ^- portfolios have fairly flat β^- pattern. Hence, the higher returns to higher downside correlation is not due to higher downside beta exposure.

Panel B of Table (1) shows the summary statistics of stocks sorted by ρ^+. In contrast to stocks sorted by ρ^-, there is no discernable pattern between the mean returns and upside correlations. However, the patterns in the β's, market capitalizations and book-to-market ratios of stocks sorted by ρ^+ are similar to the patterns found in ρ^- sorts. In particular, high ρ^+ stocks also tend to have higher betas, tend to be large stocks, and tend to be growth stocks. The last two columns list the post-formation ρ^+ and β^+ statistics. Here, both ρ^+ and β^+ increase monotonically from decile 1 to 10, but portfolio sorts by ρ^+ do not produce any pattern in their expected returns.

In summary, Table (1) shows that assets with higher downside correlations have higher returns. This result is consistent with models in which the marginal investor is more risk-averse on the downside than on the upside, and demand higher expected returns for bearing higher downside risk.

2.2 Coskewness and Cokurtosis

Table (2) shows that stocks sorted by past coskewness and past cokurtosis do not produce any discernable patterns in their expected returns. Following Harvey and Siddique (2000), we define coskewness as:

$$\text{coskew} = \frac{\text{E}[\epsilon_{i,t}\epsilon_{m,t}^2]}{\sqrt{\text{E}[\epsilon_{i,t}^2]}\text{E}[\epsilon_{m,t}^2]}, \tag{5}$$

where $\epsilon_{i,t} = r_{i,t} - \alpha_i - \beta_i MKT_t$, is the residual from the regression of $r_{i,t}$ on the contemporaneous excess market return, and $\epsilon_{m,t}$ is the residual from the regression of the market excess return on a constant.

Similar to the definition of coskewness in equation (5), we define cokurtosis as:

$$\text{cokurt} = \frac{\text{E}[\epsilon_{i,t}\epsilon_{m,t}^3]}{\sqrt{\text{E}[\epsilon_{i,t}^2]}\left(\text{E}[\epsilon_{m,t}^2]\right)^{\frac{3}{2}}}. \tag{6}$$

We compute coskewness in equation (5) and cokurtosis in equation (6) using daily data over the past year. Appendix B shows that calculating daily coskewness and cokurtosis is equivalent to calculating monthly, or any other frequency, coskewness and cokurtosis, assuming returns are drawn from infinitely divisible distributions.

Panel A of Table (2) lists the characteristics of stocks sorted by past coskewness. Like Harvey and Siddique (2000), we find that stocks with more negative coskewness have higher returns. However, the difference between the first and the tenth decile is only 1.79% per annum, which is not statistically significant at the 5% level (t-stat = 1.17). Stocks with large negative coskewness tend to have higher betas and there is little pattern in their post-formation unconditional coskewness. Panel B of Table (2) lists the summary statistics for the portfolios sorted by cokurtosis. In summary, we do not find any statistically significant reward for bearing cokurtosis risk.

We also perform (but do not report) sorts on skewness and kurtosis. We find that portfolios sorted on past skewness do have statistically significant pattern in expected returns, but the pattern is the opposite of that predicted by an investor with an Arrow-Pratt utility. Specifically, stocks with the most negative skewness have the lowest average returns. Moreover, skewness is not persistent in that stocks with high past skewness do not necessarily have high skewness in the future. Finally, we find that stocks sorted by kurtosis have no patterns in their expected returns.

2.3 Downside and Upside Betas

In Table (3), we sort stocks on the unconditional beta, the downside beta and the upside beta. Confirming many previous studies, Panel A shows that the beta does not explain the cross-section of stock returns. There is no pattern across the expected returns of the portfolio of stocks sorted on past β. The column labeled β also shows that the portfolios constructed by ranking stocks on past beta retain their beta-rankings in the post-formation period.

Panel B of Table (3) reports the summary statistics of stocks sorted by the downside beta, β^-. There is a weakly increasing, but mostly humped-shaped pattern in the expected returns of the β^- portfolios. However, the difference in the returns is not statistically significant. This is in contrast to the strong monotonic pattern we find across the expected returns of stocks sorted by downside correlation.

Both the downside beta and the downside correlation measure how an asset's return moves relative to the market's return, conditional on downside moves of the market. In order to analyze why the two measures produce different results, we perform the following decomposition. The downside beta is a function of the downside correlation and a ratio of the portfolio's downside volatility to the market's downside volatility:

$$\beta^- = \frac{\text{cov}(r_{i,t}, MKT_t | MKT_t < \overline{MKT}_t)}{\text{var}(MKT_t | MKT_t < \overline{MKT}_t)}$$
$$= \rho^- \times \frac{\sigma(r_{i,t} | MKT_t < \overline{MKT}_t)}{\sigma(r_{m,t} | MKT_t < \overline{MKT}_t)}. \tag{7}$$

We denote the ratio of the volatilities as $k^- = \sigma(r_{i,t}|MKT_t < \overline{MKT}_t)/\sigma(r_{m,t}|MKT_t < \overline{MKT}_t)$, conditioning on the downside, and a corresponding expression for k^+ for conditioning on the upside. This decomposition shows that an asset can have a high downside beta either because it has a high downside correlation or it has a high downside volatility.

The columns labeled β^- and ρ^- list summary post-formation β^- and ρ^- statistics of the decile portfolios over the whole sample. While Panel B of Table (3) shows that the post-formation β^- is monotonic for the β^- portfolios, this can be decomposed into non-monotonic effects for ρ^- and k^-. The downside correlation ρ^- increases and then decreases moving from the portfolio 1 to 10, while k^- decreases and then increases. The hump-shape in expected returns largely mirrors the hump-shape pattern in downside correlation. The two different effects of ρ^- and k^- make expected return patterns in β^- harder to detect than expected return patterns in ρ^-. In an unreported result, we find that portfolios of stocks sorted by k^- produce no discernable pattern in expected returns.

In contrast, Table (1) shows that portfolios sorted by increasing ρ^- have little pattern in the downside betas. Hence, variation in the expected returns of β^- portfolios is likely to be driven by their exposure to ρ^-. This observation is consistent with Ang and Chen (2001) who show that variations in downside beta are largely driven by variations in downside correlation. We find that sorting on downside correlation produces greater variations in returns than sorting on downside beta.

The last panel of Table (3) sorts stocks on β^+. The panel shows a relation between β^+ and ρ^+. However, just as with the lack of relation between ρ^+ and expected returns reported in Table (1), there is no pattern in the expected returns across the β^+ portfolios.

2.4 Summary and Interpretation

Stocks sorted by increasing downside risk, measured by conditional downside correlations, have increasing expected returns. In contrast, portfolios sorted by other centered higher-order moments (coskewness and cokurtosis) have little discernable patterns in returns. Although we calculate our measure conditional at a point in time and conditional on the mean market return at that time, the emphasis of the conditional downside correlation is on the asymmetry across the upside market moves and the downside market moves. If a marginal investor dislikes downside risk, why would the premium for bearing downside risk only appear in portfolios sorted by ρ^-, and not in other moments capturing left-hand tail exposure such as co-skewness? If the marginal investor's utility is kinked, skewness and other odd-centered moments may not effectively capture the asymmetric aversion to risk across upside and downside moves. On the other hand, downside correlation is a complicated function of many higher-ordered moments, including skewness, and therefore, downside correlation might serve as a better proxy for downside risk.

Downside correlation measures risk asymmetry and produces strong patterns in expected returns. However, portfolios formed by other measures of asymmetric risk, such as downside beta, do not produce strong cross-sectional differences in expected returns. One statistical reason is that downside beta involves downside correlation, plus a multiplicative effect from the ratios of volatilities, which masks the effect of downside correlation. Second, while the beta measures comovements in both the direction and the magnitude of an asset return and the market return, correlations are scaled to emphasize the comovements in only direction. Hence, our results suggest that while agents care about downside risk (a magnitude and direction effect), economic constraints which bind only on the downside (a direction effect only) are also important in producing the observed downside risk. For example, Chen, Hong and Stein (2001) examine binding short-sales constraints where the effect of a short sale constraint is a fixed cost rather than a proportional cost. Similarly, Kyle and Xiong (2001)'s wealth constraints only bind on the downside.

3 A Downside Correlation Factor

In this section, we construct a downside risk factor that captures the return premium between stocks with high downside correlations and low downside correlations. First, in Section 3.1, we show that the Fama-French (1993) model does not explain the cross-sectional variation in the returns of portfolios formed by sorting on downside correlations. Second, Section 3.2 details the construction of the downside correlation factor, which we call the CMC factor. We construct the CMC factor by going short stocks with low downside correlations, which have low expected returns, and going long stocks with high downside correlations, which have high expected returns. Finally, we show in Section 3.3 that the CMC factor does proxy for downside correlation risk by explaining the cross-sectional variations of in the returns of the ten downside correlation portfolios.

3.1 Fama and French (1993) and the Downside Correlation Portfolios

To see if the Fama and French (1993) model can price the ten downside correlation portfolios, we run the following time-series regression:

$$r_{it} = a_i + b_i MKT_t + s_i SMB_t + h_i HML_t + \epsilon_{it}, \tag{8}$$

where SMB$_t$ and HML$_t$ are the two Fama and French (1993) factors representing the size effect and the book-to-market effect, respectively. The coefficients, b_i, s_i and h_i, are the factor loadings on the market, the size factor and the book-to-market factor, respectively. We test the hypothesis that the a_i's are jointly equal to zero for all ten portfolios by using the F-test developed by Gibbons, Ross and Shanken (1989) (henceforth GRS).

Table (4) presents the results of the regression in equation (8). We find that portfolios of stocks with higher downside correlations have higher loadings on the market portfolio. That is, stocks with high downside correlations tend to be stocks with high market betas, which is consistent with the pattern of increasing betas across deciles 1-10 in Table (1). The columns labeled s and h show that the loadings on SMB and HML both decrease monotonically with increasing downside correlations. These results are also consistent with the characteristics listed in Table (1), where the highest downside risk stocks tend to be large stocks and growth stocks. Table (4) suggests that the Fama-French factors do not explain the returns on the downside risk portfolios since the relations between ρ^- and the factors go in the opposite direction than what the Fama-French model requires. In particular, stocks with high downside risk have the lowest loadings on size and book-to-market factors.

In Table (4), the intercept coefficients, a_i, represent the proportion of the decile portfolio returns left unexplained by the regression of equation (8). The intercept coefficients increase with ρ^-, so that after controlling for the Fama-French factors, high downside correlation stocks still have high expected returns. These coefficients are almost always individually significant and are jointly significantly different from zero at the 95% confidence level using the GRS test. The difference in a_i between the decile 10 portfolio and the decile 1 portfolio is 0.53% per month, or 6.55% per annum with a p-value 0.00. Hence, the variation in downside risk in the ρ^- portfolios is not explained by the Fama-French model. In fact, controlling for the market, the size factor and the book-to-market factor increases the differences in the returns from 4.91% to 6.55% per annum.

In Panel B of Table (4), we test whether this mispricing survives when we split the sample into two subsamples. We split the sample into January 1964 through December 1981 and from January 1982 through December 1999. We list the intercept coefficients, a_i, for the two subsamples, with robust t-statistics. Within each of these two sub-samples, the difference between the a_i for the tenth and first decile are large and statistically significant at the 5% level. The difference is 5.54% per annum (0.45% per month) for the earlier subsample and 7.31% per annum (0.59% per month) for the latter subsample.

3.2 Constructing the Downside Risk Factor

Table (1) shows that portfolios with higher downside correlation have higher β's and Table (4) shows that market loadings increase with downside risk. This raises the issue that the phenomenon of increasing returns with increasing ρ^- may be due to a reward for bearing higher exposures on β, rather than for greater exposures to downside risk. To investigate this, we perform a sort on ρ^-, after controlling for β. Each month, we place half of the stocks based on their β's into a low β group and the other half into a high β group. Then, within each β group, we rank stocks based on their ρ^- into three groups: a low ρ^- group, a medium ρ^- group and a high ρ^- group, with the cutoffs at 33.3% and 66.7%. This sorting procedure creates six portfolios in total.

We calculate monthly value-weighted portfolio returns for each of these 6 portfolios, and report their summary statistics in the first panel of Table (5). Within the low β group, the average returns increase from the low ρ^- portfolio to the high ρ^- portfolio, with an annualized difference of 2.40% (0.20% per month). Moving across the low β group, mean returns of the ρ^- portfolios increase, while the beta remains flat at around β = 0.66. In the high β group, we observe that the return also increases with ρ^-. The difference in returns of the high ρ^- and low ρ^- portfolios, within the high β group, is 3.24% per annum (0.27% per month), with a t-statistic of 1.98. However, the β decreases with increasing ρ^-. Therefore, the higher returns associated with portfolios with high ρ^- are not rewards for bearing higher market risk, but are rewards for bearing higher downside risk.

In Panel B of Table (5), for each ρ^- group, we take the simple average across the two β groups and create three portfolios, which we call the β-balanced ρ^- portfolios. Moving from the β-balanced low ρ^- portfolio to the β-balanced high ρ^- portfolio, mean returns monotonically increase with ρ^-. This increase is accompanied by a monotonic decrease in β from β = 0.94 to β = 0.87. Hence β is not contributing to the downside risk effect, since within each β group, increasing correlation is associated with decreasing β.

We define our downside risk factor, CMC, as the returns from a zero-cost strategy of shorting the β-balanced low ρ^- portfolio and going long the β-balanced high ρ^- portfolio, rebalancing monthly. The difference in mean returns of the β-balanced high ρ^- and the β-balanced low ρ^-

is 2.80% per annum (0.23% per month) with a t-statistic of 2.35 and a p-value of 0.02.[2]

Since we include all firms listed on the NYSE/AMEX and the NASDAQ, and use daily data to compute the higher-order moments, the impact of small illiquid firms might be a concern. We address this issue in two ways. First, all of our portfolios are value-weighted, which reduces the influence of smaller firms. Second, we perform the same sorting procedure as above, but exclude firms that are smaller than the tenth NYSE percentile. With this alternative procedure, we find that CMC is still statistically significant with an average monthly return of 0.23% and a t-statistic of 2.04. These checks show that our results are not biased by small firms.

Table (6) lists the summary statistics for the CMC factor in comparison to the market, SMB and HML factors of Fama and French (1993), the SKS coskewness factor of Harvey and Siddique (2000) and the WML momentum factor of Carhart (1997). The SKS factor goes short stocks with negative coskewness and goes long stocks with positive coskewness. The WML factor is designed to capture the momentum premium, by shorting past loser stocks and going long past winner stocks. The construction of these other factors is detailed in Appendix A.

Table (6) reports that the CMC factor has a monthly mean return of 0.23%, which is higher than the mean return of SMB (0.19% per month) and approximately two-thirds of the mean return of HML (0.32% per month). While the returns on CMC and HML are statistically significant at the 5% confidence level, the return on SMB is not statistically significant. CMC has a monthly volatility of 2.06%, which is lower than the volatilities of SMB (2.93%) and HML (2.65%). CMC also has close to zero skewness, and it is less autocorrelated (10%) than the Fama-French factors (17% for SMB and 20% for HML). The Harvey-Siddique SKS factor has a small average return per month (0.10%) and is not statistically significant. In contrast, the WML factor has the highest average return, over 0.90% per month. However, unlike the other factors, WML is constructed using equal-weighted portfolios, rather than value-weighted portfolios.

We list the correlation matrix across the various factors in Panel B of Table (6). CMC has a slightly negative correlation with the market portfolio of –16%, a magnitude less than the correlation of SMB with the market (32%) and less in absolute value than the correlation of HML with the market (–40%). CMC is positively correlated with WML (35%). The correlation matrix shows that SKS and CMC have a correlation of –3%, suggesting that asymmetric downside

[2] An alternative sorting procedure is to perform independent sorts on β and ρ^-, and take the intersections to be the 6 β/ρ^- portfolios. This procedure produces a similar result, but gives an average monthly return of 0.22% (t-stat = 1.88), which is significant at the 10% level. This procedure produces poor dispersion on ρ^- because β and ρ^- are highly correlated, so the independent sort places more firms in the low β/low ρ^- and the high β/high ρ^- portfolios, than in the low β/high ρ^- and the high β/low ρ^- portfolios. Our sorting procedure first controls for β and then sorts on ρ^-, creating much more balanced portfolios with greater dispersion over ρ^-.

correlation risk has a different effect than skewness risk.

However, Table (6) shows that CMC is highly negatively correlated with SMB (–64%). To allay fears that CMC is not merely reflecting the inverse of the size effect, we examine the individual firm composition of CMC and SMB. On average, 3660 firms are used to construct SMB each month, of which SMB is long 2755 firms and short 905 firms.[3] We find that the overlap of the firms, that SMB is going long and CMC is going short, constitutes only 27% of the total composition of SMB. Thus, the individual firm compositions of SMB and CMC are quite different. We find that the high negative correlation between the two factors stems from the fact that SMB performs poorly in the late 80's and the 90's, while CMC performs strongly over this period.

3.3 Pricing the Downside Correlation Portfolios

If the CMC factor successfully captures a premium for downside risk, then portfolios with higher downside risk should have higher loadings on CMC. To confirm this, we run (but do not report) the following time-series regression on the the portfolios sorted on ρ^-:

$$r_{it} = a_i + b_i MKT_t + c_i CMC_t + \epsilon_{it}, \tag{9}$$

where the coefficients b_i and c_i are loadings on the market factor and the downside risk factor respectively. Running the regression in equation (9) shows that the loading on CMC ranges from –1.09 for the lowest downside risk portfolio to 0.37 for the highest downside risk portfolio. These loadings are highly statistically significant. The regression produces intercept coefficients that are close to zero. In particular, the GRS test for the null hypothesis that these intercepts are jointly equal to zero, fails to reject with a p-value of 0.49.

Downside risk portfolios with low ρ^- have negative loadings on CMC. By construction, since the CMC factor shorts low ρ^- stocks, many of the stocks in the low ρ^- portfolios have short positions in the CMC factor. Similarly, the high ρ^- portfolio has a positive loading on CMC because CMC goes long high ρ^- stocks.

When we augment the regression in equation (9) with the Fama-French factors, the intercept coefficients a_i are smaller. However, the fit of the data is not much better, with the adjusted R^2's that are almost identical to the original model of around 90%. While the loadings of SMB and HML are statistically significant, these loadings still go the wrong way, as they do in Table (4). Low ρ^- portfolios have high loadings on SMB and HML, and the highest ρ^- portfolio

[3] SMB is long more firms than it is short since the breakpoints are determined using market capitalizations of NYSE firms, even though the portfolio formation uses NYSE, AMEX and NASDAQ firms.

has almost zero loadings on SMB and HML. However, the CMC factor loadings continue to be highly significant.

As a final check to ensure that the high downside correlation expected returns are not accounted for by other factors, we run a regression of CMC onto MKT, SMB, HML and WML factors in Panel C of Table (6). The constant in this regression is highly significant, and is 10 basis points per month higher than the raw mean of CMC in Panel A. The coefficient on MKT is near zero, since the construction of CMC controls for beta. The Fama-French coefficients, while significant, are negative, which reflect the wrong signs of the size and value effects on the original ρ^- portfolios. The coefficent on WML is positive and statistically significant, but small, reflecting a relation between CMC and momentum, which we examine in detail below.

That a CMC factor, constructed from the ρ^- portfolios, explains the cross-sectional variation across ρ^- portfolios is no surprise. Indeed, we would be concerned if the CMC factor could not price the ρ^- portfolios. In the next section, we use the CMC factor to help price portfolios formed on return characteristics that are not related to the construction method of CMC. This is a much harder test to pass, since the characteristics of the test assets are not necessarily related to the explanatory factors.

4 Pricing the Momentum Effect

In this section, we demonstrate that our CMC factor has partial explanatory power to price the momentum effect. We begin by presenting a series of simple time-series regressions involving CMC and various other factors in Table (7). The dependent variable is the WML factor developed by Carhart (1997), which captures the momentum premium. Model A of Table (7) regresses WML onto a constant and the CMC factor. The regression of WML onto CMC has an R^2 of 12%, and a significantly positive loading. In Model B, adding the market portfolio changes little; the market loading is almost zero and insignificant. In Model C, we regress WML onto MKT and SKS. Neither MKT nor SKS is significant, and the adjusted R^2 of the regression is zero. Therefore, WML returns are related to conditional downside correlations but do not seem to be related to skewness.

Models D and E use the Fama-French factors to price the momentum effect. Model D regresses WML onto SMB and HML. Both SMB and HML have negative loadings, and the regression has a lower adjusted R^2 than using the CMC factor alone in Model A. In this regression, the SMB loading is significantly negative (t-statistic = -3.20), but when the CMC factor is included in Model E, the loading on the Fama-French factors become insignificant, while the CMC factor continues to have a significantly positive loading.

In each of the regressions in Table (7), the intercept coefficients are significantly different from zero. Compared to the unadjusted mean return of 0.90% per month, controlling for CMC reduces the unexplained portion of returns to 0.75% per month. In contrast, controlling for SKS doesn't change the unexplained portion of returns and controlling for SMB and HML increases the unexplained portion of returns to 1.05% per month. While the WML momentum factor loads significantly onto the downside risk factor, the CMC factor alone is unlikely to completely price the momentum effect. Nevertheless, Table (7) shows that CMC has some explanatory power for WML which the other factors (MKT, SMB, HML and SKS) do not have.

The remainder of this section conducts cross-sectional tests using the momentum portfolios as base assets. Section 4.1 describes the Jegadeesh and Titman (1993) momentum portfolios. Section 4.2 estimates linear factor models using the Fama-Macbeth (1973) two-stage methodology. In Section 4.3, we use a GMM approach similar to Jagannathan and Wang (1996) and Cochrane (1996).

4.1 Description of the Momentum Portfolios

Jegadeesh and Titman (1993)'s momentum strategies involve sorting stocks based on their past J months returns, where J is equal to 3, 6, 9 or 12. For each J, stocks are sorted into deciles and held for the next K months holding periods, where K = 3, 6, 9 or 12. We form an equal-weighted portfolio within each decile and calculate overlapping holding period returns for the next K months. Since studies of the momentum effect focus on the J=6 months portfolio formation period (Jegadeesh and Titman, 1993; Chordia and Shivakumar, 2001), we also focus on the J=6 months sorting period for our cross-sectional tests. However, our results are similar for other horizons, and are particularly strong for the J=3 months sorting period.

Figure (1) plots the average returns of the 40 portfolios sorted on past 6 months returns. The average returns are shown with *'s. There are 10 portfolios corresponding to each of the K=3, 6, 9 and 12 months holding periods. Figure (1) shows average returns to be increasing across the deciles (from losers to winners) and are roughly the same for each holding period K. The differences in returns between the winner portfolio (decile 10) and the loser portfolio (decile 1) are 0.54, 0.77, 0.86 and 0.68 percent per month, with corresponding t-statistics of 1.88, 3.00, 3.87 and 3.22, for K=3, 6, 9 and 12 respectively. Hence, the return differences between winners and losers are significant at the 1% level except the momentum strategy corresponding to K=3. Figure (1) also shows the β's and ρ^- of the momentum portfolios. While the average returns increase from decile 1 to decile 10, the patterns of beta are U-shaped. In contrast, the

ρ^- of the deciles increase going from the losers to the winners, except at the highest winner decile. Therefore, the momentum strategies generally have a positive relation with downside risk exposure.[4] We now turn to formal estimations of the relation between downside risk and expected returns of momentum returns.

4.2 Fama-MacBeth (1973) Cross-Sectional Test

We consider linear cross-sectional regressional models of the form:

$$\mathrm{E}(r_{it}) = \lambda_0 + \lambda' \beta_i, \tag{10}$$

in which λ_0 is a scalar, λ is a $M \times 1$ vector of factor premia, and β_i is an $M \times 1$ vector of factor loadings for portfolio i. We estimate the factor premia, λ, test if $\lambda_0 = 0$ for various specifications of factors, and investigate if the CMC factor has a significant premium in the presence of the Fama-French factors. We first use the Fama-MacBeth (1973) two-step cross-sectional estimation procedure.

In the first step, we use the entire sample to estimate the factor loadings, β_i:

$$r_{it} = \alpha_i + F_t' \beta_i + \varepsilon_{it}, \quad t = 1, 2, ...T, \tag{11}$$

where α_i is a scalar and F_t is a $M \times 1$ vector of factors. We also examine (but do not report) factor loadings from 5-year rolling regressions and find similar results. In the second step, we run a cross-sectional regression at each time t over N portfolios, holding the β_i's fixed at their estimated values, $\hat{\beta}_i$, in equation (11):

$$r_{it} = \lambda_0 + \lambda' \hat{\beta}_i + u_{it}, \quad i = 1, 2, ...N. \tag{12}$$

The factor premia, λ, are estimated as the averages of the cross-sectional regression estimates:

$$\hat{\lambda} = \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_t. \tag{13}$$

The covariance matrix of λ, Σ_λ, is estimated by:

$$\hat{\Sigma}_\lambda = \frac{1}{T^2} \sum_{t=1}^{T} (\hat{\lambda}_t - \bar{\lambda})(\hat{\lambda}_t - \bar{\lambda})', \tag{14}$$

[4] Ang and Chen (2001) focus on correlation asymmetries across downside and upside moves, rather than the level of downside and upside correlation. They find that, relative to a normal distribution, loser portfolios have greater correlation asymmetry than winner portfolios, even though past winner stocks have a higher level of downside correlation than loser stocks.

where $\bar{\lambda}$ is the mean of λ.

Since the factor loadings are estimated in the first stage and these loadings are used as independent variables in the second stage, there is an errors-in-variables problem. To remedy this, we use Shanken's (1992) method to adjust the standard errors by multiplying $\hat{\Sigma}_\lambda$ with the adjustment factor $(1+\hat{\lambda}'\hat{\Sigma}_f^{-1}\hat{\lambda})^{-1}$, where $\hat{\Sigma}_f$ is the estimated covariance matrix of the factors F_t. In the tables, we report t-values computed using both unadjusted and adjusted standard errors.

Table (8) shows the results of the Fama-MacBeth tests. Using data on the 40 momentum portfolios corresponding to the J=6 formation period, we first examine the traditional CAPM specification in Model A:

$$\mathrm{E}(r_{it}) = \lambda_0 + \lambda_{MKT}\beta_i^{MKT}. \tag{15}$$

The fit is very poor with an adjusted R^2 of only 7%. Moreover, the point estimate of the market premium is negative.

Model B is the Fama-French (1993) specification:

$$\mathrm{E}(r_{it}) = \lambda_0 + \lambda_{MKT}\beta_i^{MKT} + \lambda_{SMB}\beta_i^{SMB} + \lambda_{HML}\beta_i^{HML}. \tag{16}$$

This model explains 91% of the cross-sectional variation of average returns but the estimates of the risk premia for SMB and HML are negative. The negative premia reflect the fact that the loadings on SMB and HML go the wrong way for the momentum portfolios.

In comparision, Model C adds CMC as a factor together with the market:

$$\mathrm{E}(r_{it}) = \lambda_0 + \lambda_{MKT}\beta_i^{MKT} + \lambda_{CMC}\beta_i^{CMC}, \tag{17}$$

and produces a R^2 of 93%, which is slightly higher than the Fama-French model. The estimated premium on CMC is 8.76% per annum (0.73 per month) and statistically significant at the 5% level. These results do not change when SMB and HML are added to equation (17) in Model D. While the estimates of the factor premia of SMB and HML are still negative, the CMC factor premium remains significantly positive and the regression produces the same R^2 of 93%.[5]

We examine the Carhart (1997) four-factor model in Model E:

$$\mathrm{E}(r_{it}) = \lambda_0 + \lambda_{MKT}\beta_i^{MKT} + \lambda_{SMB}\beta_i^{SMB} + \lambda_{HML}\beta_i^{HML} + \lambda_{WML}\beta_i^{WML}. \tag{18}$$

We find that adding WML to the Fama-French model does not improve the fit relative to the original Fama-French specification. Both models produce the same R^2 of 91%, but the WML

[5] When the ten ρ^- portfolios are used as base assets, the estimate of the CMC premium is 3.45% per annum, using only the CMC factor in a linear factor model.

premium is not statistically significant. However, when we add CMC to the Carhart four-factor model in Model F, the factor premia on WML and CMC are both become significant. Model F also has an R^2 of 93%. The fact that CMC remains significant at the 5% level (adj t-stat=2.01) in the presence of WML shows the explanatory power of downside risk. Moreover, the premium associated with CMC is of the same order of magnitude as that of WML, despite the fact that CMC is constructed using characteristics unrelated to past returns.

The downside risk factor CMC is negatively correlated with the Fama-French factors and positively correlated with WML. In estimations not reported, CMC remains significant after orthogonalizing it with respect to the other factors with little change in the magnitude or the significance levels. In particular, CMC orthogonalized with respect to either MKT or the Fama-French factors are both significant. CMC orthogonalized with respect to the Carhart four-factor model also remains significant. Therefore, we conclude that the significance of the downside risk factor CMC is not due to any information that is already captured by other factors.

Figure (2) graphs the loadings of each momentum portfolio on MKT, SMB, HML and CMC. The loadings are estimated from the time-series regressions of the momentum portfolios on the factors from the first step of the Fama-MacBeth (1973) procedure. We see that for each set of portfolios, as we go from the past loser portfolio (decile 1) to the past winner portfolio (decile 10), the loadings on the market portfolio remain flat, so that the beta has little explanatory power. The loadings on SMB decrease from the losers to the winners, except for the last two deciles. Similarly, the loadings on the HML factor also go in the wrong direction, decreasing monotonically from the losers to the winners.

In contrast to the decreasing loadings on the SMB and HML factors, the loadings on the CMC factor in Figure (2) almost monotonically increase from strongly negative for the past loser portfolios to slightly positive for the past winner portfolios. The increasing loadings on CMC across the decile portfolios for each holding period K are consistent with the increasing ρ^- statistics across the deciles in Figure (1). Winner portfolios have higher ρ^-, higher loadings on CMC, and higher expected returns. Since a linear factor model implies that the systematic variance of a stock's return is $\beta_i' \Sigma_f \beta_i$ from equation (11), the negative loadings for loser stocks imply that losers have higher downside systematic risk than winners. The negative loadings also suggest that past winner stocks do poorly when the market has large moves on the downside, while past loser stocks perform better.

4.3 GMM Cross-Sectional Estimation

In this section, we conduct asset pricing tests in the GMM framework (Hansen, 1982). In general, since GMM tests are one-step procedures, they are more efficient than two-step tests such as the Fama-MacBeth procedure. Moreover, we are able to conduct additional hypotheses tests within the GMM framework. We begin with a brief description of the procedure before presenting our results.

4.3.1 Description of the GMM Procedure

The standard Euler equation for a gross return, R_{it}, is given by:

$$\mathrm{E}(m_t R_{it}) = 1. \tag{19}$$

Linear factor models assume that the pricing kernel can be written as a linear combination of factors:

$$m_t = \delta_0 + \delta_1' F_t, \tag{20}$$

where F_t is a $M \times 1$ vector of factors, δ_0 is a scalar, and δ_1 is a $M \times 1$ vector of coefficients. The representation in equation (20) is equivalent to a linear beta pricing model:

$$\mathrm{E}(R_{it}) = \lambda_0 + \lambda' \beta_i', \tag{21}$$

which is analogous to equation (10) for excess returns. The constant λ_0 is given by:

$$\lambda_0 = \frac{1}{\mathrm{E}(m_t)} = \frac{1}{\delta_0 + \delta_1' E(F_t)},$$

the factor loadings, β_i, are given by:

$$\beta_i = \mathrm{cov}(F_t, F_t')^{-1} \mathrm{cov}(F_t, R_{it}),$$

and the factor premia, λ, are given by:

$$\lambda = -\frac{1}{\delta_0} \mathrm{cov}(F_t, F_t') \delta_1.$$

To test whether a factor j is priced, we test the null hypothesis $H_0 : \lambda_j = 0$.

Letting R_t denote an $N \times 1$ vector of gross returns $R_t = (R_{1t}, \ldots, R_{Nt})'$, and denoting the parameters of the pricing kernel as $\delta = (\delta_0, \delta_1')'$, the sample pricing error is:

$$g_T(\delta) = \frac{1}{T} \sum_{t=1}^{T} (m_t R_t - 1). \tag{22}$$

The GMM estimate of δ is the solution to

$$\min_{\delta} J = T \times g_T' W_T g_T, \tag{23}$$

where W_T is a weighting matrix. If the optimal weighting matrix, $W_T^* = S_T^{-1} = [T \cdot \text{cov}(g_T, g_T')]^{-1}$, is used, an over-identifying χ^2 test can be performed by using $J \sim \chi_k^2$, where k is the number of over-identifying restrictions.

We also use the Hansen and Jagannathan (1997) (HJ) distance measure, to compare the various models. The HJ distance can be expressed as:

$$HJ = \sqrt{g_T(\delta)' \mathrm{E}[R_t R_t']^{-1} g_T(\delta)}, \tag{24}$$

and can be interpreted as the least-square distance between a given pricing kernel and the closest point in the set of the pricing kernels that can price the base assets correctly. The HJ distance is also the maximum mispricing possible per unit of standard deviation. For example, if the HJ distance is 0.45 and the portfolio has an annualized standard deviation of 20%, then the maximum annualized pricing error is 9 percent.

The HJ distance can be estimated using the standard GMM procedure with one difference. The weighting matrix used is the inverse of the covariance matrix of the second moments of asset returns, $W_T = \mathrm{E}[R_t R_t']^{-1}$. The optimal weighting matrix cannot be used in this case since the weights are specific to each model, which makes it unsuitable for model comparisons. Hypothesis tests with the optimal weighting matrix may fail to reject a model because the model is difficult to estimate, rather than because the model produces small pricing errors. In contrast, the inverse of the covariance matrix of asset returns is invariant across models, so that the HJ distance provides an uniform measure across different models. We compute both asymptotic and small sample distributions for the HJ distance, which we detail in Appendix C.

4.3.2 Empirical Results

Table (9) presents the results of the GMM estimations. The 40 momentum $J = 6$ portfolios, together with the risk-free asset, are used as the base assets in these estimates. We first turn to Model A, the CAPM. Unlike the Fama-MacBeth estimation in Table (8), the market has a significantly positive risk premium, rather than a negative risk premium. The Fama-French model (Model B) estimates of risk premia for SMB and HML are negative, but the market risk premium is estimated to be positive. We consider the linear factor model with MKT and CMC in Model C. Both MKT and CMC command positive factor premia that are statistically significant at the 1% level. In particular, the CMC premium is estimated to be 1.00% per month, with a t-statistic of 4.75.

In Model D, which is the Fama-French model augmented with the CMC factor, factor premia for SMB and HML are still negative, although the HML premium is insignificantly different from zero. This model nests the MKT and CMC model of Model C and also nests the Fama-French model (Model B). Taking this as an unconstrained model and using its weighting matrix to re-estimate Models B and C, we can conduct two χ^2 over-identification tests. The null hypothesis of the first test is the Fama-French model, which tests for the significance of CMC given the Fama-French factors. This tests rejects the null hypothesis with a p-value of 0.02 (χ^2_1=7.99). Hence, CMC does provide additional explanatory power for the cross-section of momentum portfolios which the Fama-French model does not provide. In the second test, the null hypothesis is the linear factor model with only the MKT and the CMC factors, and the alternative hypothesis is Model D (MKT, SMB, HML and CMC). This test fails to reject, with a p-value of 0.15 (χ^2_2=2.12). Hence, given MKT and CMC, the size and the book-to-market factors provide no additional explanatory power for pricing momentum portfolios.

Model E is the Carhart four-factor model, which extends the Fama-French model by adding WML. The WML premium is significantly positive, as we would expect, since the WML factor is constructed using the momentum portfolios themselves. The Carhart model is nested by Model F, which adds CMC. This model also nests Model D, which uses MKT, SMB, HML and CMC factors. We run a χ^2 over-identification test with the null of Model D against the alternative of Model F. This tests rejects with a p-value of 0.01 (χ^2_2=6.54). Hence, we conclude that WML still has further explanatory power, in the presence of CMC, to price the cross-section of momentum portfolios. However, the premium of the CMC factor is still significant at the 1% level in the presence of the WML factor.

The last two columns of Table (9) list the results of Hansen's over-identification test (J-test) and the HJ test. Only the CAPM model is rejected using the J-test (p-value 0.03), while the remaining models cannot be rejected. However, the last column shows that none of the models can pass the HJ test. The HJ statistic is generally large, around 0.54 for every model. Both the asymptotic and small-sample p-values of the HJ test are less than 0.00%. Hence, although the downside risk factor is priced by the momentum portfolios, the pricing errors are still large and we reject that the pricing error is zero. Exposure to downside risk accounts for a statistically significant portion of momentum profits, but it cannot fully explain the momentum effect.

Finally, we graph the average pricing errors for the models in Figure (3), following Hodrick and Zhang (2001). The pricing errors are computed using the weights, $W_T = \mathrm{E}[R_t R_t']^{-1}$, which is the same weighting matrix used to compute the HJ distance. Since the same weighting matrix is used across all of the models, we can compare the differences in the pricing errors for different models. Figure (3) displays each momentum portfolio on the x-axis, where the first ten portfolios correspond to the $K = 3$ month holding period, the second ten to the $K = 6$ month holding period, the third ten to the $K = 9$ month holding period, and finally the fourth ten to the $K = 12$ holding period. The 41st asset is the risk-free asset. The figure plots two standard error bounds in solid lines, and the pricing errors for each asset in *'s.

Figure (3) shows that the CAPM has most of its pricing errors outside the two standard error bands and shows that the loser portfolios are the most difficult for the CAPM to price. The Fama-French model has most difficulty pricing past winners; the pricing errors of every highest winner portfolio lies outside the two standard error bands. The model using MKT and CMC factors is the only model that has all the pricing errors within two standard error bands.[6] Comparing the CAPM and the model with MKT and CMC factors, we see that adding CMC to the CAPM greatly helps to explain loser and winner portfolios. Adding CMC to the Fama French model or the Carhart model does not change the pricing errors of the assets very much. This is consistent with the fact that the Fama-French and Carhart models augmented with the CMC factor are still rejected, using the HJ distance. The Carhart model and the Carhart model augmented with the CMC factor produce an interesting pattern for pricing errors relative to other models. Models without WML typically under-estimate the expected returns of the winner portfolios. In contrast, models with the WML factor over-estimate the expected returns of winner stocks, and under-estimate the expected returns of loser stocks.

5 Downside Risk, Aggregate Liquidity and Macroeconomic Variables

In this section we explore the relation between downside risk, aggregate liquidity and the business cycle by investigating how the downside risk factor covaries with liquidity and macroeconomic variables. The investigation in this section should be regarded as an exploratory exercise, rather than as a formal test of the underlying economic determinants of downside risk.

5.1 Downside Risk and Liquidity Risk

A number of studies find that liquidity of the market dries up during down markets. Pástor and Stambaugh (2001) construct an aggregate liquidity measure which uses signed order flow, and

[6] Although all pricing errors for the model of MKT and CMC fall within the two standard error bands, this model does not pass the HJ distance test. The graph and the HJ distance give different results because the weighting matrix used in computing the HJ distance does not assign an equal weight to all the portfolios in the test, and the graph does not take into account any covariances between the pricing errors.

find that their liquidity measure spikes downwards during periods of extreme downward moves, such as during the October 1987 crash, and during the OPEC oil crisis. Other authors, such as Jones (2001) find that the bid-ask spreads increase with market downturns. Chordia, Roll and Subrahmanyam (2000) also find a positive association at a daily frequency between market-wide liquidity and market returns. These down markets, which coincide with systematically low liquidity, are precisely the periods which downside risk-averse investors dislike. We now try to differentiate between the effects of downside risk and aggregate liquidity risk.

To study the downside risk-liquidity relation, we follow Pástor and Stambaugh (2001) and reconstruct their aggregate liquidity measure, L. Our construction procedure is detailed in Appendix A. After constructing the liquidity measure, we assign a historical liquidity beta, β^L, at each month, for each stock listed on NYSE, AMEX and NASDAQ. This is done using monthly data over the previous 5 years from the following regression:

$$r_{it} = a_i + \beta^L L_t + b_i MKT_t + s_i SMB_t + h_i HML_t + \epsilon_{it}, \tag{25}$$

where L_t is the aggregate liquidity measure.

Since each stock i in our sample now has a downside correlation ($\rho^-_{i,t}$) and a liquidity beta ($\beta^L_{i,t}$) for each month, we can examine the unconditional relations across the two measures. First, we compute the cross-sectional correlation between $\rho^-_{i,t}$ and $\beta^L_{i,t}$ at each time t, and then average over time to obtain the average cross-sectional correlation between downside risk and liquidity. The average cross-sectional correlation is –0.0108, which is close to zero. We obtain the average time-series correlation between $\rho^-_{i,t}$ and $\beta^L_{i,t}$ by computing the correlation between these two variables for each firm across time, and then averaging across firms. The average time-series correlation is -0.0029, which is also almost zero. Hence, our measure of downside risk is almost orthogonal to Pástor and Stambaugh's measure of aggregate liquidity risk.

To further investigate the relation between downside risk and aggregate liquidity, we sort stocks into 25 portfolios. At each month, we independently sort all NYSE, AMEX and NASDAQ stocks into two quintile groups, one group by β^L and another by ρ^-. The intersection of these two quintile groups forms 25 portfolios sorted by β^L and ρ^-. To examine the effect of liquidity, controlling for downside risk, we average the 25 portfolios across the ρ^- quintiles. We call these portfolios the 'Average Liquidity Beta Portfolios'. We report the intercept coefficients from a Fama-French (1993) factor time-series regression of these portfolios in Panel A of Table (10). Consistent with Pástor and Stambaugh, portfolios with the lowest β^L have the most negative mispricing. However, none of the estimates of the intercept terms are statistically significant. Furthermore, the difference between a_5 and a_1 is positive (0.10% per month), but insignificant. Hence, controlling for downside risk, there is little relation between the cross-section of stock returns and their liquidity risk exposure.

In Panel B of Table (10), we average the 25 portfolios across the β^L quintiles. We call these portfolios the 'Average Downside Correlation Portfolios'. We observe a similar pattern in the average returns moving from low ρ^- to high ρ^- portfolios as in Table (4). There is negative mispricing in the low ρ^- portfolios and positive mispricing in the high ρ^- portfolio. The difference between a_5 and a_1 is 0.26% per month, which is statistically significant at the 1% level. Hence, even after controlling for liquidity risk using the Pástor-Stambaugh liquidity measure, there remains significant mispricing of downside risk relative to the Fama-French three factor model.

5.2 Downside Correlations and Macroeconomic Variables

To investigate the relation between downside risk and business cycle conditions, we consider six macroeconomic variables which reflect underlying economic activity and business conditions. Our first two variables are leading indicators of economic activity: the growth rate in the index of leading economic indicators (LEI) and the growth rate in the index of Help Wanted Advertising in Newspapers (HELP). We also use the growth rate of total industrial production (IP). The next three variables measure price and term structure conditions: the CPI inflation rate, the level of the Fed funds rate (FED) and the term spread between the 10-year T-bonds and the 3-months T-bills (TERM). All growth rates (including inflation) are computed as the difference in logs of the index at times t and $t-12$, where t is monthly.

To examine the connection between downside risk and macroeconomic variables, we run two sets of regressions. The first set regresses CMC on lagged macro variables, while the second set regresses macroeconomic variables on lagged CMC. The first set of regressions are of the form:

$$CMC_t = a + \sum_{i=1}^{3} b_i MACRO_{t-i} + \sum_{i=1}^{3} c_i CMC_{t-i} + \epsilon_t \qquad (26)$$

where we use various macroeconomic variables for $MACRO_t$.

Panel A of Table (11) lists the regression results from equation (26). There is no significant relation between lagged macroeconomic variables and the CMC factor, except for the first lag of LEI, which is significantly negatively related with CMC. A 1% increase in the growth rate of LEI predicts a 27 basis point decrease in the premium of the downside risk factor. However, the p-value for the joint test (in the last column of Table (11)) that all lagged LEI are equal to zero fails to reject the null with p-value=0.09. Overall, with the exception of LEI, there is little evidence of predictive power by macroeconomic variables to forecast CMC returns.

To explore if the downside risk factor predicts future movements of macroeconomic variables we run regressions of the form:

$$MACRO_t = a + \sum_{i=1}^{3} b_i CMC_{t-i} + \sum_{i=1}^{3} c_i MACRO_{t-i} + \epsilon_t. \qquad (27)$$

We also include lagged macroeconomic variables in the right hand side of the regression since most of the macroeconomic variables are highly autocorrelated. Panel B of Table (11) lists the regression results of equation (27). We report only the coefficients on lagged CMC. While the macroeconomic variables provide little forecasting power for CMC, the CMC factor has some weak forecasting ability for future macroeconomic variables. In particular, high CMC forecasts lower future economic activity (HELP, p-value = 0.00; IP, p-value = 0.03), lower future interest rates (FED, p-value = 0.01) and lower future term spreads (TERM, p-value = 0.03), where the p-values refer to a joint test that the three coefficients on lagged CMC in equation (27) are equal to zero.

In general, these results show that high CMC forecasts economic downturns. The predictions of high CMC and future low economic activity is seen directly in the negative coefficients for HELP and IP. Term spreads also tend to be lower in economic recessions. Estimates of Taylor (1993)-type policy rules on the FED over long samples, where the FED rate is a linear function of inflation and real activity, show short rates to be lower when output is low (Ang and Piazzesi, 2001). Hence, the positive correlation of high CMC with future low HELP, low IP, low TERM and low FED shows that high CMC weakly forecasts economic downturns. In other words, rewards to holding stocks with high downside risk is greater when future economic prospects turn sour.

6 Conclusion

Stocks with higher downside risk, measured by greater correlations conditional on downside moves of the market, have higher expected returns than stocks with low downside risk. The portfolio of stocks with the greatest downside correlations outperforms the portfolio of stocks with the lowest downside correlations by 4.91% per annum. This effect cannot be explained by the Fama and French (1993) model, since after controlling for the market beta, the size effect and the book-to-market effect, the difference in the returns between the highest and the lowest downside correlation portfolios increases to 6.55% per annum. To capture this downside risk effect, we construct a downside risk factor (CMC) that goes long stocks with high downside correlations and goes short stocks with low downside correlations. The CMC factor commands a statistically significant average premium of 2.80% per annum.

The factor structure in the cross-section of stock returns rewards investors for bearing assets with greater downside risk. In particular, past winner momentum portfolios have greater exposure to the downside risk factor than past loser momentum portfolios. Hence, some part of the profitability of momentum strategies can be explained as compensation for bearing greater downside risk. Arbitrageurs who engage in momentum strategies face the risk that the strategy performs poorly when the market experiences extreme downward moves. Downside risk provides some explanatory power for the cross-section of momentum returns, which the Fama and French (1993) model does not provide. In GMM cross-sectional estimations, the downside risk factor is significantly priced by the momentum portfolios, and it commands a significant risk premium. However, Hansen-Jagannathan (1997) tests reject the linear factor models with the CMC factor, indicating that exposure to downside risk is only a partial, and not a complete explanation for the momentum effect.

Since downside risk is priced and stocks' sensitivities to downside risk play a role in asset pricing, our empirical work points to the need for models that can explain the underlying economic mechanisms which generates downside correlation asymmetries. Representative agent models suggest that downside risk may arise in equilibium economies with asymmetric utility functions, such as first order risk aversion (Bekaert, Hodrick and Marshall, 1997) or loss aversion (Barberis, Huang and Santos, 2001). These equilibrium studies with asymmetric utility have been applied only at an aggregate, rather than a cross-sectional, level. Asymmetries in correlations can also be produced by economies with frictions and hidden information (Hong and Stein, 2001) or with agents facing binding wealth constraints (Kyle and Xiong, 2001). Similarly, these studies do not model large cross-sections of heterogeneous assets. Which of these explanations best explains the driving mechanism behind cross-sectional variations in downside risk remains to be explored.

Appendix

A Data and Portfolio Construction

Data

We use data from the Center for Research in Security Prices (CRSP) to construct portfolios of stocks sorted by various higher moments of returns. We confine our attention to ordinary common stocks on NYSE, AMEX and NASDAQ, omitting ADRs, REITs, closed-end funds, foreign firms and other securities which do not have a CRSP share type code of 10 or 11. We use daily returns from CRSP for the period covering January 1st, 1964 to December 31st, 1999, including NASDAQ data which is only available post-1972. We use the one-month risk-free rate from CRSP and take CRSP's value-weighted returns of all stocks as the market portfolio. All our returns are expressed as continuously compounded returns.

Higher Moment Portfolios

We construct portfolios based on correlations between asset i's excess return r_i and the market's excess return r_m conditional on downside moves of the market (ρ^-) and on upside moves of the market (ρ^+). We also constuct portfolios based on coskewness, cokurtosis, β, β conditional on downside market movements (β^-), and β conditional on upside market movements (β^+). At the beginning of each month, we calculate each stock's moment measures using the past year's daily log returns from the CRSP daily file. For the moments which condition on downside or upside movements, we define an observation at time t to be a downside (upside) market movement if the excess market return at t is less than or equal to (greater than or equal to) the average excess market return during the past one year period in consideration. We require a stock to have at least 220 observations to be included in the calculation. These moment measures are then used to sort the stocks into deciles and a value-weighted return is calculated for all stocks in each decile. The portfolios are rebalanced monthly.

SMB, HML, SKS and WML Factor Construction

The Fama and French (1993) factors, SMB and HML, are from the data library at Kenneth French's website at http://web.mit.edu/kfrench/www/data_library.html.

Harvey and Siddique (2000) use 60 months of data to compute the coskewness defined in equation (5) for all stocks in NYSE, AMEX and NASDAQ. Stocks are sorted in order of increasing negative coskewness. The coskewness factor SKS is the value-weighted average returns of firms in the top 3 deciles (with the most negative coskewness) minus the value-weighted average return of firms in the bottom 3 deciles (stocks with the most positive coskewness) in the 61st month.

Following Carhart (1997), we construct WML (called PR1YR in his paper) as the equally-weighted average of firms with the highest 30 percent eleven-month returns lagged one month minus the equally-weighted average of firms with the lowest 30 percent eleven-month returns lagged one month. In constructing WML, all stocks in NYSE, AMEX and NASDAQ are used and portfolios are rebalanced monthly.

The construction of CMC is detailed in Section 3.2.

Momentum Portfolios

To construct the momentum portfolios of Jegadeesh and Titman (1993), we sort stocks into portfolios based on their returns over the past 6 months. We consider holding period of 3, 6, 9 and 12 months. This procedure yields 4 strategies and 40 portfolios in total. We illustrate the construction of the portfolios with the example of the '6-6' strategies. To construct the '6-6' deciles, we sort our stocks based upon the past six-months returns of all stocks in NYSE and AMEX. Each month, an equal-weighted portfolio is formed based on six-months returns ending one month prior. Similarly, equal-weighted portfolios are formed based on past returns that ended one months prior, three months prior, and so on up to six months prior. We then take the simple average of six such portfolios. Hence, our first momentum portfolio consists of $1/6$ of the returns of the worst performers one month ago, plus $1/6$ of the returns of the worst performers two months ago, etc.

Liquidity Factor and Liquidity Betas

We follow Pástor and Stambaugh (2001) to construct an aggregate liquidity measure, L. Stock return and volume data are obtained from CRSP. NASDAQ stocks are excluded in the construction of the aggregate liquidity measure. The liquidity estimate, $\gamma_{i,t}$, for an individual stock i in month t is the ordinary least squares (OLS) estimate of $\gamma_{i,t}$ in the following regression:

$$r^e_{i,d+1,t} = \theta_{i,t} + \phi_{i,t}\bar{r}_{i,d,t} + \gamma_{i,t} sign\left(r^e_{i,d,t}\right) v_{i,d,t} + \epsilon_{i,d+1,t}, \quad d = 1, \ldots, D. \tag{A-1}$$

In equation (A-1), $\bar{r}_{i,d,t}$ is the raw return on stock i on day d of month t, $r^e_{i,d,t} = r_{i,d,t} - r_{m,d,t}$ is the stock return in excess of the market return, and $v_{i,d,t}$ is the dollar volume for stock i on day d of month t. The market return on day on day d of month t, $r_{m,d,t}$, is taken as the return on the CRSP value-weighted market portfolio. A stock's liquidity estimate, $\gamma_{i,t}$, is computed in a given month only if there are at least 15 consecutive observations, and if the stock has a month-end share prices of greater than \$5 and less than \$1000.

The aggregate liquidity measure, L, is computed based on the liquidity estimates, $\gamma_{i,t}$, of individual firms listed on NYSE and AMEX from August 1962 to December 1992. Only the individual liquidity estimates that meet the above criteria is used. To construct the innovations in aggregate liquidity, we follow Pástor and Stambaugh and first form the scaled monthly difference:

$$\Delta\hat{\gamma}_t = \left(\frac{m_t}{m_1}\right) \frac{1}{N} \sum_{i=1}^{N} (\gamma_{i,t} - \gamma_{i,t-1}), \tag{A-2}$$

where N is the number of available stocks at month t, m_t is the total dollar value of the included stocks at the end of month $t-1$, and m_1 is the total dollar value of the stocks at the end of July 1962. The innovations in liquidity are computed as the residuals in the following regression:

$$\Delta\hat{\gamma}_t = a + b\Delta\hat{\gamma}_{t-1} + c\left(m_t/m_1\right)\hat{\gamma}_{t-1} + u_t. \tag{A-3}$$

Finally, the aggregate liquidity measure, L_t, is taken to be the fitted residuals, $L_t = \hat{u}_t$.

To calculate the liquidity betas for individual stocks, at the end of each month between 1968 and 1999, we identify stocks listed on NYSE, AMEX and NASDAQ with at least five years of monthly returns. For each stock, we estimate a liquidity beta, β^L_i, by running the following regression using the most recent five years of monthly data:

$$r_{i,t} = \beta^0_i + \beta^L_i L_t + \beta^M_i MKT_t + \beta^S_i SMB_t + \beta^H_i HML_t + \epsilon_{i,t}, \tag{A-4}$$

where $r_{i,t}$ denotes asset i's excess return and L_t is the innovation in aggregate liquidity.

Macroeconomic Variables

We use the following macroeconomic variables from Federal Reserve Bank of St. Louis: the growth rate in the index of leading economic indicators (LEI), the growth rate in the index of Help Wanted Advertising in Newspapers (HELP), the growth rate of total industrial production (IP), the Consumer Price Index inflation rate (CPI), the level of the Fed funds rate (FED), and the term spread between the 10-year T-bonds and the 3-months T-bills (TERM). All growth rates (including inflation) are computed as the difference in logs of the index at times t and $t-12$, where t is monthly.

B Time-Aggregation of Coskewness and Cokurtosis

Since we compute all of the monthly higher moments measures using daily data, the problem of time aggregation may exist for some of the higher moments. Assuming that returns are drawn from infinitely divisible distributions, central moments at first and second order can scale. That is, an annual estimate of the mean μ and volatility σ can be estimated from means and volatilities estimated from daily data μ^d and σ^d, by the time aggregated relations $\mu = 250\mu^d$ and $\sigma = \sqrt{250}\sigma^d$. Hence, daily measures for second order moments are equivalent to their corresponding monthly measures. We now prove that daily coskewness and cokurtosis defined in equations (5) and (6) are equivalent to monthly coskewness and cokurtosis.

With the assumption of infinitely divisible distributions, cumulants scale but not central moments ("cumulants cumulate"). The central moment, μ_r, of x is defined as:

$$\mu_r = \int_{-\infty}^{\infty} (x - \mu_1)^r dF \text{ for } r = 2, 3, 4, ..., \tag{B-1}$$

integrating over the distribution of returns F, and $\mu_1 = \mathrm{E}(x)$. The product cumulants, κ_r, are the coefficients in the expansion:

$$\Phi(t) = \exp\left(\sum_{r=0}^{\infty} \kappa_r \frac{(it)^r}{r!}\right), \quad \text{(B-2)}$$

where $\Phi(t)$ is the moment generating function of a univariate normal distribution. The bivariate central moment, μ_{rs}, is defined as

$$\mu_{rs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_{10})^r (y-\mu_{01})^s dF \ \text{ for } r = 2,3,4,..., \quad \text{(B-3)}$$

where F is now the joint distribution of x and y, and $\mu_{10} \equiv \mu_1$ of x and $\mu_{01} \equiv \mu_1$ of y. The product cumulants, κ_{rs}, are coefficients in

$$\Phi(t_1,t_2) = \exp\left(\sum_{r,s=0}^{\infty} \kappa_{rs} \frac{(it_1)^r}{r!}\frac{(it_2)^s}{s!}\right), \quad \text{(B-4)}$$

where $\Phi(t_1,t_2)$ is the moment generating function of a bivariate normal distribution. Note that,

$$\begin{aligned} \mu_2 &= \kappa_2, \\ \mu_3 &= \kappa_3, \\ \mu_{21} &= \kappa_{21}, \\ \text{and} \quad \mu_{31} &= \kappa_{31} + 3\kappa_{20}\kappa_{11}. \end{aligned} \quad \text{(B-5)}$$

In the bivariate distribution, we use the first variable for the market excess return and the second variable for an individual stock's excess return. We compute all central moments, μ_r, using daily excess return. We denote all the monthly aggregate cumulants with tildes, $\tilde{\kappa_r} = (250/12)\kappa_r$ and monthly central moments with tildes, $\tilde{\mu_r}$.

We now prove that monthly coskewness is equivalent to scaled daily coskewness and monthly cokurtosis is equivalent to scaled daily cokurtosis. For coskewness, note that

$$\text{coskew}^{\text{m}} = \frac{\tilde{\mu}_{21}}{\tilde{\kappa}_{20}\sqrt{\tilde{\kappa}_{02}}}, \ \text{coskew}^{\text{d}} = \frac{\mu_{21}}{\kappa_{20}\sqrt{\kappa_{02}}}, \quad \text{(B-6)}$$

where coskew$^{\text{m}}$ is monthly coskewness and coskew$^{\text{d}}$ is daily coskewness. Since $\mu_{21} = \kappa_{21}$, coskew$^{\text{m}} = \frac{\tilde{\kappa}_{21}}{\tilde{\kappa}_{20}\sqrt{\tilde{\kappa}_{02}}} = \frac{\mu_{21}}{\sqrt{\frac{250}{12}}\kappa_{20}\sqrt{\kappa_{02}}} = \sqrt{\frac{12}{250}}$coskew$^{\text{d}}$. Therefore, monthly coskewness and daily coskewness are equivalent assuming returns are drawn from infinitely divisible distributions.

For cokurtosis:

$$\text{cokurt}^{\text{m}} = \frac{\tilde{\mu}_{31}}{\sqrt{\tilde{\kappa}_{20}^3}\sqrt{\tilde{\kappa}_{02}}}, \ \text{cokurt}^{\text{d}} = \frac{\mu_{31}}{\sqrt{\kappa_{20}^3}\sqrt{\kappa_{02}}}, \quad \text{(B-7)}$$

where cokurt$^{\text{m}}$ refers to monthly cokurtosis, cokurt$^{\text{d}}$ refers to dailly cokurtosis and $\tilde{\mu}_{31} = \tilde{\kappa}_{31} + 3\tilde{\kappa}_{20}\tilde{\kappa}_{11}$. Note that:

$$\begin{aligned} \text{cokurt}^{\text{m}} &= \frac{\tilde{\mu}_{31}}{\sqrt{\tilde{\kappa}_{20}^3}\sqrt{\tilde{\kappa}_{02}}} \\ &= \frac{\tilde{\kappa}_{31} + 3\tilde{\kappa}_{20}\tilde{\kappa}_{11}}{\sqrt{\tilde{\kappa}_{20}^3}\sqrt{\tilde{\kappa}_{02}}} \\ &= \frac{12}{250}\frac{\mu_{31}}{\sqrt{\kappa_{20}^3}\sqrt{\kappa_{02}}} \\ &= \frac{12}{250}\text{cokurt}^{\text{d}} \end{aligned} \quad \text{(B-8)}$$

The last equation follows from the fact that $\kappa_{11} = \mu_{11} = \mathrm{E}(\epsilon_{\mathrm{m,t}}\epsilon_{\mathrm{i,t}}) = 0$, where $\epsilon_{i,t} = r_{i,t} - \alpha_i - \beta_i MKT_t$, is the residual from the regression of the firm i's excess stock return r_i on the contemporaneous market excess return, and $\epsilon_{m,t}$ is the residual from the regression of the market excess return on a constant.

C Computing Hansen-Jagannathan (1997) Distances and P-values

Jagannathan and Wang (1996) derive the asymptotic distribution of the HJ distance (equation (24)), showing that the distribution of $T \times (HJ)^2$ involves a weighted sum of $(N-K-1)$ χ_1^2 statistics. The weights are the $N-K-1$ non-zero eigenvalues of:

$$A = S_T^{\frac{1}{2}} W_T^{\frac{1}{2}'} \left[I - W_T^{\frac{1}{2}} D_T (D_T' W_T D_T)^{-1} D_T' W_T^{\frac{1}{2}'} \right]^{-1} W_T^{\frac{1}{2}} S_T^{\frac{1}{2}'},$$

where $S_T^{\frac{1}{2}}$ and $W_T^{\frac{1}{2}}$ are the upper-triangular Cholesky decompositions of S_T and W_T respectively, and $D_T = \frac{\partial g_T}{\partial \delta}$. Jagannathan and Wang (1996) show that A has exactly $N-K-1$ positive eigenvalues $\theta_1, \ldots, \theta_{N-K-1}$. The asymptotic distribution of the HJ distance metric is:

$$T \times (HJ)^2 \rightarrow \sum_{j}^{N-K-1} \theta_j \chi_1^2$$

as $T \rightarrow \infty$. We simulate the HJ statistic 100,000 times to compute the asymtotic p-value of the HJ distance.

To calculate a small sample p-value for the HJ distance, we assume that the linear factor model holds and simulate a data generating process (DGP) with 432 observations, the same length as in our samples. The DGP takes the form:

$$r_{i,t} = r_{t-1}^f + \beta_i' F_t + \epsilon_{it}, \tag{C-1}$$

where $r_{i,t}$ is the return on the i-th portfolio, r_t^f is the risk-free rate, β_i is an $M \times 1$ vector of factor loadings, and F_t is the $M \times 1$ vector of factors. We assume that the risk-free rate and the factors follow a first-order VAR process. Let $X_t = (r_t^f, F_t)'$, and X_t follows:

$$X_t = \mu + AX_{t-1} + u_t, \tag{C-2}$$

where $u_t \sim N(0, \Sigma)$. We estimate this VAR system and use $\hat{\mu}$, $\hat{A}$ and $\hat{\Sigma}$ as the parameters for our factor generating process. In each simulation, we generate 432 observations of factors and the risk-free rate from the VAR system in equation (C-2). For the portfolio returns, we use the sample regression coefficient of each portfolio return on the factors, $\hat{\beta}_i$, as our factor loadings. We assume the error terms of the base assets, ϵ_t, follow IID multivariate normal distributions with mean zero and covariance matrix, $\hat{\Sigma}_r - \hat{\beta}' \hat{\Sigma}_F \hat{\beta}$, where $\hat{\Sigma}_r$ is the covariance matrix of the assets and $\hat{\Sigma}_F$ is the covariance matrix of the factors.

For each model, we simulate 5000 time-series as described above and compute the HJ distance for each simulation run. We then count the percentage of these HJ distances that are larger than the actual HJ distance from real data and denote this ratio empirical p-value. For each simulation run, we also compute the theoretic p-value which is calculated from the asymptotic distribution.

References

[1] Ahn, D., J. Conrad, and R. Dittmar, 2001, "Risk Adjustment and Trading Strategies," forthcoming *Review of Financial Studies*.

[2] Ang, A., and J. Chen, 2001, "Asymmetric Correlations of Equity Returns," forthcoming *Journal of Financial Economics*.

[3] Ang, A., and M. Piazzesi, 2001, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," working paper, Columbia University.

[4] Bansal, R., D. A. Hsieh, and S. Viswanathan, 1993, "A New Approach to International Arbitrage Pricing," *Journal of Finance*, 48, 1719-1747.

[5] Barberis, N., A. Shleifer, and R. Vishny, 1998, "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 3, 307-343.

[6] Barberis, N., Huang, M., and T. Santos, 2001, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 116, 1, 1-53.

[7] Bawa, V. S., and E. B. Lindenberg, 1977, "Capital Market Equilibrium in a Mean-Lower Partial Moment Framework," *Journal of Financial Economics*, 5, 189-200.

[8] Bekaert, G., R. J. Hodrick, and D. A. Marshall, 1997, "The Implications of First-Order Risk Aversion for Asset Market Risk Premiums," *Journal of Monetary Economics*, 40, 3-39.

[9] Carhart, M. M. 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 1, 57-82.

[10] Chen, J., Hong, H., and J. Stein, 2001, "Breadth of Ownership and Stock Returns," forthcoming *Journal of Financial Economics*.

[11] Chen, N., R. Roll, and S. Ross, 1986, "Economic Forces and the Stock Market," *Journal of Business*, 59, 383-403.

[12] Chordia, T., R. Roll, and A. Subrahmanyam, 2000, "Market Liquidity and Trading Activity," forthcoming *Journal of Finance*.

[13] Chordia, T., and L. Shivakumar, 2001, "Momentum, Business Cycle and Time-Varying Expected Returns," forthcoming *Journal of Finance*.

[14] Cochrane, J. H., 1996, "A Cross-Sectional Test of an Investment Based Asset Pricing Model," *Journal of Political Economy*, 104, 572-621.

[15] Conrad, J., and G. Kaul, 1998, "An Anatomy of Trading Strategies," *Review of Financial Studies*, 11, 489-519.

[16] Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53, 1839-1886.

[17] DeBondt, Werner F. M. and Thaler, R. H., 1987, "Further Evidence on Investor Overreaction and Stock Market Seasonality," *Journal of Finance*, 42, 557-581.

[18] Dittmar, R., 2001, "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross-Section of Equity Returns," forthcoming *Journal of Finance*.

[19] Fama, E. F., and J. D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 71, 607-636.

[20] Fama, E. F., and K. R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3-56.

[21] Fama, E. F., and K. R. French, 1996, "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51, 1, 55-84.

[22] Gervais, S., R. Kaniel, and D. Mingelgrin, 2001, "The High Volume Return Premium," forthcoming *Journal of Finance*.

[23] Ghysels, E., 1998, "On Stable Factor Structure in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?" *Journal of Finance*, 53, 457-482.

[24] Gibbons, M. R., S. A. Ross, and J. Shanken, 1989, "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 5, 1121-1152.

[25] Grundy, B. D., and S. J. Martin, 2001, "Understanding the Nature of Risks and the Sources of Rewards to Momentum Investing," *Review of Financial Studies*, 14, 1, 29-78.

[26] Gul, F., 1991, "A Theory of Disappointment Aversion," *Econometrica*, 59, 3, 667-686.

[27] Hansen, L. P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054

[28] Hansen, L. P., and R. Jagannathan, 1997, "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance*, 52, 2, 557-590.

[29] Harvey, C. R., and A. Siddique, 2000, "Conditional Skewness in Asset Pricing Tests," *Journal of Finance*, 55, 3, 1263-1295.

[30] Hirshleifer, D., 2001, "Investor Psychology and Asset Pricing," *Journal of Finance*, 56, 4, 1533-1597.

[31] Hodrick, R. J., and X. Zhang, 2001, "Evaluating the Specification Errors of Asset Pricing Models," forthcoming *Journal of Financial Economics*.

[32] Hong, H., and J. Stein, 1999, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *Journal of Finance*, 54, 6, 2143-2185.

[33] Hong, H., and J. Stein, 2001, "Differences of Opinion, Rational Arbitrage and Market Crashes," working paper, Stanford University.

[34] Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns," *Journal of Finance*, 51, 3-53.

[35] Jegadeesh, N., and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48, 65-91.

[36] Jegadeesh, N., and S. Titman, 2001, "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56, 2, 699-720.

[37] Jones, C., 2001, "A Century of Stock Market Liquidity and Trading Costs," working paper Columbia University.

[38] Kahneman, D., and A. Tversky, 1979, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263-291.

[39] Kraus, A., and R. H. Litzenberger, 1976, "Skewness Preference and the Valuation of Risk Assets," *Journal of Finance*, 31, 4, 1085-1100.

[40] Kyle, A. W., and W. Xiong, 2001, "Contagion as a Wealth Effect of Financial Intermediaries," forthcoming *Journal of Finance*.

[41] Lamont, O., C. Polk, and J. Saá-Requejo, 2001, "Financial Constraints and Stock Returns," *Review of Financial Studies*, 14, 2, 529-554.

[42] Markowitz, H., 1959, *Portfolio Selection*. New Haven, Yale University Press.

[43] Newey, W. K., and K. D. West, 1987, "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-8.

[44] Pástor, L., and R. F. Stambaugh, 2001, "Liquidity Risk and Expected Stock Returns," working paper, Wharton.

[45] Rubinstein, M., 1973, "The Fundamental Theory of Parameter-Preference Security Valuation," *Journal of Financial and Quantitative Analysis*, 8, 61-69.

[46] Shanken, J., 1992, "On the Estimation of Beta-Pricing Models," *Review of Financial Studies*, 5, 1-33.

[47] Taylor, J. B., 1993, "Discretion versus policy rules in practice." *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

Table 1: Portfolios Sorted on Conditional Correlations

Panel A: Portfolios Sorted on Past ρ^-

Portfolio	Mean	Std	Auto	β	Size	B/M	Lev	ρ^-	β^-	High–Low	t-stat
1 Low ρ^-	0.77	4.18	0.15	0.69	2.61	0.63	4.96	0.74	0.94	0.40	2.26*
2	0.88	4.34	0.17	0.81	2.92	0.62	4.71	0.80	0.97		
3	0.87	4.32	0.15	0.83	3.19	0.60	4.88	0.82	0.95		
4	0.94	4.39	0.15	0.87	3.46	0.58	4.38	0.83	0.97		
5	0.97	4.39	0.10	0.90	3.74	0.56	5.15	0.85	0.95		
6	1.00	4.45	0.09	0.94	4.04	0.53	4.49	0.90	1.01		
7	1.00	4.64	0.09	1.00	4.39	0.50	7.06	0.92	1.02		
8	1.03	4.58	0.08	1.00	4.82	0.48	4.04	0.94	1.05		
9	1.12	4.77	0.02	1.05	5.36	0.46	5.42	0.96	1.08		
10 High ρ^-	1.17	4.76	0.01	1.04	6.38	0.39	4.40	0.94	0.97		

Panel B: Portfolios Sorted on Past ρ^+

Portfolio	Mean	Std	Auto	β	Size	B/M	Lev	ρ^+	β^+	High–Low	t-stat
1 Low ρ^+	1.13	4.56	0.17	0.82	2.88	0.60	7.52	0.50	0.63	–0.06	–0.38
2	1.05	4.63	0.19	0.90	3.08	0.59	6.18	0.63	0.78		
3	1.09	4.61	0.16	0.92	3.24	0.58	4.51	0.68	0.82		
4	1.06	4.67	0.15	0.94	3.44	0.56	4.57	0.70	0.85		
5	0.99	4.62	0.14	0.95	3.66	0.54	4.47	0.76	0.91		
6	1.03	4.62	0.12	0.97	3.91	0.54	4.42	0.78	0.90		
7	1.00	4.70	0.09	1.01	4.23	0.52	4.84	0.84	0.97		
8	1.11	4.67	0.08	1.01	4.65	0.52	4.25	0.85	0.96		
9	1.12	4.63	0.07	1.02	5.27	0.48	4.71	0.92	1.02		
10 High ρ^+	1.07	4.52	0.00	1.00	6.65	0.36	4.35	0.95	1.06		

The table lists the summary statistics of the value-weighted ρ^- and ρ^+ portfolios at a monthly frequency, where ρ^- and ρ^+ are defined in equation (3). For each month, we calculate ρ^- (ρ^+) of all stocks based on daily continuously compounded returns over the past year. We rank the stocks into deciles (1–10), and calculate the value-weighted simple percentage return over the next month. We rebalance the portfolios at a monthly frequency. Means and standard deviations are in percentage terms per month. Std denotes the standard deviation (volatility), Auto denotes the first autocorrelation, and β is the post-formation beta of the portfolio with respect to the market portfolio. At the beginning of each month t, we compute each portfolio's simple average log market capitalization in millions (size) and value-weighted book-to-market ratio (B/M). The column Lev is the simple average of firms leverage ratio which is defined as the ratio of book value of asset to book value of equity. The columns labeled ρ^- (ρ^+) and β^- (β^+) show the post-formation downside (upside) correlations and downside (upside) betas of the portfolios. High–Low is the mean return difference between portfolio 10 and portfolio 1 and t-stat gives the t-statistic for this difference. T-statistics are computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. T-statistics that are significant at the 5% level are denoted by *. The sample period is from January 1964 to December 1999.

Table 2: Portfolios Sorted on Past Co-Measures

Panel A: Portfolios Sorted on Past Coskewness

Portfolio	Mean	Std	Auto	β	Coskew	High–Low	t-stat
1 Low coskew	1.18	5.00	0.09	1.06	–0.13	–0.15	–1.17
2	1.18	4.80	0.05	1.03	0.07		
3	1.13	4.71	0.08	1.02	–0.27		
4	1.16	4.75	0.04	1.04	0.20		
5	1.13	4.74	0.05	1.02	–0.13		
6	1.06	4.59	0.04	1.00	0.01		
7	1.19	4.64	0.08	1.01	0.03		
8	1.10	4.63	0.02	1.02	0.04		
9	1.07	4.54	0.03	1.00	0.18		
10 High coskew	1.03	4.44	0.03	0.96	0.15		

Panel B: Portfolios Sorted on Past Cokurtosis

Portfolio	Mean	Std	Auto	β	Cokurt	High–Low	t-stat
1 Low cokurt	1.21	4.64	0.02	1.01	–0.64	–0.18	–1.68
2	1.09	4.72	0.06	1.03	0.51		
3	1.11	4.64	0.02	1.01	0.42		
4	1.01	4.75	0.06	1.04	0.04		
5	1.04	4.58	0.02	0.99	0.32		
6	1.08	4.74	0.06	1.03	–0.22		
7	1.12	4.47	0.03	0.97	0.49		
8	1.08	4.60	0.07	0.99	–0.49		
9	1.17	4.63	0.05	1.01	–0.41		
10 High cokurt	1.03	4.58	0.07	0.99	–0.45		

The table lists the summary statistics for the value-weighted coskewness and cokurtosis portfolios at a monthly frequency. For each month, we calculate coskewness and cokurtosis of all stocks based on daily continuously compounded returns over the past year. We rank the stocks into deciles (1–10), and calculate the value-weighted simple percentage return over the next month. We rebalance the portfolios monthly. Means and standard deviations are in percentage terms per month. Std denotes the standard deviation (volatility), Auto denotes the first autocorrelation, and β is the post-formation beta of the portfolio with respect to the market portfolio. Coskew denotes the post-formation coskewness of the portfolio as defined in equation (5); cokurt denotes the post-formation cokurtosis of the portfolio as defined in equation (6). High–Low is the mean return difference between portfolio 10 and portfolio 1 and t-stat is the t-statistic for this difference. T-statistics are computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. The sample period is from January 1964 to December 1999.

Table 3: Portfolios Sorted on Past β, β^- and β^+

Panel A: Portfolios Sorted on Past β

Portfolio	Mean	Std	Auto	β	High–Low	t-stat
1 Low β	0.90	3.72	0.13	0.42	0.23	0.70
2	0.93	3.19	0.20	0.49		
3	1.01	3.33	0.18	0.59		
4	0.95	3.62	0.14	0.70		
5	1.13	3.78	0.08	0.76		
6	1.02	3.84	0.06	0.79		
7	1.00	4.37	0.07	0.93		
8	0.97	4.87	0.07	1.04		
9	1.07	5.80	0.08	1.23		
10 High β	1.13	7.63	0.05	1.57		

Panel B: Portfolios Sorted on Past β^-

Portfolio	Mean	Std	Auto	β	β^-	ρ^-	k^-	High–Low	t-stat
1 Lowβ^-	0.78	4.21	0.16	0.67	0.89	0.71	1.26	0.31	1.04
2	0.93	3.74	0.14	0.68	0.74	0.73	1.02		
3	0.99	3.71	0.09	0.73	0.82	0.83	0.98		
4	1.09	3.92	0.05	0.80	0.88	0.89	0.99		
5	1.05	4.00	0.06	0.85	0.89	0.91	0.98		
6	1.06	4.52	0.07	0.98	0.98	0.93	1.06		
7	1.11	4.82	0.04	1.04	1.02	0.92	1.11		
8	1.24	5.39	0.05	1.17	1.12	0.92	1.21		
9	1.22	6.26	0.04	1.32	1.30	0.89	1.46		
10 High β^-	1.09	7.81	0.08	1.57	1.52	0.84	1.82		

Panel C: Portfolios Sorted on Past β^+

Portfolio	Mean	Std	Auto	β	β^+	ρ^+	k^+	High–Low	t-stat
1 Low β^+	1.05	5.46	0.16	0.93	0.77	0.46	1.67	-0.05	-0.21
2	1.06	4.33	0.19	0.83	0.67	0.59	1.14		
3	1.05	4.06	0.16	0.80	0.69	0.67	1.04		
4	1.01	4.10	0.11	0.83	0.82	0.75	1.09		
5	0.98	4.03	0.13	0.84	0.79	0.75	1.05		
6	1.05	4.07	0.06	0.87	0.86	0.84	1.02		
7	1.07	4.35	0.06	0.94	0.90	0.86	1.05		
8	1.02	4.65	0.04	1.01	0.98	0.88	1.11		
9	1.12	5.25	0.05	1.12	1.13	0.86	1.31		
10 High β^+	1.00	6.77	0.06	1.41	1.45	0.80	1.81		

The table lists summary statistics for value-weighted β, β^- and β^+ portfolios at a monthly frequency, where β^- and β^+ are defined in equation (4). For each month, we calculate β (β^-, β^+) of all stocks based on daily continuously compounded returns over the past year. We rank the stocks into deciles (1–10), and calculate the value-weighted simple percentage return over the next month. We rebalance the portfolios monthly. Means and standard deviations are in percentage terms per month. Std denotes the standard deviation (volatility), Auto denotes the first autocorrelation, and β is post-formation the beta of the portfolio. The columns labeled β^- (β^+) and ρ^- (ρ^+) show the post-formation downside (upside) betas and downside (upside) correlations of the portfolios. The column labeled k^+ (k^-) lists the ratio of the volatility of the portfolio to the volatility of the market, both conditioning on the downside (upside). High–Low is the mean return difference between portfolio 10 and portfolio 1 and t-stat gives the t-statistic for this difference. T-statistics are computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. The sample period is from January 1964 to December 1999.

Table 4: Correlation Portfolios and Fama-French Factors

Panel A: Whole Sample Regression Jan 64 - Dec 99

Regression: $r_{it} = a_i + b_i MKT_t + s_i SMB_t + h_i HML_t + \epsilon_{it}$

	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
1 Low ρ^-	-0.37	0.69	0.53	0.43	-3.35	21.58	10.45	9.68	0.72
2	-0.30	0.80	0.48	0.38	-3.13	28.74	10.58	9.30	0.81
3	-0.31	0.83	0.43	0.38	-3.56	29.43	9.70	8.47	0.83
4	-0.24	0.86	0.41	0.33	-2.56	27.45	9.13	6.69	0.86
5	-0.19	0.90	0.33	0.26	-2.35	32.94	8.12	5.27	0.88
6	-0.16	0.94	0.24	0.24	-2.07	37.43	6.65	4.75	0.89
7	-0.15	1.00	0.18	0.16	-2.12	43.00	5.29	3.84	0.91
8	-0.10	1.01	0.10	0.11	-1.47	54.58	3.26	3.20	0.93
9	0.02	1.05	0.04	0.00	0.33	72.70	1.36	-0.06	0.95
10 High ρ^-	0.16	1.04	-0.15	-0.17	2.80	63.75	-5.33	-4.47	0.95

GRS = 1.92 p(GRS) = 0.04

Panel B: a_i's in Two Subsamples

	Decile										
	1	2	3	4	5	6	7	8	9	10	10-1
Jan 64 - Dec 81	-0.23	-0.21	-0.24	-0.23	-0.06	-0.12	-0.13	-0.05	0.11	0.22	0.45
t-stat	-1.82	-2.03	-2.12	-2.04	-0.73	-1.21	-1.45	-0.67	1.54	2.06	2.41
Jan 82 - Dec 99	-0.49	-0.34	-0.36	-0.25	-0.30	-0.20	-0.13	-0.12	-0.05	0.10	0.59
t-stat	-2.94	-2.19	-2.91	-1.60	-2.28	-1.74	-1.21	-1.13	-0.56	2.09	3.21

Panel A of this table shows the time-series regression of excess return r_i on factors MKT, SMB and HML. The ten ρ^- portfolios of Table (1) are used in the regression. $t()$ is the t-statistic of the regression coefficient computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. The regression R^2 is adjusted for the number of degrees of freedom. GRS is the F-statistic of Gibbons, Ross and Shanken (1989), testing the hypothesis that the regression intercept are jointly zero. $p(GRS)$ is the p-value of GRS. The sample period is from January 1964 to December 1999. Panel B reports the a's and t-statistics in the time series regression in subsamples. Column "10–1" is the difference of the a's for the 10th decile and the first decile.

Table 5: Construction of the Downside Correlation Factor

Panel A: Two (β) by Three (ρ^-) Sort

	Low ρ^-	Medium ρ^-	High ρ^-	High ρ^- - Low ρ^-
Low β	Mean = 0.86	Mean = 0.95	Mean = 1.06	Mean = 0.20
	Std = 3.83	Std = 3.61	Std = 3.34	t-stat=1.81
	β = 0.66	β = 0.69	β = 0.66	
	ρ^- = 0.76	ρ^- = 0.79	ρ^- = 0.81	
High β	Mean = 0.87	Mean = 1.02	Mean = 1.14	Mean = 0.27
	Std = 5.82	Std = 5.28	Std = 4.89	t-stat=1.98
	β = 1.23	β = 1.16	β = 1.09	
	ρ^- = 0.89	ρ^- = 0.94	ρ^- = 0.96	

Panel B: β-Balanced ρ^- Portfolios

	Low ρ^-	Medium ρ^-	High ρ^-	High ρ^- - Low ρ^-
β-balanced	Mean=0.86	Mean=0.98	Mean=1.10	Mean = 0.23
	Std=4.60	Std=4.29	Std=3.88	t-stat = 2.35
	β=0.94	β=0.93	β=0.87	
	ρ^-=0.88	ρ^-=0.92	ρ^-=0.96	

Summary statistics for the portfolios used to construct downside risk factor CMC at a monthly frequency. Each month, we rank stocks based on their β, computed from the previous year using daily data, into a low β group and a high β group, each group consisting of one half of all firms. Then, within each β group, we rank stocks based on their ρ^-, which is also computed using daily data over the past year, into three groups: a low ρ^- group, a medium ρ^- group and a high ρ^- group, with cutoff points at 33.3% and 66.7%. We compute the monthly value-weighted simple returns for each portfolio. The β-balanced groups are the equal-weighted average of the portfolios across the two β groups. T-statistics are computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. The sample period is from January 1964 to December 1999.

Table 6: Summary Statistics of the Factors

Panel A: Summary Statistics

Factor	Mean	Std	Skew	Kurt	Auto
MKT	0.55*	4.40	–0.51	5.50	0.06
SMB	0.19	2.93	0.17	3.84	0.17
HML	0.32*	2.65	–0.12	3.93	0.20
WML	0.90**	3.88	–1.05	7.08	0.00
SKS	0.10	2.26	0.69	7.45	0.08
CMC	0.23*	2.06	0.04	5.41	0.10

Panel B: Correlation Matrix

	MKT	SMB	HML	WML	SKS	CMC
MKT	1.00					
SMB	0.32	1.00				
HML	–0.40	–0.16	1.00			
WML	0.00	–0.27	–0.14	1.00		
SKS	0.13	0.08	0.03	–0.01	1.00	
CMC	–0.16	–0.64	–0.17	0.35	–0.03	1.00

Panel C: Regression of CMC onto Various Factors

	Constant	MKT	SMB	HML	WML
coef	0.33	-0.03	-0.44	-0.21	0.07
t-stat	4.02**	-1.47	-11.02**	-5.84**	2.65**

Panel A shows the summary statistics of the factors. MKT is the CRSP value-weighted returns of all stocks. SMB and HML are the size and the book-to-market factors (constructed by Fama and French (1993)), WML is the return on the zero-cost strategy of going long past winners and shorting past losers (constructed following Carhart (1997)), and SKS is the return on going long stocks with the most negative past coskewness and shorting stocks with the most positive past coskewness (constructed following Harvey and Siddique (2000)). CMC is the return on a portfolio going long stocks with the highest past downside correlation and shorting stocks with the lowest past downside correlation. The two columns show the means and the standard deviations of the factors, expressed as monthly percetages. Skew and Kurt are the skewness and kurtosis of the portfolio returns. Auto refers to first-order autocorrelation. Factors with statistically significant means at the 5% (1%) level are denoted with * (**), using heteroskedastic-robust Newey-West (1987) standard errors with 3 lags. The correlation matrix between the factors is reported in Panel B. Panel C reports the regression of CMC onto MKT, SMB, HML and WML factors, with t-statistics computed using 3 Newey-West lags. T-statistics that are significant at the 5% (1%) level are denoted with * (**). The sample period is from January 1964 to December 1999.

Table 7: Regression of WML onto Various Factors

		Constant	MKT	SMB	HML	CMC	SKS	Adj R^2
Model A:	coef	0.75				0.66		0.12
	t-stat	4.34**				4.48**		
Model B:	coef	0.72	0.05			0.68		0.12
	t-stat	4.19**	0.73			4.82**		
Model C:	coef	0.90	0.00				–0.01	0.00
	t-stat	5.33**	–0.06				–0.06	
Model D:	coef	1.05	0.02	–0.41	–0.26			0.10
	t-stat	6.25**	0.33	–3.20**	–2.19*			
Model E:	coef	0.86	0.04	–0.19	–0.15	0.47		0.13
	t-stat	4.87**	0.55	–1.18	–1.27	2.81**		

This Table shows the time-series regression of the momentum factor, WML, onto various other factors. MKT is the market, SMB and HML are Fama-French (1993) factors, CMC is the downside risk factor, and SKS is the Harvey-Siddique (2000) skewness factor. The t-stat is computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. T-statistics that are significant at the 5% (1%) level are denoted with * (**). The sample period is from January 1964 to December 1999.

Table 8: Fama-MacBeth Regression Tests of the Momentum Portfolios

	Factor Premiums λ							
	λ_0	MKT	SMB	HML	CMC	WML	R^2	Joint Sig
Model A: CAPM								
Premium (λ)	1.49	–0.62					0.07	p-val=0.32
t-stat	2.51*	–0.99						p-val(adj)=0.33
t-stat(adj)	2.49*	–0.98						
Model B: Fama-French Model								
Premium (λ)	–0.49	2.04	–0.50	–0.98			0.91	p-val=0.02*
t-stat	–0.52	1.95	–1.93	–2.55*				p-val(adj)=0.07
t-stat(adj)	–0.45	1.66	–1.65	–2.17*				
Model C: Using MKT and CMC								
Premium (λ)	–0.66	1.98			0.73		0.93	p-val=0.01**
t-stat	–1.09	2.73**			2.80**			p-val(adj)=0.03*
t-stat(adj)	–0.92	2.32*			2.38*			
Model D: Fama-French Factors and CMC								
Premium (λ)	–0.65	1.73	–0.11	–0.52	1.02		0.93	p-val=0.00**
t-stat	–0.72	1.52	–0.31	–1.02	3.02**			p-val(adj)=0.00**
t-stat(adj)	–0.57	1.22	–0.25	–0.81	2.43*			
Model E: Carhart Model								
Premium (λ)	–0.83	2.81	–0.79	–1.63		0.41	0.91	p-val=0.00**
t-stat	–0.98	2.84**	–2.11*	–2.47*		1.74		p-val(adj)=0.02*
t-stat(adj)	–0.72	2.10*	–1.56	–1.82		1.28		
Model F: Carhart Model and CMC								
Premium (λ)	–0.24	0.45	0.50	0.64	0.98	0.84	0.93	p-val=0.00**
t-stat	–0.27	0.41	1.48	1.53	2.86**	3.89**		p-val(adj)=0.01**
t-stat(adj)	–0.19	0.29	1.04	1.08	2.01*	2.74**		

This table shows the results from the Fama-MacBeth (1973) regression tests on the 40 momentum portfolios sorted by past 6 months returns. MKT, SMB and HML are Fama and French (1993)'s three factors and CMC is the downside risk factor. WML is return on the zero-cost strategy going long past winners and shorting past losers (constructed following Carhart (1997)). In the first stage we estimate the factor loadings over the whole sample. The factor premia, λ, are estimated in the second-stage cross-sectional regressions. We compute two t-statistics for each estimate. The first one is computed using the uncorrected Fama-MacBeth standard errors. The second one is computed using Shanken's (1992) adjusted standard errors. The R^2 is adjusted for the number of degrees of freedom . The last column of the table reports p-values from χ^2 tests on the joint significance of the betas of each model. The first p-value is computed using the uncorrected variance-covariance matrix, while the second one uses Shanken's (1992) correction. T-statistics that are significant at the 5% (1%) level are denoted with * (**). The sample period is from January 1964 to December 1999.

Table 9: GMM Tests of the Momentum Portfolios

	Constant	MKT	SMB	HML	CMC	WML	J-Test	HJ Test
Model A: CAPM								
Coefficient(δ)	1.01	–4.68					57.81	0.59
t-stat	70.87**	–4.14**					[0.03]*	[0.00]**
Premium (λ)		0.92						
t-stat		4.14**						
Model B: Fama-French Model								
Coefficient(δ)	1.00	–9.41	18.80	5.80			43.74	0.54
t-stat	31.15**	–6.00**	5.30**	1.16			[0.21]	[0.00]**
Premium (λ)		1.32	–1.15	–0.92				
t-stat		4.67**	–4.37**	–1.88				
Model C: Using MKT and CMC								
Coefficient(δ)	1.11	–7.75			–26.10		48.30	0.57
t-stat	33.48**	–6.00**			–5.12**		[0.12]	[0.00]**
Premium (λ)		1.13			1.00			
t-stat		4.87**			4.75**			
Model D: Fama-French Factors and CMC								
Coefficient(δ)	1.06	–9.55	13.46	0.61	–12.19		44.09	0.54
t-stat	21.05**	–5.95**	2.29**	0.11	–1.42		[0.17]	[0.00]**
Premium (λ)		1.14	–1.21	–0.43	0.90			
t-stat		3.94**	–4.71**	–1.35	4.17**			
Model E: Carhart Model								
Coefficient(δ)	1.04	–7.49	14.09	2.16		–3.73	44.80	0.51
t-stat	26.48**	–4.43**	3.46**	0.42		–1.70	[0.15]	[0.00]**
Premium (λ)		0.97	–0.99	–0.38		1.02		
t-stat		3.23**	–3.82**	–1.15		3.74**		
Model F: Carhart Model and CMC								
Coefficient(δ)	1.08	–8.43	11.31	–0.64	–11.01	–2.35	44.86	0.51
t-stat	19.62**	–4.88**	1.83	–0.11	–1.28	–1.13	[0.12]	[0.00]**
Premium (λ)		0.98	–1.12	–0.34	1.00	0.84		
t-stat		3.16**	–4.29**	–1.05	3.83**	3.63**		

This table lists the optimal GMM estimation results of the models using 40 momentum portfolios with the risk-free rate. Coefficient (δ) refers to the factor coefficients in the pricing kernel and Premia (λ) refers to the factor premia (λ) in monthly percentage terms. P-values of J and HJ tests are provided in [], with p-values of less than 5% (1%) denoted by * (**). The J-test is Hansen's (1982) χ^2 test statistics on the over-identifying restrictions of the model. HJ denotes the Hansen-Jagannathan (1997) distance measure which is defined in equation (24). Asymptotic and small-sample p-values of the HJ test are both 0.00 for all models. Statistics that are significant at 5% (1%) level are denoted by * (**). In all models, Wald tests of joint significance of all premiums are statistically significant with p-values of less than 0.01. The sample period is from January 1964 to December 1999.

Table 10: Liquidity Beta Portfolios and Downside Correlation Portfolios

Panel A: Mispricing across Average Liquidity Beta Portfolios

	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
1 Low β^L	-0.17	1.05	0.42	0.19	-1.93	40.39	9.88	3.30	0.91
2	-0.09	0.91	0.13	0.23	-1.36	49.44	4.30	5.76	0.94
3	-0.06	0.88	0.09	0.29	-1.15	49.36	3.54	8.19	0.93
4	-0.03	0.95	0.16	0.21	-0.49	36.88	4.44	4.80	0.93
5 High β^L	-0.08	1.01	0.42	0.11	-0.92	44.22	12.77	2.88	0.91

a_5-a_1 = 0.10 t-stat=0.71

Panel B: Mispricing across Average Downside Correlation Portfolios

	a	b	s	h	$t(a)$	$t(b)$	$t(s)$	$t(h)$	R^2
1 Low ρ^-	-0.18	0.81	0.54	0.45	-2.38	30.89	12.64	12.69	0.86
2	-0.17	0.91	0.44	0.37	-2.14	36.51	11.32	7.29	0.91
3	-0.12	0.98	0.26	0.23	-1.67	46.77	8.00	4.90	0.93
4	-0.05	1.02	0.11	0.10	-0.76	60.21	3.93	2.68	0.96
5 High ρ^-	0.08	1.07	-0.13	-0.13	1.80	90.83	-7.70	-4.70	0.98

a_5-a_1 = 0.26 t-stat=2.62

This table shows the time-series regression of excess return r_i on factors MKT, SMB and HML. In each month, we sort all NYSE, AMEX and NASDAQ stocks into 25 portfolios. We first sort stocks into quintiles by β^L and sort stocks into quintiles by ρ^-, where β^L is computed using equation (25) using the previous 5 years of monthly data. The intersection of these quintiles forms 25 portfolios on β^L and ρ^-. The average liquidity beta portfolios in Panel A are the liquidity beta quintiles averaged over the ρ^- quintiles. The average ρ^- portfolios in Panel B are the ρ^- quintiles averaged over the liquidity beta quintiles. The table reports the coefficients from a time-series regression of the portfolio returns onto the Fama-French (1993) factors: $r_{it} = a_i + b_i MKT_t + s_i SMB_t + h_i HML_t + \epsilon_{it}$. $t()$ is the t-statistic of the regression coefficient computed using Newey-West (1987) heteroskedastic-robust standard errors with 3 lags. The regression R^2 is adjusted for the number of degrees of freedom. January 1968 to December 1999. a_5-a_1 is the difference in the alphas a between the 5th quintile and the first quintile.

Table 11: Macroeconomic Variables and CMC

Panel A: $CMC_t = a + \sum_{i=1}^{3} b_i MACRO_{t-i} + \sum_{i=1}^{3} c_i CMC_{t-i} + \epsilon_t$

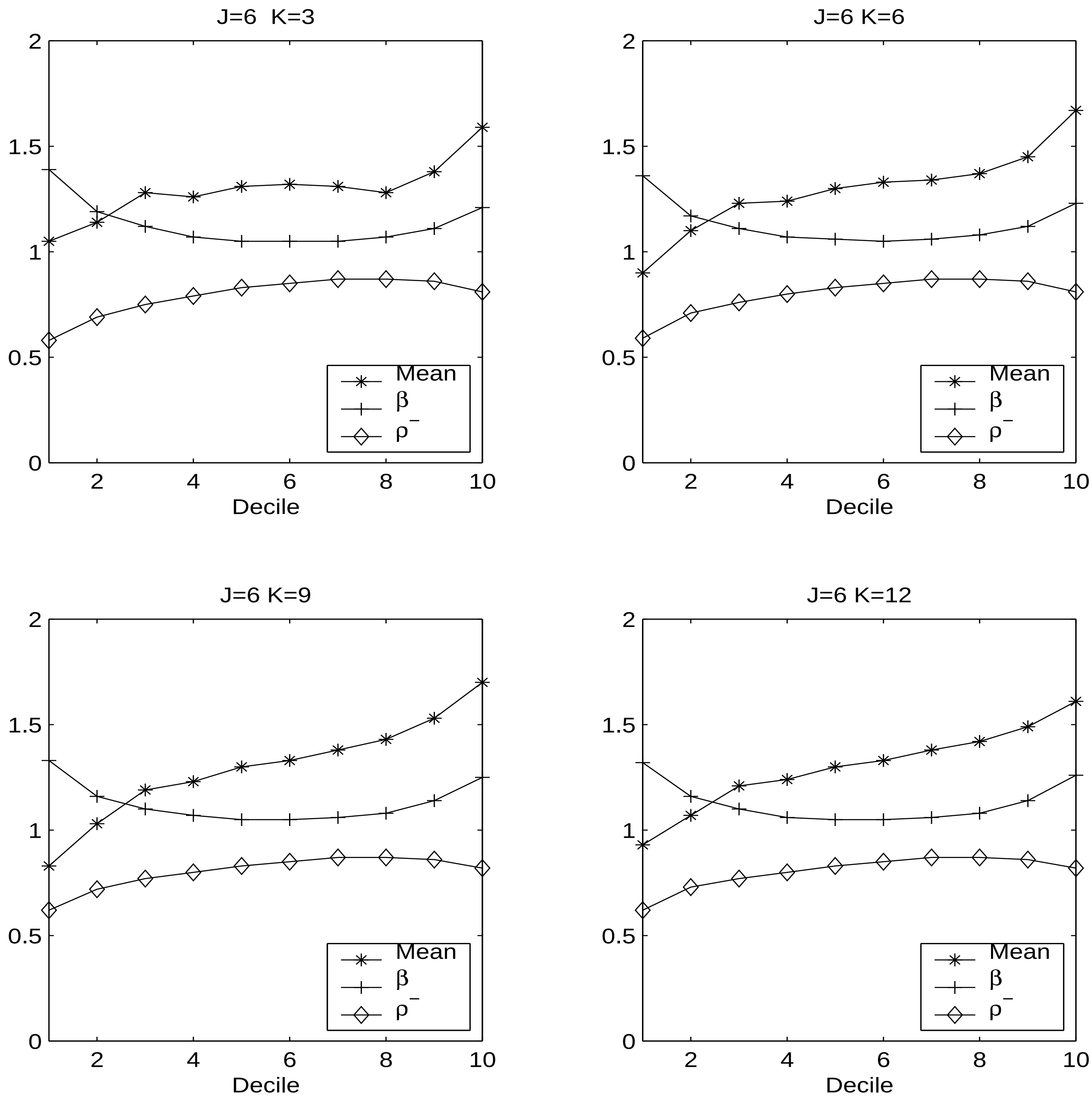
		$MACRO_{t-1}$	$MACRO_{t-2}$	$MACRO_{t-3}$	Joint Sig
LEI	coef	–0.27	0.22	0.06	0.09
	t-stat	–2.27*	1.24	0.60	
HELP	coef	–0.00	–0.04	0.05	0.24
	t-stat	–0.12	–1.26	1.97	
IP	coef	–0.13	0.18	–0.02	0.16
	t-stat	–1.39	1.43	–0.22	
CPI	coef	0.17	–0.03	–0.18	0.64
	t-stat	0.43	–0.05	–0.46	
FED	coef	0.22	–0.19	–0.02	0.48
	t-stat	1.47	–0.83	–0.12	
TERM	coef	0.10	–0.39	0.26	0.64
	t-stat	0.46	–1.11	1.10	

Panel B:
$MACRO_t = a + \sum_{i=1}^{3} b_i CMC_{t-i} + \sum_{i=1}^{3} c_i MACRO_{t-i} + \epsilon_t$

		CMC_{t-1}	CMC_{t-2}	CMC_{t-3}	Joint Sig
LEI	coef	–0.02	0.02	0.01	0.62
	t-stat	–1.04	0.73	0.31	
HELP	coef	–0.48	0.03	0.18	0.00**
	t-stat	–5.34**	0.42	1.77	
CPI	coef	–0.01	0.00	–0.01	0.51
	t-stat	–0.84	–0.17	–1.14	
IP	coef	–0.04	–0.06	0.00	0.03*
	t-stat	–1.77	–2.04*	0.03	
FED	coef	0.00	–0.03	–0.03	0.01**
	t-stat	0.30	–1.37	–2.42*	
TERM	coef	–0.02	0.02	–0.01	0.03*
	t-stat	–2.21*	1.42	–0.82	

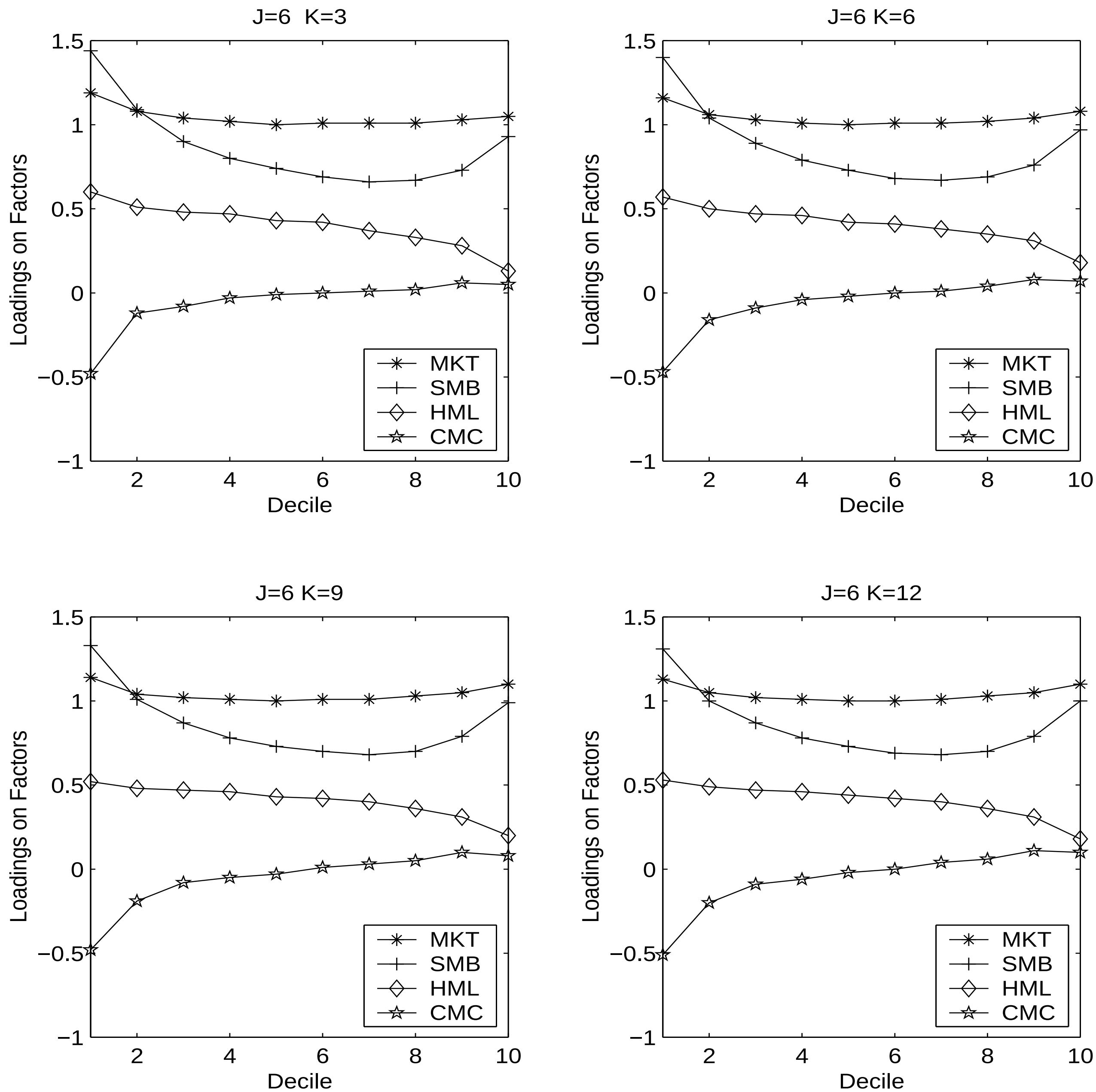
This table shows the results of the regressions between CMC and the macroeconomic variables. Panel A lists the results from the regressions of CMC on lagged CMC and lagged macroeconomic variables, but reports only the coefficients on lagged macro variables. Panel B lists the results from the regressions of macrovariables on lagged CMC and lagged macroeconomic variables, but reports only the coefficients on lagged CMC. LEI is the growth rate of the index of leading economic indicators, HELP is the growth rate in the index of Help Wanted Advertising in Newspapers, IP is the growth rate of industrial production, CPI is the growth rate of Consumer Price Index, FED is the federal discount rate and TERM is the yield spread between 10 year bond and 3 month T-bill. All growth rate (including inflation) are computed as the differences in logs of the index at time t and time $t-12$, where t is in months. FED is the federal funds rate and TERM is the yield spread between the 10 year government bond yield and the 3-month T-bill yield. All variables are expressed as percentages. T-statistics are computed using Newey-West heteroskedastic-robust standard errors with 3 lags, and are listed below each estimate. Joint Sig in Panel A denotes to the p-value of the joint significance test on the coefficients on lagged macro variables. Joint Sig in Panel B denotes the p-value of the joint significance test on the coefficients of lagged CMC. T-statistics that are significant at the 5% (1%) level are denoted with * (**). P-values of less than 5% (1%) are denoted with * (**). The sample period is from January 1964 to December 1999.

Figure 1: Average Return, β, ρ^- of Momentum Portfolios



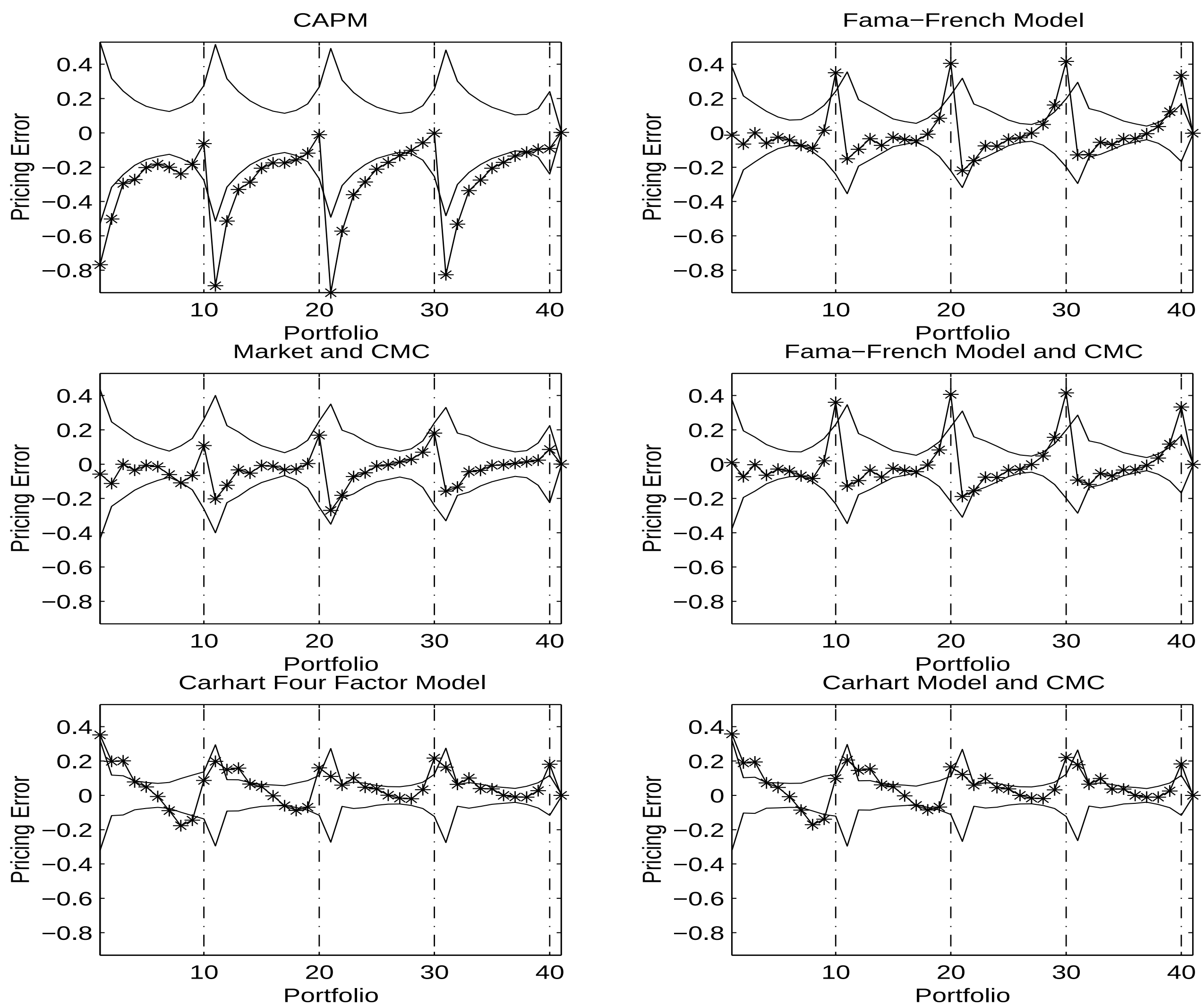
These plots show the average monthly percentage returns, β and ρ^- of the Jegadeesh and Titman (1993) momentum portfolios. J refers to formation period and K refers to holding periods. For each month, we sort all NYSE and AMEX stocks into decile portfolios based on their returns over the past J=6 months. We consider holding periods over the next 3, 6, 9 and 12 months. This procedure yields 4 strategies and 40 portfolios in total. The sample period is from January 1964 to December 1999.

Figure 2: Loadings of Momentum Portfolios on Factors



These plots show the loadings of the Jegadeesh and Titman (1993) momentum portfolios on MKT, SMB, HML and CMC. Factor loadings are estimated in the first step of the Fama-MacBeth (1973) procedure (equation (11)). J refers to formation period and K refers to holding periods. For each month, we sort all NYSE and AMEX stocks into decile portfolios based on their returns over the past J=6 months. We consider holding periods over the next 3, 6, 9 and 12 months. This procedure yields 4 strategies and 40 portfolios in total. MKT, SMB and HML are Fama and French (1993)'s three factors and CMC is the downside correlation risk factor. The sample period is from January 1964 to December 1999.

Figure 3: Pricing Errors of GMM Estimation (HJ method)



These plots show the pricing errors of various models considered in Section 4.2. Each star in the graph represents one of the 40 momentum portfolios with $J = 6$ or the risk-free asset. The first ten portfolios correspond to the $K = 3$ month holding period, the second ten to the $K = 6$ month holding period, the third ten to the $K = 9$ month holding period, and finally the fourth ten to the $K = 12$ holding period. The 41st asset is the risk-free asset. The graphs show the average pricing errors with asterixes, with two standard error bands in solid lines. The units on the y-axis are in percentage terms. Pricing errors are estimated following computation of the Hansen-Jagannathan (1997) distance.