JOB MARKET PAPER

The Momentum Gap and Return Predictability*

Simon Huang†
simon.huang@yale.edu
http://students.som.yale.edu/phd/wsh3

January 23, 2014

Abstract

Momentum strategies have historically delivered high Sharpe ratios and large positive alphas. However, returns to these strategies also display significant time-variation that is not very well understood. I show that expected momentum returns vary negatively and monotonically with the formation period return difference between past winners and losers, which I term the *momentum gap*. A one standard deviation increase in the momentum gap predicts a 1.29 percent decrease in the monthly momentum return after controlling for existing predictors. The momentum gap remains a significant predictor in out-of-sample tests. Conditional momentum strategies using the momentum gap yield substantially higher Sharpe ratios and lower skewness than the unconditional strategy. These findings are less consistent with the notion that past return proxies for the loading on a priced risk factor. I find evidence to support the alternative hypothesis that momentum is a mispricing phenomenon and that the momentum gap measures momentum arbitrage activity.

*I would especially like to thank my advisors: Will Goetzmann (chair), Nick Barberis, Andrew Metrick, and Jake Thomas. I would also like to thank Sriya Anbil, Brad Barber, Hui Chen, James Choi, Sean Hundtofte, Peter Kelly, Charles Lee, Stefan Lewellen, Alan Moreira, Tyler Muir, Marina Niessner, Terry Odean, Scott Richardson, Geert Rouwenhorst, Clemens Sialm, Matt Spiegel, Jeff Wurgler, Lei Xie, Hongjun Yan, Frank Zhang, as well as seminar participants at Yale University. I gratefully acknowledge financial support of the Whitebox Advisors Doctoral Fellowship. All remaining errors are my own. Click here for the most recent version of this paper.

†Ph.D. Candidate, Financial Economics. Yale School of Management, 135 Prospect Street, New Haven, CT 06520-8200. (415) 623-0776.

1 Introduction

The profitability of momentum strategies, as documented by Jegadeesh and Titman (1993), is an important asset pricing anomaly because it implies high variability of marginal utility across states of nature and does not appear to compensate for systematic risk (Fama and French (1996)). At the same time, returns to these strategies exhibit significant time-variation. While there is a large literature on the cross-sectional determinants of momentum profits, much less is understood about this time-variation.[1] In this paper, I identify a new and significant determinant of variation in expected momentum returns. I then use my results to shed light on the source of the anomaly.

Risk-based theories of momentum typically argue that stocks with higher past returns are more exposed to a risk factor.[2] Assuming that exposure to this factor carries a positive price of risk, past winners should have higher expected returns than past losers. However, momentum is a cross-sectional phenomenon. If a stock's past return proxies for its loading on the risk factor and the price of risk is constant, expected momentum returns should vary positively with the difference in returns between winners and losers in the prior period. In this paper, I test this hypothesis.[3]

I find that expected momentum returns vary *negatively* with the formation period return difference between past winners and losers, which I term the momentum gap. The average adjusted return of my baseline momentum strategy varies monotonically from 2.23 percent per month, when the lagged momentum gap is small, to -0.13 percent per month, when the lagged momentum gap is large. The predictive power of the momentum gap remains economically and statistically significant after controlling for market return, market volatility, and market illiquidity.[4] Conditional on these controls, a one standard deviation increase

[1] I give a fuller discussion of the related literature in Section 2.

[2] See, e.g., Berk, Green, and Naik (1999), Johnson (2002), and Sagi and Seasholes (2007). Empirically, Liu and Zhang (2006) contend that the growth rate of industrial production could be one such risk factor.

[3] In the related literature on the value anomaly, Cohen, Polk, and Vuolteenaho (2003) have documented an effect that is consistent with this idea: expected value returns vary positively with the book-to-market difference between value and growth stocks. They name their measure the value spread.

[4] See Cooper, Gutierrez, and Hameed (2004), Wang and Xu (2011), and Avramov, Cheng, and Hameed

in the momentum gap is associated with a 1.29 percent decrease in the monthly adjusted return of the momentum strategy. My regression-based tests employ a bootstrap methodology to ensure that the inferences are robust to the predictive regression biases highlighted in Stambaugh (1999) and Ferson, Sarkissian, and Simin (2003). The momentum gap has significant predictive power for both the long and short legs of the momentum trade, thus ruling out that the results are driven by the infrequent yet large-scale reversals of past losers documented by Daniel and Moskowitz (2013).

A recent emphasis in the return predictability literature is on the out-of-sample performance of predictors (Campbell and Thompson (2007); Welch and Goyal (2007)). Even though momentum strategies are popular, to the best of my knowledge there is as yet no comprehensive study on the out-of-sample performance of momentum predictors. I examine the out-of-sample predictive power of the momentum gap and other predictors in the literature using the same method, time-period, and estimation frequency. Formal tests show that the momentum gap has significant out-of-sample predictive power. In fact, it delivers the highest out-of-sample $\overline{R}^2$ among the predictors I examine. Conditional momentum strategies using the momentum gap in real time yield substantially higher Sharpe ratios and lower skewness than the unconditional strategy.

The fact that expected momentum returns vary negatively with the momentum gap does not necessarily imply that the existing risk-based theories are falsified, nor that risk can be ruled out as an explanation for momentum. However, it does suggest that if risk were to explain momentum, the price of the risk factor that drives momentum profits must vary through time. Chordia and Shivakumar (2002) document that momentum returns are strong in economic expansions but non-existent in recessions. This procyclical nature of expected momentum returns, though, remains an obstacle for risk-based explanations, which usually rely on a *countercyclical* price of risk.[5]

I set out three alternative hypotheses to explain the negative relation between expected

(2013).

[5]See Liu and Zhang (2006) and Liu and Zhang (2013).

momentum returns and the momentum gap. Hypothesis 1 is that the empirical relation is spurious. While robust in the panel of U.S. stocks, it is simply be the outcome of an elaborate data snooping process. Hypothesis 2 is that the momentum gap's predictive power is driven by its relation to the business cycle. Once macroeconomic variables known to predict returns are accounted for, the momentum gap will lose its predictive power. Hypothesis 3 is that momentum is a mispricing phenomenon and the momentum gap reflects the degree to which arbitrageurs are trading the strategy. A large momentum gap indicates the presence of many arbitrageurs and near-complete convergence of prices to fundamental values, thereby explaining why the trade is less profitable going forward.[6]

I find no evidence to support Hypothesis 1 or Hypothesis 2. Using a sample of international stocks from the 21 largest markets excluding the U.S., I find a negative country-specific relation between expected momentum returns and the momentum gap in 20 of the countries. In 13 of these countries, the estimated coefficient is significant. In no country is the relation significantly positive. Moreover, in the U.S., where data is more readily available, I find that the momentum gap holds its predictive power in the presence of the dividend yield, default spread, term spread, short-term interest rate, and industrial production growth. In contrast, none of the macroeconomic variables can predict momentum returns after controlling for the momentum gap.

For Hypothesis 3, I find some evidence that the momentum gap measures strategy-level arbitrage activity for momentum. Using data on institutional investors and short sellers, I show that the momentum gap is significantly related to these groups' exposures to momentum. The momentum gap is higher when institutional investors hold more past winners than losers. Similarly, the momentum gap is also higher when aggregate short interest is higher among past losers than among past winners. In addition, using the change

[6]Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999) provide behavioral explanations for momentum. Lou and Polk (2013) also propose a measure of momentum arbitrage activity. Their measure primarily works at a longer horizon (12 to 24 months) and is based on the high-frequency abnormal return correlation among stocks on which a typical momentum strategy would speculate.

in the momentum gap as a proxy for strategy-level capital flow, I find a positive performance-flow relationship. This is consistent with the idea that real-world arbitrageurs managing other people's money experience capital inflows after good performance.[7]

Basic economic intuition suggests that, as more money is brought to bear against a given trading strategy, any abnormal returns it delivers must be reduced and eventually eliminated. Stein (2009)'s crowded-trade model agrees with this intuition, but argues that the elimination of abnormal returns does not necessarily imply a decrease in volatility unrelated to fundamentals. It is built on two key assumptions: 1) no individual arbitrageur knows exactly how much arbitrage capital is available in the aggregate at any point in time; and 2) for the trading strategy in question, arbitrageurs do not have an independent estimate of fundamental value. The model predicts that arbitrage activity can be destabilizing when trading becomes too crowded. Momentum is a classic example of this unanchored strategy. Consistent with the model's prediction, I find that subsequent momentum returns are more negatively skewed and leptokurtic when the momentum gap is large.

I also explore other implications of the momentum gap's predictive power. Moskowitz and Grinblatt (1999) argue that industry momentum drives individual stock momentum. While the literature has documented a number of predictors for individual stock momentum, there is little evidence that industry momentum is predictable. A natural inquiry into the link between these phenomena, therefore, is whether the momentum gap can also predict industry momentum. This bears out in the data. Furthermore, the momentum gap is the only significant variable in a multivariate predictive regression for industry momentum that also includes market return, market volatility, and market liquidity as explanatory variables.

The momentum literature has documented significant cross-sectional heterogeneity in momentum profits. For instance, Hong, Lim, and Stein (2000) find that the profitability of momentum strategies declines with firm size, consistent with predictions of behavioral theories. Lee and Swaminathan (2000) document that momentum profits increase in trading vol-

[7]Such a positive performance-flow relationship is featured in many limits-to-arbitrage models. See, e.g., Shleifer and Vishny (1997) and Gromb and Vayanos (2010).

ume, which they argue can help reconcile intermediate-horizon momentum and long-horizon reversal. I investigate the predictability of momentum for subgroups of stocks sorted by these characteristics using the momentum gap and find that its predictive power is pervasive. The momentum gap is a significant predictor in all subgroups sorted by size and trading volume.

A battery of additional tests confirms the robustness of my results. For example, the momentum gap essentially accounts for the evidence in Stivers and Sun (2010) that short-horizon factor return dispersion (*FRD*) negatively predicts momentum. In multivariate regressions, the predictive ability of *FRD* is subsumed by the momentum gap. Using one-way sorts, I find that momentum returns are not significantly higher when *FRD* is low than when it is high. In two-way sorts, high *FRD* actually precedes *high* momentum returns when the momentum gap is small.

The remainder of the paper is organized as follows. Section 2 briefly reviews related literature. Section 3 describes the data and methodology. Section 4 establishes the main results, while Secton 5 explores interpretations of the results. Section 6 provides further analysis, including robustness tests. Section 7 concludes.

2 Related Literature

Momentum strategies buy past winners and sell past losers. Motivated by the observation that mutual funds tend to buy stocks that have recently increased in price, Jegadeesh and Titman (1993) systematically investigate the profitability of such strategies for U.S. common stocks. For each portfolio formation month t, they form portfolios based on cumulative stock returns from month $t-J$ to $t-1$ and hold them for K months. They examine strategies for J ranging from 3 to 12 months and K also ranging from 3 to 12 months. Using data from 1965 to 1989, they find that medium-term past winners significantly outperform past losers for all horizons considered and by as much as 1.49 percent a month.

Unlike a number of other anomalies examined, Fama and French (1996) find that their three-factor model cannot explain momentum profits. After controlling for the Fama and French (1993) factors, to which it is negatively correlated, momentum yields monthly alphas of 1.74 percent. This deepens the momentum puzzle. The inability of pre-existing asset pricing models to explain momentum has led researchers to add it as an additional risk factor (Carhart (1997)).

Momentum does not seem to be a spurious result of data snooping as its effects are also present when other basis assets are used. Rouwenhorst (1998) documents momentum in developed markets stocks while Rouwenhorst (1999) finds evidence of momentum in emerging markets. Using a U.S. sample, Moskowitz and Grinblatt (1999) demonstrate that industry momentum strategies are just as profitable as individual stock momentum strategies. Together, Moskowitz, Ooi, and Pedersen (2012) and Asness, Moskowitz, and Pedersen (2013) show that momentum is also present in exchange traded futures contracts, cross-country equity index futures, cross-country bonds, currencies, and commodity futures.

There is an ongoing theoretical debate as to what causes momentum. Using a risk-based framework, Berk, Green, and Naik (1999), Johnson (2002), and Sagi and Seasholes (2007) contend that past winners are riskier so that momentum is due to time-varying expected returns. On the other hand, a number of behavioral models based on well-documented psychological evidence also yield momentum as an implication. Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999), and Grinblatt and Han (2005) propose investor underreaction to news as an explanation of momentum. De Long et al. (1990), Daniel, Hirshleifer, and Subrahmanyam (1998), and Barberis and Shleifer (2003) argue that momentum could be the result of overreaction to information by market participants.

This paper is closely related to several other papers in the literature that investigate the time series predictability of momentum returns. Cooper, Gutierrez, and Hameed (2004) argue that the extent to which investors are affected by psychological biases that cause momentum depends on the market state. Using a lookback window of three years, they show that the mean momentum profit following positive market returns is significantly higher than the mean momentum profit following negative market returns. Taking a hint from momentum returns in 2008-2009, Wang and Xu (2011) find that high market volatility forecasts low momentum profits. In contemporaneous work, Avramov, Cheng, and Hameed (2013) argue that market illiquidity negatively covaries with investor overconfidence and find that it is a useful predictor of momentum payoffs. While these papers provide valuable evidence on the predictability of momentum returns, my approach is different and can shed some light on the possible source of momentum profits. I explicitly control for these existing predictors in my statistical tests to ensure that the momentum gap is not subsumed by them.

In addition, a number of recent articles have explored the time-varying nature of the risk of momentum strategies. Daniel and Moskowitz (2013) study the infrequent episodes when momentum strategies experience strong and persistent strings of negative returns. They find that these "momentum crashes" are mostly the result of sharp reversals of past losers and are forecastable using a combination of market return and market volatility. In related research, Daniel, Jagannathan, and Kim (2012) demonstrate that a hidden Markov model in which the market moves between latent states can also forecast tail risk in momentum strategies. This paper shows that the momentum gap is another important variable that can forecast momentum crashes. In fact, the momentum gap is also effective in forecasting returns to the long leg of momentum strategies.

3 Data and Methodology

I obtain data from several sources to construct the main sample that runs from 1926 to 2012. I collect stock returns, dividends, trading volume, and the short-term interest rate from the Center for Research in Security Prices (CRSP). The Fama and French (1993) factors as well as returns on portfolios formed on size and book-to-market are from Kenneth French.[8] Moody's AAA and BAA corporate bond yields, the Industrial Production Index, and the VIX are from Federal Reserve Economic Data (FRED). I obtain data on institutional holdings from

[8] I thank Kenneth French for providing the data on his website.

the Thomson-Reuters Institutional Holdings (13F) Database. I extract quarterly holdings beginning in the first quarter of 1980 and ending in the third quarter of 2012. Short interest data for the period from February 1973 to June 2012 is from COMPUSTAT for NYSE and AMEX stocks,[9] and directly from the exchange for NASDAQ stocks.

My formulation of the momentum strategy is standard (Fama and French (1996); Carhart (1997); Asness, Moskowitz, and Pedersen (2013)). I start with all stocks listed on the NYSE, AMEX, and NASDAQ with share code 10 or 11. This eliminates closed-end funds, real estate investment trusts, American Depository Receipts, foreign companies, primes, and scores. To mitigate the impact of any microstructure biases, I exclude stocks with a price below $1 (penny stocks). At the beginning of each month t, stocks are sorted into ten deciles based on their cumulative returns from month $t-12$ to $t-2$ (the formation period) using NYSE breakpoints. The momentum strategy goes long a value-weighted portfolio of stocks in the top decile and sells short a value-weighted portfolio of stocks in the bottom decile. To correct returns for delisting bias, I use the adjustment procedure proposed in Shumway (1997) and Shumway and Warther (1999) when a stock's delisting return is missing in the CRSP dataset.

I consider two measures of the momentum gap that follow directly from the construction of the momentum strategy. The idea is to parsimoniously capture the formation period return difference between past winners and losers. The first measure (*MomentumGap*) is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. The second measure (*MomentumGap2*) is the difference between the 90th and 10th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. The first measure, *MomentumGap*, is preferred a priori because it is more robust to outliers.

I conduct both in-sample and out-of-sample tests of momentum return predictability at the monthly frequency in this paper. In-sample regression coefficients are estimated using ordinary least squares while their statistical significance is assessed using bootstrapped

[9]FT Interactive is the original source of this data.

p-values. My bootstrap methodology is derived from the return predictability literature (Goetzmann and Jorion (1993); Baker, Taliaferro, and Wurgler (2006); Welch and Goyal (2007)). It imposes the null hypothesis of no predictability and assumes the data generating process to be

$$\begin{aligned} r_{t+1} &= \alpha + \qquad\quad u_{t+1}, \\ x_{t+1} &= \mu + \rho \cdot x_t + v_{t+1}, \end{aligned} \tag{1}$$

where r denotes momentum return, for example, and x is a candidate predictor variable such as *MomentumGap*. μ and ρ are estimated by maximum likelihood using the full sample of observations. Both residuals u and v are saved for sampling. I then randomly choose the initial observation from actual data and proceed to generate 10,000 bootstrapped time series by drawing with replacement from the residuals. This bootstrap procedure preserves the autocorrelation structure of the predictor variable and the cross-correlation structure of the two residuals. It is therefore robust to the predictive regression biases studied in Stambaugh (1999) and Ferson, Sarkissian, and Simin (2003).

By conducting out-of-sample tests, I ask whether the predictive regressions or the historical sample mean have delivered better out-of-sample momentum return forecasts. For both methods, the forecast for month t uses only data available up to and including month $t-1$. My out-of-sample statistics are defined as

$$\begin{aligned} R^2 &= 1 - \frac{\text{MSE}_A}{\text{MSE}_N}, \\ \overline{R}^2 &= 1 - (1 - R^2) \cdot \frac{T - k - 1}{T - k - p - 1}, \\ \Delta\text{RMSE} &= \sqrt{\text{MSE}_N} - \sqrt{\text{MSE}_A}, \\ \text{MSE-F} &= (T - k) \cdot \frac{\text{MSE}_N - \text{MSE}_A}{\text{MSE}_A}, \end{aligned} \tag{2}$$

where MSE_A is the mean square error of the predictive regression (the alternative hypothesis), MSE_N is the mean square error of the historical sample mean (the null hypothesis), T is the length of the sample period, k is the length of the training period (set to 30 months throughout this article), and p is the number of predictors in the regression. Out-of-sample statistical significance is assessed using MSE-F statistic by McCracken (2007), which tests for equal MSE of the unconditional forecast and the conditional forecast.[10] One-sided critical values of MSE-F statistic are obtained from McCracken (2007).

4 Main Results

Figure 1 plots the two momentum gap measures. These time series are standardized for ease of comparison because, by construction, *MomentumGap* is always smaller than *MomentumGap2*. The figure makes it clear that these two measures capture very similar information. They are correlated at 94 percent in the full sample. I focus on *MomentumGap* for the rest of this paper, but the results are quantitatively similar if I use *MomentumGap2* instead. Another important observation from this figure is that peaks in the momentum gap tend to occur around the time of momentum crashes. Daniel and Moskowitz (2013) document these devastating episodes for momentum investors. The benchmark momentum strategy has cumulative returns of -91.6 percent for the two months from August 1932 to September 1932, and -67.2 percent from April 2009 to May 2009. We will ultimately see that, while the momentum gap correctly predicts many momentum crashes, it also predicts momentum returns over the relatively tranquil period between the Great Depression and the Great Recession.

Table 1 provides summary statistics for the momentum gap and the existing momentum predictors. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is

[10] The inferences are similar with MSE-T, the Diebold and Mariano (1995) t-statistic modified by Harvey, Leybourne, and Newbold (1997). I use MSE-F in this paper because Clark and McCracken (2001) find that MSE-F has higher power than MSE-T.

the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. All returns are in logs so as to be consistent with the later predictive regressions, which have log returns as the dependent variables. This, in turn, follows from the return predictability literature.[11] In addition, if underlying returns are lognormal, using simple returns on the left hand side of predictive regressions could confound predictability of the variance with predictability of the mean. The results are quantitatively similar if the predictive regressions have simple returns as the dependent variables instead.

The momentum gap shows substantial time-variation. For example, one can see from Panel B of Table 1 that its mean over the period from 1927 to 1940 is 48.40, which is over 80 percent higher than its mean over the period from 1941 to 1960. The momentum gap does not have a clear trend, however, as Figure 1 shows that it has a peak in 2009 as well. Consistent with the dramatic changes in trading technology that have occurred over the last century, market illiquidity does show a downward trend – both its mean and standard deviation since 2001 are only tiny fractions of what they are at the beginning of the sample. Panel D shows that the momentum gap is significantly correlated with other predictor variables, which highlights the importance of investigating them jointly.

Panel E of Table 1 reports 12-month autocorrelations; it shows that all of the predictor variables are highly persistent. Persistent predictors effectively reduce the number of independent observations in the sample. Stambaugh (1999) shows that, when the regressor is persistent and its innovation is correlated with the regression disturbance, inference using ordinary least squares can be biased. In related research, Ferson, Sarkissian, and Simin (2003) warn that if returns are noisy realizations of an autoregressive expected return process, spurious regression bias of the variety studied by Granger and Newbold (1974) can arise in the presence of highly persistent regressors. Therefore, I employ simulations constructed so as to maintain the dynamics of regressions with persistent dependent variables in the in-sample tests.

[11] See, e.g., Cochrane (2007) and Welch and Goyal (2007).

4.1 In-Sample Predictive Regressions

I focus on the Fama and French (1993) three-factor adjusted returns of the momentum strategy in the empirical analysis that follows because I want to distinguish novel predictability effects from well-known comovement.[12] The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression

$$R_t = \alpha + \beta_{MKT} \cdot MKT_t + \beta_{SMB} \cdot SMB_t + \beta_{HML} \cdot HML_t + \epsilon_t. \tag{3}$$

Using adjusted returns on the left hand side of predictive regressions also makes the bootstrap procedure used to compute p-values easier to interpret.

Table 2 presents adjusted returns of the momentum strategy by the lagged momentum gap. I sort all months in the sample into quintiles according to the lagged momentum gap. Following Stambaugh, Yu, and Yuan (2012), I break down the momentum strategy to a long leg and a short leg. However, I define the long leg to be the strategy that buys past winners and sells the market portfolio, and the short leg to be the strategy that buys the market portfolio and sells past losers. The idea is that both legs should contribute positively to the overall profitability of the momentum strategy.

We see a negative and monotonic relation between adjusted momentum returns and the lagged momentum gap in Table 2. When the momentum gap is small, momentum earns large alphas of 2.23 percent a month. On the other hand, when the momentum gap is large, momentum loses 0.13 percent a month.[13] The difference is 2.36 percent a month and significant at the 1 percent level. It is also interesting to note that while this relation is monotonic, it appears to be non-linear – the difference between quintile 1 and 2 is substantially smaller than the difference between quintile 4 and 5. This motivates a quadratic specification which I investigate in my out-of-sample tests.

[12] See Cooper, Gutierrez, and Hameed (2004) and Stambaugh, Yu, and Yuan (2012). The results are quantitatively similar if I use unadjusted returns as the dependent variables and control for contemporaneous factor returns instead.

[13] We will later see that momentum returns are extremely left-skewed and leptokurtic during these periods.

I run predictive regressions of adjusted momentum returns on the predictor variables to control for competing effects. Table 3 presents the results. Each row contains the results from one regression. In Panel A, the dependent variable is the adjusted return of the momentum strategy. In Panel B and Panel C, the dependent variables are the adjusted return of the long and short legs of the momentum strategy, respectively.

As Panel A of Table 3 shows, momentum returns are highly predictable using the momentum gap. The p-values are consistently small. A one standard deviation increase in the momentum gap is associated with a 1.29 percent decrease (-0.12×10.76) in the monthly adjusted return of the momentum strategy. The sizable economic magnitude of the effect is also evident from the adjusted R^2 of 3.35 percent in Regression (1), which is extremely high for a predictive regression at the monthly frequency.[14] Compared to the results from Regression (2), the momentum gap alone has a higher adjusted R^2 than market return, market volatility, and market illiquidity combined. Regression (3) includes as regressors all of the predictor variables. The coefficient and the p-value of the momentum gap hardly change. In contrast, the effects of market return and market illiquidity are weakened.

Comparing Panel B and Panel C, we see that the momentum gap is a significant predictor for returns to both the long and short legs of the momentum strategy. This rules out that the results are driven by the infrequent yet devastating reversals of past losers that characterize momentum crashes. Its predictive power seems to be stronger for the short leg: the coefficient is almost twice as large and the adjusted R^2 is more than 1 percent higher. Interestingly, market return and market volatility cannot predict the long leg of momentum with or without the presence of the momentum gap. The literature has documented that many investors engage in momentum trading.[15] Most of these investors are long-only and do not have a mandate to sell stocks short. Results in these two panels, therefore, show that the momentum gap has bigger practical significance compared to market return and market

[14]The adjusted R^2 is higher still if I use unadjusted returns as the dependent variables.

[15]See, e.g., Grinblatt, Titman, and Wermers (1995), Nofsinger and Sias (1999), and Wahal and Badrinath (2002).

volatility.

4.2 Out-of-Sample Tests

This paper has so far demonstrated a robust in-sample relation between expected momentum returns and the momentum gap. A recent emphasis in the return predictability literature is on the out-of-sample performance of predictors (Campbell and Thompson (2007); Welch and Goyal (2007)), so a natural concern is whether this relation is stable enough to extend out-of-sample. Evidence of out-of-sample predictive power would suggest that the momentum gap could potentially be used to form conditional momentum strategies that yield enhanced performance. That, in turn, could shed some light on explanations for momentum that attribute its high (unconditional) profits to compensation for bearing downside risk.

Table 4 shows how well each of the predictor variables performs out-of-sample. The dependent variable in this subsection is always unadjusted momentum return because factor exposures used previously are calculated using the full sample and are not known in real time. Panel A of Table 4 provides the results when a linear regression specification is used to form conditional forecasts, i.e., momentum return is assumed to have a linear univariate relation with the lagged value of the variable. We observe in Table 2 a non-linear relation between momentum returns and the lagged momentum gap. Therefore, I examine a quadratic regression specification in Panel B.[16] In each month after the training period (the first 30 months of the sample), I calculate the forecast errors for both the historical sample mean and the regression forecast (the regression coefficients are estimated using only data available up to that point). These forecast errors are then used to compute the out-of-sample statistics according to Equation (2).

The results reveal that the momentum gap has robust out-of-sample forecasting power. With a linear specification, it achieves an out-of-sample $\overline{R}^2$ of 0.58 percent and a RMSE

[16] Cooper, Gutierrez, and Hameed (2004) note that, for market return, the relation with expected momentum returns is non-linear as well so their out-of-sample test for market return as a momentum predictor also uses a quadratic specification.

improvement of 0.03. McCracken (2007)'s MSE-F test shows that this is significant at the 1 percent level. By way of comparison, Welch and Goyal (2007) report that, for return on the S&P 500 index at the monthly frequency, the highest out-of-sample $\overline{R}^2$ achieved is 0.22 percent by the term spread.[17] For the existing predictors, market return and market illiquidity also show significant out-of-sample performance. Market return is significant at the 10 percent level while market illiquidity is significant at the 1 percent level. With a quadratic specification, the momentum gap dominates the existing predictors, regardless of which specification they use. Consistent with Cooper, Gutierrez, and Hameed (2004), market return shows better performance with a quadratic specification and achieves an out-of-sample $\overline{R}^2$ of 0.5 percent. Market volatility, on the other hand, does not deliver positive out-of-sample performance in either case.

I formulate two simple conditional momentum strategies that exploit the momentum gap's predictive power in real time and Figure 2 illustrates the results. WML$^\bigstar$ represents the conditional strategy that takes a position in momentum at the beginning of month t unless the momentum gap is ranked in the top quintile. WML$^\spadesuit$ represents the conditional strategy that takes a position in momentum at the beginning of month t unless a negative return is predicted in a quadratic predictive regression using the momentum gap. Both conditional strategies use an expanding look-back window with an initial length of 30 months.

Performance of the two conditional momentum strategies deepens the momentum puzzle. Both strategies deliver economically significant improvement in the Sharpe ratio and downside risk. For example, WML$^\bigstar$ yields a Sharpe ratio that is more than 50 percent higher (0.79 vs. 0.52) and a skewness that is substantially lower (-2.51 vs. -6.33) than the unconditional momentum strategy. Harvey and Siddique (2000) argue that conditional skewness is an economically important determinant of the cross-sectional variation of expected returns, and that it could help explain momentum profits. The evidence presented here, however, seems to suggest that high profitability does not have to go hand in hand with high down-

[17] See Table 3 of their paper for details.

side risk for the momentum strategy. In related research, Barroso and Santa-Clara (2013) use a different approach and show that the risk-managed momentum strategy has a Sharpe ratio that is almost twice as high as the unconditional momentum strategy. In untabulated analysis, I find that my approach using the momentum gap can be profitably combined with Barroso and Santa-Clara's risk management method to enhance the momentum strategy.

Overall, I do not find a positive relation between expected momentum returns and the momentum gap. To the contrary, I find the relation to be negative, monotonic, and significant. The predictive power of the momentum gap uniformly exceeds that of competing variables and holds up out-of-sample. It would be useful to understand what drives this puzzling relation. I address this issue in depth in the next section.

5 Interpreting the Results

Consider a risk-based model of momentum in which returns are generated by the following factor structure

$$r_{i,t} = E[r_{i,t}] + \sum_{j=1}^{J} \beta_{i,j} \cdot f_{j,t} + \theta_{i,t-1} \cdot f_{G,t} + \epsilon_{i,t}, \tag{4}$$

$$\epsilon_{i,t} \sim N(0, \sigma_{ei}^2),$$

$$f_{j,t} \sim N(0, 1),$$

where $\beta_{i,j}$ is the loading of firm i on factor j and $f_{j,t}$ is the return on factor j at time t. In this equation I separate out $\theta_{i,t-1}$, firm i's loading on the growth factor, and $f_{G,t}$, the return on the growth factor at time t. In this factor pricing model, expected returns are a linear function of all factor loadings:

$$E[r_{i,t}] = r_{f,t} + \sum_{j=1}^{J} \beta_{i,j} \cdot \lambda_j + \theta_{i,t-1} \cdot \lambda_G. \tag{5}$$

Here, firm i's past return proxies for $\theta_{i,t-1}$, the loading on the growth factor. The premium associated with the growth factor, λ_G, is positive, meaning that firms that load on the growth factor, i.e., past winners, earn a positive risk premium.

This simplified model is consistent with the theoretical work of Berk, Green, and Naik (1999), Johnson (2002), and Sagi and Seasholes (2007), as well as Liu and Zhang (2006)'s interpretation of these models. In particular, Liu and Zhang argue that the growth rate of industrial production is a priced risk factor that captures growth-related risk. They find that past winners have temporarily higher loadings on this factor than past losers. Moreover, they show that in many specifications, this macroeconomic risk factor explains more than half of momentum profits.

If the price of the growth factor is constant, this model implies that expected momentum returns should vary positively with the difference in returns between winners and losers in the prior period, i.e., the momentum gap. The estimate for b should be positive in the following predictive regression

$$r_t^{WML} = a + b\,(\theta_{t-1}^{W} - \theta_{t-1}^{L}) + \epsilon_t, \tag{6}$$

where r_t^{WML} is the adjusted return of the momentum strategy, θ_{t-1}^{W} is a measure of the winners' past returns, and θ_{t-1}^{L} is a measure of the losers' past returns. However, the empirical facts presented in Section 4 are at odds with this hypothesis. This suggests that, if risk were to explain momentum, the price of the growth factor must vary through time. Chordia and Shivakumar (2002) document that momentum profits are strong in economic expansions but are non-existent in recessions. This procyclical nature of expected momentum profits, though, remains elusive for many risk-based models that rely on a *countercyclical* price of risk.[18]

There are nevertheless several plausible explanations for the results:

[18]See Liu and Zhang (2006) and Liu and Zhang (2013).

Hypotheis 1. The negative relation between expected momentum returns and the momentum gap, found in the database of U.S. stocks, is spurious. It is simply be the result of an elaborate data snooping process.

Hypotheis 2. The momentum gap, as a measure of the cross-sectional dispersion of stock returns, is driven by the macroeconomy. Its predictive power for momentum returns is subsumed by business cycle variables known to predict returns.

Hypotheis 3. Momentum is a mispricing phenomenon, and as such, attracts arbitrageurs. The momentum gap measures the degree to which these rational speculators are trading the strategy. A large momentum gap indicates near-complete convergence of prices to values, thereby explaining why the trade is less profitable in the future.

In the remainder of this section, I examine the evidence for each of these hypotheses.

5.1 Data Snooping

One explanation for the results is that they are spurious. The data snooping story predicts that in out-of-sample tests, the momentum gap will fail to predict momentum returns.[19] Data-snooping bias can never be ruled out, but following the momentum literature,[20] I gather a large sample of international stocks after the main results in Section 4 have been obtained. I examine in this subsection the predictive power of the momentum gap in an international dataset consisting of returns to momentum strategies in the 21 largest markets excluding the U.S. for the period from July 1989 to August 2013.[21] These countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, and the United Kingdom.

[19] This assumes that even the out-of-sample tests done in Section 4 are in-sample.

[20] See, e.g., Rouwenhorst (1998) and Rouwenhorst (1999).

[21] Chaves (2012) provides detailed description of the data and methodology. I thank Denis Chaves for providing this data.

Table 5 presents the results. For each country and each month, stocks above the median market capitalization are sorted into five quintiles based on their cumulative returns from month $t-12$ to $t-2$. Each country-specific momentum strategy goes long an equally-weighted portfolio of stocks in the top quintile and sells short an equally-weighted portfolio of stocks in the bottom quintile. I adjust these country-specific momentum returns using the regional Fama and French (1993) factors described in Fama and French (2012). I calculate the country-specific momentum gap measures in a way similar to my U.S. measure and run a univariate predictive regression for each country.

The evidence suggests that the U.S. evidence is unlikely to be simply due to chance. 20 of the 21 point estimates are negative and 13 of them are significant at the 5% level. The magnitude of the point estimates are similar to that from the U.S. sample. For example, the coefficient in the U.K. is also -0.12. The point estimate is positive only in New Zealand but it is not statistically significant.

5.2 Macroeconomic Risk

The momentum gap is also a measure of the cross-sectional dispersion of stock returns, albeit over a unique horizon – past one year skipping the most recent month. While it is not a standard macroeconomic variable, we see in Figure 1 that it varies through the business cycle. Chordia and Shivakumar (2002) argue that intertemporal variations in macroeconomic factors (and presumably risk) are the main sources of momentum profits. They show that a standard set of macroeconomic variables can predict momentum returns.[22] The macroeconomic story predicts that, once macro variables known to predict market returns are accounted for, the momentum gap will lose its predictive power.

Table 6 shows predictive regressions using the momentum gap and a set of macroeconomic instruments. Following Chordia and Shivakumar (2002), I include the dividend yield, default

[22] Cooper, Gutierrez, and Hameed (2004) demonstrate that Chordia and Shivakumar's results may be sensitive to the exclusion of penny stocks. Moreover, Griffin, Ji, and Martin (2003) find that momentum profits bear basically no significant relation to standard macroeconomic factors in a sample of 17 international markets.

spread, term spread, and short-term interest rate. The growth rate of industrial production is featured prominently in Liu and Zhang (2006) so I include it as well. For the purpose of comparison, I repeat the results from Table 3 when the momentum gap is the sole predictor.

Standard macroeconomic variables do not seem to be able to drive out the momentum gap's predictive ability. In Panel A, where the dependent variable is adjusted momentum return, the coefficient on the momentum gap drops slightly but remains highly significant. On the other hand, neither the dividend yield, default spread, term spread, short-term interest rate, nor the growth rate of industrial production show significant relation to expected momentum return in the presence of the momentum gap.

5.3 Mispricing and Arbitrage

Behavioral patterns underlie the models by Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999), who focus on imperfect formation and revision of investor expectations in response to new information. In these models, momentum profits are the result of mispricing by less sophisticated market participants. Theories of efficient markets that go back to Friedman (1953) argue that arbitrageurs will quickly come in and these profits will dissipate. However, these deviations from fundamental value may persist because of frictions that limit arbitrage activities (Shleifer and Vishny (1997)). Furthermore, Stein (2009) shows that arbitrageurs can be destabilizing when the trade gets crowded.

In this view, the momentum gap is a measure of arbitrage activity for the momentum trade. When arbitrageurs buy past winners and sell past losers, their trades systematically impact asset prices and are reflected in the momentum gap. A large momentum gap indicates the presence of many arbitrageurs and fairly complete convergence of asset prices to fundamental values so the trade is less profitable going forward.[23] We do not see a strong

[23]Lou and Polk (2013) propose another measure of momentum arbitrage activity. Their measure primarily works at a longer horizon (12 to 24 months) and is based on the high-frequency abnormal return correlation among stocks on which a typical momentum strategy would speculate.

trend in the momentum gap in Figure 1. This might seem to be inconsistent with the casual observation that the professional arbitrage industry has grown tremendously over the last 30 years. However, the stock market has also grown much more liquid over this period so it is conceivable that momentum traders' market impacts have stayed relatively constant.

To relate the momentum gap to momentum arbitrage activity, I introduce two variables that proxy for sophisticated investors' exposures to momentum. The first variable is the difference in aggregate institutional ownership between past winners and losers. Gompers and Metrick (2001) document that large institutions have been playing an increasingly important role in the stock market. Some of them are known to employ momentum strategies (e.g., Grinblatt and Titman (1989)). This measure would capture the extent to which institutional investors are pursuing the momentum strategy. Instead of just the institutional ownership of past winners, this identification uses both the tendency of momentum strategies to bet on winners and bet against losers. One drawback to the institutional ownership data is that only long positions are included while momentum is a long-short strategy. To remedy this, I use the difference in aggregate short interest between past losers and winners as a complementary variable.[24] This is arguably a better proxy than the institutional ownership variable because intuitively only the more skilled institutional investors sell stocks short.

Panel A of Table 7 presents results from time series regressions of the form

$$y_t = \beta_0 + \beta_1 \cdot WMLInstitutionalOwnership_{t-1} + \beta_2 \cdot LMWShortInterest_{t-1} + \epsilon_t. \quad (7)$$

In Regression (1), the dependent variable is the momentum gap. In Regression (2), the dependent variable is the residual momentum gap, which is purged of the influences of the existing momentum predictors as well as business cycle variables examined earlier.

The evidence is generally supportive of the idea that the momentum gap captures momentum arbitrage activity. More institutional capital invested in the momentum trade enlarges

[24]Hanson and Sunderam (2013) develop a more sophisticated methodology to draw inference from the short interest data and using their measure could potential yield even stronger results.

the momentum gap. Consistent with the view that short sellers form a more elite subset of institutional investors, the short interest variable shows stronger effects. Furthermore, when the dependent variable is the residual momentum gap, which should better reflect the possible effects of arbitrage trading, both independent variables show higher statistical significance. This makes it less likely that the results are driven by the business cycle. In analyses not reported, neither of my proxy variables trends, thus ruling out spurious regression bias as a driver of my results.

Shleifer and Vishny (1997)'s limits to arbitrage model features a positive performance-flow relationship. The idea is that agency problems lead investors to withdraw equity capital from arbitrageurs following poor returns. In Panel B of Table 7, I estimate monthly regressions for changes in the momentum gap on lagged momentum return with lagged market return as a control. For reasons similar to Panel A, Regression (2) uses as its dependent variable change in the residual momentum gap.

The results show that changes in the momentum gap strongly follow momentum profits. Each percentage point of momentum profits is associated with 0.19 percent increase in the momentum gap. On the other hand, the market return variable is not significant. Nevertheless, these two variables together explain more than 27 percent of the variations in changes in the momentum gap. The literature has found little evidence that momentum profits forecast changes in economic activity (Liew and Vassalou (2000)). We see here that momentum return can significantly predict changes in a measure of the cross-sectional dispersion of stock returns.

I also investigate an implication of Stein (2009)'s model, which predicts that arbitrage activity can be destabilizing when trading becomes too crowded. It is built on two key features. First, no individual arbitrageur knows exactly how much arbitrage capital is available in the aggregate at any point in time. Second, for the trading strategy in question, arbitrageurs do not have an independent estimate of fundamental value. Momentum is a classic example of such an unanchored strategy. Furthermore, Daniel and Moskowitz (2013) show that returns to momentum strategies are characterized by infrequent yet highly devastating crashes. Table 8 presents higher moments of momentum return as a function of the lagged momentum gap.

Consistent with the idea that a large momentum gap indicates crowded momentum trading, Table 8 shows that momentum returns have the largest departure from normality when the momentum gap is large. On the other hand, when the momentum gap is small, the skewness in momentum return is only -0.26, which is comparable to the unconditional skewness of market return.

6 Additional Analysis

6.1 Industry Momentum

In addition to individual stock momentum, I also consider industry momentum, where the basis assets are industry portfolios instead of individual stocks. Moskowitz and Grinblatt (1999) identify industry momentum as an important source of individual stock momentum profits.[25] I am not aware of evidence in the literature that shows industry momentum is predictable, but I investigate here whether the underlying mechanism that drives the momentum gap's predictive ability for individual stock momentum is also present for industry momentum.

Following Fama and French (1997), each stock in my sample is assigned to a value-weighted industry portfolio at the end of June each year based on its four-digit SIC code at that time. At the beginning of each month t, these industries are sorted into five equal-weighted portfolios based on their cumulative returns from month $t-12$ to $t-1$. I do not skip the most recent month in the formation period because, as shown in Moskowitz and Grinblatt (1999), industry portfolios do not exhibit reversals in the short-term (at the

[25]In Barberis and Shleifer (2003), categorical thinking together with extrapolative expectations explain industry momentum.

one-month horizon) as individual stocks do.[26] Table 9 repeats the predictive regressions from Table 3 with adjusted returns of the industry momentum strategy as the dependent variables.

Even though industry momentum is less profitable than individual stock momentum in my sample,[27] its expected returns vary substantially with the momentum gap. The p-values are consistently small, though not as small as those in Table 3. A one standard deviation increase in the momentum gap is associated with a 0.65 percent decrease (-0.06×10.76) in the monthly adjusted return of the industry momentum strategy. This economic magnitude is about half as large as for individual stock momentum. The adjusted R^2 of 1.65 percent is still sizable for a predictive regression at the monthly frequency. In contrast, neither market return, market volatility, nor market illiquidity shows any predictive ability for industry momentum.

6.2 Cross-Sectional Heterogeneity

The large momentum literature has documented significant cross-sectional heterogeneity in momentum profits. For example, Hong, Lim, and Stein (2000) find that the profitability of momentum strategies declines with firm size, consistent with predictions of behavioral theories. Lee and Swaminathan (2000) document that momentum profits increase in trading volume, which they argue can help reconcile intermediate-horizon "underreaction" and long-horizon "overreaction" effects. In this subsection, I examine time-variation in expected momentum returns for subgroups of stocks sorted by these characteristics. The motivation for this analysis is twofold. First, if time-variation in momentum profits is in part driven by the amount of arbitrage capital brought to bear against the strategy, then there is reason to believe that the effect would be stronger for large, liquid stocks. Second, the fact that there is substantial cross-sectional heterogeneity in unconditional momentum returns raises the question of whether my main results are driven entirely by a subset of stocks.

[26] See Jegadeesh (1990) and Lehmann (1990).

[27] Its average adjusted return is 1.06 percent per month versus individual stock momentum's 1.75 percent.

Table 10 repeats the predictive regressions from Table 3 with adjusted returns of the momentum strategy formed across size groups as the dependent variables. Following Fama and French (2008), stocks are sorted independently into three size groups (Micro, Small, and Big) and momentum quintiles. Similarly, Table 11 repeats the predictive regressions from Table 3 with adjusted returns of the momentum strategy formed across trading volume groups as the dependent variables. Following Lee and Swaminathan (2000), stocks are sorted independently into three trading volume groups (Low, Medium, and High) and momentum quintiles.

The results in these tables clearly show that the predictive power of the momentum gap is pervasive across size and trading volume groups. There is very little evidence of cross-sectional heterogeneity in the relation between expected momentum returns and the momentum gap – the coefficients and p-values are surprisingly similar across the groups. In contrast, none of the competing variables is consistently significant at the 1% level. For instance, market illiquidity cannot predict momentum in small stocks with or without the presence of the momentum gap.

6.3 Robustness Tests

Table 12 investigates variations to my empirical methodology to ensure that my central result relating the momentum gap and subsequent momentum returns is robust. I mainly repeat the predictive regressions from Table 3 under alternative scenarios. Unless a control is specified, the momentum gap is the only predictor variable, so only its coefficients and p-values are shown. For comparison, the first row shows the baseline results that are also reported in Table 3. A one standard deviation increase in the momentum gap is associated with a 1.29 percent decrease (-0.12×10.76) in the monthly adjusted return of the momentum strategy. We can reject the null hypothesis that this effect is zero as the associated p-value is significant at any conventional level.

Row 2 and row 3 show the results of the same analysis for two halves of the sample period (1927-1969 and 1970-2012). We see that the predictive power of the momentum gap is robust. It is significant in both subperiods, each of which spans more than 40 years. The result is stronger in the second subperiod, with the coefficient more than twice as large as the coefficient in the first subperiod. This is consistent with the intuition that sophisticated investors such as hedge funds and momentum trading by these specialists have become significantly more prominent over the last 40 years.

In row 4, I consider the case when observations from 1932 and 2009 are excluded. As documented by Daniel and Moskowitz (2013), momentum experiences the biggest crashes in those two years, mainly driven by reversals of past losers. Therefore, it is reassuring to see that while the magnitude of the coefficients are smaller, the momentum gap has significant predictive power outside of the biggest momentum crashes.

Row 5 includes results where past winners and losers are redefined to be stocks in momentum deciles 9 and 2, respectively. The magnitude of the coefficients are smaller, consistent with the idea that momentum traders focus their attention on the extreme winners and losers. I examine the momentum gap's predictive ability for the risk-managed momentum strategy as documented by Barroso and Santa-Clara (2013) in row 6. Interestingly, the coefficients on the long and short legs become much more balanced.

In Row 7, I control for the VIX, which restricts the sample period to 1990-2012. The momentum gap is still significant at the 5% level. The magnitude of the coefficients are similar to those in row 3, which covers a somewhat similar sample period. In row 8, I similarly control for factor return dispersion, as constructed by Stivers and Sun (2010). The coefficients of the momentum gap are not adversely affected at all; in fact, they become slightly larger.

The results on the last row of Table 12 deserve more attention. While testing Zhang (2005)'s theory of the value premium, Stivers and Sun (2010) nevertheless find that factor return dispersion (FRD) can negatively predict momentum returns. Their main measure is the three-month moving-average of the cross-sectional standard deviations of returns across the 100 Fama-French portfolios formed on size and book-to-market. They attribute this predictive ability to *FRD*'s role as a countercyclical state variable. While they are motivated by risk-based theories, they do not make the observation that the procyclical nature of momentum profits remains a challenge to these theories.[28]

By construction, the momentum gap is mechanically related to *FRD* – they are both measures of the cross-sectional dispersion of stock returns. The momentum gap focus on the intermediate horizon, which, as documented by Jegadeesh and Titman (1993), is the horizon at which return continuation works. *FRD*, as constructed by Stivers and Sun (2010), focus on the one-month horizon though they always smooth it by calculating a moving-average. In the multivariate predictive regressions in row 8 of Table 12, the momentum gap drives away all of *FRD*'s predictive ability. Nevertheless, I use the more robust sorting method to investigate *FRD* in Table 13.

In Panel A of Table 13, I repeat the analysis in Table 2 using *FRD* as the sorting variable. We see that the relation between adjusted momentum returns and lagged *FRD* is highly non-linear. The average adjusted momentum return is highest at 2.58% per month in *FRD* quintile 2. It drops to 1.41% in quintile 3, but increases to 1.48% in quintile 4. As a result, the Small-Minus-Large difference is only 1.23% and not statistically significant. By comparison, the Small-Minus-Large difference when the momentum gap is the sorting variable is about twice as large at 2.36% and statistically significant at any conventional level.

I perform a two-way sort in Panel B. I first sort the months in my sample using the momentum gap into three groups (Small, Medium, and Large). Then I further sort the months within these groups using *FRD* into five groups.[29] I am effectively testing for *FRD*'s predictive ability after controlling for the momentum gap. Similar to the results I obtain from the multivariate predictive regressions, there is no evidence of *FRD*'s predictive ability for momentum returns in the presence of the momentum gap. In fact, when the momentum

[28]See Liu and Zhang (2006) and Liu and Zhang (2013).

[29]The results are similar for other numbers of groups and are not shown for brevity.

gap is small, high *FRD* actually precedes *high* momentum returns.

7 Conclusion

In this paper, I test a simple hypothesis derived from risk-based theories of momentum, namely that expected momentum returns vary positively with the formation period return difference between past winners and losers. I term this measure the momentum gap and I find that it *negatively* predicts momentum returns. A battery of robustness tests confirms this main result. These findings are therefore less consistent with the notion that a stock's past return proxies for its loading on a priced risk factor.

I find evidence to support the alternative hypothesis that momentum is a mispricing phenomenon and that the momentum gap reflects the degree to which arbitrageurs are trading the strategy. A large momentum gap indicates the presence of many arbitrageurs and near-complete convergence of prices to fundamental values, thereby explaining why the trade is less profitable going forward. Using the change in the momentum gap as a proxy for strategy-level capital flow, I document a positive performance-flow relationship. Furthermore, I find that when the momentum gap is large, the unanchored momentum trade exhibits significantly heightened instability. This is consistent with the crowded-trade model of Stein (2009). The momentum gap could therefore also be a useful policy tool for detecting systemic risks in real time.

References

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Andrews, Donald W. K., 1991, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817–858.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and Momentum Everywhere, *Journal of Finance* 68, 929–985.

Avramov, Doron, Si Cheng, and Allaudeen Hameed, 2013, Time-Varying Momentum Payoffs and Illiquidity, *Working paper* .

Baker, Malcolm, Ryan Taliaferro, and Jeffrey Wurgler, 2006, Predicting Returns with Managerial Decision Variables: Is There a Small-Sample Bias?, *Journal of Finance* LXI, 1711–1730.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barroso, Pedro, and Pedro Santa-Clara, 2013, Momentum has its moments, *Working paper* .

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal Investment, Growth Options, and Security Returns, *Journal of Finance* LIV, 1553–1607.

Campbell, John Y., and Samuel B. Thompson, 2007, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* 21, 1509–1531.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* LII, 57–82.

Chaves, Denis B., 2012, Eureka! A Momentum Strategy that Also Works in Japan, *Working paper* .

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, Business Cycle, and Time-varying Expected Returns, *Journal of Finance* LVII, 985–1019.

Clark, Todd E., and Michael W. McCracken, 2001, Tests of equal forecast accuracy and encompassing for nested models, *Journal of Econometrics* 105, 85–110.

Cochrane, John H., 2007, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* 21, 1533–1575.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2003, The Value Spread, *Journal of Finance* LVIII, 609–641.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market States and Momentum, *Journal of Finance* LIX, 1345–1365.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreactions, *Journal of Finance* LIII, 1839–1885.

Daniel, Kent, Ravi Jagannathan, and Soohun Kim, 2012, Tail Risk in Momentum Strategy Returns, *Working paper* .

Daniel, Kent, and Tobias Moskowitz, 2013, Momentum Crashes, *Working paper* .

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive Feedback Investment Strategies and Destabilizing Rational Speculation, *Journal of Finance* XLV, 379–395.

Diebold, Francis X., and Roberto S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 134–144.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance* LI, 55–84.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting Anomalies, *Journal of Finance* LXIII, 1653–1678.

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin, 2003, Spurious Regressions in Financial Economics?, *Journal of Finance* LVIII, 1393–1413.

Friedman, Milton, 1953, The Case for Flexible Exchange Rates, in Milton Friedman, ed., *Essays in Positive Economics* (University of Chicago Press, Chicago, IL).

Goetzmann, William N., and Philippe Jorion, 1993, Testing the Predictive Power of Dividend Yields, *Journal of Finance* XLVIII, 663–679.

Gompers, Paul A., and Andrew Metrick, 2001, Institutional Investors and Equity Prices, *Quarterly Journal of Economics* February, 229–259.

Granger, C. W. J., and P. Newbold, 1974, Spurious Regressions in Econometrics, *Journal of Econometrics* 2, 111–120.

Griffin, John M., Xiuqing Ji, and J. Spencer Martin, 2003, Momentum Investing and Business Cycle Risk: Evidence from Pole to Pole, *Journal of Finance* LVIII, 2515–2547.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311–339.

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings, *Journal of Business* 62, 393.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior, *American Economic Review* 85, 1088–1105.

Gromb, Denis, and Dimitri Vayanos, 2010, Limits of Arbitrage, *Annual Review of Financial Economics* 2, 251–275.

Hanson, Samuel G., and Adi Sunderam, 2013, The Growth and Limits of Arbitrage : Evidence from Short Interest, *Working paper* .

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional Skewness in Asset Pricing Tests, *Journal of Finance* LV, 1263–1295.

Harvey, David, Stephen Leybourne, and Paul Newbold, 1997, Testing the Equality of Prediction Mean Squared Errors, *International Journal of Forecasting* 13, 281–291.

Hong, Harrison, Terence Lim, and Jeremy C Stein, 2000, Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies, *Journal of Finance* LV, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance* LIV, 2143–2184.

Jegadeesh, Narasimhan, 1990, Evidence of Predictable Behavior of Security Returns, *Journal of Finance* XLV, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* XLVIII, 65–91.

Johnson, Timothy C., 2002, Rational Momentum Effects, *Journal of Finance* LVII, 585–608.

Lee, Charles M. C., and Bhaskaran Swaminathan, 2000, Price Momentum and Trading Volume, *Journal of Finance* LV, 2017–2069.

Lehmann, Bruce N., 1990, Fads, Martingales, and Market Efficiency, *Quarterly Journal of Economics* CV, 1–28.

Liew, Jimmy, and Maria Vassalou, 2000, Can book-to-market , size and momentum be risk factors that predict economic growth?, *Journal of Financial Economics* 57, 221–245.

Liu, Laura Xiaolei, and Lu Zhang, 2006, Momentum Profits, Factor Pricing, and Macroeconomic Risk, *Review of Financial Studies* 21, 2417–2448.

Liu, Laura Xiaolei, and Lu Zhang, 2013, A Model of Momentum, *Working paper* .

Lou, Dong, and Christopher Polk, 2013, Comomentum: Inferring Arbitrage Activity from Return Correlations, *Working paper* .

McCracken, Michael W., 2007, Asymptotics for out of sample tests of Granger causality, *Journal of Econometrics* 140, 719–752.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do Industries Explain Momentum ?, *Journal of Finance* LIV, 1249–1290.

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

Nofsinger, John R., and Richard W. Sias, 1999, Herding and Feedback Trading by Institutional and Individual Investors, *Journal of Finance* LIV, 2263–2295.

Rouwenhorst, K. Geert, 1998, International Momentum Strategies, *Journal of Finance* LIII, 267–284.

Rouwenhorst, K. Geert, 1999, Local Return Factors and Turnover in Emerging Stock Markets, *Journal of Finance* LIV, 1439–1464.

Sagi, Jacob S., and Mark S. Seasholes, 2007, Firm-specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389–434.

Shleifer, Andrei, and Robert W. Vishny, 1997, The Limits of Arbitrage, *Journal of Finance* LII, 35–55.

Shumway, Tyler, 1997, The Delisting Bias in CRSP Data, *Journal of Finance* LII, 327–340.

Shumway, Tyler, and Vincent A. Warther, 1999, The Delisting Bias in CRSPs Nasdaq Data and Its Implications for the Size Effect, *Journal of Finance* LIV, 2361–2379.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Stein, Jeremy C., 2009, Sophisticated Investors and Market Efficiency, *Journal of Finance* 64, 1517–1548.

Stivers, Chris, and Licheng Sun, 2010, Cross-Sectional Return Dispersion and Time Variation in Value and Momentum Premiums, *Journal of Financial and Quantitative Analysis* 45, 987–1014.

Wahal, Sunil, and S.G. Badrinath, 2002, Momentum Trading by Institutions, *Journal of Finance* LVII, 2449–2478.

Wang, Kevin Q., and Jianguo Xu, 2011, Market Volatility and Momentum, *Working paper* .

Welch, Ivo, and Amit Goyal, 2007, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455–1508.

White, Halbert, 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and A Direct Test for Heteroskedasticity, *Econometrica* 48, 817–838.

Zhang, Lu, 2005, The Value Premium, *Journal of Finance* LX, 67–103.

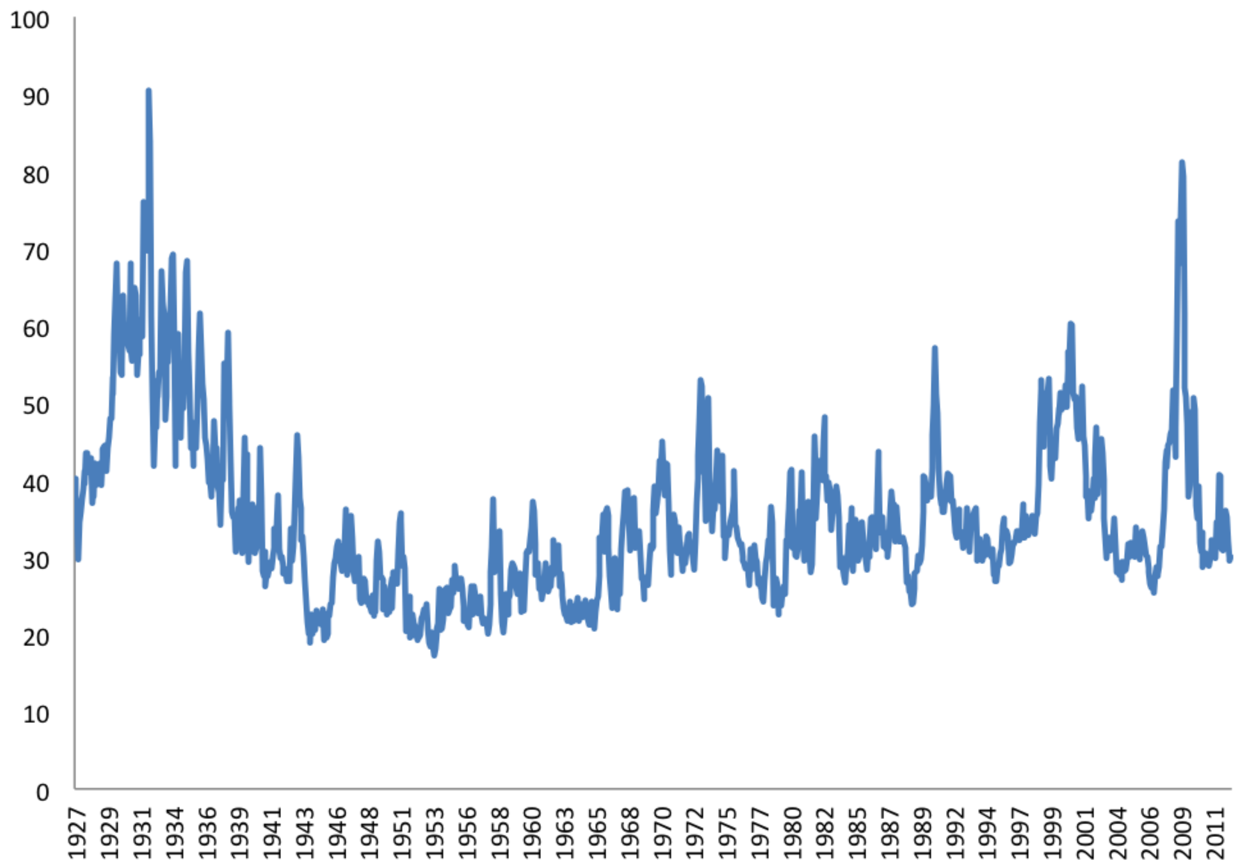

Figure 1: Measures of the Momentum Gap

This figure shows the standardized time series of two momentum gap measures. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MomentumGap2* is the difference between the 90th and 10th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$.

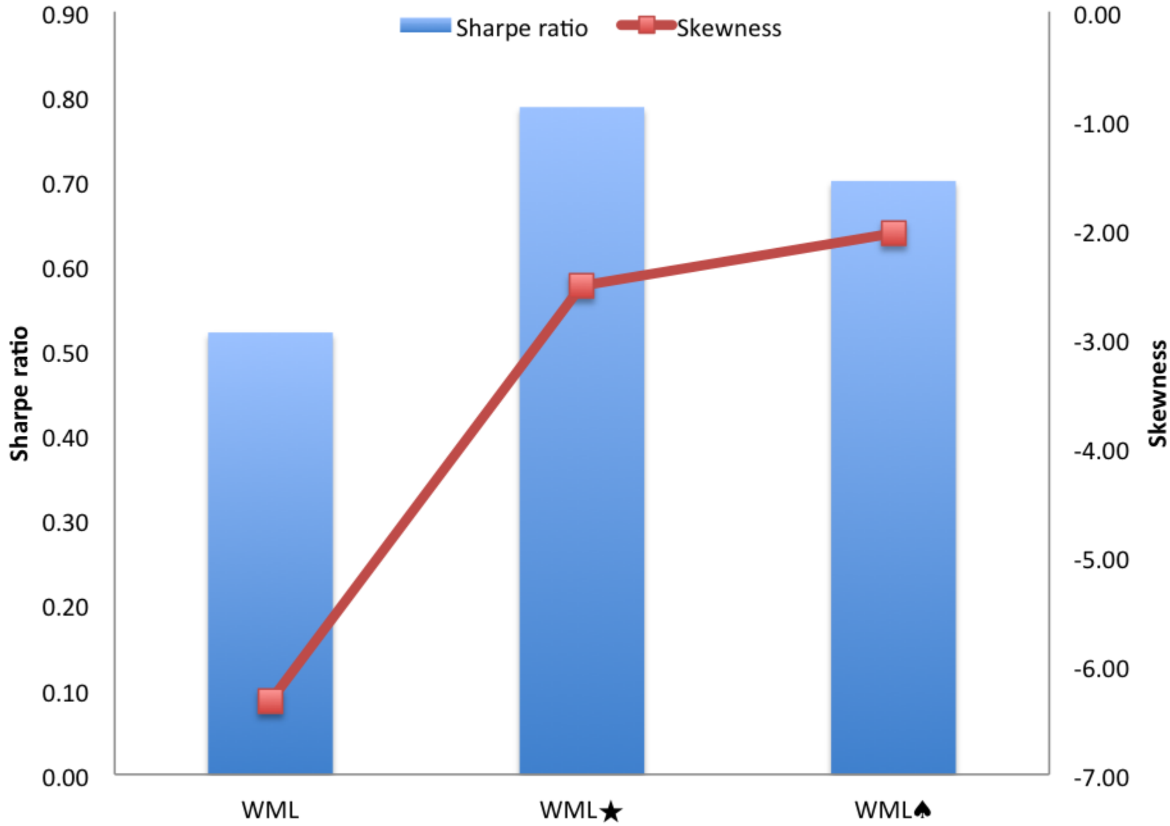

Figure 2: Unconditional vs. Conditional Momentum Strategies
This figure compares the Sharpe ratios and skewness of the unconditional momentum strategy and two conditional strategies based on the momentum gap. At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. WML represents the momentum strategy, which buys past winners and sells past losers. The momentum gap is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. WML$^{\bigstar}$ represents the conditional strategy that takes a position in WML at the beginning of month t unless the momentum gap is ranked in the top quintile. WML$^{\spadesuit}$ represents the conditional strategy that takes a position in WML at the beginning of month t unless a negative return is predicted in a predictive regression using the momentum gap. Both conditional strategies use an expanding look-back window with initial length of 30 months. The sample period is 1926 to 2012.

Table 1: Summary Statistics

This table presents summary statistics for the main predictor variables. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. All returns are in logs. Panel A presents summary statistics of these variables for the full sample. Panel B and Panel C present the subsample means and standard deviations, respectively. Panel D presents the correlations. Panel E presents the 12-month autocorrelations. The sample period is 1926 to 2012.

	MomentumGap	*MarketReturn*	*MarketVolatility*	*MarketIlliquidity*
Panel A: Summary Statistics				
Mean	34.85	28.02	4.87	0.93
SD	10.76	35.16	2.33	2.67
Min	17.21	-168.31	2.02	0.00
Median	32.18	32.94	4.25	0.15
Max	90.53	110.22	14.95	34.85
Panel B: Subsample Means				
1927-1940	48.40	13.29	8.10	4.07
1941-1960	26.21	38.49	3.95	0.90
1961-1980	31.12	21.71	4.15	0.21
1981-2000	35.84	46.58	4.33	0.04
2001-2012	38.00	7.37	4.71	0.01
Panel C: Subsample SDs				
1927-1940	11.67	62.76	3.71	5.46
1941-1960	4.81	23.50	1.15	1.06
1961-1980	6.29	19.83	0.99	0.17
1981-2000	6.78	15.40	1.13	0.02
2001-2012	10.89	28.66	1.42	0.01
Panel D: Correlations				
MomentumGap		-0.37	0.56	0.45
MarketReturn			-0.56	-0.59
MarketVolatility				0.62
MarketIlliquidity				
Panel E: Autocorrelations				
MomentumGap$_{t+12}$	0.62			
MarketReturn$_{t+12}$		0.65		
MarketVolatility$_{t+12}$			0.86	
MarketIlliquidity$_{t+12}$				0.75

Table 2: Momentum Returns by the Lagged Momentum Gap

This table presents the average Fama and French (1993) three-factor adjusted returns of the momentum strategy by the lagged momentum gap. At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. "Long+Short" represents the momentum strategy, which buys past winners and sells past losers. "Long" represents the long leg of the momentum strategy, which buys past winners and sells the market portfolio. "Short" represents the short leg of the momentum strategy, which buys the market portfolio and sells past losers. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. All returns are in logs. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). t-statistics are in parentheses. Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

Rank	Num. Obs.	Avg. *MomentumGap*	Long+Short	Long	Short
Small	207	23.26	2.23	0.96	1.31
			(7.89)	(5.62)	(8.07)
2	206	28.80	1.90	0.64	1.29
			(6.42)	(3.85)	(6.92)
3	206	32.28	1.93	0.71	1.28
			(5.20)	(3.34)	(6.04)
4	206	37.70	1.50	0.61	0.98
			(3.09)	(2.83)	(3.10)
Large	207	52.17	-0.13	-0.02	0.14
			(-0.15)	(-0.05)	(0.25)
Small-Large			2.36	0.98	1.16
			[0.01]	[0.01]	[0.05]

Table 3: Predictive Regressions of Momentum Returns

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. In Panel A, the dependent variable is the Fama and French (1993) three-factor adjusted return of the momentum strategy, which buys past winners and sells past losers. In Panel B, the dependent variable is the adjusted return of the long leg of the momentum strategy, which buys past winners and sells the market portfolio. In Panel C, the dependent variable is the adjusted return of the short leg of the momentum strategy, which buys the market portfolio and sells past losers. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

	MomentumGap	*MarketReturn*	*MarketVolatility*	*MarketIlliquidity*	$\overline{R}^2$
		Panel A: Long+Short			
(1)	-0.12				3.35
	[0.000]				
(2)		0.02	0.21	-0.39	3.20
		[0.010]	[0.159]	[0.003]	
(3)	-0.12	0.02	0.46	-0.32	5.22
	[0.000]	[0.033]	[0.008]	[0.005]	
		Panel B: Long			
(1)	-0.04				1.69
	[0.001]				
(2)		0.00	0.02	-0.15	1.77
		[0.302]	[0.579]	[0.008]	
(3)	-0.03	0.00	0.08	-0.13	2.41
	[0.004]	[0.439]	[0.126]	[0.010]	
		Panel C: Short			
(1)	-0.07				2.71
	[0.000]				
(2)		0.02	0.17	-0.20	2.60
		[0.002]	[0.100]	[0.008]	
(3)	-0.08	0.02	0.34	-0.15	4.60
	[0.000]	[0.007]	[0.005]	[0.017]	

Table 4: Out-of-Sample Forecasting Performance

This table presents statistics on out-of-sample forecast errors for momentum return forecasts at the monthly frequency. At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. The out-of-sample $\overline{R}^2$ is defined in Equation (2). A star next to it is based on significance of MSE-F statistic by McCracken (2007), which tests for equal MSE of the unconditional forecast and the conditional forecast. One-sided critical values of MSE-F statistic are obtained from McCracken (2007). One, two, and three stars denote significance levels at 10%, 5%, and 1%, respectively. ΔRMSE is the RMSE difference between the unconditional forecast and the conditional forecast for the same sample/forecast period. Positive numbers signify superior out-of-sample conditional forecast. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. Panel A presents results where momentum return is assumed to be linear in the lagged predictor. Panel B presents results where momentum return is assumed to be quadratic in the lagged predictor. The sample period is 1926 to 2012.

	Panel A: Linear Specification		Panel B: Quadratic Specification	
	$\overline{R}^2$	ΔRMSE	$\overline{R}^2$	ΔRMSE
MomentumGap	0.58***	0.03	5.45***	0.28
MarketReturn	0.13*	0.01	0.50***	0.03
MarketVolatility	-2.20	-0.10	-13.27	-0.62
MarketIlliquidity	2.52***	0.13	-5.32	-0.25

Table 5: International Predictive Regressions of Momentum Returns

This table presents results from country-specific predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + \epsilon_t.$$

At the beginning of each month t, stocks are sorted into five quintiles based on their cumulative returns from month $t-12$ to $t-2$. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. The dependent variable is the regional Fama and French (1993) three-factor adjusted return of the momentum strategy, which buys past winners (top quintile) and sells past losers (bottom quintile). The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is July 1989 to August 2013.

Country	*MomentumGap*	Country	*MomentumGap*
Australia	-0.07	Japan	-0.05
	[0.006]		[0.125]
Austria	-0.16	Netherlands	-0.10
	[0.001]		[0.007]
Belgium	-0.11	New Zealand	0.02
	[0.025]		[0.768]
Canada	-0.19	Norway	-0.08
	[0.010]		[0.048]
Denmark	-0.09	Portugal	-0.06
	[0.055]		[0.150]
Finland	-0.11	Singapore	-0.13
	[0.035]		[0.009]
France	-0.12	Spain	-0.13
	[0.010]		[0.103]
Germany	-0.01	Sweden	-0.13
	[0.436]		[0.017]
Greece	-0.10	Switzerland	-0.06
	[0.016]		[0.192]
Ireland	-0.23	United Kingdom	-0.12
	[0.002]		[0.002]
Italy	-0.08		
	[0.086]		

Table 6: Predictive Regressions Controlling for Macroeconomic Variables

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *DIV* is the lagged dividend yield of the CRSP value-weighted index. *DEF* is the lagged default spread. *TERM* is the lagged term spread. *YLD* is the lagged short-term interest rate. *MP* is the lagged industrial production growth. In Panel A, the dependent variable is the Fama and French (1993) three-factor adjusted return of the momentum strategy, which buys past winners and sells past losers. In Panel B, the dependent variable is the adjusted return of the long leg of the momentum strategy, which buys past winners and sells the market portfolio. In Panel C, the dependent variable is the adjusted return of the short leg of the momentum strategy, which buys the market portfolio and sells past losers. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

	MomentumGap	*DIV*	*DEF*	*TERM*	*YLD*	*MP*	$\bar{R}^2$
	Panel A: Long+Short						
(1)	-0.12						3.35
	[0.000]						
(2)	-0.10	-0.16	-0.43	-0.04	0.09	0.13	3.70
	[0.000]	[0.236]	[0.810]	[0.850]	[0.582]	[0.599]	
	Panel B: Long						
(1)	-0.04						1.69
	[0.001]						
(2)	-0.02	0.02	-0.45	0.01	0.02	0.03	1.74
	[0.026]	[0.831]	[0.114]	[0.881]	[0.915]	[0.669]	
	Panel C: Short						
(1)	-0.07						2.71
	[0.000]						
(2)	-0.07	-0.17	0.03	-0.05	0.07	0.10	3.21
	[0.000]	[0.078]	[0.422]	[0.849]	[0.452]	[0.606]	

Table 7: The Momentum Gap as a Measure of Arbitrage Activity

Panel A presents results from time series regressions of the form

$$y_t = \beta_0 + \beta_1 \cdot WMLInstitutionalOwnership_{t-1} + \beta_2 \cdot LMWShortInterest_{t-1} + \epsilon_t.$$

WMLInstitutionalOwnership is the percentage difference in aggregate institutional ownership between past winners and losers. *LMWShortInterest* is the percentage difference in aggregate short interest between past losers and winners. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MomentumGap*$\perp$ is the residual from regressing *MomentumGap* on *MarketReturn*, *MarketVolatility*, *MarketIllquidity*, the dividend yield, default spread, term spread, short-term interest rate, and industrial production growth. t-statistics are computed using heteroskedasticity-and-autocorrelation-consistent standard errors of Andrews (1991) and are in parentheses. The sample period is March 1980 to June 2012.

Panel B presents results from time series regressions of the form

$$z_t = \gamma_0 + \gamma_1 \cdot MomentumReturn_{t-1} + \gamma_2 \cdot MarketReturn_{t-1} + \epsilon_t.$$

MomentumReturn is the return to the momentum strategy. *MarketReturn* is the return on the CRSP value-weighted index. t-statistics are computed using heteroskedasticity-consistent standard errors of White (1980) and are in parentheses. The sample period is March 1927 to December 2012.

Panel A: Determinants of the Momentum Gap

	Dep. Var.	$WMLInstitutionalOwnership_{t-1}$	$LMWShortInterest_{t-1}$	$\overline{R}^2$	Num. Obs.
(1)	*MomentumGap*	0.13	1.62	7.87	388
		(0.92)	(2.67)		
(2)	*MomentumGap*$_\perp$	0.27	1.22	10.85	388
		(2.47)	(2.94)		

Panel B: Performance-Flow Relationship

	Dep. Var.	$MomentumReturn_{t-2}$	$MarketReturn_{t-2}$	$\overline{R}^2$	Num. Obs.
(1)	*ΔMomentumGap*	0.05		1.57	1029
		(3.15)			
(2)	*ΔMomentumGap*	0.04	-0.05	2.04	1029
		(2.46)	(-1.83)		

Table 8: Higher Moments of Momentum Returns

This table presents higher moments of returns to the momentum strategy by the lagged momentum gap. At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. "Long+Short" represents the momentum strategy, which buys past winners and sells past losers. "Long" represents the long leg of the momentum strategy, which buys past winners and sells the market portfolio. "Short" represents the short leg of the momentum strategy, which buys the market portfolio and sells past losers. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. All returns are in logs. Panel A presents the results for skewness. Panel B presents the results for kurtosis. Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

Rank	Num. Obs.	Avg. *MomentumGap*	Long+Short	Long	Short
Panel A: Skewness					
Small	207	23.26	-0.26	0.37	-0.60
2	206	28.80	-0.33	-0.41	-0.30
3	206	32.28	-0.48	0.02	-0.65
4	206	37.70	-2.67	-0.70	-2.78
Large	207	52.17	-3.83	-1.09	-2.51
Small-Large			3.57	1.46	1.91
			[0.00]	[0.01]	[0.01]
Panel B: Kurtosis					
Small	207	23.26	4.47	3.49	5.76
2	206	28.80	3.09	4.11	3.36
3	206	32.28	3.67	3.75	4.00
4	206	37.70	20.59	4.34	22.23
Large	207	52.17	27.66	7.47	13.76
Small-Large			-23.18	-3.99	-8.01
			[0.01]	[0.01]	[0.08]

Table 9: Predictive Regressions of Industry Momentum Returns

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

Following Fama and French (1997), each stock is assigned to an industry portfolio at the end of June each year based on its four-digit SIC code at that time. At the beginning of each month t, these industries are sorted into five portfolios based on their cumulative returns from month $t-12$ to $t-1$. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. In Panel A, the dependent variable is the Fama and French (1993) three-factor adjusted return of the industry momentum strategy, which buys past winner industries and sells past loser industries. In Panel B, the dependent variable is the adjusted return of the long leg of the industry momentum strategy, which buys past winner industries and sells the market portfolio. In Panel C, the dependent variable is the adjusted return of the short leg of the industry momentum strategy, which buys the market portfolio and sells past loser industries. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is June 1926 to December 2012.

	MomentumGap	*MarketReturn*	*MarketVolatility*	*MarketIlliquidity*	$\overline{R}^2$
		Panel A: Long+Short			
(1)	-0.06				1.65
	[0.001]				
(2)		0.00	-0.08	-0.19	1.78
		[0.630]	[0.668]	[0.031]	
(3)	-0.04	0.00	0.00	-0.17	2.17
	[0.013]	[0.786]	[0.689]	[0.052]	
		Panel B: Long			
(1)	-0.02				0.42
	[0.027]				
(2)		0.00	-0.03	-0.05	0.53
		[0.700]	[0.814]	[0.078]	
(3)	-0.01	0.00	-0.01	-0.05	0.52
	[0.164]	[0.797]	[0.807]	[0.102]	
		Panel C: Short			
(1)	-0.04				1.50
	[0.000]				
(2)		0.00	-0.06	-0.11	1.39
		[0.698]	[0.568]	[0.091]	
(3)	-0.03	0.00	0.00	-0.10	1.79
	[0.022]	[0.845]	[0.837]	[0.140]	

Table 10: Predictive Regressions of Momentum Across Size Groups

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

Following Fama and French (2008), Micro stocks are below the 20th percentile of NYSE market capitalization, Small stocks are between the 20th and 50th percentiles, and Big stocks are above the median. Stocks are also sorted (independently) into momentum quintiles using breakpoints calculated from NYSE Small and Big stocks. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. In Panel A, the dependent variable is the Fama and French (1993) three-factor adjusted return of the momentum strategy that buys Micro winners and sells Micro losers. In Panel B, the dependent variable is the adjusted return of the momentum strategy that buys Small winners and sells Small losers. In Panel C, the dependent variable is the adjusted return of the momentum strategy that buys Big winners and sells Big losers. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

	MomentumGap	*MarketReturn*	*MarketVolatility*	*MarketIlliquidity*	$\overline{R}^2$
			Panel A: Micro		
(1)	-0.09				2.40
	[0.000]				
(2)		0.02	0.14	-0.43	4.35
		[0.034]	[0.389]	[0.001]	
(3)	-0.06	0.01	0.27	-0.39	5.04
	[0.009]	[0.060]	[0.096]	[0.003]	
			Panel B: Small		
(1)	-0.07				2.21
	[0.000]				
(2)		0.02	0.12	-0.08	1.38
		[0.002]	[0.240]	[0.186]	
(3)	-0.08	0.02	0.29	-0.04	3.32
	[0.000]	[0.009]	[0.014]	[0.350]	
			Panel C: Big		
(1)	-0.09				3.04
	[0.000]				
(2)		0.01	0.13	-0.32	2.40
		[0.159]	[0.173]	[0.000]	
(3)	-0.09	0.01	0.32	-0.27	4.35
	[0.000]	[0.323]	[0.009]	[0.001]	

Table 11: Predictive Regressions of Momentum Across Trading Volume Groups

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

Following Lee and Swaminathan (2000), stocks are independently sorted into three trading volume groups and momentum quintiles using NYSE breakpoints. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. *MarketReturn* is the lagged three-year return on the CRSP value-weighted index. *MarketVolatility* is the lagged three-year monthly return volatility of the CRSP value-weighted index. *MarketIlliquidity* is the lagged value-weighted average of the Amihud (2002) stock-level illiquidity measure for all NYSE and AMEX stocks. In Panel A, the dependent variable is the Fama and French (1993) three-factor adjusted return of the momentum strategy that buys low volume winners and sells low volume losers. In Panel B, the dependent variable is the adjusted return of the momentum strategy that buys medium volume winners and sells medium volume losers. In Panel C, the dependent variable is the adjusted return of the momentum strategy that buys high volume winners and sells high volume losers. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

	MomentumGap	*MarketReturn*	*MarketVolatility*	*MarketIlliquidity*	$\overline{R}^2$
			Panel A: Low		
(1)	-0.10				3.06
	[0.000]				
(2)		0.01	0.07	-0.34	2.52
		[0.237]	[0.386]	[0.006]	
(3)	-0.09	0.01	0.26	-0.28	4.14
	[0.000]	[0.419]	[0.027]	[0.010]	
			Panel B: Medium		
(1)	-0.09				2.88
	[0.000]				
(2)		0.01	0.13	-0.19	1.32
		[0.029]	[0.166]	[0.010]	
(3)	-0.10	0.01	0.35	-0.13	3.78
	[0.000]	[0.090]	[0.007]	[0.029]	
			Panel C: High		
(1)	-0.09				2.89
	[0.000]				
(2)		0.01	0.15	-0.33	2.56
		[0.072]	[0.172]	[0.001]	
(3)	-0.09	0.01	0.34	-0.27	4.36
	[0.000]	[0.166]	[0.011]	[0.001]	

Table 12: Robustness of Main Results

This table presents results from predictive regressions of the form

$$r_t = a + b \cdot MomentumGap_{t-1} + c' \cdot D_{t-1} + \epsilon_t.$$

At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. Unless a control is specified, *MomentumGap* is the only predictor variable, so only its coefficients and p-values are shown. In the column "Long+Short", the dependent variable is the Fama and French (1993) three-factor adjusted return of the momentum strategy, which buys past winners and sells past losers. In the column "Long", the dependent variable is the adjusted return of the long leg of the momentum strategy, which buys past winners and sells the market portfolio. In the column "Short", the dependent variable is the adjusted return of the short leg of the momentum strategy, which buys the market portfolio and sells past losers. Row 1 shows the baseline results that are also reported in Table 3. Row 2 and Row 3 show results of the same analysis for two halves of the sample period. Row 4 considers the case when observations from 1932 and 2009 are excluded. In Row 5, past winners and losers are redefined to be stocks in momentum deciles 9 and 2, respectively. Row 6 shows results for the risk-managed momentum strategy. In Row 7, I control for the VIX. In Row 8, I control for factor return dispersion, which is calculated as the lagged three-month moving-average of the cross-sectional standard deviations of returns across the 100 Fama-French portfolios formed on size and book-to-market. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

	Long+Short	Long	Short
1. Full sample	-0.12	-0.04	-0.07
	[0.000]	[0.001]	[0.000]
2. Subsample: 1927-1969	-0.09	-0.03	-0.05
	[0.002]	[0.015]	[0.005]
3. Subsample: 1970-2012	-0.22	-0.06	-0.14
	[0.003]	[0.020]	[0.005]
4. Subsample: exclude 1932 and 2009	-0.05	-0.03	-0.02
	[0.029]	[0.010]	[0.217]
5. Long decile 9 and short decile 2	-0.09	-0.02	-0.06
	[0.000]	[0.003]	[0.000]
6. Risk-managed momentum	-0.06	-0.03	-0.03
	[0.000]	[0.000]	[0.002]
7. Control for VIX	-0.33	-0.05	-0.26
	[0.063]	[0.351]	[0.049]
8. Control for factor return dispersion	-0.13	-0.05	-0.08
	[0.000]	[0.000]	[0.000]

Table 13: Momentum Returns and Factor Return Dispersion

This table presents the Fama and French (1993) three-factor adjusted returns of the momentum strategy by lagged factor return dispersion (*FRD*). At the beginning of each month t, stocks are sorted into ten portfolios based on their cumulative returns from month $t-12$ to $t-2$. "Long+Short" represents the momentum strategy, which buys past winners and sells past losers. "Long" represents the long leg of the momentum strategy, which buys past winners and sells the market portfolio. "Short" represents the short leg of the momentum strategy, which buys the market portfolio and sells past losers. *FRD* is the lagged three-month moving-average of the cross-sectional standard deviations of returns across the 100 Fama-French portfolios formed on size and book-to-market. *MomentumGap* is the difference between the 75th and 25th percentiles of the distribution of cumulative stock returns from month $t-12$ to $t-2$. All returns are in logs. In Panel A, a one-way sort is performed using *FRD*. In Panel B, a two-way sort is performed: first by *MomentumGap*, then by *FRD*. The adjusted return is defined as the sum of α and the fitted value of ϵ_t in the full-period regression in Equation (3). t-statistics are in parentheses. Two-sided bootstrapped p-values are in brackets. The sample period is 1926 to 2012.

Panel A: One-Way Sort by *FRD*

Rank	Num. Obs.	Avg. *FRD*	Long+Short	Long	Short
Small	207	2.17	1.60	0.49	1.14
			(5.29)	(2.98)	(5.90)
2	206	2.66	2.58	1.05	1.59
			(7.12)	(5.90)	(6.86)
3	206	3.16	1.41	0.54	0.93
			(3.47)	(2.62)	(3.78)
4	206	4.02	1.48	0.48	1.07
			(3.13)	(2.10)	(3.32)
Large	207	7.77	0.36	0.34	0.28
			(0.45)	(1.04)	(0.52)
Small-Large			1.23	0.16	0.86
			[0.16]	[0.67]	[0.14]

Panel B: Two-Way Sort by *MomentumGap-FRD*

Rank of *FRD*	Small *MomentumGap*			Medium *MomentumGap*			Large *MomentumGap*		
	Long+Short	Long	Short	Long+Short	Long	Short	Long+Short	Long	Short
Small	1.93	0.65	1.31	1.09	0.19	0.93	1.87	0.88	1.10
	(4.42)	(2.52)	(4.81)	(2.20)	(0.69)	(2.71)	(2.02)	(2.54)	(1.68)
2	3.25	1.35	1.95	2.09	0.74	1.42	0.40	0.22	0.29
	(7.20)	(5.07)	(7.07)	(3.28)	(2.05)	(3.96)	(0.38)	(0.55)	(0.39)
3	2.01	0.90	1.14	2.93	1.11	1.89	-0.83	-0.57	-0.05
	(4.36)	(3.33)	(4.10)	(4.96)	(3.23)	(5.57)	(-0.58)	(-1.08)	(-0.04)
4	1.53	0.52	1.03	0.97	0.37	0.66	1.94	0.38	1.79
	(3.17)	(1.75)	(3.81)	(1.49)	(1.01)	(1.91)	(1.49)	(0.71)	(2.17)
Large	2.65	1.18	1.53	0.83	0.60	0.32	-0.40	0.18	-0.29
	(4.78)	(3.61)	(4.49)	(0.96)	(1.36)	(0.59)	(-0.26)	(0.28)	(-0.30)
S-L	-0.72	-0.53	-0.22	0.26	-0.41	0.61	2.28	0.70	1.39
	[0.30]	[0.20]	[0.62]	[0.79]	[0.42]	[0.34]	[0.20]	[0.33]	[0.23]