Sources of Momentum Profits: Evidence on the Irrelevance of Characteristics*

Pavel Bandarchuk   Jens Hilscher
State Street Global Advisors   Brandeis University

September 2011

Abstract

Several recent studies document that sorting stocks first on certain stock-level characteristics and then on past returns results in elevated momentum profits. We show that such strategies enhance momentum profits simply by trading in stocks with more extreme past returns. Adjusted for this effect, elevated momentum profits resulting from characteristics (size, R^2, turnover, age, analyst coverage, analyst forecast dispersion, market-to-book, price, illiquidity, credit rating) disappear almost entirely. Interaction patterns have been used to support behavioral and limits-to-arbitrage explanations of momentum; our findings imply that explanations of momentum should instead focus on the link between momentum profits and extreme past returns.

JEL Classification: G11, G12, G14

Keywords: Momentum, past returns, volatility, stock-level characteristics, double
        sorts

---

1 Introduction

It is well established that past winners outperform past losers (e.g. Jegadeesh and Titman, 1993).[1] There is, however, no widely accepted explanation for high and significant momentum profits. A large part of the literature has favored behavioral and information-based explanations: Momentum in stock returns may be the result of investors having a tendency to herd, under-react to information, trade securities too infrequently, or pay too much attention to recent performance (e.g. Barberis et al., 1998; Daniel et al., 1998; Grinblatt and Han, 2005).

Several recent studies document that momentum profits are higher for stocks with certain characteristics. Momentum returns are higher for stocks that are small and have low analyst coverage (Hong et al., 2000), high analyst forecast dispersion (Zhang, 2006; Verardo, 2009), low return R^2 (Hou et al., 2006), and high market-to-book ratios (Daniel and Titman, 1999). Since these characteristics are commonly used to proxy for information uncertainty and limits to arbitrage, these findings are often interpreted as evidence in support of behavioral explanations of momentum. Other studies document elevated momentum profits for stocks with low-grade credit ratings (Avramov et al., 2007) and high turnover (Lee and Swaminathan, 2000).

In this paper we show that characteristic screens lead to elevated momentum profits because they produce more extreme momentum sorts. We demonstrate that the ability of characteristics to enhance momentum profits is the result of two empirical patterns: (1) stocks with extreme characteristics tend to have more extreme past returns; (2) more extreme past returns result in higher momentum profits (see e.g. Jegadeesh and Titman, 1993; Fama and French, 1996). It is therefore not the characteristic screens per se that are responsible for elevated profits; rather, characteristic interaction patterns are the result of more extreme past return sorts. This suggests that explanations of the momentum anomaly that are based on evidence that characteristic screens enhance momentum profits should be reconsidered.

We begin by showing the effect of sorting on characteristics and past returns in a simple model. If stocks have different levels of return volatility a conditional double sort on volatility and past returns results directly in more extreme past returns: High volatility past winners

---

[1]Momentum profits are economically significant, they are robust, and they have continued to be high even after becoming well-documented (Jegadeesh and Titman, 2001; Schwert, 2003).

have more extreme past returns than full sample winners. Sorting on a characteristic that proxies for volatility, e.g. size, results in a similar pattern. We also show that an independent sort does not eliminate the problem: High volatility stocks continue to have more extreme past returns.

We next show that characteristics are correlated with volatility and extreme past returns. Stocks with more volatile and more extreme past returns are more likely to be small stocks, have low return R^2, be young, illiquid, have low analyst coverage, high analyst forecast dispersion, high levels of market-to-book ratios, low share prices, recently high turnover, and low-grade credit ratings.[2] Thus, given the link between past returns and momentum profits (Jegadeesh and Titman, 1993; Fama and French, 1996), we expect double sorts of characteristics and past returns to produce enhanced profits. Following the literature[3] we double sort stocks using characteristics and past returns and confirm that such strategies trade stocks with more extreme past returns and result in choose stocks with higher momentum profits.

Some of the characteristics that we sort on, in particular size and market-to-book, are known to be related to risk. It is therefore important to analyze risk-adjusted returns of momentum strategies. Grundy and Martin (2001) point out that if returns are generated by a factor model, winner and loser stocks' factor exposures vary with formation period factor realizations.[4] Correcting returns using constant factor exposures is thus not appropriate; instead, risk-adjustment of momentum returns must take into account the dynamic nature of factor loadings. We follow Korajczyk and Sadka (2004) to calculate risk-adjusted returns. Consistent with Grundy and Martin (2001) we find that momentum profits remain significant after correcting for dynamic variation in the three Fama and French (1993) factors. We also find that the effect on momentum returns of sorting stocks first by characteristics is robust to measuring risk-adjusted returns.

We next provide three sets of results which indicate that characteristic screens lead to

---

[2] We postpone a detailed discussion of the relevant studies to Section 3, where we introduce each of the characteristics in turn.

[3] See e.g. Hong et al. (2000), Lee and Swaminathan (2000), Zhang (2006), Avramov et al. (2007), and Arena, et al. (2008) among others.

[4] For example, if the market return was large and positive during the formation period, then winner stocks tend to be high beta stocks. Importantly, the opposite is true in a market downturn.

enhanced momentum profits primarily by trading in stocks with more extreme past returns. First, to isolate the incremental effect of the characteristics over and above volatility, we regress each characteristic on volatility and construct levels of residual characteristics. We find that differences in momentum returns between high and low residual characteristics are substantially smaller than unadjusted return differences in all cases. None of the differences remain positive and statistically significant except for age, for which the effect is cut in half. Second, we adjust momentum profits by directly controlling for the empirical relationship between past returns and momentum profits. Adjusted returns for extreme characteristic quintiles are substantially reduced and in most cases lose statistical significance. In other words, disregarding information about characteristics and using only past returns to sort stocks into portfolios results in momentum profits that are very close to and in several cases higher than returns resulting from double sorts.

Third, we implement a regression-based test of our hypothesis. We run regressions of momentum profits on characteristics, volatility, and extreme past returns (following Fama and MacBeth, 1973). Characteristics, volatility, and extreme past returns are all related to momentum profits, but they are also related to each other. In a setting with correlated explanatory variables, a regression-based approach seems natural if we want to control for different drivers of profits (Fama and French, 2008). We first run a regression of momentum profits on each of the characteristics and confirm the apparent link between extreme characteristics and momentum profits. When we correct for volatility and extreme past returns the relationship between characteristics and profits is substantially reduced in all cases, while the effect of extreme past returns and volatility remains present and statistically significant.

Previous studies emphasize that the ability of characteristic screens to enhance momentum profits supports behavioral or limits-to-arbitrage explanations of momentum. Our findings call for a reappraisal of these conclusions and suggest that a focus on the link between extreme past returns and momentum profits may be more appropriate (e.g. Grundy and Martin, 2001). Indeed, the evidence is consistent with Vayanos and Woolley (2010), who present a rational explanation of momentum that links momentum to high idiosyncratic volatility and therefore

3

to extreme past returns.[5]

The rest of the paper is organized as follows. Section 2 presents a theoretical framework that shows the consequences of double sorting on past returns. Section 3 discusses the data, the link between characteristics, past returns, and volatility, as well as risk-adjusted returns. In section 4 we perform conditional double sorts on characteristics and past returns and adjust profits for volatility and extreme past returns. Section 5 concludes.

2 Consequences of Double-Sorting: Analytical Investigation of Conditional Expectations of Past Returns

A standard momentum strategy sorts stocks into groups based on past returns. Given that more extreme past returns are related to higher momentum profits, when evaluating momentum profits for a given strategy it is important to keep track of past returns.

Before proceeding, we briefly show empirically that low volatility stocks have a lower past return spread than high volatility stocks. In Figure 1 we sort stocks into quintiles Q1 (low volatility) to Q5 (high volatility).[6] For each quintile we then calculate winner and loser cutoffs, which are past return levels so that 20% of the distribution lies above the winner cutoff and 20% lies below the loser cutoff. We then calculate the average past return for losers and winners, which are indicated as the top and bottom point of each vertical bar. It is apparent from the figure that cutoffs and average past returns are much more extreme for high volatility stocks than they are for low volatility stocks. In this section we now demonstrate this effect analytically.

---

[5]Consistent with this prediction, Lou (2011) finds that flows into individual stocks can explain 50% of momentum.

[6]We discuss our data and provide variable definitions in sections 3 and 4.

2.1 SORTING ON VOLATILITY AND PAST RETURNS

A stock is classified as a winner if its past return r is higher than a cutoff c, which is typically defined so that a certain percentage of stocks are classified as winners and $\Pr(r > c) = Q$ e.g. a quintile sort where $Q = 20\%$. Conditional on being a winner the expected past return is equal to

$$E[r|r \text{ is a top quintile winner}] = E[r|r > c].$$

In a volatility-past-return double sort, stocks are classified as high volatility winners. If a stock is in the top volatility quintile and it is a winner in that quintile, then the stock's expected past return is given by

$$E[r|\sigma \text{ is in top quintile } AND \ r \text{ is a top quintile winner}] = E[r|\sigma > c_\sigma, r > c_{r,\sigma}]$$

where σ denotes the return standard deviation, c_σ is the volatility cutoff, and $c_{r,\sigma}$ is the return cutoff. In the case of a conditional quintile double sort $\Pr(\sigma > c_\sigma) = \Pr(r > c_{r,\sigma}|\sigma > c_\sigma) = 20\%$.

In order to quantify the effects of single and double sorts we assume that returns are distributed normally: $r \sim N\left(0, \sigma^2\right)$.[7] Conditional on a level of volatility σ, and a share Q of winners, $E[r|r > c, \sigma] = \frac{1}{Q} n\left(\frac{c}{\sigma}\right)\sigma$ where $n(.)$ denotes the standard normal p.d.f. (for details please see the appendix). Since $n\left(\frac{c}{\sigma}\right)\sigma$ is increasing in σ concentrating on high volatility winners leads to higher expected past returns. We can now compare past returns for two groups of winners: high volatility winners and winners resulting from a single (unconditional) sort on past returns only. Since we are comparing sorts based on the same percentile cutoff, e.g. quintile past return sorts, it follows that $Q = \Pr(r > c_{r,\sigma}|\sigma > c_\sigma) = \Pr(r > c)$. It is then straightforward to show that past returns for high volatility winners are larger than past

---

[7]We assume that expected returns are equal to zero for simplicity. This assumption is reasonable given that sorting stocks on past returns results in past return spreads that are very large relative to possible differences in mean returns.

returns for full universe winners:

$$E\left[r|\sigma > c_\sigma, r > c_{r,\sigma}\right] > E\left[r|r > c\right]$$

There are two effects at work here. First, high volatility returns are more extreme than full universe returns. Since the conditional expected past return is increasing in σ it follows that for the same return cutoff $c_{r,\sigma}$ $E\left[r|\sigma > c_\sigma, r > c_{r,\sigma}\right] > E\left[r|r > c_{r,\sigma}\right]$. Second, high volatility stocks have more extreme past returns. Since in a conditional double sort the probability of lying above the cutoff is fixed, the past return cutoff for high volatility stocks is higher than the full universe cutoff $(c_{r,\sigma} > c)$, which then implies that $E\left[r|r > c_{r,\sigma}\right] > E\left[r|r > c\right]$. This restriction is dropped in an independent double sort for which the high volatility return cutoff is set equal to the full universe cutoff c. However, since the first effect remains, an independent sort does not eliminate the effect of high volatility on conditional past returns, though it does reduce the size of the difference.

We now provide an illustrative example of the consequences of double-sorting. If we assume that volatility has a uniform distribution, $\sigma \epsilon \left[\sigma_{\min}, \sigma_{\max}\right]$ we can solve for the conditional expected past return:[8]

$$
\begin{aligned}
E\left[r|r > c\right] &= \frac{E\left[n\left(\frac{c}{\sigma}\right)\sigma\right]}{\Pr\left(r > c\right)} \\
&= \frac{1}{Q\sqrt{2\pi}\left(\sigma_{\max} - \sigma_{\min}\right)}\frac{1}{4}\left[2\sigma^2\exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right) - c^2\Gamma\left(0, \frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right)\right]_{\sigma_{\min}}^{\sigma_{\max}}
\end{aligned}
\tag{1}
$$

where $Q = \Pr\left(r > c\right)$ and $\Gamma\left(.,.\right)$ is the incomplete *Gamma* function. We can use this expression either for the single sort (including the entire distribution of volatility) or for the double sort, for which we choose $\sigma_{\min}$ so that 20% of the volatility distribution lies above it. To match the mean and standard deviation of volatility to the data (mean 6-month volatility of 28.4% with a cross-sectional standard deviation of 11.2%) we assume that $\sigma_{\min} = 9.0\%$ and $\sigma_{\max} = 47.8\%$. For the double sort (including stocks with volatility between $c_\sigma = 40\%$ and $\sigma_{\max} = 47.8\%$).

---

[8]The assumption of a uniform distribution for volatility results in an analytical expression for the conditional expected past return. The results are similar if we assume a truncated normal distribution for volatility (see next section).

6

Given these inputs we can find the quintile single sort return cutoff of $c = 21.7\%$[9] and the conditional expected past return from Equation (1) of $E[r|r > c] = 42.1\%$. For high volatility stocks the return cutoff is $c_{r,\sigma} = 36.9\%$ and the conditional expected past return is $E[r|\sigma > c_\sigma, r > c_{r,\sigma}] = 61.5\%$, which is much larger than the single sort expected return.[10]

Another illustrative quantity is the single-sort percentile cutoff that makes single-sort and double-sort past returns equal:

$$E[r|\sigma \text{ is in top quintile } AND \ r \text{ is a top quintile winner}] = E\left[r|r \text{ is a } Q^{match} \text{ percent winner}\right]$$

$$E[r|\sigma > c_\sigma, r > c_{r,\sigma}] = E\left[r|r > c_{r,\sigma}^{match}\right].$$

The single-sort percentile Q^{match} and cutoff $c_{r,\sigma}^{match}$ are chosen so that the above equalities hold. Intuitively, the percentile tells us how extreme a single sort needs to be in order to be comparable to a double sort. In our example the double sort expected return is 61.5%, which is achieved with a single-sort cutoff of $c_{r,\sigma}^{match} = 42.6\%$ that classifies the top $Q^{match} = \Pr\left(r > c_{r,\sigma}^{match}\right) = 7.5\%$ as winners. This is, of course, a much more extreme sort than a quintile single sort.

We also briefly consider the case of an independent double sort. An independent sort reduces the impact on past returns but it does not eliminate it. The return cutoff is $c = 21.7\%$ (the single-sort return cutoff) and the conditional expected return is $E[r|\sigma > c_\sigma, r > c] = 49.9\%$, larger than the single sort expected return of 42.1%. Going forward, we focus on conditional sorts since these are the ones commonly used in the context of characteristic and past return sorts.

---

[9]Note that since volatility varies across stocks, the overall distribution of returns is not standard, which means that we need to find the cutoff numerically: $c = \arg\text{solve}\left\{0.2 = \frac{1}{(\sigma_{\max} - c_{\min})} \int_{c_{\min}}^{\sigma_{\max}} \left(1 - N\left(\frac{c}{\sigma}\right)\right) d\sigma\right\}$, where $N(.)$ denotes the cumulative standard normal c.d.f.

[10]The results we report in this section are based on numerical integration. We have checked that the results can also be obtained using Monte Carlo simulation.

2.2 SORTING ON SIZE (OR ANOTHER CHARACTERIS-TIC) AND PAST RETURNS

One of the most common characteristics used as a momentum interaction variable is size (e.g. Hong et al., 2000). If returns can be explained by a factor structure (e.g. Fama and French, 1993), then higher loadings on the small stock factor SMB (Fama and French, 1996) means that small stocks have more extreme returns than large stocks. In addition, small stocks also have higher idiosyncratic volatility. Size is therefore correlated with volatility so that the logic of the effects of double sorting apply: Sorting on small stocks represents a sort on volatility and past returns are more extreme for small stocks than for large stocks.

In order to analyze a size-past-return sort more rigorously we assume that size and volatility follow a bivariate normal distribution.[11] This approach incorporates characteristics such as size that are correlated with volatility, and also represents a more realistic assumption regarding the distribution of volatility.

We can quantify the relationship of past returns and size. We assume that $\log(size)$ and volatility σ are distributed bivariate normal and use the empirical mean and standard deviation of volatility (28.4% and 11.2%) and the correlation with $\log(size)$ of -31%. Without loss of generality, we $s = -\frac{\log(size) - \mu_s}{\nu_s}$, where μ_s and ν_s are the mean and standard deviation of $\log(size)$. This way s is distributed normally and is positively correlated with volatility. We provide expressions for the conditional expectations in the appendix.

Given our assumptions, the single-sort return cutoff and conditional expected returns are $c = 22.0\%$ and $E[r|r > c] = 42.1\%$.[12] The double sort cutoff is $c_{r,s} = 26.3\%$ and the expected return is $E[r|s > c_s, r > c_{r,s}] = 48.5\%$. We can also find the percent cutoff in a single sort that

---

[11] We take into account that volatility cannot be negative. The discussion in the appendix reflects this constraint.

[12] The return cutoff is different than in the previous example since we now assume that volatility has a truncated normal distribution.

would result in the same level of average past returns as the double sort:

$$E\left[r|s \text{ is in bottom quintile } AND \ r \text{ is a top quintile winner}\right] = E\left[r|r \text{ is a } Q \text{ percent winner}\right]$$
$$E\left[r|s > c_s, r > c_{r,s}\right] = E\left[r|r > c_{r,s}^{match}\right].$$

Given the inputs, classifying the top 14.6% as winners ($\Pr\left(r > c_{r,s}^{match}\right) = 14.6\%$) results in the same conditional expected return (48.5%) as a size-past-return quintile double sort.

More generally, any characteristic that is correlated with volatility will produce more extreme past returns. If we assume that a characteristic and volatility are distributed bivariate normal, then the only difference in two double sorts for two characteristics is the correlation between the characteristics and volatility.

3 Extreme Past Returns, Characteristics, and Momentum Profits

We now explore the relationship between characteristics, volatility, and extreme past returns empirically. We discuss the return data, our measures of return dispersion, and introduce the characteristics that we analyze.

3.1 DATA

We calculate momentum profits using monthly stock return data from the Center for Research in Security Prices (CRSP). The sample period is from 1964 to 2008.[13] Following the literature (e.g. Hong et al. 2000; Brav et al. 2010), we include only stocks with share codes equal to 10 and 11, dropping from our analysis ADRs, closed-end funds, and REITs. We include only observations that at the beginning of the holding period had stock prices above $5 and drop firms with market capitalization in the lowest NYSE plus AMEX decile. We also exclude stocks

---

[13]We choose June 1963 as the starting point since it represents the start of daily industry returns on Professor Ken French's website. We use these returns to construct our measure of idiosyncratic volatility and R^2.

that do not have enough consecutive non-missing observations to be placed in a momentum portfolio at least once. Our final data set contains a total of 17,940 stocks and 3,187 stocks per month on average.

Each month, we calculate a stock's past return as the cumulative return over the previous six months $(t-6, t-1)$, that is $\Pi_{k=1}^{6} (1 + r_{i,t-k}) - 1$, where r_{it} is the month t, firm i equity holding period return. We then sort firms into portfolios at the end of month t, and report equally-weighted portfolio returns for month $(t+1)$.[14] We include the standard one-month lag between the formation and holding period so that our results are not influenced by the impact of bid-ask bounce or short-term reversals (Jegadeesh, 1990; Lehmann, 1990). Furthermore, since momentum profits derive partly from the low performance of loser stocks, we adjust all monthly stock returns by using CRSP de-listing returns whenever available (see Eisdorfer, 2008).[15]

3.2   MEASURING RETURN DISPERSION

Throughout the paper we use two stock-level measures of extreme past returns. We first measure past returns directly: For each stock and each month t we define 'momentum strength' as the absolute value of the difference between the $(t-6, t-1)$ stock return and the median return over that period:

$$Mom\_strength_{it} = \exp\left(|r_{i,t-6,t-1} - r_{median,t-6,t-1}|\right) - 1 \tag{2}$$

where $r_{i,t-6,t-1}$ and $r_{median,t-6,t-1}$ denote $(t-6, t-1)$ log returns. Momentum strength measures directly the extent to which past returns are extreme. More extreme losers and winners

---

[14]We do not use overlapping momentum portfolios as in Jegadeesh and Titman (1993) since we are not primarily concerned with reducing trading costs but are instead interested in the characteristics of stocks in different momentum portfolios. Grundy and Martin (2001), Ang et al. (2006), and Fama and French (2008) also use non-overlapping portfolios.

[15]We include de-listing returns for accuracy of results. Single-sort and double-sort momentum returns are robust to not adjusting for de-listing returns. Shumway and Warther (1999) point out that observations with de-listing codes but no de-listing returns may need to be adjusted. Due to the size and share price cutoffs such an adjustment has no effect on our results since there are virtually no instances of missing de-listing returns.

will have higher levels of momentum strength and strategies with a higher past return spread will be trading in stocks with higher momentum strength. We use log returns in order to make returns of extreme winners and extreme losers more comparable. For example, when sorting stocks into deciles there is a large difference in the absolute value of average past returns of winners and losers (77% and 30%) while the difference in momentum strength is much smaller (51% and 55%).

Volatility also measures the extent to which past returns are extreme. Following Hou et al. (2006) we construct stock return idiosyncratic volatility ($IVOL$) from a regression that includes the market return and industry returns. We use the 49 industry portfolio returns and industry classifications using firm-level Standard Industrial Classification (SIC) codes that are provided on Professor Ken French's website. Since volatility may vary over time we calculate estimates on a rolling basis. At the end of each month, using data for the previous 52 weeks,[16] we estimate the following regression:

$$r_{it} = a_i + b_i r_{mt} + \gamma_i r_{It} + e_{it} \tag{3}$$

where r_{it} is the log return of stock i at time t, r_{mt} is the log return on the CRSP value-weighted market index, a_i is a firm-specific constant, and r_{It} is the log return for the industry to which stock belongs. We implement the model using weekly data and Wednesday-to-Wednesday returns.[17] $IVOL$ is the annualized standard deviation of residual returns.

We note that our results do not depend on the specific measure of volatility that we employ. We also compute the standard deviation of log returns, which has an average monthly rank correlation of 96% with $IVOL$.[18] To make our results more easily comparable with the existing literature we use $IVOL$ to capture return dispersion and extreme returns (along with

---

[16]We require a minimum of 26 weekly observations in order to report a valid estimate of volatility.

[17]Using weekly data helps to mitigate the impact of micro-structure effects and at the same time allows us to include a large number of observations.

[18]This high level reflects the fact that at the stock level the return R^2 of, for example, a market model is very low and that therefore high idiosyncratic volatility stocks are the same ones as those that have high overall return volatility. This also means that our results will be robust to other ways of calculating $IVOL$ (e.g. using the Fama-French 3-factor model). Stocks with high $IVOL$ will again be stocks with high total volatility.

momentum strength). Going forward we will use both momentum strength and volatility to capture stocks with high return dispersion and extreme past returns. As we show below, our results are robust to how we measure extreme past returns.

3.3 CHARACTERISTICS DEFINITIONS AND RELATED LITERATURE

Much of the literature has considered characteristics one at a time; we show that characteristics are correlated with each other and with volatility. We consider the following characteristics: (1) size, (2) R^2, (3) turnover, (4) age, (5) analyst coverage, (6) analyst forecast dispersion, (7) market-to-book ratio, (8) price, (9) illiquidity, and (10) credit rating. We measure size, R^2, turnover, price, and illiquidity using data from CRSP. We add data to construct market-to-book, credit ratings, and IPO dates (to construct company age) from COMPUSTAT. Data on analyst coverage and analyst forecast dispersion is from I/B/E/S.

We expect characteristics to be linked to volatility and extreme past returns. For example, small stocks have higher level of idiosyncratic volatility, consistent with small companies being less diversified and more exposed to firm-specific shocks. In addition, the return standard deviation of small stocks is also higher because they have higher small cap (SMB) factor loadings. Small stocks are more likely to be young and, especially if they have high idiosyncratic volatility, to have low return R^2. Volatility will be closely related to turnover, and credit rating,[19] and if growth and glamor stocks tend to be small firms with low current profitability and more intangible assets, then high market-to-book firms will have high volatility.[20] In section 3.4 we discuss correlations between characteristics, volatility and extreme past returns further.

Given the links between characteristics, volatility, and size, our analytical framework im-

---

[19]Karpoff (1987) surveys the (primarily time-series) literature on the links between price changes and trading volume. To the extent that credit rating reflects default probability and credit spreads, we expect rating and volatility to be closely related (Merton, 1974).

[20]Market-to-book is negatively correlated with size (-21.4%) and positively correlated with volatility (19.4%); we discuss this in section 3.4. Consistent with these patterns, at the portfolio level, growth stocks have higher volatility than value stocks (Fama and French, 1996).

plies that conditional double sorts will produce differences in past returns across characteristics and therefore momentum profits. Before analyzing double sort returns, we discuss the construction of each of the firm-level characteristics in more detail and briefly summarize the related literature as well as arguments that are often made to link these characteristics to information uncertainty and limits to arbitrage.

We note that, while we can replicate the characteristic interaction patterns found in the previous literature, we do not interpret these patterns as necessarily supporting behavioral explanations of momentum; in the next section, we will present evidence that interaction patterns are the result of momentum strategies based on more extreme past returns, as discussed in the analytical section.

(1) **Size** is the equity market capitalization of the firm. It is sometimes interpreted as a proxy for information uncertainty. If a company faces costs of information dissemination or if investors face costs of information gathering (and if these costs are fixed), information quality will be lower for smaller stocks. In addition, given the relatively lower cost of establishing sizable positions in large firms, investors may have more of an incentive to gather information for larger firms. Jegadeesh and Titman (1993), Hong et al. (2000), and Zhang (2006) find that momentum profits are higher among smaller stocks.

(2) $\mathbf{R}^2$: We calculate return R^2 using the same specification we use to calculate idiosyncratic volatility. Our calculation follows Hou et al. (2006) who develop a model in which investor behavior is related to R^2 and who show that the momentum anomaly is more prominent for low R^2 stocks. Return R^2 has been argued to be a proxy for quality of information, e.g. Teoh et al. (2009) argue that prices of low R^2 stocks incorporate less information about their future fundamentals. They suggest that the pattern arises from behavioral biases that are amplified by information uncertainty.[21]

(3) **Turnover**: We compute average monthly share turnover as the monthly number of shares traded divided by the number of shares outstanding averaged over the past 12 months.[22]

---

[21] An alternative interpretation is that low R^2 reflects a high level of firm-specific information. Morck et al. (2000) and Durnev et al. (2003) argue that low R^2 is associated with a high information quality environment.

[22] At least six months of data is required to compute turnover. For NASDAQ stocks, we divide the

Lee and Swaminathan (2000) show that companies with high past turnover have higher momentum returns. This relationship is inconsistent with the idea that turnover is a proxy for limits to arbitrage since momentum profits are higher (not lower) for high turnover stocks. They explain this finding by showing that high turnover stocks exhibit glamour characteristics. It may, therefore, be the case that turnover is higher for stocks that are more difficult to value, which means that we can interpret turnover as a proxy of information uncertainty.

(4) **Age** is the number of years since the firm's IPO date, if available. Otherwise, we calculate age as the number of years since the firm first appears in CRSP. Age may also be a proxy for information uncertainty (Zhang, 2006) if there is more information available for firms which have been in operation for a long time or belong to mature industries. For example, technology firms are likely young and investors may have both less firm-specific information and less skill to value stocks in the industry.

(5) **Analyst coverage** is defined as the number of earnings estimates in I/B/E/S. Analyst coverage is used as a momentum interaction variable by Hong et al. (2000) and Zhang (2006). Firm specific news may be more readily available to investors if a firm has more analysts covering it. Analyst coverage may therefore be related to theories of underreaction (e.g. Hong and Stein, 1999). We include firms with at least one analyst.

(6) **Forecast dispersion**: We measure the standard deviation of earnings per share (EPS) forecasts scaled by the mean EPS forecast. Forecast dispersion can be interpreted as a measure of uncertainty or heterogeneous beliefs. Zhang (2006) considers behavioral models of momentum and uses forecast dispersion as one proxy for uncertainty. Verardo (2009) uses forecast dispersion as a proxy for heterogeneous beliefs, which may be related to price drift (Allen et al., 2006)). Following Diether et al. (2002) we use data from the I/B/E/S summary files to measure forecast dispersion. Data on analyst coverage and forecast dispersion is available starting January 1976. We include firms with at least two forecasts.

(7) **Market-to-book**: We measure book equity following Davies et al. (2000), and use the detailed definition provided in Cohen et al. (2003). Following Fama and French (1996) we do

---

number of shares traded by two to avoid double-counting of dealer trades (see Gould and Kleidon, 1994).

not include negative book equity observations. We then scale market equity capitalization by book equity. Asness (1997) and Daniel and Titman (1999) point out that momentum profits are high for high market-to-book stocks. Sagi and Seasholes (2007) confirm the relationship by implementing a conditional double sort on market-to-book and past returns. This pattern can be linked to a behavioral interpretation of momentum: If high market-to-book stocks' value derives mainly from difficult-to-value growth options, then market-to-book may be viewed as another proxy for uncertainty.

(8) **Price**, (9) **Illiquidity**: We use the stock's price per share as another characteristic. Following Amihud (2002) we measure illiquidity as the absolute value of weekly log returns divided by weekly dollar trading volume, averaged over the previous 52 weeks.[23] Both price and illiquidity can be argued to be related to transaction costs and limits to arbitrage (Stoll, 2000; Amihud, 2002). It is therefore natural to expect that the momentum anomaly will be more prominent for stocks with low prices and high levels of illiquidity. However, we are not aware of studies documenting that either of these variables enhance momentum profits; in Section 4 we show that unadjusted momentum profits are indeed higher for illiquid stocks and those with low prices.

(10) **Credit rating**: We add to these measures the firm's credit rating, which Avramov et al. (2007) find to be strongly related to momentum (profits are higher for stocks with a low credit rating). We use S&P long-term issuer credit rating from COMPUSTAT and assign a numerical score from 1 (AAA) to 22 (for D or SD rating).[24] The average of the monthly median rating is BBB and 61% of firm months have an investment grade rating (BBB- or higher).

The samples for which the different characteristics are available are not all the same, e.g. analyst coverage and forecast dispersion are only available starting in 1976. We have checked that average returns from a decile past return sort are similar across all the samples.

---

[23]We require at least 26 weeks of data to compute illiquidity. To calculate weekly dollar volume we multiply daily shares traded by the closing price and sum the daily volume over the week.

[24]Specifically: AAA = 1, AA+ = 2, AA = 3, AA- = 4, A + = 5, A = 6, A- = 7, BBB+ = 8, BBB = 9, BBB- = 10, BB+ = 11, BB = 12, BB- = 13, B+ = 14, B = 15, B- = 16, CCC+ = 17, CCC = 18, CCC- = 19, CC = 20, C = 21, and D/SD = 22.

Data on credit rating is only available from December 1985 and companies that have an available credit rating are larger. We briefly compare the rating sample to the full sample. The total number of unique rated stocks is 3,393, with an average monthly count of 1,048, which is about 32% of the average monthly count of the full sample. The median rated company is 9 times larger than the median unrated company and rated companies represent 79% of the market capitalization of the total universe of stocks. Consistent with Avramov et al. (2007) we find that rated companies have average monthly returns and momentum payoffs comparable to those of unrated companies.[25]

To control for outliers, we winsorize the data for all of the characteristics, as well as $IVOL$, past returns, and momentum strength at the 0.5% and 99.5% levels, which means that each month we replace the smallest 0.5% of the cross-sectional distribution with the $0.5th$ percentile and the largest 0.5% with the $99.5th$ percentile.

3.4   RETURNS AND CHARACTERISTICS OF PORTFOLIOS WITH EXTREME PAST RETURNS

A direct implication of our analytical framework is that recent winners and losers have higher than average levels of volatility. We briefly formalize this intuition: A single sort of stocks into groups based on their past returns uses the entire cross section of stocks and in particular includes stocks with high and low volatility. Let $\Pr(\sigma)$ denote the unconditional volatility p.d.f. Conditional on being a winner, the conditional volatility p.d.f. is $\Pr(\sigma|r > c)$. We can compare the two distributions by considering the relative likelihood: $\frac{\Pr(\sigma|r>c)}{\Pr(\sigma)} = \frac{\Pr(r>c|\sigma)}{\Pr(r>c)}$ (from Bayes rule), where the cutoff c and the denominator are fixed (e.g. for a decile sort). If returns are normally distributed $\Pr(r > c|\sigma) = \left(1 - N\left(\frac{c}{\sigma}\right)\right)$, which is increasing in σ. Thus, relative to the unconditional distribution, recent winners are more likely to have high volatility. If volatility and characteristics are correlated, recent winners and losers also have more extreme characteristics. We now check these predictions in the data.

---

[25]The decile strategy momentum payoffs are 1.55% (t-statistic = 3.9) for rated stocks and 1.71% (t-stat = 4.4) for unrated stocks. Both of these are close to 1.61%, the return from a decile momentum strategy for the full universe of stocks.

In Table I we sort stocks into ten deciles depending on their past returns (P1, losers, to P10, winners) and report summary statistics for each portfolio. Panel A reports return statistics for the 10 portfolios: momentum returns, risk-adjusted returns (we discuss these in detail in the next sub-section), past returns, momentum strength, and return volatility. Winners outperform losers by 1.6% per month and the same is true for risk-adjusted returns. Past losers underperformed past winners by 108% and both groups have more extreme returns than stocks in the middle of the past return distribution. Momentum strength of recent winners and losers is equal to 53.1% (P1, P10) compared to 2.6% (P5, P6). This U-shaped pattern is also present for volatility: Portfolios P1 and P10 have volatility of 44.9%, substantially higher than the 26.9% volatility of stocks in the middle of the distribution (P5, P6).

Importantly, more extreme past returns are associated with higher momentum profits. Consistent with Jegadeesh and Titman (1993) and Fama and French (1996) there is a monotonic relationship between momentum returns (row 1) and past returns (row 3). In other words, a more extreme momentum sort, long P10 and short P1, has a higher momentum return (1.61%) than does a strategy based on five portfolios, long the top quintile winners, short the bottom quintile losers (1.11%, t-statistic 5.56). The pattern is also present for more extreme momentum sorts: Using a 25-portfolio sort (trading in stocks in the top and bottom 4% of the past return distribution) results in a return of 2.20% (t-statistic 7.28), and a strategy based on a 50-portfolio sort has a return of 2.43% (t-statistic 7.21).

Panel B reports average levels of characteristics for each of the 10 portfolios. Since some of the variables (e.g. size) have somewhat skewed distributions we report averages of monthly medians. For all characteristics there is a pronounced U-shaped or inverted U-shaped pattern across momentum-sorted portfolios. Extreme momentum stocks (P1, P10) are 35% smaller, have 15% lower return R^2, 82% higher turnover, are 39% younger, have 23% lower analyst coverage, twice as large analyst forecast dispersion, 51% higher market-to-book ratios, 26% lower prices, and 96% higher illiquidity than firms in the middle of the past return distribution (P5, P6). Extreme past return stocks also have substantially higher levels of credit risk (credit rating score of 12 (BBB) compared to a score of 8 (A+)). In Table II we sort stocks on volatility and show that stocks with high volatility have similar characteristics to those with extreme

past returns.

Table III provides additional evidence of a link between extreme past returns and characteristics. For each month we calculate the rank correlation of momentum strength and volatility with each of the characteristics and report average correlations. We also report the share of months for which the correlation is significant at the 1% level. There is a consistently high and statistically significant correlation between volatility and each of the characteristics and the same is true for extreme past returns (momentum strength) and characteristics.

3.5 MOMENTUM PROFITS AND RISK-ADJUSTED RETURNS

Extreme past return stocks have extreme characteristics, in particular they tend to be small and have high market-to-book ratios, both of which are known to be related to expected returns (see e.g. Fama and French, 1992). It is therefore important to risk-adjust momentum returns. Grundy and Martin (2001) show that a momentum strategy by construction has time-varying factor exposures that depend on factor realizations during the formation period. The intuition is easy to see in the context of a single factor model: If the market has a positive return during the formation period, then winners will tend to be high beta stocks while losers are more likely low beta stocks. The opposite is the case for a momentum strategy after a down market. The changing factor loadings of a momentum strategy mean that it is not appropriate to assume that loadings are constant through time.

We follow Korajczyk and Sadka (2004) who propose a straightforward way to calculate dynamic Fama and French (1993) three factor alphas. Table I Panel A reports risk adjusted returns (second line). Consistent with the findings in Grundy and Martin (2001), risk-adjustment does not eliminate momentum profits. Recent winners (P10) have an alpha of 1.15% and outperform recent losers (P1) that have an alpha of -0.43%. The difference is equal to 1.58%, which is close to the average monthly momentum return of 1.61%. The difference in alphas of winners and losers is highly statistically significant and has a t-statistic of 9.06. Since risk-adjusting reduces volatility, the difference in alphas is more significant than the momentum return, which has a t-statistic of 6.46.

Controlling for risk is particularly important when implementing double sorts, which we do

18

in the next section. Sorting stocks first using a characteristic, then on past returns, implies that extreme characteristic momentum portfolios may also have extreme levels of factor loadings. We therefore report alphas along with average portfolio returns. However, consistent with the results in Table I, we find that patterns in momentum returns and risk-adjusted returns are very similar.

4   Sources of Momentum Profits

Having established a link between characteristics, volatility, and extreme past returns and confirmed that momentum profits are higher for stocks with more extreme past returns, we now turn our attention to the main focus of the paper: Do characteristic screens truly enhance momentum? In this section we show that characteristics empirically 'enhance' profits because sorting on characteristics results in momentum strategies with more extreme past returns. The least extreme returns are screened out and the resulting momentum strategy produces higher momentum profits.

We begin by following the methodology of the existing literature and report results from conditional double sorts (Table IV). Every month, we sort stocks into quintiles first by one of the characteristics and then by past returns. For each quintile of the characteristic (Q1 to Q5) we report average monthly returns for the momentum long-short portfolio (P5-P1), as well as for past winners (P5), and past losers (P1). For each characteristic we also report the difference in momentum long-short returns between Q5 and Q1 ('Return diff') and the difference in risk-adjusted returns; below the momentum returns we report t-statistics in parentheses. The final two columns of the table report the difference in average momentum strength and volatility between Q5 and Q1 for the stocks included in the long-short (P5-P1) portfolio.

A double sort of volatility and past returns results in a more extreme momentum sort and higher momentum profits. The average momentum return for stocks in the highest volatility quintile is equal to 1.96% per month, while momentum returns in the lowest volatility quintile are close to zero. The difference between high and low volatility quintile momentum profits is 1.93% and has a t-statistic of 9.36. Risk-adjusted returns are 1.92% higher for Q5 relative

to Q1. High-volatility momentum portfolios have much higher momentum strength than low-volatility portfolios; the differences is equal to 59.2%.

We perform the same exercise for the ten characteristics. To make results more easily comparable we order stocks depending on the characteristic's correlation with volatility: ascending order for positive correlation (turnover, forecast dispersion, market-to-book, illiquidity, credit rating), descending order for negative correlation (size, R^2, age, analyst coverage, price). This way momentum returns are everywhere higher for stocks in quintile Q5 (high volatility and extreme past returns) than they are for stocks in quintile Q1.

Similar to volatility, sorting on characteristics produces differences in momentum profits. For all of the characteristics momentum profits for stocks within the high characteristic quintile (Q5) are larger than momentum profits for Q1.[26] For most of the characteristics, momentum profits increase almost monotonically as we move from Q1 to Q5. For example, momentum returns are equal to 0.53% per month for large stocks, 1.17% for stocks in the middle 20% of the size distribution (Q3), and 1.34% per month for small stocks. The difference in returns is largest for credit rating (2.43% per month), and weakest for R^2 (0.16%), the only characteristic for which the return difference is not statistically significant.[27] The patterns in magnitude and statistical significance of risk-adjusted returns closely track those of standard momentum returns.

Consistent with our analytical framework we find that the more closely a characteristic proxies for volatility, the higher the past return spread and the higher the momentum profits. Sorting first on characteristics produces more extreme past returns: for each of the ten characteristics, the Q5 long-short momentum portfolio contains stocks with more extreme and more

---

[26]We have also checked that, consistent with Avramov et al. (2007), average momentum returns are higher for firms with non-investment grade ratings (2.51% (t-statistic 5.0)) than for firms with investment grade ratings (0.44% (t-statistic 1.4)).

[27]Our results are different from Hou et al. (2006) who find that a double sort on R^2 and momentum produces significant differences in momentum returns for high and low R^2 stocks. There are three reasons for this: first, Hou et al. (2006) report value-weighted returns, second, they use NYSE R^2 quintile break points (which means that a large fraction of AMEX and NASDAQ stocks are in the lowest R^2 quintile), third, they use full-sample R^2 estimates. We instead report equally-weighted returns, use quintile breakpoints for the entire set of stocks each period, and calculate rolling estimates for R^2 (as we do for all other characteristics). However, we have verified that if we make these adjustments we can replicate their results.

volatile past returns than Q1. Indeed the correlation of momentum strength differences and correlations of characteristics with volatility (i.e. the correlation of 'Mom str diff' (Table IV) and '$IVOL$' (Table III)) is equal to 0.99. This relationship is then reflected in the patterns in momentum profits: Characteristics for which volatility and momentum strength differences are small (e.g. R^2) also have lower momentum return differences than those for which volatility and momentum strength differences are large (e.g. credit rating). The correlation between momentum return differences and momentum strength differences is 0.78.

Given the link between past returns and momentum profits it is not surprising that characteristic sorts produce high returns. Characteristics not only select stocks which deliver high momentum *profits*, but they also select stocks with the most extreme *volatility* and *past returns*.

4.1 CONTROLLING FOR VOLATILITY

We now calculate momentum profits adjusting for the effect of volatility. Our aim is to measure the *incremental* ability of the characteristic to interact with momentum. We isolate the part of the characteristic that is unrelated to volatility by calculating residual characteristics from linear regressions of characteristics on volatility. We then sort stocks into quintiles using residual characteristics and calculate momentum profits.

The relationship between characteristics and volatility may be non-linear and we therefore, whenever appropriate, transform characteristics and volatility before running regressions. For each of the characteristics we explore the relationship by considering 25 volatility-sorted portfolios. For characteristics we consider taking logs and for $IVOL$ we consider the portfolio rank (1 to 25). For each characteristic we choose the specification (characteristic or log(characteristic), average $IVOL$ or $IVOL$ rank) that has the highest 25 portfolio regression R^2. The 25-portfolio regression R^2 is equal to 98.5% on average, reflecting the close to linear relationships.[28]

Each month we run cross-sectional regressions of characteristics on volatility for 25 $IVOL$-

---

[28]The specifications are log(characteristic)-$IVOL$ rank for size, turnover, age, analyst coverage, price, credit rating; log(characteristic)-$IVOL$ for R^2, market-to-book; and characteristic-$IVOL$ for illiquidity and forecast dispersion. Hong et al. (2000) and Campbell et al. (2008) also use monotonic transformations to calculate residual characteristics using a linear model.

sorted portfolios[29] and compute residual characteristics by applying the monthly portfolio regression estimates to the stock-level data. This procedure allows us to isolate the effect of the characteristic on momentum profits that is in excess of the effect of volatility. If enhanced momentum profits are driven mainly by volatility, sorting by residual characteristics should result in a substantial reduction of returns or no enhancement of returns.

Table V reports the results. For each characteristic we sort stocks into quintiles using residual characteristics. As in Table IV, Q1 contains the stocks with low residual characteristics and Q5 contains the stocks with high residual characteristics, except for size, R^2, age, analyst coverage, and price, for which we use descending order. For each quintile (Q1 to Q5) we report momentum profits long the quintile of winners, short the quintile of losers; the final two columns report differences in momentum profits and risk-adjusted momentum returns between Q5 and Q1. These differences are directly comparable to the momentum return differences in Table IV.

In all cases the difference between momentum profits for high and low characteristics (Q5 and Q1) is substantially reduced and for all characteristics except age the enhancing effect is no longer statistically significant. These patterns are in contrast to the pronounced enhancing ability of characteristics documented in Table IV. When sorting by residual size, R^2, turnover, analyst coverage, and illiquidity, the difference between Q5 and Q1 momentum profits is either virtually equal to zero (below 0.2%) or slightly negative and not statistically significant. For market-to-book and credit rating the difference in momentum profits is equal to 0.30% and 0.27% respectively, though neither difference is statistically significant and both are smaller than the unadjusted differences of 0.75% and 2.45%. For analyst forecast dispersion and share price the interaction effects actually reverse.[30] Adjusting for volatility cuts the enhancing effect

---

[29]We run the regression on 25 portfolios in order to reduce noise.

[30]For most of the variables the relationship between volatility and the characteristic has the expected sign each month, e.g. in every period high volatility stocks are small. However, for R^2 (4% of the time), illiquidity (13% of the time) and market-to-book (8% of the time) the sign of the relationship switches. We have checked that our results are robust to setting the coefficient to zero during those times, that is we constrain the regression coefficient to be of the same sign throughout. Also, for illiquidity, the relationship with volatility is much weaker between 1997 and 2008. This pattern does not affect our results since during this period illiquidity (not adjusted for volatility) loses its ability to enhance momentum profits.

of age in half; the return difference declines from 1.52% to 0.79%. Patterns in magnitudes and significance levels are again similar to those of risk-adjusted returns.

Since age is the only effect for which the return difference remains positive and statistically significant we briefly explore it further. We find that the effect of residual age is driven mainly by very young firms; if we exclude firms with age less than 3 years, the effect of residual age on momentum profits disappears.[31] Higher momentum profits for very young firms may therefore be the result of IPO underperformance (see e.g. Ritter, 1991).

To summarize: Having established that characteristics are closely related to volatility we find that this relationship almost entirely explains characteristics' ability to enhance momentum profits. It is thus not characteristics that enhance momentum but instead extreme past returns that lead to elevated profits.

4.2   CONTROLLING FOR EXTREME PAST RETURNS

We also ask if the results are robust to instead adjusting directly for the relationship between past returns and forward returns. Table VI reports the results. For reference, column 1 reports momentum profits for the highest quintiles for each characteristic (the numbers are the same as in Table IV, Q5). In column 2 we report risk-adjusted returns to confirm that extreme characteristic momentum alphas are significantly different from zero. As before, we find that momentum returns and risk-adjusted returns are similar.[32]

The momentum returns reported in Table I demonstrate that there is a direct relationship between past returns and returns over the next month. We adjust for this effect directly by assuming a linear relationship between past returns and forward returns and have checked that this assumption is reasonable. Using time-series averages of monthly 25 past-return-sorted

---

[31]We find that the enhancing ability of age is not driven by the inclusion of NASDAQ stocks; if we exclude these, the effect remains. It is also not driven by the possible mis-measurement of age; including only stocks for which we have the IPO date does not affect the results.

[32]As an additional check, in unreported results we calculate Fama and French (1993) 3-factor alphas and Carhart (1997) 4-factor alphas. We find that, consistent with Fama and French (1996), the 3-factor model does not explain momentum profits; momentum alphas are large and significant for extreme characteristics. For all characteristics 4-factor alphas are substantially smaller, about half as large as 3-factor alphas.

portfolios we run one regression of average forward returns $(t+1)$ on the log of average past returns $(t-6, t-1)$. We then calculate predicted momentum profits; that is, we compute the average past return spread between winner and loser portfolios within each characteristic group and compute predicted momentum profits based on the regression estimates. In column 3 of Table VI we report returns adjusted for predicted momentum returns. If it is the level of past returns (and not the extreme level of the characteristic) that is related to momentum profits, we expect momentum returns adjusted in this way to be substantially smaller than the unadjusted returns resulting from conditional double sorts using characteristics.

Consistent with the results when adjusting for volatility, momentum profits adjusted for past returns are significantly reduced. Momentum profits for extreme characteristics are 1.5% on average (average of column 1 returns), while adjusted returns are 0.2% on average (average of column 3 returns). In all cases adjusted returns are much smaller than unadjusted returns. Adjusted returns are indistinguishable from zero in all cases except for age (t-statistic of 1.92 and thus marginally significant) and credit rating (t-statistic of 2.34). Adjusted returns for low age (young) firms are much smaller (0.45%) than unadjusted returns (1.77%). For credit rating, adjusted returns are reduced from 2.4% to 1.3%. The small remaining effect of credit rating on momentum profits may be driven by low-grade credit rating stocks being more distressed and distressed stocks having a tendency to underperform (see Campbell et al., 2008). We conclude that most of the variation in momentum profits that is present when implementing a double-sort using characteristics and past returns can be explained by variation in past returns.

In our theoretical discussion in Section 2 we demonstrate that sorting on a proxy for volatility will directly result in more extreme past returns for high characteristic winners and losers. We now adjust for this effect directly. For each characteristic we calculate the past return spread, the difference between the past return for winners and the past return for losers. We then find the percentile of the single sort past return distribution that matches the past return spread.

For example, a quintile double sort on size and past return results in a past return spread for small firms of 85%. Implementing a single sort, the past return spread between the top 17% (winners) and the bottom 17% (losers) of the past return distribution matches this past

return spread (we find the percentile with the closest past return spread). Therefore the appropriate benchmark for a double sort using size is the $17th$ percentile single-sort momentum strategy. Using the notation from the theoretical discussion, we find $c_{r,s}^{match}$ such that $E\left[r|s < c_s, r > c_{r,s}\right] = E\left[r|r > c_{r,s}^{match}\right]$. We then check the percentile of the single sort past return distribution to find $Q_s^{match} = \Pr\left(r > c_{r,s}^{match}\right) = 17\%$.[33]

For each of the 10 characteristics we report the matched percentile (column 4). In all cases quintile double sorts represent more extreme momentum sorts. The matched percentiles range from 17% (size, R^2, analyst coverage) to 7% (credit rating). The average of the percentiles is equal to 13.4%. For each characteristic we calculate the momentum returns for the matched percentile cutoff and report the difference between extreme characteristic returns and matched single-sort returns (column 5).[34] We find that double-sort returns are almost the same as matched momentum returns; across all ten characteristics the difference is equal to 0.07% per month on average. In other words, using double sorts does not produce profits that are higher than single sort profits, so long as past returns are the same.

Using the methodology in Table VI we perform another check: It is well documented (e.g. Jegadeesh and Titman, 1993; Grundy and Martin, 2001) that the momentum strategy is much weaker in January and often reverses sign. It is possible that some of the characteristics we consider are able to select stocks that are not affected by the documented January reversal of momentum and that this is the reason they outperform a single sort momentum strategy. We find that this is not the case: our results are robust to excluding Januaries.

4.3   REGRESSION-BASED APPROACH

Another way to control for the effect of past returns on momentum profits is to use a regression-based framework.[35] Characteristics, volatility, and extreme past returns are correlated which

---

[33]We note that, though not perfect, our analytical model (section 2) reflects the empirical patterns well: The model implied matched percentile cutoff is 14.6% compared to an actual matched cutoff of 17%.

[34]Since the sample is not the same across all characteristics (e.g. credit rating, analyst coverage), we peform the matching and return procedure for each sample separately.

[35]Sorting stocks into portfolios has benefits and drawbacks, but in the case of correlated drivers of momentum using a regression-based approach will shed more light on what drives variation in returns.

makes a regression-based framework a natural choice to explore their effects on profits. We run Fama-MacBeth regressions of monthly stock-level momentum profits first on characteristics only and then controlling for momentum strength and volatility.

Momentum profits result from winners outperforming and losers underperforming, in other words, profits result from relative return continuation. We define winners and losers as having above or below median past returns, $R_{t-6,t-1} - R_{median,t-6,t-1}$, and measure forward returns relative to the median, $R_{t+1} - R_{median,t+1}$.[36] The stock's momentum profit is then equal to the stock's forward return multiplied by a winner/loser dummy equal to 1 if the stock was a winner and -1 if the stock was a loser:

$$R_{mom,t+1} = (R_{t+1} - R_{median,t+1}) \times sign\left(R_{t-6,t-1} - R_{median,t-6,t-1}\right). \tag{4}$$

Stock-level momentum profits $R_{mom,t+1}$ are positive if either the stock has an above-median forward return and was a winner or if the stock has a below-median forward return and was a loser. Both of these cases (in which past and forward returns are of the same sign) contribute to a momentum strategy being profitable. If the return switches sign between formation and holding period, momentum profits are negative.

We regress momentum profits on log(characteristic), $IVOL$ rank (1 to 25), and momentum strength. We check that momentum profits are close to linear in these variables in order to reduce potential biases from regression misspecification. Table VII reports the results.

We first run univariate regressions of momentum profits on characteristics only (column 1). For each of the characteristics we find an effect consistent with the one documented in Table IV: Momentum profits are high for firms that are small, have low R^2, high turnover, are young, have low analyst coverage, high analyst forecast dispersion, high market-to-book ratios, low share price, are illiquid, and have high-risk credit ratings (junk). We then control for the effect of past returns by including momentum strength and $IVOL$ rank together with the characteristic. If momentum profits are high because past returns are extreme, then the

---

Fama and French (2008) discuss benefits and drawbacks of using sorts and regressions when exploring what drives cross-sectional returns. See also Berk (2000) for a more general discussion of sorts.

[36]To make results comparable with our analysis in Table IV we calculate medians within characteristic quintile.

magnitude of the coefficients on characteristics should be reduced. This is exactly what we find. Coefficients decline substantially for all characteristics and in some cases (R^2, price, and turnover), the coefficient switches sign. The effect of age remains statistically significant, though controlling for past returns cuts the coefficient in half (from -0.154 to -0.079). Credit rating remains marginally significant, but controlling for past returns reduces the size of the coefficient by two thirds (from 0.437 to 0.144).[37]

For all regressions both momentum strength and $IVOL$ are statistically significant. This means that it is mainly variation in past returns that explains variation in momentum profits, not variation in characteristics. The evidence reported in Table VII is consistent with our previous results and supportive of our hypothesis that differences in past returns are an important source of variation in momentum profits.

We can also draw a conclusion regarding the methodology of double-sorting: Using double sorts to investigate the effects of characteristics on momentum profits may not be informative if characteristics are correlated with extreme past returns. Characteristics, past returns, and volatility are all highly correlated and it is a priori not clear what explains variation in momentum profits. In a regression it is possible to sort out the effects of correlated variables and, given the results in Table VII, we can conclude that for most of the characteristics higher momentum profits are the result of trading in stocks with more extreme past returns.

5   Conclusion

Several recent studies have documented that momentum profits are more pronounced for stocks with certain characteristics. In this paper we demonstrate that there is a common channel that can explain these patterns: Characteristic screens result in elevated momentum profits by excluding stocks with less extreme past returns from a momentum strategy. Characteristic interaction returns are thus easily matched by a more extreme momentum sort.

There are two main implications of our findings. First, a search for 'enhanced' momentum strategies needs to account for variation in past returns. In fact, we demonstrate that there

---

[37]As discussed, the remaining effect may be the result of the underperformance of distressed stocks.

is virtually no benefit from conditioning a momentum strategy on stock-level characteristics. This means that, from an investment perspective, focusing on past returns is sufficient when aiming to maximize momentum profits.

Second, existing explanations of momentum, in particular those based on the apparent ability of characteristic screens to enhance momentum profits, need to be reappraised. Many of the interaction characteristics are common proxies for information uncertainty and limits-to-arbitrage and thus the literature has enlisted the evidence to support behavioral explanations of momentum. However, the empirical patterns may also be explained by rational models. Our results imply that an explanation of momentum needs to take as a starting point the link between volatility, past returns, and momentum profits. Vayanos and Woolley (2010) present a rational model of delegated portfolio management in which stocks with high idiosyncratic volatility are predicted to have higher momentum profits. Another possibility is that stock returns are generated by a factor model. In such a framework more extreme past returns may lead to increased momentum profits (Grundy and Martin, 2001). The existing evidence on momentum interaction effects therefore need not support behavioral explanations of momentum.

Of course, volatility itself has been used as a measure of limits to arbitrage (e.g. Shleifer and Vishny, 1997) and it is also interpreted as a proxy for information uncertainty (e.g. Zhang, 2006). However, Brav, et al. (2010) question the extent to which high volatility and limits-of-arbitrage explanations can be applied to anomalous stock returns more generally. Shen (2008) challenges the idea of using volatility as a proxy for information uncertainty arguing that prices of high idiosyncratic volatility stocks may in fact be more informative. We note also that more direct proxies of information uncertainty such as analyst coverage and analyst forecast dispersion are not associated with more extreme momentum returns once their relationship to past returns is taken into account.

Our paper also provides a framework for understanding studies that have documented an increasing number of momentum 'sub-anomalies.' A reader of the existing literature might well pose the question: Do the interaction effects of different characteristics represent many effects (one for each characteristic) or are they the result of one effect (there is a common channel that explains elevated momentum profits)? We show that one effect – variation in past returns

28

– can explain the observed patterns. There is thus scope for shrinking the dimensionality of explanations of characteristic interaction patterns: The relationship between past returns and momentum profits can itself explain the momentum interaction effects.

A   Appendix

A.1   VOLATILITY AND PAST RETURNS

In this appendix we analyze past return single sorts and conditional double sorts on characteristics and past returns. Given dispersion in volatility σ, conditional on being a winner $(r > c)$, and assuming a joint discrete distribution of volatility and returns $\Pr(r, \sigma)$, we can write the conditional expected return as:

$$
\begin{aligned}
E\left[r|r > c\right] &= \sum_{\sigma} \sum_{r>c} r \frac{\Pr(r, \sigma)}{\Pr(r > c)} \\
E\left[r|r > c\right] &= \sum_{\sigma} E\left[r|r > c, \sigma\right] \frac{\Pr(r > c|\sigma) \Pr(\sigma)}{\Pr(r > c)}.
\end{aligned}
\tag{5}
$$

We assume that conditional on volatility σ returns are distributed normally, $r \sim N\left(0, \sigma^2\right)$. The conditional expected return is then given by (see Johnson and Kotz, 1970, p.81; Grundy and Martin, 2001, p.73):

$$
E\left[r|r > c, \sigma\right] = \frac{n\left(\frac{c}{\sigma}\right) \sigma}{1 - N\left(\frac{c}{\sigma}\right)}
$$

where $n(.)$ is the normal p.d.f. and $N(.)$ is the normal c.d.f. Also, conditional on volatility, the probability of a return above the cutoff c is

$$
\Pr\left(r > c|\sigma\right) = \left(1 - N\left(\frac{c}{\sigma}\right)\right).
$$

This means that, conditional on volatility σ, sorting stocks into the top group with probability Q, e.g. the top quintile, means that $E\left[r|r > c, \sigma\right] = \frac{1}{Q} n\left(\frac{c}{\sigma}\right) \sigma$. Not surprisingly, past returns are more extreme for high volatility stocks since $n\left(\frac{c}{\sigma}\right) \sigma$ is increasing in σ:

$$
\begin{aligned}
\frac{\partial}{\partial \sigma} n\left(\frac{c}{\sigma}\right) \sigma &= \frac{\partial}{\partial \sigma} \frac{\exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right) \sigma}{\sqrt{2\pi}} \\
&= \frac{\exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right) \frac{c^2}{\sigma^3} \sigma}{\sqrt{2\pi}} > 0.
\end{aligned}
$$

Sorting on past returns also has an effect on volatility (and characteristics). The probability weights in Equation (5) used to sum over different conditional expectations $E[r|r > c, \sigma]$ are $\frac{\Pr(r>c|\sigma)\Pr(\sigma)}{\Pr(r>c)}$, where the denominator is fixed and positive (equal to Q). We can compare these weights to using the unconditional probability distribution of volatility $\Pr(\sigma)$ by considering their ratio $\frac{\Pr(r>c|\sigma)\Pr(\sigma)}{\Pr(r>c)\Pr(\sigma)} = \frac{\Pr(r>c|\sigma)}{\Pr(r>c)} = \frac{\left(1-N\left(\frac{c}{\sigma}\right)\right)}{\Pr(r>c)}$, which is increasing in σ since

$$
\begin{aligned}
\frac{\partial}{\partial \sigma} \Pr(r > c|\sigma) &= \frac{\partial}{\partial \sigma}\left(1 - N\left(\frac{c}{\sigma}\right)\right) \\
&= -n\left(\frac{c}{\sigma}\right)\left(-\frac{c}{\sigma^2}\right) > 0.
\end{aligned}
$$

Thus when sorting on past returns $(r > c)$ relatively more weight is placed on observations with higher volatility. An implication of this is that if a characteristic is correlated with volatility a sort on past returns will produce a spread in characteristics also (Table I).

To find the conditional expected return we can now plug into the conditional expectation (from above):

$$
\begin{aligned}
E[r|r > c] &= \frac{1}{\Pr(r > c)} \sum_{\sigma} \frac{n\left(\frac{c}{\sigma}\right)\sigma}{1 - N\left(\frac{c}{\sigma}\right)}\left(1 - N\left(\frac{c}{\sigma}\right)\right)\Pr(\sigma) \\
&= \frac{1}{\Pr(r > c)} \sum_{\sigma} n\left(\frac{c}{\sigma}\right)\sigma \Pr(\sigma) \\
&= \frac{E\left[n\left(\frac{c}{\sigma}\right)\sigma\right]}{\Pr(r > c)}.
\end{aligned}
$$

Assuming that volatility has a uniform distribution, $\sigma \epsilon [\sigma_{\min}, \sigma_{\max}]$, the conditional expected return is given by

$$
\begin{aligned}
E[r|r > c] &= \frac{1}{\Pr(r > c)} \int_{\sigma_{\min}}^{\sigma_{\max}} n\left(\frac{c}{\sigma}\right)\frac{\sigma}{\sigma_{\max} - \sigma_{\min}}d\sigma \\
&= \frac{1}{\Pr(r > c)}\frac{1}{\sqrt{2\pi}(\sigma_{\max} - \sigma_{\min})}\int_{\sigma_{\min}}^{\sigma_{\max}} \exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right)\sigma d\sigma.
\end{aligned}
$$

This integral has an analytical solution in terms of the incomplete *Gamma* function $\Gamma(.,.)$

(from Mathematica):

$$\int_{\sigma_{\min}}^{\sigma_{\max}} \exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right)\sigma d\sigma = \frac{1}{4}\left[2\sigma^2\exp\left(-\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right) - c^2\Gamma\left(0,\frac{1}{2}\left(\frac{c}{\sigma}\right)^2\right)\right]_{\sigma_{\min}}^{\sigma_{\max}}.$$

A.2 CHARACTERISTICS AND PAST RETURNS

In order to introduce characteristics as proxies for volatility we assume that characteristics and volatility are distributed bivariate normal.

A.2.1 *Mean and Standard Deviation of Volatility*

We assume that $\sigma \sim N\left(\mu_\sigma, \nu_\sigma^2\right)$. For the example discussed in section 2 we choose parameters to match the data. We therefore first find values for μ_σ and ν_σ so that the truncated mean and standard deviation of volatility matches the data (the mean volatility is 28.4%, the standard deviation is 11.2%). Since (see above)

$$E\left[\sigma|\sigma > 0\right] = \mu_\sigma + \frac{n\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\nu_\sigma}{1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)}$$

and (see Johnson and Kotz, 1970, p.83)

$$var\left(\sigma|\sigma > 0\right) = \nu_\sigma^2\left(1 + \frac{\frac{-\mu_\sigma}{\nu_\sigma}n\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)}{1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)} - \left(\frac{n\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)}{1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)}\right)^2\right)$$

we can numerically find $\mu_\sigma = 0.2817$, $\nu_\sigma = 0.1149$ which results in $E\left[\sigma|\sigma > 0\right] = 28.4\%$, $var\left(\sigma|\sigma > 0\right) = 11.2\%$.

A.2.2 *Single Sort*

In the case of a single sort we only need to consider the unconditional distribution of volatility, which is normal: $\sigma \sim N\left(\mu_\sigma, \nu_\sigma^2\right)$ with p.d.f. $\frac{1}{\sqrt{2\pi}\nu_\sigma}\exp\left(-\frac{1}{2}\left(\frac{\sigma - \mu_\sigma}{\nu_\sigma}\right)^2\right)$. We integrate from $\sigma = 0$ to ∞ and rescale the p.d.f. by $\Pr\left(\sigma > 0\right) = 1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)$.

We first find the cutoff c. Given a level of volatility σ, $\Pr(r > c|\sigma) = \left(1 - N\left(\frac{c}{\sigma}\right)\right)$. To find $\Pr(r > c|\sigma > 0)$ we integrate over the possible values of σ and solve for c:

$$
\begin{aligned}
Q &= \Pr(r > c|\sigma > 0) \\
&= \frac{\int_0^\infty \left(1 - N\left(\frac{c}{\sigma}\right)\right) \exp\left[-\frac{1}{2}\left(\frac{\sigma - \mu_\sigma}{\nu_\sigma}\right)^2\right] d\sigma}{\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right)\sqrt{2\pi}\nu_\sigma}.
\end{aligned}
$$

The single-sort expected return is then given by:

$$
\begin{aligned}
E[r|r > c, \sigma > 0] &= \frac{1}{Q}E\left[n\left(\frac{c}{\sigma}\right)\sigma\right] \\
&= \frac{\int_0^\infty \exp\left[-\frac{1}{2}\left(\left(\frac{c}{\sigma}\right)^2 + \left(\frac{\sigma - \mu_\sigma}{\nu_\sigma}\right)^2\right)\right]\sigma\, d\sigma}{Q\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right)2\pi\nu_\sigma},
\end{aligned}
\tag{6}
$$

which we calculate numerically. To be precise we write $E[r|r > c, \sigma > 0]$ but we note here that we are not explicit about this constraint when discussing the results in section 2.

A.2.3   *Characteristic Past Return Double Sort*

For characteristic double sorts we calculate the return cutoff and the conditional expected return by integrating over values of the characteristic that are included in the momentum strategy. We assume that the characteristic and volatility are distributed bivariate normal. To make the analysis applicable to any characteristic we assume, without loss of generality, that the characteristic is distributed standard normal and positively correlated with volatility. For example, if $\log(size)$ and volatility are distributed bivariate normal we can define $s = -\frac{\log(size) - \mu_s}{\nu_s}$ where μ_s and ν_s are mean and standard deviation of $\log(size)$. Denoting the correlation between s and σ by ρ we then have

$$
\begin{pmatrix} \sigma \\ s \end{pmatrix} \sim N\left(\begin{array}{c} \mu_\sigma \\ 0 \end{array}, \begin{array}{cc} \nu_\sigma^2 & \rho\nu_\sigma \\ \rho\nu_\sigma & 1 \end{array}\right).
$$

We first find the cutoff c_s such that $\Pr\left(s > c_s | \sigma > 0\right) = 20\%$. The unconditional distribution of s is normal but since s is correlated with volatility we need to use the joint distribution to find the size cutoff c_s:[38]

$$
\begin{aligned}
Q_s &= \Pr\left(s > c_s | \sigma > 0\right) \\
&= \frac{\int_{s>c_s} \int_{\sigma>0} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)^2 - 2\rho\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)\left(\frac{s-\mu_s}{\nu_s}\right) + \left(\frac{s-\mu_s}{\nu_s}\right)^2\right]\right)}{\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right) 2\pi\nu_\sigma\nu_s\sqrt{1-\rho^2}}.
\end{aligned}
$$

Since we are focusing on the sub-universe of observations with high characteristic and positive volatility we need to scale the conditional p.d.f. when calculating the return cutoff $c_{r,s}$ and the conditional expected return using the probability:

$$
\begin{aligned}
\Pr\left(s > c_s, \sigma > 0\right) &= \Pr\left(s > c_s | \sigma > 0\right)\Pr\left(\sigma > 0\right) \\
&= Q_s\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right).
\end{aligned}
$$

We can now find the return cutoff $c_{r,s}$:

$$
\begin{aligned}
Q &= \Pr\left(r > c_{r,s} | s > c_s, \sigma > 0\right) \\
&= \frac{\int_{s>c_s} \int_{\sigma>0} \left(1 - N\left(\frac{c}{\sigma}\right)\right) \exp\left(-\frac{\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)^2 - 2\rho\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)\left(\frac{s-\mu_s}{\nu_s}\right) + \left(\frac{s-\mu_s}{\nu_s}\right)^2}{2(1-\rho^2)}\right) d\sigma ds}{2\pi\nu_\sigma\nu_s\sqrt{1-\rho^2}Q_s\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right)}
\end{aligned}
$$

---

[38] If x and y are bivariate normal the p.d.f. is given by (see Casella and Berger, 1990, p.167)

$$
f(x,y) = \frac{\exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right)}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}.
$$

and then the conditional expected return:

$$
E\left[r|s > c_s, r > c_{r,s}, \sigma > 0\right]
$$

$$
= \frac{\int_{s>c_s} \int_{\sigma>0} \sigma \exp\left(-\frac{1}{2}\left(\frac{c_{r,s}}{\sigma}\right)^2 - \frac{\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)^2 - 2\rho\left(\frac{\sigma-\mu_\sigma}{\nu_\sigma}\right)\left(\frac{s-\mu_s}{\nu_s}\right) + \left(\frac{s-\mu_s}{\nu_s}\right)^2}{2(1-\rho^2)}\right) d\sigma ds}{QQ_s\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right)\sqrt{2\pi}2\pi\nu_\sigma\nu_s\sqrt{1-\rho^2}}.
$$

Note that $\Pr\left(r > c_{r,s}, s > c_s, \sigma > 0\right) = \Pr\left(r > c_{r,s}|s > c_s, \sigma > 0\right)\Pr\left(s > c_s, \sigma > 0\right) = QQ_s\left(1 - N\left(\frac{-\mu_\sigma}{\nu_\sigma}\right)\right)$.

The two remaining statistics that we report in the text are $c_{r,s}^{match}$ such that $E\left[r|s > c_s, r > c_{r,s}, \sigma > 0\right] = E\left[r|r > c_{r,s}^{match}, \sigma > 0\right]$ and $\Pr\left(r > c_{r,s}^{match}\right)$. We solve for $c_{r,s}^{match}$ using the expression for single sort conditional expected return given in Equation (6) and taking into account that we need to scale by $\Pr\left(r > c_{r,s}^{match}|\sigma > 0\right)$, which itself depends on $c_{r,s}^{match}$. We also calculate $\Pr\left(r > c_{r,s}^{match}\right)$ using the single sort probability of lying above a cutoff.

Instead of solving the integrals numerically, another possibility is to use simulations. We have checked that our results are the same using such a methodology.

References

Allen, F., Morris S., and Shin H.S. (2006) Beauty Contests and Iterated Expectations in Asset Markets, *Review of Financial Studies*, **19**, 719-752.

Amihud, Y. (2002) Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets*, **5**, 31-56.

Ang, A., Hodrick, R., Xing, Y., and Zhang X. (2006) The Cross-Section of Volatility and Expected Returns, *Journal of Finance*, **56**, 259-299.

Arena, M.P., Haggard, S.K., and Yan, X.S. (2008) Price Momentum and Idiosyncratic Volatility, *Financial Review*, **48**, 159-190.

Asness, C.S., (1997) The Interaction of Value and Momentum Strategies, *Financial Analysts Journal*, **53**, 29-36.

Avramov, D., Chordia, T., Jostova, G., and Philipov, A. (2007) Momentum and Credit Rating, *Journal of Finance*, **62**, 407-427.

Barberis, N., Shleifer, A., and Vishny, R. (1998) A Model of Investor Sentiment, *Journal of Financial Economics*, **49**, 307-343.

Berk, J.B., (2000) Sorting Out Sorts, *Journal of Finance*, **55**, 2503-2520.

Brav, A., Heaton, J.B., and Li, S. (2010) The Limits of the Limits of Arbitrage, *Review of Finance*, **14**, 157-187.

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

Carhart, M., (1997) On Persistence in Mutual Fund Performance, *Journal of Finance*, **52**, 57-82.

Casella, G. and Berger, R.L. (1990) *Statistical Inference*, Duxbury Press (Wadsworth Inc.), Belmont California.

Cohen, R.B., Polk, C., and Vuolteenaho, T. (2003) The Value Spread, *Journal of Finance*, **58**, 609-641.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998) Investor Psychology and Security Market Under- and Over-reactions, *Journal of Finance*, **53**, 1839-1885.

Daniel, K. and Titman, S. (1999) Market Efficiency in an Irrational World, *Financial Analysts Journal*, **55**, 28-40.

Davis, J.L., Fama, E.F., and French, K.R. (2000) Characteristics, Covariances, and Average Returns: 1929 to 1997, *Journal of Finance*, **55**, 389-406.

Diether, K.B., Malloy, C.J., and Scherbina, A. (2002) Difference of Opinion and the Cross-section of Stock Returns, *Journal of Finance*, **57**, 2113-2141.

Durnev, A., Morck, R., Yeung, B., and Zarowin, P. (2003) Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing? *Journal of Accounting Research*, **41**, 797-836.

Eisdorfer, A. (2008) Delisted Firms and Momentum Profits, *Journal of Financial Markets*, **11**, 160-179.

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

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

Fama, E.F. and French, K.R. (1996) Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance*, **51**, 55-84.

Fama, E.F. and French, K.R. (2008) Dissecting Anomalies, *Journal of Finance*, **63**, 1653-1678.

Fama, E.F. and MacBeth, J. (1973) Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy*, **71**, 921-956.

Gould, J.F. and Kleidon, A.W. (1994) Market Maker Activity on Nasdaq: Implications for Trading Volume, *Stanford Journal of Law, Business and Finance*, **1**, 1-17.

Grinblatt, M. and Han, B. (2005) Prospect Theory, Mental Accounting, and Momentum, *Journal of Financial Economics*, **78**, 311-339.

Grundy, B.D. and Martin, J.S. (2001) Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing, *Review of Financial Studies*, **14**, 29-78.

Hong, H., Lim, T. and Stein, J.C. (2000) Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies, *Journal of Finance*, **55**, 265-295.

Hong, H. and Stein, J.C. (1999) A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance*, **54**, 2143-2184.

Hou, K., Peng, L., and Xiong, W. (2006), R^2 and Price Inefficiency, unpublished working paper, Ohio State University, Baruch College, Princeton University.

Jegadeesh, N. (1990) Evidence of Predictable Behavior of Security Returns, *Journal of Finance*, **45**, 881-898.

Jegadeesh, N. and Titman, S. (1993) Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance*, **48**, 65-93.

Jegadeesh, N. and Titman, S. (2001) Profitability of Momentum Strategies: an Evaluation of Alternative Explanations, *Journal of Finance*, **56**, 699-720.

Johnson, N.L. and Kotz, S. (1970) *Distributions in Statistics: Continuous Univariate Distributions*, Vol. 1, John Wiley & Sons, New York.

Karpoff, J.M. (1987) The Relation Between Price Changes and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis*, **22**, 109-126.

Korajczyk, R.A. and Sadka, R. (2004) Are Momentum Profits Robust to Trading Costs? *Journal of Finance*, **59**, 1039-1082.

Lee, C.M.C. and Swaminathan, B. (2000) Price Momentum and Trading Volume, *Journal of Finance*, **55**, 2017-2069.

Lehmann, B.N. (1990) Fads, Martingales, and Market Efficiency, *Quarterly Journal of Economics*, **105**, 1-28.

Lou, D. (2011), A Flow-Based Explanation for Return Predictability, unpublished working paper, London School of Economics.

Merton, R.C. (1974) On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, **29**, 449-70.

Morck, R., Yeung, B., and Yu, W. (2000) The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements? *Journal of Financial Economics*, **58**, 215-260.

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

Sagi, J.S. and Seasholes, M.S. (2007) Firm-specific Attributes and the Cross-section of Momentum, *Journal of Financial Economics*, **84**, 389-434.

Schwert, G.W., (2003). Anomalies and Market Efficiency, in G.M. Constantinides, M. Harris and R.M. Stulz (eds.), *Handbook of the Economics of Finance*, Vol. 1b. North-Holland, Amsterdam, pp. 939-974.

Shen, J. (2008), Idiosyncratic Volatility: Information or Noise? unpublished working paper, University of New South Wales.

Shleifer, A. and Vishny, R. (1997) The Limits of Arbitrage, *Journal of Finance*, **52**, 35-55.

Shumway, T. and Warther, V.A. (1999) The Delisting Bias in CRSP's Nasdaq Data and its Implications for the Size Effect, *Journal of Finance*, **54**, 2361–2379.

Stoll, H.R. (2000) Friction, *Journal of Finance*, **55**, 1479-1514.

Teoh, S.H., Yang, Y.G., and Zhang, Y. (2009), R-Square and Market Efficiency, unpublished working paper, UC Irvine, Chinese University of Hong Kong.

Vayanos, D. and Woolley, P. (2010), An Institutional Theory of Momentum and Reversal, unpublished working paper, London School of Economics.

Verardo, M. (2009) Heterogeneous Beliefs and Momentum Profits, *Journal of Financial and Quantitative Analysis*, **44**, 795-822.

Zhang, X.F., (2006) Information Uncertainty and Stock Returns, Journal of Finance, 61, 105-136.
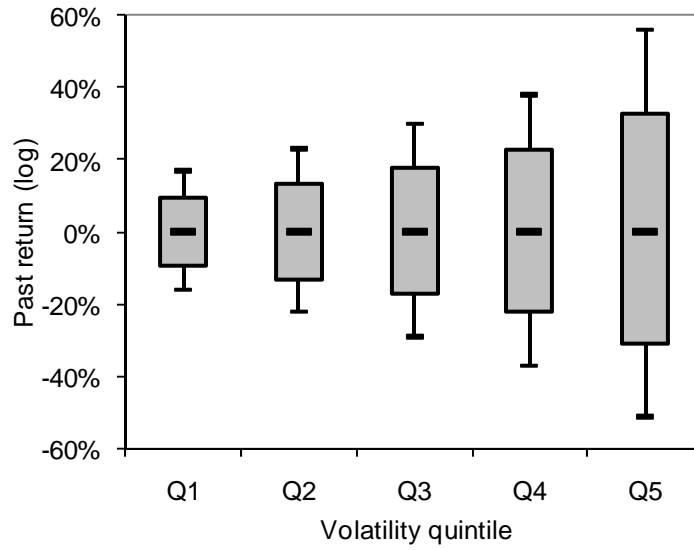
Figure 1. Distribution of past returns for portfolios sorted on volatility. This figure plots winner and loser cutoffs and average past returns for a conditional double sort on *IVOL* and past returns. Past returns are cumulative (*t-6,t-1*) holding period returns and are de-medianed within volatility quintile. For each *IVOL* quintile (Q1=low volatility, Q5=high volatility) we report average log past returns for losers (P1) and winners (P5), indicated in the figure as the top and bottom end points of each distribution. The top and bottom end point of the shaded areas are the winner and loser cutoffs.

Table I. Returns and characteristics of portfolios sorted on past returns

This table reports statistics for returns (Panel A) and characteristics (Panel B) of portfolios sorted on past returns *(t-6,t-1)*: Equally-weighted return for month *(t+1)*, risk adjusted return (using the conditional Fama and French (1993) 3-factor model as implemented by Korajczyk and Sadka (2004)), past return, absolute value of log past return in excess of the cross-sectional median, exponentiated (Momentum strength), annualized idiosyncratic volatility estimated using a 52-week rolling regression of firm return on the market and industry return (*IVOL*), equity market capitalization in billions of dollars (Size), R^2 of regressions of returns on industry and market returns, average monthly turnover over the previous 12 months, number of years since IPO or number of years since the company first appears in CRSP (Age), analyst coverage from I/B/E/S, standard deviation of earnings per share (EPS) forecasts divided by absolute value of mean EPS (Forecast dispersion), market-to-book ratio, price per share, average absolute value of the weekly log return divided by weekly trading volume over the previous 52 weeks, multiplied by 10^6 (Illiquidity), and S&P long-term issuer credit rating (e.g. AAA=1, BBB=9, 22=D/SD rating). We report time series averages of monthly means (return, risk-adjusted return, past return) and medians (momentum strength, IVOL, and characteristics). Past returns, momentum strength, and characteristics are winsorized at the 0.5% and 99.5% levels. The sample runs from 1964 to 2008, except analyst coverage and forecast dispersion (starts January 1976), and credit rating (starts December 1985). The sample includes only stocks with share codes equal to 10 or 11. We exclude stocks that at the beginning of the formation period have price below $5 or that are in the lowest market-capitalization decile.

	P1 (L)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (W)
Panel A: Return statistics										
Return	0.17	0.74	0.93	0.99	1.11	1.13	1.15	1.26	1.35	1.78
Risk-adjusted return	-0.43	0.13	0.31	0.35	0.47	0.52	0.55	0.64	0.79	1.15
Past return	-30.5	-15.2	-7.6	-1.8	3.5	8.9	15.0	22.9	35.4	77.1
Momentum strength	51.5	25.8	15.1	8.1	2.6	2.6	8.3	15.5	26.7	54.8
IVOL	44.8	34.2	30.0	27.8	26.9	26.9	27.7	29.8	33.9	45.1
Panel B: Characteristics										
Size	0.12	0.16	0.19	0.21	0.22	0.23	0.24	0.24	0.23	0.17
R^2	17.4	18.4	18.7	18.8	18.9	18.9	18.7	18.4	17.6	14.8
Turnover	0.62	0.43	0.37	0.34	0.33	0.33	0.35	0.38	0.44	0.59
Age	7.0	9.8	11.1	11.8	11.8	11.5	11.1	10.5	9.3	7.2
Analyst coverage	4.6	4.9	5.2	5.2	5.3	5.4	5.3	5.1	4.7	3.7
Forecast dispersion	0.11	0.07	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.06
Market-to-book	1.51	1.47	1.46	1.47	1.51	1.57	1.68	1.83	2.14	3.12
Price	12.1	15.9	18.3	20.0	21.4	22.2	22.8	22.9	22.4	20.3
Illiquidity	0.10	0.09	0.08	0.07	0.07	0.07	0.07	0.08	0.10	0.17
Credit rating	12.3	10.1	8.9	8.3	8.0	8.0	8.3	8.8	10.0	12.3

Table II . Returns and characteristics of portfolios sorted on volatility

This table reports time-series averages of median momentum strength and *IVOL* (Panel A), as well as characteristics (Panel B) of portfolios sorted on volatility. Momentum strength, *IVOL*, and characteristics are winsorized at the 0.5% and 99.5% levels.

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
		Panel A: Return statistics				
IVOL	17.2	25.0	32.3	41.5	57.4	40.2
Mom strength	9.5	11.5	14.9	20.2	31.9	22.4
		Panel B: Characteristics				
Size	0.73	0.31	0.20	0.14	0.09	-0.63
R^2	25.9	20.7	18.5	16.7	12.9	-13.0
Turnover	0.28	0.32	0.40	0.50	0.67	0.39
Age	17.6	13.7	10.0	7.4	5.5	-12.1
Analyst coverage	9.6	5.9	4.7	3.8	3.1	-6.5
Forecast dispersion	0.03	0.04	0.05	0.07	0.11	0.08
Market-to-book	1.43	1.54	1.61	1.83	2.45	1.0
Price	30.1	25.2	19.6	15.0	10.9	-19.2
Illiquidity	0.02	0.04	0.08	0.15	0.25	0.24
Credit rating	6.6	8.8	11.3	12.9	14.0	7.3

Table III . Correlations of characteristics with momentum strength and volatility

This table reports average monthly rank correlations of *IVOL* and momentum strength with characteristics. We also report the share of months for which the rank correlation is significant at the 1% level.

	IVOL	Share significant	Mom strength	Share significant
Size	-0.43	99%	-0.13	86%
R^2	-0.25	95%	-0.06	61%
Turnover	0.42	100%	0.24	100%
Age	-0.36	100%	-0.15	97%
Analyst coverage	-0.30	97%	-0.07	66%
Forecast dispersion	0.37	100%	0.20	96%
Market-to-book	0.19	82%	0.12	80%
Price	-0.52	100%	-0.17	93%
Illiquidity	0.38	98%	0.09	74%
Credit rating	0.71	100%	0.29	100%

Table IV. Momentum profits of portfolios sorted on volatility and characteristics

This table reports average returns for portfolios sorted first by *IVOL* or characterstic (Q1 to Q5), then by past returns (P1 to P5). For size, R^2, age, analyst coverage, and price we sort stocks in descending order. We report average monthly returns for past losers (P1), past winners (P5), and the momentum strategy portfolios (P5-P1). We also report the difference between high and low characteristic momentum returns (Return diff) and risk-adjusted return differences, as well as t-statistics for return differences (in parentheses). In the final two columns we report the average differences between high (Q5) and low (Q1) characteristics of momentum strength (Mom str diff) and idiosyncratic volatility (IVOL diff) for the stocks included in the momentum strategy (P1 and P5).

		Q1	Q2	Q3	Q4	Q5	Return diff	Risk adjusted	Mom str diff	IVOL diff
IVOL	P5-P1	0.04	0.52	0.91	1.44	1.96	1.93	1.92	59.2	46.2
		(0.37)	(4.28)	(6.03)	(7.57)	(8.16)	(9.36)	(8.51)		
	P5	1.14	1.46	1.79	1.87	1.63				
	P1	1.10	0.94	0.88	0.43	-0.33				
Size	P5-P1	0.53	0.92	1.17	1.42	1.34	0.81	0.95	16.9	20.7
		(2.37)	(3.98)	(5.29)	(6.97)	(7.66)	(4.37)	(4.49)		
	P5	1.18	1.44	1.53	1.71	1.83				
	P1	0.65	0.52	0.36	0.30	0.49				
R^2	P5-P1	0.90	1.13	1.25	1.12	1.06	0.16	0.28	9.4	15.2
		(3.38)	(4.74)	(6.11)	(6.08)	(6.50)	(0.76)	(1.20)		
	P5	1.40	1.58	1.74	1.58	1.46				
	P1	0.50	0.455	0.50	0.468	0.40				
Turnover	P5-P1	0.60	0.90	0.96	1.29	1.47	0.88	0.80	38.1	20.5
		(5.33)	(6.19)	(5.48)	(5.72)	(5.63)	(3.76)	(3.42)		
	P5	1.45	1.62	1.67	1.65	1.37				
	P1	0.85	0.72	0.70	0.35	-0.11				
Age	P5-P1	0.26	0.63	1.03	1.37	1.77	1.52	1.45	24.7	19.1
		(1.62)	(3.41)	(4.89)	(6.25)	(7.54)	(9.04)	(6.53)		
	P5	1.20	1.48	1.60	1.76	1.68				
	P1	0.95	0.86	0.57	0.39	-0.09				
Analyst coverage	P5-P1	0.63	0.95	1.10	1.47	1.43	0.81	0.85	10.6	13.8
		(2.04)	(3.36)	(4.34)	(6.33)	(6.23)	(3.61)	(2.94)		
	P5	1.28	1.49	1.61	1.90	1.79				
	P1	0.66	0.54	0.51	0.43	0.36				
Forecast dispersion	P5-P1	0.09	0.62	0.91	1.45	1.37	1.29	1.27	29.7	17.6
		(0.45)	(2.76)	(3.46)	(4.96)	(4.46)	(5.74)	(4.20)		
	P5	1.14	1.37	1.61	1.79	1.38				
	P1	1.05	0.76	0.70	0.33	0.01				
Market-to-book	P5-P1	0.89	0.56	0.97	1.43	1.64	0.75	0.59	19.7	11.5
		(5.18)	(3.27)	(5.03)	(6.64)	(6.93)	(3.45)	(2.52)		
	P5	1.59	1.47	1.50	1.66	1.70				
	P1	0.70	0.91	0.53	0.23	0.06				
Price	P5-P1	0.79	0.96	1.11	1.37	1.28	0.49	0.58	23.5	24.8
		(3.70)	(5.00)	(5.99)	(6.80)	(5.76)	(2.56)	(2.59)		
	P5	1.51	1.54	1.62	1.65	1.55				
	P1	0.72	0.58	0.52	0.28	0.27				
Illiquidity	P5-P1	0.54	0.92	1.30	1.31	1.18	0.65	0.73	13.6	17.8
		(2.36)	(4.00)	(5.89)	(6.17)	(6.07)	(3.03)	(3.25)		
	P5	1.14	1.32	1.55	1.69	1.81				
	P1	0.60	0.40	0.26	0.38	0.62				
Credit rating	P5-P1	-0.04	-0.06	0.29	1.18	2.39	2.43	2.46	42.9	30.7
		(0.18)	(0.21)	(1.00)	(3.47)	(5.39)	(6.64)	(5.62)		
	P5	0.83	0.99	1.04	1.50	1.53				
	P1	0.88	1.04	0.75	0.32	-0.86				

Table V. Momentum profits of portfolios sorted on residual characteristics (adjusted for volatility)

This table reports momentum returns by quintile of residual characteristics: We sort first into quintiles (Q1 to Q5) of residual characteristics; the residuals are calculated using the regressions of log(characteristic) on IVOL rank for size, turnover, age, analyst coverage, price, credit rating, log(characteristic) on IVOL for stock return R^2 and market-to-book, and characteristic on IVOL for illiquidity and forecast dispersion. For each residual characteristic quintile we then calculate momentum profits using a momentum strategy long the quintile of winners, short the quintile of losers (P5-P1). We report average monthly momentum returns, the difference in momentum returns between high and low residual characteristics (Return difference), and the difference in risk-adjusted returns (Risk adj return diff), as well as t-statistics (in parentheses).

| | Momentum returns (P5-P1) by residual characteristic quintile | | | | | | |
	Q1	Q2	Q3	Q4	Q5	Return difference	Risk adj return diff
Size	0.84	1.21	1.36	1.51	0.84	0.00	0.20
	(3.36)	(4.96)	(6.12)	(7.86)	(5.75)	(0.01)	(0.98)
R^2	1.27	1.11	1.14	1.03	1.01	-0.26	-0.12
	(4.67)	(4.66)	(5.68)	(5.83)	(6.77)	(1.24)	(0.52)
Turnover	1.02	1.21	1.04	1.12	1.12	0.10	0.02
	(6.72)	(6.15)	(4.93)	(4.75)	(4.23)	(0.49)	(0.07)
Age	0.73	0.73	1.12	1.35	1.52	0.79	0.75
	(3.60)	(3.55)	(5.17)	(6.31)	(6.98)	(5.06)	(3.44)
Analyst coverage	0.94	0.96	1.18	1.39	1.12	0.18	0.24
	(2.94)	(3.36)	(4.72)	(5.98)	(5.31)	(0.74)	(0.80)
Forecast dispersion	1.63	1.14	0.50	0.44	1.08	-0.54	-0.42
	(5.78)	(4.45)	(2.50)	(2.09)	(3.79)	(2.51)	(1.24)
Market-to-book	1.14	0.76	0.94	1.16	1.44	0.30	0.27
	(5.75)	(3.83)	(4.71)	(5.27)	(6.13)	(1.47)	(1.14)
Price	1.31	1.37	1.28	1.30	0.73	-0.58	-0.38
	(5.79)	(6.96)	(6.55)	(6.42)	(3.26)	(2.93)	(1.68)
Illiquidity	1.29	0.93	0.86	0.95	1.26	-0.04	-0.13
	(6.18)	(4.91)	(4.45)	(4.32)	(6.30)	(0.18)	(0.59)
Credit rating	0.31	0.99	1.24	1.30	0.58	0.27	0.22
	(0.99)	(3.12)	(3.27)	(3.24)	(2.25)	(1.06)	(0.59)

Table VI. Momentum profits of portfolios adjusted for past returns

This table reports momentum returns, adjusted returns, and matched past return percentiles for stocks with extreme characteristics. Column 1 reports long-short momentum returns (P5-P1) for the top quintile of characteristic (returns in column 1 correspond to Q5 in Table 4). Column 2 reports risk-adjusted returns. Column 3 reports returns adjusted for predicted momentum returns: We adjust predicted returns based on a regression of average momentum returns on log past returns. Column 4 reports the percentile of a momentum single sort for which the past return spread (the difference between winner and loser past returns) matches the double sort past return spread. For example, for size the difference in past returns of the single-sort top 17% winners and bottom 17% losers matches the double-sort (small firm) past return spread. Column 5 reports the difference between double sort momentum returns and matched single-sort momentum returns. We report t-statistics in parentheses.

	Extreme Characteristic momentum return	Risk adjusted return	Adjusted for predicted momentum return	Matched percentile	Adjusted for matched past return spread
Size	1.34	1.43	0.13	17th	0.10
	(7.67)	(10.61)	(0.75)		(0.79)
R^2	1.06	1.12	-0.12	17th	-0.19
	(6.50)	(9.38)	(0.72)		(1.49)
Turnover	1.47	1.46	-0.01	9th	-0.16
	(5.63)	(6.84)	(0.02)		(1.67)
Age	1.77	1.75	0.45	14th	0.38
	(7.54)	(9.82)	(1.92)		(4.36)
Analyst coverage	1.43	1.47	0.24	17th	0.15
	(6.23)	(8.97)	(1.04)		(1.13)
Forecast dispersion	1.37	1.42	0.11	13th	0.02
	(4.46)	(5.44)	(0.36)		(0.18)
Market-to-book	1.64	1.51	0.38	10th	0.07
	(6.93)	(8.35)	(1.62)		(0.58)
Price	1.28	1.28	-0.08	15th	-0.05
	(5.76)	(7.10)	(0.36)		(0.48)
Illiquidity	1.18	1.23	0.03	15th	-0.11
	(6.07)	(8.96)	(0.03)		(0.74)
Credit rating	2.39	2.46	1.03	7th	0.46
	(5.39)	(6.50)	(2.34)		(2.53)

Table VII . Regressions of momentum profits on characteristics, past returns, and volatility

This table reports results from Fama-MacBeth regressions of momentum profits on log(characteristics), momentum strength, and *IVOL* rank. Momentum profit is the stock's forward return in excess of its group median return, multiplied by a winner/loser dummy equal to 1 if the stock was a winner and -1 if the stock was a loser. Thus winners with positive forward returns and losers with negative forward returns both have positive momentum profits. Medians are calculated within characteristic quintile. To control for outliers all explanatory variables are winsorized at the 0.5% and 99.5% levels. We run monthly regressions for our sample (same as in Table IV) and report average coefficients and *t*-statistics (reported in parentheses) based on standard errors of average coefficients for two sets of regression specifications: characteristic only and controlling for momentum strength and *IVOL* .

Variables included	Char only	Controlling for IVOL and Momentum Strength		
	Char	Char	Mom Strength	IVOL
Size	-0.070	-0.011	0.011	0.016
	(6.18)	(0.88)	(7.23)	(5.66)
R^2	-0.018	0.022	0.013	0.015
	(1.01)	(1.27)	(7.82)	(5.94)
Turnover	0.082	-0.038	0.013	0.012
	(3.29)	(2.04)	(8.83)	(5.19)
Age	-0.154	-0.079	0.013	0.008
	(8.67)	(5.52)	(7.94)	(3.20)
Analyst coverage	-0.105	-0.041	0.012	0.015
	(4.34)	(1.75)	(6.57)	(4.62)
Forecast dispersion	0.104	0.034	0.008	0.015
	(5.20)	(1.88)	(4.40)	(4.27)
Market-to-book	0.140	0.041	0.011	0.011
	(5.05)	(1.63)	(7.14)	(4.20)
Price	-0.049	0.155	0.007	0.031
	(1.64)	(4.66)	(5.04)	(9.72)
Illiquidity	0.044	0.006	0.013	0.014
	(4.48)	(0.60)	(7.74)	(5.09)
Credit rating	0.437	0.144	0.013	0.012
	(6.94)	(2.61)	(4.99)	(2.44)