Flow-induced Trades and Asset Pricing Factors*

Shiyang Huang Yang Song Hong Xiang

First Version: March 2019

(Extremely Preliminary, Comments are Welcome)

Abstract

We show that mutual funds' flow-induced trades significantly influence returns and co-movement among 50 well-known asset pricing factors (anomalies). Mutual fund investors are ignorant about both systematic and idiosyncratic risks when allocating capital among funds. We measure the *non-fundamental* demand shocks to each factor by aggregating mutual funds' flow-induced trading of individual stocks underlying the factor. We show that flow-induced demand shifts largely determine factor return dynamics and that the expected (co)variance of flow-induced trades of factors strongly forecasts factor return (co)variance. Our results indicate that these factors are heavily exposed to flow-driven "noise trader" risk, which we further show is significantly priced. The flow-driven effects on factor return dynamics can partially explain factor momentum and underperformance of large-sized mutual funds relative to small funds.

*Huang and Xiang are with Faculty of Business and Economics, the University of Hong Kong, and Song is with Foster School of Business, University of Washington. Send correspondence to huangsy@hku.hk, songy18@uw.edu, and hongxiang18@hku.hk.

1 Introduction

In recent decades, asset pricing literature has discovered dozens of factors (anomalies) that arguably explain the average returns among individual stocks. While risk-based explanations are often offered, there is considerable debate over the economic plausibility and statistical reliability of these factors.[1]

In this paper, we provide a new perspective on this collection of asset pricing factors. We show that mutual funds' flow-induced trades significantly determine returns and comovement among a large set of well-studied factors, despite that fund flow movement is hard to be reconciled with a risk-based theory with rational agents. In other words, these factors are heavily exposed to non-fundamental "risk" that is due to flow-driven demand shifts of mutual funds. We also provide evidence that this non-fundamental risk is priced as liquidity suppliers and other investors demand higher average premia across factors when factors are susceptible to higher aggregate flow-driven demand shocks. The flow-driven effects on factor return dynamics partially account for factor momentum (Gupta and Kelly, 2019; Arnott, Clements, Kalesnik, and Linnainmaa, 2019)) and underperformance of large mutual funds relative to small-sized funds (Chen, Hong, Huang, and Kubik, 2004; Pollet and Wilson, 2008).

Our analysis is motivated by two recent observations in the mutual fund literature. First, mutual fund flows are largely uninformative and non-fundamental as investors barely use any asset pricing model when allocating capital among mutual funds. For example, Ben-David, Li, Rossi, and Song (2018) show that mutual fund investors rely on easy-to-follow signals (e.g. Morningstar Ratings) and largely ignore both systematic risks and idiosyncratic volatilities for their investment decisions.[2] Second, mutual funds' flow-induced trades have a large price impact

[1] See, for example, Cochrane (2011), Nagel (2013), McLean and Pontiff (2016), Harvey, Liu, and Zhu (2016), and Hou, Xue, and Zhang (2018). We use factors and anomalies interchangablely in this paper.

[2] Ben-David, Li, Rossi, and Song (2018) show that these ratings do not account for fund exposures to any systematic factor, e.g., the aggregate market, and explain less than 4% variation in return volatility. Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016) also show that mutual fund investors don't account for

on individual stock prices (Coval and Stafford, 2007; Frazzini and Lamont, 2008; Lou, 2012). The goal of this paper is to examine the link between mutual funds' flow-induced trades and asset pricing factors. We hypothesize and then verify that these non-fundamental flow-driven demand shocks strongly influence both returns and comovement among a large set of well-studied factors in the literature.

To this end, we replicate 50 characteristic-based asset pricing factors[3] following similar steps in Hou, Xue, and Zhang (2018). To construct a characteristic-based factor, we sort all NYSE-AMEX-NASDAQ stocks into quintile portfolios with NYSE breakpoints for that characteristic. The factor return is the spread between the value-weighted returns of the top-quintile and the bottom-quintile stocks. We use NYSE breakpoints and value-weighted returns to mitigate the impact of microcaps, as highlighted by Hou, Xue, and Zhang (2018).

To measure mutual funds' flow-driven demand shocks to a given factor π, we use a bottom-up approach in two steps: First, we estimate mutual fund flow-induced trading for each individual stock using the FIT measure of Lou (2012). In a nutshell, FIT measures the magnitude of flow-driven trading by the aggregate mutual fund industry on a particular shock in a given quarter. In the second step, we estimate factor-level flow-induced trading by value-weighting FIT of the stocks underlying the long leg and the short leg of factor π, respectively. The difference between the value-weighted FIT of the long and short legs is our measure of flow-induced trades of factor π, which we denote as FITOF_{π}. In total, we use 208,419 fund-quarter observations with 4,999 active equity mutual funds in the US from 1980 to 2017.

We start our analysis by examining the extent to which flow-induced trades of factors (FITOF) can affect contemporaneous factor returns. In panel regressions, we find that a one-standard-deviation increase in FITOF is significantly associated with an increase of 1.18% in

the Fama-French size and value factors. Song (2018) shows that investors' irrational behaviors lead to a significant mismatch between managerial skill and scale of actively managed mutual funds.

[3] See the appendix for the complete list of factors.

quarterly factor returns ($t > 3$). The magnitude is economically meaningful as it is more than twice as much as the average quarterly factor return in our sample (0.48% per quarter). The effects have similar magnitudes even after controlling for factor fixed effects and time fixed effects. In short, these results indicate that the flow-driven price impact does not wash out at the factor level, but rather largely determines factor returns.

If mutual fund flow-induced demand shocks are non-fundamental (Ben-David, Li, Rossi, and Song, 2018), the aforementioned flow-driven return patterns would revert over time. This is indeed what we find in the data. For example, when we sort factors by the past-eight-quarter FITOF, the top-third factors that have experienced the highest past flow-induced trades significantly *underperform* the bottom-third factors with the lowest past flows.[4] The average Fama-French-Carhart-alpha spread between these two groups is 1.95% over the next quarter ($t =$ 4.00). We also confirm that this flow-driven return reversal cannot be explained by statistical mean-reversion in factor returns.[5] Moreover, we find that this flow-driven factor return reversal can explain the underperformance of large-sized mutual funds relative to small funds (Chen, Hong, Huang, and Kubik, 2004; Pollet and Wilson, 2008).

Having demonstrated a large flow-driven effect on factor return dynamics, we then link variation of flow-induced trading to factor volatility and factor comovement.[6] Forecasting the variance-covariance matrix of factor returns is also of practical importance as one may use this information to construct portfolios of factors (Moskowitz, 2003), especially given the growing popularity of factor or smart-beta investing. To this end, we extend the idea of Greenwood and Thesmar (2011) to estimate the expected variance and expected covariance of flow-induced trading of factors, which we call "factor fragility" and "factor co-fragility," respectively.

[4]The flow-driven return reverals are robust if we use different time windows.

[5]When we sort factors by the past-eight-quarter factor returns, we don't find any difference in subsequent returns.

[6]Greenwood and Thesmar (2011), Lou (2012), and Anton and Polk (2014) show that mutual fund ownership structure plays an important role in determining volatilities and comovement among individual stocks.

We find that factor fragility strongly forecasts factor return variance, while factor co-fragility strongly forecasts factor return covariance. Here, factor volatility and covariance are calculated by daily factor returns over the subsequent quarter. For example, in the univariate panel regression with quarterly frequency, a one-standard-deviation increase in the normalized co-fragility (normalized by the square root of fragility) is associated with an increase of 56% of a standard deviation in the future correlation between factors. Moreover, factor co-fragility alone can explain 20% variation in factor return comovement. We also show that the results are robust if we further control for the lagged factor (co)variance and factor-pair and time fixed effects. In short, our results indicate that factor variation and comovement are influenced by demand shifts of retail mutual fund investors. In the spirit of De Long, Shleifer, Summers, and Waldmann (1990), these factors are heavily exposed to "noise trader" risk.

A natural question is whether the aggregate non-fundamental flow-driven risk is "priced" by liquidity providers and other sophisticated investors. Following the argument of Pollet and Wilson (2010), an increase in aggregate flow-driven risk, other things equal, is likely to be revealed by an increase in the comovement among flow-driven trades of different factors.[7] Thus, if the flow-driven noise trader risk is being priced, then the average covariance of flow-induced trades of factors should positively predict future average factor premia. This is what we find in the data. In the time-series regressions, we find that average factor co-fragility significantly and positively forecast average factor premia over the next quarter. More importantly, a horse-race test shows that factor co-fragility can fully subsume the power of average factor return covariance in predicting future factor premia.

We also discuss other implications of flow-driven effects on factor return dynamics. Gupta and Kelly (2019) and Arnott, Clements, Kalesnik, and Linnainmaa (2019) show that factors

[7]Pollet and Wilson (2010) use the average covariance between daily stock returns to forecast subsequent average stock returns. They argue that when assets have high covariance, investors face high risk in their portfolio and require high return.

exhibit return momentum similar to that found in stock returns (Jegadeesh and Titman, 1993). We provide a flow-based explanation of factor momentum. Intuitively, for those winning factors, stocks in their long legs are performing well recently, generating superior performance of mutual funds that concentrate in these stocks. Because investors hardly account for any factor exposures (Ben-David, Li, Rossi, and Song, 2018), these stocks are likely to experience more flow-induced buyings, leading to higher factor returns in the near future.[8]

To examine the extent to which mutual fund flow-induced trading can explain factor return momentum, we measure the expected flow-induced trades of factors ($\mathbb{E}$[FITOF]) based on predicted fund flows using past fund returns and past flows as predictors. Then we conduct a horse race between past factor returns and $\mathbb{E}$[FITOF] to forecast future factor returns. After controlling for $\mathbb{E}$[FITOF], the positive coefficients on lagged factor returns drop by 25% to 40%, depending on the length of the holding period. In other words, the mutual funds' flow-driven effect on factor returns is an important driver of factor momentum.

This paper contributes to the asset pricing literature in several ways. First, our paper is closely related to the recent literature that investigates the seemingly high dimensionality of cross-sectional asset pricing models. Examples include Baz, Granger, Harvey, Le Roux, and Rattray (2015), Harvey, Liu, and Zhu (2016), Harvey (2017), McLean and Pontiff (2016), Hou, Xue, and Zhang (2018), Kozak, Nagel, and Santosh (2017), Kelly, Pruitt, and Su (2018), Feng, Giglio, and Xiu (2019), among some others.[9] We offer a new perspective by emphasizing the important influence of the non-fundamental mutual funds' flow-driven trades on returns and comovement among these factors. Our results highlight that asset pricing factors are heavily exposed to non-fundamental risk or "noisy trader risk" in the terminology of De Long, Shleifer, Summers, and Waldmann (1990).

[8]Mutual fund managers typically scale up old positions with new inflows or liquidate existing positions with outflows. Lou (2012) shows that persistence in fund flows can partially explain stock price momentum.

[9]See Feng, Giglio, and Xiu (2019) for a more comprehensive litereature review.

A vast empirical literature shows that investor demand unrelated to fundamentals can impact asset prices. For example, Coval and Stafford (2007), Frazzini and Lamont (2008), and Lou (2012) show that mutual fund flow-induced demand shocks have considerable price impact on individual stock prices.[10] In an earlier paper, Huang, Song, and Xiang (2019) connect mutual fund flow-induced trades and the Fama-French size and value factors. This paper complements the prior literature by comprehensively analyzing the impact of non-fundamental demand shocks on a large collection of asset pricing factors.

The rest of the paper is organized as follows. Section 2 introduces the dataset, the set of asset pricing factors, and our measure of flow-induced trading of factors (FITOF). Section 3 analyzes how demand shifts induced by the uninformative fund flows affect factor return dynamics. Section 4 links (co)variance of fund-induced trades of factors to factor return (co)variance and shows that the non-fundamental flow-driven risk is significantly priced. Section 5 studies implications of the flow-driven return dynamics on factor momentum and mutual fund performance. Section 6 provides a concluding remark. Robustness checks and supplementary results are in the appendices.

2 Data and Methodology

In this section, we describe the dataset, the list of asset pricing factors (anomalies), and how we estimate mutual fund flow-induced trades of factors.

2.1 Factor construction

We use CRSP and Compustat to construct 50 asset pricing factors largely following Linnainmaa and Roberts (2018) and Arnott, Clements, Kalesnik, and Linnainmaa (2019). Table 1 shows the list of factors. Our sample stocks include all ordinary common shares (CRSP share code 10

[10] Other examples include Shleifer (1986), Harris and Gurel (1986), Mitchell, Pedersen, and Pulvino (2007), Gârleanu, Pedersen, and Poteshman (2009), Barber, Odean, and Zhu (2009), Greenwood and Vayanos (2010), Foucault, Sraer, and Thesmar (2011), and Li (2018), among many others.

or 11) listed on NYSE, AMEX, and NASDAQ. To avoid microstructure issues, we follow Hou, Xue, and Zhang (2018) to form quintile portfolios based on NYSE breakpoints. Five factors are constructed based on discrete portfolio sorting variables (e.g., debt issuance status). Then, we long and short the two extreme quintiles and track value-weighted returns of the long-short portfolio.[11]

Specifically, the universe of factors consists of 38 annually or quarterly rebalanced factors based on firm fundamentals and 12 monthly rebalanced factors based on stock prices, returns, or trading volume. For annually rebalanced factors, at June-end of each calendar year, we sort all stocks into quintiles based on NYSE breakpoints of sorting variables (e.g., book-to-market ratio) at the fiscal year ending in the previous calendar year, and then we track value-weighted portfolio returns from this July to next June. For quarterly rebalanced factors, we skip one quarter between the data date and the start of the portfolio holding period to ensure all information is available upon portfolio formation. Specifically, at the end of each quarter, we form quintile portfolios based on sorting variables as of fiscal quarter ending in the previous calendar quarter, and then we hold the portfolios in the next calendar quarter. For monthly rebalanced factors, we form quintile portfolios at the end of each month using stock trading data in and before this month-end, and then we hold the portfolios in the next month.

2.2 Mutual fund flow-induced trades of factors

To measure flow-induced trades of each of the 50 factors, we first estimate flow-induced trades of individual stocks. To this end, we merge the Thomson Reuters CDA/Spectrum database with the CRSP Survivorship-bias-free mutual fund database. In particular, we obtain mutual funds' holding data from the CDA/Spectrum database. Mutual funds' total net assets (TNA), monthly net returns (after fee), and annual expense ratios are from the CRSP database. For mutual funds

[11] Hou, Xue, and Zhang (2018) replicate 447 factors (anomalies) through forming decile portfolios based NYSE breakpoints. Our focus is not replication of factor returns, and instead, we form quintile portfolios to avert conducting analysis on thin testing portfolios.

Table 1: **List of factors.** This table presents the list of factors studied in this paper. The details of factor construction are in the Appendix.

Fundamentals based factors		26	Operating profitability
1	Abnormal capital investment	27	Piotroski's F-score
2	Accruals	28	Profit margin
3	Asset Growth	29	QMJ Profitability
4	Book-to-market	30	Return on assets
5	Cash-based profitability	31	Return on equity
6	Cashflow-to-price	32	Sales growth
7	Change in asset turnover	33	Sales-to-price
8	Debt issuance	34	Sales-minus-inventory growth
9	Distress risk	35	Size
10	Earnings-to-price	36	Sustainable growth
11	Enterprise multiple	37	Total external financing
12	Five-year share issuance	38	Altman's Z-score
13	Gross profitability	**Non-Fundamentals based factors**	
14	Growth in Inventory	39	52-week high
15	Industry Concentration	40	Firm age
16	Industry adjusted CAPX growth	41	Heston and Sadka's seasonality
17	Investment growth	42	High-volume return premium
18	Investment-to-assets	43	Idiosyncratic volatility
19	Investment-to-capital	44	Intermediate momentum
20	Leverage	45	Long-term reversals
21	M/B and accruals	46	Market beta
22	Net operating assets	47	Maximum daily return
23	Net working capital changes	48	Momentum
24	Ohlson's O-score	49	Nominal Price
25	One-year share issuance	50	Short-term reversals

with multiple share classes, we use the sum of TNA across all share classes as the TNA of the fund, and we take TNA-weighted average net returns and expense ratios across all share classes. We compute mutual fund monthly gross returns (before fee) as the sum of monthly net returns and 1/12 of the annual expense ratio.

We focus on actively-managed equity mutual funds. Specifically, we filter out non-equity funds based on investment objective codes reported in the CDA/Spectrum database and CRSP mutual fund database.[12] In addition, we require the ratio of common stock holdings to TNA to be between 80% and 105% on average over the sample period. Finally, we exclude fund-quarter observations with less than $1 million TNA. Our fund sample includes $4,999$ distinct US domestic equity funds with $208,419$ fund-quarter observations during 1980-2017.

We take two steps to construct the stock-level flow-induced trading. We first calculate quarterly mutual fund flows, which is defined as the percentage change of total net assets after adjusting for appreciation of fund holdings (Sirri and Tufano (1998)):

$$\text{Flow}_{k,t} = \frac{\text{TNA}_{k,t} - \text{TNA}_{k,t-1} \times (1 + R_{k,t})}{\text{TNA}_{k,t-1}},$$

where $\text{TNA}_{k,t}$ is the total net assets of fund k at the end of quarter t and $R_{k,t}$ is the gross return of fund k in quarter t.

Second, we measure quarterly aggregate mutual fund trading of each individual stock in

[12] We follow Kacperczyk, Sialm, and Zheng (2008) to screen funds in the following steps. First, we screen funds by investment objectives reported by CDA/Spectrum database. We exclude funds with Investment Objective Codes in 1,5,6,or 7, in CDA/Spectrum database. Then, we screen funds by investment objectives reported by CRSP mutual funds database. For funds with non-missing "Type of Securities Mainly Held by Fund" variable (policy variable), we remove those with policy in C&I, Bal, Bonds, Pfd, B&P, GS, MM, or TFM. We then require remaining funds to have Lipper Classification Code in EIEI, G, LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, SCVE, or Missing. For funds with missing Lipper Classification Code, we require them to have Strategic Objective Insight Code in AGG, GMC, GRI, GRO, ING, SCG, or missing. If a fund has both missing Lipper Classification Code and Strategic Objective Insight Code, we screen them through Wiesenberger Fund Type Code and retain funds with objective codes in G, G-I, AGG, GCI, GRI, GRO, LTG, MCG, SCG, or Missing.

response to fund flows. We use the flow-induced-trading (FIT) measure of Lou (2012):

$$\text{FIT}_{j,t} = \frac{\sum_k \text{Shares}_{k,j,t-1} \times \text{Flow}_{k,t} \times \text{PSF}}{\sum_k \text{Shares}_{k,j,t-1}}, \tag{1}$$

where $\text{Shares}_{k,j,t-1}$ is the number of shares of stocks j held by fund k at the end of quarter $t-1$, $\text{Flow}_{k,t}$ is the percentage flow of fund k in quarter t, and PSF is the partial scaling factor. The scaling factor reflects how fund managers, on average, increase and liquidate their holdings in response to capital inflows and outflows, respectively. Lou (2012) estimates PSF to be 0.970 for outflows and 0.858 for inflows, and we use the same estimates of PSF in our study. Moreover, we use FIT rather than the entire realized trading of mutual funds because FIT only captures those trades that are driven by the demand shifts from mutual fund investors, which are largely ignorant about fundamentals (Ben-David, Li, Rossi, and Song, 2018).

Based on stock-level flow-induced trading, we measure flow-induced trading of a factor π as the value-weighted average FIT of stocks in the factor's long leg minus the value-weighted average FIT of stocks in the short leg. That is,

$$\text{FITOF}_{\pi,t} = \sum_{j \in \mathcal{N}_L^\pi} w_{j,t}^\pi \text{FIT}_{j,t} - \sum_{j \in \mathcal{N}_S^\pi} w_{j,t}^\pi \text{FIT}_{j,t}, \tag{2}$$

where $\mathcal{N}_L^\pi$ and $\mathcal{N}_S^\pi$ are the set of stocks consisting of the long-leg and short-leg of factor π at time t, respectively, and $w_{j,t}^\pi$ is the weight of stock j in factor π. In short, FITOF measures the flow-induced trades of the long-leg stocks relative to the flow-induced trades of the short-leg stocks.

Panel A of Table 2 shows that the 25th and 75th percentiles of the stock-level FIT are -1.95% and 3.02%, respectively. This suggests that, in response to retail investors' demand shifts, mutual funds adjust their stock holdings relative to their existing holdings at a scale between -1.95% and 3.02% within a quarter in the 25th to 75th percentile range. From Panel B of Table 2, the 25th and 75 percentiles of FITOF are -0.53% and 0.55%, respectively.

Table 2: **Summary Statistics.** This table reports the summary statistics of our mutual fund and stock datasets. The sample period is 1980Q2-2017Q4. Panel A reports the summary statistics of the US equity mutual funds in our study. # Funds is the number of distinct mutual funds in each period. TNA is the average fund total net assets (in million $). Equity Holdings is the average total market capitalization of the common stock holdings (in million $). Market Held (%) is the average percentage of the US common stocks held by the mutual funds in our sample. Panel B reports the stock and factor characteristics. Size and book-to-market ratio of our sample stocks are shown in NYSE percentiles. Stock-level flow-induced trading (FIT) is defined in (1). Flow-induced trading of factor (FITOF) is defined as the value-weighted FIT of a factor's long-leg stocks minus that of the short-leg stocks in (2). The list of factors is shown in Table 1.

Panel A: Summary statistics of mutual funds

Period	#Funds	TNA		Equity Holdings		Market Held (%)
		Median	Mean	Median	Mean	
1980-1984	370	64.43	159.62	39.00	111.31	2.62
1985-1989	610	79.18	264.52	57.13	198.09	4.33
1990-1994	1,453	71.81	299.02	47.60	215.16	7.51
1995-1999	2,699	110.79	698.13	69.79	537.19	13.19
2000-2004	3,461	120.31	837.97	71.42	655.13	15.35
2005-2009	3,636	172.73	1,097.92	104.06	799.67	18.52
2010-2014	2,875	297.49	1,664.22	157.97	1,172.03	18.61
2014-2017	2,216	479.88	2,757.40	287.47	2,210.29	20.07

Panel B: Summary statistics of stocks and factors

Variables	Mean	SD	Q1	Median	Q3
Stock-level:					
Size	0.3105	0.2923	0.0547	0.2143	0.5190
Book-to-Market	0.4839	0.3033	0.2088	0.4801	0.7534
FIT	0.0157	0.1196	−0.0195	0.0017	0.0302
Factor-level:					
Annualized Return	0.0192	0.2590	−0.1053	0.0196	0.1450
FITOF	0.0003	0.0177	−0.0053	0.0001	0.0055

3 Flow-induced trading and factor return dynamics

In this section, we examine how mutual fund flow-induced demand shifts influence factor returns. Although some earlier work (e.g., Coval and Stafford (2007) and Lou (2012)) documents that mutual fund flows generate price pressure on individual stocks, it is not clear whether flow-induced price impact would cancel out at the factor level, given that factors are well-diversified long-short stock portfolios. In the following analysis, we show that the non-fundamental flow-induced trades significantly determine contemporaneous factor returns, followed by strong reversals over long term.

3.1 Flow-induced trading and contemporaneous factor returns

We start by showing a significant contemporaneous relationship between flow-induced trading of factors and factor returns. Based on mounting evidence that mutual fund flows generate positive price pressure on individual stocks, we hypothesize that a factor should have higher returns when its long-leg stocks experience more flow-induced buying relative to the stocks of the short-leg. As described in Section 2, we gauge flow-induced demand shocks to each factor by aggregating mutual funds' flow-induced trading of individual stocks following equation (2). We then estimate the contemporaneous correlation between quarterly factor returns and FITOF through panel regressions. Table 3 reports the results.

Column (1) of Table 3 is based on univariate regression of factor returns on FITOF. As one can see, the coefficient estimate is 0.68 with a t-statistic of 3.29. It implies that a one-standard-deviation increase in FITOF is associated with an increase of 118 bps in quarterly factor return. This magnitude is economic meaningful as it is more than twice as much as the average quarterly factor returns, which is around 48 bps per quarter (see Table 2). In Columns (2) to (6), we further control for factor fixed effects, time fixed effects, and the predictors of market return in Welch and Goyal (2007). The coefficient estimates are similar and are highly significant across

Table 3: **Flow-induced trading of factors and contemporaneous factor returns.** This table reports panel regressions of quarterly factor returns on the contemporaneous flow-induced trading of factor (FITOF). The sample period is 1980Q2-2017Q4. The independent variable FITOF is the value-weighted flow-induced trading (FIT) of a factor's long-leg stocks minus that of the short-leg stocks (see (2)). Control variables include dividend yield, earnings-price ratio, market volatility, book-to-market ratio, three-month T-bill rate, long-term yield, net equity expansion, consumer price index, investment-to-capital ratio, and default yield spread as in Welch and Goyal (2007). Standard errors are double clustered by factor and time. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

DepVAR: Ret	(1)	(2)	(3)	(4)	(5)	(6)
FITOF	0.68***	0.66***	0.73***	0.72***	0.67***	0.73***
	(3.29)	(3.23)	(3.26)	(3.19)	(3.20)	(3.17)
Controls	No	Yes	No	Yes	No	No
Factor FE	No	No	Yes	Yes	No	Yes
Time FE	No	No	No	No	Yes	Yes
No. Obs.	7,550	7,550	7,550	7,550	7,550	7,550
Adjusted R^2	0.03	0.07	0.04	0.07	0.08	0.09

all difference specifications.

Ben-David, Li, Rossi, and Song (2018) show that mutual fund investors are largely ignorant about both systematic and idiosyncratic risks when allocating capital among funds. If mutual fund flow-induced trading is largely non-fundamental, the aforementioned positive impact of flow-induced trading on factor returns would revert subsequently. This is indeed what we find in the next section.

3.2 Flow-induced factor return reversals

In this section, we document a strong flow-induced factor return reversal over longer horizons. To see this, at the end of each quarter t, we sort the 50 factors into three groups based on the average FITOF over the past eight quarters (including quarter t). The top and bottom groups contain 15 factors and the middle group has 20 factors.[13] We then calculate the equally-weighted monthly returns for each factor portfolio over quarter $t+1$. We evaluate return performance of each of the three factor portfolios and report the results in Table 4. In the untabulated exercise,

[13] If we sort factors into finer groups, we get even stronger results.

we also form portfolios based on the past four-quarter or past twelve-quarter FITOF, and we get similar results.

Table 4: **Performance of factor portfolios sorted on the past-eight-quarter FITOF.** Panel A reports monthly performance of the three factor portfolios formed by the past-eight-quarter flow-induced trading of factor (FITOF). FITOF is the value-weighted flow-induced trading (FIT) of a factor's long-leg stocks minus that of the short-leg stocks (see (2)). Specifically, at the end of each quarter t, we sort the 50 factors into three groups (15/20/15) based on their average FITOF over the past eight quarters (including quarter t). We then track the equally-weighted returns for the three groups of factors in quarter $t+1$. Panel B reports monthly performance of the three factor portfolios formed by the past-eight-quarter factor returns. That is, at the end of each quarter t, we sort the 50 factors into three groups (15/20/15) based on their average returns over the past eight quarters (including quarter t). We then track the equally-weighted returns for the three groups of factors in quarter $t+1$. t-statistics are computed based on standard errors with Newey-West correction for twelve lags. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Sort on past-eight-quarter FITOF

Portfolio	Avg Ret	CAPM Alpha	FFC Alpha	FF5+UMD Alpha
Low	0.53***	0.68***	0.52***	0.32***
	(4.07)	(4.87)	(5.62)	(3.91)
Medium	0.19***	0.28***	0.20***	0.10**
	(2.92)	(4.65)	(2.99)	(1.97)
High	−0.03	−0.01	−0.13	−0.08
	(−0.32)	(−0.10)	(−1.50)	(−0.76)
Low-High	0.56***	0.69***	0.65***	0.40**
	(2.90)	(3.31)	(4.00)	(2.31)

Panel B: Sort on past-eight-quarter return

Portfolio	Avg Ret	CAPM Alpha	FFC Alpha	FF5+UMD Alpha
Low	0.25**	0.30**	0.38***	0.23***
	(2.24)	(2.48)	(3.84)	(3.00)
Medium	0.19***	0.27***	0.13***	0.06
	(3.19)	(4.05)	(2.98)	(1.32)
High	0.25**	0.38***	0.11	0.08
	(2.20)	(3.48)	(1.00)	(0.86)
Low-High	−0.01	−0.08	0.28	0.14
	(−0.04)	(−0.39)	(1.55)	(0.92)

Panel A of Table 4 shows significant return spreads between factors with high and low past FITOF. For example, the group of factors with low past FITOF, on average, earns a significantly positive FFC four-factor alpha of 52 bps per month ($t = 5.62$). In contrast, the group of factors with high past FITOF earns a negative average monthly four-factor alpha of -13 bps. The four-factor alpha spread between the two groups is 65 bps per month (7.8% on an annual basis), with a t-statistic of 4.0. The results are similar when portfolio returns are evaluated using average return, CAPM alpha, or alpha adjusted by the Fama-French five factors and the momentum factor.

We emphasize that this flow-driven long-term return reversal is not caused by statistical mean-reversion in factor returns. In panel B of Table 4, we sort factors based on the past-eight-quarter factor returns and we don't find any difference in subsequent factor returns.

We also confirm the flow-driven long-term reversal using panel regressions. That is, we regress quarterly factor returns on the past-eight-quarter FITOF and we report the results in Table 5. As one can see, past FITOF is a significant negative predictor of factor returns in the future. Based on Column (4) with both factor and time fixed effects controlled, a one-standard-deviation increase in the past eight-quarter FITOF is associated with an average decrease in factor returns by 81 basis points over the next quarter.

In sum, the results in this section indicate that mutual fund flow-driven price pressure does not wash out at the factor level. Rather, these non-fundamental demand shifts are statistically strong and economically significant drivers of factor return dynamics.

4 Flow-induced trading, factor volatility, and factor comovement

In this section, we link variation in flow-induced trades to factor return volatility and factor comovement.[14] To this end, we extend the method of Greenwood and Thesmar (2011) to estimate the variance-covariance matrix of flow-induced trades of factors. We find that (i)

[14]Given the growing popularity of factor investing (Ang (2014)), forecasting the variance-covariance matrix of factor returns is of practical importance, as one can potentially use it to construct factor portfolios (Moskowitz (2003)).

Table 5: **Regressions of factor returns on the past-eight-quarter** FITOF**.** This table reports the regressions of quarterly factor returns on the average flow-induced trading of factor (FITOF) over the past eight quarters. FITOF is the value-weighted flow-induced trading (FIT) of a factor's long-leg stocks minus that of the short-leg stocks. The independent variable is the average FITOF over the past eight quarters. Columns (1)-(4) report the panel regression results, where the t-statistics are computed based on standard errors double clustered by factor and time. Column (5) reports the Fama-Macbeth regression result, where t-statistics are computed based on standard errors with Newey-West correction for twelve lags. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

DepVar: Ret	(1)	(2)	(3)	(4)	(5)
Past 8-QTR FITOF	−0.21**	−0.20**	−0.27**	−0.25**	−0.46***
	(−2.20)	(−2.18)	(−2.40)	(−2.40)	(−3.68)
Factor FE	No	No	Yes	Yes	-
Time FE	No	Yes	No	Yes	-
No. Obs.	21,450	21,450	21,450	21,450	21,450
Adjusted R^2	0.00	0.05	0.01	0.05	0.14
Regression Method	Panel	Panel	Panel	Panel	FM

expected volatilities of flow-induced trading strongly forecast factor volatilities and (*ii*) return comovement between two factors is higher when they are expected to have more correlated flow-driven trades. In the spirit of De Long, Shleifer, Summers, and Waldmann (1990), these factors are significantly exposed to "noise trader" risk. Moreover, we provide evidence that this non-fundamental risk is indeed being "priced" by liquidity suppliers and other sophisticated investors (e.g. quant funds).

4.1 Construction of factor fragility and factor co-fragility

We use a bottom-up approach to estimate the variance-covariance matrix of flow-driven trades of factors. Similar to Greenwood and Thesmar (2011) (GT), we assume the following relationship between mutual fund flow-induced trading and return of stock j:

$$r_{j,t} = \alpha_j + \lambda_j \frac{\sum_k \text{Shares}_{k,j,t-1} f_{k,t} \text{PSF}}{\sum_k \text{Shares}_{k,j,t-1}} + \varepsilon_{j,t}. \tag{3}$$

Here, $r_{j,t}$ is the return of stock j in quarter t, $\text{Shares}_{k,j,t-1}$ is the number of shares of stocks j held by fund k at the end of quarter $t-1$, $f_{k,t}$ is the percentage flow of fund k in quarter t, and PSF is the partial scaling factor as in (1). α_j and λ_j are two parameters. In our implementation, we assume that $\lambda_j = \lambda \sum_k \text{Shares}_{k,j,t-1}/\text{Shrout}_{j,t-1}$, where λ is the unconditional price impact factor and $\text{Shrout}_{j,t-1}$ is shares outstanding of stock j at the end of quarter $t-1$. The error term, $\varepsilon_{j,t}$, has a conditional mean of zero and may capture other sources of variation of returns (e.g., news about fundamentals).

Since factors are effectively value-weighted portfolios of stocks, returns of factor π can be expressed as:

$$r_{\pi,t} = \sum_j \mu^{\pi}_{j,t-1} r_{j,t}, \tag{4}$$

where $\mu^{\pi}_{j,t-1}$ is the weight of stock j in factor π in quarter t.[15] Plugging equation (3) into (4), we get

$$r_{\pi,t} = \sum_j \mu^{\pi}_{j,t-1}\alpha_j + \lambda \left(\sum_k w^{\pi}_{k,t-1} f_{k,t} \right) \text{PSF} + \sum_j \mu^{\pi}_{j,t-1}\varepsilon_{j,t}, \tag{5}$$

where $w^{\pi}_{k,t-1} = \sum_j \mu^{\pi}_{j,t-1}\text{Shares}_{k,j,t-1}/\text{Shrout}_{j,t-1}$ can be regarded as the weight of mutual fund k in factor π in quarter t.

Based on equation (5), the conditional variance and covariance of $r_{\pi,t+1}$ at the end of quarter t are

$$\text{Var}_t(r_{\pi,t+1}) = \lambda^2 W^{\pi\prime}_t E_t(\Omega_{t+1}) W^{\pi}_t \text{PSF}^2 + \text{Var}_t\left(\sum_j \mu^{\pi}_{j,t}\varepsilon_{j,t+1} \right) \tag{6}$$

and

$$\text{Cov}_t(r_{\pi_1,t+1}, r_{\pi_2,t+1}) = \lambda^2 W^{\pi_1\prime}_t E_t(\Omega_{t+1}) W^{\pi_2}_t \text{PSF}^2 + \text{Cov}_t\left(\sum_j \mu^{\pi_1}_{j,t}\varepsilon_{j,t+1}, \sum_j \mu^{\pi_2}_{j,t}\varepsilon_{j,t+1} \right), \tag{7}$$

respectively. Here, $E_t(\Omega_{t+1})$ is the conditional variance-covariance matrix of mutual fund

[15] For a long-leg stock, $\mu^{\pi}_{j,t}$ simply equals its original weight in the long leg. For a short-leg stock, $\mu^{\pi}_{j,t}$ is its original weight in the short leg multiplied by negative one.

flows in quarter $t+1$ and $W_t^\pi = \left(w_{1,t}^\pi, \ldots, w_{K,t}^\pi\right)$ is the vector of mutual fund weights in factor π. Because the factors are well-diversified portfolios, $\mathrm{Var}_t\left(\sum_j \mu_{j,t}^\pi \varepsilon_{j,t+1}\right) \approx 0$ and $\mathrm{Cov}_t\left(\sum_j \mu_{j,t}^{\pi_1} \varepsilon_{j,t+1}, \sum_j \mu_{j,t}^{\pi_2} \varepsilon_{j,t+1}\right) \approx 0$.

Similar to GT, we define "factor fragility" of factor π in quarter t as

$$G_t^\pi = W_t^{\pi\prime} E_t(\Omega_{t+1}) W_t^\pi. \tag{8}$$

Likewise, we define co-fragility between factor π_1 and factor π_2 to be

$$G_t^{\pi_1,\pi_2} = W_t^{\pi_1\prime} E_t(\Omega_{t+1}) W_t^{\pi_2}. \tag{9}$$

As one can see from (6) and (7), factor fragility and co-fragility essentially measure the expected variance and expected covariance of flow-driven trades of factors, respectively. Moreover, factor fragility and factor co-fragility depend on mutual fund ownership concentration and the expected variance-covariance matrix of mutual fund flows.

Table 6: **Summary statistics of factor fragility and factor co-fragility.** This table reports the summary statistics of factor fragility and factor co-fragility. Because mutual fund holdings are at quarterly frequency, we use 38 annually- or quarterly-rebalanced factors. The sample period is 1981Q1-2017Q4. Panel A reports factor fragility and factor return volatility, where factor fragility is defined in (8) and factor return volatility is measured as the standard deviation of daily factor returns in a given quarter. Panel B reports pairwise factor return comovement and factor co-fragility. Factor return comovement is measured by covariance or correlation of daily factor return in a given quarter, and factor co-fragility is defined in (9). Normalized Co-fragility between two factors is defined as the co-fragility divided by the product of square root fragility of the two factors.

Variables	Mean	SD	Q1	Median	Q3
Panel A: Factor level					
Standard deviation of daily returns	0.0060	0.0042	0.0037	0.0049	0.0068
Square root of Fragility	0.0008	0.0009	0.0003	0.0005	0.0010
Panel B: Factor-pair level					
Covariance of daily return (10^{-6})	0.2899	4.1099	-0.4525	0.1002	0.7590
Correlation of daily return	0.0608	0.4127	-0.2478	0.0666	0.3767
Co-fragility (10^{-6})	0.1063	2.4459	-0.1190	0.0065	0.1940
Normalized Co-fragility	0.0561	0.9702	-0.7129	0.1212	0.8256

To estimate $E_t(\Omega_{t+1})$, we calculate the variance-covariance matrix of mutual fund flows using observations in the most recent four quarters (including quarter t). Since mutual fund holdings are at quarter frequency, we focus on the 38 annually or quarterly rebalanced factors in the following analyses of factor fragility and factor co-fragility. We report the summary statistics in Table 6.

4.2 Factor fragility and factor volatility

We first examine how factor fragility, the measure of expected variance of flow-induced trades of factors, can forecast factor return variance over the next quarter. In the regressions, the dependent variable is the one-quarter-ahead factor volatility σ^{π}_{t+1}, estimated by the standard deviation of daily factor returns. The independent variable of interest is the square root of factor fragility $\sqrt{G^{\pi}_t}$. We also control for lagged factor volatilities, factor fixed effects, and time fixed effects. Table 7 reports the results.

Under all specifications, we find that the square root of factor fragility, $\sqrt{G_t}$, positively and significantly forecasts the one-quarter-ahead factor volatility. For example, in the univariate regression (Column (1)), a one-standard-deviation increase in $\sqrt{G_t}$ is associated with an increase in σ_{t+1} of 0.14% ($t = 5.66$), which is about 42% of its sample standard deviation. In other words, variation in flow-induced trading is an important determinant of factor return variation.

4.3 Factor co-fragility and factor return comovement

In this section, we show that factor co-fragility, the expected covariance of flow-induced trades of factors, strongly forecasts factor return comovement. To see this, we estimate the

Table 7: **Factor fragility and factor return volatility.** This table reports regressions of factor volatility on the square root of lagged factor fragility. We use 38 annually- or quarterly-rebalanced factors as listed in the Appendix. The sample period is from 1981Q1-2017Q4. Factor volatility is measured by the standard deviation of daily factor returns within a given quarter. Factor fragility, G_t, is defined in (8) and we use the square root of fragility $\sqrt{G_t}$ in the regressions. Columns (1)-(3) report the results based on panel regressions, where t-statistics are computed based on standard errors clustered by factor and by time. Columns (4) and (5) report the results based on the Fama-Macbeth regressions, where t-statistics are computed based on standard errors with Newey-West correction for four lags. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

DepVar: σ_{t+1}	(1)	(2)	(3)	(4)	(5)
$\sqrt{G_t}$	1.53***	0.58***	0.22**	1.78***	0.25***
	(5.66)	(3.73)	(2.31)	(14.21)	(2.98)
σ_t			0.60***		0.51***
			(8.85)		(17.22)
σ_{t-1}			0.09		0.21***
			(1.66)		(6.36)
σ_{t-2}			-0.03		0.12***
			(-0.66)		(3.89)
σ_{t-3}			0.00		0.07**
			(0.07)		(2.46)
Factor FE	No	Yes	Yes	-	-
Time FE	No	Yes	Yes	-	-
No. Obs.	5,586	5,586	5,586	5,586	5,586
Adjusted R^2	0.15	0.71	0.83	0.19	0.73
Regression method	Panel	Panel	Panel	FM	FM

following two regressions:[16]

$$\rho_{\pi_1,\pi_2,t+1} = \alpha + \beta \frac{G_{\pi_1,\pi_2,t}}{\sqrt{G_{\pi_1,t}}\sqrt{G_{\pi_2,t}}} + \gamma Z_{\pi_1,\pi_2,t} + \epsilon_{\pi_1,\pi_2,t+1}, \tag{10}$$

and

$$\sigma_{\pi_1,\pi_2,t+1} = \alpha + \beta G_{\pi_1,\pi_2,t} + \gamma Z_{\pi_1,\pi_2,t} + \epsilon_{\pi_1,\pi_2,t+1}, \tag{11}$$

where $\sigma_{\pi_1,\pi_2,t+1}$ and $\rho_{\pi_1,\pi_2,t+1}$ are the covariance and correlation between factors π_1 and π_2, respectively, and are estimated based on daily factor returns over quarter $t+1$. $G_{\pi_1,\pi_2,t}$ is the co-fragility between factors π_1 and π_2 in (9). We also consider a set of control variables, $Z_{\pi_1,\pi_2,t}$, that include the lagged factor return covariance (correlation) and the factor pairwise difference in size, book-to-market, and momentum.[17] We further control for lagged factor return covariance (correlation), factor-pair fixed effects, and time fixed effects. Table 8 presents the results.

Panel A and Panel B report results of regressions (10) and (11), respectively. For easy interpretation, all variables are normalized to have a standard deviation of one. Across all specifications, factor co-fragility significantly forecasts factor comovement. The economic magnitude is also meaningful. For example, Column (1) of Panel A shows that a one-standard-deviation increase in the normalized factor co-fragility is positively associated with an increase of 55% of a standard deviation in factor return correlation over the next quarter. The explanatory power is also high as the univariate R^2 is around 20% at the quarterly frequency. Factor co-fragility remains highly statistically significant even after we include factor-pair and time fixed effects and various controls.

In sum, the results in Section 3 and Sections 4.2 and 4.3 indicate that these asset pricing

[16]To be consistent with (7), when the dependent variable is factor covariance, the independent variable is co-fragility $G_{\pi_1,\pi_2,t}$. When the dependent variable is $\rho_{\pi_1,\pi_2,t+1}$, we normalize the co-fragility by the product of square root of fragility $\sqrt{G_{\pi_1,t}}$ and $\sqrt{G_{\pi_2,t}}$.

[17]We construct factor pairwise characteristics difference as follows. First, following Anton and Polk (2014), we construct stock-level NYSE percentile ranking of characteristics. Second, we take value-weighted NYSE percentile rankings for long leg and short leg, and compute factor-level NYSE percentile ranking as the long-short difference. The factor pair-level difference is the absolute value of difference in factor-level NYSE percentile ranking of characteristics.

Table 8: **Factor co-fragility and factor return comovement.** This table reports regressions of factor pairwise return comovement on the lagged factor co-fragility ($G_{\pi_1,\pi_2,t}$). The sample period is 1981Q1-2017Q4 and we use all factor pairs among 38 annually- or quarterly-rebalanced factors listed in the Appendix. Panel A reports the results of regression (10), where the dependent variable $\rho_{\pi_1,\pi_2,t+1}$ is the correlation of daily returns between factor π_1 and π_2 over quarter $t+1$. The main independent variable is the co-fragility between the two factors, $G_{\pi_1,\pi_2,t}$, normalized by the product of $\sqrt{G_{\pi_1,t}}$ and $\sqrt{G_{\pi_2,t}}$. Panel B reports the results of regression (11), where the dependent variable $\sigma_{\pi_1,\pi_2,t+1}$ is the covariance of daily returns between factor π_1 and π_2 over quarter $t+1$. The main independent variable is the co-fragility between the two factors ($G_{\pi_1,\pi_2,t}$). Control variables include lagged correlation, lagged covariance, and factor pairwise differences in size, book-to-market, and momentum. Columns (1)-(5) are based on panel regressions and t-statistics in parentheses are clustered by factor-pair and by time. Columns (6) and (7) are based on Fama-Macbeth regressions and t-statistics in parentheses are computed with Newey-West correction with four lags. All variables are standardized to have a standard deviation of one. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Predict factor return correlation by normalized co-fragility

DepVar: $\rho_{\pi_1,\pi_2,t+1}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\frac{G_{\pi_1,\pi_2,t}}{\sqrt{G_{\pi_1,t}}\sqrt{G_{\pi_2,t}}}$	0.55***	0.27***	0.56***	0.27***	0.12***	0.57***	0.14***
	(21.56)	(12.64)	(21.53)	(12.61)	(6.76)	(29.75)	(8.60)
$\rho_{\pi_1,\pi_2,t}$					0.62***		0.76***
					(24.81)		(49.70)
Size Diff					−0.03		−0.04***
					(−1.25)		(−7.55)
BM Diff					0.02		0.02***
					(1.49)		(3.52)
MOM Diff					0.01		0.00
					(0.65)		(0.28)
Pair FE	No	Yes	No	Yes	Yes	-	-
Time FE	No	No	Yes	Yes	Yes	-	-
No. Obs.	102,638	102,638	102,638	102,638	102,638	102,638	102,638
Adjusted R^2	0.19	0.47	0.20	0.48	0.68	0.22	0.70

Panel B: Predict factor return covariance by factor co-fragility

DepVar: $\sigma_{\pi_1,\pi_2,t+1}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$G_{\pi_1,\pi_2,t}$	0.60***	0.50***	0.60***	0.49***	0.22**	1.75***	0.40***
	(4.31)	(3.41)	(4.37)	(3.44)	(2.08)	(7.90)	(4.08)
$\sigma_{\pi_1,\pi_2,t}$					0.64***		0.86***
					(6.10)		(20.03)
Size Diff					−0.02		−0.01**
					(−0.94)		(−2.44)
BM Diff					0.00		0.00
					(0.37)		(0.86)
MOM Diff					0.03**		−0.00
					(2.06)		(−0.09)
Pair FE	No	Yes	No	Yes	Yes	-	-
Time FE	No	No	Yes	Yes	Yes	-	-
No. Obs.	102,638	102,638	102,638	102,638	102,638	102,638	102,638
Adjusted R^2	0.11	0.17	0.13	0.18	0.53	0.24	0.71

factors (anomalies) included in our study are exposed to significant non-fundamental risk, which is due to flow-driven demand shifts from mutual funds.

4.4 The flow-driven non-fundamental risk is priced

Is the aggregate non-fundamental risk due to flow movement "priced" by liquidity suppliers or other sophisticated investors? We provide affirmative evidence in this section.

To answer this question, we resort to the idea of Pollet and Wilson (2010) (PW) and we use the average factor co-fragility to proxy the aggregate flow-driven non-fundamental risk. Following the argument of PW, an increase in aggregate flow-driven risk, other things equal, is likely to be revealed by an increase in the comovement among flow-driven trades of different factors. As a result, if the aggregate non-fundamental flow-driven risk is indeed being priced, then the average factor co-fragility should positively forecast future average returns across factors.

To see this, we estimate time-series regressions of future average premia across factors on the average factor co-fragility. Column (1) of Table 9 shows that average co-fragility positively and significantly predicts the one-quarter-ahead average factor premia. A one-standard-deviation increase in the average co-fragility is significantly associated with an increase of 44 bps in the average factor premia over the next quarter ($t = 2.23$). The predictive power is even stronger if we use the past-four-quarter average co-fragility as the predictor (Column (5)). A one-standard-deviation increase of that predicts an increase of 61 bps in the average factor premia over the next quarter ($t = 3.96$). These magnitudes are economically meaningful as the average factor return is 48 bps per quarter throughout our sample.

We also use the average covariance of factor returns to forecast future average factor premia (Columns (2) and (6)). We find that the average factor return covariance also significantly forecasts the one-quarter-ahead average factor returns. This result is similar to that of Pollet and Wilson (2010), who show that the average covariance of individual stocks predicts average

Table 9: **Average factor co-fragility and future average premium.** This table reports the time-series regressions of average factor returns on the lagged average factor co-fragility defined in (9). The sample period is 1981Q1-2017Q4. The dependent variable is the average quarterly returns of 38 annually- or quarterly-rebalanced factors in a given quarter $t+1$. In Columns (1) to (4), the independent variable is the average pairwise factor co-fragility in a given quarter t. In Columns (5) to (8), the independent variable is the past-four-quarter average co-fragility across all factor pairs. The control variables include the lagged average factor covariance, the lagged average factor variance, the market excess returns, the market dividend-price ratio, and the average factor returns. t-statistics in parentheses are computed based on standard errors with Newey-West correction for four lags. For easy interpretation, all independent variables are standardized by their sample standard deviations. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Avg Factor Ret_{t+1}	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Avg Co-fragility_t	0.44**		0.36*	0.46**				
	(2.23)		(1.94)	(2.06)				
Avg $\text{Co-fragility}_{t-3,t}$					0.61***		0.67**	0.82**
					(3.96)		(2.47)	(2.56)
Avg Covariance_t		0.37**	0.23	0.54				
		(2.51)	(1.63)	(1.53)				
Avg $\text{Covariance}_{t-3,t}$						0.32*	−0.10	−0.02
						(1.83)	(−0.35)	(−0.06)
MKTRF_t				0.21				0.15
				(1.04)				(0.78)
D/P_t				−0.23**				−0.27**
				(−2.03)				(−2.31)
Avg Variance_t				−0.40				−0.28
				(−1.19)				(−1.17)
Avg Factor Ret_t				0.12				0.09
				(0.44)				(0.41)
No. Obs.	146	146	146	146	146	146	146	146
Adjusted R^2	0.05	0.04	0.06	0.09	0.11	0.03	0.11	0.13

stock returns.

However, in the horse-race tests (Columns (3) and (7)), we find that the average co-fragility can subsume the predictive power of the average factor covariance, but not vice versa. That is, the coefficients of the average covariance become insignificant after controlling for the average factor co-fragility. In Columns (4) and (8), we further control for the average variance of factor returns, market excess returns, market dividend-price ratio, and average factor return. The predictive power of the average factor co-fragility remains statistically and economically significant.

In short, our results have demonstrated that the asset pricing factors included in this study are susceptible to non-fundamental risk due to demand shifts from mutual fund investors and that this "noise trader" risk is being priced by liquidity providers or other sophisticated investors.

5 Implications of flow-driven factor return dynamics

In this section, we discuss two implications of the flow-induced effects on factor return dynamics. First, we analyze the role of mutual fund flow-induced trading in driving factor momentum. Second, we link the flow-induced long-term reversals of factor returns in Section 3 to the underperformance of large-sized mutual funds relative to small-sized funds.

5.1 A flow-based explanation of factor momentum

Recent work of Gupta and Kelly (2019) and Arnott, Clements, Kalesnik, and Linnainmaa (2019) shows that return momentum also exists at the factor level. That is, factors performing well in the recent past continue to outperform poorly-performing factors in the next one to three months. Moreover, factor momentum cannot be explained by either stock momentum (Jegadeesh and Titman, 1993) or industry momentum (Grinblatt and Moskowitz, 1999).[18]

[18] Arnott, Clements, Kalesnik, and Linnainmaa (2019) show that factor momentum subsumes industry momentum, not vice versa.

In this section, we propose a flow-based explanation for factor momentum. The intuition is as follows. For those winning factors, stocks in their long legs perform well recently, so do mutual funds that concentrate in these stocks. Because mutual fund investors don't account for any factor exposures and chase unadjusted returns (Ben-David, Li, Rossi, and Song, 2018), these funds are likely to receive additional capitals, with which they largely scale up existing holdings.[19] Due to the positive price pressure, stocks in the long leg of the winning factors will experience higher returns in the near future. Consequently, the winning factors continue to perform well. The opposite effect would apply to those losing factors.

To examine the extent to which our proposed mechanism can explain factor momentum, we follow the similar idea of Lou (2012) and we run a horse-race test between the expected flow-induced trading of factors ($\mathbb{E}$[FITOF]) and recent factor returns in predicting future factor returns. Based on (1) and (2), the expected flow-induced trades of factor π is given by

$$\mathbb{E}_t[\text{FITOF}_{\pi,t+1}] = \sum_{j\in\mathcal{N}_L^\pi} w_{j,t}^\pi \mathbb{E}_t\left[\text{FIT}_{j,t+1}\right] - \sum_{j\in\mathcal{N}_S^\pi} w_{j,t}^\pi \mathbb{E}_t\left[\text{FIT}_{j,t+1}\right], \tag{12}$$

where

$$\mathbb{E}_t[\text{FIT}_{j,t+1}] = \frac{\sum_k \text{Shares}_{k,j,t} \times \mathbb{E}_t[\text{Flow}_{k,t+1}] \times \text{PSF}}{\sum_i \text{Shares}_{k,j,t}}$$

and $\mathbb{E}_t[\text{Flow}_{k,t+1}]$ is the expected flow to mutual fund k at the end of quarter t. In particular, we use past 12-month fund performance and flows to predict future fund flows.

Table 10 reports the results of the horse-race tests using Fama-Macbeth regressions. Specifically, in Panel A of Table 10, the dependent variable is the factor return at month $m+1$ and the independent variables include the factor return at month m and the most recent $\mathbb{E}$[FITOF] before month $m+1$. ($\mathbb{E}$[FITOF] is estimated at a quarterly frequency.) In Panel B, the dependent variable is the factor return at quarter $t+1$ and the independent variables include the factor return at quarter t and $\mathbb{E}$[FITOF] at the end of quarter t.

[19] See the discussions in Section 2.

Table 10: **Expected flow-induced trading and factor momentum.** This table reports the Fama-Macbeth regressions of factor returns on past factor returns and the expected flow-induced trading of factors ($\mathbb{E}$[FITOF]) defined in (12). The sample period is from 1980-2017. We use the 50 factors listed in Table 1. In Panel A, the dependent variable is factor return in month $m+1$ and the independent variables are the factor return in month m and the most recent $\mathbb{E}$[FITOF] as of month m end. In Panel B, the dependent variable is the factor return in quarter $t+1$ and the independent variables are the factor return in quarter t and $\mathbb{E}$[FITOF] at quarter t end. The t-statistics in Panel A and Panel B are computed based on standard errors with Newey-West correction for four lags and one lag, respectively. *, **, *** indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A: Monthly frequency			
DepVar: Ret_{m+1}	(1)	(2)	(3)
Ret_m	0.12***		0.09***
	(4.44)		(3.83)
$\mathbb{E}$[FITOF]		0.37**	0.37***
		(2.24)	(2.68)
No. Obs.	22,600	22,600	22,600
Adjusted R^2	0.20	0.20	0.34
Panel B: Quarterly frequency			
DepVar: Ret_{t+1}	(1)	(2)	(3)
Ret_t	0.10**		0.06
	(2.26)		(1.60)
$\mathbb{E}$[FITOF]		0.37**	0.23
		(2.19)	(1.36)
No. Obs.	7,500	7,500	7,500
Adjusted R^2	0.16	0.20	0.31

From the first two columns of Panel A, one can see that the past one-month factor return and $\mathbb{E}$[FITOF], each alone, positively predicts factor returns in the next month. For example, a one-standard-deviation increase in $\mathbb{E}$[FITOF] is associated with an increase in monthly factor returns of 0.34% ($t = 2.24$). If we include both the past-one-month factor return and $\mathbb{E}$[FITOF], the coefficient estimate of the past one-month return drops by more than 25%, while the coefficient estimate of $\mathbb{E}$[FITOF] is unchanged. The first two columns of Panel B show that both the past-one-quarter factor return and $\mathbb{E}$[FITOF] positively forecast factor returns over the next quarter. Comparing Columns (1) and (3), we find that $\mathbb{E}$[FITOF] subsumes 40% of the coefficient estimate of the past one-quarter return (0.10 vs 0.06).

In short, the horse-race tests suggest that the flow-based mechanism can partially explain factor momentum.

5.2 The underperformance of large-sized funds relative to small funds

Despite the debate of whether a fund's performance decreases with its own size,[20] in the cross-sections, large-sized mutual funds underperform small funds (Chen, Hong, Huang, and Kubik, 2004; Pollet and Wilson, 2008). In this section, we offer a possible explanation of this empirical pattern based on the long-term flow-induced factor return reversals documented in Section 3.

To investigate how flow-induced factor reversals can explain the underperformance of large-sized funds relative to small funds, we take several steps: First, we construct return spread between small-sized and large-sized mutual funds. Each month we sort mutual funds into quintiles based on fund size at previous month-end and we calculate the equal-weighted gross returns over the next month.[21] The return difference between smallest and largest quintile is dubbed

[20] See, for example, Reuter and Zitzewitz (2015), Pástor, Stambaugh, and Taylor (2015), and Harvey and Liu (2017b).

[21] We retain mutual funds with CRSP objective code starting with "ED" and require mutual funds to have TNA greater than \$5 million and have 80%-105% of their TNA invested in equity as of portfolio formation.

as FSMB ("fund small-minus-big"). Second, we construct return spread between the two factor portfolios with low and high past-eight-quarter FITOF as in Section 3. The return spread is dubbed as FLMH ("factor low-minus-high"). Specifically, FLMH is the monthly return of a portfolio that longs the 15 factors with the lowest FITOF over the past eight quarters and shorts the 15 factors with the highest past FITOF. We then run spanning regressions of FSMB on various factors including FLMH. Table 11 reports the results.

Table 11: **Flow-induced factor return reversals and mutual fund performance.** This table reports spanning regressions of fund return spread on various factors. The sample period is 1982-2017. Each month we sort equity funds into quintiles based on their assets under management (AUM) at the previous month end and track the equal-weighted gross returns of the quintile portfolios. We define FSMB, "fund small-minus-big," as the return spread between the bottom-AUM and the top-AUM quintiles in each month. The monthly return of the "factor low-minus high," FLMH, is calculated as in Section 3. Specifically, each month we long the 15 factors with the lowest past-eight-quarter FITOF and we short the 15 factors with the highest average FITOF over the past eight quarters. We define FLMH as the return of this long-short factor portfolios. Other right-hand-side factors are the Fama and French (2015) five factors (MKTRF, SMB, HML, RMW, and CMA) and the Carhart (1997) momentum factor (UMD). The intercepts are shown in basis points. The parameter estimates with 90% significance are highlighted in bold.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DepVar:	FSMB	FSMB	FSMB	FSMB	FSMB	FSMB	FLMH	FLMH	FLMH
Intercept	**5.70**	3.60	**4.14**	1.27	**4.17**	2.88	**46.93**	**47.68**	**31.29**
	(2.05)	(1.34)	(1.87)	(0.60)	(2.10)	(1.48)	(3.06)	(3.51)	(2.17)
FLMH		**0.04**		**0.05**		**0.03**			
		(3.27)		(5.34)		(3.88)			
FSMB							**1.30**	**2.14**	**1.62**
							(3.07)	(4.70)	(3.22)
MKTRF			**-0.02**	**-0.01**	**-0.02**	**-0.02**		-0.04	0.00
			(-2.48)	(-2.01)	(-2.66)	(-2.61)		(-0.90)	(0.04)
SMB			**0.10**	**0.11**	**0.12**	**0.12**		**-0.52**	**-0.35**
			(5.17)	(6.15)	(9.94)	(9.85)		(-5.28)	(-3.25)
HML			**0.08**	**0.06**	**0.04**	**0.03**		0.10	0.04
			(5.98)	(5.44)	(3.09)	(2.74)		(1.27)	(0.41)
RMW					**0.07**	**0.05**			**0.41**
					(4.70)	(2.96)			(4.45)
CMA					0.02	0.02			0.06
					(1.39)	(1.22)			(0.49)
UMD					**-0.05**	**-0.04**			-0.03
					(-8.69)	(-7.89)			(-0.67)
No. Obs.	429	429	429	429	429	429	429	429	429
Adjusted R^2	0.00	0.05	0.34	0.41	0.49	0.52	0.05	0.26	0.31

Column (1) of Table 11 shows that the mean of FSMB is 5.7 bps per month ($t = 2.05$). It means that, on average, the top-quintile-sized funds underperform the bottom-quintile-sized funds by 70 bps each year. After controlling for FLMH (Column (2)), the intercept largely drops and becomes statistically insignificant with only 3.6 bps per month ($t = 1.34$). From Columns (3) to (6), if we measure fund performance based on the Fama-French three-factor alpha or the six-factor alpha, we still find that FLMH can largely explain the performance difference between the two fund groups. Besides, FSMB cannot subsume FLMH (Columns (7)-(9)). The intercepts of FLMH are statistically significant in all specifications. In fact, FSMB can only explain 9 bps in FLMH (the raw return of FLMH is 56 bps per month).

In short, the results in Table 11 suggest that the flow-driven factor return reversal of FITOF can largely explain the performance difference between small-sized and large-sized mutual funds.

6 Conclusion

In this paper, we study how mutual funds' trades induced by non-fundamental flow movement affect returns and comovement among 50 well-studied asset pricing "factors." Using a bottom-up approach, we estimate mutual funds' flow-induced trading of factors and have several novel findings: First, flow-induced trades positively affect contemporaneous factor returns, followed by strong reversals over longer horizons. This indicates that the flow-driven price impact (Coval and Stafford, 2007; Lou, 2012) does not cancel out at the factor level but rather largely influence factor return dynamics. We then link variation in flow-induced trades to factor volatility and factor comovement. We find that the expected volatility of flow-driven demand shifts strongly forecasts factor return volatility and the expected correlation of flow-induced trades of factors strongly predicts factor correlation. Our results indicate that these asset pricing factors are heavily exposed to non-fundamental risk that is due to mutual funds' flow-driven demand shifts.

We further explore the asset pricing implications of our findings. In particular, we show that the flow-driven non-fundamental risk is significantly priced by liquidity providers or other sophisticated investors who trade these factors. We also show that the flow-driven effects on factor return dynamics can partially account for the recently documented factor momentum and can largely explain the underperformance of large mutual funds relative to small-sized funds.

References

Abarbanell, Jeffery S., and Brian J. Bushee, 1998, Abnormal returns to a fundamental analysis strategy, *Accounting Review* pp. 19–45.

Ang, Andrew, 2014, *Asset management: A systematic approach to factor investing* (Oxford University Press).

———, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Anton, Miguel, and Christopher Polk, 2014, Connected stocks, *Journal of Finance* 69, 1099–1127.

Arnott, Robert D., Mark Clements, Vitali Kalesnik, and Juhani T. Linnainmaa, 2019, Factor momentum, .

Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen, 2017, Quality minus junk, Working paper. Available at SSRN: `https://ssrn.com/abstract=2312432`.

Asness, Clifford S., Antti Ilmanen, Ronen Israel, and Tobias J. Moskowitz, 2015, Investing with style, *Journal of Investment Management* 13, 27–63.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics* 99, 427–446.

Ball, Ray, Joseph Gerakos, Juhani T. Linnainmaa, and Valeri Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, *Journal of Financial Economics* 121, 28–45.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barbee Jr., William C., Sandip Mukherji, and Gary A. Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size?, *Financial Analysts Journal* 52, 56–60.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

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

———, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Barry, Christopher B., and Stephen J. Brown, 1984, Differential information and the small firm effect, *Journal of Financial Economics* 13, 283–294.

Bartov, Eli, and Myungsun Kim, 2004, Risk, mispricing, and value investing, *Review of Quantitative Finance and Accounting* 23, 353–376.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663–682.

Baz, Jamil, Nicolas Granger, Campbell R. Harvey, Nicolas Le Roux, and Sandy Rattray, 2015, Dissecting investment strategies in the cross section and time series, Working paper. Available at SSRN: `https://ssrn.com/abstract=2695101`.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2018, What do mutual fund investors really care about?, Working paper.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

——— , and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.

Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507–528.

Birru, Justin, and Baolian Wang, 2016, Nominal price illusion, *Journal of Financial Economics* 119, 578–598.

Bogousslavsky, Vincent, 2016, Infrequent rebalancing, return autocorrelation, and seasonality, *Journal of Finance* 71, 2967–3006.

Bollen, Nicolas P.B., and Jeffrey A. Busse, 2001, On the timing ability of mutual fund managers, *Journal of Finance* 56, 1075–1094.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan, 2006, The relation between corporate financing activities, analysts' forecasts and stock returns, *Journal of Accounting and Economics* 42, 53–85.

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

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577–2603.

Chan, Yeung Lewis, and Leonid Kogan, 2002, Catching up with the joneses: Heterogeneous preferences and the dynamics of asset prices, *Journal of Political Economy* 110, 1255–1285.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28, 212–246.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, An alternative three-factor model, Working paper. Available at SSRN: `https://ssrn.com/abstract=1418117`.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Christoffersen, Susan E.K., and Mikhail Simutin, 2017, On the demand for high-beta stocks: Evidence from mutual funds, *Review of Financial Studies* 30, 2596–2620.

Cochrane, John H., 2011, Presidential address: Discount rates, *Journal of finance* 66, 1047–1108.

Cohen, Randolph B., Christopher Polk, and Bernhard Silli, 2010, Best ideas, Working paper. Available at SSRN: `https://ssrn.com/abstract=1364827`.

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

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2013, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2, 1–48.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Del Guercio, Diane, and Jonathan Reuter, 2013, Mutual fund performance and the incentive to generate alpha, *Journal of Finance* 69, 1673–1704.

Dichev, Ilia D., 1998, Is the risk of bankruptcy a systematic risk?, *Journal of Finance* 53, 1131–1147.

Duffie, Darrell, 2010, Presidential address: Asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237–1267.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: The impact of prices on takeovers, *Journal of Finance* 67, 933–971.

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

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

——— , 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

———, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

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

———, 2010, Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance* 65, 1915–1947.

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

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

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2019, Taming the factor zoo: A test of new factors, Working paper. Available at SSRN: `https://ssrn.com/abstract=2934020`.

Ferreira, Miguel A., Aneel Keswani, António F Miguel, and Sofia B. Ramos, 2013, The determinants of mutual fund performance: A cross-country study, *Review of Finance* 17, 483–525.

Financial Times, 2017, Smart beta funds pass $1tn in assets, `https://www.ft.com/content/bb0d1830-e56b-11e7-8b99-0191e45377ec` Accessed: 2018-09-11.

Foucault, Thierry, David Sraer, and David J. Thesmar, 2011, Individual investors and volatility, *Journal of Finance* 66, 1369–1406.

Frazzini, Andrea, and Owen A. Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

French, Kenneth R., 2008, Presidential address: The cost of active investing, *Journal of Finance* 63, 1537–1573.

Froot, Kenneth, and Melvyn Teo, 2008, Style investing and institutional investors, *Journal of Financial and Quantitative Analysis* 43, 883–906.

Gârleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2009, Demand-based option pricing, *Review of Financial Studies* 22, 4259–4299.

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

Gervais, Simon, Ron Kaniel, and Dan H. Mingelgrin, 2001, The high-volume return premium, *Journal of Finance* 56, 877–919.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

Gompers, Paul A., and Andrew Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229–259.

Greenwood, Robin, Samuel G. Hanson, and Gordon Y. Liao, 2018, Asset price dynamics in partially segmented markets, *Review of Financial Studies* 31, 3307–3343.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Greenwood, Robin, and Dimitri Vayanos, 2010, Price pressure in the government bond market, *American Economic Review* 100, 585–90.

Grinblatt, Mark, and Tobias J. Moskowitz, 1999, Do industries explain momentum?, *Journal of Finance* 54, 1249–1290.

Gruber, Martin J., 2011, Another puzzle: The growth in actively managed mutual funds, in *Investments And Portfolio Performance* . pp. 117–144 (World Scientific).

Guercio, Diane Del, and Jonathan Reuter, 2014, Mutual fund performance and the incentive to generate alpha, *Journal of Finance* 69, 1673–1704.

Gupta, Tarun, and Bryan T. Kelly, 2019, Factor momentum everywhere, Working paper. Available at SSRN: https://ssrn.com/abstract=3300728.

Harris, Lawrence, and Eitan Gurel, 1986, Price and volume effects associated with changes in the s&p 500 list: New evidence for the existence of price pressures, *Journal of Finance* 41, 815–829.

Harvey, Campbell R., 2017, Presidential address: The scientific outlook in financial economics, *Journal of Finance* 72, 1399–1440.

——— , and Yan Liu, 2017a, Decreasing returns to scale, fund flows, and performance, Working paper. Available at SSRN: https://ssrn.com/abstract=2990737.

——— , 2017b, Does scale impact skill?, Working paper. Available at SSRN: https://ssrn.com/abstract=2872385.

——— , and Heqing Zhu, 2016, ... and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Haugen, Robert A., and Nardin L. Baker, 1996, Commonality in the determinants of expected stock returns, *Journal of Financial Economics* 41, 401–439.

Heston, Steven L., and Ronnie Sadka, 2008, Seasonality in the cross-section of stock returns, *Journal of Financial Economics* 87, 418–445.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hou, Kewei, and David T. Robinson, 2006, Industry concentration and average stock returns, *Journal of Finance* 61, 1927–1956.

Hou, Kewei, Chen Xue, and Lu Zhang, 2018, Replicating anomalies, forthcoming, *Review of Financial Studies.*

Huang, Shiyang, Yang Song, and Hong Xiang, 2019, Fragile factor premia, Working paper. Available at SSRN: https://ssrn.com/abstract=3312837.

Ilmanen, Antti, 2011, *Expected returns: An investor's guide to harvesting market rewards* (John Wiley & Sons).

Investment Company Institute, 2017, *Investment company fact book (2017)* (Investment Company Institute).

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

———, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

Kelly, Bryan, Seth Pruitt, and Yinan Su, 2018, Characteristics are covariances: A unified model of risk and return, Working paper. https://www.nber.org/papers/w24540.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2017, Shrinking the cross section, Working paper. https://www.nber.org/papers/w24070.

———, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kumar, Alok, 2009, Dynamic style preferences of individual investors and stock returns, *Journal of Financial and Quantitative Analysis* 44, 607–640.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lettau, Martin, Sydney C. Ludvigson, and Paulo Manoel, 2018, Characteristics of mutual fund portfolios: Where are the value funds?, Working paper. https://www.nber.org/papers/w25381.

Lewis, Ryan, 2015, Finding fund scale, Working paper, University of Colorado.

Li, Jiacui, 2018, Slow-moving liquidity provision and flow-driven common factors in stock returns, Working paper. Available at SSRN: https://ssrn.com/abstract=2909960.

Linnainmaa, Juhani T., and Michael R. Roberts, 2018, The history of the cross-section of stock returns, *Review of Financial Studies* 31, 2606–2649.

Lockwood, Larry, and Wikrom Prombutr, 2010, Sustainable growth and stock returns, *Journal of Financial Research* 33, 519–538.

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

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

Loughran, Tim, and Jay W. Wellman, 2011, New evidence on the relation between the enterprise multiple and average stock returns, *Journal of Financial and Quantitative Analysis* 46, 1629–1650.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008, The new issues puzzle: Testing the investment-based explanation, *Review of Financial Studies* 21, 2825–2855.

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549–572.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

McQuade, Timothy J., 2018, Stochastic volatility and asset pricing puzzles, Working paper. Available at SSRN: `https://ssrn.com/abstract=3222902`.

Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007, Slow moving capital, *American Economic Review* 97, 215–220.

Moskowitz, Tobias J., 2003, An analysis of covariance risk and pricing anomalies, *Review of Financial Studies* 16, 417–457.

Nagel, Stefan, 2013, Empirical cross-sectional asset pricing, *Annual Review of Financial Economics* 5, 167–199.

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

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

———, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* pp. 109–131.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Piotroski, Joseph D., 2000, Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research* 38, 1–52.

Pollet, Joshua M., and Mungo Wilson, 2008, How does size affect mutual fund behavior?, *Journal of Finance* 63, 2941–2969.

———, 2010, Average correlation and stock market returns, *Journal of Financial Economics* 96, 364–380.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921–945.

Ratcliffe, Ronald, Paolo Miranda, and Andrew Ang, 2017, Capacity of smart beta strategies: A transaction cost perspective, *Journal of Index Investing* 8, 39–50.

Reuter, Jonathan, and Eric Zitzewitz, 2015, How much does size erode mutual fund performance? a regression discontinuity approach, Working paper.

Ritter, Jay R., 1991, The long-run performance of initial public offerings, *Journal of Finance* 46, 3–27.

Schwert, G. William, 2003, Anomalies and market efficiency, *Handbook of the Economics of Finance* 1, 939–974.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41, 579–590.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* pp. 289–315.

Soliman, Mark T., 2008, The use of dupont analysis by market participants, *The Accounting Review* 83, 823–853.

Song, Yang, 2018, The mismatch between mutual fund scale and skill, Working paper. Available at SSRN: `https://ssrn.com/abstract=3065700`.

Spiegel, Matthew, and Hong Zhang, 2013, Mutual fund risk and market share-adjusted fund flows, *Journal of Financial Economics* 108, 506–528.

Spiess, D. Katherine, and John Affleck-Graves, 1999, The long-run performance of stock returns following debt offerings, *Journal of Financial Economics* 54, 45–73.

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

Stein, Jeremy C., 2009, Presidential address: Sophisticated investors and market efficiency, *Journal of Finance* 64, 1517–1548.

Teo, Melvyn, and Sung-Jun Woo, 2004, Style effects in the cross-section of stock returns, *Journal of Financial Economics* 74, 367–398.

Thomas, Jacob K., and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, K.C. John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Wahal, Sunil, and M. Deniz Yavuz, 2013, Style investing, comovement and return predictability, *Journal of Financial Economics* 107, 136–154.

Welch, Ivo, and Amit Goyal, 2007, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655–1695.

Xing, Yuhang, 2008, Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Yan, Xuemin Sterling, 2008, Liquidity, investment style, and the relation between fund size and fund performance, *Journal of Financial and Quantitative Analysis* 43, 741–767.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Appendix

A Factor definition

This section describes definition of sorting variables used to construct the 50 factors studied in this paper. Variables indicated in bold are corresponding Compustat items. The general rules for factor construction can be found in Section 2.1. For factors constructed on discrete sorting variables, we do not use NYSE breakpoints to form long-short portfolios. We describe the specific methods for constructing those factors below.

1. Fundamentals-based Factors

We construct 38 fundamentals-based factors as annually-rebalanced long-short portfolios. The only two exceptions are "Distree risk" and "Ohlson's O-socre", which are constructed as quarterly long-short portfolios.

Abnormal capital investment: Titman, Wei, and Xie (2004).

$$Abnormal\,capital\,investment = \frac{CE_t}{(CE_{t-1} + CE_{t-2} + CE_{t-3})/3} - 1,$$

where CE_t is a firm's capital expenditure (**CAPX**) in year t scaled by its sales (**SALE**) in year t-1. Long low abnormal cpital investment stocks.

Accruals: Sloan (1996).

$$Accruals = ((\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD - \Delta TP) - Dep)/Average\,total\,assets,$$

where ΔCA = Change in current assets(**ACT**), $\Delta Cash$ = Change in cash and cash equivalents (**CHE**), ΔCL = Change in current liabilities (**LCT**), ΔSTD = Change in debt included in current liabilities (**DLC**), ΔTP = Change in income taxes payable (**TXP**), Dep = Depreciation and amortization expense (**DP**), and $Average\,total\,assets$ = the average of total assets (**AT**) at current and previous fiscal year-end. Long low accruals stocks.

Asset growth: Cooper, Gulen, and Schill (2008).

$$Asset\,growth_t = (Total\,Assets_t - Total\,Assets_{t-1})/Total\,Assets_{t-1},$$

where $Total\,Assets_t$ is measured by compustat item **AT**. Long low asset growth stocks.

Book-to-market: Fama and French (1992).

$$Book - to - market_t = Book\,equity_t/Market\,equity_t,$$

Book equity = book value of stockholders's equity (**SEQ**) + deferred taxes and investment tax credit (**TXDITC**) - book value of preferred stock. We replace missing **TXDITC** with zero. Depending on data vailability, we use redemption (**PSTKRV**), liquidation (**PSTKL**), par value (**PSTK**), and zero in that order as the book value of preferred stock. The market equity is the market capitalization at the end of calendar year t. Long high book-to-market stocks.

Cash-based profitability: Ball, Gerakos, Linnainmaa, and Nikolaev (2016).

$$\begin{aligned}Cash - based\,operating\,profitability = [&Operating\,profitability - \Delta Account\,receivable(\mathbf{RECT})\\ &-\Delta Inventory(\mathbf{RECT}) - \Delta Prepaid\,expenses(\mathbf{XPP}) + \Delta Deferred\,revenue(\mathbf{DRC}+\mathbf{DRLT})\\ &+\Delta Trade\,accounts\,payable(\mathbf{AP}) + \Delta Accrued\,expenses(\mathbf{XACC})]/Average\,total\,assets,\end{aligned}$$

where operating proftability=Revenue(**REVT**)-Cost of goods sold(**COGS**)-Reported sales, general, and administrative expenses(**XSGA**-**XRD**). Missing values in **XPP**, **DRC**, **DRLT**, **AP**, **XACC**, **XSGA**, and **XRD** are replaced by zero. Long high cash-based profitability stocks.

Cashflow-to-price: Lakonishok, Shleifer, and Vishny (1994).

$$Cashflow - to - price = (\mathbf{IB}+\mathbf{DP})/market\,equity,$$

where market equity is December-end market capitalization. Long high cashflow-to-price stocks.

Change in asset turnover: Soliman (2008).

$$Asset\,turnover_t = Asset\,turnover_t - Asset\,turnover_{t-1},$$

where asset turnover = Sales(**SALE**) / Average NOA. NOA = Operating assets - Operating liabilities, where Operating assets = Total assets(**AT**) - Cash and short-term investments (**CHE**); Operating libility = Total assets(**AT**) - Long- and short-term portions of debt (**DLTT+DLC**). Long stocks with high change in asset turnover.

Debt issuance: Spiess and Affleck-Graves (1999). Debt issuance is a *binary variable* that equals one if Compustat item long-term debt issuance (**DLTIS**) is greater than 0, and it equals zero elsewhere. Long non-debt issuance stocks.

Distress risk: Campbell, Hilscher, and Szilagyi (2008). A detailed description of how to construct distress risk (failure probability) can be found in Chen, Novy-Marx, and Zhang (2011). By definition, this is a quarterly-rebalanced factor. Long low distress risk stocks.

Earnings-to-price: Basu (1977).

$$Earnings - to - price_t = \mathbf{IB}_t / ME_t,$$

where ME is market capitalization at December-end. Long high earnings-to-price stocks.

Enterprise multiple: Loughran and Wellman (2011).

$$Enterprise\,multiple = Enterprise\,value/\mathbf{OIBDP},$$

where enterprise value is the June-end market value of equity plus total debt (**DLC+DLTT**) plus preferred stocks value (**PSTKRV**) minus cash and short-term investments (**CHE**). Negative EM and financial firms are excluded. Long low enterprise multiple stocks.

Five-year share issuance: Daniel and Titman (2006).

$$Issue - 5yr_t = log(adjshrout_t/adjshrout_{t-5}),$$

where the split-adjusted shares outstanding is measured by Compustat items at each fiscal-year end: adjshrout = **CSHO**×**ADJEX_F**. Long stocks with low five-year share issuance.

Gross profitability Novy-Marx (2013).

$$Gross\,profitability = [Revenues(\mathbf{REVT}) - Cost\,of\,goods\,sold(\mathbf{COGS})]/total\,assets(\mathbf{AT}).$$

Long high gross profitability stocks.

Growth in inventory: Thomas and Zhang (2002).

$$Growth\,in\,inventory_t = \frac{\Delta\mathbf{INVT}_t}{(\mathbf{AT}_t + \mathbf{AT}_{t-1})/2}.$$

Long stocks with low growth in inventory.

Industry concentration Hou and Robinson (2006). We first compute industry concentration measured by Herfindahl index:

$$Herfindahl_j = \sum_i {S_{ij}}^2,$$

where ${S_{ij}}^2$ is the market share of firm i in industry j. The market share is computed based on net sales (**SALE**), and iundustries are defined through three-digit SIC classifications. Firms in regulated industries are excluded: Railroads (SIC code 4011) through 1980; Trucking (4210 and 4213) through 1980; Airlines (4512) through 1978; Telecommunications (4812 and 4813) through 1982; Gas and eletric utilities (4900 to 4939). Industries with less than 20 member firms are excluded to avoid forming thin portfolios.

We form long-short stock portfolio based on industry concentration as follows. Each June we sort industries into quintiles based on Herfindahl index calculated above. We long firms whose industries are in the bottom quintile of Herfindahl index and short firms whose industries are in the top quintile of Herfindahl index.

Industry-adjusted CAPX growth: Abarbanell and Bushee (1998). We first compute unadjusted CAPX growth:

$$CAPX\,growth_t = (\mathbf{CAPX_t} - \frac{\mathbf{CAPX_{t-1}} + \mathbf{CAPX_{t-2}}}{2})/(\frac{\mathbf{CAPX_{t-1}} + \mathbf{CAPX_{t-2}}}{2}).$$

Industry-adjusted CAPX growth is calculated as unadjusted CAPX growth minus the industry-average (equal-weighted) unadjusted CAPX growth. Industries are defined through two-digit SIC code. Long stocks with low industry-adjusted CAPX growth.

Investment growth: Xing (2008).

$$Investment\,growth_t = \frac{\mathbf{CAPX}_t}{\mathbf{CAPX}_{t-1}} - 1.$$

Long stocks with low investment growth.

Investment-to-assets: Lyandres, Sun, and Zhang (2008).

$$Investment - to - asset_t = \frac{\Delta\mathbf{PPENT}_t + \Delta\mathbf{INVT}_t}{\mathbf{AT}_{t-1}}.$$

Long stocks with low investment-to-assets.

Investment-to-capital: Xing (2008).

$$Investment - to - capital_t = \mathbf{CAPX}_t/\mathbf{PPENT}_t.$$

Long stocks with low investment-to-capital.

Leverage: Bhandari (1988).

$$Leverage = \mathbf{DLTT}/ME,$$

where ME is market capitalization at December-end. Long stocks with high leverage ratio.

M/B and accruals: Bartov and Kim (2004). This is a discrete signal based on book-to-market ratio and accruals. Each June-end, stocks are sorted into book-to-market terciles and accruals terciles based on NYSE 30th and 70th percentiles, where book-to-market and accruals are defined as in book-to-market factor and accruals factor above. The composite signal is set to one for stocks that are in both the highest book-to-market tercile and the lowest accruals tercile, and it is set to zero for stocks that are in both the lowest book-to-market tercile and the highest accruals tercile. We long stocks with signal = 1 and short stocks with signal = 0.

Net operating assets: Hirshleifer, Hou, Teoh, and Zhang (2004).

$$NOA_t = \frac{Operating\,assets_t - Operating\,liabilities_t}{\mathbf{AT}_{t-1}},$$

where Operating assets = Total assets(**AT**) - Cash and short-term investments(**CHE**), and Operating liability = Total assets(**AT**) - Long- and short-term portions of debt(**DLTT**+**DLC**) - Minority interest(**MIB**) - Preferred stocks(**PSTKRV**) - Common equity(**CEQ**). Long stocks with low net operating assets.

Net working capital changes: Soliman (2008).

$$Net\,WC\,Changes_t = \frac{\Delta(\mathbf{ACT}_t - \mathbf{CHE}_t) - \Delta(\mathbf{LCT}_t - \mathbf{DLC}_t)}{\mathbf{AT}_t}.$$

Long stocks with low net working capital changes.

Ohlson's O-score: Ohlson (1980). A detailed description of how to construct O-score can be found in Chen, Novy-Marx, and Zhang (2011). By definition, this is a quarterly-rebalanced factor. Long stocks with low O-score.

One-year share issuance: Pontiff and Woodgate (2008). This is a discrete signal based on one-year share issuance computed as below:

$$Issue - 1yr_t = log(adjshrout_t/adjshrout_{t-1}),$$

where the split-adjusted shares outstanding is computed based on CRSP shares outstanding and adjustment factor as of June-end of year t: adjshrout = $shrout \times cfacshr$. We long stocks with positive one-year share issuance and short stocks with non-positive one-year share issuance.

Operating profitability: Fama and French (2015).

$$Operating\,profitability = \frac{\textbf{REVT} - \textbf{COGS} - \textbf{XSGA} - \textbf{XINT}}{BE},$$

where BE is defined as in book-to-market factor above. Long stocks with high operating profitability.

Piotroski's F-score: Piotroski (2000). Piotroski's F-score is a composite score defined as the sum of 9 individual signals that correspond to ROA, change in ROA, operating cash flow, accruals, change in gross margin ratio, change in asset turnover, change in leverage ratio, change in current ratio, and equity issuance. Detailed definition of the 9 individual signals can be found in Section 2.3 of Piotroski (2000). Long high F-score stocks.

Profit margin: Soliman (2008).

$$Profit\,margin = \textbf{OIADP}/\textbf{SALE}.$$

Long high profit margin stocks.

QMJ profitability: Asness, Frazzini, and Pedersen (2017).

$$QMJ = Z_{GPOA} + Z_{ROE} + Z_{ROA} + Z_{CFOA} + Z_{GMAR} + Z_{ACC},$$

where Z_{GPOA}, Z_{ROE}, Z_{ROA}, Z_{CFOA}, Z_{GMAR}, and Z_{ACC} are cross-sectional z-scores of gross profits over assets, return on equity, return on assets, cashflow over assets, gross margin, and low accruals. Detailed definition of the six profitability measures can be found in the Appendix of Asness, Frazzini, and Pedersen (2017). Long high QMJ stocks.

Return on assets: Haugen and Baker (1996).

$$ROA = \mathbf{IB}/\mathbf{AT}.$$

Long high ROA stocks.

Return on assets: Haugen and Baker (1996).

$$ROE = \mathbf{IB}/BE,$$

where BE is book equity defined in the book-to-market factor above. Long high ROE stocks.

Sales growth: Lakonishok, Shleifer, and Vishny (1994).

$$Sales\, growth_t = 5 \times rank_t + 4 \times rank_{t-1} + 3 \times rank_{t-2} + 2 \times rank_{t-3} + rank_{t-4},$$

where $rank_t$ is the cross-sectional decile ranking of a firm's sales growth (percentage change in **SALE**) in fiscal year ending in calendar year t. We require firms to have 6 years of observations at portfolio formation.Long high sales growth stocks.

Sales-to-price: Barbee Jr., Mukherji, and Raines (1996).

$$Sales - to - price = \mathbf{REVT}/ME,$$

where ME is December-end market capitalization. Long high sales-to-price stocks.

Sales-minus-inventory growth: Abarbanell and Bushee (1998).

$$Sales - inventory\, growth_t = \frac{\mathbf{REVT}_t}{(\mathbf{REVT}_{t-1} + \mathbf{REVT}_{t-2})/2} - \frac{\mathbf{INVT}_t}{(\mathbf{INVT}_{t-1} + \mathbf{INVT}_{t-2})/2}.$$

Long stocks with high sales-minus-inventory growth.

Size: Banz (1981). Firm size is measured by December-end market capitalization. Long small stocks.

Sustainable growth: Lockwood and Prombutr (2010).

$$Sustainable\,growth_t = (BE_t - BE_{t-1})/BE_{t-1}$$

where BE is defined as in the book-to-market factor above. Long low sustainable growth stocks.

Total external financing: Bradshaw, Richardson, and Sloan (2006).

$$Total\,external\,financing = [(\frac{adjshrout_t}{adjshrout_{t-1}} - 1) \times ME_t + \Delta\mathbf{DLC}_t + \Delta\mathbf{DLTT}_t - \mathbf{DVC}_t]/\mathbf{AT}_t,$$

where the split-adjusted shares outstanding is measured by Compustat items at each fiscal-year end: adjshrout = **CSHO**×**ADJEX_F**, and ME is market capitalization at fiscal year-end. Long stocks with low total external financing.

Altman's Z-score: Dichev (1998).

$$Z - score = 1.2 \times \frac{\mathbf{ACT} - \mathbf{LCT}}{\mathbf{AT}} + 1.4 \times \frac{\mathbf{RE}}{\mathbf{AT}} + 3.3 \times \frac{\mathbf{EBIT}}{\mathbf{AT}} + 0.6 \times \frac{ME}{\mathbf{LT}} + 1.0 \times \frac{\mathbf{REVT}}{\mathbf{AT}},$$

where ME is December-end market capitalization. Long stocks with high z-score.

2. Return-based Factors

Below are definitions for factors constructed based on stock returns, stock prices, or trading volume. All factors below are constructed as monthly-rebalanced long-short portfolios.

52-week high: George and Hwang (2004). Nearness to 52-week high is computed as month-end close price devided by the highest price during the 12-month period that ends on this month. Long stocks with high nearness to 52-week high.

Firm age: Barry and Brown (1984). Firm age is computed as the number of month that the firm has been in CRSP. Long stocks with larger firm age.

Heston and Sadka's seasonality: Heston and Sadka (2008). In each month, the sorting variable is the average returns in the same calendar month over the previous 20 years. Long stocks with higher past returns in the same calendar month.

High-volume return premium: Gervais, Kaniel, and Mingelgrin (2001). At each month end, we classify stocks into high (low) volume group if its trading volume (share-based) on the month-end trading day is at the top (bottom) volume decile when pooled with its trading volumes over the past 50 trading days. We then long (short) stocks in highest (lowest) volume decile in the next month.

Idiosyncratic Volatility: Ang, Hodrick, Xing, and Zhang (2006). For each stock-month, we regress daily stock returns on daily Fama-French three factors in the month and compute idiosyncratic volatility as standard deviation of residuals from the regression. Long low idiosyncratic volatility stocks.

Intermediate Momentum: Novy-Marx (2012). At the end of each month $t-1$, we form portfolios based on cumulative returns from month t-12 to month t-7 and hold the portfolios in month t. Long stocks with high cumulative returns from month t-12 to month t-7.

Long-term reversal: De Bondt and Thaler (1985). At the end of each month $t-1$, we form portfolios based on cumulative returns from month t-60 to month t-13 and hold the portfolios in month t. Long stocks with high cumulative returns from month t-60 to month t-13.

Market beta: Fama and MacBeth (1973). At each month-end, we estimate market beta with respect to CRSP equal-weighted return using monthly returns over past 60 months. Long high beta stocks.

Max daily return: Bali, Cakici, and Whitelaw (2011). At each month-end, we form portfolios based on maximum daily return in the month and hold the portfos over next month. Long stocks with high maximum daily return.

Momentum: Jegadeesh and Titman (1993). At the end of each month $t-1$, we form portfolios based on cumulative returns from month t-6 to month t-2 and hold the portfolios in month t. Long stocks with high cumulative returns from month t-6 to month t-2.

Nominal price: Birru and Wang (2016). The sorting variable is month-end share price.

Long stocks with low share price.

Short-term reversal: Jegadeesh (1990). The sorting variable is return in current month. Long stocks with low returns in current month.