# Flow-Based Asset Pricing: A Factor Framework of Cross-Sectional Price Impacts*

Yu An† Yinan Su‡ Chen Wang§

This version: July 24, 2023
First version: May 2, 2022

**Abstract**

We propose that noise trading flows impact cross-sectional asset prices through systematic risk factors. In our model, asset-level flows, when aggregated at the factor level, drive fluctuations in factor prices and risk premia. In the cross-section, asset prices change with factor-level price impacts in accordance with the factor pricing model. Empirically, our model explains the price impacts and cross impacts between underlying assets with a few risk factors. Moreover, the model-implied trading strategy, designed to optimally exploit the reversion in price impacts, delivers strong and robust investment outcomes.

**Keywords**: cross-section, factor, flow, price impact, risk

**JEL Codes**: G11, G12

*We thank Daniel Andrei, Federico Bandi, Hank Bessembinder, Andrew Chen, Thummim Cho, Zhi Da, Darrell Duffie, Daniel Green, Robin Greenwood, Zhiguo He, Ben Hébert, Shiyang Huang, Bryan Kelly, Ralph Koijen, Badrinath Kottimukkalur, Serhiy Kozak, Arvind Krishnamurthy, Jiacui Li, Xin Liu, Dong Lou, Jun Pan, Andrey Pankratov, Jonathan Parker, Paolo Pasquariello, Nagpurnanand Prabhala, Tarun Ramadorai, Alessandro Rebucci, Oleg Rytchkov, Paul Schultz, Dongho Song, Yang Song, Zhaogang Song, Semih Üslü, Adrien Verdelhan, Wei Wu, Zeyu Zheng, and conference and seminar participants at Notre Dame, JHU Carey, Fed Board, RUC-VUW Joint Virtual Research Workshop, Wolfe Research QES 6th NYC Quant Conference, Campbell & Company, MFA, Southern Methodist University, FMCG Conference, DC junior finance conference, SoFiE early-career scholars conference, and CICF for helpful comments. All errors are our own.

†Carey Business School, Johns Hopkins University; yua@jhu.edu.
‡Carey Business School, Johns Hopkins University; ys@jhu.edu.
§Mendoza College of Business, University of Notre Dame; chen.wang@nd.edu.

# 1 Introduction

The interaction between noise traders and sophisticated investors plays a pivotal role in shaping asset prices. Noise trading is a major source of the asset pricing phenomenon that violates the canonical factor pricing framework. Empirically, one strand of the literature treats flow from noise traders as an asset characteristic and finds that sorting on it generates anomalous returns.[1] In another strand, researchers find quantitatively large price multipliers of the return of an individual asset or a factor portfolio to its respective flow.[2]

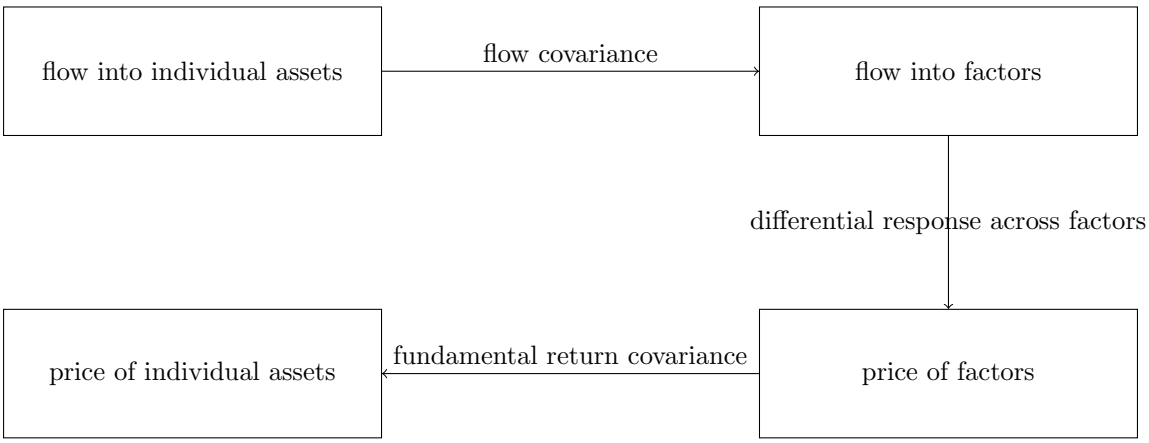
The theoretical rationale explaining how noise trading can generate asset price fluctuations beyond the canonical factor model hinges on the limited risk-bearing capacity of sophisticated investors. These investors, tasked with absorbing the flows generated by noise traders, experience changes in the equilibrium holdings of risky assets. Consequently, asset prices adjust to compensate these sophisticated investors for absorbing extra risks.

A central challenge in the existing literature is the absence of a framework that integrates two critical elements underpinning cross-sectional asset pricing with noise trading. On the one hand, the limited risk-bearing capacity of sophisticated marginal investors causes asset prices to respond to noisy flows. On the other hand, these marginal investors are still present in the market to enforce some degrees of the law of one price, such that factor pricing remains a prominent feature of cross-sectional asset pricing.

In this paper, we tackle this challenge by providing a new framework that unifies price impacts and factor pricing. Our model, depicted in Figure 1, posits that noisy flows impact cross-sectional asset prices through systematic risk factors. Firstly, we utilize empirical evidence highlighting the commonality in noisy flows across various assets, indicating that noise traders tend to buy or sell different assets bearing similar characteristics in a collective manner. Consequently, we employ the flow covariance structure to aggregate individual asset flows into factor-level flows. Secondly, market clearing dictates that the assets that

[1] See, e.g., Coval and Stafford (2007), Lou (2012), and Barber, Huang, Odean, and Schwarz (2022).
[2] See, e.g., Chang, Hong, and Liskovich (2015), Koijen and Yogo (2019), and Gabaix and Koijen (2022).

**Figure 1. Factor model of price impacts**




noise traders purchase must be sold by marginal investors. As a result, factor flows shift marginal investors' portfolio holdings and their exposure to fundamental risk factors. In equilibrium, each factor's price response generated by its own flow can vary, depending on the willingness of marginal investors to accommodate risk for that particular factor. Lastly, we note that different assets share exposure (i.e., $\beta$) to fundamental risk factors. Therefore, as factor prices fluctuate, individual asset prices adjust in accordance with their respective risk exposures to preserve the law of one price.

In short, given the interconnected covariance structures of noisy flows and fundamental returns, investigating the price impact of flows by assessing each asset or factor portfolio in isolation is inadequate. Rather, it is essential to consider the price impacts for the cross-section of assets under a structured model like ours. Our cross-sectional analysis is grounded in portfolio theory and factor pricing.

Our framework is versatile by accommodating flexible choices of flow data, test assets, and factors. We demonstrate the effectiveness of our model by estimating it using the most commonly used measures in existing literature. We measure noisy flows using U.S. equity mutual fund flow-induced trading, a widely used measure in the literature.[3] This measure is based on two conventional assumptions: (1) mutual fund investors primarily consist of

[3]This measure has been employed in, e.g., Coval and Stafford (2007), Froot and Ramadorai (2008), Greenwood and Thesmar (2011), Lou (2012), Ben-David, Li, Rossi, and Song (2022a), and Li (2022).

uninformed retail investors; and (2) their purchases or sales of mutual fund shares prompt mutual funds to buy or sell individual stocks in proportion to lagged holdings. The test assets are the Fama and French (1993) 5×5 size and book-to-market double-sorted portfolios, and the factors include market (MKT), small-minus-big (SMB), and high-minus-low (HML). Our analysis employs a monthly sample spanning from 2000 through 2020.

We develop a two-stage procedure to estimate the factor model of price impacts, drawing inspiration from the well-known Fama-MacBeth regression. In the first stage, we perform asset-by-asset time-series regressions, where each test asset's flows are regressed on the three-factor flows. This regression estimates the factor structure of flows and allows us to infer how factor flows change marginal investors' portfolio holdings and risk exposures at the 5×5 asset level. The second stage runs a panel regression of asset returns on the product of factor flow and the quantity of risk exposure. This regression estimates the extent to which marginal investors are willing to accommodate risk for each factor's flow.

The innovation in our two-stage procedure stems from the relationship between the first and second stages. In the first-stage regression, we estimate the betas of test asset flows with respect to factor flows, which we call "flow betas." The central idea is that *flow betas serve as portfolio weights*. For instance, consider the MKT factor. All else equal, a one-dollar flow into the MKT factor increases the flow into a test asset by an amount equal to the flow beta. Since marginal investors absorb these flows, market clearing implies that a one-dollar MKT flow reduces marginal investors' holdings in the test asset by the same amount equal to the flow beta. In other words, the flow betas represent the changes in marginal investors' 5×5 portfolio holdings caused by one dollar of MKT flow. Consequently, in the second-stage regression, we use flow betas as portfolio weights to compute the covariance between the factor's and asset's fundamental-driven returns. The product of this return covariance and factor flow represents the "flow-induced risk" absorbed by the marginal investors. Our approach, which involves using flow betas as portfolio weights, relates to the recent work of Alekseev, Giglio, Maingi, Selgrad, and Stroebel (2022), who estimate the betas of mutual

fund portfolio changes in response to heat shocks and use these quantity betas as portfolio weights to hedge against climate risk.

The empirical evidence supports our initial hypothesis that noisy flows impact cross-sectional asset prices through systematic risk factors. First, we find that each unit of flow-induced risk generates a greater price response in the SMB and HML factors than in the MKT. This finding aligns with the idea that investors may have distinct investment mandates for different style portfolios. For example, an insurance company focusing exclusively on large-cap stocks can absorb MKT flows but cannot elastically absorb SMB flows, as doing so would require trading small-cap stocks. Second, we find that directly regressing returns on flows for the 5×5 test assets accounts for about 8% of return variations, while our model, using only three free parameters, explains 7% of return variations. Third, we estimate the price impacts and cross impacts between the 5×5 assets and find that these estimates closely align with our model-implied values. Notably, the return covariance in our model naturally implies that two assets with more similar fundamentals would have a greater cross impact. This ability to fit cross impacts through return covariance differentiates our framework from the characteristics-based demand system of Koijen and Yogo (2019).

Ultimately, our estimation provides a structured model depicting how the prices of the 5×5 assets change as a function of their flows. With this, we apply our estimates to construct the optimal strategy to generate profit from these flows. The underlying idea is that noisy flows push prices away from fundamental values, so flow-induced price impacts should revert over the long term. While the literature has estimated price impacts at individual stock, factor, or market levels, it remains unclear how to construct the mean-variance optimal portfolio to profit from the flow, because different stocks and factors are interconnected through the covariance structures of return and flow.

We employ our model to derive the optimal trading strategy, which exhibits two key features. First, because we have demonstrated that three factors explain the cross-section of price impacts for the 5×5 assets, our strategy is built upon mean-variance optimization

for these three factors as opposed to all test assets. Optimizing at the factor level helps eliminate potential noises inherent in asset-level estimation, leading to a more consistent reversion of out-of-sample prices. Second, our strategy differs from a simplistic strategy that indiscriminately trades against all flows, because we have shown that the compensation for absorbing flow-induced risk differs among factors. Consequently, the optimal strategy assigns greater weight to factors that offer higher risk compensation.

Our trading strategy that capitalizes on flow can significantly enhance traditional fundamental investing approaches, which typically overlook flow-related information. To empirically support this claim, we combine our strategy with a comprehensive list of firm characteristics-based anomaly portfolios (which serve as proxies for fundamental investing). We find that the improvements in investment outcomes on top of these anomaly portfolios are wide-ranging and economically significant. Out of the 154 Fama-French and Jensen, Kelly, and Pedersen (2021) portfolios, 140 or 91% experience a positive increase in the out-of-sample Sharpe ratio. The average change in the annualized Sharpe ratio is 0.3. The results are robust across a series of alternative empirical specifications.

The remainder of this paper is organized as follows. Section 2 reviews related literature. Section 3 provides a theoretical foundation for the factor model of price impacts. Section 4 outlines the empirical framework employed for model estimation. Section 5 applies this empirical framework to analyze mutual fund flows and Fama-French portfolios. Section 6 concludes. The appendices provide supplementary results and robustness checks.

## 2 Related Literature

Our paper lies at the intersection of noise trading literature and factor pricing literature. Early theoretical works, such as De Long, Shleifer, Summers, and Waldmann (1990), Campbell and Kyle (1993), and Daniel, Hirshleifer, and Subrahmanyam (2001) model noise traders as an additional source of risk. More closely related, Kozak, Nagel, and Santosh (2018) argue that due to the arbitrage activity of marginal investors, asset returns should exhibit a factor

structure even in the presence of noise traders. Building upon their intuition, our contribution is developing a factor framework that jointly incorporates flows, risks, and returns.

Our paper's primary objective is to estimate the price impacts between various underlying assets through risk factors, with a unique focus on cross-sectional connections. This approach differentiates our work from earlier studies that document the impact of noise trading on factor prices (Teo and Woo, 2004; Huang, Song, and Xiang, 2021; Ben-David, Li, Rossi, and Song, 2022a; Li, 2022; Li and Lin, 2022). Additionally, our research stands out in the context of factor-level price impact estimation, as it addresses a crucial definitional issue that has long persisted in the literature.

Traditionally, the literature measures how a 1% change in quantity affects the price (refer to Gabaix and Koijen (2022) for a summary). While this definition is applicable to individual assets or aggregate markets, it is mathematically ill-defined for factors like SMB. This issue arises because the total quantity of SMB is zero, rendering it impossible to divide by zero and nonsensical to say "buying 1% more of SMB" (for detailed arguments, see Section 3).

Our approach diverges from the conventional method by estimating the price impact of one unit of risk induced by a change in quantity. This difference is not merely technical; it is deeply intertwined with the core economic channel of our paper. Price impacts occur because flow-absorbing investors are averse to the *risk* induced by flows, as opposed to being simply "inelastic" to the flows.

The demand-based asset pricing literature has largely focused on estimating how flows into a specific asset or factor portfolio influence its own price. A smaller subset of studies has explored the cross impacts on other assets or factors (Boulatov, Hendershott, and Livdan, 2013; Pasquariello and Vega, 2015; Huang, Song, and Xiang, 2021; Chaudhary, Fu, and Li, 2023). However, relying on a reduced-form analysis to investigate these cross impacts overlooks the structures of the underlying economic models. Additionally, one would need to estimate $N^2$ pairs of cross impacts for $N$ assets, not to mention when portfolios of these assets are also considered.

Our factor model addresses this issue by connecting the cross-section of price impacts through risk factors. Empiricists only need to estimate the price impacts for a few factors. The cross impact between any pair of assets and factors can then be derived through the covariance structure of flows and returns. Empirically, we demonstrate that our three-parameter model effectively fits the cross impacts between the 5×5 test assets.

We contribute to the cross-sectional asset pricing literature by integrating flows into the classical portfolio tools. While there is a long-standing tradition of studying the factor structure of returns, our framework uniquely examines the factor structure of both flows and returns. Recently, a growing body of literature has been exploring new methods for forming asset-pricing factors using firm-level characteristics or trading signals (e.g., Harvey, Liu, and Zhu, 2016; Kelly, Pruitt, and Su, 2019; Giglio and Xiu, 2021; Kelly, Malamud, and Pedersen, 2021). In this context, we consider flow information as a distinct characteristic that warrants special attention, as it reflects the equilibrium portfolio changes of marginal investors due to market clearing. Crucially, we demonstrate that the model-implied strategy can enhance the investment performance of a broad array of characteristics-sorted portfolios by optimally incorporating flow information.

Finally, a large strand of literature investigates the asset pricing implications of commonality in trading quantity. Hasbrouck and Seppi (2001) first document that flows into the cross-section of stocks exhibit a factor structure using NYSE's TAQ data. Dou, Kogan, and Wu (2022) and Kim (2020) demonstrate that the commonality in flows in and out of mutual funds is a risk factor for expected stock returns. Lo and Wang (2000) show that trading volume exhibits a factor structure, and Alvarez and Atkeson (2018) show that trading volume is a priced risk factor. Balasubramaniam, Campbell, Ramadorai, and Ranish (2021) find that Indian households' stock holdings exhibit a factor structure. Unlike these papers, in our setting, common flows into the cross-section of stocks are not risk factors but create price impacts by changing factors' risk premia.

## 3 Theoretical Foundation

This section presents the theoretical foundation for the factor model of price impacts. The key assumption is that different groups of marginal investors accommodate noisy flows into different factor portfolios, diverging from the common assumption that the same marginal investors accommodate noisy flows into various factors. Our assumption is motivated by the evidence that stock markets exhibit heterogeneous elasticity for different portfolios.[4] Our assumption naturally generates such differential price responses for different factors, which then, through the law of one price (LOOP), leads to the factor model of price impacts.

Our model consists of two periods, $t = 0$ and $t = 1$, with a gross risk-free rate $R_F$. The model features $N$ assets.[5] Each asset, denoted by $n = 1, 2, \ldots, N$, has a payoff $X_n$ at time $t = 1$, with the payoff vector represented as $\mathbf{X} = (X_1, X_2, \ldots, X_N)^\top$. We assume that the payoff $\mathbf{X}$ is jointly normally distributed and exhibits an exact $K$-factor structure spanned by the factors: $\mathbf{b}_1^\top \mathbf{X}, \mathbf{b}_2^\top \mathbf{X}, \ldots, \mathbf{b}_K^\top \mathbf{X}$. Specifically, $\mathbf{b}_k = (b_{1,k}, b_{2,k}, \ldots, b_{N,k})^\top$ denotes the portfolio shares of factor $k$, where one share of factor portfolio holds $b_{n,k}$ shares of asset $n$.

At time 0, noise traders have an inelastic demand to buy $q_k$ shares of factor $k$. We assume that the $K$ factors have both uncorrelated payoffs and uncorrelated flows, i.e., $\text{cov}(\mathbf{b}_k^\top \mathbf{X}, \mathbf{b}_j^\top \mathbf{X}) = 0$ and $\text{cov}(q_k, q_j) = 0$ for $k \neq j$. This assumption is without loss of generality because factors with correlated payoffs and flows can always be rotated to uncorrelated ones. Appendix A.1 provides technical details of this rotation.

Let $f_n$ denote the total flow into asset $n$. Because $q_k$ shares of flow into factor $k$ lead to $q_k b_{n,k}$ shares of flow into asset $n$, it follows that $f_n = \sum_{k=1}^{K} q_k b_{n,k}$, meaning that asset flow exhibits a $K$-factor structure. The following proposition summarizes this relationship.

[4] See, e.g., Gabaix and Koijen (2022) and Li and Lin (2022).

[5] We use bold font notation for matrices and vectors, and $\mathbf{A}^\top$ to denote the transpose of matrix $\mathbf{A}$.

**PROPOSITION 1.** *The relationship between asset flows and factor flows is*

$$\begin{pmatrix} f_1 \\ f_2 \\ \cdots \\ f_N \end{pmatrix} = q_1 \begin{pmatrix} b_{1,1} \\ b_{2,1} \\ \cdots \\ b_{N,1} \end{pmatrix} + q_2 \begin{pmatrix} b_{1,2} \\ b_{2,2} \\ \cdots \\ b_{N,2} \end{pmatrix} + \cdots + q_K \begin{pmatrix} b_{1,K} \\ b_{2,K} \\ \cdots \\ b_{N,K} \end{pmatrix}. \tag{1}$$

Marginal investors on the other side of the market accommodate these noisy flows and determine asset prices based on their optimality conditions. We assume that for each factor $k$, a continuum of marginal investors with a total mass $\mu_k$ and a CARA risk-aversion parameter $\gamma_k$ absorbs the flow $q_k$. As discussed, our model diverges from traditional setups in that $\mu_k$ and $\gamma_k$ can differ across factors, thereby leading to a differential response to flow. Intuitively, our model accounts for the potential segmentation of flow-absorbing investors for these factors that exhibit uncorrelated payoffs and flows.

We now determine the time-0 price of the $k$-th factor, denoted as a function $P_k(q_k)$ of the flow $q_k$. In equilibrium, the factor price $P_k(q_k)$ is determined such that each marginal investor finds it optimal to buy $-q_k/\mu_k$ shares of the factor, clearing the market. That is,

$$-q_k/\mu_k = \arg\max_y \mathbb{E}[-\exp(-\gamma_k((S_k/\mu_k + y)\mathbf{b}_k^\top \mathbf{X} - yR_F P_k(q_k))], \tag{2}$$

where $y$ represents the change in each marginal investor's holding in factor $k$, and $S_k$ is the total amount outstanding of factor $k$. The first-order condition of (2) implies that

$$P_k(q_k) = \lambda_k(q_k - S_k)\text{var}(\mathbf{b}_k^\top \mathbf{X}) + \frac{\mathbb{E}(\mathbf{b}_k^\top \mathbf{X})}{R_F}, \tag{3}$$

where we define the parameter $\lambda_k = \gamma_k/(\mu_k R_F)$. As one can see, $\lambda_k$ determines the factor-level price response for each unit of factor risk $\text{var}(\mathbf{b}_k^\top \mathbf{X})$ induced by the factor flow $q_k$. Naturally, a greater risk aversion $\gamma_k$ or a smaller mass $\mu_k$ of marginal investors for a given factor leads to a larger price response $\lambda_k$.

We apply the LOOP to price individual assets relative to factors. That is, if two portfolios have the same payoff at time 1, they should have the same price at time 0. We denote the time-0 price of asset $n$ as $P_n(\mathbf{f})$, where the vector of asset flows in Proposition 1 is denoted as $\mathbf{f} = (f_1, f_2, \ldots, f_N)^\top$. Denote the vector of asset prices as $\mathbf{P}(\mathbf{f}) = (P_1(\mathbf{f}), P_2(\mathbf{f}), \ldots, P_N(\mathbf{f}))^\top$. The LOOP implies that

$$\mathbf{P}(\mathbf{f}) = \sum_{k=1}^{K} \frac{\text{cov}(\mathbf{X}, \mathbf{b}_k^\top \mathbf{X})}{\text{var}(\mathbf{b}_k^\top \mathbf{X})} P_k(q_k) = \sum_{k=1}^{K} \left( \lambda_k (q_k - S_k) \text{cov}(\mathbf{X}, \mathbf{b}_k^\top \mathbf{X}) + \frac{\text{cov}(\mathbf{X}, \mathbf{b}_k^\top \mathbf{X}) \mathbb{E}(\mathbf{b}_k^\top \mathbf{X})}{\text{var}(\mathbf{b}_k^\top \mathbf{X}) R_F} \right). \tag{4}$$

Next, we simplify the asset prices in (4) to obtain an empirically implementable factor model. Define *price impacts* as the time-0 percentage price change with and without flow $\mathbf{f}$,

$$\Delta \mathbf{p}(\mathbf{f}) = \left( \frac{P_1(\mathbf{f}) - P_1(\mathbf{0})}{P_1(\mathbf{0})}, \frac{P_2(\mathbf{f}) - P_2(\mathbf{0})}{P_2(\mathbf{0})}, \ldots, \frac{P_N(\mathbf{f}) - P_N(\mathbf{0})}{P_N(\mathbf{0})} \right)^\top. \tag{5}$$

Define *fundamental returns* as the asset returns from time 0 to 1 in the absence of flow,

$$\mathbf{R}_0 = \left( \frac{X_1}{P_1(\mathbf{0})}, \frac{X_2}{P_2(\mathbf{0})}, \ldots, \frac{X_N}{P_N(\mathbf{0})} \right)^\top. \tag{6}$$

We can then simplify equation (4) as

$$\Delta \mathbf{p}(\mathbf{f}) = \sum_{k=1}^{K} \lambda_k q_k \text{cov}(\mathbf{R}_0, \mathbf{b}_k^\top \mathbf{X}). \tag{7}$$

So far, the asset flows $\mathbf{f}$, factor flows $q_k$, and portfolio shares $\mathbf{b}_k$ are measured in the unit of shares. To further simplify (7), we change the unit of measurement to dollar value relative to the asset prices $\mathbf{P}(\mathbf{0})$, consistent with standard empirical practice.[6] Although we still employ the same notation throughout the paper for $\mathbf{f}$, $q_k$, and $\mathbf{b}_k$, they are now all measured

[6]Measurement in dollars is consistent with standard empirical practice for two reasons. First, the cross-sectional asset pricing literature uses dollar weights (e.g., Fama-French portfolio weights). Second, the demand-based asset pricing literature computes dollar flows using asset prices before the flow has arrived (see, e.g., Coval and Stafford, 2007 and Lou, 2012).

in dollars.[7] By implementing this unit change, we can further simplify equation (7).

**PROPOSITION 2.** *The factor model of price impacts is*

$$\Delta \mathbf{p}(\mathbf{f}) = \sum_{k=1}^{K} \lambda_k q_k \text{cov}(\mathbf{R}_0, \mathbf{b}_k^\top \mathbf{R}_0). \tag{8}$$

Proposition 2 shows that the equilibrium price impacts $\Delta \mathbf{p}(\mathbf{f})$ for the cross-section of assets are influenced by the new parameter $\lambda_k$, the factor flows $q_k$, and the quantity of risk exposure to the factors, represented by $\text{cov}(\mathbf{R}_0, \mathbf{b}_k^\top \mathbf{R}_0)$.

We highlight several unique features of our model that differentiate it from the literature. First, the cross-asset price impacts in our model explicitly depend on the fundamental return covariance, $\text{cov}(\mathbf{R}_0, \mathbf{b}_k^\top \mathbf{R}_0)$. This feature implies that flow into an asset results in a more substantial impact on assets with higher return covariance. By incorporating this aspect, our model can more accurately capture cross-asset price impacts present in the data.

Second, the new parameter $\lambda_k$ deserves special attention in terms of its unit. We multiply equation (8) with the factor's portfolio weights $\mathbf{b}_k$ and obtain

$$\lambda_k = \frac{\mathbf{b}_k^\top \Delta \mathbf{p}(\mathbf{f})}{q_k \text{var}(\mathbf{b}_k^\top \mathbf{R}_0)}. \tag{9}$$

On the right-hand side, the denominator $q_k \text{var}(\mathbf{b}_k^\top \mathbf{R}_0)$ is the total amount of risk induced by the factor flow. The numerator $\mathbf{b}_k^\top \Delta \mathbf{p}(\mathbf{f})$ is the factor-level price impact. Economically, $\lambda_k$ quantifies the impact of one unit of risk, induced by the flow, on the price for each factor $k$. Consequently, we term this new parameter $\lambda_k$ as *the price of flow-induced risk*.

The literature uses price elasticity $(dP/P)/(dQ/Q)$ to measure how a 1% change in quantity impacts the price. In contrast, in our setting, $\lambda_k$ in (9) differs by calculating how a single unit of risk, induced by the flow, impacts the price. There are two key distinctions:

[7] The conversion from the unit of shares to the dollar value proceeds as follows: the asset flow becomes $f_n \to f_n P_n(\mathbf{0})$, the flow factor becomes $q_k \to q_k P_k(0)$, and the portfolio weights become $b_n \to b_n P_n(\mathbf{0})/P_k(0)$. The portfolio relationship described in Proposition 1 remains unaltered after this conversion.

- Economically: Price impacts in our model arise because flow-absorbing investors are averse to the risk induced by flows. This differs from the traditional concept of price elasticity, which is based on the idea that investors are "inelastic" to flows per se.

- Technically: For long-short portfolios, the traditional measure $(dP/P)/(dQ/Q)$ is not a well-defined mathematical object. This is because the total quantity for long-short portfolios is zero ($Q = 0$), and division by zero is not allowed in mathematics.[8]

To put it simply, the newly introduced measure, $\lambda_k$, serves a different purpose than the traditional price elasticity measure. It is designed to capture the impact of risk on price, specifically for flow-absorbing investors, and is mathematically well-defined for long-short portfolios where the traditional measure fails.

Finally, the assumptions required to derive the factor model (8) can be relaxed. This flexibility includes permitting approximate, instead of exact, $K$-factor structures for payoffs and flows, endogenously generating the segmentation across the $K$ factors, and accommodating the dynamic autocorrelations of flows. An (2023) and An and Zheng (2023) consider these theoretical generalizations. In particular, even though our model is formulated within a static context, the cross-sectional factor structure in Proposition 2 remains the same when marginal investors anticipate variations in future flows. Consequently, in the dynamic setting, the empirical estimation also follows the same procedure that we describe next.[9]

## 4 Empirical Framework

Having established a theoretical basis for the factor model of price impacts in Section 3, this section presents the general estimation procedure.

[8]The total quantity $Q$ is computed similarly to the portfolio flow $dQ$ in Proposition 1, which projects the asset-level amount outstanding onto a set of portfolio weights. In most empirical studies, the market factor is included and its portfolio weights align perfectly with the asset-level outstanding $Q$. Hence, as usually anticipated, the $Q$ of the market factor equals the sum of all assets. Yet for all other portfolios, $Q = 0$.

[9]The only divergence from the static setting lies in the theoretical interpretation of $\lambda_k$. In the static model, $\lambda_k = \gamma_k/(\mu_k R_F)$, with $\gamma_k$ and $\mu_k$ representing the risk aversion and the mass of marginal investors of factor $k$. In the dynamic model, $\lambda_k$ also depends on the autocorrelations of flows (An and Zheng, 2023).

## 4.1 Constructing Portfolio Flows

Estimating the factor model of price impacts requires selecting $N$ test assets and $K$ factors, both of which can be portfolios of $M$ underlying stocks.[10] Therefore, we start by describing how to aggregate stock-level flows into portfolio flows.

The observable data consists of a panel of flow $f_{m,t}^{\{\text{stock}\}}$ into stock $m$ at time $t$. We aim to aggregate these stock-level flows into $K$ portfolios, where $w_{m,k}$ represents the weight of stock $m$ in portfolio $k$. (The aggregation to the $N$ test assets is conducted similarly.) As shown by Proposition 1, portfolio flows first go into portfolios and are then allocated to individual stocks according to the portfolio weights. Thus, to aggregate stock-level flows to portfolio flows, one should perform a cross-sectional regression of stock flows $f_{m,t}^{\{\text{stock}\}}$ on portfolio weights $w_{m,1}, w_{m,2}, \ldots, w_{m,K}$ in each period $t$, with the regression coefficients being the portfolio flows $q_{1,t}, q_{2,t}, \ldots, q_{K,t}$.

Formulating this procedure precisely, we represent the portfolio weight $w_{m,k}$ as an $M \times K$ matrix $\mathbf{W}$. The flow $q_{k,t}$ into portfolio $k$ at time $t$ is constructed as follows

$$\begin{pmatrix} q_{1,t} \\ q_{2,t} \\ \cdots \\ q_{K,t} \end{pmatrix} = (\mathbf{W}^\top \mathbf{W})^{-1} \mathbf{W}^\top \begin{pmatrix} f_{1,t}^{\{\text{stock}\}} \\ f_{2,t}^{\{\text{stock}\}} \\ \cdots \\ f_{M,t}^{\{\text{stock}\}} \end{pmatrix}. \tag{10}$$

It is worth noting that our method for aggregating portfolio flows differs from the literature's common approach that defines $q_{k,t} := \sum_{m=1}^{M} w_{m,k} f_{m,t}^{\{\text{stock}\}}$. The key economic intuition is that quantity is not the same as price, so it is not appropriate to simply apply the formula for aggregating portfolio returns to portfolio flows. Instead, our formula (10) is based on the portfolio flow theory presented in Proposition 1.

[10] Our framework is not limited to stocks. We use "stocks" because our empirical application uses stocks.

## 4.2 Estimation Procedure

In this section, we present a two-stage regression procedure for estimating our factor model.

The estimation involves $N$ test assets, $K$ factors, and $T$ periods. The data inputs consist of the return $r_{n,t}$ and flow $f_{n,t}$ of test asset $n$ at time $t$, along with the flow $q_{k,t}$ of factor $k$ at time $t$. These variables are defined for $n = 1, 2, \ldots, N$, $k = 1, 2, \ldots, K$, and $t = 1, 2, \ldots, T$. The construction of test asset and factor flows follows the approach detailed in Section 4.1. The data input does not require $\mathbf{b}_k = (b_{1,k}, b_{2,k}, \ldots, b_{N,k})^\top$, which is the portfolio weights of factor $k$ in terms of the $N$ test assets. The first-stage regression estimates these weights.

In Section 3, our model requires the measurement of flows in dollar amounts for cross-sectional analysis, such that $\mathbf{b}_k$ can be interpreted as portfolio weights corresponding to each dollar invested. However, the model does not impose any intertemporal constraints on flow normalization. In our empirical application, we normalize flows using the total stock market capitalization from the preceding period. This approach accommodates the increasing total market capitalization observed in the data and economically implies that the risk-bearing capacity of marginal investors is proportional to the total market capitalization. To streamline the subsequent regression analysis, we also eliminate the unconditional time-series mean of $r_{n,t}$, $f_{n,t}$, and $q_{k,t}$ for each test asset and factor.

The data generating process assumes that the price impact, as described in our model (8), occurs repeatedly over time. In each period $t$, flows $f_{n,t}$ arrive, leading to price impacts across all assets. Each asset $n$ also experiences a fundamental-driven return fluctuation in period $t$, denoted as $\xi_{n,t}$. The theoretical foundation of (8) assumes that the $K$ factors exhibit uncorrelated flows and fundamental returns. However, the $K$ factors in the data may not meet this condition, so a rotation is needed. The original factors are still denoted by $q_{k,t}$ (flows) and $\mathbf{b}_k$ (portfolio weights), while the rotated factors—those that have uncorrelated flows and fundamental returns—are denoted by $\tilde{q}_{k,t}$ and $\tilde{\mathbf{b}}_k$.

Given Proposition 2, the observed return $r_{n,t}$ of asset $n$ in period $t$ is modeled as follows

$$r_{n,t} = \sum_{k=1}^{K} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}. \tag{11}$$

Special attention should be paid to $\boldsymbol{\xi}_t = (\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{N,t})^\top$, the fundamental returns of the $N$ assets in period $t$. This term replaces $\mathbf{R}_0$ in the theoretical model (8). In equation (11), $\boldsymbol{\xi}_t$ serves a dual purpose. First, it represents the fundamental-return component of $r_{n,t}$. Second, it influences the price impact component by determining the quantity of risk exposure through the term $\text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$.

In the subsequent discussion, we present the two-stage procedure for estimating (11).

### 4.2.1 First-Stage Regression

The first stage of our regression analysis draws upon the portfolio flow theory outlined in Proposition 1. It involves performing individual time-series regressions for each asset, where the flow of an asset $f_{n,t}$ is regressed on the contemporaneous factor flows $q_{k,t}$,

$$f_{n,t} = \sum_{k=1}^{K} b_{n,k} q_{k,t} + e_{n,t}. \tag{12}$$

This regression yields the flow beta $b_{n,k}$, representing the flow of asset $n$ in response to the flow of factor $k$.

The crucial insight is that flow betas $b_{n,k}$ are the portfolio weights of factor $k$ in terms of the $N$ test assets. To see this, note that regression (12) implies that, all else equal, a one-dollar increase in factor-$k$ flow results in an increase of $\$b_{n,k}$ in asset-$n$ flow. Consequently, market clearing dictates that a one-dollar increase in factor-$k$ flow leads to a decrease of $\$b_{n,k}$ in marginal investors' holdings of asset $n$. Therefore, the flow betas $\mathbf{b}_k = (b_{1,k}, b_{2,k}, \ldots, b_{N,k})^\top$ represent the changes in marginal investors' holdings of the $N$ assets caused by a one-dollar

factor-$k$ flow. This interpretation aligns flow betas with the concept of portfolio weights.[11]

The residual $e_{n,t}$ in the first-stage regression (12) represents the idiosyncratic asset-level flows that are not explained by factor flows. Our factor model (11) does not use these idiosyncratic flows. Hence, if these idiosyncratic flows contribute minimally to the flow variation in the cross-section, our model provides an adequate approximation by relying on a few important factors.[12] Empirically, we indeed find a high $R^2$ for the first-stage regression. Additionally, the empirical test for the price impacts of the idiosyncratic flows is theoretically linked to the mean-variance optimal strategy that profits from flows, generalizing the well-known Gibbons, Ross, and Shanken (1989) test. Appendix A.3 provides the theoretical details, and we empirically apply our factor model to construct the optimal strategy.

### 4.2.2 Second-Stage Regression

In this section, we present the second stage, which involves implementing panel regression for the data generating process (11). The regression estimates the price of flow-induced risk $\lambda_k$ for each factor. A couple of points require attention.

First, the factor flows $\tilde{q}_{k,t}$ and portfolio weights $\tilde{\mathbf{b}}_k$ used in (11) are rotated from the data input $q_{k,t}$ and estimated $\mathbf{b}_k$ from the first-stage regression. This procedure involves finding some $K \times K$ matrix $\mathbf{O}$, such that the rotated factors have both uncorrelated fundamental returns and flows. Appendix A.1 provides technical details for constructing $\mathbf{O}$.

Second, the fundamental return $\boldsymbol{\xi}_t$ is not directly observable. This return is not only an input into the regression (11), contributing to the construction of the quantity of risk $\text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$, but also forms the regression residual. To tackle this issue, we resort to an iterative algorithm.[13] Initially, we set the fundamental return equal to the observed asset return (i.e., $\xi_{n,t} = r_{n,t}$) and carry out the second-stage regression. Following this, we use the

[11] The portfolio weights $b_{n,k}$ obtained in the first-stage regression should not be confused with the original portfolio weights $w_{m,k}$ that construct factor flows in (10). One set of weights pertains to the $N$ test assets, while the other is associated with the $M$ underlying stocks.

[12] This theoretical claim is made precise in An (2023) by generalizing the Ross (1976) APT argument.

[13] We have also tried to use long-horizon returns as a proxy for fundamental returns when constructing the quantity of risk, and this alternative approach produces empirical results similar to those in the paper.

regression residual, which corresponds to the model-implied fundamental return, as the new $\xi_{n,t}$. This procedure is repeated until $\xi_{n,t}$ reaches convergence.

Third, a consistent estimate for $\lambda_k$ requires $\text{cov}(\tilde{q}_{k,t}, \xi_{n,t}) = 0$, meaning that the factor flow $\tilde{q}_{k,t}$ is uncorrelated with the concurrent fundamental return $\xi_{n,t}$ of the assets. In our empirical application, we address the endogeneity concern by carefully selecting $\tilde{q}_{k,t}$ to be noisy shocks. We also implement robustness checks using alternative OLS and IV methods.

# 5 Empirical Application

In this section, we apply the empirical framework in Section 4 to the standard data (mutual fund flows) and standard factors (the Fama-French factors).

## 5.1 Data and Empirical Measures

We measure noisy flows using the mutual fund flow-induced trading, as proposed by Coval and Stafford (2007) and Lou (2012). We employ the Fama-French 5×5 portfolios, which are sorted based on size and book-to-market equity ratios, as our test assets. We then measure the returns $r_{n,t}$ and flows $f_{n,t}$ for these portfolios. Additionally, we measure the flows $q_{k,t}$ into the Fama-French three factors. In what follows, we provide a detailed explanation of the flow construction methodology.

First, we compute mutual fund flows in dollar amounts following standard procedures. In particular, we retrieve monthly mutual fund returns and characteristics from the CRSP Survivorship-Bias-Free Mutual Fund database, in addition to quarterly holdings from the Thomson/Refinitiv Mutual Fund Holdings Data (S12). Our sample period spans from 2000 through September 2020.[14] Our mutual fund sample comprises both active and passive mutual funds. We denote the total net assets (TNA) of mutual fund $m$ at the end of month

[14]The mutual fund industry witnessed significant growth and sustained inflows throughout the 1990s (Lou, 2012; Ben-David, Li, Rossi, and Song, 2022a). In the post-2000 era, the monthly flows of mutual funds have generally maintained relative stability, prompting us to commence our sample period from 2000, aligning with Gabaix and Koijen (2022).

$t$ as $\text{TNA}_{m,t}$, and the mutual fund's net-of-fee return in month $t$ as $r^{\text{fund}}_{m,t}$. The mutual fund flow in dollar amount is defined as follows:

$$f^{\text{fund}}_{m,t} = \text{TNA}_{m,t} - \text{TNA}_{m,t-1}(1 + r^{\text{fund}}_{m,t}). \tag{13}$$

We conduct a cross-validation of mutual funds' monthly returns and TNA obtained from the CRSP database with corresponding data from Morningstar and Thomson/Refinitiv. In the process, we manually rectify several data input inaccuracies. Detailed information regarding this correction process is provided in Appendix C.

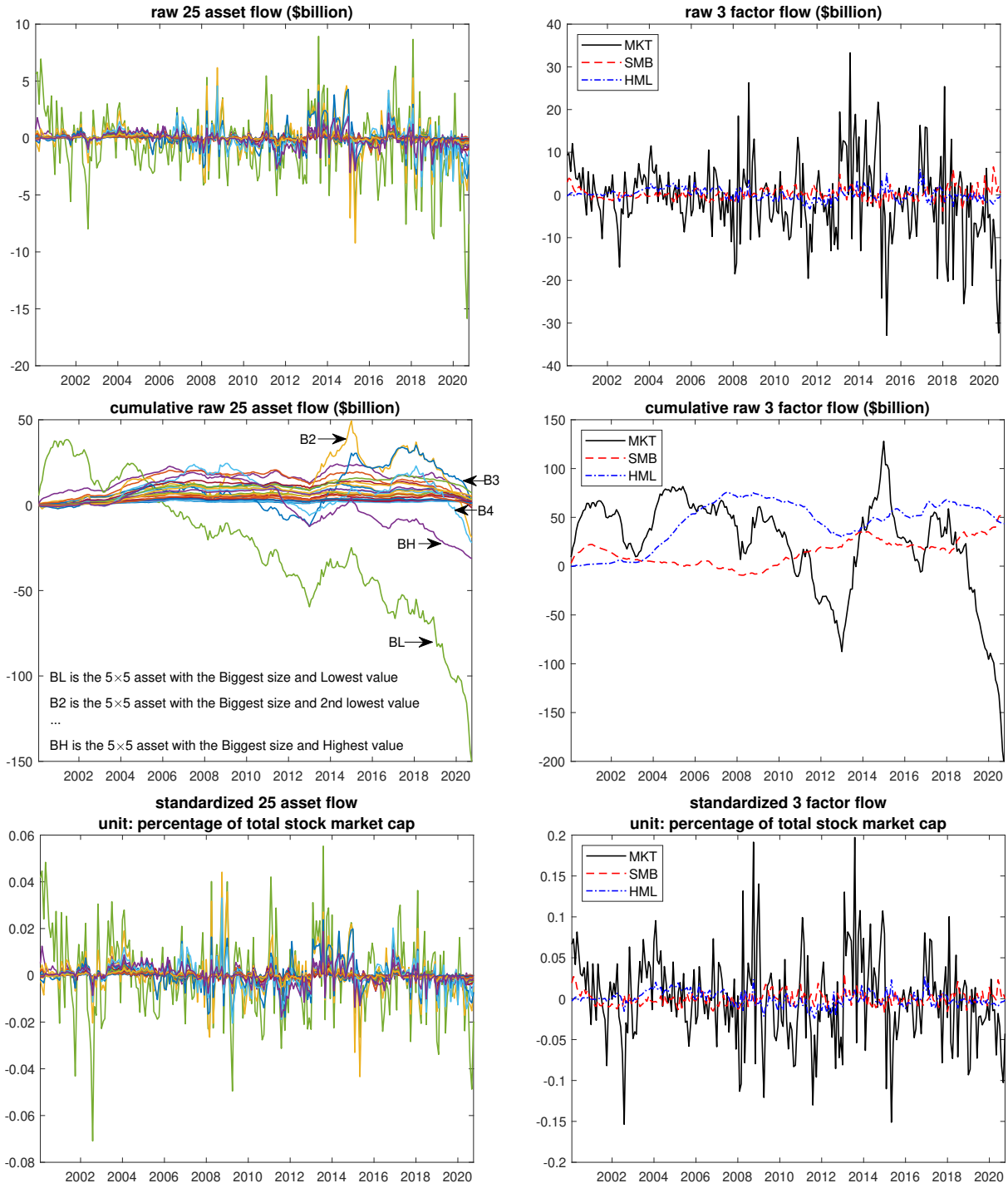
Second, we translate mutual fund flows into stock-level flows, adhering to the established assumption in the literature that mutual funds buy or sell stocks in proportion to their prior holdings. Crucially, we employ the two-quarter-lagged mutual fund holding to transform mutual fund flow into stock-level flows. For instance, we use the fund holding information from Q4 of the preceding year for mutual fund flows occurring in April, May, and June. This lag is based on two considerations. Firstly, the flow-induced trading construction formulated by Lou (2012) utilizes lagged mutual fund holding. The lag is used because while uninformed retail investors primarily drive mutual fund flows, mutual fund managers may possess private information.[15] The lag ensures that the constructed stock-level flows represent the non-discretionary component of mutual fund flows, which are induced by retail trading. Secondly, we impose a two-quarter lag to ensure that the holding information is observable for the out-of-sample tradable strategy.[16]

Third, we use stock-level flows and equation (10) to construct the Fama-French MKT, SMB, and HML flows $q_{k,t}$ and 5×5 test asset flows $f_{n,t}$. The top two panels of Figure 2 illustrate these monthly flows. The cumulative sum of test assets and factor flows are

[15] Frazzini and Lamont (2008) and Ben-David, Li, Rossi, and Song (2022b) provide evidence supporting the theory of uninformed mutual fund investors.

[16] The holding information is reported with a maximum statutory delay of 45 days (Christoffersen, Danesh, and Musto, 2015), which implies that Q1 holding information may not be observable in April. To remain conservative, we use the holding information from Q4 of the previous year for the flows in April, May, and June. Our results remain robust if we choose to alternatively use Q1 holding information.

**Figure 2. Time series of 25 test asset flows and three factor flows**




Note: In the first two figures, we plot the monthly flows for the test assets and factors. In the subsequent two figures in the middle, we plot the cumulative sum of the flows for both test assets and factors. In the final two figures at the bottom, we standardize the test asset and factor flows by dividing them by the total stock market capitalization from the previous month and then subtracting the unconditional time-series mean. The sample period extends from January 2000 to September 2020.

displayed in the two middle panels. We observe that large companies demonstrate substantial fluctuations in flows due to their considerable market capitalization. Additionally, we notice a strong commonality among the flows of the 25 assets. The MKT flows display the most significant fluctuation, but the SMB and HML flows also show a sizable variation.

As discussed in Section 4.2, we standardize the test asset and factor flows by dividing them by the total stock market capitalization from the previous month and then subtracting the unconditional time-series mean. The bottom two panels of Figure 2 plot the standardized test asset and factor flows, which are used in subsequent regressions. The MKT factor can experience inflows and outflows as large as 0.2% of the total stock market capitalization within a month.[17] The SMB and HML flows display lesser variations. The pairwise correlations between the flows of MKT and SMB, MKT and HML, and HML and SMB are 0.11, 0.25, and -0.11, respectively.

## 5.2 First-Stage Regression

Table 1 presents the results of the first-stage regression, as detailed in Section 4.2.1,

$$f_{n,t} = b_{n,\text{MKT}} q_{\text{MKT},t} + b_{n,\text{SMB}} q_{\text{SMB},t} + b_{n,\text{HML}} q_{\text{HML},t} + e_{n,t}, \tag{14}$$

In this time-series regression, the flow $f_{n,t}$ of each asset $n$ is regressed against the factor flows $q_{\text{MKT},t}$, $q_{\text{SMB},t}$, and $q_{\text{HML},t}$ to estimate the values of $b_{n,\text{MKT}}$, $b_{n,\text{SMB}}$, and $b_{n,\text{HML}}$. The regression $R^2$ is provided in the upper-middle panel of Table 1, ranging from 50% for smaller companies to 80% for larger ones. This high $R^2$ value implies that factor flows account for a substantial proportion of the common variations in asset flows. This finding supports our approach of using three factors to estimate the price impacts of the 5×5 assets.

As elaborated in Section 4.2.1, flow betas are economically interpreted as portfolio

[17] Flow into the MKT factor induced by mutual fund trading is less volatile than aggregate flow into mutual funds (which Appendix Figure A.1 shows) for two reasons. First, mutual funds do not invest 100% in stocks. Second, by our construction (10), flow into the MKT factor is generally less volatile than the sum of flows into all stocks when idiosyncratic flows are present.

**Table 1. First-stage regression: asset flows on factor flows**

| | market cap. weight $w_n$ | | | | | regression $R^2$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Low | 2 | 3 | 4 | High | Low | 2 | 3 | 4 | High |
| Small | 0.0042 | 0.0034 | 0.0040 | 0.0045 | 0.0044 | 48.03% | 52.44% | 52.61% | 55.12% | 52.13% |
| 2 | 0.0081 | 0.0072 | 0.0071 | 0.0065 | 0.0042 | 59.38% | 63.04% | 58.81% | 62.10% | 63.67% |
| 3 | 0.0142 | 0.0130 | 0.0107 | 0.0091 | 0.0060 | 60.50% | 58.31% | 62.55% | 57.68% | 62.68% |
| 4 | 0.0371 | 0.0272 | 0.0190 | 0.0148 | 0.0112 | 57.79% | 58.16% | 54.72% | 66.86% | 67.10% |
| Big | 0.3651 | 0.1790 | 0.1079 | 0.0844 | 0.0478 | 85.68% | 85.68% | 84.56% | 87.44% | 73.30% |

| | MKT flow beta $b_{n,\text{MKT}}$ | | | | | SMB flow beta $b_{n,\text{SMB}}$ | | | | | HML flow beta $b_{n,\text{HML}}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Low | 2 | 3 | 4 | High | Low | 2 | 3 | 4 | High | Low | 2 | 3 | 4 | High |
| Small | 0.0022 | 0.0019 | 0.0027 | 0.0039 | 0.0033 | 0.0357 | 0.0298 | 0.0378 | 0.0357 | 0.0331 | 0.0258 | 0.0275 | 0.0377 | 0.0465 | 0.0464 |
| 2 | 0.0077 | 0.0073 | 0.0071 | 0.0073 | 0.0049 | 0.0844 | 0.0739 | 0.0649 | 0.0570 | 0.0378 | 0.0567 | 0.0697 | 0.0683 | 0.0681 | 0.0439 |
| 3 | 0.0139 | 0.0125 | 0.0096 | 0.0081 | 0.0046 | 0.1120 | 0.0848 | 0.0588 | 0.0455 | 0.0278 | 0.0895 | 0.1072 | 0.0948 | 0.0858 | 0.0591 |
| 4 | 0.0392 | 0.0252 | 0.0161 | 0.0143 | 0.0129 | 0.1153 | 0.0359 | 0.0003 | 0.0024 | -0.0019 | 0.1045 | 0.1352 | 0.0976 | 0.0879 | 0.0855 |
| Big | 0.3252 | 0.1780 | 0.1169 | 0.0849 | 0.0575 | 0.0140 | -0.2877 | -0.2329 | -0.2607 | -0.1041 | -0.3386 | -0.1532 | 0.0654 | 0.2350 | 0.2007 |

| sum of $b_{n,\text{MKT}}$ coefficients | | | sum of $b_{n,\text{SMB}}$ coefficients | | | sum of $b_{n,\text{HML}}$ coefficients | | |
|---|---|---|---|---|---|---|---|---|
| all | positive | negative | all | positive | negative | all | positive | negative |
| 0.97 | 0.97 | 0.00 | 0.10 | 0.99 | -0.89 | 1.45 | 1.94 | -0.49 |

| | $t(b_{n,\text{MKT}})$ | | | | | $t(b_{n,\text{SMB}})$ | | | | | $t(b_{n,\text{HML}})$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Small | 4.64 | 4.34 | 4.80 | 6.60 | 4.74 | 8.65 | 9.40 | 7.73 | 6.07 | 5.37 | 8.18 | 9.85 | 9.96 | 12.44 | 11.20 |
| 2 | 7.64 | 8.48 | 8.13 | 8.63 | 8.31 | 10.40 | 8.66 | 7.63 | 6.85 | 6.75 | 9.96 | 14.25 | 13.65 | 12.81 | 14.46 |
| 3 | 8.21 | 8.34 | 8.12 | 6.83 | 6.24 | 10.55 | 7.20 | 5.65 | 5.11 | 5.00 | 10.32 | 12.56 | 16.26 | 13.29 | 15.54 |
| 4 | 10.38 | 9.39 | 8.44 | 10.30 | 9.29 | 5.47 | 2.52 | 0.02 | 0.29 | -0.25 | 5.16 | 8.83 | 10.32 | 12.06 | 10.93 |
| Big | 21.37 | 22.34 | 25.74 | 22.49 | 9.98 | 0.23 | -9.04 | -9.11 | -15.62 | -5.92 | -4.49 | -3.61 | 2.79 | 12.17 | 6.66 |

Note: We run asset-by-asset time-series regressions (14) of asset flows on factor flows. The 5×5 assets are sorted based on size (small to big) and book-to-market equity (low to high). We report the regression $R^2$, point estimates and t-statistics for flow betas $b_{n,\text{MKT}}$, $b_{n,\text{SMB}}$, and $b_{n,\text{HML}}$, as well as the market capitalization weight $w_n$ of the 5×5 assets. To obtain $w_n$, we calculate the ratio of the market capitalization of asset $n$ to the total stock market for each month. We then take a simple average of these values over time. The t-statistics are calculated from heteroskedasticity-robust standard errors.

weights. We now empirically investigate this relationship factor-by-factor.

First, the middle-left panel of Table 1 displays the MKT flow beta $b_{n,\text{MKT}}$. For instance, a $b_{\text{BL,MKT}}$ value of 0.3252 implies that an increase of \$1 in the MKT flow results in a \$0.3252 increase in the BL asset flow. The top-left panel of Table 1 provides the market capitalization weight $w_n$ of the 5×5 assets, calculated as the time-series average of the ratio of asset $n$'s market capitalization to the total stock market capitalization. The value $w_n$ serves as a proxy for the weights of the Fama-French MKT portfolio on the 5×5 test assets.

In line with the theory, for all assets $n$, we observe that the flow beta $b_{n,\text{MKT}}$ is closely aligned with the Fama-French MKT weight $w_n$. All $b_{n,\text{MKT}}$ values are positive, and their sum approximates one. This finding aligns with the intuitive understanding that, on average, mutual funds hold the market portfolio. However, the significance of this finding lies in demonstrating that our new methods of constructing factor flows and interpreting flow betas indeed bear out in the data.

Second, the middle-middle panel of Table 1 presents the SMB flow beta $b_{n,\text{SMB}}$. We notice that $b_{n,\text{SMB}}$ resembles the Fama-French small-minus-big construction. For instance, $b_{n,\text{SMB}}$ is positive for small companies, while for large companies, $b_{n,\text{SMB}}$ is negative. Furthermore, the sum of all positive $b_{n,\text{SMB}}$ values is 0.99, and the sum of all negative $b_{n,\text{SMB}}$ values is $-0.89$. This means the absolute value of both sums is nearly one, once again aligning empirical evidence with theoretical interpretation.

Third, the middle-right panel of Table 1 presents the HML flow beta $b_{n,\text{HML}}$. We find that only the BL and B2 assets have negative flow betas, while all other assets have positive flow betas. The sum of positive $b_{n,\text{HML}}$ values is nearly 2, and the sum of negative $b_{n,\text{HML}}$ values is approximately $-0.5$. While these flow beta estimates broadly align with the Fama-French HML construction, there are significant numerical differences. What could explain this notable disparity? The key lies in the fact that the BL and B2 assets, representing large-growth companies, collectively account for over 50% of the total stock market capitalization. Therefore, the empirical evidence suggests that trading flows along the value direction essen-

tially represent a contrast between large-growth companies and all other companies. This substantially deviates from the Fama-French 2×3 construction.

In summary, the first-stage regression yields the flow betas $b_{n,k}$, which represent the relationship between the flow of asset $n$ and the flow of factor $k$. Empirically, these flow betas align with the theoretical interpretation of portfolio weights. With these portfolio weights in hand, we can now proceed to execute the second-stage regression.

## 5.3 Second-Stage Regression

Table 2 presents the results of the second-stage regression, as detailed in Section 4.2.2,

$$r_{n,t} = \sum_{k \in \{\text{MKT,SMB,HML}\}} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}. \tag{15}$$

Recall that one needs to rotate the MKT, SMB, and HML factors to have uncorrelated flows and fundamental returns. For rotated factors, $\tilde{q}_{k,t}$ represents the flows, and $\tilde{\mathbf{b}}_k$ denotes the portfolio weights.[18] The details of the rotation can be found in Appendix Table A.1.

The first column of Table 2 presents the estimated price of flow-induced risk $\lambda_k$. This measurement represents the price elasticity of each factor in response to one unit of risk induced by the flow. As discussed in Section 3, we use this new measure because the conventional price elasticity metric is not applicable to long-short factor portfolios.

To interpret the estimated $\lambda_{\text{MKT}} = 9.54$, we refer to the equilibrium condition (3) from the theory section: $\lambda_k = \gamma_{\text{MKT}}/(\mu_{\text{MKT}} R_F)$, where $\gamma_{\text{MKT}}$ and $\mu_{\text{MKT}}$ denote the risk aversion and mass of the MKT factor's marginal investors, respectively, and $R_F$ represents the gross risk-free rate. Assuming $\gamma_{\text{MKT}} \approx 3$ and $R_F \approx 1$, we obtain $\mu_{\text{MKT}} \approx 0.3$, meaning that 0.3% of the total stock market capitalization actively responds to flow.[19] While this figure may

[18] The purpose of the rotation is to ensure a strict alignment between the empirical regression and the theoretical foundation presented in Section 3. However, even without implementing this rotation, the empirical results obtained from running the second-stage regression are quite similar.

[19] Recall that the unit of flow is expressed as a percentage of the total stock market capitalization. Our choice of risk aversion $\gamma_{\text{MKT}} \approx 3$ is derived from equation (9.6) in Cochrane (2009), combined with the fact that the Sharpe ratio of the stock market is roughly 0.5 and the volatility is 0.16.

**Table 2. Second-stage regression: returns on factor flows × quantity of risk exposure**

| | total return OLS | intraday return OLS | intraday return IV |
|---|---|---|---|
| $\lambda_{\text{MKT}}$ | 9.54 | 5.99 | 6.76 |
| | (14.48) | (12.07) | (2.83) |
| $\lambda_{\text{SMB}}$ | 109.93 | 64.42 | 151.56 |
| | (9.84) | (6.19) | (3.49) |
| $\lambda_{\text{HML}}$ | 65.21 | 42.13 | 136.94 |
| | (2.59) | (1.90) | (1.20) |

Note: In this table, we run the second-stage regression of 5×5 asset returns on the product of factor flows and the quantity of risk to estimate the price of flow-induced risk. The unit of flow is expressed as a percentage of the total stock market capitalization, and the quantity of risk is expressed in terms of the annualized variance in returns. The first two columns display the OLS estimation results using total returns and intraday (open-to-close) returns. The third column outlines the IV estimation results using intraday returns, in which each factor flow is instrumented by the factor's concurrent overnight (close-to-open) return and the difference between one-month and half-year lagged flows. The figures in parentheses represent the t-statistics, computed from heteroskedasticity-robust standard errors.

seem small, it actually aligns with Gabaix and Koijen (2022), who find that the market is approximately 100 times less elastic than theoretical models suggest. Translated into our terms, this means that less than 1% of the market actively responds to flow as mean-variance optimizing marginal investors.

Performing the same calculation for the SMB and HML factors suggests that the market is even less elastic for these factors. Specifically, our estimates suggest that 0.03% and 0.05% of the total stock market capitalization actively respond to SMB and HML flows, respectively. These empirical findings lend support to our theoretical assumption that different factors have distinct segments of flow-absorbing investors. One possible explanation for these results lies in the differing investment mandates. For instance, an insurance company that solely focuses on large-cap stocks could absorb MKT flows but would not be able to elastically absorb SMB flows, which would require trading in small-cap stocks.

Lastly, we address potential endogeneity issues concerning the estimates of $\lambda_k$. For an unbiased estimate, it is essential that the factor flow $\tilde{q}_{k,t}$ is uncorrelated with the concurrent fundamental return $\xi_{n,t}$. By constructing noisy flows $\tilde{q}_{k,t}$ using lagged fund holdings, we partially alleviate this endogeneity concern. This is because $\tilde{q}_{k,t}$ serves as a proxy for the

portion of mutual fund flows that are mechanically driven by retail investors' buying and selling of mutual fund shares, rather than by the discretionary allocation of mutual fund managers. It is unlikely that retail investors are informed about $\xi_{n,t}$.

To address potential endogeneity concerns arising from mutual fund investors chasing concurrent returns, we perform additional robustness checks. We follow the methodology of Li (2022), where we substitute asset return $r_{n,t}$ with intraday return $r_{n,t}^{\{\text{intraday}\}}$, which is the monthly aggregation of all open-to-close returns for each trading day.[20] The underlying premise is that intraday returns are more likely to reflect institution trading, while overnight (close-to-open) returns more likely reflect retail trading (Lou, Polk, and Skouras, 2019, 2022; Bogousslavsky and Muravyev, 2021). To isolate the price impacts driven by mutual fund flows, we utilize the intraday return $r_{n,t}^{\{\text{intraday}\}}$ as the dependent variable in regression (15). The regression results are reported in the second column of Table 2. The estimated $\lambda_k$ is similar to, albeit smaller than, those in the first column.

Moreover, we supplement the intraday return OLS estimate with an IV strategy. The first instrument for $\tilde{q}_{k,t}$ is the factor's concurrent overnight return, represented by $\sum_{m=1}^{25} \tilde{b}_{m,k} r_{m,t}^{\{\text{overnight}\}}$. The IV relevance condition is met as factor flows are positively correlated with concurrent night returns, attributable to the return-chasing behavior of mutual fund investors. Furthermore, given that overnight and intraday returns are non-overlapping, the IV exclusion restriction is likely to be satisfied. The second instrument is $\tilde{q}_{k,t-1} - \tilde{q}_{k,t-6}$, which represents the difference between lagged flows over one month and half a year. The relevance of this instrument stems from the serial correlation observed in factor flows. The exclusion comes from the unlikelihood that lagged flows would provide information about the fundamental return $\xi_{n,t}$. Importantly, using the difference in lagged flows, $\tilde{q}_{k,t-1} - \tilde{q}_{k,t-6}$, as the instrument, rather than the lagged flow $\tilde{q}_{k,t-1}$ itself, enables us to difference out potential confounding channels, such as the possibility of lagged flow inducing future price reversion.

[20] Following Lou, Polk, and Skouras (2019), we construct the intraday and overnight returns on each date $t$ as $r_t^{\{\text{intraday}\}} = \text{close}_t/\text{open}_t - 1$ and $r_t^{\{\text{overnight}\}} = (1 + r_t^{\{\text{close-to-close}\}})/(1 + r_t^{\{\text{intraday}\}}) - 1$. We source the price data from CRSP and adjust daily close-to-close returns for corporate actions such as stock splits.

The third column of Table 2 reports the IV results.[21] The point estimates largely align with those presented in the first two columns.

In summary, the second-stage regression yields $\lambda_k$, which represents the price elasticity of each factor in response to one unit of risk induced by the flow. Empirically, the substantial variation in $\lambda_k$ values supports our theoretical premise that different factors have distinct segments of flow-absorbing marginal investors.

## 5.4 Evaluating Model Fit

Ultimately, our model aims to explain the price impacts and cross impacts between the Fama-French 5×5 assets. Therefore, in this section, we show that our estimated model indeed fits the patterns of price impacts along size and value directions.

First, we investigate the ability of our model to account for the $R^2$ value in the return-on-flow regression. As a benchmark, we carry out a time-series regression of each asset's return $r_{n,t}$ against its respective flow $f_{n,t}$,

$$r_{n,t} = \eta_n f_{n,t}/w_n + \epsilon_{n,t}. \tag{16}$$

In this regression, we normalize the flow by the asset's market capitalization weight $w_n$, ensuring that our estimated multipliers $\eta_n$ align with standard price impact multipliers.

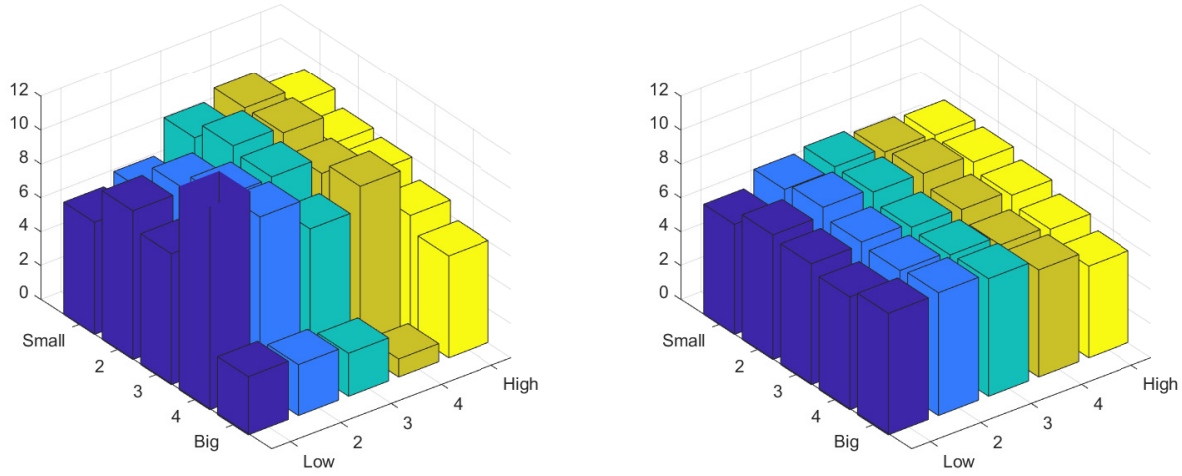
Panel A of Figure 3 reports the $R^2$ values from regression (16) for each 5×5 asset. The raw data analysis indicates an average $R^2$ value of approximately 8% across all assets. In contrast, Panel B reports the $R^2$ values derived from our factor-model regression (15) for each 5×5 asset.[22] We find an average $R^2$ value of about 7%, which broadly aligns with the $R^2$ values obtained from the raw data.[23]

[21] The IV first-stage regression results are provided in Appendix Table A.2.

[22] For each asset $n$, the $R^2$ is computed as the ratio of the variance of model-implied price impact $\sum_{k\in\{\text{MKT,SMB,HML}\}} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$ to the variance of $r_{n,t}$.

[23] We have also tried to regress an asset's return $r_{n,t}$ on the 5×5 flow into all test assets, instead of solely on its own flow $f_{n,t}$. This alternative approach, however, encounters issues of overfitting and multicollinearity, given the strong commonality exhibited by the 5×5 flow. Despite these issues, we find that the $R^2$ value for

**Figure 3. Regression $R^2$ of raw and model-implied price impact**




**(A) raw $R^2$** **(B) model-implied $R^2$**

| | mean | std | P25 | median | P75 |
|---|---|---|---|---|---|
| raw $R^2$ | 7.89% | 2.79% | 6.85% | 8.78% | 9.93% |
| model-implied $R^2$ | 6.99% | 0.55% | 6.65% | 7.10% | 7.35% |

Note: In panel (A), we report the $R^2$ values from regression (16), where we regress return on flow for each 5×5 asset. In panel (B), we report the $R^2$ values from our factor-model regression (15) for each 5×5 asset. The table at the bottom provides the summary statistics for these $R^2$ values.

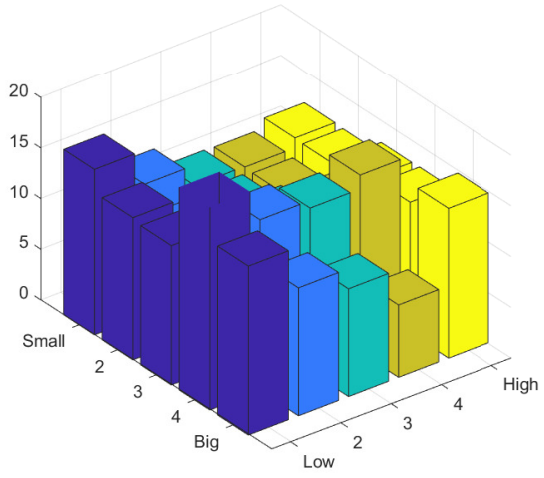
Second, we investigate the ability of our model to account for the price multipliers in the return-on-flow regression. Panel A of Figure 4 reports the price multiplier $\eta_n$ from regression (16) for each 5×5 asset.[24] We observe that the raw price multipliers for the majority of assets generally range from 10 to 15. Panel B reports the model-implied multipliers. Here, we use the model-implied price impact $\sum_{k\in\{\text{MKT,SMB,HML}\}} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$ as the left-hand side of regression (16). The model-implied price multiplier is typically around 10, which aligns with the raw price multiplier for most assets, albeit the magnitudes are slightly smaller.

Furthermore, as elaborated in Section 3, the use of return covariances enables our factor model (15) to accurately capture cross-asset price impacts. We now empirically validate this

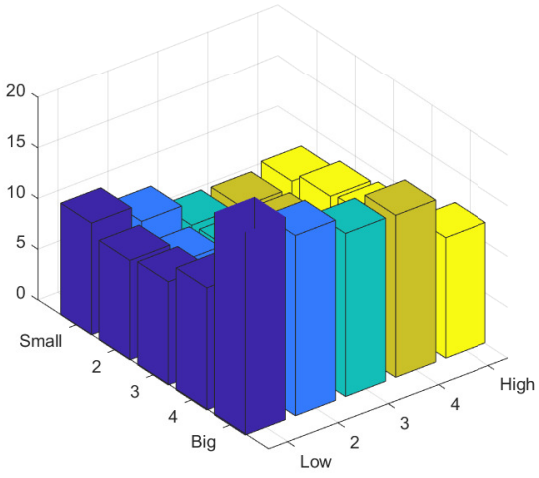
this alternative regression only increases to about 20%. This result suggests that our factor model continues to explain a significant portion of the $R^2$, even when compared with an artificially inflated benchmark.

[24]The heteroskedasticity-robust t-statistics are around 4 for most assets.

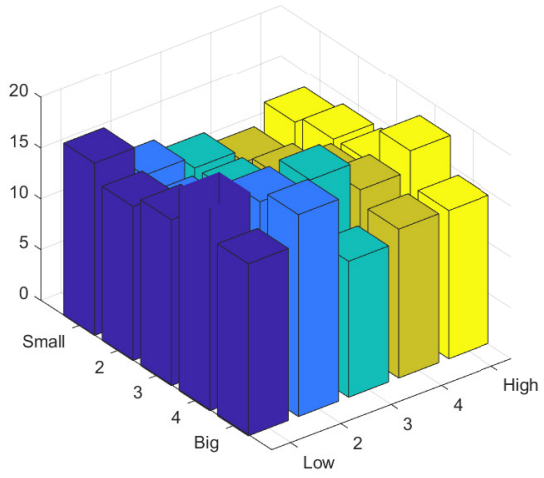
**Figure 4. Multipliers of raw and model-implied price impact**



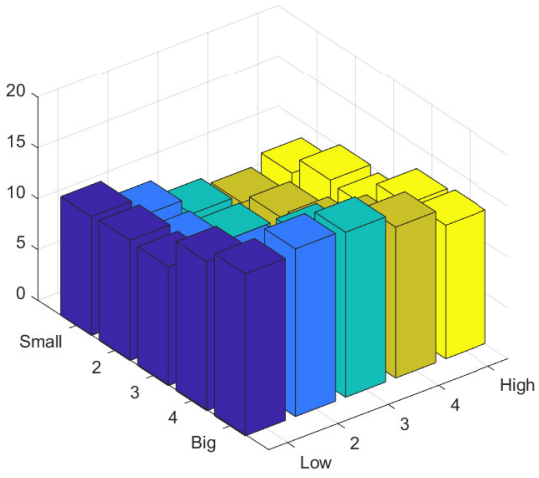

**(A) raw price multiplier**




**(B) model-implied price multiplier**




**(C) raw cross multiplier**




**(D) model-implied cross multiplier**

| | mean | std | P25 | median | P75 |
|---|---|---|---|---|---|
| raw price multiplier | 13.24 | 3.02 | 10.99 | 12.62 | 15.11 |
| model-implied price multiplier | 10.49 | 3.81 | 7.66 | 9.75 | 11.41 |
| raw cross multiplier | 15.18 | 2.47 | 13.33 | 14.58 | 16.93 |
| model-implied cross multiplier | 11.38 | 2.78 | 9.12 | 11.68 | 13.13 |

Note: In panel (A), we report the price multiplier from regression (16), where we regress return on flow for each 5×5 asset. In panel (B), we report the model-implied multiplier, where we regress the model-implied price impact on flow for each 5×5 asset. Panels (C) and (D) replicate the methods of panels (A) and (B) respectively, but with one key distinction: we use the average flow into the neighboring 5×5 assets as the regression's right-hand side. The table at the bottom provides the summary statistics for these multipliers.

point. Panel C of Table 4 reports the raw cross-impact multipliers. For each 5×5 asset $n$, we regress its return on the average flow into the neighboring 5×5 assets,

$$r_{n,t} = \phi_n \left( \sum_{m \text{ neighboring to } n} f_{m,t}/w_m \right) \Big/ (\text{number of } m \text{ neighboring to } n) + \epsilon_{n,t}. \tag{17}$$

The result reveals substantial cross multipliers, averaging around 15. Intriguingly, most raw cross multipliers exceed the raw price multipliers shown in Panel A. These significant cross multipliers originate from the highly correlated flows and fundamental returns of neighboring 5×5 assets, which are sorted based on specific characteristics. This evidence highlights that cross multipliers constitute a vital feature of the data that ought not to be overlooked.

Panel D reports the model-implied cross multipliers, demonstrating our model's capability to capture cross impacts in the data. As in the previous analysis, we use the model-implied price impact as the left-hand side of regression (17). The model-implied cross multiplier generally hovers around 11, which aligns with the raw cross multiplier for most assets, albeit the magnitudes are smaller.

In summary, this section empirically shows that our factor model well explains the price impacts and cross impacts among the Fama-French 5×5 assets. The evidence supports our initial hypothesis that noisy flows impact cross-sectional asset prices through risk factors.

## 5.5 Trading Strategy

In this section, we apply our factor model to construct the optimal strategy to profit from the flow. We show that our strategy can enhance traditional fundamental investing approaches.

### 5.5.1 Strategy Construction

We begin by delving into the rationale behind our strategy. The underlying idea is that noisy flows distort prices from their fundamental values, implying that flow-induced price impacts should see long-term reversions. Intuitively, the mean-variance optimal reversion

strategy should trade against the portfolio that maximizes the ratio of its price impact to its risk. Consequently, we designate this portfolio as the Maximum-Price-Impact-Ratio (MPIR) portfolio. The associated strategy, which shorts this portfolio, is henceforth referred to as the MPIR strategy. Appendix A.2 offers a comprehensive formulation of the mean-variance optimization of price impacts.

Given that our factor model (15) robustly characterizes the 5×5 assets' price impacts using only three factors, the model-implied MPIR strategy takes a simple form,

$$\sum_{k\in\{\text{MKT,SMB,HML}\}} -\lambda_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k. \tag{18}$$

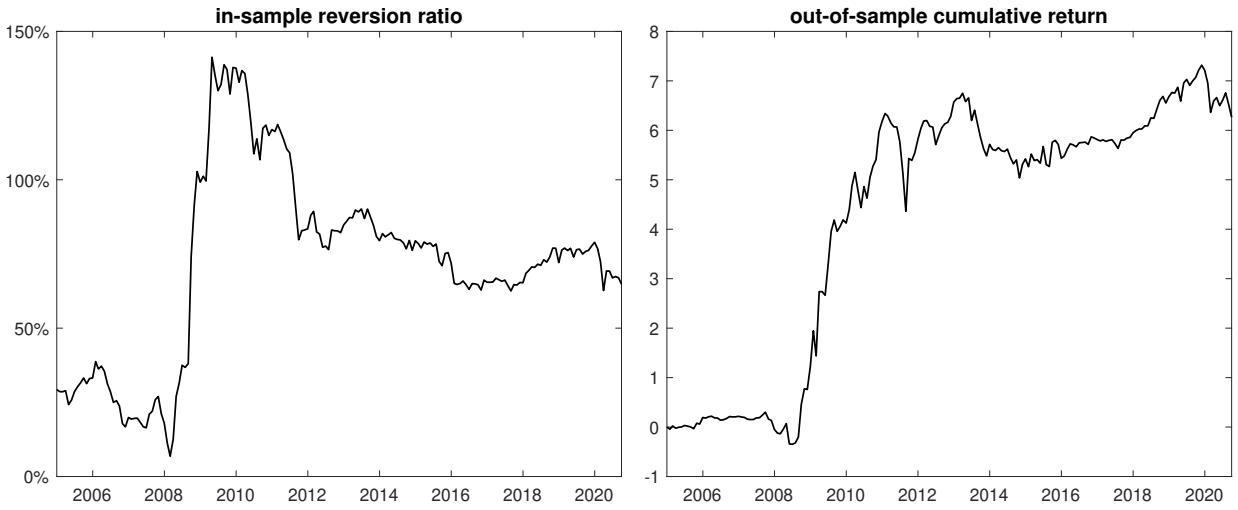
The MPIR strategy shorts the factor portfolio $\tilde{\mathbf{b}}_k$ if there is a positive inflow $\tilde{q}_{k,t}$ and longs if there is an outflow. The dynamic strategy changes every month $t$ depending on the factor flow $\tilde{q}_{k,t}$. The magnitude of the long/short position is proportional to the estimated $\lambda_k$.

Three features of the MPIR strategy warrant discussion. First, even though the MPIR strategy is formed using the 5×5 assets, its portfolio weights rely exclusively on the three factors' estimated $\lambda_k$. Importantly, it does not use the 5×5 multipliers estimated in Figure 4. The rationale is that factor-level optimization, unlike asset-level optimization, delivers more consistent out-of-sample price reversion. This is akin to the conventional asset pricing approach, which utilizes factors to estimate the mean-variance optimal portfolio.

Second, the MPIR strategy differs from a simplistic strategy that indiscriminately trades against all flows, which can be represented by $\sum_{k\in\{\text{MKT,SMB,HML}\}} -\tilde{q}_{k,t}\tilde{\mathbf{b}}_k$. This distinction arises because the compensation $\lambda_k$ for absorbing flow-induced risk varies across factors, as demonstrated in Section 5.3. Therefore, the optimal strategy assigns greater weight to factors that offer higher compensation per unit of risk.

Third, although the flow in month $t$ distorts the price within the same month, the price might not revert entirely in month $t+1$, potentially taking a longer period. Consequently, to construct the optimal portfolio for month $t+1$, we implement a staggered strategy that

**Figure 5. MPIR strategy's in-sample reversion and out-of-sample cumulative return**




Note: In the left panel of the figure, we display $\text{REV}^{\{t\}}$, which represents the ratio of the average one-month-forward reversion to the average model-implied price impact of the staggered MPIR strategy. This ratio is calculated for each training window that begins in January 2000 and ends in month $t$. The right panel illustrates the cumulative returns of the MPIR strategy throughout the out-of-sample period, extending from January 2005 to September 2020.

combines the portfolios (18) from the previous six months, from $t-5$ to $t$, using equal weights. Section 5.5.3 later shows the robustness of altering the initial and final months of the staggered strategy.

Our initial assessment aims to determine the effectiveness of the MPIR strategy within the sample. This assessment utilizes expanding estimation windows, beginning from January 2000 and ending in December 2004, with the out-of-sample testing period extending from January 2005 to September 2020. With each training window, the strategy is implemented, and the strategy's in-sample price reversion is computed. Subsequently, the ratio of the average reversion to the average model-implied price impact is calculated (for the precise formula, refer to Appendix A.4). Ideally, this reversion ratio should be 100%.

The left panel of Figure 5 presents the reversion ratio $\text{REV}^{\{t\}}$ corresponding to each training window that ends in month $t$. In most of the windows, this ratio falls below 100%, signifying that the model-implied price impact does not achieve full reversion. Notably, the reversion has exhibited considerable improvement following the financial crisis of 2008. This aligns with existing literature findings that after the crisis, marginal investors demonstrate

reduced willingness to undertake substantial risk (Du, Tepper, and Verdelhan, 2018).

The fact that the in-sample reversion is less than 100% suggests a potential overestimation of the concurrent price impact in our model. To address this, for each training window, we normalize the estimated $\lambda_k$ by $\text{REV}^{\{t\}}$ if $\text{REV}^{\{t\}}$ is less than 100%. These adjusted $\lambda_k$ values, which are normalized using in-sample data, are then employed to construct the out-of-sample MPIR strategy for the succeeding month $t+1$. Section 5.5.3 later investigates the robustness of our approach when this normalization is not implemented.

The right panel of Figure 5 plots the cumulative returns of the MPIR strategy. For the out-of-sample period extending from January 2005 to September 2020, the strategy delivers an annualized Sharpe ratio of 0.46. The performance of the MPIR strategy markedly improves during the 2008 financial crisis, aligning with the theory that marginal investors demand higher risk compensation during periods of crisis, which tends to enhance the performance of reversion strategies (Nagel, 2012).

### 5.5.2 MPIR Strategy Improves Fundamental Investing

While the MPIR strategy exhibits a commendable Sharpe ratio, it is crucial to note that the strategy is not constructed with the primary intention of maximizing the Sharpe ratio. Instead, the objective is to form an optimal trading strategy that capitalizes on flow. Our strategy is useful because it then enhances fundamental investing, which typically overlooks flow-related information. Appendix A.2 provides a theoretical foundation for this assertion, and in this section, we substantiate it empirically.

We denote the MPIR strategy's excess return in month $t+1$ as $\tilde{r}^*_{t+1}$. Fundamental investing is proxied using the 154 anomaly portfolios from Jensen, Kelly, and Pedersen (2021), which include the Fama-French-Carhart six factors among a broad array of other firm characteristics-based anomaly portfolios.[25] We denote the excess return of anomaly portfolio $j$ in month $t+1$ as $r_{j,t+1}$. The excess return of the combined portfolio $j$ in month

[25]We incorporate the 153 anomaly portfolios from Jensen, Kelly, and Pedersen (2021), along with the market excess return.

$t+1$ is defined as

$$r^*_{j,t+1} := r_{j,t+1} + w_{j,t}\tilde{r}^*_{t+1}, \tag{19}$$

where $w_{j,t}$ denotes the mean-variance optimal mixing ratio, which is estimated in-sample (refer to Appendix A.4 for detailed formula).

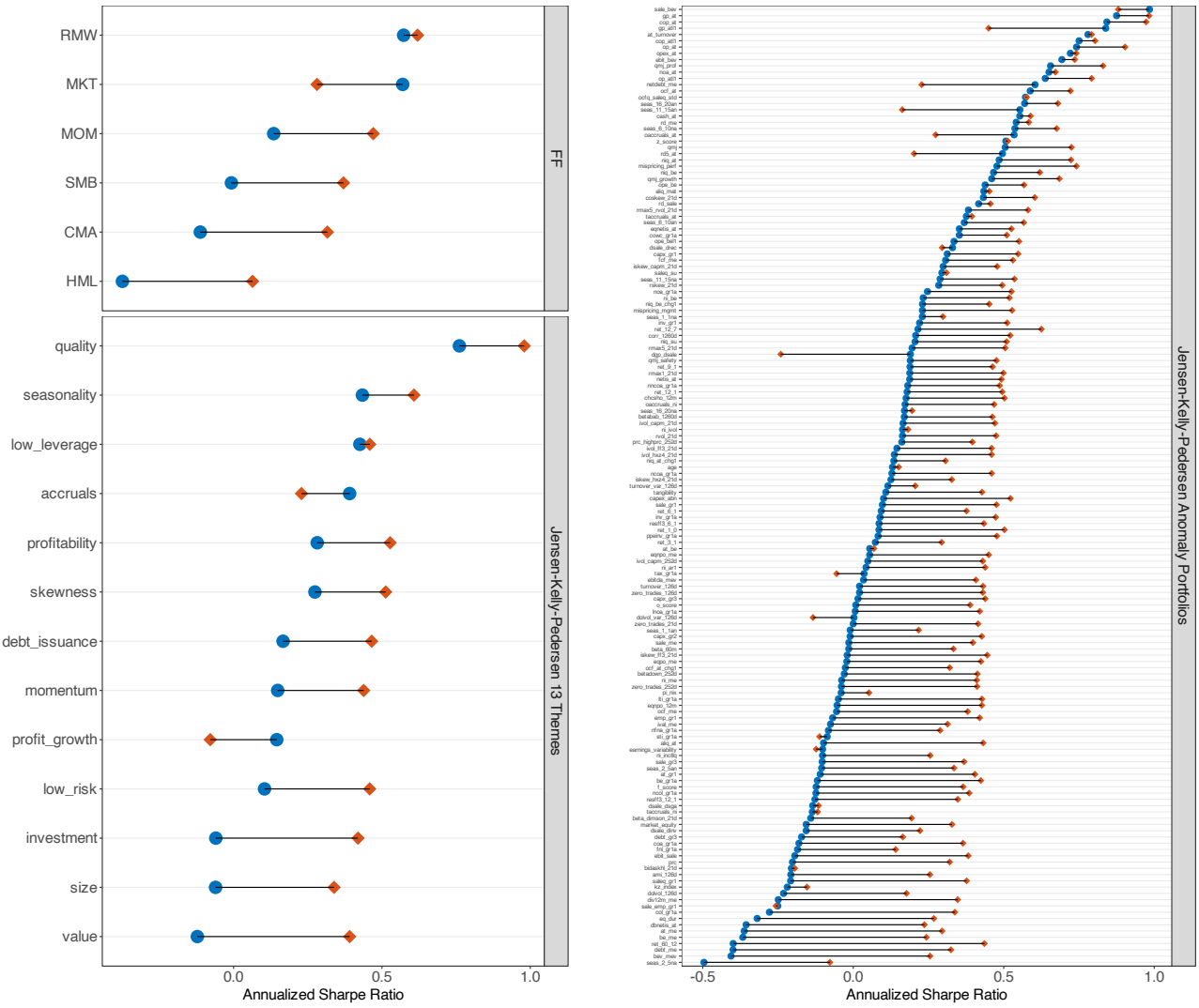
Figure 6 demonstrates the improvement in the Sharpe ratio. The diagram is divided into three panels, showcasing the Fama-French-Carhart six factors, the Jensen-Kelly-Pedersen 13 themes, and individual anomaly portfolios respectively. It is evident that the Sharpe ratios of the combined portfolios (represented by red diamonds) exceed those of the original anomaly portfolios (represented by blue dots) across the spectrum. Out of the 154 portfolios, 140, or 91%, exhibit a positive increase in the Sharpe ratio.[26] The average change in the Sharpe ratio is 0.26, and the median change is 0.30. This empirical evidence shows that our optimal profit-from-flow strategy enhances fundamental investing.

Because our MPIR strategy capitalizes on the subsequent reversal of price impacts, it is natural to compare its performance with other reversal strategies. We consider two popular reversal strategies: the short-term reversal from Jegadeesh (1990) and the long-term reversal from De Bondt and Thaler (1985). Both strategies determine portfolio allocation solely based on past returns. Appendix Figure A.2 replicates the out-of-sample exercise in Figure 6, this time replacing the MPIR strategy with the aforementioned reversal strategies. The result shows mean and median Sharpe ratio improvements of -0.02 and 0.00 for the short-term reversal, and -0.11 and -0.14 for the long-term reversal, respectively. This suggests that our MPIR strategy not only offers a different approach but also significantly augments the efficiency of well-established reversal strategies.

From a practical standpoint, the MPIR strategy is less susceptible to issues related to turnover and transaction costs. First, our strategy forms staggered portfolios based on the

[26]The decrease in the Sharpe ratio for a small number of anomalies can be attributed to the weight $w_{j,t}$ becoming notably large during periods when the MPIR return $\tilde{r}^*_{t+1}$ is negative. This situation arises because $w_{j,t}$ is divided by the Sharpe ratio of anomaly $j$ up to month $t$, in order to ensure proper risk weighting as described in equation (A.33) in the appendix. Consequently, if the historical Sharpe ratio is near zero, the weight $w_{j,t}$ can increase substantially.

**Figure 6. MPIR strategy increases out-of-sample Sharpe ratio of anomaly portfolios**




| | number of anomalies | mean | std | P25 | median | P75 |
|---|---|---|---|---|---|---|
| Sharpe ratio change | 154 | 0.26 | 0.23 | 0.11 | 0.30 | 0.41 |

Note: The blue dots in the figure represent the out-of-sample annualized Sharpe ratio of the Jensen, Kelly, and Pedersen (2021) 154 portfolios, including the Fama-French-Carhart six factors and a large list of other firm characteristics-based anomaly portfolios. These anomaly portfolios are also organized into 13 thematic categories. The red diamonds represent the Sharpe ratio of the portfolio that optimally combines the original anomaly portfolio with our MPIR strategy. The table at the bottom presents the summary statatics of the changes in the Sharpe ratio between the original anomaly portfolios and the combined portfolios. The expanding windows span from January 2000 to December 2004, and the out-of-sample testing period extends from January 2005 to September 2020.

previous six months' flow, so only one-sixth of the positions need adjustment each month. Second, by its very design, our strategy provides liquidity to flows. Generally speaking, liquidity provision strategies are more likely to earn rather than pay bid-ask spreads.

### 5.5.3 Robustness

We show that the MPIR strategy is robust to alternative specifications.

Our baseline strategy trades against the average factor flows from the past six months. However, there may be concerns about the accessibility of monthly flow information at the start of the subsequent month. Moreover, the decision to look back six months, as opposed to a different time period, might be questioned. We now demonstrate that our findings remain robust to these concerns.

In the top two panels of Table 3, we report the Sharpe ratio of the MPIR strategy and the average Sharpe ratio change for the 154 anomaly portfolios for different skip and lookback months. For a large set of alternative parameters, the investment results align closely with our baseline specification, which corresponds to the (0,6) position in the table. Importantly, even if we skip two months (i.e., using the flow information up to January at the end of March), our strategy still performs reasonably well.

In the bottom two panels of Table 3, we replicate the analysis from the top two panels, however without applying normalization based on the in-sample reversion ratio. The investment outcomes are similar, though slightly diminished, indicating that the in-sample normalization is indeed useful.

## 6 Conclusion

In conclusion, our paper proposes that noise trading flows impact cross-sectional asset prices through systematic risk factors. In our model, asset-level flows, when aggregated at the factor level, drive fluctuations in factor prices and risk premia. In the cross-section, asset prices

**Table 3. Robustness to skip, lookback period, and normalization**

**Panel A: MPIR strategy Sharpe ratio**

| skip month \ lookback month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.38 | 0.54 | 0.55 | 0.49 | 0.50 | 0.47 | 0.40 | 0.34 | 0.32 | 0.19 | 0.10 | 0.07 |
| 1 | | 0.28 | 0.34 | 0.34 | 0.37 | 0.33 | 0.22 | 0.06 | 0.03 | 0.02 | -0.01 | -0.10 |
| 2 | | | 0.28 | 0.25 | 0.30 | 0.25 | 0.11 | -0.06 | -0.07 | -0.33 | -0.27 | -0.13 |

**Panel B: average Sharpe ratio change for anomalies**

| skip month \ lookback month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.21 | 0.32 | 0.32 | 0.27 | 0.27 | 0.26 | 0.21 | 0.18 | 0.17 | 0.08 | 0.03 | 0.02 |
| 1 | | 0.12 | 0.16 | 0.16 | 0.18 | 0.16 | 0.09 | -0.01 | -0.01 | -0.01 | -0.02 | -0.03 |
| 2 | | | 0.13 | 0.11 | 0.14 | 0.11 | 0.02 | -0.08 | -0.06 | -0.19 | -0.08 | -0.03 |

**Panel C: MPIR strategy Sharpe ratio (without in-sample normalization)**

| skip month \ lookback month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.42 | 0.37 | 0.47 | 0.39 | 0.38 | 0.36 | 0.25 | 0.16 | 0.16 | 0.10 | 0.11 | 0.10 |
| 1 | | 0.10 | 0.28 | 0.24 | 0.24 | 0.22 | 0.11 | 0.02 | 0.03 | -0.03 | -0.02 | -0.01 |
| 2 | | | 0.35 | 0.26 | 0.27 | 0.23 | 0.09 | -0.02 | -0.00 | -0.07 | -0.05 | -0.04 |

**Panel D: average Sharpe ratio change for anomalies (without in-sample normalization)**

| skip month \ lookback month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.23 | 0.17 | 0.24 | 0.19 | 0.18 | 0.17 | 0.09 | 0.04 | 0.05 | 0.00 | 0.00 | 0.01 |
| 1 | | -0.06 | 0.10 | 0.07 | 0.08 | 0.07 | -0.01 | -0.07 | -0.06 | -0.10 | -0.09 | -0.08 |
| 2 | | | 0.16 | 0.10 | 0.11 | 0.08 | -0.01 | -0.08 | -0.07 | -0.11 | -0.10 | -0.09 |

Note: In the top two panels, we report the Sharpe ratio of the MPIR strategy and the average Sharpe ratio change for the 154 anomaly portfolios for different skip and lookback months. Our base specification, which trades against the flows from the past six months, corresponds to the (0,6) position. The top two panels apply normalization by the in-sample reversion ratio, consistent with our base specification, while the bottom two panels do not. The out-of-sample testing period spans from January 2005 to September 2020.

change with factor-level price impacts in accordance with the factor pricing model. Empirically, our model explains the price impacts and cross impacts between underlying assets with a few risk factors. Moreover, the model-implied trading strategy, designed to optimally exploit the reversion in price impacts, delivers strong and robust investment outcomes.

## References


Alekseev, Georgij, Stefano Giglio, Quinn Maingi, Julia Selgrad, and Johannes Stroebel, 2022, A quantity-based approach to constructing climate risk hedge portfolios, Working paper, NYU.

Alvarez, Fernando, and Andrew Atkeson, 2018, The risk of becoming risk averse: A model of asset pricing and trade volumes, Working paper, University of Chicago.

An, Yu, 2023, Flow-based arbitrage pricing theory, Working paper, Johns Hopkins University.

An, Yu, and Zeyu Zheng, 2023, A dynamic equilibrium factor model of price impacts, Working paper, Johns Hopkins University.

Balasubramaniam, Vimal, John Y Campbell, Tarun Ramadorai, and Benjamin Ranish, 2021, Who owns what? A factor model for direct stockholding, *Journal of Finance* Forthcoming.

Barber, Brad M, Xing Huang, Terrance Odean, and Christopher Schwarz, 2022, Attention-induced trading and returns: Evidence from Robinhood users, *Journal of Finance* 77, 3141–3190.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022a, Ratings-driven demand and systematic price fluctuations, *Review of Financial Studies* 35, 2790–2838.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022b, What do mutual fund investors really care about? *Review of Financial Studies* 35, 1723–1774.

Bogousslavsky, Vincent, and Dmitriy Muravyev, 2021, Who trades at the close? implications for price discovery and liquidity, Working paper, Boston College.

Boulatov, Alex, Terrence Hendershott, and Dmitry Livdan, 2013, Informed trading and portfolio returns, *Review of Economic Studies* 80, 35–72.

Campbell, John Y, and Albert S Kyle, 1993, Smart money, noise trading and stock price behaviour, *Review of Economic Studies* 60, 1–34.

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.

Chaudhary, Manav, Zhiyu Fu, and Jian Li, 2023, Corporate bond multipliers: Substitutes matter, Working paper, Columbia Business School.

Christoffersen, Susan Kerr, Erfan Danesh, and David K Musto, 2015, Why do institutions delay reporting their shareholdings? evidence from form 13F, Working paper, University of Toronto.

Cochrane, John H, 2009, *Asset pricing: Revised edition* (Princeton university press).

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

Daniel, Kent D, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* 56, 921–965.

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.

Dou, Winston, Leonid Kogan, and Wei Wu, 2022, Common fund flows: Flow hedging and factor pricing, *Journal of Finance* Forthcoming.

Du, Wenxin, Alexander Tepper, and Adrien Verdelhan, 2018, Deviations from covered interest rate parity, *Journal of Finance* 73, 915–957.

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

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.

Froot, Kenneth A, and Tarun Ramadorai, 2008, Institutional portfolio flows and international investments, *Review of Financial Studies* 21, 937–971.

Gabaix, Xavier, and Ralph SJ Koijen, 2022, In search of the origins of financial fluctuations: The inelastic markets hypothesis, Working paper, Harvard University.

Gibbons, Michael R, Stephen A Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 1121–1152.

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129, 1947–1990.

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

Harvey, Campbell R, Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Hasbrouck, Joel, and Duane J Seppi, 2001, Common factors in prices, order flows, and liquidity, *Journal of Financial Economics* 59, 383–411.

Huang, Shiyang, Yang Song, and Hong Xiang, 2021, Noise trading and asset pricing factors, Working paper, The University of Hong Kong.

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

Jensen, Theis Ingerslev, Bryan T Kelly, and Lasse Heje Pedersen, 2021, Is there a replication crisis in finance? *Journal of Finance* Forthcoming.

Kelly, Bryan T, Semyon Malamud, and Lasse H Pedersen, 2021, Principal portfolios, Working paper, Yale University.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Kim, Minsoo, 2020, Fund flows, liquidity, and asset prices, Working paper, University of Melbourne.

Koijen, Ralph SJ, and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Li, Jiacui, 2022, What drives the size and value factors? *Review of Asset Pricing Studies* 12, 845–885.

Li, Jiacui, and Zihan Lin, 2022, Prices are less elastic at more aggregate levels, Working paper, University of Utah.

Lo, Andrew W, and Jiang Wang, 2000, Trading volume: definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257–300.

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

Lou, Dong, Christopher Polk, and Spyros Skouras, 2019, A tug of war: Overnight versus intraday expected returns, *Journal of Financial Economics* 134, 192–213.

Lou, Dong, Christopher Polk, and Spyros Skouras, 2022, The day destroys the night, night extends the days, Working paper, LSE.

Moreira, Alan, and Tyler Muir, 2017, Volatility-managed portfolios, *Journal of Finance* 72, 1611–1644.

Nagel, Stefan, 2012, Evaporating liquidity, *Review of Financial Studies* 25, 2005–2039.

Pasquariello, Paolo, and Clara Vega, 2015, Strategic cross-trading in the US stock market, *Review of Finance* 19, 229–282.

Ross, Stephen A, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–60.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

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

Warther, Vincent A., 1995, Aggregate mutual fund flows and security returns, *Journal of Financial Economics* 39, 209–235.

# Appendix

The appendices provide additional theoretical and empirical details.

## A Technical Details

In this section, we provide technical details omitted in the main text.

### A.1 Details for Rotation

The goal is to find a $K \times K$ matrix $\mathbf{O}$, such that the rotated factors defined using

$$(\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_K) := (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K)\mathbf{O}, \tag{A.1}$$

$$(\tilde{q}_{1,t}, \tilde{q}_{2,t}, \ldots, \tilde{q}_{K,t})^\top := \mathbf{O}^{-1}(q_{1,t}, q_{2,t}, \ldots, q_{K,t})^\top. \tag{A.2}$$

have both uncorrelated fundamental returns and flows, i.e., $\text{cov}(\tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t, \tilde{\mathbf{b}}_j^\top \boldsymbol{\xi}_t) = 0$ and $\text{cov}(\tilde{q}_{k,t}, \tilde{q}_{j,t}) = 0$ for any $k \neq j$. The calculations for the rotation remain the same whether we use the payoff $\mathbf{X}$ and flow measured in share amounts, or the fundamental return $\boldsymbol{\xi}_t$ and flow measured in dollar amounts. Therefore, for simplicity, we opt for the latter setup as in for the data generating process (11).

We write portfolio weights of the $K$ factors as an $N \times K$ matrix $\mathbf{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K)$. We write the $K$ factors' flows as a $K \times 1$ vector $\mathbf{q}_t = (q_{1,t}, q_{2,t}, \ldots, q_{K,t})^\top$. In matrix form, the conditions for uncorrelated fundamental returns and flows become

$$\mathbf{O}^\top \mathbf{B}^\top \text{var}(\boldsymbol{\xi}_t)\mathbf{B}\mathbf{O} = \mathbf{I}_K, \tag{A.3}$$

$$\mathbf{O}\boldsymbol{\Pi}\mathbf{O}^\top = \text{var}(\mathbf{q}_t), \tag{A.4}$$

where $\boldsymbol{\Pi} = \text{diag}(\pi_1, \pi_2, \ldots, \pi_K)$ is some $K \times K$ diagonal matrix.

To obtain $\mathbf{O}$, we first carry out Cholesky decomposition and obtain

$$\mathbf{B}^\top \mathrm{var}(\boldsymbol{\xi}_t)\mathbf{B} = \mathbf{U}^\top \mathbf{U}, \tag{A.5}$$

where $\mathbf{U}$ is an $K \times K$ upper triangular matrix. Second, we carry out eigenvalue decomposition

$$\left(\mathbf{U}\mathrm{var}(\mathbf{q}_t)\mathbf{U}^\top\right)\mathbf{G} = \mathbf{G}\boldsymbol{\Pi}, \tag{A.6}$$

where $\boldsymbol{\Pi} = \mathrm{diag}(\pi_1, \pi_2, \ldots, \pi_K)$, and $\mathbf{G}$ is an orthogonal $K \times K$ matrix satisfying $\mathbf{G}^{-1} = \mathbf{G}^\top$. We claim that $\mathbf{O} = \mathbf{U}^{-1}\mathbf{G}$ satisfies the conditions (A.3) and (A.4).

*Proof.* First, we see that

$$\mathbf{O}^\top \mathbf{B}^\top \mathrm{var}(\boldsymbol{\xi}_t)\mathbf{B}\mathbf{O} = \mathbf{G}^\top (\mathbf{U}^\top)^{-1}\mathbf{U}^\top \mathbf{U}\mathbf{U}^{-1}\mathbf{G} = \mathbf{I}_K. \tag{A.7}$$

Second, we have by (A.6),

$$\mathbf{U}\mathrm{var}(\mathbf{q}_t)\mathbf{U}^\top \mathbf{U}\mathbf{O} = \mathbf{U}\mathbf{O}\boldsymbol{\Pi}. \tag{A.8}$$

Eliminating the invertible matrix $\mathbf{U}$ on both sides, we obtain

$$\mathrm{var}(\mathbf{q}_t)\mathbf{U}^\top \mathbf{U}\mathbf{O} = \mathbf{O}\boldsymbol{\Pi}. \tag{A.9}$$

Note that

$$\mathbf{O}^\top \mathbf{U}^\top \mathbf{U}\mathbf{O} = \mathbf{G}^\top \mathbf{G} = \mathbf{I}_K. \tag{A.10}$$

Therefore, we have

$$\mathrm{var}(\mathbf{q}_t)(\mathbf{O}^\top)^{-1} = \mathbf{O}\boldsymbol{\Pi}, \tag{A.11}$$

which proves (A.4). □

## A.2 Price Impact Ratio

In this section, we present the theory of mean-variance optimization for price impacts, using a newly developed concept known as the price impact ratio. We show that the optimal strategy for capitalizing on flow improves fundamental investing.

We follow the notations in Section 3. Recall that under flow $\mathbf{f}$, the asset returns from time 0 to 1 are denoted as

$$\mathbf{R}(\mathbf{f}) = \left(\frac{X_1}{P_1(\mathbf{f})}, \frac{X_2}{P_2(\mathbf{f})}, \ldots, \frac{X_N}{P_N(\mathbf{f})}\right)^\top. \tag{A.12}$$

In particular, the fundamental returns $\mathbf{R}(\mathbf{0})$ are asset returns when there are no flows.

For any portfolio $\mathbf{c} = (c_1, c_2, \ldots, c_N)^\top$, where $c_n$ is the dollar amount invested in asset $n$ when asset prices are $P_n(\mathbf{0})$, we define the *price impact ratio* of portfolio $\mathbf{c}$ in the economy with flow $\mathbf{f}$ as

$$\mathrm{PIR}(\mathbf{c}, \mathbf{f}) := \frac{\mathbf{c}^\top \Delta\mathbf{p}(\mathbf{f})}{\sigma(\mathbf{c}^\top \mathbf{R}(\mathbf{0}))}, \tag{A.13}$$

where $\mathbf{c}^\top \Delta\mathbf{p}(\mathbf{f})$ is the portfolio's price impact and $\sigma(\mathbf{c}^\top \mathbf{R}(\mathbf{0}))$ is the fundamental-return risk. In this definition, the denominator uses $\mathbf{R}(\mathbf{0})$, the assets' fundamental returns when there are no flows, instead of $\mathbf{R}(\mathbf{f})$. This approach ensures that the measurement of portfolio risk is not contaminated by flow-induced changes in time-0 asset prices. The following proposition shows that the price impact ratio is flow-induced changes in the Sharpe ratios.

**PROPOSITION A.1.** *We have*

$$R_F \mathrm{PIR}(\mathbf{c}, \mathbf{f}) = \mathrm{SR}(\mathbf{c}, \mathbf{0}) - \mathrm{SR}(\mathbf{c}, \mathbf{f}), \tag{A.14}$$

*where the Sharpe ratio of portfolio* $\mathbf{c}$ *in the economy with flow* $\mathbf{f}$ *is defined as*

$$\mathrm{SR}(\mathbf{c}, \mathbf{f}) := \frac{\mathbb{E}[(\mathbf{W}(\mathbf{f})\mathbf{c})^\top (\mathbf{R}(\mathbf{f}) - R_F \mathbf{1})]}{\sigma((\mathbf{W}(\mathbf{f})\mathbf{c})^\top \mathbf{R}(\mathbf{f}))}, \tag{A.15}$$

*with the $N \times N$ diagonal matrix* $\mathbf{W}(\mathbf{f}) := \mathrm{diag}(P_1(\mathbf{f})/P_1(\mathbf{0}), P_2(\mathbf{f})/P_2(\mathbf{0}), \ldots, P_N(\mathbf{f})/P_N(\mathbf{0}))$. *Recall that the portfolio weights* $\mathbf{c} = (c_1, c_2, \ldots, c_N)^\top$ *are in the unit of dollar amounts invested in asset* $n$ *when asset prices are* $P_n(\mathbf{0})$. *When asset prices change from* $P_n(\mathbf{0})$ *to* $P_n(\mathbf{f})$ *with flow* $\mathbf{f}$*, the dollar amounts need to change from* $\mathbf{c}$ *to* $\mathbf{W}(\mathbf{f})\mathbf{c}$ *for the same portfolio.*

*Proof.* We note that $\mathbf{W}(\mathbf{f})\mathbf{R}(\mathbf{f}) = \mathbf{R}(\mathbf{0})$. Therefore, we have

$$\begin{aligned}\mathrm{SR}(\mathbf{c}, \mathbf{0}) - \mathrm{SR}(\mathbf{c}, \mathbf{f}) &= \frac{\mathbb{E}[\mathbf{c}^\top(\mathbf{R}(\mathbf{0}) - R_F\mathbf{1})]}{\sigma(\mathbf{c}(\mathbf{f})^\top\mathbf{R}(\mathbf{0}))} - \frac{\mathbb{E}[\mathbf{c}^\top(\mathbf{R}(\mathbf{0}) - R_F\mathbf{W}(\mathbf{f})\mathbf{1})]}{\sigma(\mathbf{c}(\mathbf{f})^\top\mathbf{R}(\mathbf{0}))} \\ &= R_F\frac{\mathbb{E}[\mathbf{c}^\top(\mathbf{W}(\mathbf{f})\mathbf{1} - \mathbf{1})]}{\sigma(\mathbf{c}(\mathbf{f})^\top\mathbf{R}(\mathbf{0}))} = R_F\frac{\mathbf{c}^\top\Delta\mathbf{p}(\mathbf{f})}{\sigma(\mathbf{c}^\top\mathbf{R}(\mathbf{0}))} = R_F\mathrm{PIR}(\mathbf{c}, \mathbf{f}). \qquad \text{(A.16)}\end{aligned}$$

□

Next, we consider the mean-variance optimal portfolio of price impacts and show that this portfolio improves fundamental investing. We define the maximum-price-impact-ratio (MPIR) portfolio under flow $\mathbf{f}$ as

$$\tilde{\mathbf{c}}^*(\mathbf{f}) := \mathrm{var}(\mathbf{R}(\mathbf{0}))^{-1}\Delta\mathbf{p}(\mathbf{f}) \in \arg\max_{\mathbf{c}\in\mathbb{R}^N} \mathrm{PIR}(\mathbf{c}, \mathbf{f}). \qquad \text{(A.17)}$$

The maximum-Sharpe-ratio portfolio under flow $\mathbf{f}$ is defined in the standard way as

$$\mathbf{c}^*(\mathbf{f}) := \mathbf{W}(\mathbf{f})^{-1}\mathrm{var}(\mathbf{R}(\mathbf{f}))^{-1}\mathbb{E}[\mathbf{R}(\mathbf{f}) - R_F\mathbf{1}] \in \arg\max_{\mathbf{c}\in\mathbb{R}^N} \mathrm{SR}(\mathbf{c}, \mathbf{f}). \qquad \text{(A.18)}$$

The MPIR portfolio is not simply the portfolio with the largest price impacts but also depends on risks. Intuitively, we want to find two assets that either have opposite price impacts but positively correlated risks (so we can long one and short the other) or have similar price impacts but weakly or negatively correlated risks (so we can diversify). Trading against the MPIR portfolio offers the best price-impact-versus-risk trade-off for a liquidity provider. Theorem A.1 makes precise this intuition (see Appendix B.1 for a proof).

**THEOREM A.1** (**Two-portfolio separation**). *We have*

$$\underbrace{\mathbf{c}^*(\mathbf{f})}_{\substack{\text{max. Sharpe ratio portfolio with flow}}} = \overbrace{\mathbf{c}^*(\mathbf{0})}^{\substack{\text{max. Sharpe ratio portfolio without flow}}} - R_F \underbrace{\tilde{\mathbf{c}}^*(\mathbf{f})}_{\substack{\text{max. price impact ratio portfolio}}}. \tag{A.19}$$

*The return volatility of portfolio* $\mathbf{c}^*(\mathbf{0})$ *equals the maximum Sharpe ratio without flow. The return volatility of portfolio* $\tilde{\mathbf{c}}^*(\mathbf{f})$ *equals the maximum price impact ratio.*

Equation (A.19) shows that the maximum-Sharpe-ratio portfolio $\mathbf{c}^*(\mathbf{f})$ under flow $\mathbf{f}$ can be separated into two. The first portfolio $\mathbf{c}^*(\mathbf{0})$ maximizes the Sharpe ratio in the same economy but without flow or, equivalently, maximizes the Sharpe ratio driven by the fundamental returns $\mathbf{R}(\mathbf{0})$. In other words, the fundamental-investing portfolio $\mathbf{c}^*(\mathbf{0})$ completely ignores the flow information. The second portfolio $\tilde{\mathbf{c}}^*(\mathbf{f})$ maximizes the price impact ratio under flow $\mathbf{f}$. Shorting the MPIR portfolio is the most efficient strategy to provide liquidity to flows.

Longing the fundamental-investing portfolio $\mathbf{c}^*(\mathbf{0})$ and shorting the MPIR portfolio $\tilde{\mathbf{c}}^*(\mathbf{f})$ maximize the overall Sharpe ratio. Because of diversification benefits, the two portfolios' risk exposures are roughly proportional to their respective fundamental-based Sharpe ratio and flow-based price impact ratio. That is, in periods with greater noisy flows, marginal investors optimally allocate greater risk exposure to liquidity provision; in periods with smaller flows, marginal investors allocate greater risk exposure to fundamental investing.

## A.3 Flow-Based Gibbons-Ross-Shanken Test

Under our model (11), the cross-sectional variations in price impacts rely solely on factor flows, with idiosyncratic flows exerting no influence. We show that testing this restriction gauges the efficiency of our factor model in approximating the cross-section of price impacts.

We run the second-stage regression (11), supplementing it with extra terms for idiosyncratic flows,

$$r_{n,t} = \sum_{m=1}^{N} a_{n,m} e_{m,t} + \sum_{k=1}^{K} \lambda_k \tilde{q}_{k,t} \mathrm{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}. \tag{A.20}$$

The idiosyncratic flow $e_{m,t}$ of asset $m$ at time $t$ is the residual from the first-stage regression (12). The coefficient $a_{n,m}$ measures the *residual price impact* on asset $n$ created by the idiosyncratic flow into asset $m$. Our factor model implies the null hypothesis

$$H_0 : a_{n,m} = 0 \text{ for all } n \text{ and } m. \tag{A.21}$$

That is, the idiosyncratic flow into any asset $m$ should not generate price impacts for any asset $n$, including itself. Let $\hat{\mathbf{a}}$ be the $N^2 \times 1$ vector of parameter estimates $\hat{a}_{n,m}$. As period $T$ tends to infinity, the asymptotic $\chi^2$ test statistic for the null hypothesis is

$$T\hat{\mathbf{a}}^\top \frac{\text{var}(\hat{\mathbf{a}})^{-1}}{T}\hat{\mathbf{a}} \sim \chi^2_{N^2}. \tag{A.22}$$

To understand the economics of the $\chi^2$ test statistic, we study the cross-section of price impacts. Under our factor model (11), the price impact of asset $n$ at time $t$ is $\delta_{n,t} := \sum_{k=1}^K \hat{\lambda}_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$, where we use $\hat{\lambda}_k$ to denote the estimates of $\lambda_k$. As defined in (A.13), the price impact ratio of any portfolio $\mathbf{c} \in \mathbb{R}^N$ at time $t$ is $\mathbf{c}^\top \boldsymbol{\delta}_t / \sigma(\mathbf{c}^\top \boldsymbol{\xi}_t)$, where the cross-section of price impact is denoted as $\boldsymbol{\delta}_t = (\delta_{1,t}, \delta_{2,t}, \ldots, \delta_{N,t})^\top$, and the $N$ assets' fundamental returns at time $t$ is $\boldsymbol{\xi}_t$. The numerator is the price impact of portfolio $\mathbf{c}$ at time $t$, and the denominator is the fundamental risk. The maximum price impact ratio (MPIR) across all portfolios is

$$\max_{\mathbf{c} \in \mathbb{R}^N} \frac{\mathbf{c}^\top \boldsymbol{\delta}_t}{\sigma(\mathbf{c}^\top \boldsymbol{\xi}_t)} = \sqrt{\boldsymbol{\delta}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \boldsymbol{\delta}_t}. \tag{A.23}$$

We define the time-series average of the model-implied squared MPIR as

$$\theta^2 := \frac{1}{T} \sum_{t=1}^T \boldsymbol{\delta}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \boldsymbol{\delta}_t. \tag{A.24}$$

Similarly, the price impact under the model (A.20) with residual price impacts[27] is $\check{\delta}_{n,t} := \sum_{m=1}^N \hat{a}_{n,m} e_{m,t} + \sum_{k=1}^K \hat{\lambda}_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)$, with $\check{\boldsymbol{\delta}}_t = (\check{\delta}_{1,t}, \check{\delta}_{2,t}, \ldots, \check{\delta}_{N,t})^\top$. We define the

[27] Appendix B.2 proves that the estimated $\hat{\lambda}_k$ remain unchanged in regressions (11) and (A.20).

time-series average of the realized squared MPIR as

$$\check{\theta}^2 := \frac{1}{T}\sum_{t=1}^{T} \check{\boldsymbol{\delta}}_t^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\check{\boldsymbol{\delta}}_t. \tag{A.25}$$

The following theorem connects the MPIR to the residual impacts of idiosyncratic flows (see Appendix B.3 for a proof).

**THEOREM A.2.** *Assuming that $e_{n,t}$, $\tilde{b}_{n,k}$, and $\tilde{q}_{k,t}$ are observed and the fundamental return $\boldsymbol{\xi}_t$ is i.i.d. over time, we have, almost surely as $T$ tends to infinity,*

$$\hat{\mathbf{a}}^\top \frac{\mathrm{var}(\hat{\mathbf{a}})^{-1}}{T}\hat{\mathbf{a}} = \check{\theta}^2 - \theta^2. \tag{A.26}$$

Equation (A.26) shows that as period $T$ tends to infinity, the $\chi^2$ test statistic for residual price impacts in (A.22) equals the difference between the realized and model-implied average squared MPIR. This result generalizes Gibbons, Ross, and Shanken (1989), who show that the $\chi^2$ test statistic for anomaly expected returns equals the difference between the realized and model-implied squared maximum Sharpe ratio. Simply put, Gibbons, Ross, and Shanken (1989) theoretically connect the idiosyncratic risks to the maximum Sharpe ratio, and we connect the idiosyncratic flows to the maximum price impact ratio. We make the assumptions on observed regressors and i.i.d. fundamental returns to avoid econometric complications. We leave the generalization of the Shanken (1992) correction for future research.

## A.4 Details for the MPIR strategy

In this appendix, we provide details on constructing the MPIR strategy and how to combine it with fundamental investing.

First, using equation (18), we define the staggered MPIR strategy in month $s$ as

$$\tilde{\mathbf{c}}_s^{\{t\}} := \sum_{k\in\{\mathrm{MKT,SMB,HML}\}} -\lambda_k^{\{t\}}\bar{q}_{k,s}^{\{t\}}\tilde{\mathbf{b}}_k^{\{t\}}. \tag{A.27}$$

The superscript $\{t\}$ indicates that $\tilde{\mathbf{c}}_s^{\{t\}}$ is estimated using the training window up to month $t$. The term $\lambda_k^{\{t\}}$ represents the estimated price of flow-induced risk of factor $k$, $\bar{q}_{k,s}^{\{t\}}$ represents the average flow into factor $k$ over past six months $s-5, s-4, \ldots, s$, and $\tilde{\mathbf{b}}_k^{\{t\}}$ represents the portfolio weights of factor $k$.

Second, we compute the model-implied price impact ratio and actual price reversion of the staggered MPIR portfolio $\tilde{\mathbf{c}}_s^{\{t\}}$. By equation (A.13), the price impact ratio in month $s$ is

$$\kappa_s^{\{t\}} := \sqrt{\sum_{k\in\{\text{MKT,SMB,HML}\}} \left(\lambda_k^{\{t\}} \bar{q}_{k,s}^{\{t\}}\right)^2}. \tag{A.28}$$

Here, we apply the condition $(\tilde{\mathbf{b}}_k^{\{t\}})^\top \text{var}(\boldsymbol{\xi})^{\{t\}} \tilde{\mathbf{b}}_k^{\{t\}} = 1$ and $(\tilde{\mathbf{b}}_k^{\{t\}})^\top \text{var}(\boldsymbol{\xi})^{\{t\}} \tilde{\mathbf{b}}_l^{\{t\}} = 0$ for $k \neq l$. Similarly, the price reversion of the staggered MPIR portfolio $\tilde{\mathbf{c}}_s^{\{t\}}$ in month $s+1$, normalized by its fundamental risk, is

$$\check{\kappa}_s^{\{t\}} := \frac{(\tilde{\mathbf{c}}_s^{\{t\}})^\top \tilde{\mathbf{r}}_{s+1}}{\sqrt{(\tilde{\mathbf{c}}_s^{\{t\}})^\top \text{var}(\boldsymbol{\xi})^{\{t\}} \tilde{\mathbf{c}}_s^{\{t\}}}} = \frac{(\tilde{\mathbf{c}}_s^{\{t\}})^\top \tilde{\mathbf{r}}_{s+1}}{\sqrt{\sum_{k\in\{\text{MKT,SMB,HML}\}} \left(\lambda_k^{\{t\}} \bar{q}_{k,s}^{\{t\}}\right)^2}}, \tag{A.29}$$

where $\tilde{\mathbf{r}}_{s+1}$ is the demeaned return of the 25 test assets in month $s+1$.

Third, we define the ratio of the average reversion to the average model-implied price impact as

$$\text{REV}^{\{t\}} = \frac{\sum_{s=6}^{t-1} \check{\kappa}_s^{\{t\}}}{\sum_{s=6}^{t-1} \kappa_s^{\{t\}}}. \tag{A.30}$$

The summation over month $s$ starts from 6 because $\bar{q}_{k,s}^{\{t\}}$ is the average flow over the past six months and ends at $t-1$ because $\mathbf{r}_t$ is the last observable return using the training window ending in month $t$.

Fourth, as discussed in the main text, we normalize the MPIR strategy by the in-sample estimated reversion ratio,

$$\tilde{\mathbf{c}}_t^* = \min\left(\text{REV}^{\{t\}}, 1\right) \tilde{\mathbf{c}}_t^{\{t\}}, \tag{A.31}$$

and the MPIR strategy's excess return in month $t+1$ is

$$\tilde{r}^*_{t+1} = (\tilde{\mathbf{c}}^*_t)^\top (\mathbf{r}_{t+1} - r_{F,t+1}), \tag{A.32}$$

where $\mathbf{r}_{t+1}$ is the 25 test assets' return in month $t+1$ and $r_{F,t+1}$ is the net risk-free rate.

Finally, we combine the MPIR strategy with fundamental investing using equation (19). The mixing ratio is given by

$$w_{j,t} := \max\left(\frac{\mathrm{VOL}_j^{\{t\}}}{\mathrm{SR}_j^{\{t\}}}, 0\right)(1 + r_{F,t+1}), \tag{A.33}$$

according to Theorem A.1. The term $\mathrm{SR}_j^{\{t\}}$ represents the Sharpe ratio of anomaly portfolio $j$, and $\mathrm{VOL}_j^{\{t\}}$ represents the return standard deviation. Both terms are calculated utilizing the same training windows as those used in the MPIR estimation. We normalize the MPIR strategy by $\mathrm{VOL}_j^{\{t\}}/\mathrm{SR}_j^{\{t\}}$, rather than normalizing portfolio $j$ by $\mathrm{SR}_j^{\{t\}}/\mathrm{VOL}_j^{\{t\}}$. We do so to avoid normalizing portfolio $j$ by the inverse of volatility in the time series, because Moreira and Muir (2017) show that such normalization per se increases the Sharpe ratio. Because the estimated Sharpe ratio $\mathrm{SR}_j^{\{t\}}$ could be negative, we bound the scaling by zero. That is, we leave $r_{j,t+1}$ unchanged if portfolio $j$ has a negative historical Sharpe ratio up to month $t$.

# B Proofs

In this appendix, we provide proofs.

## B.1 Proof of Theorem A.1

We note that $\mathbf{W}(\mathbf{f})\mathbf{R}(\mathbf{f}) = \mathbf{R}(\mathbf{0})$. Therefore, we have

$$\mathrm{var}(\mathbf{R}(\mathbf{f}))^{-1} = \mathbf{W}(\mathbf{f})\mathrm{var}(\mathbf{R}(\mathbf{0}))^{-1}\mathbf{W}(\mathbf{f}). \tag{A.34}$$

Second, we have

$$\mathbf{R}(\mathbf{0}) - R_F\mathbf{1} + R_F\mathbf{1} = \mathbf{W}(\mathbf{f})(\mathbf{R}(\mathbf{f}) - R_F\mathbf{1} + R_F\mathbf{1}), \tag{A.35}$$

which simplifies to

$$\mathbf{R}(\mathbf{f}) - R_F\mathbf{1} = \mathbf{W}(\mathbf{f})^{-1}(\mathbf{R}(\mathbf{0}) - R_F\mathbf{1} - R_F\Delta\mathbf{p}(\mathbf{f})). \tag{A.36}$$

Taking expectations on both sides, we have

$$\mathbb{E}[\mathbf{R}(\mathbf{f})] - R_F\mathbf{1} = \mathbf{W}(\mathbf{f})^{-1}(\mathbb{E}[\mathbf{R}(\mathbf{0})] - R_F\mathbf{1} - R_F\Delta\mathbf{p}(\mathbf{f})). \tag{A.37}$$

Therefore, by equations (A.34) and (A.37), we have

$$\begin{aligned}\mathbf{c}^*(\mathbf{f}) &= \mathbf{W}(\mathbf{f})^{-1}\mathrm{var}(\mathbf{R}(\mathbf{f}))^{-1}(\mathbb{E}[\mathbf{R}(\mathbf{f})] - R_F\mathbf{1}) \\ &= \mathrm{var}(\mathbf{R}(\mathbf{0}))^{-1}(\mathbb{E}[\mathbf{R}(\mathbf{0})] - R_F\mathbf{1} - R_F\Delta\mathbf{p}(\mathbf{f})) = \mathbf{c}^*(\mathbf{0}) - R_F\tilde{\mathbf{c}}^*(\mathbf{f}).\end{aligned} \tag{A.38}$$

The return volatility of portfolio $\mathbf{c}^*(\mathbf{0})$ is

$$\sigma(\mathbf{c}^*(\mathbf{0})^\top\mathbf{R}(\mathbf{0})) = \sqrt{\mathbf{c}^*(\mathbf{0})^\top\mathrm{var}(\mathbf{R}(\mathbf{0}))\mathbf{c}^*(\mathbf{0})} = \sqrt{\mathbb{E}[\mathbf{R}(\mathbf{0}) - R_F\mathbf{1}]^\top\mathrm{var}(\mathbf{R}(\mathbf{0}))^{-1}\mathbb{E}[\mathbf{R}(\mathbf{0}) - R_F\mathbf{1}]}, \tag{A.39}$$

which equals the maximum Sharpe ratio without flow by definition (A.18). Similarly, the return volatility of portfolio $\tilde{\mathbf{c}}^*(\mathbf{f})$ is

$$\sigma(\tilde{\mathbf{c}}^*(\mathbf{f})^\top\mathbf{R}(\mathbf{0})) = \sqrt{\tilde{\mathbf{c}}^*(\mathbf{f})^\top\mathrm{var}(\mathbf{R}(\mathbf{0}))\tilde{\mathbf{c}}^*(\mathbf{f})} = \sqrt{\Delta\mathbf{p}(\mathbf{f})^\top\mathrm{var}(\mathbf{R}(\mathbf{0}))^{-1}\Delta\mathbf{p}(\mathbf{f})}, \tag{A.40}$$

which equals the maximum price impact ratio by definition (A.17).

## B.2 Proof of the Simplifying Regression

We show that the panel regression (A.20),

$$r_{n,t} = \sum_{m=1}^{N} a_{n,m} e_{m,t} + \sum_{k=1}^{K} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}, \tag{A.41}$$

can be separated into two regressions. The first asset-by-asset time-series regression

$$r_{n,t} = \sum_{m=1}^{N} a_{n,m} e_{m,t} + \xi_{n,t}. \tag{A.42}$$

obtains the same $a_{n,m}$ as regression (A.41). The second panel regression

$$r_{n,t} = \sum_{k=1}^{K} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}, \tag{A.43}$$

obtains the same $\lambda_k$ as regression (A.41).

To see this fact, note that because the idiosyncratic flow $e_{m,t}$ is the residual of the first-stage regression (12), we know by construction that $\sum_{t=1}^{T} q_{k,t} e_{m,t} = 0$. Because each $\tilde{q}_{k,t}$ is a linear combination of $q_{1,t}, q_{2,t}, \ldots, q_{K,t}$, we know that $\sum_{t=1}^{T} \tilde{q}_{k,t} e_{m,t} = 0$. We rewrite the panel regression (A.41) in vector form as

$$\mathbf{r} = \sum_{n=1}^{N} \sum_{m=1}^{N} a_{n,m} \mathbf{e}_{n,m} + \sum_{k=1}^{K} \lambda_k \mathbf{y}_k + \boldsymbol{\xi}, \tag{A.44}$$

where the $NT \times 1$ vector $\mathbf{r}$ is

$$\mathbf{r} = (r_{1,1}, r_{1,2}, \ldots, r_{1,T}, r_{2,1}, r_{2,2}, \ldots, r_{2,T}, \ldots, r_{N,1}, r_{N,2}, \ldots, r_{N,T})^\top. \tag{A.45}$$

Each vector $\mathbf{e}_{n,m}$ is an $NT \times 1$ vector with only the $(n-1)T + 1$-th to $nT$-th entry ranging

from $e_{m,1}$ to $e_{m,T}$ and all other entries equaling zero. Each $\mathbf{y}_k$ is an $NT \times 1$ vector

$$\mathbf{y}_k = \left( \underbrace{\tilde{q}_{k,1}\text{cov}(\xi_{1,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t), \ldots, \tilde{q}_{k,T}\text{cov}(\xi_{1,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)}_{T \text{ terms}}, \ldots, \underbrace{\tilde{q}_{k,1}\text{cov}(\xi_{N,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t), \ldots, \tilde{q}_{k,T}\text{cov}(\xi_{N,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t)}_{T \text{ terms}} \right)^\top. \tag{A.46}$$

The vector $\boldsymbol{\xi}$ simply stacks all error terms $\xi_{n,t}$.

Note that, for any $n$, $m$, and $k$, we have

$$\mathbf{e}_{n,m}^\top \mathbf{y}_k = \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) \sum_{t=1}^{T} \tilde{q}_{k,t} e_{m,t} = 0, \tag{A.47}$$

where the last equality uses the first step in the proof. As a result, to estimate coefficient $a_{n,m}$ and $\lambda_k$, it suffices to run separate regressions

$$\mathbf{r} = \sum_{n=1}^{N} \sum_{m=1}^{N} a_{n,m} \mathbf{e}_{n,m} + \boldsymbol{\xi} \text{ and } \mathbf{r} = \sum_{k=1}^{K} \lambda_k \mathbf{y}_k + \boldsymbol{\xi}. \tag{A.48}$$

This first regression further reduces to the asset-by-asset time-series regression (A.42).

## B.3 Proof of Theorem A.2

First, we simplify the $\chi^2$ test statistics in (A.26). We write a $1 \times N$ vector $\mathbf{e}_t = (e_{1,t}, e_{2,t}, \ldots, e_{N,t})$ and an $N \times 1$ vector $\mathbf{a}_n = (a_{n,1}, a_{n,2}, \ldots, a_{n,N})^\top$, and we write regression (A.20) as

$$r_{n,t} = \mathbf{e}_t \mathbf{a}_n + \sum_{k=1}^{K} \lambda_k \tilde{q}_{k,t} \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) + \xi_{n,t}. \tag{A.49}$$

We define the $T \times N$ matrix $\mathbf{e} = (\mathbf{e}_1; \mathbf{e}_2; \ldots; \mathbf{e}_T)$ and the $N \times 1$ vector $\boldsymbol{\xi}_n = (\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,T})^\top$. As shown in Appendix B.2, we can run an asset-by-asset time-series

regression to obtain the point estimator of $\mathbf{a}_n$ as

$$\hat{\mathbf{a}}_n = (\mathbf{e}^\top \mathbf{e})^{-1}\mathbf{e}^\top \begin{pmatrix} r_{n,1} \\ r_{n,2} \\ \cdots \\ r_{n,T} \end{pmatrix} = \mathbf{a}_n + (\mathbf{e}^\top \mathbf{e})^{-1}\mathbf{e}^\top \boldsymbol{\xi}_n, \tag{A.50}$$

where we use the fact that, for any $n = 1, 2, \ldots, N$ and $k = 1, 2, \ldots, K$, $\sum_{t=1}^{T} e_{n,t}\tilde{q}_{k,t} = 0$, because $e_{n,t}$ is the residual from the first-stage regression.

Therefore, we have, for any $m$ and $n$,

$$\text{cov}(\hat{\mathbf{a}}_n, \hat{\mathbf{a}}_m) = (\mathbf{e}^\top \mathbf{e})^{-1}\mathbf{e}^\top \text{cov}(\boldsymbol{\xi}_n, \boldsymbol{\xi}_m)\mathbf{e}(\mathbf{e}^\top \mathbf{e})^{-1} = (\mathbf{e}^\top \mathbf{e})^{-1}\boldsymbol{\Sigma}_\xi(n, m), \tag{A.51}$$

because $\boldsymbol{\xi}_t$ is i.i.d. over time. The term $\boldsymbol{\Sigma}_\xi(n, m)$ is the $(n, m)$-th element of the cross-sectional variance-covariance matrix of $\boldsymbol{\xi}_t$. When constructing the $\chi^2$ test statistic, we use the asymptotically consistent sample covariance matrix $\hat{\boldsymbol{\Sigma}}_\xi$ for the true $\boldsymbol{\Sigma}_\xi$. We denote $\hat{\mathbf{a}} = (\hat{\mathbf{a}}_1; \hat{\mathbf{a}}_2; \ldots; \hat{\mathbf{a}}_N)$ as the $N^2 \times 1$ vector of parameter estimates. By equation (A.51), we have

$$\text{var}(\hat{\mathbf{a}}) = \hat{\boldsymbol{\Sigma}}_\xi \otimes (\mathbf{e}^\top \mathbf{e})^{-1}, \tag{A.52}$$

where $\otimes$ represents the Kronecker product. Therefore, we have

$$\hat{\mathbf{a}}^\top \frac{\text{var}(\hat{\mathbf{a}})^{-1}}{T}\hat{\mathbf{a}} = \hat{\mathbf{a}}^\top \left(\hat{\boldsymbol{\Sigma}}_\xi^{-1} \otimes (\mathbf{e}^\top \mathbf{e}/T)\right)\hat{\mathbf{a}}. \tag{A.53}$$

Under the null hypothesis of $\mathbf{a} = \mathbf{0}$, we have

$$\hat{\mathbf{a}}^\top \left( \hat{\mathbf{\Sigma}}_\xi^{-1} \otimes (\mathbf{e}^\top \mathbf{e}/T) \right) \hat{\mathbf{a}} = \sum_{n=1}^{N} \sum_{m=1}^{N} \left( (\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top \boldsymbol{\xi}_n \right)^\top \hat{\mathbf{\Sigma}}_\xi^{-1}(n,m) (\mathbf{e}^\top \mathbf{e}/T) (\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top \boldsymbol{\xi}_m \quad \text{(A.54)}$$

$$= \frac{1}{T} \sum_{n=1}^{N} \sum_{m=1}^{N} \hat{\mathbf{\Sigma}}_\xi^{-1}(n,m) \boldsymbol{\xi}_n^\top \mathbf{e} (\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top \boldsymbol{\xi}_m$$

$$= \frac{1}{T} \sum_{n=1}^{N} \sum_{m=1}^{N} \hat{\mathbf{\Sigma}}_\xi^{-1}(n,m) \boldsymbol{\psi}_n^\top \boldsymbol{\psi}_m \quad \text{(A.55)}$$

$$= \frac{1}{T} \begin{pmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \\ \cdots \\ \boldsymbol{\psi}_N \end{pmatrix}^\top \left( \hat{\mathbf{\Sigma}}_\xi^{-1} \otimes \mathbf{I}_T \right) \begin{pmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \\ \cdots \\ \boldsymbol{\psi}_N \end{pmatrix}, \quad \text{(A.56)}$$

where we define

$$\boldsymbol{\psi}_n = \mathbf{e} (\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top \boldsymbol{\xi}_n, \quad \text{(A.57)}$$

as the projection of $\boldsymbol{\xi}_n$ onto the idiosyncratic flow space. In step (A.54), we use block matrix multiplication for every $N$ elements and $\hat{\mathbf{\Sigma}}_\xi^{-1}(n,m)$ is the $(n,m)$-th element of $\hat{\mathbf{\Sigma}}_\xi^{-1}$. Step (A.55) is because the projection matrix $\mathbf{e}(\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top$ is idempotent. In step (A.56), we use the block-matrix multiplication in the reverse direction, with each $\boldsymbol{\psi}_n$ as a $T \times 1$ vector.

We define $\boldsymbol{\psi}_t = (\psi_{1,t}, \psi_{2,t}, \ldots, \psi_{N,t})^\top$. In this way, $\boldsymbol{\psi}_n$ is the time-series variation in $\psi_{n,t}$ for a given asset $n$, and $\boldsymbol{\psi}_t$ is the cross-sectional variation in $\psi_{n,t}$ for a given time $t$. By rearranging $\boldsymbol{\psi}_n$ into $\boldsymbol{\psi}_t$, we have

$$\frac{1}{T} \begin{pmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \\ \cdots \\ \boldsymbol{\psi}_N \end{pmatrix}^\top \left( \hat{\mathbf{\Sigma}}_\xi^{-1} \otimes \mathbf{I}_T \right) \begin{pmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \\ \cdots \\ \boldsymbol{\psi}_N \end{pmatrix} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\psi}_t^\top \hat{\mathbf{\Sigma}}_\xi^{-1} \boldsymbol{\psi}_t. \quad \text{(A.58)}$$

Because $\boldsymbol{\xi}_t$ is i.i.d. over time, the strong law of large numbers implies that the sample

covariance matrix $\hat{\mathbf{\Sigma}}_\xi$ converges to the true $\mathbf{\Sigma}_\xi$ almost surely as $T$ tends to infinity. Therefore, in the limit of $T$ tending to infinity, we have almost surely,

$$\hat{\mathbf{a}}^\top \frac{\text{var}(\hat{\mathbf{a}})^{-1}}{T} \hat{\mathbf{a}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\psi}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \boldsymbol{\psi}_t. \tag{A.59}$$

Next, we transform the squared MPIR in (A.26). We have defined $\check{\boldsymbol{\delta}}_t = (\check{\delta}_{1,t}, \check{\delta}_{2,t}, \ldots, \check{\delta}_{N,t})^\top$ as the cross-section of price impacts at time $t$ in the main text. We now define $\check{\boldsymbol{\delta}}_n = (\check{\delta}_{n,1}, \check{\delta}_{n,2}, \ldots, \check{\delta}_{n,T})^\top$. Using equations (A.50) and (A.57), we have

$$\check{\boldsymbol{\delta}}_n = \mathbf{e}(\mathbf{e}^\top \mathbf{e})^{-1} \mathbf{e}^\top \boldsymbol{\xi}_n + \sum_{k=1}^{K} \hat{\lambda}_k \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) \tilde{\mathbf{q}}_k = \boldsymbol{\psi}_n + \sum_{k=1}^{K} \hat{\lambda}_k \text{cov}(\xi_{n,t}, \tilde{\mathbf{b}}_k^\top \boldsymbol{\xi}_t) \tilde{\mathbf{q}}_k, \tag{A.60}$$

where $\tilde{\mathbf{q}}_k = (\tilde{q}_{k,1}, \tilde{q}_{k,2}, \ldots, \tilde{q}_{k,T})^\top$. In this way, $\check{\boldsymbol{\delta}}_n$ is the time-series variation in $\check{\delta}_{n,t}$ for a given asset $n$, and $\check{\boldsymbol{\delta}}_t$ is the cross-sectional variation in $\check{\delta}_{n,t}$ for a given time $t$. Thus, we have

$$\check{\boldsymbol{\delta}}_t = \boldsymbol{\psi}_t + \text{var}(\boldsymbol{\xi}_t) \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k. \tag{A.61}$$

The realized squared MPIR at time $t$ is

$$\begin{aligned}
&\check{\boldsymbol{\delta}}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \check{\boldsymbol{\delta}}_t \\
=&\left( \boldsymbol{\psi}_t^\top + \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k^\top \text{var}(\boldsymbol{\xi}_t) \right) \text{var}(\boldsymbol{\xi}_t)^{-1} \left( \boldsymbol{\psi}_t + \text{var}(\boldsymbol{\xi}_t) \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k \right) \\
=&\boldsymbol{\psi}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \boldsymbol{\psi}_t + 2\boldsymbol{\psi}_t^\top \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k + \left( \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k^\top \right) \text{var}(\boldsymbol{\xi}_t) \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k \\
=&\boldsymbol{\psi}_t^\top \text{var}(\boldsymbol{\xi}_t)^{-1} \boldsymbol{\psi}_t + 2\boldsymbol{\psi}_t^\top \sum_{k=1}^{K} \hat{\lambda}_k \tilde{q}_{k,t} \tilde{\mathbf{b}}_k + \sum_{k=1}^{K} \hat{\lambda}_k^2 \tilde{q}_{k,t}^2,
\end{aligned} \tag{A.62}$$

where, in the last step, we use $\tilde{\mathbf{B}}^\top \mathrm{var}(\boldsymbol{\xi}_t)\tilde{\mathbf{B}} = \mathbf{I}_K$. The time-series average is

$$\begin{aligned}\check{\theta}^2 &= \frac{1}{T}\sum_{t=1}^{T} \check{\boldsymbol{\delta}}_n^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\check{\boldsymbol{\delta}}_n \\ &= \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\psi}_t^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\boldsymbol{\psi}_t + \frac{2}{T}\sum_{t=1}^{T}\boldsymbol{\psi}_t^\top \sum_{k=1}^{K}\hat{\lambda}_k \tilde{q}_{k,t}\tilde{\mathbf{b}}_k + \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\hat{\lambda}_k^2\tilde{q}_{k,t}^2. \end{aligned} \tag{A.63}$$

Note that for any $n = 1, 2, \ldots, N$ and $k = 1, 2, \ldots, K$, we have

$$\begin{aligned}\sum_{t=1}^{T}\tilde{q}_{k,t}\psi_{n,t} &= \sum_{t=1}^{T}\tilde{q}_{k,t}\mathbf{e}_t(\mathbf{e}^\top\mathbf{e})^{-1}\mathbf{e}^\top\boldsymbol{\xi}_n \\ &= \left(\sum_{t=1}^{T}\tilde{q}_{k,t}e_{1,t}, \sum_{t=1}^{T}\tilde{q}_{k,t}e_{2,t}, \ldots, \sum_{t=1}^{T}\tilde{q}_{k,t}e_{N,t}\right)(\mathbf{e}^\top\mathbf{e})^{-1}\mathbf{e}^\top\boldsymbol{\xi}_n = 0. \end{aligned} \tag{A.64}$$

Therefore, we know that

$$\check{\theta}^2 = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\psi}_t^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\boldsymbol{\psi}_t + \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\hat{\lambda}_k^2\tilde{q}_{k,t}^2. \tag{A.65}$$

A similar calculation gives

$$\theta^2 = \frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K}\hat{\lambda}_k^2\tilde{q}_{k,t}^2. \tag{A.66}$$

Therefore, we have

$$\check{\theta}^2 - \theta^2 = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\psi}_t^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\boldsymbol{\psi}_t. \tag{A.67}$$

Using (A.59), we have almost surely in the limit of $T$ tending to infinity,

$$\hat{\mathbf{a}}^\top \frac{\mathrm{var}(\hat{\mathbf{a}})^{-1}}{T}\hat{\mathbf{a}} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{\psi}_t^\top \mathrm{var}(\boldsymbol{\xi}_t)^{-1}\boldsymbol{\psi}_t = \check{\theta}^2 - \theta^2. \tag{A.68}$$

# C Construction and Cleaning of Mutual Fund Flows

We present the details involved in constructing and cleaning mutual fund flows.

Our primary data source is the CRSP Survivorship-Bias-Free Mutual Fund database.

We start with all funds' return and total net assets (TNA) data at the share-class level. A mutual fund may include multiple share classes. We first drop observations with no valid CRSP identifier, `crsp_fundno`. A few fund-share classes report multiple TNA in a given month. These are likely data duplicates, so we keep only the first observation of the month. We end up with 8,591,018 share-class×month observations. In what follows, we discuss the cleaning steps for returns and TNA separately at the share-class level. After cleaning, we aggregate the share-class level data to the fund level.

## C.1 Return Cleaning

We first correct data errors in monthly net returns, `mret`.

First, we address extremely positive returns. We study the case in which a particular fund share has returns greater than 100% and has existed for more than one year.[28] We manually check the entire time series of each share class in this subsample. Most of these extreme returns reflect misplaced decimal points, which confound returns in decimal and percentage formats. For these cases, we divide the faulty returns by 100.

Second, we address extreme negative returns. Similarly, we study the case in which a particular fund share has existed for more than one year and has returns lower than −50%. With extremely negative returns, we need to distinguish data errors from significantly negative returns before a fund's closure. Thus, we manually check only the subsample of negative returns that occur at least one year prior to the last observation of a closed fund. We manually check whether these extreme returns reflect data-input errors for this subsample. For the cases with misplaced decimal points, we divide the faulty returns by 100.

[28] We use the one-year threshold because mutual fund returns and TNAs during the first year are sometimes inaccurate in the CRSP database. For example, returns and TNAs can be stale or reported using a placeholder number such as 0.1. We address these issues by cross-checking with the alternative database.

## C.2 TNA Cleaning

Unlike many prior studies that construct percentage mutual fund flows, we study dollar-amount flows to preserve the cross-sectional relative magnitudes. The mutual fund size distribution features a very long right tail. Winsorizing the extreme dollar-amount TNA likely removes both valid large values and input errors, generating significant bias. We devise an algorithm to identify and correct erroneous observations of TNA:

1. Using the sample with corrected returns, we calculate dollar flows as in (13) at the share-class level.

2. We study the top and bottom 0.5% of all dollar flows.[29] We manually check this subsample's TNA time series of all share classes. We identify several common error types:

   - Misplaced decimal points (usually by hundredths or thousandths).
   - Stale TNA observations from CRSP, typically when a fund reorganizes its share class offering (e.g., adding a new share class and moving assets into a single share class).
   - CRSP sometimes sets TNA $= 0.1$ for the first few months of a new fund or a new share class.

   We correct the misplaced decimal issue immediately. For funds suffering from the latter two problems, we obtain their TNA from Morningstar.[30] Morningstar's TNA data (`Net_Assets_ShareClass_Monthly`) suffer to a lesser extent from these issues than CRSP's TNA data. We conclude by further cross-checking other third-party vendors (e.g., Yahoo Finance and Bloomberg Terminal). Hence, when a fund's CRSP TNA deviates more than 50% from its Morningstar TNA, we use the Morningstar TNA.

[29] The choice of the top and bottom 0.5% is motivated by the distribution of dollar flows, where most extreme values tend to occur at these tails.

[30] We merge the CRSP and Morningstar databases using a fund's ticker.

3. We repeat the previous steps one more time to ensure that we have accounted for most, if not all, extreme errors.

4. We compare our cleaned TNA with total assets (`assets`) from Thomson/Refinitiv Holdings data.[31] Following Coval and Stafford (2007) and Lou (2012), we drop observations whenever our cleaned TNA deviate more than 50% from `assets` from Thomson/Refinitiv.

Using cleaned return and TNA data, we calculate dollar flows at the share-class level using (13). We obtain a fund's flows by adding up the flows of all share classes in the fund. The final sample contains 1,613,579 fund×month observations.

## C.3 Cross-Validating the Data-Cleaning Procedure

We cross-validate our data-cleaning procedure. We compute the aggregate mutual fund flows in dollar amounts each month. We compare our aggregate flow measures with alternative sources, including the Investment Company Institute (ICI) and the Flow of Funds (FoF).

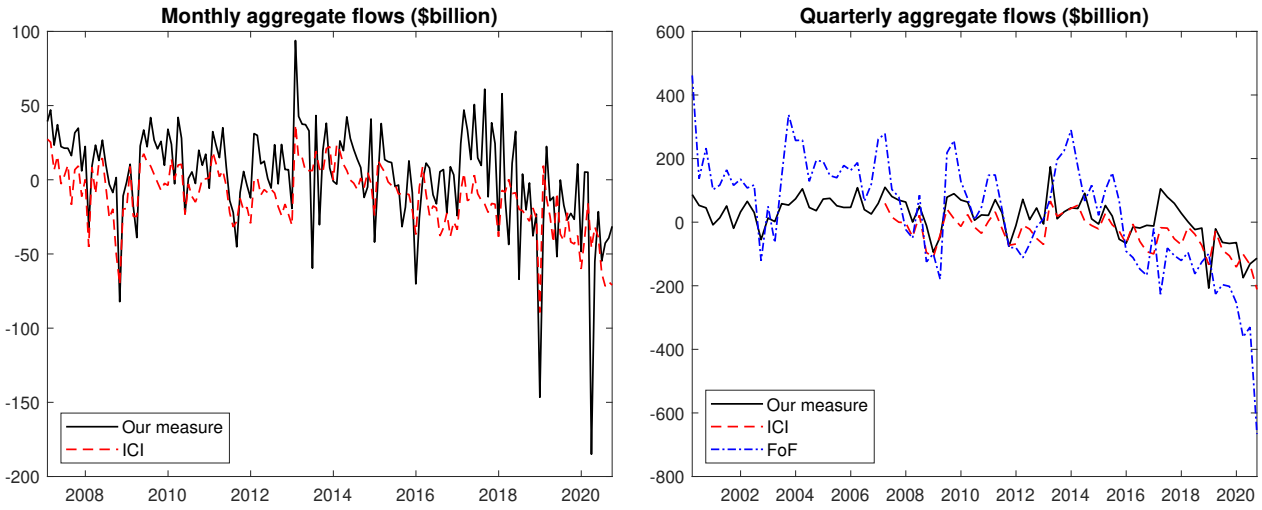
The ICI provides aggregate monthly mutual fund flows. We obtain a version of ICI aggregate flows data from 2007 to 2020. We use the ICI's Total Equity mutual fund flows, which feature a close coverage scope to mutual funds in our sample.

The FoF data (now known as "Financial Accounts of the United States - Z.1") are published quarterly by the Federal Reserve Board. We use mutual fund flow (Line 28) of Corporate Equities (Table 223) and unadjusted flows (FU). We use the December 2021 vintage of the data because the Federal Reserve revises historical FoF data every quarter.

Figure A.1 plots the time series of aggregate mutual fund flows from various sources. The left panel shows the monthly time series of our measure and ICI flow, and the right panel shows the quarterly time series of all three sources. Our measure of aggregate mutual fund flows is broadly consistent with the other two sources. The correlation between our

[31] We merge the two databases via the linking table MFLINKS, which WRDS provides.

**Figure A.1. Time series of aggregate mutual fund flows from various sources**




Note: The left panel plots the monthly time series of our measure of aggregate mutual fund flows and Investment Company Institute (ICI) flows. The right panel plots the quarterly time series of our measure, ICI flows, and Flow of Funds (FoF) flows.

aggregate flow measure and ICI flow is 0.68 at the monthly level and 0.80 at the quarterly level. The correlation between our measure and FoF flow is 0.55 at the quarterly level.

The differences in Figure A.1 between the three measures of aggregate flows likely reflect differences in mutual fund coverage. Specifically, the ICI flow tracks virtually all U.S. equity mutual funds that invest in both domestic and world equity markets.[32] The FoF flow, sourced from unpublished ICI data, aggregates unadjusted flows into and out of all U.S. mutual funds (including variable annuity long-term mutual funds). The FoF flow is calculated based on mutual fund assets in common stock, preferred stock, and rights and warrants.[33] In comparison, our mutual fund sample contains U.S. mutual funds that CRSP covers. CRSP collects historical data from various sources.[34] Due to the nature of the data collection process, CRSP's coverage is smaller than ICI's coverage.

[32]The ICI is a trade association for the mutual fund industry, and virtually all U.S. mutual funds are ICI members (Warther, 1995).

[33]See https://www.federalreserve.gov/apps/fof/SeriesAnalyzer.aspx?s=FA653064100&t=F.223&suf=Q.

[34]The sources include the Fund Scope Monthly Investment Company Magazine, the Investment Dealers Digest Mutual Fund Guide, Investor's Mutual Fund Guide, the United and Babson Mutual Fund Selector, and the Wiesenberger Investment Companies Annual Volumes.

**Table A.1. Model rotation**

**original factors**

| | **flow** | | | | **fundamental return** | | | |
|---|---|---|---|---|---|---|---|---|
| | correlation | | | std | correlation | | | std |
| | MKT | SMB | HML | | MKT | SMB | HML | |
| MKT | 1 | 0.11 | 0.25 | $5.34 \times 10^{-2}$ | 1 | 0.41 | 0.91 | 0.14 |
| SMB | 0.11 | 1 | -0.11 | $0.83 \times 10^{-2}$ | 0.41 | 1 | 0.56 | 0.12 |
| HML | 0.25 | -0.11 | 1 | $0.87 \times 10^{-2}$ | 0.91 | 0.56 | 1 | 0.28 |

**rotated factors**

| | **flow** | | | | **fundamental return** | | | |
|---|---|---|---|---|---|---|---|---|
| | correlation | | | std | correlation | | | std |
| | MKT | SMB | HML | | MKT | SMB | HML | |
| MKT | 1 | 0 | 0 | $0.85 \times 10^{-2}$ | 1 | 0 | 0 | 1 |
| SMB | 0 | 1 | 0 | $0.10 \times 10^{-2}$ | 0 | 1 | 0 | 1 |
| HML | 0 | 0 | 1 | $0.07 \times 10^{-2}$ | 0 | 0 | 1 | 1 |

| | Low | 2 | 3 | 4 | High |
|---|---|---|---|---|---|
| Small | 0.03 | 0.03 | 0.04 | 0.05 | 0.05 |
| 2 | 0.08 | 0.09 | 0.08 | 0.08 | 0.06 |
| 3 | 0.14 | 0.14 | 0.11 | 0.10 | 0.06 |
| 4 | 0.30 | 0.22 | 0.15 | 0.13 | 0.12 |
| Big | 1.80 | 0.97 | 0.71 | 0.60 | 0.44 |

**rotated $\tilde{b}_{MKT}$**

| | Low | 2 | 3 | 4 | High |
|---|---|---|---|---|---|
| Small | 0.30 | 0.28 | 0.36 | 0.38 | 0.37 |
| 2 | 0.65 | 0.66 | 0.61 | 0.56 | 0.37 |
| 3 | 0.89 | 0.85 | 0.68 | 0.58 | 0.39 |
| 4 | 0.70 | 0.59 | 0.31 | 0.29 | 0.27 |
| Big | -5.24 | -4.29 | -2.23 | -1.17 | -0.20 |

**rotated $\tilde{b}_{SMB}$**

| | Low | 2 | 3 | 4 | High |
|---|---|---|---|---|---|
| Small | -0.14 | -0.06 | -0.06 | 0.02 | 0.05 |
| 2 | -0.39 | -0.18 | -0.10 | -0.03 | -0.03 |
| 3 | -0.44 | -0.02 | 0.15 | 0.21 | 0.19 |
| 4 | -0.65 | 0.51 | 0.63 | 0.55 | 0.59 |
| Big | -6.97 | -0.82 | 1.27 | 3.36 | 1.96 |

**rotated $\tilde{b}_{HML}$**

Note: The top panel reports the correlations and standard derivations of flows and fundamental returns of the original MKT, SMB, and HML factors. The bottom panel reports the rotated MKT, SMB, and HML factors. The unit of flow is expressed as a percentage of the total stock market capitalization. The three figures at the end illustrate the portfolio weights of these rotated factors.

# D Additional Empirical Results

In this appendix, we provide additional empirical results.

Table A.1 illustrates the process of rotating MKT, SMB, and HML factors from their original counterparts. As depicted in the top-left panel, the MKT flow's volatility is approximately six times that of the SMB and HML flows. Although the pairwise correlation between these factors' flow is small, it is not zero. In the top-right panel, we present the correlations and standard deviations of the factors' fundamental returns $\mathbf{b}_k^\top \boldsymbol{\xi}_t$. These returns exhibit

**Table A.2. IV first-stage regression**

| | MKT flow | SMB flow | HML flow |
|---|---|---|---|
| concurrent night return | $89.03\times10^{-4}$ | $-5.72\times10^{-4}$ | $-1.68\times10^{-4}$ |
| | (2.35) | (-1.29) | (-0.49) |
| lag-1 flow − lag-6 flow | 0.10 | 0.15 | 0.17 |
| | (1.70) | (3.74) | (3.27) |
| constant | $-5.01\times10^{-4}$ | $-0.19\times10^{-4}$ | $0.25\times10^{-4}$ |
| | (-0.92) | (-0.29) | (0.59) |
| regression $R^2$ | 4.90% | 4.81% | 3.85% |
| F-statistics | 6.19 | 6.06 | 4.81 |

Note: We report the IV first-stage regression results, with each factor's flow being regressed against its concurrent night return and the difference between one-month and six-month lagged flows. The figures in parentheses represent the t-statistics, computed from heteroskedasticity-robust standard errors.

high correlation, particularly between the MKT and HML factors. This high correlation arises from the discrepancy between the estimated SMB and HML portfolio weights $\mathbf{b}_k$ and the original Fama-French portfolio weights, which is discussed in detail in Section 5.2.

The bottom panel displays the rotated MKT, SMB, and HML factors, which exhibit uncorrelated flows and fundamental returns. We also provide the portfolio weights of these rotated factors in the three figures at the end. Upon examination of these portfolio weights, we note that the rotated factors maintain the same interpretation as the original factors.

Table A.2 presents the IV first-stage regression results for the third column of Table 2. In this analysis, each factor's flow is regressed on its corresponding concurrent night return and the difference between one-month and six-month lagged flows. The flow into the MKT factor is found to significantly chase concurrent night returns. On the other hand, the flows into SMB and HML factors exhibit a more pronounced response to the differences in lagged flows. The F-statistics for these regressions hover around six for these factors.

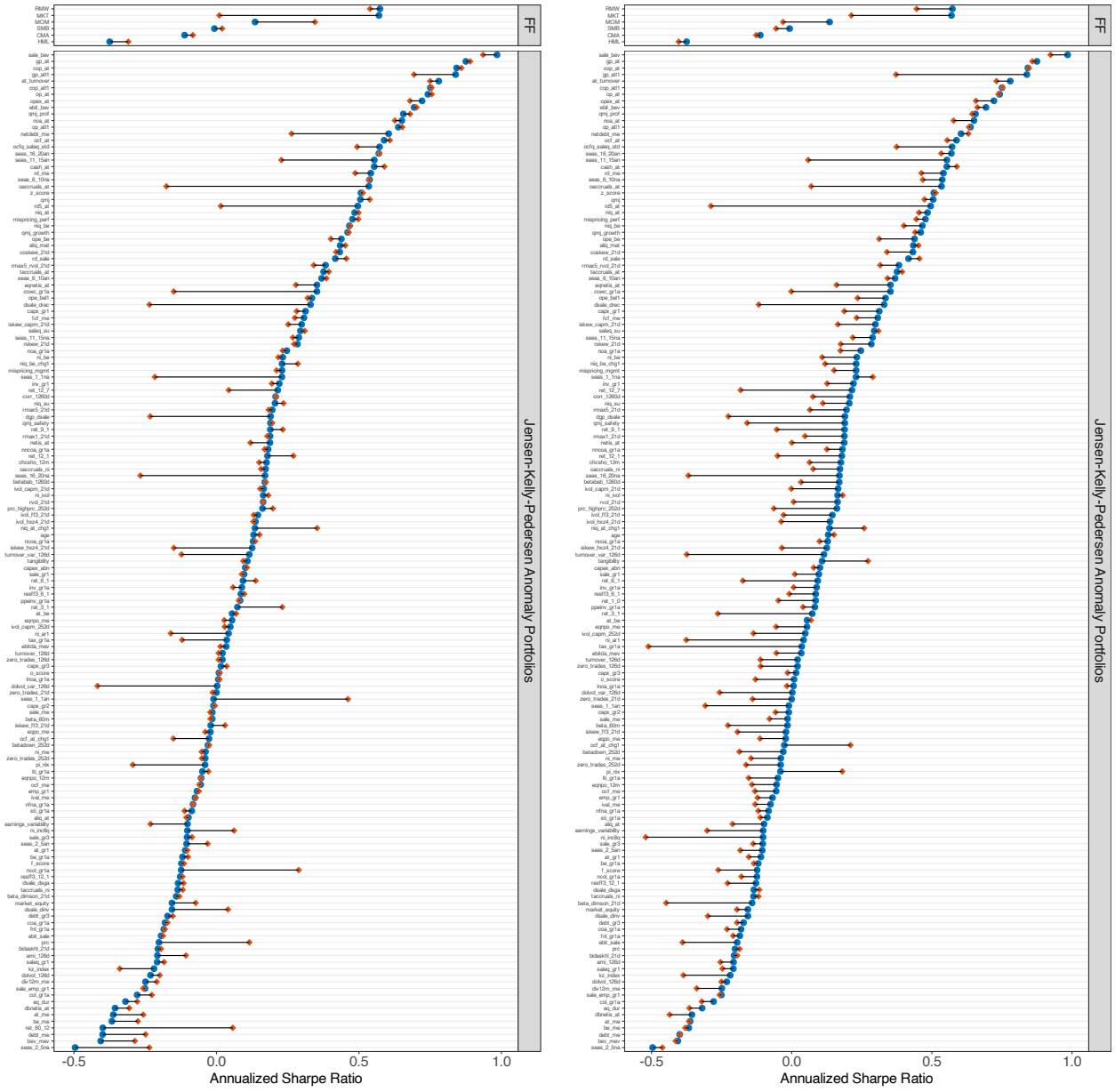
In Figure A.2, we reproduce the out-of-sample exercise presented in Figure 6, substituting the MPIR strategy with two alternate reversal strategies: the short-term reversal from Jegadeesh (1990) and the long-term reversal from De Bondt and Thaler (1985). In each month, we scale each reversal strategy so that the combined portfolio allocates risk between

the original anomaly and the reversal strategy proportional to their respective Sharpe ratios. The formula is

$$r^*_{j,t+1} = r_{j,t+1} + w_{j,t} r_{\text{rev},t+1} \text{ with } w_{j,t} := \max\left(\frac{\text{VOL}_j^{\{t\}}}{\text{SR}_j^{\{t\}}}, 0\right) \max\left(\frac{\text{SR}_{\text{rev}}^{\{t\}}}{\text{VOL}_{\text{rev}}^{\{t\}}}, 0\right). \tag{A.69}$$

Here, $\text{SR}_j^{\{t\}}$ and $\text{SR}_{\text{rev}}^{\{t\}}$ represent the Sharpe ratios of the anomaly portfolio $j$ and reversal strategy, while $\text{VOL}_j^{\{t\}}$ and $\text{VOL}_{\text{rev}}^{\{t\}}$ represent their respective return standard deviations. The analysis reveals mean and median Sharpe ratio improvements of -0.02 and 0.00 for the short-term reversal strategy and -0.11 and -0.14 for the long-term reversal strategy, respectively. Thus, it is evident that our MPIR strategy does more than merely replicate well-established reversal strategies in the literature.

**Figure A.2. Alternative reversal strategies and the Sharpe ratios of anomaly portfolios**




**(A) Short-term reversal** **(B) Long-term reversal**

Note: The blue dots in the figure represent the out-of-sample annualized Sharpe ratio of the Jensen, Kelly, and Pedersen (2021) 154 portfolios, including the Fama-French-Carhart six factors and a large list of other firm characteristics-based anomaly portfolios. The red diamonds represent the Sharpe ratio of the portfolio that combines the original portfolio with two alternative reversal strategies: short-term reversal from Jegadeesh (1990) (panel A) and long-term reversal from De Bondt and Thaler (1985) (panel B). The expanding windows span from January 2000 to December 2004, and the out-of-sample testing period extends from January 2005 to September 2020.