What Explains Momentum?
A Perspective From International Data

Amit Goyal
Narasimhan Jegadeesh
Avanidhar Subrahmanyam*

February 2022

*Amit Goyal is from Swisss Finance Institute at the University of Lausanne, email: Amit.Goyal@unil.ch, Tel: +41-21-692-3676; Narasimhan Jegadeesh is from Goizueta Business School, Emory University, email: jegadeesh@emory.edu, Tel: +1-404-384-3503; and Avanidhar Subrahmanyam is from Anderson School, UCLA, email: asubrahm@anderson.ucla.edu, Tel: +1-310-825-5355. We thank Michael Brennan, Andrea Eisfeldt, Valentin Haddad, and Siew Hong Teoh for helpful comments. All errors are our own.

Abstract

What Explains Momentum?
A Perspective From International Data

There is as yet no consensus on why equity markets permit momentum, although the literature proposes several explanations. Our analysis uses out-of-sample international data to consider a "horse race" across existing empirical proxies for momentum rationales used by earlier studies. Our central finding in cross-sectional analyses is that the proxy for the frog-in-the-pan (FIP) hypothesis, which posits that due to limited attention, investors underreact to information that arrives gradually rather than in concentrated doses, consistently wins. Also, internationally, momentum is stronger in less volatile markets and in up-markets. The FIP proxy indicates that information flows more gradually during these market states, implying additional support for the hypothesis.

1 Introduction

The market efficiency debate is central to the field of finance, and continues unabated. A significant contribution to this debate is the extensive evidence of stock market momentum, which is the tendency of stocks' relative performance to be predictable from their relative performance in the past three to twelve months. This return pattern, uncovered by Jegadeesh and Titman (1993, JT), directly contradicts the notion of weak form market efficiency (that stock prices have no memory of past prices) and hence understanding its source is of fundamental importance.

Fama and French (1996) are unable to capture momentum via their three-factor model (Fama and French, 1993). Motivated by this finding, in recent years, much additional work has sought to uncover the drivers of momentum and the literature proposes a large number of rationales for this phenomenon. Given the numerous explanations for momentum that have accumulated,[1] identifying those explanations that are robust is a vital issue in financial economics.

We consider a broad set of rationales for momentum *out of sample*. Asness, Moskowitz, and Pedersen (2013), Griffin, Ji, and Martin (2003), and Rouwenhorst (1998) show that the profitability of momentum strategies readily extends to non-US markets. Accordingly, we test whether the proposed momentum explanations within earlier literature hold internationally. Since empirical tests of such explanations mostly use US data, our analysis serves as an out-of-sample test of the findings. Further, we minimize our judgment calls by using empirical proxies for explanations that are used in the original studies. Additionally, our tests have the virtue of examining all the rationales on an equal footing that uses the same methodology and the same data. Such a basis is better able to assess the available explanations.

The empirical tests of explanations for momentum fall into two broad categories. One

[1] See Jegadeesh and Titman (2011) and Subrahmanyam (2018) for reviews of the momentum literature.

category tests explanations that involve cross-sectional variations in momentum. In this category, a specific variable is proposed to explain conditional variations in the *cross-sectional strength* of the momentum effect, and this variable usually proxies for a specific explanation for momentum. In the second category, the literature tests for differences in momentum *over time* depending on the state of the market (direction or volatility), and such differences are used to support a proposed rationale. We examine the robustness of explanations proposed across both these categories.

One prominent stream of research in the cross-sectional category tests behavioral theories of momentum. For example, Daniel, Hirshleifer, and Subrahmanyam (1998, DHS) present a model where investor overconfidence leads to momentum.[2] The idea here is that overconfidence builds as investors receive public signals that confirm their initial trading decisions but does not subside equally when they receive disconfirming ones. This leads to momentum due to continuing overreaction, on average. Daniel and Titman (1999, DT) argue that overconfidence plays a bigger role in the valuation of a firm's uncertain growth options and hence propose book-to-market as a proxy for overconfidence. Lee and Swaminathan (2000, LS) argue that the degree of investor overconfidence is reflected in stocks' turnover[3] and hence use this quantity as a proxy for overconfidence. These latter papers present evidence supporting the overconfidence hypothesis with US data. We test the robustness of the overreaction hypothesis with international data using the same proxies that these papers use.

Another behavioral explanation for momentum is the Hong and Stein (1999, HS) slow diffusion of information hypothesis. In HS (1999), investors in one category condition their demands on the private information they receive but not on market prices, and investors in another category do not receive private information and they condition their demand only on market prices. HS (1999) show that news diffuses slowly through the investing population

[2]To avoid repeating long references, we abbreviate oft-reused references by author initials. Appendix Table A1 provides a complete list of full references and their corresponding abbreviations.

[3]In Odean (1998) overconfidence implies greater trading activity. This is because over-assessing one's signal precision (a consequence of overconfidence) implies larger positions for any given signal.

in their model, which results in momentum. We follow Hong, Lim, and Stein (2000, HLS) and use the number of analysts following a stock after controlling for firm size as the proxy for the speed of information diffusion.

The other behavioral explanations that we examine are the Da, Gurun, and Warachka (2014, DGW) frog-in-the pan (FIP) hypothesis and the anchoring bias hypothesis of George and Hwang (2004, GH). The former explanation proposes that momentum arises because investors underreact to small bits of information due to limited attention (Hirshleifer and Teoh, 2003), while reacting appropriately to large chunks. DGW (2014) propose a proxy for the discreteness of information arrival and show that momentum is inversely related to this proxy. GH (2004) propose that the ratio of current prices to their 52-week high is related to the degree of underreaction to news, as investors are anchored to that high. They provide evidence in support of their hypothesis. We use the measures proposed by these authors to examine the robustness of their findings out of sample.

Turning to rationales based on risk-pricing, Sagi and Seasholes (2007, SS) build on Johnson (2002) and propose a rational explanation for momentum. Johnson's (2002) model uses the notion that prices are convex in growth rates, and hence growth rate risk is higher with higher growth rates, and is therefore higher for winners and lower for losers. In his model, momentum profit is the risk premium compensation for bearing the risk of the winner minus loser portfolios. Johnson's (2002) structural model of time-varying risk and risk premiums and its extension in SS (2007) is difficult to estimate. Instead, SS (2007) use real options as the source of growth rate risk, and develop comparative statics to derive predictions about how momentum profits vary across stocks with different characteristics and test those predictions. SS (2007) use the cost of goods sold as an inverse proxy for real options.[4] They also use volatility of sales growth as a real options proxy but many of the firms in our sample do not have sufficient data to compute sales volatility. In SS (2007), bigger sales volatility also implies bigger return volatility and SS (2007) note that momentum increases with return

[4] In SS (2007), firms with low costs benefit more from real options, leading to greater momentum.

volatility as well. We examine the relation between momentum and return volatility.[5]

We conduct a cross-sectional "horse race" across all of the preceding proxies for momentum explanations, using Fama and MacBeth (1973) regressions as well as portfolio-based analyses. We also use regression regularization approaches (i.e., penalized regressions) to examine the competing explanations. Across all of our tests, we find remarkably robust support for the proxy used by DGW (2014) for testing the FIP hypothesis. The news discreteness variable is strongly significant in both emerging and non-US developed markets and is of a sign consistent with DGW (2014). We find some modest but inconsistent support for the overreaction hypothesis using book-to-market ratio as the proxy in some tests. We do not find any support for the other hypotheses.

Our next set of tests examine the time-series relation between momentum profits and market states. Cooper, Gutierrez, and Hameed (2004, CGH) find that momentum profits are higher in up markets than in down markets.[6] They attribute this finding to the notion that investors are more overconfident in up-markets. The logic is that investors who face shorting constraints receive more confirming signals for their buy trades in up-markets. Further, Wang and Xu (2015, WX) uncover that momentum profits are lower in high volatility states. WX (2015) hypothesize that investors become overly risk-averse in highly volatile markets and "over-sell" losers. The subsequent recovery of losers results in the poor performance of momentum in high volatility states. We test whether these empirical results are robust out of sample.

Our investigation finds that momentum profits are bigger and significant in up-markets and in less volatile markets within our international setting, thus confirming CGH (2004) and WX (2015). We also find that the DGW (2014) measure of discrete news arrival is lower in up markets and in low volatility states, suggesting that the time-series findings

[5]In a different rationale for the momentum-volatility link, Zhang (2006) proposes that biases which cause underreaction have a bigger impact when there is more uncertainty.

[6]A closely-related finding is that of Antoniou, Doukas, and Subrahmanyam (2013), who show that momentum profits are higher in periods of optimistic sentiment (Baker and Wurgler, 2006). However, we do not examine the robustness of this hypothesis because sentiment proxies are not available internationally.

might obtain because information arrives less discretely in such market states as per DGW (2014). Additionally, we do not find that momentum profits reverse in the long run, which supports an underreaction story, rather than one based on continuing overreaction. In sum, the arguments of DGW (2014) receive robust support in our out of sample tests.[7]

We emphasize that throughout our work, we consider only the empirical proxies for the theories that have already been identified in the literature and do not experiment with new proxies. So, while it is possible there are proxies for theories other than those proposed in the literature, we take first things first by examining the out of sample evidence with existing proxies.[8]

In closing the introduction, we hasten to add that we do not intend in the slightest to critique any of the papers involved in our empirical work. Indeed, considerable theoretical and empirical insight has been contributed by many colleagues on the topic of why markets permit momentum. Nonetheless, while we acknowledge these insights, we believe there is value in documenting the out-of-sample international evidence on explanations for momentum, holding constant the research methodology and the dataset.[9]

[7]Müller and Müller (2020) analyze variation in momentum profits at the country level. We instead investigate variation in momentum across individual stocks within an international setting. Our approach allows us to investigate the cross-sectional evidence on US momentum out of sample.

[8]For example, Barberis, Shleifer, and Vishny (1998) propose the representativeness bias as an explanation for momentum, and Hong, Stein, and Yu (2007) suggest that investors use overly-simplified models to evaluate stocks, and make persistent forecast errors, which also leads to momentum. Since the empirical literature does not directly consider proxies for their theories, we do not consider their rationales in this paper.

[9]We note that there are other debates surrounding momentum. For example, Moskowitz, Ooi, and Pedersen (2012) find time-series momentum, where long-short portfolios are formed solely based on the sign of past returns rather than their cross-sectional ranks. Goyal and Jegadeesh (2018) show that because of the return premium is on average positive for risky assets, the time series strategy has a bigger exposure on the long side than on the short side, but otherwise both time-series momentum and cross-sectional momentum are virtually the same. Novy-Marx (2012) proposes that momentum profits arise not because recent winners continue to outperform recent losers, but because good performers over the past seven to 12 months continue to outperform bad performers over that same period. Goyal and Wahal (2015) demonstrate that no such "echo" obtains in an international setting comprising more than thirty countries. They show that the US-based result of Novy-Marx (2012) is due to a reversal in the second month prior to portfolio formation. Chordia and Shivakumar (2002) argue that momentum profits can be accounted for by the business cycle, but Griffin, Ji, and Martin (2003) find only modest support for the dependence of momentum profits on the business cycle in international markets. Ehsani and Linnainmaa (2021) show that factor momentum subsumes individual stock momentum (see also Kelly, Moskowitz, and Pruitt, 2021), whereas Falck, Rej, and Thesmar (2020) indicate that this conclusion depends on whether the most recent month is excluded from portfolio formation. We eschew these issues as they do not focus on the underlying drivers of momentum.

The remainder of this paper is organized as follows. Section 2 describes our dataset, and lays out the cross-sectional empirical tests. Section 3 considers the central evidence based on regression analysis. Section 4 performs some robustness checks on the regressions. Section 5 runs a horse race among characteristics with penalized regressions. Section 6 considers the cross-sectional evidence based on portfolio-based analyses. Section 7 considers the evidence on the time-series of momentum profits. Section 8 concludes. The Appendix to this paper contains some ancillary tables, prefixed with the letter 'A.'

2 Data and Cross-Sectional Regression Method

This section first describes the data that we use and then presents an overview of our tests. We then discuss the hypotheses proposed to explain momentum and the empirical proxies that are used to empirically test the cross-sectional implications of these hypotheses. Broadly, these hypotheses imply that the underlying phenomenon that leads to momentum falls in one of the following categories: (i) Continuing overreaction to information that is corrected in the long run, (ii) pure underreaction to information due to various cognitive limitations, and (iii) time-varying expected returns. The section next presents the methodology used for conducting the out-of-sample tests that examine the robustness of these hypotheses. The actual tests appear in Section 3.

2.1 Data

We obtain data for all countries in the MSCI Developed (ex-US) and the MSCI Emerging markets index. There are a total of 22 developed markets and 27 emerging markets in the MSCI indexes for which we are able to get necessary data.[10] The stock market data are from

[10]The developed countries are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Israel, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, and the United Kingdom. The emerging countries are Argentina, Brazil, Chile, China, Colombia, Czech Republic, Egypt, Greece, Hungary, India, Indonesia, Korea, Kuwait, Malaysia, Mexico, Pakistan, Peru, Philippines, Poland, Qatar, Russia, Saudi Arabia, South Africa, Taiwan, Thailand, Turkey,

Datastream and the annual accounting data from Worldscope. Appendix Table A2 provides details on the accounting variables.

For each country, we download data for both listed and delisted companies which have exchange code (EXDSCD) as that of the primary exchange of that country, for which the type of instrument (TYPE) is equity, the indicator ISINID identifies the equity as the primary security, the geography code (GEOGN) identifies the home or listing country of the equity as the same country, and the currency of the equity is the same as that of the country.[11] We exclude depository receipts (DRs), REITS, and preferred stocks, and apply filters described in Tables B.1 and B.2 of Griffin, Kelly, and Nardari (2010).

Because of potential data errors in Datastream and Worldscope, we use data cleaning procedures as in Ince and Porter (2006), Griffin, Kelly, and Nardari (2010), Hou, Karolyi, and Kho (2011), Lee (2011), and Jacobs and Müller (2020). Specifically, we proceed as follows. We download all data in US dollars with five decimal places to minimize return errors stemming from currency conversions. If the gross return in any month is greater than 300% and the product of gross returns in two consecutive months is less than 50% then we set returns in both months as missing. The equivalent numbers for daily returns are 100% and 20%. We discard all daily returns exceeding 100% and all monthly returns exceeding 200%. We also exclude micro-cap stocks (stocks in the top 97% of the market capitalization of each country). For each period (day or month), we winsorize returns in each country at the 0.1% and 99.9% levels. If 90% or more of stocks have zero returns in a period for a country, we set all of them to missing.

and the United Arab Emirates.

[11]We use both Toronto and TSX Ventures for Canada, Shanghai and Shenzhen for China, Deutsche Boerse and Xetra for Germany, BSE and National Stock Exchange for India, Tokyo and Osaka for Japan, and the Korea main exchange as well as KOSDAQ for South Korea as primary exchanges.

2.2 Momentum in international markets

This subsection documents momentum in our international markets sample. The momentum variable that we use is a stock's return over the previous 12 months, excluding the previous month. Specifically, for stock i the momentum variable for month t is the return from month $t-12$ to $t-2$. We country neutralize this return by subtracting its cross-sectional mean across all stocks in our sample from that country. We then rank stocks based on country-neutralized returns and assign each stock to one of ten momentum deciles. Because we country-neutralize the momentum variable, country-specific returns do not affect a stock's decile rank. We define the value-weighted portfolio of stocks in the winner decile minus stocks in the loser decile as the WML hedge portfolio, and we rebalance it monthly.

Table 1 presents the results of the WML hedge portfolio, which is the long-short portfolio formed across extreme winners and losers. The hedge portfolio returns across All ex-US, Developed ex-US, and Emerging markets are, on average, 10.64%, 10.23%, and 8.93% respectively, and are remarkably consistent across the three markets. The medians are slightly greater than the means, and momentum profits are more volatile and more negatively skewed for Developed ex-US. Thus, our updated sample results confirm earlier international momentum evidence in Griffin, Ji, and Martin (2003) and Rouwenhorst (1998).

2.3 An overview of the regression-based tests

The literature proposes a number of behavioral and rational hypotheses to explain momentum and these hypotheses are tested using various empirical proxies. We examine the robustness of prominent hypotheses in international data using the same proxies that are used with US data.

We use the following cross-sectional regression to test whether these hypotheses are robust internationally:

$$R_{i,t} = \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times X_{i,t-1} \times MOM_{i,t-1} + e_{i,t}, \qquad (1)$$

where $R_{i,t}$ is the return of stock i in month t, MOM is the momentum variable used in Table 1, and X is one of the explanatory variables for momentum that are described below. The $t-1$ subscript implies that all right-hand variables are computed at a one month lag. Note that for MOM, the computation stops at $t-2$ to skip the monthly reversal which might arise due to illiquidity or bid-ask issues; this is as per convention (Brennan, Chordia, and Subrahmanyam, 1998).[12] For convenience, we will often drop the time subscripts from these right-hand variables henceforth. We process the variables MOM and X on the right-hand through the following steps every month (i) we winsorize at the 0.5% and 99.5% levels, (ii) we country neutralize by subtracting the cross-sectional mean across all stocks in our sample from that country, and (iii) we standardize across the entire sample to have zero mean and unit standard deviation. The processed variables are used in the interaction terms. To minimize concerns about transaction costs and thin liquidity of some international markets, we include only non-microcap stocks. These are defined as stocks in the top 97% of the market capitalization of each country, as in Fama and French (2017).

As described in the introduction, we use proxies for momentum explanations, i.e., the X variables in Equation (1), from existing literature. Below, we describe how we construct them empirically. We discuss the underlying hypotheses in more detail within later sections.

- Book-to-market ratio (B/M): B/M is the ratio of the book value of equity to the market value. The Appendix Table A2 describes the precise formula we use to compute B/M.
- Turnover (Turn): We compute turnover as the number of shares traded in a month divided by shares outstanding as of the end of the previous month.

[12] While we use the most current values of the X variables (measured at month $t-1$) in Regression (1), measuring these instead at month $t-2$ makes no substantive difference to our conclusions.

- Residual Analysts (ResAnly): We compute residual analysts as in HLS (2000). Specifically, we cross-sectionally regress the log of one plus the number of analysts covering a stock on the log market capitalization of that stock each month, using the full sample. ResAnly is the residual from this regression.

- 52-week high (52wHi): We compute 52wHi for each stock each month as the ratio of the stock price at the end of the previous month to its highest price over the previous 12 months.

- Information discreteness (ID): Following DGW (2014), we define ID as follows:

$$\mathrm{ID}_{i,t-1} = \mathrm{sign}(\mathrm{PRET}) \times [\%\mathrm{neg} - \%\mathrm{pos}], \tag{2}$$

  where %pos and %neg are the percentage of daily returns that are positive and negative, respectively, and PRET is the past 11-month return.

- Cost of Goods Sold (COGS): COGS is the ratio of the cost of goods sold divided by the total assets as of the previous year.

- Return Volatility (RetVol): This is the standard deviation of daily returns over the previous 12 months for stocks with at least 100 days of return data. If volatility is greater than 300%, we suspect error in the data and set it to a missing value.

In Table 2, we present summary statistics for the explanatory variables by each region. We present statistics for the number of analysts, rather than ResAnly, as the former is more informative. We observe that mean turnover tends to be higher while COGS tends to be lower in emerging markets. These markets also tend to be more volatile. Other variables are not materially different across the three groups we consider.

3 Cross-Sectional Tests of Momentum Explanations

This section examines the robustness of the hypotheses proposed to explain momentum. As a starting point, we fit Equation (1) with only MOM as the independent variable, without any interaction variables. We estimate monthly cross-sectional regressions and we use the Fama and MacBeth (1973) approach to obtain the coefficients and standard errors. Column (1) in Panel A of Table 3 presents the results. The coefficients on MOM are 0.217, 0.277, and 0.135 for All ex-US, Developed ex-US, and Emerging markets, respectively. These coefficients are all statistically significant and confirm the evidence of momentum in Table 1.

3.1 Overconfidence

DHS (1998) present a behavioral model to explain momentum. Investors in DHS (1998) are subject to a self-attribution bias whereby they attribute profitable investments to their own skills and unprofitable ones to chance. As a result, investors become overconfident about the precision of their private signals over time and they overweight their private information when they value stocks. DHS (1998) show that this behavioral bias results in momentum due to a continuing overreaction.

DT (1999) hypothesize that the impact of overconfidence is likely to be stronger when it is harder to determine the intrinsic value of a firm. They argue that firms with bigger growth options relative to their assets in place are likely harder to value than firms with smaller growth options. Because the book value of a stock is the accounting value of assets-in-place, DT (1999) use B/M as an observable proxy for overconfidence. They report that momentum profits are bigger for growth firms than for value firms.

We fit Equation (1) with B/M as the interaction variable to test the robustness of the DT (1999) evidence outside the US. Column (2) of Panel A of Table 3 presents the results by regions. The interaction coefficients in All ex-US, Developed ex-US, and Emerging markets are -0.034, -0.037, and -0.043, respectively. However, the statistical significance of these interaction coefficients is small with t-statistics of only around -1.3. These results indicate that there is no reliable evidence that momentum is stronger for growth stocks than for value stocks. The coefficient on B/M is significantly positive in all regions. Therefore, consistent with the evidence in Fama and French (1992), B/M explains cross-sectional differences in returns, although it does not explain cross-sectional differences in momentum.[13]

LS (2000) document a positive relation between momentum and turnover. They note that many of the characteristics of high turnover stocks are similar to those of growth stocks, and those of low turnover stocks are similar to those of value stocks. LS (2000) suggest that turnover could also be a proxy for overconfidence, based on Odean (1998). We fit Equation (1) with turnover as the interaction variable to test the robustness of LS (2000) evidence. Column (3) in Panel A of Table 3 presents the regression estimates. The negative relation between returns and turnover is consistent with the evidence in Datar, Naik, and Radcliffe (1998). The interaction coefficient is -0.028 in All ex-US and -0.047 in Developed ex-US; the former is insignificant even at the 10% level, and the latter is significant at the 5% level. The interaction coefficient is economically small (close to zero) and statistically insignificant in Emerging markets. These estimates indicate that the positive relation between momentum and turnover that LS (2000) find is not robust outside the US. Turnover by itself is significantly negatively related to returns in All ex-US and Emerging markets (the respective coefficients are -0.232 and -0.398) but not in Developed ex-US markets (coefficient of -0.056).

3.2 Slow diffusion of information

HS (1999) present a model which assumes that investors process only a limited set of information. Investors in one cohort use only the price history to compute a stock's intrinsic value

[13] One could argue that book values are often noisy in international contexts. Even if this is so, we do find a baseline value effect in our setting; it is only the interaction of B/M with momentum that is not robust. B/M has also been used as an inverse proxy for real options availability in Berk, Green, and Naik (1999). As we do not find that the effect of B/M on momentum is robust, we sidestep the need to examine the specific rationale for this variable.

and in another use information about the stock's fundamentals but overlook its price history. Their model also assumes that information about fundamentals diffuses gradually and reaches different investors at different times. These assumptions differentiate the HS (1999) model from a rational expectations model where investors use all available information.

HLS (2000) empirically test the predictions of HS (1999). HLS (2000) hypothesize that the speed of information diffusion would be related to the extent of analyst coverage of a firm. Because more analysts cover large firms than small firms, HLS (2000) regress analyst coverage against firm size and use the residual number of analysts as the proxy for speed of information diffusion. They report that momentum is stronger for firms with smaller residual analyst coverage, which is consistent with the prediction of HS (1999).

We fit Equation (1) with ResAnly as the interaction variable and column (4) in Panel A of Table 3 presents the results. The interaction coefficients are *positive* and statistically significant in all three samples. Therefore, the international evidence does not support the slow diffusion of information hypothesis using residual analyst coverage as a proxy for the speed of information dissemination. Instead, in our empirical setting, higher residual analyst following tends to be associated with higher momentum.[14]

3.3 52-Week high

GH (2004) suggest that the anchoring bias could provide an explanation for momentum. They note that results in experimental economics research that are surveyed in Kahneman, Slovic, and Tversky (1982) find that subjects tend to use anchors to guide their assessment of unknown quantities. In the context of momentum, GH (2004) suggest that investor may use the 52-week high price for a stock as their anchor and, therefore, perceive stocks with prices near 52-week highs as expensive relative to stocks with prices farther away. Such a behavioral

[14] We also find that residual analyst coverage by itself is a strong positive predictor of returns in all samples. The coefficients on residual analyst coverage are 0.242, 0.267, and 0.194 for All ex-US, Developed ex-US, and Emerging markets, respectively, all strongly statistically significant. This finding requires more investigation in follow-up research.

bias would lead to an undervaluation of near 52-week high stocks and overvaluation of away from 52-week high ones. GH (2004) use the ratio of the price at the end of the previous month and the high price over the past 52 weeks as a measure of nearness to the 52-week high. They report that this measure explains a large portion of momentum in the U.S.

The effect of the 52-week high variable is directional. Specifically, high values of this ratio imply high returns (because investors should underreact to good news in this case), and low values should imply low returns (because investors should underreact to bad news in this case). Note that the correlation between nearness to 52-week high and MOM is likely to be large because past winners are likely to be closer to the 52-week high and past losers are likely to be farther away. So the key test of GH (2004) is whether their 52-week high variable explains cross-sectional variation in returns out-of-sample, and if so, whether it supplants momentum appreciably. As such, it is the coefficient of X, and how much its inclusion attenuates the effect of MOM, that are of greater interest here than the interaction term $MOM \times X$. However, the interaction might matter if investors' perception of overvaluation when the price is higher than the 52 week high dominates their perception of undervaluation when the reverse is true. We might expect this to be the case if high values of the ratios are more salient because investors are loss averse (Coval and Shumway, 2005). For this reason, and for consistency, we include the interaction term as well.

Column (5) in Panel A of Table 3 reports the regression results. We find that the interaction term is significant at the 5% level in the Developed ex-US region (coefficient of 0.058) but insignificant in the other regions. The coefficient on nearness to the 52-week high by itself is small and statistically insignificant in all regions. The coefficient on return momentum is barely changed in the presence of the 52-week high variable. Hence, we find very modest support for the 52-week high hypothesis as an explanation for momentum outside the U.S.

3.4 The frog-in-the-pan hypothesis

DGW (2014) propose the frog-in-the-pan hypothesis (FIP) as an explanation for momentum. They hypothesize that investors are inattentive to information that arrives continuously in small amounts, which results in underreaction to such information and results in momentum profits. FIP predicts that momentum would be bigger for stocks with continuous information flow than for stocks with discrete information flow.

DGW (2014) observe that ID as defined in Equation (2) should be bigger when the information flow is discrete. Intuitively, returns are equally likely to be positive or negative when information flow is continuous and a large difference in the frequency of positive and negative returns suggests more concentrated information flow of the corresponding sign. DGW (2014) find bigger momentum for stocks with more continuous information.

We fit Equation (1) with ID as the interaction variable, and column (6) in Panel A of Table 3 reports the regression estimates. The interaction coefficients are -0.161 in All ex-US, -0.125 in Developed ex-US, and -0.188 in Emerging markets. All these coefficients are strongly statistically significant. Further, the coefficient of momentum attenuates by at least 20% in each of the three cases when the interaction of MOM with ID is included. For example, for All ex-US markets, coefficient on MOM is 0.217 in univariate regression but only 0.165 in column (6) with ID and interaction term as additional variables, an attenuation of 24%. Therefore, the FIP hypotheses is robust internationally.

3.5 Time-varying growth options

The explanations considered so far suggest that behavioral biases explain momentum. In contrast, Johnson (2002) and SS (2007) present models where the risk premium varies through time, and momentum is a compensation for the bigger risk exposures that past winners face. Intuitively, SS (2007) consider firms with safe assets and growth options, and the firm value is a sum of these two parts. Firms become winners when the value of their growth options increase and become a bigger fraction of their value. Because the firms as whole now becomes riskier, they command a bigger risk premium. In the SS (2007) model, the relative value of growth options is bigger for low cost of goods sold (COGS) firms than for high COGS firms because their operational leverage is bigger. Therefore, the growth options hypothesis predicts that momentum would be bigger for low COGS firms and SS (2007) find empirical support this hypothesis.

We test robustness of the real options hypothesis in international markets using COGS as the interaction variable in Equation (1). Column (7) in Panel A of Table 3 presents the regression results. We find that the interaction coefficient is negative in All ex-US and Developed ex-US, and positive in Emerging markets. However, all these coefficients are statistically insignificantly different from zero. The insignificant interactions in all regions suggest that the growth option hypothesis is not robust internationally.

SS (2007) also suggest that revenue volatility is also a proxy for growth options. Because we need a reasonable number of data quarters to compute revenue volatility, the sample size of firms for which we are able to compute this variable is too small for any meaningful power of the tests. Instead we use daily return volatility computed over the last 12 months. SS (2007) note that in their model, momentum also increases in return volatility. This latter variable is also used by Zhang (2006) in a different context: as a proxy for uncertainty, which he argues increases the level of underreaction. The results are in column (8) in Panel A of Table 3. We find that the interaction coefficients in All ex-US, Developed ex-US, and Emerging markets are -0.029, -0.069, and -0.077, respectively. The negative coefficients for Developed ex-US and Emerging markets are both statistically significantly different from zero. These results suggest that return volatility interacts negatively with momentum. Therefore, we find no evidence that such volatility is related positively to the strength of the momentum effect.

4 Robustness Checks

In this section, we conduct ancillary analysis to check the robustness of our results to risk adjustment, additional cross-sectional controls, an alternative definition of momentum, and different sub-periods.

4.1 Risk-adjusted returns

Our baseline FM regressions do not include any controls for risk. Even though risk-based explanations for momentum are elusive, we adjust returns for risk and check their relation to the variables of interest. Specifically, we use the Brennan, Chordia, and Subrahmanyam (1998) procedure for the risk-adjustment. We compute the month t factor loadings for each stock with the following time-series regression:

$$R_{i,s} = a + b'_{i,t} f_s + e_{i,s}, \quad \text{for } s = t - 36 \text{ to } t - 1, \tag{3}$$

where f_s is the month s realization of the five factors in the Fama and French (2017) five-factor model.[15] We then compute risk-adjusted returns $R_{i,t} - \hat{b}'_{i,t} f_t$ where $\hat{b}_{i,t}$ is the factor sensitivity estimate from the time-series Equation (3) and use these risk-adjusted returns in the following regression:

$$R_{i,t} - \hat{b}'_{i,t} f_t = \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times X_{i,t-1} \times MOM_{i,t-1} + e_{i,t}, \tag{4}$$

Panel B of Table 3 reports the regression estimates of Equation (4). The results in Panel B are similar to the corresponding results in Panel A with a few exceptions. The most prominent difference is that the interaction coefficient with B/M becomes negative

[15]We use separate factors for developed and emerging markets and use the factor model for the stock corresponding to its region. These factors are obtained from Ken French's website (`http://tinyurl.com/bdfn35ze`). An alternative set of factors are provided in Hou, Xue, and Zhang (2015). These latter factors are not available at the international level (see also Novy-Marx, 2015), so we use the ones in Fama and French (2017) instead.

and statistically significant for All ex-US and Developed ex-US markets; this evidence is consistent with DT (1999). However, because B/M-based HML is one of the risk factors in the left-hand-side of Equation (4), the correlation between the measurement error in the HML loading estimate and the characteristic could also spuriously contribute to the coefficient. All other conclusions from Section 3 are largely unchanged. In particular, ID continues to significantly explain cross-sectional differences in momentum as the interaction coefficient on ID in column (6) of Panel B of Table 3 is negative and statistically significant for all regions.

Chordia and Shivakumar (2006) argue that a factor based on earnings surprises (PMN) can capture momentum profits. In Table A3 within the internet appendix, we add PMN computed from international data as an additional factor when risk-adjusting returns. Due to patchy availability of analysts-related forecasts and quarterly earnings data, we compute standardized unexpected earnings as simply the change in annual earnings scaled by the most recent market price. This approach is similar to that of Livnat and Mendenhall (2006). We find that the results are virtually unchanged after adding this factor, suggesting that our measure of earnings momentum is not related to return momentum in our out-of-sample international context.[16]

4.2 Additional controls and alternative momentum returns

We now modify Equation (1) to include additional controls:

$$R_{i,t} = \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times X_{i,t-1} \times MOM_{i,t-1} + \gamma_{4,t} \times Z_{i,t-1} + e_{i,t}, \quad (5)$$

where Z's are new controls. These include asset growth (Cooper, Gulen, and Schill, 2008), gross profitability (Novy-Marx, 2013), one month reversal (Jegadeesh, 1990), book/market,

[16] Avramov, Chordia, Jostova, and Philipov (2007) relate distress risk to momentum returns. Because we do not have international bond ratings data, we use the annual bankruptcy predictor developed by Campbell, Hilscher, and Szilagyi (2008) (which in turn, is based on Altman, 1968) as an additional X variable and find that it plays no role in explaining momentum. Results are available upon request.

and size (market capitalization), as of a given month.[17] We process the additional Z variables on the right-hand following the same steps as those for MOM and X described after Equation (1).

Panel C of Table 3 reports results with additional controls. We do not tabulate the coefficients on control variables for brevity. Compared to Panel A, we lose about one-third of the stocks for which we do not have sufficient accounting data to calculate control variables. Nevertheless, results in Panel C are similar to those in Panel A. Once again the robust result that emerges is that of the ID variable, and as in Panel B, growth stocks have more momentum in the All ex-US group, whereas B/M does not interact significantly with momentum in the other two groups.

In the Appendix Table A4, we present the analog of Table 3 (Panel C) using lagged returns from the second to seventh month, MOM', to measure momentum instead of MOM. We find that the results are materially unchanged (other panels of Table 3 yield similar results). The robust picture that emerges is that ID is a consistent and significant explanatory variable for momentum.

4.3 Subsample results

Our sample period includes the tech bubble, the 2008 financial crisis, and the beginnings of the COVID crisis. To assess whether our results are affected by such outlier events, we examine the results in two equal subsamples, 1993-2006, and 2007-2020. We present the results for Equation (5) in Table 4. Panel A reports results for 1993 to 2006 while Panel B reports the results for 2007 to 2020.

Since this exercise cuts the number of time-series observations for the FM coefficients in half, we expect power to be lower in regressions reported in Table 4. Nevertheless, in every one of the six cases (three regions and two subsamples), the coefficient of ID interacted

[17] We include B/M only in the specifications that do not already include it as an X variable.

with momentum is negative and significant, pointing to the robustness of this explanatory variable.

Among other results for the subsamples, book/market and turnover tend to be positively and negatively associated with future returns in both subperiods, for emerging markets, but this relation is less robust for the other regions. The interaction of B/M with momentum is statistically significant in the first half of the subsample but not so in the second half. Patterns in the other interaction coefficients are not robust to various regions or over time.

5 Multivariate Analysis with Penalized Regressions

Our tests so far analyze the X variables one at a time. We next check the marginal explanatory power of these variables and their interactions to explain momentum in a kitchen sink regression. Since the dangers of overfitting loom large, we now employ penalized regressions, in addition to multivariate OLS regressions.[18] Our regression setup is

$$R_{i,t} = \gamma_0 + \gamma_1 \times MOM_{i,t-1} + \gamma_2' \times X_{i,t-1} + \gamma_3' \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t}, \tag{6}$$

where X is now the vector of all standardized explanatory variables. We use Lasso and Elastic net to run Equation (6). These regressions take the general form of minimizing the following loss function

$$\mathcal{L}(\gamma, \lambda, \rho) = \sum_{i,t}(y_{i,t} - \gamma' x_{i,t-1})^2 + \lambda(1-\rho)\sum_j \gamma_j + 0.5\lambda\rho\sum_j \gamma_j^2, \tag{7}$$

where λ and ρ are additional hyperparameters. $\rho = 0$ corresponds to Lasso and $\rho = 1$ corresponds to ridge regressions (see Hastie, Tibshirani, and Friedman, 2009). As Lasso imposes a penalty related to the absolute values of the coefficients, it tends to completely

[18]See Gu, Kelly, and Xiu (2020) and Han, He, Rapach, and Zhou (2019) for other applications of these techniques to the cross-section of stock returns.

eliminate some variables from the model, allowing for sparse selection of variables. On the other hand, ridge regression shrinks coefficients towards zero without necessarily setting them to zero. We choose Lasso ($\rho = 0$) and Elastic net ($\rho = 0.5$) for our specifications. These two techniques are the simplest and most parsimonious amongst commonly-used machine learning techniques. We choose the hyperparameter λ via ten-fold cross-validation. To ensure comparability with Lasso and Elastic net, in using OLS, we use panel regressions instead of FM. The model is fit over the training sample that is the first half of the sample period, viz. 1993 to 2006.

In Table 5, we present the coefficients using standard OLS, Lasso, and elastic nets. We find that the interaction coefficient of ID with momentum barely shrinks across the three procedures. The biggest shrinkage occurs across the interaction coefficient with residual analyst coverage. The interactive coefficient with the 52-week high also shrinks and is not included by Lasso in one case. The interactive coefficient with RetVol is not included by Lasso in All ex-US. The coefficients of B/M and Turn generally do not shrink appreciably across the three procedures.

Table A5 in the online appendix provides the FM coefficients *during the training period* for all of the predictive variables and their interactions with MOM. The appendix confirms that ID is the only robust interaction with momentum (it is significant across all three regions). The interaction of MOM with BM is not significant for emerging markets, but is significant for the other two regions.

We next analyze the predictive power of the three procedures in an out-of-sample (OOS) setting. Our forecasting period is the second half of the sample period, viz. 2007 to 2020. We do not refit OLS, LASSO, or Elastic net on a rolling or an expanding window basis. Therefore, the OOS period is a true testing period. For each of the procedures and for each region, we obtain a forecast of the returns as $\hat{R}_{i,t}$ using coefficients from Table 5 and the most recent $X_{i,t-1}$. Following Gu, Kelly, and Xiu (2020), we calculate the OOS-R^2 as

$$\text{OOS-}R^2 = 1 - \frac{\sum_{(i,t)} \left(R_{i,t} - \hat{R}_{i,t}\right)^2}{\sum_{(i,t)} R_{i,t}^2}, \tag{8}$$

where we take forecast errors over all stocks over the entire OOS period in the numerator and raw (not demeaned) returns as the denominator. We also present the mean-squared error (MSE) and the mean absolute error (MAE) for the three samples. Table 6 presents the results. We find that OOS-R^2's, MSE, and MAE are similar across the three procedures. Using the Diebold and Mariano (1995) test, we are unable to reject the hypothesis that the OOS-R^2's are different from each other.

Overall, the conclusion from Section 3 is that ID wins the horse race conducted via penalized multivariate regressions, which reinforces the earlier conclusions from univariate analyses. The significance of ID interacted with momentum persists when other explanatory variables for momentum are included, and the coefficient on this interactive variable barely changes when we use $\mathcal{L}_1$ and $\mathcal{L}_2$ penalties via Lasso and Elastic net.

6 Portfolio Returns

To investigate the economic significance of the cross-sectional explanatory power of the relation between momentum and the variables that we examine, we calculate profits for double-sorted portfolios. These sorts also lend perspective to the issue of how the explanatory variables affect the profitability of momentum strategies. Specifically, we sort by the explanatory variable X and by the momentum variable MOM into terciles. The sorting is done every month and we hold the portfolios for one month.

6.1 Hedge Portfolio Returns

We calculate the hedge portfolio return, winners minus losers (WML) for each tercile of X, and present the results in Table 7. The table then reports the difference in WML returns, denoted ΔWML, from high X tercile to low X tercile. Because WML is the momentum profit in each of the X terciles, ΔWML is the incremental effect of the X variable on momentum profits. We perform independent sorts as well as sequential sorts (where we sort first on X and then on momentum return). While both kinds of sorts examine the relation between future returns and momentum returns controlling for the X variable, the method of controlling is different. Independent sorts effectively compare an unconditional relation between future returns and momentum returns so that ΔWML is the difference in unconditional effect of momentum returns in high and low X terciles. This method is close in spirit to our earlier FM regressions. Sequential sorts compare the conditional relation between future returns and momentum returns, controlling for values of X and, thus, are of equal interest.

Panel A of Table 7 reports the results for independent sorts. For ID, ΔWML is statistically significantly negative at the 10% level for Developed ex-US and at the 5% level for All ex-US and Emerging markets. For example, in Emerging markets ΔWML equals -8.88%, which indicates that the momentum profit in the low ID tercile is 8.88% higher than that for the high ID tercile. When B/M is the X variable, ΔWML is statistically significantly negative in All ex-US, indicating that momentum is more profitable for growth stocks than for value stocks. However, ΔWML is statistically insignificant for the other regions, and when X represents any of the other variables.

Panel B of Table 7 presents the results for sequential sorts. ΔWML is again statistically insignificant when any of the variables other than ID are used as X. For X = ID, ΔWML is statistically significant in all regions at the 5% level. The magnitude of ΔWML is also economically large, at around 6% for All ex-US and Developed ex-US markets and at around 9% for Emerging markets, showing that high values of ID exert a large influence on attenuating the importance of momentum for explaining future returns.[19]

Our surmise is that again, apart from the significant average returns for the ID variable, none of the other variables reliably explain cross-sectional differences in momentum. Thus, these portfolio results are consistent with the regression results.[20]

6.2 Do momentum profits reverse in the long run?

JT (1993) present evidence that momentum profits tend to reverse in the long run, although Jegadeesh and Titman (2001) find that the reversals are weak in recent years. Several models including HS (1999) and DHS (1998) predict long-term reversals following short-term momentum. However, pure underreaction explanations for momentum, such as those of DGW (2014), do not imply long-term reversal. Hence, to further pin down causes of momentum, it is of importance to explore whether the evidence of long-term reversal in momentum profits is robust in international data.

There is another reason to examine reversals in momentum portfolios. Specifically, the flows hypothesis of Vayanos and Woolley (2013) and Lou (2012) suggests that momentum profits are due to uninformed fund flows. Specifically, these papers argue that momentum obtains due to trend-chasing behavior of institutions such as mutual funds.[21] Note that if momentum arises due to fund flows that are largely uninformed, the resulting price pressure exerted by fund flows should reverse in the long run.[22] Thus, examining whether momentum

[19]As we note in Section 3.3, in the case of the 52-week high variable, it is the X variable that is of at least equal interest relative to the interaction of momentum with X. We have verified that extreme sorts on this X variable alone do not yield a significant return spread.

[20]Bandarchuk and Hilscher (2013) argue that sequential sorts on characteristics, and then on momentum, simply sort on extreme realizations of past returns, and therefore have challenges in isolating the effect of characteristics on momentum. Their observation applies to sequential sorts, but not to independent sorts or to regressions. We get similar results with all three methods, so that our overall conclusions are not subject to the bias that they discuss.

[21]This extrapolative behavior could, in turn, could reflect individual investors' biases; see Barberis, Greenwood, Jin, and Shleifer (2018).

[22]On the other hand, if informed institutions play a role in explaining momentum, then this explanation should be subsumed by another already-considered X variable that represents underreaction to news, such

profits reverse can potentially also shed light on the flows hypothesis.

Motivated by the above observations, we examine the long-run performance of momentum portfolios. Table A6 in the internet appendix presents momentum profits over horizons ranging from one month to five years. They do suggest evidence of momentum at the usual (3-12 month) horizons but there is no evidence that these momentum profits reverse in the long run. Specifically, all the coefficients beyond the 12-month horizon are insignificant. Hence the international (out of sample) evidence is more supportive of pure underreaction hypotheses such as that of DGW (2014), as opposed to models suggesting that momentum is due to overreaction to past information, and is therefore followed by reversals.[23]

6.3 Cross-sectional tests: Overall conclusion

The analysis above considers a comprehensive set of empirical findings that propose rationales for momentum. Among the many proxies for momentum explanations considered, the most robust proxy one is the one used by DGW (2014) for testing the frog-in-the-pan hypothesis, which indicates that due to limited attention, investors underreact to small bits of information flowing in about the firm, but react in a timely manner to large discrete pieces of information. Thus, a variable related to the discreteness of news arrival is inversely related to the strength of the momentum effect. We do not find robust evidence in support of the other hypotheses proposed to explain momentum.

as ID or analyst following.

[23] We do not have high quality fund flows data at the monthly horizon within our international context, so we cannot test the fund flows hypothesis directly. In Table A7 within the internet appendix, we compute quarterly changes in institutional holdings as an additional X variable using *FactSet* (we also present the coefficients for other X variables and their interactions for comparison). The coverage from *FactSet* is not comprehensive, leading to a substantial reduction in sample size. We find that changes in percentage holdings are not significantly related to momentum, and the significance of ID continues to prevail in this smaller sample.

7 The Time-Series of Momentum Profits

This section examines the robustness of the evidence in the US that finds predictable variations in momentum profits over time. We focus on two major findings: that aggregate momentum profits depend on (i) the sign and (ii) the volatility of market states.

7.1 Up and down markets

CGH (2004) show that momentum is stronger following up markets than following down markets in the US. They attribute this finding to the notion that overconfidence is higher in rising markets. The idea is that investors are net long in markets and are likely to have received a sequence of positive signals confirming their long positions in up markets, thus building their overconfidence.

We investigate if CGH (2004) results are robust out of sample. We examine momentum profits in up and down markets internationally using the following regression:

$$\mathrm{WML}_t = \gamma_1 \times \mathrm{UP}_{t-1} + \gamma_2 \times \mathrm{DOWN}_{t-1} + e_t, \tag{9}$$

where UP is a dummy variable that equals one for an up market and zero otherwise. DOWN is defined analogously for down markets. Following CGH (2004), UP equals unity if the market return over the previous 36 months is positive and DOWN equals unity if this return is negative. We use the MSCI All-Country ex-US, World ex-US, and Emerging total return indices as the market return proxies for All ex-US, Developed ex-US, and Emerging markets, respectively. The up market and down market coefficients represent momentum profits during the two states in Equation (9).

Table 8 presents the regression estimates. We find that momentum profits in up markets are significantly positive at 14.16% but marginally negative at −1.94% during down markets for the All ex-US region. The momentum profits in all the other regions are also significantly positive in up markets and marginally negative in down markets. Thus, the CGH (2004) results are robust internationally.

We next examine the effect of market states on winners and losers separately. While CGH (2004) do not make any predictions on this issue, nonetheless, to gain additional empirical insight, we replace the dependent variable WML in Equation (9) with returns on winner and loser portfolios separately. We report the regression estimates within additional columns in Table 8. In All ex-US, the difference between returns in up and down markets is −27.01% for losers and −10.92% for winners. The difference is significant for losers but not for winners. The results are similar in the Developed ex-US region as well. In Emerging markets, the return difference for losers is −10.72% compared with 2.17% for winners. Although the point estimate of the difference is bigger in magnitude for losers than for winners, they are both insignificant. Overall, the momentum strategy is profitable in up markets but not in down markets, confirming the empirical findings of CGH (2004). In a further finding, this phenomenon is stronger for losers.

7.2 High and low volatility

WX (2015) find that momentum profits are bigger when market volatility is low than when it is high. They find that the relation between momentum profits and market volatility is mainly due to the asymmetric performance of loser stocks. They argue that losers in high volatility states because more likely become distressed in such states. Hence, investors "over-sell" losers, and the subsequent price recovery of these stocks results in low momentum profits.[24]

We examine the relation between momentum and market volatility with the regression

$$\mathrm{WML}_t = \gamma_1 \times \mathrm{HIVOL}_{t-1} + \gamma_2 \times \mathrm{LOVOL}_{t-1} + e_t, \tag{10}$$

[24]Stivers and Sun (2010) find that cross-sectional return dispersion explains the time-series of momentum profits; however, WX (2015) find that accounting for market volatility fully captures this effect.

where HIVOL is a dummy variable that equals one in a high volatility market and zero otherwise and LOVOL is an analogously defined variable for a low volatility market. We classify the market as being in a high volatility state if the standard deviation of daily market returns over the previous 12 months is greater than that the previous 36 months and in a low volatility state otherwise. We classify the state of market volatility relative to the past three-year volatility because unreported analysis suggests a secular decline in market volatility across regions during our sample period. We compute market standard deviation for each region using daily return data for the corresponding MSCI index.

We fit Equation (10) separately with WML_t, winner returns and loser returns as dependent variables and present the results in Table 8 (right panel). Momentum profits are 17.69%, 15.75%, and 15.69% during low volatility periods and 1.12%, 2.99%, and −0.49% during high volatility periods in All ex-US, in Developed ex-US, and in Emerging markets, respectively. These profits during low volatility periods are significant in all regions but insignificant during high volatility periods.

The difference in returns across the two states for losers are 18.59% in All ex-US, 12.74% in Developed ex-US, and 11.96% in Emerging markets, compared with 2.01%, −0.02%, and −4.22% for winners. Although the return difference for losers is only significant in All ex-US, the point estimates of the differences are bigger for losers than winners in all regions. Therefore, the difference between the performance of momentum in high and low volatility states are also driven largely by the differential performance of losers. Overall, our finding confirms the empirical conclusion of WX (2015) that momentum does poorly when market volatility is high, and that this phenomenon is driven by losers.[25]

In Table 9, we present the regressions of Table 8 for the two halves of the full sample. In general, the results are robust. There is a decline in significance related to the volatility states result for Developed ex-US in the second half, but significance remains in all other

[25]Daniel and Moskowitz (2016) also find that momentum is negatively related to volatility in continental Europe and Japan, but they do not consider emerging markets.

cases. Thus, the two central momentum results on market states based on direction and volatility are generally robust across regions and across time.

7.3 Interpretation of the time-series results

Our cross-sectional results indicate support for the DGW (2014) FIP hypothesis while the results on up and down markets support CGH (2004), who suggest that up markets represent greater overconfidence, thus leading to more momentum, based on the arguments of DHS (1998). However, the overconfidence-based explanation implies a long-term reversal in momentum profits, for which we do not find empirical support (see Table A6). This leaves open a puzzle that needs an attempt at reconciliation.

Accordingly, we propose an alternative to the CGH (2004) interpretation of their finding, which is that the diffusion of news might be slower in up versus down markets. In other words, the nature of the ID variable (capturing FIP) might differ across market states. To address this issue, we compute the value-weighted ID variable at the country level, and then compute how it differs across up and down markets on average. The results appear in Panel A of Table A8 within the internet appendix. The evidence clearly indicates that ID is lower in up markets, which means news flows are more lumpy in down markets and news diffuses more gradually and slowly in up markets.

It is possible that the WX (2015) results are also influenced by the differential behavior of ID across high and low volatility states. Thus, Panel B of Table A8 computes how ID differs across these states. Again the point estimate of ID is lower in low volatility markets, which means news diffuses more gradually in such markets. The difference in ID across volatility states is significant for Developed ex-US and All ex-US, but not for emerging markets.

We note that the explanation given by WX (2015) for their finding is that losers are "oversold" in volatile markets, and the recovery of losers implies that the losing leg of momentum fails in more volatile markets. In unreported analyses, we also find that ID is lower for winners relative to losers in up-markets and low-volatility states, which supports the finding that high momentum profits in low volatility periods emanate from winners.[26] It is possible that both the ID-based explanation as well as that of WX (2015) operate simultaneously in explaining why momentum profits are higher in low volatility states.

In sum, the cross-sectional evidence indicates that slow and gradual diffusion of news (as proxied inversely by ID) is the cause of cross-sectional momentum. Further, our time-series findings confirm the results of CGH (2004) that momentum is higher in up-markets and of WX (2015) that momentum is also higher in less volatile periods. Since news also diffuses more slowly in these market states, the FIP hypothesis of DGW (2014) best explains cross-sectional variations in momentum and is consistent with time-series variations in momentum. We leave the issue of *why* ID differs across market states for future research.

8 Conclusion

Momentum represents a simple violation of weak-form efficiency, and is termed the "premier" anomaly in equity returns by Fama and French (2008). Why do markets permit such memory in stock prices? While copious research has been devoted to this key question, there is as yet no prevailing consensus on the issue. In this paper, we conduct a horse race of proxies for momentum explanations by considering an out-of-sample analysis using non-US data. For this analysis, we separately consider developed and emerging markets. We consider several previously-considered proxies for rationales based on overconfidence, slow diffusion of news, the anchoring bias, information discreteness, and risk premia that vary with past returns. We minimize subjectivity on our part by only using empirical proxies developed in the original papers that consider momentum explanations. Our goal is not to critique the many insights on momentum that have accumulated, but to examine the explanations using the same methodology and a common dataset.

[26]Further, upon including market return states and volatility states within the same specification, we find that these are two distinct phenomena, so that one does not subsume the other.

Our horse race has a clear winner out-of-sample. The most robust rationale for cross-sectional momentum in our international setting is the frog-in-the-pan hypothesis of DGW (2014). The "discrete news" variable proposed by these authors is by far the most significant in explaining momentum. As such, we can propose that the explanation for momentum that best fits the available international evidence is that investors underreact to news that dribbles out slowly, as opposed to news releases in discrete chunks. Momentum profits show little evidence of reversal in the long-run, supporting this underreaction explanation.

In other analysis, we find robust international evidence that momentum is stronger in up-markets and in low volatility markets, as documented by CGH (2004) and WX (2015) for the US. Using the proxy used by DGW (2014), we also find that news diffuses less discretely during periods of low volatility and in up markets, which is also consistent with the FIP hypothesis that DGW (2014) propose to explain cross-sectional differences in momentum.

References

Altman, Edward, 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589–609.

Antoniou, Constantinos, John A. Doukas, and Avanidhar Subrahmanyam, 2013, Cognitive dissonance, sentiment, and momentum, *Journal of Financial and Quantitative Analysis* 48, 245–275.

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

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2007, Momentum and credit rating, *Journal of Finance* 62, 2503–2520.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Bandarchuk, Pavel, and Jens Hilscher, 2013, Sources of momentum profits: Evidence on the irrelevance of characteristics, *Review of Finance* 17, 809–845.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2018, Extrapolation and bubbles, *Journal of Financial Economics* 129, 203–227.

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

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Brennan, Michael J., Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns, *Journal of Financial Economics* 49, 345–373.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899–2939.

Chordia, Tarun, and Lakshmanan Shivakumar, 2002, Momentum, business cycle, and time-varying expected returns, *Journal of Finance* 57, 985–1019.

Chordia, Tarun, and Lakshmanan Shivakumar, 2006, Earnings and price momentum, *Journal of Financial Economics* 80, 627–656.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of returns, *Journal of Finance* 63, 1609–1651.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

Coval, Joshua D., and Tyler Shumway, 2005, Do behavioral biases affect prices?, *The Journal of Finance* 60, 1–34.

Da, Zhi, Umit G. Gurun, and Mitch Warachka, 2014, Frog in the pan: Continuous information and momentum, *Review of Financial Studies* 27, 2171–2218.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and over-reactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221–247.

Daniel, Kent, and Sheridan Titman, 1999, Market efficiency in an irrational world, *Financial Analysts Journal* 55, 28–40.

Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

Diebold, Francis X., and Roberto S. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 13, 253–265.

Ehsani, Sina, and Juhani T. Linnainmaa, 2021, Factor momentum and the momentum factor, *Journal of Finance,* forthcoming.

Falck, Antoine, Adam Rej, and David Thesmar, 2020, Is factor momentum more than stock momentum? *Available at SSRN* 3688983.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

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* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653–1678.

Fama, Eugene F., and Kenneth R. French, 2017, International tests of a five-factor asset pricing model, *Journal of Financial Economics* 123, 441–463.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

George, Thomas J., and Chuan-Yang Hwang, 2004, The 52-week high and momentum investing, *Journal of Finance* 59, 2145–2176.

Goyal, Amit, and Narasimhan Jegadeesh, 2018, Cross-sectional and time-series tests of return predictability: What is the difference? *Review of Financial Studies* 31, 1784–1824.

Goyal, Amit, and Sunil Wahal, 2015, Is momentum an echo? *Journal of Financial and Quantitative Analysis* 50, 1237–1267.

Griffin, John M., Xiuqing Ji, and J. Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Griffin, John M., Patrick J. Kelly, and Federico Nardari, 2010, Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* 23, 3225–3277.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *Review of Financial Studies* 33, 2223–2273.

Han, Yufeng, Ai He, David Rapach, and Guofu Zhou, 2019, Expected stock returns and firm characteristics: E-lasso, assessment, and implications, *Available at SSRN #3185335* .

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, 2009, *The elements of statistical learning* (Springer, New York, NY).

Hirshleifer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337–386.

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* 55, 265–295.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hong, Harrison, Jeremy C. Stein, and Jialin Yu, 2007, Simple forecasts and paradigm shifts, *Journal of Finance* 62, 1207–1242.

Hou, Kewei, G. Andrew Karolyi, and Bong-Chan Kho, 2011, What factors drive global stock returns? *Review of Financial Studies* 24, 2527–2574.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Ince, Ozgur S., and R. Burt Porter, 2006, Individual equity return data from thomson datastream: Handle with care! *Journal of Financial Research* 29, 463–479.

Jacobs, Heiko, and Sebastian Müller, 2020, Anomalies across the globe: Once public, no longer existent? *Journal of Financial Economics* 135, 213–230.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 36, 699–720.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2011, Momentum, *Annual Review of Financial Economics* 3, 493–509.

Johnson, Timothy C., 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Kahneman, Daniel, Paul Slovic, and Amos Tversky, 1982, *Judgment under uncertainty: Heuristics and biases* (Cambridge University Press).

Kelly, Bryan T., Tobias J. Moskowitz, and Seth Pruitt, 2021, Understanding momentum and reversal, *Journal of Financial Economics* 140, 726–743.

Lee, Charles M. C., and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017–2069.

Lee, Kuan-Hui, 2011, The world price of liquidity risk, *Journal of Financial Economics* 99, 136–161.

Livnat, Joshua, and Richard R. Mendenhall, 2006, Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts, *Journal of Accounting Research* 44, 177–205.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

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

Müller, Birgit, and Sebastian Müller, 2020, Cross-country composite momentum, *Available at SSRN #3240609* .

Novy-Marx, Robert, 2012, Is momentum really momentum? *Journal of Financial Economics* 103, 429–453.

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

Novy-Marx, Robert, 2015, How can a q-theoretic model price momentum? *National Bureau of Economic Research Working Paper* .

Odean, Terrance, 1998, Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887–1934.

Rouwenhorst, K. Geert, 1998, International momentum strategies, *Journal of Finance* 53, 267–284.

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

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.

Subrahmanyam, Avanidhar, 2018, Equity market momentum: A synthesis of the literature and suggestions for future work, *Pacific-Basin Finance Journal* 51, 291–296.

Vayanos, Dimitri, and Paul Woolley, 2013, An institutional theory of momentum and reversal, *Review of Financial Studies* 26, 1087–1145.

Wang, Kevin Q., and Jianguo Xu, 2015, Market volatility and momentum, *Journal of Empirical Finance* 30, 79–91.

Zhang, X. Frank, 2006, Information uncertainty and stock returns, *Journal of Finance* 61, 105–137.

Table 1: Momentum outside the US

This table presents momentum profits outside the US. We form momentum portfolios based on stock returns over the previous 12 months, excluding the previous month. Specifically, for stock i the momentum variable for month t is the return from month $t-12$ to $t-2$. We country neutralize the momentum variable by subtracting its cross-sectional mean across all stocks in our sample from that country. We then rank stocks based on country-neutralized returns and assign each stock to one of ten momentum deciles. The WML hedge portfolio is long the value-weighted portfolio of stocks in the winner decile short the corresponding loser decile. The table reports annualized returns from the WML portfolio in percent. The sample excludes microcap stocks (stocks not in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	All ex-US	Developed ex-US	Emerging
Mean	10.64	10.23	8.93
Median	13.86	12.60	9.16
StdDev	65.24	73.32	61.48
Skewness	−0.64	−0.74	−0.44
Kurtosis	9.22	9.81	4.94
Minimum	−405.60	−468.59	−270.77
Maximum	258.58	295.74	217.58

Table 2: Summary statistics

This table presents summary statistics for variables that have been proposed to explain momentum profits. The variables are defined in Section 2.

	B/M	Turn	Anly	52wHi	ID	COGS	RetVol
				All ex-US			
5th percentile	0.105	0.001	0.000	0.500	−0.147	0.052	0.070
Median	0.523	0.04	0.231	0.895	−0.044	0.532	0.302
Mean	0.708	0.095	3.415	0.844	−0.044	0.648	0.305
95th percentile	1.900	0.376	16.595	1.000	0.056	1.695	0.645
StdDev	0.665	0.156	5.875	0.174	0.062	0.521	0.198
# stocks	7,022	7,142	10,951	10,818	10,951	6,324	10,950
				Developed ex-US			
5th percentile	0.111	0.001	0.000	0.546	−0.138	0.049	0.070
Median	0.586	0.033	0.190	0.950	−0.043	0.567	0.228
Mean	0.753	0.053	3.615	0.876	−0.044	0.678	0.260
95th percentile	1.922	0.168	17.625	1.000	0.05	1.765	0.595
StdDev	0.660	0.069	6.195	0.160	0.058	0.539	0.186
# stocks	3,874	3,856	6,982	6,887	6,982	3,462	6,982
				Emerging			
5th percentile	0.103	0.001	0.000	0.449	−0.160	0.059	0.084
Median	0.480	0.065	0.653	0.795	−0.044	0.458	0.400
Mean	0.728	0.153	2.893	0.773	−0.046	0.580	0.406
95th percentile	2.134	0.580	13.56	1.000	0.066	1.534	0.721
StdDev	0.806	0.241	4.943	0.180	0.069	0.480	0.191
# stocks	3,293	3,484	4,148	4,110	4,148	3,080	4,148

Table 3: Fama-MacBeth regressions of future returns on past momentum return and explanatory variables

This table presents the results of Fama-MacBeth cross-sectional regressions. Panel A runs the regression:

$$R_{i,t} = \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t},$$

where $MOM_{i,t-1}$ is the return from month $t-12$ to $t-2$ and X represents each of variables listed in the top row of Table 2. We process the variables MOM and X on the right-hand through the following steps every month (i) we winsorize at the 0.5% and 99.5% levels, (ii) we country neutralize by subtracting the cross-sectional mean across all stocks in our sample from that country, and (iii) we standardize across the entire sample to have zero mean and unit standard deviation. The processed variables are used in the interaction terms. Panel B runs the regression:

$$R_{i,t} - \beta_i' F_t = \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t},$$

where the betas on the left-hand-side are calculated using the full sample for each stock from a five-factor Fama and French (2017) model. We use separate factors for developed and emerging markets and use the factor model for the stock corresponding to its region. Panel C runs the regression:

$$\begin{aligned} R_{i,t} = {} & \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} \\ & + \gamma_{4,t} \times Size_{i,t-1} + \gamma_{5,t} \times MOM_{i,t-1} \times Size_{i,t-1} + \gamma_{6,t} \times (B/M)_{i,t-1} + \gamma_{7,t} \times (GP/AT)_{i,t} \\ & + \gamma_{8,t} \times ATG_{i,t-1} + \gamma_{9,t} \times R_{i,t-1} + e_{i,t}, \end{aligned}$$

with additional controls on the right-hand-side. We then run the above regressions separately for stocks in different regions. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. We omit the coefficients on the control variables in Panel C for brevity. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
				Panel A: No controls				
				All ex-US				
MOM	0.217	0.290	0.268	0.217	0.308	0.165	0.216	0.304
	(3.32)	(4.72)	(4.46)	(3.75)	(6.32)	(2.70)	(3.40)	(6.06)
X	—	0.283	−0.232	0.242	−0.063	0.001	0.049	−0.008
		(6.72)	(−5.74)	(3.25)	(−0.60)	(0.04)	(3.04)	(−0.08)
MOM × *X*	—	−0.034	−0.028	0.118	0.017	−0.161	−0.009	−0.029
		(−1.33)	(−1.55)	(3.55)	(0.60)	(−5.56)	(−0.55)	(−1.19)
#stocks	9,782	7,001	7,122	9,782	9,768	9,780	6,299	9,781
Adj-R^2	1.0	1.3	1.2	2.2	3.2	1.4	0.9	3.2
				Developed ex-US				
MOM	0.277	0.352	0.341	0.271	0.397	0.230	0.285	0.415
	(3.59)	(4.57)	(4.61)	(4.21)	(7.20)	(3.22)	(3.72)	(7.24)
X	—	0.261	−0.056	0.267	−0.073	−0.007	0.044	−0.008
		(5.47)	(−1.04)	(2.90)	(−0.58)	(−0.19)	(2.39)	(−0.07)
MOM × *X*	—	−0.037	−0.047	0.078	0.058	−0.125	−0.018	−0.069
		(−1.22)	(−2.43)	(2.03)	(1.96)	(−4.11)	(−0.94)	(−2.56)
#stocks	5,987	3,870	3,852	5,987	5,978	5,987	3,457	5,987
Adj-R^2	1.7	2.1	2.1	4.2	5.4	2.5	1.4	5.4
				Emerging				
MOM	0.135	0.231	0.191	0.163	0.214	0.077	0.136	0.234
	(2.23)	(3.88)	(3.58)	(2.77)	(4.39)	(1.30)	(2.20)	(4.37)
X	—	0.369	−0.398	0.194	−0.080	0.018	0.084	0.028
		(7.47)	(−7.49)	(4.04)	(−0.88)	(0.61)	(3.71)	(0.35)
MOM × *X*	—	−0.043	0.003	0.117	0.008	−0.188	0.023	−0.077
		(−1.26)	(0.11)	(3.12)	(0.19)	(−4.60)	(0.99)	(−2.22)
#stocks	3,973	3,285	3,468	3,973	3,969	3,972	3,067	3,973
Adj-R^2	0.6	1.1	1.2	1.0	1.9	1.0	0.7	1.9

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
Panel B: Risk-adjusted returns on the left-hand-side								
	All ex-US							
MOM	0.169	0.213	0.218	0.175	0.271	0.141	0.169	0.239
	(4.48)	(5.32)	(5.58)	(5.01)	(8.98)	(3.95)	(4.17)	(6.50)
X	—	0.206	−0.211	0.181	−0.110	−0.028	0.026	0.010
		(8.10)	(−6.84)	(4.63)	(−1.88)	(−1.42)	(1.84)	(0.20)
$MOM \times X$	—	−0.053	−0.033	0.102	0.026	−0.108	−0.004	−0.032
		(−2.65)	(−2.05)	(3.83)	(1.10)	(−4.85)	(−0.25)	(−1.48)
#stocks	9,761	6,982	7,102	9,761	9,747	9,760	6,281	9,761
Adj-R^2	0.4	0.7	0.7	0.9	1.3	0.7	0.5	1.2
	Developed ex-US							
MOM	0.218	0.268	0.281	0.221	0.352	0.204	0.227	0.320
	(5.04)	(5.48)	(6.09)	(5.83)	(10.52)	(5.12)	(4.75)	(7.87)
X	—	0.198	−0.063	0.184	−0.120	−0.047	0.009	0.017
		(7.08)	(−1.71)	(3.87)	(−1.82)	(−2.03)	(0.59)	(0.31)
$MOM \times X$	—	−0.062	−0.050	0.059	0.061	−0.068	−0.015	−0.051
		(−2.60)	(−2.78)	(1.90)	(2.47)	(−3.14)	(−0.86)	(−2.19)
#stocks	5,979	3,862	3,844	5,979	5,970	5,979	3,450	5,979
Adj-R^2	0.8	1.1	1.2	1.7	2.2	1.2	0.8	2.0
	Emerging							
MOM	0.092	0.163	0.145	0.118	0.172	0.045	0.099	0.176
	(2.06)	(3.31)	(3.29)	(2.61)	(4.40)	(0.99)	(1.99)	(3.81)
X	—	0.284	−0.345	0.155	−0.112	0.008	0.081	0.019
		(6.89)	(−7.60)	(4.53)	(−1.75)	(0.31)	(3.81)	(0.38)
$MOM \times X$	—	−0.046	−0.003	0.117	0.023	−0.157	0.032	−0.076
		(−1.46)	(−0.10)	(3.51)	(0.61)	(−4.22)	(1.45)	(−2.34)
#stocks	3,960	3,273	3,455	3,960	3,956	3,959	3,055	3,960
Adj-R^2	0.4	0.8	1.0	0.6	1.2	0.8	0.5	1.0

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
Panel C: Controls on the right-hand-side								
	All ex-US							
MOM	0.283	0.259	0.342	0.301	0.255	0.242	0.280	0.367
	(5.16)	(4.30)	(6.49)	(5.62)	(6.08)	(4.56)	(5.13)	(7.75)
X	—	0.302	−0.206	0.118	0.180	−0.029	0.038	−0.171
		(7.68)	(−5.40)	(2.99)	(2.01)	(−1.31)	(2.42)	(−2.44)
MOM × *X*	—	−0.057	−0.042	0.072	0.055	−0.164	−0.004	−0.037
		(−2.24)	(−2.35)	(2.53)	(1.58)	(−5.37)	(−0.22)	(−1.40)
#stocks	6,102	6,102	5,849	6,102	6,101	6,102	6,073	6,102
Adj-R^2	2.1	2.3	2.6	2.4	3.2	2.5	2.2	3.2
	Developed ex-US							
MOM	0.332	0.315	0.389	0.355	0.305	0.294	0.331	0.459
	(5.21)	(4.30)	(6.43)	(5.80)	(6.45)	(4.69)	(5.17)	(8.74)
X	—	0.287	−0.032	0.134	0.178	−0.051	0.025	−0.170
		(6.48)	(−0.64)	(2.59)	(1.71)	(−2.22)	(1.51)	(−2.15)
MOM × *X*	—	−0.060	−0.046	0.047	0.114	−0.149	−0.014	−0.086
		(−1.95)	(−2.30)	(1.43)	(3.16)	(−4.84)	(−0.77)	(−3.03)
#stocks	3,370	3,370	3,209	3,370	3,369	3,370	3,354	3,370
Adj-R^2	3.2	3.5	4.0	3.7	4.7	3.6	3.4	4.8
	Emerging							
MOM	0.236	0.226	0.310	0.266	0.209	0.189	0.230	0.327
	(4.15)	(3.68)	(5.89)	(4.62)	(4.22)	(3.30)	(4.02)	(5.95)
X	—	0.384	−0.390	0.120	0.172	0.017	0.083	−0.156
		(7.51)	(−7.56)	(3.06)	(1.94)	(0.57)	(3.46)	(−2.21)
MOM × *X*	—	−0.026	0.012	0.098	0.002	−0.169	0.028	−0.036
		(−0.71)	(0.43)	(2.19)	(0.04)	(−3.56)	(1.12)	(−0.90)
#stocks	2,944	2,944	2,850	2,944	2,944	2,944	2,931	2,944
Adj-R^2	2.0	2.2	2.6	2.3	2.9	2.4	2.1	2.8

Table 4: Fama-MacBeth regressions of future returns on past momentum returns and explanatory variables: Subsamples

This table presents the results of Fama-MacBeth cross-sectional regressions similar to those in Panel C of Table 3:

$$\begin{aligned}R_{i,t} &= \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} \\ &+\gamma_{4,t} \times Size_{i,t-1} + \gamma_{5,t} \times MOM_{i,t-1} \times Size_{i,t-1} + \gamma_{6,t} \times (B/M)_{i,t-1} + \gamma_{7,t} \times (GP/AT)_{i,t-1} \\ &+\gamma_{8,t} \times ATG_{i,t-1} + \gamma_{9,t} \times R_{i,t-1} + e_{i,t}.\end{aligned}$$

We then run the above regressions separately for stocks in different regions. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. We omit the coefficients on the control variables for brevity. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2006 in Panel A, and 2007 to 2020 in Panel B.

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
Panel A: Sample period is 1993 to 2006								
	All ex-US							
MOM	0.402	0.347	0.433	0.424	0.395	0.375	0.398	0.480
	(5.61)	(4.51)	(6.25)	(6.13)	(6.58)	(5.67)	(5.58)	(8.56)
X	—	0.440	−0.138	0.124	0.128	−0.088	0.038	−0.144
		(7.13)	(−2.60)	(2.61)	(1.01)	(−2.62)	(1.53)	(−1.38)
MOM × *X*	—	−0.129	−0.035	0.022	0.084	−0.174	0.010	−0.041
		(−3.69)	(−1.26)	(0.65)	(1.76)	(−4.25)	(0.40)	(−0.99)
#stocks	4,411	4,411	4,160	4,411	4,410	4,411	4,385	4,411
Adj-R^2	2.4	2.6	2.9	2.7	3.6	2.7	2.5	3.6
	Developed ex-US							
MOM	0.396	0.358	0.446	0.418	0.399	0.370	0.394	0.515
	(4.56)	(3.73)	(5.38)	(5.01)	(5.86)	(4.58)	(4.52)	(7.72)
X	—	0.404	−0.018	0.115	0.140	−0.099	0.016	−0.158
		(5.95)	(−0.27)	(2.12)	(0.98)	(−2.47)	(0.63)	(−1.43)
MOM × *X*	—	−0.109	−0.056	0.025	0.090	−0.193	0.006	−0.079
		(−2.71)	(−1.65)	(0.69)	(1.66)	(−4.20)	(0.25)	(−1.70)
#stocks	2,977	2,977	2,777	2,977	2,977	2,977	2,960	2,977
Adj-R^2	3.6	3.9	4.5	4.0	5.2	4.2	3.8	5.2
	Emerging							
MOM	0.409	0.358	0.415	0.456	0.366	0.375	0.397	0.477
	(5.13)	(4.07)	(5.69)	(5.70)	(4.81)	(4.67)	(4.92)	(6.38)
X	—	0.590	−0.326	0.130	0.107	−0.008	0.134	−0.115
		(6.89)	(−4.17)	(2.09)	(0.78)	(−0.18)	(3.08)	(−1.01)
MOM × *X*	—	−0.114	0.068	0.035	0.048	−0.189	0.061	−0.038
		(−1.65)	(1.36)	(0.44)	(0.63)	(−2.57)	(1.34)	(−0.52)
#stocks	1,677	1,677	1,630	1,677	1,676	1,677	1,668	1,677
Adj-R^2	1.8	2.0	2.4	2.1	2.7	2.2	2.0	2.7

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
Panel B: Sample period is 2007 to 2020								
	All ex-US							
MOM	0.164	0.172	0.252	0.179	0.116	0.110	0.163	0.255
	(2.01)	(1.86)	(3.20)	(2.22)	(2.05)	(1.35)	(2.00)	(3.39)
X	—	0.167	−0.274	0.113	0.232	0.030	0.037	−0.198
		(3.54)	(−5.01)	(1.78)	(1.83)	(1.08)	(1.99)	(−2.09)
$MOM \times X$	—	0.014	−0.048	0.121	0.025	−0.155	−0.017	−0.033
		(0.38)	(−2.18)	(2.66)	(0.51)	(−3.40)	(−0.86)	(−1.01)
#stocks	7,773	7,773	7,518	7,773	7,771	7,773	7,741	7,773
Adj-R^2	1.8	2.0	2.3	2.2	2.8	2.2	1.9	2.8
	Developed ex-US							
MOM	0.270	0.273	0.332	0.293	0.211	0.220	0.268	0.404
	(2.89)	(2.46)	(3.77)	(3.27)	(3.26)	(2.29)	(2.87)	(4.99)
X	—	0.170	−0.046	0.153	0.215	−0.004	0.035	−0.182
		(3.07)	(−0.61)	(1.73)	(1.41)	(−0.19)	(1.55)	(−1.61)
$MOM \times X$	—	−0.012	−0.037	0.070	0.137	−0.106	−0.034	−0.092
		(−0.25)	(−1.67)	(1.26)	(2.89)	(−2.58)	(−1.33)	(−2.84)
#stocks	3,758	3,758	3,636	3,758	3,757	3,758	3,743	3,758
Adj-R^2	2.8	3.1	3.5	3.5	4.2	3.1	3.0	4.3
	Emerging							
MOM	0.083	0.109	0.216	0.097	0.070	0.024	0.081	0.194
	(1.05)	(1.29)	(2.90)	(1.21)	(1.11)	(0.30)	(1.02)	(2.47)
X	—	0.201	−0.447	0.112	0.230	0.040	0.038	−0.193
		(3.58)	(−6.54)	(2.24)	(2.00)	(1.00)	(1.62)	(−2.20)
$MOM \times X$	—	0.051	−0.037	0.154	−0.040	−0.151	0.000	−0.034
		(1.46)	(−1.18)	(3.25)	(−0.63)	(−2.46)	(−0.00)	(−0.88)
#stocks	4,068	4,068	3,932	4,068	4,067	4,068	4,051	4,068
Adj-R^2	2.1	2.3	2.8	2.4	3.0	2.6	2.2	3.0

Table 5: Penalized regressions of future returns on past momentum return and explanatory variables

We run the regression:

$$R_{i,t} \;=\; \gamma_0 + \gamma_1 \times MOM_{i,t-1} + \gamma_2 \times X_{i,t-1} + \gamma_3 \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t},$$

where $MOM_{i,t-1}$ is the return from month $t-12$ to $t-2$ and X represents each of variables listed in the top row of Table 2. We process the variables MOM and X on the right-hand through the following steps every month (i) we winsorize at the 0.5% and 99.5% levels, (ii) we country neutralize by subtracting the cross-sectional mean across all stocks in our sample from that country, and (iii) we standardize across the entire sample to have zero mean and unit standard deviation. The processed variables are used in the interaction terms. The column 'OLS' runs panel regressions. The column 'LASSO' runs LASSO regressions and the column 'ENet' runs elastic net regressions (with $\rho = 0.5$). We use 10-fold cross-validation for LASSO and ENet. Coefficients selected to be zero by LASSO or ENet are highlighted in boldface. We then run the above regressions separately for stocks in different regions. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2006.

	All ex-US			Developed ex-US			Emerging		
	OLS	LASSO	ENet	OLS	LASSO	ENet	OLS	LASSO	ENet
MOM	0.411	0.403	0.400	0.461	0.431	0.447	0.423	0.397	0.399
MOM × B/M	−0.096	−0.094	−0.093	−0.130	−0.117	−0.123	−0.001	**0.000**	**0.000**
MOM × Turnover	−0.078	−0.075	−0.073	−0.108	−0.102	−0.106	0.007	**0.000**	**0.000**
MOM × ResAnly	−0.005	**0.000**	**0.000**	−0.017	0.000	−0.009	0.069	0.048	0.050
MOM × 52wHi	0.034	0.029	0.027	−0.015	**0.000**	−0.004	0.056	0.046	0.047
MOM × ID	−0.241	−0.238	−0.237	−0.249	−0.245	−0.247	−0.251	−0.244	−0.244
MOM × COGS/TA	−0.024	−0.018	−0.016	−0.037	−0.022	−0.031	0.017	0.006	0.007
MOM × RetVolat	0.000	**0.000**	**0.000**	−0.047	−0.025	−0.034	−0.014	−0.004	−0.004
B/M	0.449	0.440	0.436	0.432	0.412	0.423	0.532	0.506	0.508
Turnover	−0.107	−0.102	−0.100	0.027	0.007	0.018	−0.327	−0.314	−0.315
ResAnly	0.122	0.112	0.108	0.094	0.075	0.086	0.201	0.175	0.177
52wHi	0.078	0.073	0.071	0.005	0.004	0.006	0.099	0.088	0.089
ID	−0.091	−0.085	−0.083	−0.094	−0.078	−0.086	−0.069	−0.056	−0.057
COGS/TA	0.061	0.055	0.053	0.051	0.036	0.044	0.125	0.111	0.112
RetVolat	−0.146	−0.144	−0.143	−0.223	−0.208	−0.217	−0.013	−0.007	−0.008
Intercept	0.772	0.774	0.775	0.853	0.850	0.850	0.804	0.806	0.806

Table 6: Penalized regressions of future returns on past momentum return and explanatory variables: Out-of-sample performance

We run penalized regressions as in Table 5. Using the coefficient estimates from the training period (1993 to 2006), we calculate forecast errors for the out-of-sample period of 2007 to 2020. We calculate OOS-R^2 as:

$$\text{OOS-}R^2 = 1 - \frac{\sum_{(i,t)} \left(R_{i,t} - \hat{R}_{i,t}\right)^2}{\sum_{(i,t)} R_{i,t}^2},$$

where $R_{i,t}$ is the realized return and $\hat{R}_{i,t}$ is the forecasted return using OLS, LASSO, or elastic net (ENet). We calculate OOS-MSE as

$$\text{OOS-MSE} = \frac{1}{T}\sum_t \frac{1}{N_t}\sum_i \left(R_{i,t} - \hat{R}_{i,t}\right)^2.$$

We calculate OOS-MAE as

$$\text{OOS-MAE} = \frac{1}{T}\sum_t \frac{1}{N_t}\sum_i |R_{i,t} - \hat{R}_{i,t}|.$$

In each case, the coefficient estimates are not updated over the out-of-sample period. The R^2 is reported in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country).

	All ex-US			Developed ex-US			Emerging		
	OLS	LASSO	ENet	OLS	LASSO	ENet	OLS	LASSO	ENet
OOS-R^2	0.363	0.368	0.370	0.189	0.198	0.193	0.581	0.593	0.592
OOS-MSE	0.017	0.017	0.017	0.015	0.015	0.015	0.020	0.020	0.020
OOS-MAE	0.090	0.090	0.090	0.083	0.083	0.083	0.098	0.098	0.098

Table 7: Double-sorted portfolio alphas on past momentum return and explanatory variables

Each month we sort stocks into tercile portfolios based on last 11-month returns skipping the most recent month, MOM, and an X variable. The stocks are independently sorted in Panel B and sequentially sorted in Panel B (where we first sort on X and then on MOM). The portfolios are value-weighted and rebalanced monthly. For all sorts, we country neutralize by subtracting the cross-sectional country (not regional) mean of that variable. We calculate the winner minus loser (WML) portfolio for each tercile of X. The table reports the difference in WML returns, ΔWML, across high and low X terciles. The returns are annualized and t-statistics are reported in parentheses below the returns. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	B/M	Turn	ResAnly	52wHi	ID	COGS	RetVol
Panel A: Independent sorts							
All ex-US							
ΔWML	−4.77	−1.31	1.12	−0.67	−4.82	−0.54	3.85
	(−2.04)	(−0.56)	(0.50)	(−0.24)	(−2.17)	(−0.32)	(1.27)
Developed ex-US							
ΔWML	−3.52	−2.81	0.01	−0.43	−4.20	0.17	4.19
	(−1.45)	(−1.18)	(0.01)	(−0.14)	(−1.74)	(0.09)	(1.16)
Emerging							
ΔWML	−3.47	−0.14	4.13	−4.11	−8.88	1.333	3.29
	(−1.07)	(−0.05)	(1.42)	(−0.94)	(−2.54)	(0.46)	(0.86)
Panel B: Sequential sorts							
All ex-US							
ΔWML	−4.38	−2.09	3.44	−0.77	−6.50	−0.30	7.75
	(−1.66)	(−0.83)	(1.58)	(−0.32)	(−2.87)	(−0.18)	(2.45)
Developed ex-US							
ΔWML	−4.26	−2.82	2.58	−0.64	−6.14	0.43	6.60
	(−1.59)	(−1.10)	(1.01)	(−0.23)	(−2.39)	(0.23)	(1.92)
Emerging							
ΔWML	−5.40	0.46	4.80	−4.77	−9.49	1.73	6.15
	(−1.63)	(0.15)	(1.64)	(−1.42)	(−2.84)	(0.58)	(1.61)

Table 8: Time-series determinants of momentum

This table describes the results of the time series:

$$R_t = \gamma_1 \times State1_{t-1} + \gamma_2 \times State2_{t-1} + e_t,$$

where R is the loser, or winner, or winner minus loser portfolio constructed by sorting on last 11 month returns (excluding the most recent month) and State are dummy variables indicating macroeconomic state in the previous month. UP (DOWN) is equal to one if the market return over the last 36 months is positive (negative), and zero otherwise. We use MSCI All-Country ex-US, World ex-US, and Emerging total return indices as the proxies for market return for All ex-US, Developed ex-US, and Emerging markets, respectively. HIVOL (LOVOL) is equal to one if the market volatility over the last 12 months is higher (lower) than the market volatility over the past 36 months, and zero otherwise. Market volatility is calculated using daily data. We use MSCI All-Country ex-US, World ex-US, and Emerging price (not total return) indices as the market return proxies for All ex-US, Developed ex-US, and Emerging markets, respectively. The table reports the annualized slopes (in percent) from the above regression together with their t-statistics. The sample consists of only non-microcap stocks (those in the top 97 of the market capitalization of each country). The sample period is 1993 to 2020.

State	#obs	L	W	WML	State	#obs	L	W	WML
				All ex-US					
UP	262	−3.88	10.27	14.16	HIVOL	140	12.65	13.77	1.12
		(−0.79)	(2.57)	(3.52)			(1.87)	(2.51)	(0.20)
DOWN	71	23.12	21.19	−1.94	LOVOL	193	−5.94	11.75	17.69
		(2.44)	(2.76)	(−0.25)			(−1.03)	(2.52)	(3.79)
DIFF		−27.01	−10.92	16.09	DIFF		18.59	2.01	−16.58
		(−2.52)	(−1.26)	(1.85)			(2.09)	(0.28)	(−2.30)
				Developed ex-US					
UP	258	−4.37	10.22	14.59	HIVOL	141	9.02	12.02	2.99
		(−0.84)	(2.67)	(3.21)			(1.27)	(2.32)	(0.49)
DOWN	75	22.47	18.25	−4.22	LOVOL	192	−3.72	12.04	15.75
		(2.33)	(2.57)	(−0.50)			(−0.61)	(2.71)	(2.98)
DIFF		−26.84	−8.03	18.81	DIFF		12.74	−0.02	−12.76
		(−2.45)	(−1.00)	(1.96)			(1.36)	(−0.00)	(−1.57)
				Emerging					
UP	237	0.48	13.03	12.56	HIVOL	141	10.46	9.98	−0.49
		(0.09)	(2.49)	(3.15)			(1.52)	(1.47)	(−0.09)
DOWN	96	11.20	10.87	−0.33	LOVOL	192	−1.50	14.19	15.69
		(1.34)	(1.32)	(−0.05)			(−0.25)	(2.44)	(3.56)
DIFF		−10.72	2.17	12.89	DIFF		11.96	−4.22	−16.18
		(−1.08)	(0.22)	(1.74)			(1.32)	(−0.47)	(−2.39)

Table 9: Time-series determinants of momentum: Subsamples

This table describes the results of the time series:

$$R_t = \gamma_1 State1_{t-1} + \gamma_2 State2_{t-1} + e_t,$$

where R is the loser, or winner, or winner minus loser portfolio constructed by sorting on last 11 month returns (excluding the most recent month) and State are dummy variables indicating the state of the market in the previous month. UP (DOWN) is equal to one if the market return over the last 36 months is positive (negative), and zero otherwise. We use MSCI All-Country ex-US, World ex-US, and Emerging total return indices as the proxies for market return for All ex-US, Developed ex-US, and Emerging markets, respectively. HIVOL (LOVOL) is equal to one if the market volatility over the last 12 months is higher (lower) than the market volatility over the past 36 months, and zero otherwise. Market volatility is calculated using daily data. We use MSCI All-Country ex-US, World ex-US, and Emerging price (not total return) indices as the market return proxies for All ex-US, Developed ex-US, and Emerging markets, respectively. The table reports the annualized slopes (in percent) from the above regression together with their t-statistics. The sample consists of only non-microcap stocks (those in the top 97 of the market capitalization of each country). The sample period is 1993 to 2006 in Panel A and from 2007 to 2020 in Panel B.

State	#obs	L	W	WML	State	#obs	L	W	WML
Panel A: Sample period is 1993 to 2006									
All ex-US									
UP	133	−2.72	14.64	17.36	HIVOL	58	10.96	19.53	8.57
		(−0.45)	(2.75)	(2.87)			(1.19)	(2.41)	(0.94)
DOWN	32	13.78	26.76	12.98	LOVOL	107	−5.20	15.62	20.82
		(1.11)	(2.46)	(1.05)			(−0.77)	(2.62)	(3.09)
DIFF		−16.50	−12.12	4.38	DIFF		16.16	3.92	−12.25
		(−1.19)	(−1.00)	(0.32)			(1.41)	(0.39)	(−1.08)
Developed ex-US									
UP	133	0.18	14.16	13.98	HIVOL	59	12.11	17.31	5.20
		(0.03)	(2.60)	(2.05)			(1.23)	(2.12)	(0.51)
DOWN	32	11.31	23.23	11.93	LOVOL	106	−3.10	15.14	18.24
		(0.85)	(2.09)	(0.86)			(−0.42)	(2.48)	(2.39)
DIFF		−11.12	−9.07	2.05	DIFF		15.20	2.17	−13.04
		(−0.75)	(−0.73)	(0.13)			(1.24)	(0.21)	(−1.02)
Emerging									
UP	108	6.58	19.58	13.00	HIVOL	59	12.65	15.00	2.35
		(0.93)	(2.65)	(2.12)			(1.33)	(1.50)	(0.28)
DOWN	57	0.67	9.04	8.36	LOVOL	106	0.03	16.46	16.43
		(0.07)	(0.89)	(0.99)			(0.00)	(2.20)	(2.67)
DIFF		5.91	10.55	4.64	DIFF		12.62	−1.46	−14.09
		(0.49)	(0.84)	(0.44)			(1.06)	(−0.12)	(−1.37)

State	#obs	L	W	WML	State	#obs	L	W	WML
Panel B: Sample period is 2007 to 2020									
All ex-US									
UP	128	−5.49	5.59	11.08	HIVOL	81	13.43	9.45	−3.97
		(−0.70)	(0.93)	(2.10)			(1.35)	(1.25)	(−0.60)
DOWN	39	30.79	16.62	−14.17	LOVOL	86	−6.86	6.95	13.81
		(2.16)	(1.53)	(−1.48)			(−0.71)	(0.95)	(2.13)
DIFF		−36.28	−11.03	25.25	DIFF		20.29	2.51	−17.78
		(−2.23)	(−0.89)	(2.31)			(1.46)	(0.24)	(−1.91)
Developed ex-US									
UP	124	−9.50	6.00	15.50	HIVOL	81	6.56	8.19	1.63
		(−1.16)	(1.10)	(2.57)			(0.64)	(1.21)	(0.22)
DOWN	43	30.78	14.54	−16.24	LOVOL	86	−4.48	8.20	12.68
		(2.22)	(1.57)	(−1.59)			(−0.45)	(1.25)	(1.72)
DIFF		−40.28	−8.55	31.74	DIFF		11.04	−0.01	−11.05
		(−2.50)	(−0.80)	(2.67)			(0.77)	(−0.00)	(−1.05)
Emerging									
UP	128	−4.96	7.02	11.98	HIVOL	81	8.55	5.52	−3.04
		(−0.63)	(0.94)	(2.30)			(0.85)	(0.59)	(−0.46)
DOWN	39	26.58	13.54	−13.04	LOVOL	86	−3.38	11.40	14.78
		(1.86)	(1.00)	(−1.38)			(−0.35)	(1.25)	(2.32)
DIFF		−31.54	−6.52	25.02	DIFF		11.93	−5.88	−17.81
		(−1.93)	(−0.42)	(2.32)			(0.85)	(−0.45)	(−1.95)

Appendix: Extra Tables

Table A1: References and Abbreviations

Paper	Abbreviation
Jegadeesh and Titman (1993)	JT (1993)
Daniel, Hirshleifer, and Subrahmanyam (1998)	DHS (1998)
Daniel and Titman (1999)	DT (1999)
Lee and Swaminathan (2000)	LS (2000)
Hong and Stein (1999)	HS (1999)
Hong, Lim, and Stein (2000)	HLS (2000)
George and Hwang (2004)	GH (2004)
Da, Gurun, and Warachka (2014)	DGW (2014)
Sagi and Seasholes (2007)	SS (2007)
Cooper, Gutierrez, and Hameed (2004)	CGH (2004)
Wang and Xu (2015)	WX (2015)

Table A2: Worldscope variables

Code	Name
WC01001	Sales
WC01051	COGS
WC02999	Total assets
WC03063	Income taxes payable
WC03263	Deferred taxes
WC03351	Total liabilities
WC03451	Preferred stock
WC03501	Common equity
WC03995	Shareholder equity
Book equity	[(Shareholder equity) or (Common equity + Preferred stock*) or (Total assets − Total liabilities)] + (Deferred Taxes* − Preferred stock*)
Gross profit	Sales − COGS

In general, we do not replace missing values with zero. However, if a variable is starred in the above list, then it is set to zero if missing.

Table A3: Fama-MacBeth regressions of future returns on past momentum return and explanatory variables: Risk-adjusting returns with additional PMN factor

This table presents the results of Fama-MacBeth cross-sectional regressions similar to those in Panel B of Table 3:

$$R_{i,t} - \beta_i' F_t \quad = \quad \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t},$$

except that we include an additional factor on the left-hand-side to risk-adjust returns. Betas are calculated using the full-sample for each stock from a five-factor Fama and French (2017) model and a PMN factor. The PMN factor is constructed as follows. Each month we sort stocks into decile portfolios based on standardized unexpected earnings (SUE), which are defined as changes in annual earnings divided by the most recent price. The portfolios are value weighted and rebalanced monthly. We calculate the PMN factor as the decile 10 minus the decile 1 portolio. We use separate factors for developed and emerging markets and use the factor model for the stock corresponding to its region. In the regression equation, $MOM_{i,t-1}$ is the return from month $t-12$ to $t-2$ and X represents each of variables listed in the top row of Table 2. We process the variables MOM and X on the right-hand through the following steps every month (i) we winsorize at the 0.5% and 99.5% levels, (ii) we country neutralize by subtracting the cross-sectional mean across all stocks in our sample from that country, and (iii) we standardize across the entire sample to have zero mean and unit standard deviation. The processed variables are used in the interaction terms. We run the regressions separately for stocks in different regions. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. We omit the coefficients on the control variables in Panel C for brevity. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	(1) —	(2) B/M	(3) Turn	(4) ResAnly	(5) 52wHi	(6) ID	(7) COGS	(8) RetVol
				All ex-US				
MOM	0.168	0.210	0.217	0.174	0.270	0.140	0.168	0.237
	(4.52)	(5.28)	(5.60)	(5.04)	(9.03)	(3.98)	(4.19)	(6.48)
X	—	0.204	−0.212	0.181	−0.111	−0.027	0.026	0.011
		(8.19)	(−6.83)	(4.70)	(−1.92)	(−1.40)	(1.88)	(0.24)
$MOM \times X$	—	−0.053	−0.032	0.101	0.026	−0.111	−0.004	−0.031
		(−2.68)	(−1.99)	(3.81)	(1.08)	(−4.98)	(−0.26)	(−1.45)
#stocks	9,761	6,982	7,102	9,761	9,747	9,759	6,280	9,760
Adj-R^2	0.4	0.7	0.7	0.9	1.3	0.7	0.5	1.2
				Developed ex-US				
MOM	0.216	0.262	0.278	0.219	0.351	0.203	0.225	0.317
	(5.08)	(5.41)	(6.09)	(5.86)	(10.58)	(5.15)	(4.75)	(7.87)
X	—	0.195	−0.062	0.186	−0.122	−0.047	0.010	0.019
		(7.16)	(−1.68)	(3.96)	(−1.88)	(−2.03)	(0.64)	(0.37)
$MOM \times X$	–	−0.064	−0.050	0.059	0.062	−0.071	−0.015	−0.051
		(−2.70)	(−2.80)	(1.91)	(2.50)	(−3.23)	(−0.85)	(−2.18)
#stocks	5,979	3,862	3,844	5,979	5,969	5,978	3,449	5,979
Adj-R^2	0.7	1.1	1.2	1.7	2.2	1.2	0.7	2.0
				Emerging				
MOM	0.094	0.168	0.148	0.118	0.173	0.045	0.099	0.174
	(2.12)	(3.42)	(3.38)	(2.64)	(4.44)	(1.01)	(2.00)	(3.78)
X	—	0.284	−0.346	0.149	−0.109	0.010	0.080	0.020
		(6.94)	(−7.61)	(4.33)	(−1.71)	(0.39)	(3.75)	(0.41)
$MOM \times X$	—	−0.045	0.003	0.113	0.022	−0.162	0.033	−0.072
		(−1.42)	(0.09)	(3.34)	(0.58)	(−4.31)	(1.45)	(−2.20)
#stocks	3,960	3,273	3,455	3,960	3,956	3,959	3,055	3,960
Adj-R^2	0.4	0.7	0.9	0.6	1.2	0.8	0.5	1.0

Table A4: Fama-MacBeth regressions of future returns on past 6-month momentum return and explanatory variables

This table presents the results of Fama-MacBeth cross-sectional regressions similar to those in Panel C of Table 3:

$$\begin{aligned} R_{i,t} &= \gamma_{0,t} + \gamma_{1,t} \times MOM'_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM'_{i,t-1} \times X_{i,t-1} \\ &\quad + \gamma_{4,t} \times Size_{i,t-1} + \gamma_{5,t} \times MOM'_{i,t-1} \times Size_{i,t-1} + \gamma_{6,t} \times (B/M)_{i,t-1} + \gamma_{7,t} \times (GP/AT)_{i,t-1} \\ &\quad + \gamma_{8,t} \times ATG_{i,t-1} + \gamma_{9,t} \times R_{i,t-1} + e_{i,t}, \end{aligned}$$

except that we use MOM' defined as the return from month $t-7$ to $t-2$. We then run the above regressions separately for stocks in different regions. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. We omit the coefficients on the control variables for brevity. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	—	B/M	Turn	ResAnly	52wHi	ID	COGS	RetVol
				All ex-US				
MOM'	0.284	0.262	0.342	0.303	0.253	0.242	0.281	0.371
	(5.20)	(4.36)	(6.53)	(5.69)	(6.08)	(4.59)	(5.17)	(7.86)
X	—	0.300	−0.211	0.125	0.187	−0.029	0.038	−0.185
		(7.66)	(−5.55)	(3.19)	(2.09)	(−1.33)	(2.48)	(−2.65)
$MOM' \times X$	—	−0.053	−0.042	0.072	0.051	−0.168	−0.003	−0.039
		(−2.11)	(−2.45)	(2.57)	(1.50)	(−5.55)	(−0.21)	(−1.50)
#stocks	6,102	6,102	5,849	6,102	6,101	6,102	6,073	6,102
Adj-R^2	2.2	2.3	2.7	2.5	3.2	2.5	2.3	3.2
				Developed ex-US				
MOM'	0.332	0.319	0.389	0.356	0.300	0.292	0.331	0.460
	(5.24)	(4.38)	(6.46)	(5.85)	(6.43)	(4.70)	(5.20)	(8.80)
X	—	0.286	−0.042	0.140	0.187	−0.050	0.028	−0.190
		(6.50)	(−0.83)	(2.71)	(1.80)	(−2.21)	(1.71)	(−2.41)
$MOM' \times X$	—	−0.053	−0.047	0.050	0.107	−0.155	−0.016	−0.086
		(−1.72)	(−2.41)	(1.54)	(3.03)	(−5.10)	(−0.88)	(−3.12)
#stocks	3,370	3,370	3,209	3,370	3,369	3,370	3,354	3,370
Adj-R^2	3.3	3.6	4.1	3.8	4.8	3.7	3.4	4.9
				Emerging				
MOM'	0.235	0.230	0.310	0.267	0.204	0.186	0.228	0.330
	(4.17)	(3.75)	(5.93)	(4.66)	(4.15)	(3.27)	(4.03)	(6.00)
X	—	0.381	−0.389	0.129	0.183	0.017	0.082	−0.170
		(7.51)	(−7.60)	(3.31)	(2.07)	(0.55)	(3.45)	(−2.42)
$MOM' \times X$	—	−0.017	0.007	0.106	−0.002	−0.172	0.027	−0.038
		(−0.47)	(0.24)	(2.43)	(−0.05)	(−3.65)	(1.08)	(−0.97)
#stocks	2,944	2,944	2,850	2,944	2,944	2,944	2,931	2,944
Adj-R^2	2.0	2.2	2.7	2.3	2.9	2.5	2.1	2.9

Table A5: Fama-MacBeth regressions of future returns on past momentum returns and a full set of explanatory variables

This table presents the results of Fama-MacBeth cross-sectional regressions.

$$R_{i,t} \quad = \quad \gamma_{0,t} + \gamma_{1,t} \times MOM_{i,t-1} + \gamma_{2,t} \times X_{i,t-1} + \gamma_{3,t} \times MOM_{i,t-1} \times X_{i,t-1} + e_{i,t},$$

where where $MOM_{i,t-1}$ is the return from month $t-12$ to $t-2$ and X represents each of variables listed in the top row of Table 2. We process the variables MOM and X on the right-hand through the following steps every month (i) we winsorize at the 0.5% and 99.5% levels, (ii) we country neutralize by subtracting the cross-sectional mean across all stocks in our sample from that country, and (iii) we standardize across the entire sample to have zero mean and unit standard deviation. The processed variables are used in the interaction terms. We run the regressions separately for stocks in different regions. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2006.

	All ex-US	Developed ex-US	Emerging
MOM	0.512	0.639	0.403
	(7.71)	(8.85)	(3.94)
MOM × B/M	−0.093	−0.110	−0.111
	(−3.13)	(−3.34)	(−1.86)
MOM × Turnover	−0.039	−0.050	0.066
	(−1.45)	(−1.67)	(1.40)
MOM × ResAnly	0.007	−0.013	0.077
	(0.27)	(−0.45)	(1.08)
MOM × 52wHi	0.026	−0.021	0.040
	(0.58)	(−0.40)	(0.52)
MOM × ID	−0.143	−0.095	−0.202
	(−3.60)	(−2.29)	(−2.90)
MOM × COGS/TA	−0.018	−0.015	0.021
	(−0.82)	(−0.63)	(0.49)
MOM × RetVolat	−0.031	−0.105	−0.053
	(−0.68)	(−2.11)	(−0.59)
B/M	0.355	0.334	0.446
	(5.83)	(5.11)	(6.04)
Turnover	−0.136	0.011	−0.356
	(−3.55)	(0.24)	(−5.74)
ResAnly	0.094	0.064	0.132
	(2.49)	(1.58)	(2.19)
52wHi	−0.110	−0.191	0.033
	(−1.05)	(−1.57)	(0.28)
ID	−0.049	−0.037	−0.009
	(−1.77)	(−1.11)	(−0.20)
COGS/TA	0.037	0.025	0.136
	(1.39)	(0.91)	(3.25)
RetVolat	−0.193	−0.274	−0.052
	(−2.40)	(−3.22)	(−0.60)
#stocks	4,353	2,840	1,792
Adj-R2	4.2	6.2	3.8

Table A6: Longer horizon portfolio returns and alphas on past momentum returns

Each month we sort stocks into decile portfolios based on last 11-month returns (skipping the most recent month). The portfolios are value weighted and rebalanced monthly. We country neutralize by subtracting the cross-sectional country (not regional) mean of past returns. We calculate the extreme winner minus the extreme loser portfolio returns. These post-formation returns are calculated over horizons ranging from one month to 60 months. The horizon is indicated in the first column. For horizons of multiple months, we use overlapping portfolio approach of Jegadeesh and Titman (1993). We report the returns as well as alphas based on the five-factor Fama and French (2017) model. We use separate factors for developed and emerging markets and use the factor model for the stock corresponding to its region. The returns and alphas are annualized and t-statistics are reported in parentheses below the returns. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	Returns			Five-factor alphas		
Months	All ex-US	Developed ex-US	Emerging	All ex-US	Developed ex-US	Emerging
1	10.64	10.23	8.93	8.06	7.26	10.05
	(2.98)	(2.55)	(2.65)	(2.24)	(1.86)	(2.71)
1 to 6	6.60	7.07	6.43	4.36	5.12	6.93
	(2.17)	(2.05)	(2.23)	(1.45)	(1.54)	(2.20)
7 to 12	0.25	1.64	3.44	−1.52	1.59	5.09
	(0.10)	(0.63)	(1.27)	(−0.68)	(0.65)	(1.83)
13 to 24	−0.07	−1.16	2.22	0.75	0.07	4.51
	(−0.04)	(−0.56)	(1.00)	(0.46)	(0.04)	(2.08)
25 to 36	2.34	1.39	1.55	3.34	1.74	3.44
	(1.34)	(0.69)	(0.75)	(2.00)	(0.91)	(1.71)
37 to 48	0.06	1.02	0.62	1.19	2.56	2.99
	(0.04)	(0.60)	(0.31)	(0.83)	(1.53)	(1.45)
49 to 60	−0.75	−1.53	−0.71	−0.12	−0.35	1.29
	(−0.44)	(−0.87)	(−0.32)	(−0.07)	(−0.20)	(0.58)

Table A7: Fama-MacBeth regressions of future returns on past momentum return and explanatory variables including fund flows

This table presents the results of Fama-MacBeth cross-sectional regressions as in Panel A of Table 3. We add one more X variable to these regressions, namely ΔFundFlow. This is calculated as the percentage change in the institutional fund holding (the ratio of the number of shares held by institutions to the total number of shares outstanding). The rest of the procedure remains the same. The table reports the time-series averages of coefficients together with their t-statistics in parentheses. The last two rows in each subpanel report the average number of stocks and the average adjusted-R^2 in percent. The sample consists of only non-microcap stocks (those in the top 97% of the market capitalization of each country). The sample period is 1993 to 2020.

	—	B/M	Turn	ResAnly	52wHi	ID	COGS	RetVol	ΔFundFlow
					All ex-US				
MOM	0.217	0.290	0.268	0.217	0.308	0.165	0.216	0.304	0.191
	(3.32)	(4.72)	(4.46)	(3.75)	(6.32)	(2.70)	(3.40)	(6.06)	(2.53)
X	—	0.283	−0.232	0.242	−0.063	0.001	0.049	−0.008	0.011
		(6.72)	(−5.74)	(3.25)	(−0.60)	(0.04)	(3.04)	(−0.08)	(0.76)
MOM × *X*	—	−0.034	−0.028	0.118	0.017	−0.161	−0.009	−0.029	−0.014
		(−1.33)	(−1.55)	(3.55)	(0.60)	(−5.56)	(−0.55)	(−1.19)	(−0.62)
#stocks	9,782	7,001	7,122	9,782	9,768	9,780	6,299	9,781	2,128
Adj-R^2	1.0	1.3	1.2	2.2	3.2	1.4	0.9	3.2	1.3
					Developed ex-US				
MOM	0.277	0.352	0.341	0.271	0.397	0.230	0.285	0.415	0.240
	(3.59)	(4.57)	(4.61)	(4.21)	(7.20)	(3.22)	(3.72)	(7.24)	(2.96)
X	–	0.261	−0.056	0.267	−0.073	−0.007	0.044	−0.008	0.004
		(5.47)	(−1.04)	(2.90)	(−0.58)	(−0.19)	(2.39)	(−0.07)	(0.20)
MOM × *X*	—	−0.037	−0.047	0.078	0.058	−0.125	−0.018	−0.069	−0.023
		(−1.22)	(−2.43)	(2.03)	(1.96)	(−4.11)	(−0.94)	(−2.56)	(−0.97)
#stocks	5,987	3,870	3,852	5,987	5,978	5,987	3,457	5,987	1,475
Adj-R^2	1.7	2.1	2.1	4.2	5.4	2.5	1.4	5.4	1.8
					Emerging				
MOM	0.135	0.231	0.191	0.163	0.214	0.077	0.136	0.234	0.081
	(2.23)	(3.88)	(3.58)	(2.77)	(4.39)	(1.30)	(2.20)	(4.37)	(0.94)
X	—	0.369	−0.398	0.194	−0.080	0.018	0.084	0.028	0.007
		(7.47)	(−7.49)	(4.04)	(−0.88)	(0.61)	(3.71)	(0.35)	(0.16)
MOM × *X*	—	−0.043	0.003	0.117	0.008	−0.188	0.023	−0.077	−0.011
		(−1.26)	(0.11)	(3.12)	(0.19)	(−4.60)	(0.99)	(−2.22)	(−0.25)
#stocks	3,973	3,285	3,468	3,973	3,969	3,972	3,067	3,973	930
Adj-R^2	0.6	1.1	1.2	1.0	1.9	1.0	0.7	1.9	0.9

Table A8: ID and market states

This table describes the results of the contemporaneous time series:

$$\text{IDvw}_t = \gamma_1 \times State1_t + \gamma_2 \times State2_t + e_t,$$

where IDvw is the value-weighted ID variable. Each month, we first winsorize ID at 0.5% level. Second, we country neutralize ID by subtracting the cross-sectional country (not regional) mean of ID. Third, ID is standardized to have zero mean and unit standard deviation. We then calculate the value-weighted average using last month's market capitalization to calculate IDvw. The *State* variables denote market states. UP (DOWN) is equal to one if the market return over the last 36 months is positive (negative), and zero otherwise. We use MSCI All-Country ex-US, World ex-US, and Emerging total return indices as market return proxies for All ex-US, Developed ex-US, and Emerging markets, respectively. HIVOL (LOVOL) is equal to one if the market volatility over the last 12 months is higher (lower) than the market volatility over the past 36 months, and zero otherwise. Market volatility is calculated using daily data. We use MSCI All-Country ex-US, World ex-US, and Emerging price (not total return) indices as the proxies for market return for All ex-US, Developed ex-US, and Emerging markets, respectively. The table reports slopes from the above regression together with their t-statistics. The sample consists of only non-microcap stocks (those in the top 97 of the market capitalization of each country). The sample period is 1993 to 2020.

	All ex-US	Developed ex-US	Emerging
	Panel A: States are up and down markets		
UP	0.0050	0.0059	0.0013
	(5.00)	(4.74)	(1.58)
DOWN	0.0102	0.0107	0.0098
	(5.24)	(4.59)	(7.31)
DIFF	−0.0051	−0.0047	−0.0085
	(−2.34)	(−1.80)	(−5.33)
	Panel B: States are high and low volatility		
HIVOL	0.0113	0.0135	0.0044
	(8.44)	(8.28)	(3.80)
LOVOL	0.0024	0.0022	0.0033
	(2.07)	(1.55)	(3.39)
DIFF	0.0089	0.0113	0.0010
	(5.06)	(5.27)	(0.67)