Stress Testing Spillover Risk in Mutual Funds [*]

Agostino Capponi[†]   Paul Glasserman[‡]   Marko Weber[§]

June 23, 2023

Abstract

We develop a framework to quantify the vulnerability of mutual funds to fire-sale spillover losses. We account for the first-mover incentive that results from the mismatch between the liquidity offered to redeeming investors and the liquidity of assets held by the funds. In our framework, the negative feedback loop between investors' redemptions and price impact from asset sales leads to an aggregate change in funds' NAV, which is determined as a fixed point of a nonlinear mapping. We show that a higher concentration of first movers increases the aggregate vulnerability of the system, as measured by the ratio between endogenous losses due to fund redemptions and exogenous losses due to initial price shocks only. When calibrated to U.S. mutual funds, our model shows that, in stressed market scenarios, spillover losses are significantly amplified through a nonlinear response to initial shocks that results from the first-mover incentive. Higher spillover losses provide a stronger incentive to redeem early, further increasing fire-sale losses and the transmission of shocks through overlapping portfolio holdings.

Key words: mutual funds, liquidity mismatch, fire-sale externalities, first-mover incentive, systemic risk.

JEL Classification: G01, G23, G28

1   Introduction

The mutual fund industry has experienced strong growth in the past decade and holds an increasingly large portion of financial assets. As such, the possibility of a threat to financial stability from the mutual fund sector has become a prominent concern for regulators. In particular, the liquidity transformation provided by open-end funds has been identified as a potential source of vulnerability: investors may redeem their fund shares at the end-of-day net asset value (NAV), even if the fund holds illiquid assets that can only be liquidated over multiple days and at distressed prices. Referring to funds that hold less liquid assets, former Bank of England Governor Mark Carney

famously stated in June 2019 that "these funds are built on a lie, which is that you can have daily liquidity for assets that fundamentally aren't liquid."

Mutual funds have been implicated in the "taper tantrum" of 2013 and in the disruption of bond markets early in the Covid-19 period. The liquidity mismatch between shareholder claims and fund holdings has resulted in the collapse of individual funds (prominent examples include the Third Avenue Focused Credit fund and the Woodford Equity Income Fund), leading to concerns for the broader impact on financial stability. The mutual fund structure creates a first-mover advantage among investors, because investors who withdraw early are shielded from the adverse impact of asset liquidation. This first-mover advantage can produce a run on a fund that amplifies fire-sale losses to other investors.

The objective of our study is to build a framework to quantify ex ante the vulnerability of mutual funds to fire sales, accounting for the first-mover incentive created by the liquidity mismatch. Our model reflects the fact that investors' redemptions are paid at an NAV that has not yet accounted for the cost of subsequent asset liquidations incurred to meet redemption requests. Furthermore, building on Capponi et al. (2020), we posit that some investors redeem fund shares in anticipation of the impact that their (and other investors') redemptions have on *future* fund performance, instead of responding only to realized shocks. We refer to these investors as first movers, and their inclusion is the key feature that distinguishes our analysis from prior work on the financial stability implications of the mutual fund structure. Funds that hold illiquid assets are more sensitive to the impact of fire sales, and their investors have a stronger incentive to exit the fund early. Early redemptions in turn increase the cost of remaining invested in the fund, and prompt additional redemptions. This creates a downward spiral of investor withdrawals, price impact, and investment losses that can substantially amplify an initial price shock.

We apply the framework to quantify the vulnerability of mutual funds in the United States to spillover losses. We take institutional investors as a proxy for first movers — the investors that exploit the liquidity mismatch. This premise is in keeping with the Security and Exchange Commission's (SEC) regulatory treatment of retail and institutional money market funds (MMFs).[1]

We measure the aggregate vulnerability of mutual funds using the Spillover Loss Ratio (SLR), defined as the ratio between spillover losses and the initial losses due to an exogenous shock. We show that the first-mover incentive creates a nonlinear dependence of spillover losses on exogenous asset shocks, and this nonlinear relation has a compounding effect on losses. In more detail, we construct a *systemicness matrix* to quantify the relation between an exogenous shock and the drop in value of fund shares due to ensuing redemptions. If the spectral radius of this matrix is well

---

[1]As stated in the SEC Release No. IC-34441, "institutional investors frequently scrutinize liquidity levels in money market funds [...] facilitating rapid redemptions when a fund's liquidity begins to decline." Since 2014, institutional prime and municipal MMFs "are required to use a floating NAV because their investors have historically made the heaviest redemptions in times of market stress and are more likely to act on the incentive to redeem if a fund's stable price per share is higher than its market-based value". The SEC proposed rule "Money Market Fund Reforms", released in December 2021, suggests that these institutional funds may also be required to adopt swing pricing, a provision aimed at mitigating the first-mover advantage, because institutional investors are more likely to exploit this advantage.

below unity, then the first-mover incentive is immaterial; as the spectral radius approaches one, the first-mover incentive becomes stronger, and spillover losses become increasingly large compared to a system with no first movers.

The nonlinearity stemming from the first-mover advantage has implications for financial stability. First, a concentration of first movers in fewer funds increases the system's vulnerability. As a consequence, fund liquidity management measures that unintentionally alter the distribution of first movers across funds, e.g., by prompting them to migrate and concentrate into fewer funds, might increase the fragility of the system. For example, patchy adoption of swing pricing (a tool to remove the first-mover incentive) may inadvertently reduce the system's ability to withstand shocks, instead of strengthening it. Second, because spillover losses do not scale linearly with model inputs, small changes in asset liquidity or investor base can substantially alter the vulnerability of the financial system. This implies that historical evidence on mutual fund resilience may severely underestimate or fail to predict future fragility. The same asset shock may cause spillover losses of different magnitudes in different market environments. Third, the nonlinearity reinforces fire-sale contagion across mutual funds and asset classes. Forced liquidations can spread losses across funds and assets through overlapping portfolios. As the prospects of widespread contagion increase, so does the incentive to redeem early.

Our work provides a new framework to design macroprudential stress tests and measure vulnerability. Prior studies have analyzed the mechanism that renders mutual funds vulnerable to runs (Allen et al. (2009) and Gennaioli et al. (2013)), and provided empirical evidence for this fragility (e.g., Chen et al. (2010), Goldstein et al. (2017), Jiang et al. (2022)). The empirical study of Johnson (2004) shows that short-term fund shareholders pay for less liquidity than they demand, and thus impose liquidity costs on the long-term shareholders because of the liquidity mismatch. Our work differs from most prior studies because its focus is on measuring the impact of the first-mover incentive created by the mutual fund structure.

Our paper is related to models of fire sales caused by propagation of shocks across balance sheets of constrained banks (see Greenwood et al. (2015), Duarte and Eisenbach (2021), and Capponi and Larsson (2015)). In these models, banks liquidate part of their holdings in response to an exogenous shock to satisfy leverage requirements. The spillover losses due to deviation of market prices from fundamentals are a measure of the banking sector's vulnerability to fire sales.

The studies of Fricke and Fricke (2021) and Cetorelli et al. (2016) adapt the banking fire-sales model of Greenwood et al. (2015) to mutual funds. They conclude that vulnerability to spillover losses is significantly lower for mutual funds. These studies recognize that poor fund performance leads to forced sales and depressed prices, but they do not account for the amplifying effect of funds' liquidity mismatch — the mismatch between the liquidity promised to the funds' investors and the liquidity of the funds' assets. This liquidity mismatch can create greater fire-sale losses through mutual fund ownership than would be incurred if investors held the funds' assets directly. From the perspective of financial stability, it is the key feature that differentiates mutual fund investing from direct ownership of the fund's assets. Hence, macroprudential frameworks that

3

do not incorporate the first-mover advantage, such as those discussed above, may underestimate mutual fund vulnerability.

Choi et al. (2020) study the impact of fire sales caused by fund flows in the corporate bond market. They conclude that the impact of fire sales is low because corporate bond funds maintain significant liquidity buffers to manage redemptions. The bond liquidity measure of Chernenko and Sunderam (2020) is also based on the observation that cash buffers can counterbalance low market liquidity. Cash buffers can help mitigate costly liquidations, but funds still sell non-negligible amounts of illiquid assets — for every 1% of outflows, corporate bond holdings decrease by 0.84%[2] — and cash buffers are eventually depleted. As explained above and emphasized later, we cannot extrapolate from the historically low impact of fire sales triggered by fund flows because the first-mover advantage is highly nonlinear in periods of market stress and low liquidity.

Ma et al. (2022) find that selling pressure from mutual funds was a major determinant of the increase in Treasury yields early in the Covid-19 period. Fixed-income funds holding illiquid assets experienced larger outflows and were forced to sell some of their liquid assets, which depressed prices of Treasury securities.

Schmidt et al. (2016) compare flow patterns in money market mutual funds around the collapse of Lehman Brothers in September 2008. They provide evidence that large and more sophisticated institutional investors had a stronger reaction to negative shocks than retail investors. Their study lends support to our identification of first movers as institutional investors. It also supports our conclusion that a higher proportion of first movers reinforces strategic complementarities in mutual fund redemptions.

Capponi et al. (2020) characterize the optimal swing pricing charge to first movers, in a model that accounts for the feedback loop between asset illiquidity, mutual fund performance, and redemption flow. We build on that framework here, extending it to multiple assets and multiple funds, studying the resulting fixed-point problem, and analyzing the effect of the distribution of first-movers across funds. An empirical study by Jin et al. (2022) analyzes how swing pricing can mitigate mutual fund vulnerability. Our study complements theirs by showing that an uneven adoption of swing pricing may lead to a higher concentration of first-mover investors in fewer funds, and thus weaken financial stability.

The rest of the paper is organized as follows. In Section 2, we use municipal bond fund data to motivate our use of institutional investors as a proxy for first movers. In Section 3, we develop our framework and specify the measure of mutual fund vulnerability. In Section 4, we apply the model to a dataset of mutual fund portfolio holdings. We conclude in Section 5. Additional results are relegated to the Appendix.

---

[2]Li et al. (2020) conduct a similar investigation to Choi et al. (2020), but on municipal bonds. They conclude that fire sales due to fund outflows have a significant impact on prices. During the Covid-19 period, bonds held by municipal funds fell more than bonds held primarily by retail investors. Yield spreads between the two types of bonds persisted even after market conditions reverted to normal, suggesting the presence of a fire-sale premium for bonds held by mutual funds.
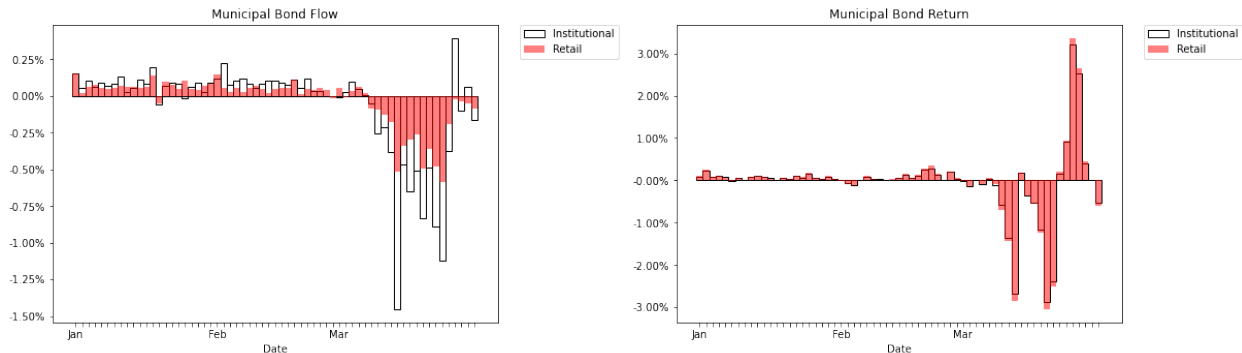
Figure 2.1: Aggregate daily flows (left panel) and average daily return (right panel) for institutional and retail fund share classes in U.S. open-end municipal bond funds during Q1 2020. We source data from the Morningstar database. Municipal bond funds posted positive returns after the Fed announced that the Money Market Mutual Fund Liquidity Facility would accept certain U.S. municipal bonds as eligible collateral on March 20, 2020.

2 Evidence from Municipal Bond Funds

As motivation for our framework, we use municipal bond funds data to test the hypothesis that institutional investors react faster to a drop in bond prices than retail investors. Municipal bonds are less liquid than many other assets held by funds, and therefore the effect of funds' liquidity transformation is stronger. This evidence lends support to our choice of using institutional investors as a proxy for first-mover investors.

As the Covid-19 shock hit financial markets in March 2020, municipal bond funds experienced a spike in outflows. While average returns for institutional and retail fund share classes were virtually indistinguishable, institutional investors were significantly more likely to run for the exit (see Figure 2.1). Institutional investors are arguably more active in monitoring market conditions and have the technical skills to anticipate how selling pressure exacerbates the impact of a market shock on prices. As a result, we expect them to be more likely to withdraw funds early.

We run the following panel regression using daily data on U.S. open-end municipal bond funds from the Morningstar database for Q1 2020, a period that covers the Covid-19 market shock:

$$Flow_{i,t} = \alpha + \beta_1 Return_{i,t-1} + \beta_2 Return_{i,t-1} \times I_{\{Return_{i,t-1}<0\}}$$
$$+ \beta_3 I_{\{Return_{i,t-1}<0\}} + \gamma Controls_{i,t} + \varepsilon_{i,t}.$$

Here, $Flow_{i,t}$ is the flow for fund share class i on day t, as reported in the Morningstar database. The $Controls_{i,t}$ variables are the lagged flow (the flow over day $t-1$) and Log(TNA) (the logarithm of total net assets held by the fund on day t). Summary statistics for the fund share classes in our sample are reported in Table A.1. This specification is adapted from the one in Goldstein et al. (2017), with the main difference that we use returns rather than excess returns over a sector benchmark. We use returns because our focus is on the systemic implications of mutual fund flows, and in particular on the impact of outflows for an entire sector of funds, rather than just of

outflows from underperforming funds. Returns are computed after accounting for paid fees. Flows and returns are measured in percent.

We run the regression separately for institutional and retail share classes. The regression results in Table 2.1 show that negative returns are associated with outflows that are nearly 50% larger for institutional investors (compare $0.106 = -0.011 + 0.117$ for institutional share classes with $0.073 = -0.013 + 0.086$ for retail fund share classes). These findings support our hypothesis that institutional investors are more reactive to negative market shocks. In Appendix A, we show that our results are robust to the observation frequency, i.e., they remain qualitatively the same if we use quarterly data from the CRSP database.[3] Our results are also consistent with Schmidt et al. (2016), who show that outflows from institutional investors are stronger than those from retail investors. Their study focuses on money market funds during 2008.

Table 2.1: Results from regressing flows on previous day returns for municipal bond funds. We use daily data from Morningstar for institutional fund share classes (left column) and for retail fund share classes (right column) during Q1 2020. The dependent variable is the proportional flow of fund share classes on day t. The return variable is the daily return on day $t-1$. Lagged Flow is the flow on day $t-1$. Log(TNA) is the natural logarithm of total net assets on day $t-1$. Flows are winsorized at the 1st and 99th percentiles.

	Institutional	Retail
Constant	0.058***	0.019***
	(0.01)	(0.00)
Return	-0.011**	-0.013***
	(0.01)	(0.00)
Return$\times I_{\{Return<0\}}$	0.117***	0.086***
	(0.01)	(0.01)
$I_{\{Return<0\}}$	-0.075***	-0.052***
	(0.01)	(0.01)
Lagged Flow	0.216***	0.166***
	(0.01)	(0.00)
Log(TNA)	-0.001***	-0.001
	(0.00)	(0.00)
N	26,260	49,259
Adj. R^2	0.098	0.057

$^{***}p < 0.01,\ ^{**}p < 0.05,\ ^{*}p < 0.1$

3 Framework

We begin with the design of a reference model that does not account for the first-mover incentive. The sequence of events is as follows: (1) Asset prices are subject to an exogenous initial shock; (2) investors redeem shares in response to funds' (negative) returns; (3) funds liquidate assets to repay redeeming investors; (4) forced sales drive down market prices; (5) further fund redemptions

---

[3]Unlike Morningstar, the CRSP database does not include daily mutual fund flow data.

and asset sales are triggered, i.e., steps (2)–(4) are repeated. This reference framework is related to that proposed by Cetorelli et al. (2016) and Fricke and Fricke (2021) for mutual funds based on the banking model of Greenwood et al. (2015). Our reference model differs primarily in accounting for multiple rounds of share redemptions and asset sales.

We then extend the reference model to the full model, which accounts for the liquidity mismatch in the mutual fund structure. The full model differs from the reference model in two crucial aspects: some investors are fast and redeem before the fund liquidates assets, and thus get repaid at an NAV that does not yet account for liquidation costs; and those investors respond not only to realized returns but also to anticipated liquidation costs that will result from further redemptions by other investors.

We assume that a fund liquidates assets in proportion to its holdings. This is the most commonly adopted liquidation strategy in the fire sales literature (e.g., Greenwood et al. (2015), Duarte and Eisenbach (2021)), and the one implicitly assumed by the SEC in its Proposed Rule "Money Market Fund Reforms".[4] Jiang et al. (2021) also find that funds tend to liquidate proportionally in stressed scenarios, in order to prevent the liquidity level of their portfolio from deteriorating excessively. Furthermore, funds often have mandates that restrict them from deviating widely from a target mix of assets.

We use lowercase letters to denote quantities for individual funds or assets, and uppercase letters to denote vectors or matrices that summarize quantities for multiple funds or assets. The system consists of N mutual funds, indexed by $i \in \{1, \ldots, N\}$, and K assets, indexed by $k \in \{1, \ldots, K\}$. We use a_i to denote the dollar value of fund i's asset holdings, and A to denote the $N \times N$ diagonal matrix with entries $A_{ii} = a_i$. The weight of asset k in fund i's portfolio, m_{ik}, is the ratio of the dollar value of fund i's holdings in asset k and a_i, and M is the $N \times K$ matrix of portfolio weights. The asset holdings of each fund i are divided into q_i^0 identical portfolio units. One portfolio unit comprises a pro rata amount of each security, i.e., a portfolio unit of fund i consists of m_{ik} shares of each asset k. We normalize the initial price of a share of each asset to \$1. Hence, by construction, the initial value p_i^0 of a unit of fund i's portfolio is equal to \$1. For each fund i, there are n_i^0 outstanding shares. The initial value of a share of fund i, s_i^0, is also normalized to \$1. Therefore, $a_i = n_i^0 = q_i^0$.

3.1 Reference Model without First Movers

We outline the sequence of events and actions in the reference model of mutual funds where no first mover is present. Throughout the paper, we use $\top$ to denote the transpose of a matrix.

1. Exogenous shock and investors' redemptions. The assets are hit by negative shocks $\Delta F^0 := (\Delta f_1^0, \ldots, \Delta f_K^0)^\top$. The magnitude of shock Δf_k^0 is smaller than the price of asset k, so asset prices

---

[4]The swing factor in the Proposed Rule is to be computed as the cost of liquidating a pro rata amount of each security in the fund's portfolio.

remain positive. The value of a portfolio unit of fund i decreases by

$$\Delta p_i^0 = \sum_{k=1}^{K} m_{ik} \Delta f_k^0. \tag{1}$$

Therefore, the change in value of each fund's portfolio is given by the vector $\Delta P^0 = M \Delta F^0$. The change in value of a share of fund i is

$$\Delta s_i^0 = \frac{q_i^0}{n_i^0} \Delta p_i^0 = \Delta p_i^0. \tag{2}$$

Let U be the $N \times N$ diagonal matrix with $U_{ii} = \frac{q_i^0}{n_i^0}$. In vector form, the change in fund share value is $\Delta S^0 = U \Delta P^0 = U M \Delta F^0$. Because $n_i^0 = q_i^0$, U is the identity matrix[5].

We assume a linear relation between fund performance and net fund flow. Let b_i be the flow-to-performance sensitivity of fund i, i.e., following a change in fund i's share value Δs_i^0, investors redeem

$$\Delta w_i^0 := -a_i \cdot b_i \cdot \Delta s_i^0 \tag{3}$$

shares of the fund. B is the $N \times N$ diagonal matrix with $B_{ii} = b_i$. In vector form, $\Delta W^0 = -AB \Delta S^0$ is the number of redeemed shares per fund.

2. Asset liquidation. Funds liquidate assets to raise cash to repay redeeming investors. We assume that funds sell their holdings proportionately to their portfolio weights. In other words, each fund sells some number of its portfolio units. This pro rata liquidation strategy is the most commonly adopted assumption in the fire-sale literature;[6] it posits that funds aim to hold the same portfolio mix before and after asset liquidation. In Appendix C, we discuss the model with a pecking order liquidation strategy.

Each fund i sells Δq_i^0 units of its portfolio to meet Δw_i^0 redemptions, with Δq_i^0 determined by

$$\Delta q_i^0 \cdot (p_i^0 + \Delta p_i^0) = \Delta w_i^0 \cdot (s_i^0 + \Delta s_i^0);$$

the expression on the left is the cash raised through the sale, and the expression on the right is the cash required. Because $p_i^0 = s_i^0$ and $\Delta p_i^0 = \Delta s_i^0$, it follows that $\Delta q_i^0 = \Delta w_i^0$. Since fund i sells $m_{ik} \Delta q_i^0$ shares of asset k, the total number of shares of asset k liquidated across funds is $\sum_{j=1}^{N} m_{jk} \Delta q_j^0$. In vector form, ΔQ^0 is the number of sold portfolio units per fund, and $M^\top \Delta Q^0$ is the number of sold shares per asset across all funds.

3. Price impact. Asset liquidation has a linear impact on asset prices. After a sale of Δh shares of asset k, the price of asset k declines by $l_k \cdot \Delta h$. L is the $K \times K$ diagonal matrix with price impact

---

[5]In the full model, the number of portfolio units and that of fund shares may instead deviate, and it is therefore more convenient to express quantities using the matrix U.

[6]Greenwood et al. (2015), Duarte and Eisenbach (2021), Fricke and Fricke (2021) make this assumption.

coefficients $L_{kk} = l_k$.

The number of shares of asset k sold by all funds is $\sum_{j=1}^{N} m_{jk}\Delta q_j^0$, so the price of asset k declines by $l_k \sum_{j=1}^{N} m_{jk}\Delta q_j^0$. The change in value of a portfolio unit of fund i due to liquidation costs is then

$$\Delta p_i^1 = -\sum_{k=1}^{K} m_{ik} l_k \sum_{j=1}^{N} m_{jk}\Delta q_j^0.$$

In vector form, $\Delta P^1 = -MLM^\top \Delta Q^0$. Hence, the change in value of fund i's share due to liquidation costs is

$$\Delta s_i^1 = \frac{(q_i^0 - \Delta q_i^0)(p_i^0 + \Delta p_i^0 + \Delta p_i^1)}{n_i^0 - \Delta w_i^0} - s_i^0 - \Delta s_i^0.$$

Since $p_i^0 = s_i^0$, $\Delta p_i^0 = \Delta s_i^0$, $q_i^0 = n_i^0$ and $\Delta q_i^0 = \Delta w_i^0$, we obtain that $\Delta s_i^1 = \frac{q_i^0}{n_i^0}\Delta p_i^1$.[7] Hence, in vector form, $\Delta S^1 = UMLM^\top UAB\Delta S^0$.

4. Further rounds of redemptions and asset liquidation. The change in funds' share values due to the price impact of fire sales triggers further redemptions. Investors redeem an amount $\Delta W^1 = -AB\Delta S^1$ of additional fund shares, funds liquidate $\Delta Q^1 = -UAB\Delta S^1$ portfolio units, which in turn drives down the value of each portfolio unit by $\Delta P^2 = MLM^\top UAB\Delta S^1$, and results in the fund share change in value $\Delta S^2 = UMLM^\top UAB\Delta S^1$. The total fund share value change due to both fire sales and the initial exogenous shock is $\Delta S^\infty := \sum_{n=0}^{\infty} \Delta S^n$, where $\Delta S^n = (UMLM^\top UAB)^n \Delta S^0$ is the change in value after the n-th round of redemptions. (Recall that in the reference model, U is the identity matrix. We have included it here in preparation for the full model.)

3.2   Full Model

In this section, we describe the steps and actions in the full model, which accounts for the presence of first movers in the funds. We refer to all other investors as second movers. Recall that B is the $N \times N$ diagonal matrix with $B_{ii} = b_i$, where b_i is fund i's flow-to-performance sensitivity, and L is the $K \times K$ diagonal matrix with price impact coefficients $L_{kk} = l_k$.

1. First movers' redemptions. Following the initial negative shock $\Delta F^0 = (\Delta f_1^0, \ldots, \Delta f_K^0)^\top$ to asset prices, the value of fund i's portfolio unit declines by Δp_i^0 in (1), and the funds's NAV also declines by Δs_i^0 in (2). We write Δs_i^∞ for the total change in fund i's NAV due both to the initial exogenous shock and subsequent fire sales. We do not yet know Δs_i^∞; it will be determined as a fixed point as we iteratively update the funds' NAVs through subsequent rounds of redemptions and liquidations. We write Δs_i^* for an initial guess of the total NAV change Δs_i^∞.

The proportion of first movers among fund i's investors is π_i, and Π is the $N \times N$ diagonal matrix with $\Pi_{ii} = \pi_i$. Fund i's first movers withdraw their investments in response to the anticipated (as

---

[7]In the reference model, $\frac{q_i^0}{n_i^0} = 1$. We include this coefficient for notational consistency with the full model.

yet unrealized) NAV change Δs_i^* and redeem

$$\Delta w_i^{fm} := -a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^* \tag{4}$$

fund shares. In vector form, $\Delta W^{fm} = -A\Pi B \Delta S^*$ is the number of shares redeemed by first movers. Equation (4) captures the key feature of first movers: they anticipate that liquidation costs will drive down the fund's NAV, and they redeem shares in anticipation of this decline. In contrast, the redemption orders in (3) respond only to the realized decline Δs_i^0.

2. *Asset liquidation to repay first movers.* When mutual fund investors redeem shares, they receive a price per share equal to the NAV at the end of the day that they submitted their redemption orders. As the fund sells assets to meet these redemptions, it incurs liquidation costs that are borne by investors who remain in the fund. In particular, first movers do not bear the liquidation costs they impose on the fund. Each share of fund i redeemed by first movers is repaid at the price $s_i^0 + \Delta s_i^0$, and fund i sells Δq_i^{fm} units of its portfolio to meet first movers' redemptions. Since funds sell assets in proportion to their initial allocations, the total amount of shares of asset k liquidated across all funds is $\sum_{j=1}^{N} m_{jk} \Delta q_j^{fm}$, and the price of a share of asset k declines by $l_k \sum_{j=1}^{N} m_{jk} \Delta q_j^{fm}$. The cash raised by each fund i from asset sales is $\Delta q_i^{fm} \cdot (p_i^0 + \Delta p_i^{fm})$, where

$$\Delta p_i^{fm} = \Delta p_i^0 - \sum_{k=1}^{K} m_{ik} l_k \sum_{j=1}^{N} m_{jk} \Delta q_j^{fm} \tag{5}$$

is the change in value of fund i's portfolio unit due to both the exogenous shock (reflected in Δp_i^0) and asset liquidation (reflected in the double sum in (5)). In vector form, ΔQ^{fm} is the number of portfolio units sold to repay first movers, and $\Delta P^{fm} = \Delta P^0 - MLM^\top \Delta Q^{fm}$ is the resulting price change. In order to meet first movers' redemptions, the number of portfolio units Δq_i^{fm} sold by fund i must satisfy

$$\Delta q_i^{fm} \cdot (p_i^0 + \Delta p_i^{fm}) = \Delta w_i^{fm} \cdot (s_i^0 + \Delta s_i^0). \tag{6}$$

The expression on the left is the cash raised through the sale, and the expression on the right is the cash required to redeem Δw_i^{fm} fund shares. Hence, the vector ΔQ^{fm} is the solution to the system

$$Diag[\Delta Q^{fm}](P^0 + \Delta P^{fm}) = Diag[\Delta W^{fm}](S^0 + \Delta S^0), \tag{7}$$

where $Diag[x]$ is the diagonal matrix whose j-diagonal entry is x_j. Recall that ΔW^{fm} and, as a consequence, ΔQ^{fm} and ΔP^{fm} are functions of the (as yet unknown) total NAV change ΔS^*.[8]

3. *NAV change due to first movers' redemptions.* The share price $s_i^0 + \Delta s_i^0$ received by first

---

[8]In our numerical calculations, we truncate (5) and (6) so that prices never become negative and funds never sell more assets than they own. In our theoretical analysis in Appendix B, we show that these caps are unnecessary for sufficiently small price impact coefficients $(l_k)_k$.

movers does not incorporate the liquidation costs they generate because $s_i^0 + \Delta s_i^0 = p_i^0 + \Delta p_i^0 > p_i^0 + \Delta p_i^{fm}$, and therefore $\Delta q_i^{fm} > \Delta w_i^{fm}$ in (6). As a result,

$$n_i^{fm} := n_i^0 - \Delta w_i^{fm} \geq q_i^0 - \Delta q_i^{fm} =: q_i^{fm}. \tag{8}$$

Here, n_i^{fm} is the number of fund shares remaining after the first-mover redemptions, and q_i^{fm} is the number of portfolio units remaining after the asset sales used to meet these redemptions. Fund i's NAV after first movers' redemptions is $s_i^{fm} = \frac{q_i^{fm}}{n_i^{fm}}(p_i^0 + \Delta p_i^{fm})$, which is the ratio of the fund's assets to the number of fund shares outstanding. The change in NAV observed by remaining investors is $\Delta s_i^{fm} = s_i^{fm} - s_i^0$. Let U^{fm} be the $N \times N$ diagonal matrix with diagonal entries

$$U_i^{fm} = \frac{q_i^{fm}}{n_i^{fm}}. \tag{9}$$

The NAV change due to both the exogenous shock and first movers' redemptions is

$$\Delta S^{fm} = U^{fm}(P^0 + \Delta P^{fm}) - S^0, \tag{10}$$

where the vectors P^0 and S^0 are, respectively, the initial value of a portfolio unit and of a fund share. The NAV change ΔS^{fm} is a function of ΔS^*.

4. Second movers' redemptions. The remaining iterations mirror the reference model. Fund i's second movers observe the NAV change Δs_i^{fm} and redeem $\Delta w_i^{0,sm} = -a_i(1 - \pi_i)b_i\Delta s_i^{fm}$ fund shares, which parallels (3). In vector form, $\Delta W^{0,sm} = -A(1-\Pi)B\Delta S^{fm}$. Following the same steps as in the reference model, redemptions force funds to sell assets, further depressing asset prices and fund NAVs. More precisely, the impact of second movers' redemptions on each fund's NAV is

$$\Delta S^{1,sm} = U^{fm}MLM^\top U^{fm}A(I - \Pi)B\Delta S^{fm}.$$

This NAV change triggers further rounds of redemptions by second movers. The total change in each fund's NAV is

$$\Delta S^\infty = \sum_{n=0}^\infty \Delta S^{n,sm} \tag{11}$$

where $\Delta S^{n,sm} = (U^{fm}MLM^\top U^{fm}A(I - \Pi)B)^n\Delta S^{fm}$.

5. Total NAV change. The total NAV change $\Delta S^\infty(\Delta S^*) = \sum_{n=0}^\infty \Delta S^{n,sm}$ computed in the previous steps depends on the initial guess ΔS^* through ΔS^{fm}. But recall that we assume that first movers correctly anticipate the full NAV impact of the initial shock and subsequent liquidations. This holds when $\Delta S^* = \Delta S^\infty(\Delta S^*)$; that is, when the anticipated NAV impact is a fixed point of the mapping defined by (11). The next proposition establishes the existence of such a fixed point.

Proposition 1. *Assume that M has nonnegative entries, each price impact coefficient l_k is suffi-*

11

ciently small, $b_i < 1$ for each i, and $\Delta s_0^i > -s_0^i$ for each i.[9] *Then there exists a fixed point of the mapping $\Delta S^* \to \Delta S^\infty(\Delta S^*)$ defined in step 5 of the above procedure.*

3.3 Aggregate Vulnerability Measure

We measure the aggregate vulnerability of the mutual fund sector as the total amplification of losses through the sector. We measure this amplification through the ratio between the endogenous losses, due to fund redemptions and fire sales, and the exogenous losses caused by the initial shock only. Formally, we define the Spillover Loss Ratio as

$$SLR := \frac{\sum_i a_i \Delta s_i^{sl}}{\sum_i a_i \Delta s_i^0},$$

where the sum is over funds, a_i is fund i's asset value, and $\Delta s_i^{sl} := \Delta s_i^\infty - \Delta s_i^0$ is the NAV change due exclusively to the feedback loop between fund redemptions and fire sales of assets needed to meet these redemptions.

We impose a cap on both the number of portfolio units that each fund can sell and the price impact imposed on each asset. A fund cannot sell more portfolio units than it owns, so the total number of liquidated portfolio units Δq_i is capped at a_i. A fund *fails* if it liquidates all of its assets. Furthermore, asset prices cannot become negative as a result of price impact from sales.

3.4 First Mover Concentration and NAV Change

We quantify analytically how the distribution of first movers across funds impacts the total change in NAV. In Proposition 2, we show that a higher concentration exacerbates the feedback loop between fund redemptions and asset sales, and imposes a higher downward impact on the NAV. Technical details about the mathematical set-up and the proof of the proposition are relegated to Appendix D.

Proposition 2. *Consider two funds holding identical portfolios, both subject to an initial negative shock Δs^0. Let $\frac{\bar\pi}{2} \in (0, \frac{1}{2}]$ be the proportion of first movers in the system, and let $\pi \in \left(\frac{\bar\pi}{2}, \bar\pi\right)$ be the proportion of first movers in the first fund. The proportion of first movers in the second fund is $\bar\pi - \pi$. If the price impact is sufficiently small, then for all π there exists a fixed point $\Delta S^*(\pi) = (\Delta s_1^*(\pi), \Delta s_2^*(\pi))^\top$ of the mapping $\Delta S^* \to \Delta S^\infty(\Delta S^*)$ such that $\Delta s_1^*(\pi) + \Delta s_2^*(\pi)$ is decreasing in π. Since $\Delta s^0 < 0$, this implies that the spillover loss ratio is increasing in π.*

As stated in the proposition, the aggregate exposure of the funds to redemption and fire sales is minimized if first movers are evenly distributed between the two funds. This result has implications for policies aimed at mitigating first-mover externalities. It warns that a regulatory intervention that unintentionally alters the distribution of first movers across funds could adversely affect financial stability. Consider, for example, the use of swing pricing, in which a fund's NAV is adjusted so

---

[9]An initial shock $\Delta s_0^i = -s_0^i$ implies that fund i's asset holdings become worthless, and its fire sales are inconsequential. Hence, the fixed point of the system can be computed without including fund i.

that redeeming or purchasing investors bear the trading costs that result from their transactions. In a mutual fund system where only half of the funds implement swing pricing, first movers might migrate to similar funds that do not implement it, and as a result make the system more fragile, based on Proposition 2. This suggests that while swing pricing reduces fire-sale losses at the individual fund level and — if widely adopted — also for the whole system, it may be less effective or even damaging if it is only implemented by a small group of funds.

4   Mutual Fund Aggregate Vulnerability

In this section, we apply the model to the system of U.S. mutual funds and estimate the system's Spillover Loss Ratio from data.

4.1   Data Description

We use quarterly mutual fund holding data from the CRSP Survivor-Bias-Free US Mutual Fund Database spanning the period Q1 2010 through Q4 2020. For each date, we remove from the database ETFs, funds with missing information, and funds with less than \$5 million in total net assets. The database includes total net asset value of each fund, and groups each fund's holdings into the twelve asset classes listed in Table 4.1. We divide funds into nine types, according to their CRSP Style Code. The types are equity domestic (ED), equity foreign (EF), fixed income municipal (IU), fixed income corporate (IC), fixed income government (IG), fixed income foreign (IF), other fixed income (I), mixed fixed income and equity (M), and other (O). For each type, we work with the 100 largest funds, and we combine the holdings for the remaining funds into a single aggregate fund.[10] Hence, for each quarter the system consists of at most 909 funds, and we use these funds to construct the matrices A and M from the CRSP data.

We use price impact parameters estimated under stressed trading conditions by Bouveret and Yu (2021).[11] To account for time varying liquidity, we construct a price impact matrix L_t that depends on time t. The parameters in Table 4.2 pin down the matrix $L_t = L^*$ at the initial date of our analysis (the benchmark date), which is Q1 2010. The price impact matrix is then renormalized by the size of the financial sector on subsequent dates to capture the idea that the pool of potential buyers of fund assets varies over time. For this calculation, we follow a similar approach to Duarte and Eisenbach (2021). As a proxy for the wealth w_t of potential buyers of liquidated assets, we take the value of assets held by the U.S. financial sector and U.S. households minus the value of mutual fund shares they hold. We source this data from the "Financial Accounts of the United

---

[10]We have verified that aggregating funds at different levels of granularity does not significantly affect our results. Aggregation may even understate vulnerability, because it removes the first-mover heterogeneity within each aggregated fund.

[11]Greenwood et al. (2015) assume that a net trade of 10 billion euros leads to a price change of 10 basis points, regardless of the liquidated asset. Duarte and Eisenbach (2021) consider heterogeneous price impact parameters implied by the Net Stable Funding Ratio of the Basel III regulatory framework.

States".[12] The price impact matrix at date t is $L_t = \frac{w^*}{w_t} L^*$, where w^* is the value of w_t at the benchmark date.

The CRSP database classifies every fund share class as either institutional or retail. We measure the proportion of first-mover investors π_i in fund i as the proportion of total net assets held by institutional share classes within fund i. This identification is supported by the empirical evidence and discussion in Section 2. We will also investigate the effect of varying the proportion of first movers. Observe that our measure π_i depends on the quarter t.

Prior research has studied the relationship between fund flows and performance. For example, Franzoni and Schmalz (2017) find that the sensitivity of flow to performance strongly depends on the state of the market and can range from 20% to around 70%. These estimates cannot disentangle the direct response measured by the coefficient b in our model from the combined effect of first- and second-mover redemptions. We will therefore examine the impact of different values of b, holding this parameter constant across funds.

We apply shocks of different magnitudes to different asset classes, based on their relative volatilities. For example, to translate a 10% drop in stock prices to an equally severe shock to municipal bonds, we would use a drop of 3.981%, based on the relative volatilities in Table 4.2. To calculate the relative volatilities, we use daily returns during Q1 2020 (the Covid-19 shock) on representative ETFs for each asset class. We use the Vanguard Total Stock Market ETF (VTI) for common and preferred stocks and other equities; the iShares Convertible Bond ETF (ICTV) for convertible bonds; the Vanguard Total Bond Market Index Fund ETF (BND) for corporate bonds and other fixed-income securities; the iShares National Municipal Bond ETF (MUB) for municipal bonds; the iShares US Treasury Bond ETF (GOVT) for government bonds; and the iShares MBS ETF (MBB) for mortgage-backed securities and asset-backed securities. No shock is applied to the "Other Securities" class.

4.2 Mutual Fund Vulnerability in the Reference Model

We begin by measuring spillover losses in the reference model without first movers and then measure the impact of accounting for first movers. Through portfolio overlap, as reflected in M, fire sales can spread from one asset to another. We refer to the matrix $MLM^\top AB$ as the *systemicness matrix*. The total change in each fund's share value is then given by the vector

$$\sum_{n=0}^{\infty} (MLM^\top AB)^n \Delta S^0,$$

At each round of redemptions, the vector of shocks is multiplied by the systemicness matrix. If its spectral radius is smaller than 1, then the spillover losses of each round are smaller than losses from the previous round of redemptions. If instead the spectral radius is larger than 1, the vector

---

[12]The corresponding codes are FL794090005 (Domestic financial sectors; total financial assets), FL154090005 (Households and nonprofit organizations; total financial assets), FL793064205 (Domestic financial sectors; mutual fund shares; asset), FL153064205 (Households and nonprofit organizations; mutual fund shares; asset).

	Domestic Equity	Foreign Equity	FI Corporate	FI Foreign	FI Government
Total assets ($ billions)	6,372	2,189	152	273	242
Institutional investors (percent)	37.51	46.72	36.16	53.24	47.76
Portfolio shares (percent):					
Cash	1.95	1.90	1.86	4.62	1.63
Common Stocks	87.39	84.66	0.37	0.10	0.03
Preferred Stocks	0.18	0.97	0.40	0.05	0.01
Convertible Bonds	0.13	0.04	0.65	0.09	0.00
Corporate Bonds	1.55	0.88	55.11	21.56	5.14
Municipal Bonds	0.05	0.03	3.09	1.65	0.30
Government Bonds	2.61	1.57	18.99	63.20	74.44
Asset-Backed Securities	0.26	0.06	4.41	1.53	5.85
Mortgage-Backed Securities	0.57	0.13	10.32	1.73	9.84
Other Equities	1.95	7.12	0.09	0.07	0.00
Other Fixed-Income Securities	0.32	0.12	2.45	1.48	1.51
Other Securities	3.03	2.52	2.25	3.92	1.25

	FI Muni	FI Other	Mixed FI & Equity	Other
Total assets ($ billions)	605	1,912	1,746	351
Institutional investors (percent)	23.29	50.49	25.95	49.65
Portfolio shares (percent):				
Cash	1.25	2.15	2.67	5.96
Common Stocks	0.04	0.47	51.29	4.94
Preferred Stocks	0.02	0.29	0.62	0.10
Convertible Bonds	0	0.29	1.46	0.34
Corporate Bonds	0.17	41.67	14.74	13.07
Municipal Bonds	97.40	1.53	0.64	0.44
Government Bonds	0.10	23.00	13.30	8.48
Asset-Backed Securities	0.02	7.75	1.70	11.38
Mortgage-Backed Securities	0.02	15.29	4.43	26.23
Other Equities	0.01	0.17	2.40	0.27
Other Fixed-Income Securities	0.40	4.64	1.98	22.22
Other Securities	0.56	2.75	4.78	6.55

Table 4.1: Summary of the balance sheet data used to compute aggregate vulnerability. The table shows average total net assets, proportion of assets held by institutional fund share classes, and aggregate portfolio composition for each fund type over the period from Q1 2010 to Q4 2020.

Asset Class	Price Impact	Relative Volatility
Cash	0	0
Common Stocks	2.8×10^{-13}	1
Preferred Stocks	2.8×10^{-13}	1
Convertible Bonds	7.7×10^{-13}	0.8710
Corporate Bonds	7.7×10^{-13}	0.3169
Municipal Bonds	14.5×10^{-13}	0.3981
Government Bonds	0.3×10^{-13}	0.1905
Asset-Backed Securities	0.5×10^{-13}	0.1829
Mortgage-Backed Securities	0.5×10^{-13}	0.1829
Other Equities	2.8×10^{-13}	1
Other Fixed-Income Securities	0.3×10^{-13}	0.3169
Other Securities	0	0

Table 4.2: A price impact of 10^{-13} indicates that a \$10 billion net trade leads to a price decline of 10 basis point. The second column is the relative daily volatility, over Q1 2020, of an ETF representative of each asset class compared to that of equity.

of NAV shocks ΔS^0 can get amplified in each iteration. The spectral radius of the systemicness matrix is therefore a measure of aggregate fund exposure to fire sales caused by redemptions.

The systemicness matrix can be decomposed into three factors, analogously to the decomposition of aggregate vulnerability in Duarte and Eisenbach (2021): $MLM^\top$ is the *illiquidity concentration*, A is the *size* of the system, B is the *flow-to-performance sensitivity*. The (i,j) entry of the illiquidity concentration matrix $MLM^\top$, $\sum_{k=1}^{K} l_k m_{ik} m_{jk}$, is the liquidity-weighted portfolio overlap of funds i and j. Recall that the entries of the diagonal matrices A and B are, respectively, the net total assets and flow-to-performance sensitivity of each fund. The spectral radius is therefore larger, and the system more vulnerable, if large funds with high flow-to-performance sensitivity have significant portfolio overlap on illiquid assets.

The magnitude of the Spillover Loss Ratio is directly related to the value of the spectral radius. In the right panel of Figure 4.1, we compute the SLR for an initial exogenous shock of -5% multiplied by the relative volatilities in Table 4.2 for each asset class. Spillover losses dwarf initial losses if the spectral radius is close to 1 or larger.[13] Moreover, in recent years, the spectral radius has exceeded 1 for large, yet plausible, values of flow-to-performance sensitivity.

Investors that hold their assets directly, rather than through a mutual fund, may also liquidate them if their portfolios are subject to a negative shock. As a result, they would drive down asset prices. If holding a portfolio directly or through a fund does not affect investors' sensitivity to performance, the spillover losses quantified using the reference model would remain in the absence of mutual fund intermediation. However, as we demonstrate in the next section, spillover losses would be greater if the assets are intermediated by the fund, after accounting for the first-mover advantage.

---

[13]Observe that spillover losses are finite because of the imposed caps discussed in Section 3.3.
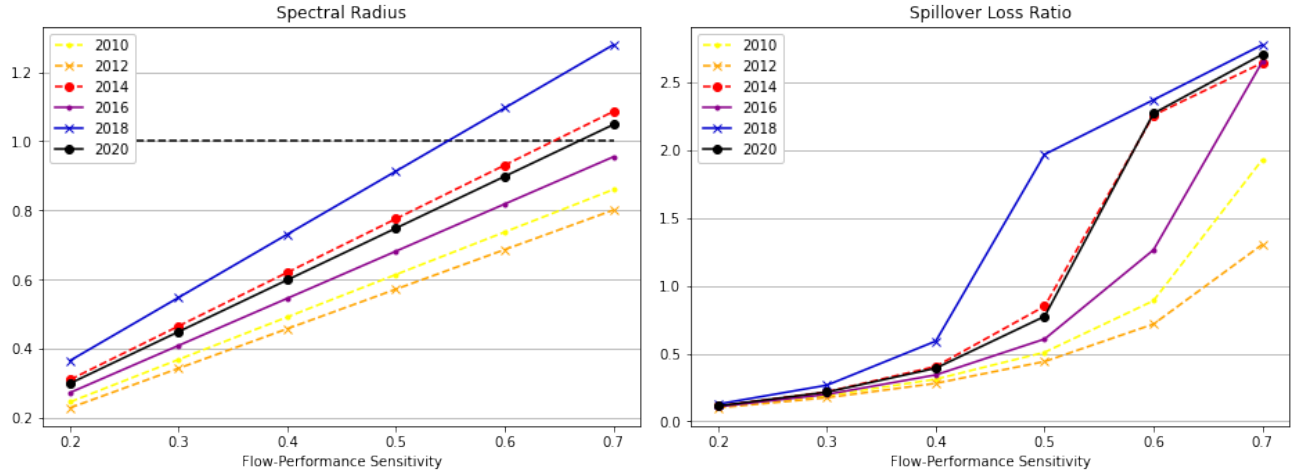
Figure 4.1: Flow-to-performance sensitivity is assumed constant across funds, data refer to the first quarter of each year. The left panel shows the spectral radius of the systemicness matrix for different values of flow-to-performance sensitivity and different years. The right panel shows the Spillover Loss Ratio for different values of flow-to-performance sensitivity and different years.

4.3 Impact of First-Mover Advantage

We now quantify the share of spillover losses that can be attributed to funds' liquidity mismatch and the resulting first-mover advantage.

4.3.1 Spillover Loss Ratio over Time

The presence of first-mover investors exacerbates the vulnerability of a fragile system but has minimal impact on a resilient one. In fact, first movers have limited incentive to exit a fund early, if asset liquidation costs are low, i.e., when the spectral radius of the systemicness matrix is significantly below 1. However, the first-mover advantage has a strong destabilizing effect on a system that is already vulnerable: if first movers expect funds to face significant spillover losses, then they benefit from redeeming their fund shares early, accelerating a systemic fire-sale spiral. In Figure 4.2, we compute the SLR with and without first movers for an initial price change of -5% multiplied by the relative volatilities in Table 4.2 for each asset. If in the absence of first movers the system would be resilient to spillover losses, e.g., if flow-to-performance sensitivity is low, then the impact of first movers is negligible. However, the fragility of a system that is moderately vulnerable without first movers may deteriorate significantly when accounting for the first-mover advantage. As shown in Figure 4.2, after the year 2017 and assuming a flow-to-performance sensitivity of 45%, the SLR in the full model is often at least twice as large as in the reference model.

4.3.2 Contributing Factors to Spillover Losses

The vulnerability of the mutual fund system is sensitive to several factors (see Figure 4.3). The first factor is the size of the U.S. mutual fund industry relative to the whole U.S. financial sector.
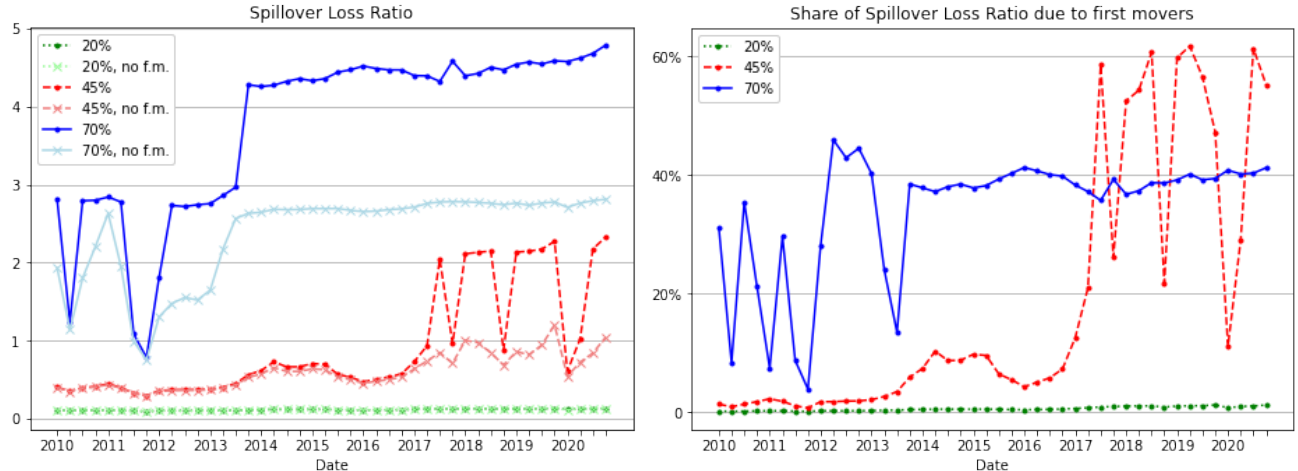
Figure 4.2: The left panel shows the Spillover Loss Ratio over time for different values of flow-to-performance sensitivity with/without first movers. The right panel shows the proportion of Spillover Loss Ratio due to the presence of first movers.

Over time, funds have accounted for an increasingly large share of the whole financial market. To see this, compare the first quarter of 2010 when assets held by mutual funds accounted for less than 12% of all financial assets, with the last quarter of 2019 when this proportion grew to more than 16%.

A second factor is the concentration of fund holdings in illiquid assets. Let a^m be the aggregate asset value held by mutual funds and a^{tot} the total value of assets in the whole financial system. The matrix $C := \frac{a^{tot}}{a^m} \cdot MLM^\top A$ quantifies the impact that portfolio overlap in illiquid assets has on each fund. Notice that this matrix is independent of the size of the system: the entries of $\frac{A}{a^m}$ are the weights of each fund in the system, and the entries of $a^{tot}L$ are (approximately) a size-independent measure of each asset's illiquidity. (Recall that our specification of price impact is such that assets are more liquid as the size of the whole financial system increases.) We measure the amplification effect due to portfolio concentration in illiquid assets using the spectral radius of the matrix C: it is the largest asset price change triggered by any vector of initial NAV shocks of a specified size. Notice that accounting only for the impact of C on the initial vector of shocks ΔS^0 does not capture vulnerability due to portfolio commonality. This is because we consider multiple rounds of redemptions and fire sales and, in each round, the vector of realized shocks across asset classes may be different compared to the previous round. As seen from the top right panel of Figure 4.3, the impact of illiquidity concentration on the system's vulnerability has increased steadily since 2013.

A third factor is the propensity of investors to redeem fund shares in response to a decline in fund NAV. The stronger the reaction of investors to negative NAV shocks, the more vulnerable the system to asset fire sales. As shown in Goldstein et al. (2017), funds that hold more illiquid assets have a higher sensitivity of outflows to bad performance. Even if our analysis assumes that the flow-to-performance sensitivity b_i is the same across funds, a fund holding illiquid assets is subject to more redemptions after a negative initial shock than a fund holding liquid assets because
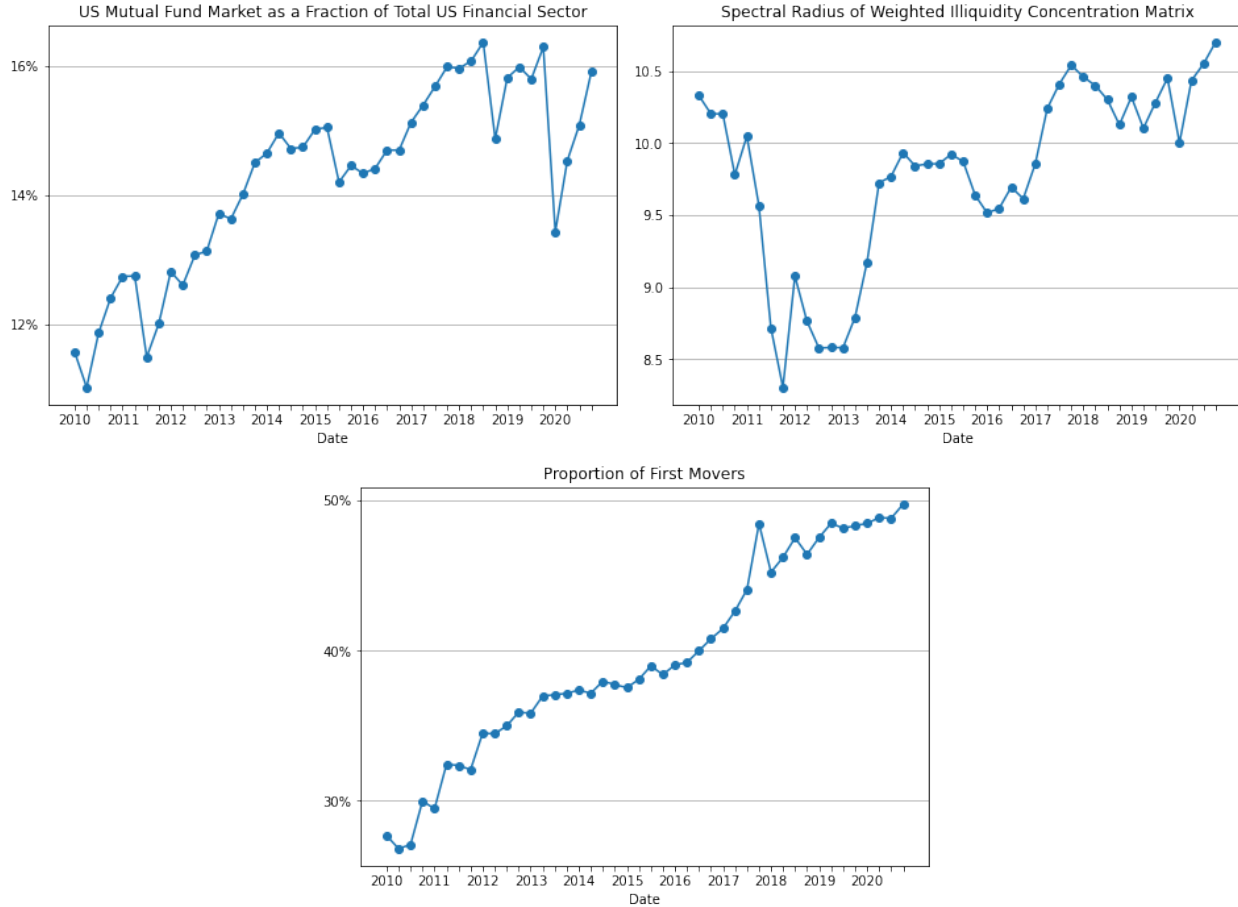
18

Figure 4.3: The left panel plots the size of the U.S. mutual fund industry relative to the whole U.S. financial sector over time. The right panel plots the spectral radius of the matrix $C = \frac{a^{tot}}{a^m} \cdot MLM^\top A$ over time. The bottom panel plots the proportion of assets held by institutional fund share classes (our proxy for first movers) over time. The systemicness matrix is defined as $\frac{a^m}{a^{tot}} CB$. Therefore, differences in the magnitudes of the spectral radius of C and that of the systemicness matrix are due to the relative size of the mutual fund industry $\frac{a^m}{a^{tot}}$ and to the flow-to-performance sensitivity matrix B (set as a multiple of the identity matrix in all examples in the paper).

first movers anticipate the higher spillover losses and therefore have a stronger incentive to redeem early.[14] Hence our findings are consistent with those in Goldstein et al. (2017).

The fourth factor is the proportion of institutional investors among holders of fund shares. Early redemptions by first movers increase asset liquidation pressure and, hence, spillover losses. The presence of more first movers creates additional feedback effects, as other first movers account for their withdrawals and hence redeem additional fund shares. It can be seen from the bottom graph of Figure 4.3 that the fraction of assets held in institutional fund share classes, our proxy for the proportion of first movers, has increased to nearly 50% in the year 2020. Even though we presented

---

[14]Even in the absence of first movers, redemptions would be higher for illiquid funds because of the feedback loop between price declines and redemptions. The presence of first-movers significantly amplifies the feedback loop between spillover losses and number of redemptions.
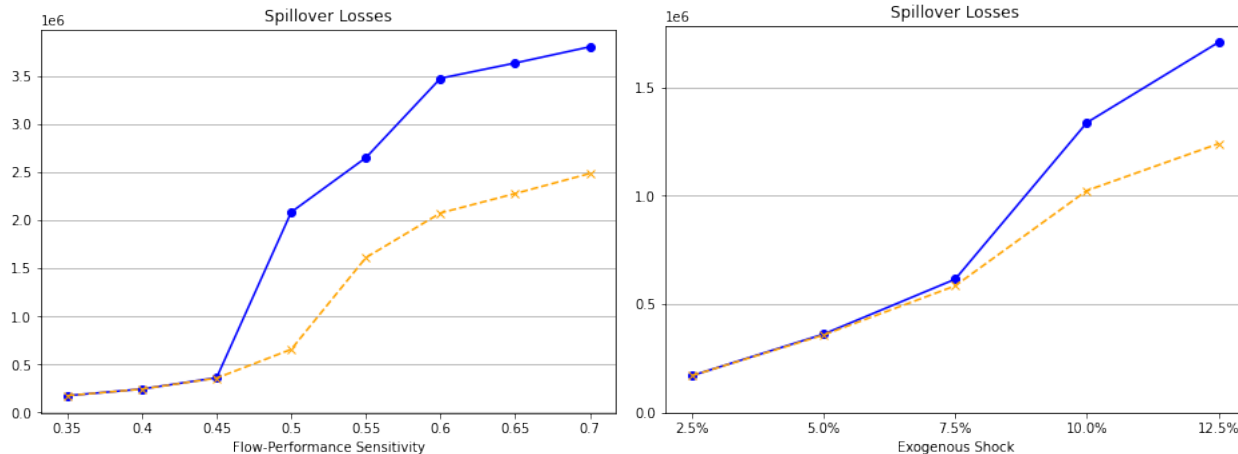
Figure 4.4: Spillover losses when first movers are concentrated in half of the funds (solid line) and when first movers are evenly distributed across funds (dashed line). In the left panel, the exogenous shock applied to each asset equal is equal to a price change of -5% multiplied by its relative volatility specified in Table 4.2. In the right panel, flow-to-performance sensitivity is set to 45% and the exogenous shock to each asset is obtained by multiplying the price change on the x-axis by each asset's relative volatility. We use fund holdings data from Q1 2020.

the proportion of first movers as a separate factor that affects the system's vulnerability, we cannot disentangle the impact of illiquidity concentration from that of first movers. This is because we consider a system with first mover heterogeneity: fragility is magnified if funds holding concentrated portfolios have a higher proportion of first movers. Even in an otherwise homogeneous system, if first movers are concentrated in fewer funds, the system would be more fragile (as discussed in Section 4.3.3).

4.3.3 Nonlinearity of Spillover Losses due to First Movers

We demonstrate how the nonlinearity introduced by first-mover incentives exacerbates the impact of first-mover concentration and initial shocks on spillover losses.

To analyze the impact of first-mover distribution across funds, we split every fund into two identical funds, each holding half of the assets of the original fund. We compare two system configurations for the distribution of first-mover investors. In the first configuration, we set the proportion of first movers in every fund equal to 50%. In the second configuration, for each pair of identical funds, the first fund's shares are entirely owned by first movers, and the second fund's shares are owned only by second movers. Hence, the total number of first movers is the same across the two configurations, but in the second configuration first movers are concentrated in just half of the funds in the system.

Figure 4.4 illustrates the spillover losses for each of these two configurations using fund holdings data from Q1 2020. The system in which first movers are highly concentrated on fewer funds is more fragile than the system in which first movers are evenly distributed across funds, consistent with Proposition 2. The higher fragility is explained by the nonlinearity in spillover losses created

by the first-mover advantage: the feedback between fire sales and fund redemptions is stronger in funds with a high proportion of first movers, and the resulting downward pressures imposed on asset prices may also hit funds without first movers. The difference in vulnerability between the system with first mover concentration and the system in which first movers are evenly spread across funds is small in the market scenarios where the flow sensitivity to performance is low. In these market scenarios, the incentive to run is small, and thus fire-sale losses are not impacted much by the distribution of first movers in the system.

We next study the amplification of initial shocks created by redeeming first movers. In the reference model, given the linear assumptions on price impact and flow sensitivity to performance, spillover losses scale linearly with the size of the initial exogenous shock; the SLR increases in proportion to the initial shock. But the reference model fails to capture the incentive to run observed with first movers. Figure 4.5 shows that spillover losses grow faster and nonlinearly in the size of the initial exogenous market shock once we account for the first-mover advantage.

In Figure 4.5, assets are subject to initial shocks ranging from −2.5% to −15% times their corresponding relative volatilities in Table 4.2. Consider, first, the results using parameters for 2020. As we increase the the exogenous shock from 7.5% to 12.5%, the SLR for the reference model increases linearly, as expected. Over the same range, the SLR accounting for first movers grows far more, with an inflection in the growth rate at a 7.5% shock. The SLR growth plateaus at larger initial shocks because of the caps we impose on price impact and on the quantity of assets that funds can sell.

For the years 2012 and 2016, we observe little impact of first movers on spillover losses in Figure 4.5. This can be explained by the measures plotted in Figure 4.3, where it can be seen that mutual funds accounted for a smaller percentage of the U.S. financial sector, had smaller portfolio overlap in illiquid assets, and had a lower proportion of first movers. This comparison shows that simply extrapolating from the environment in 2012 and 2016 would miss the greater vulnerability of the system in 2020.

This is the main takeaway from Figure 4.5. In a sufficiently fragile system, there is a critical size for the initial shock that triggers a wave of redemptions that magnify spillover losses substantially. As a consequence, spillover losses in ordinary times are not a good gauge of the aggregate vulnerability of the system or of the magnitude of potential spillover losses in a heavily stressed economy.

4.3.4 Portfolio Commonality, First Movers, and Asset Price Contagion

Price shocks can spread across mutual funds and asset classes through portfolio commonality. A fire sale by one fund drives down the share price of other funds holding the same assets; and a price drop in one asset class may force a fund to liquidate other assets, driving down their prices, in order to meet redemptions. These contagion effects and selling by first movers can reinforce each other.

To study the joint impact of portfolio commonality and first movers on financial fragility, for
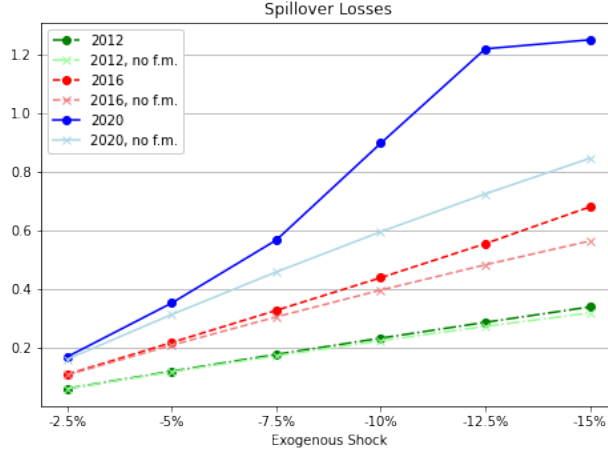
Figure 4.5: Spillover losses (in trillions of dollars) for different initial exogenous shocks with/without first movers. Flow-to-performance sensitivity is set equal to 45%.

each fund type i[15] we consider a benchmark system in which funds of type i are not connected to other funds in the system. In such a system, asset liquidation by funds of type i do not impact others in the system, and vice versa. We then compare the benchmark with the original interconnected system, both with and without first movers.

The figure shows that isolating a fund type from the rest of the system can significantly reduce the total spillover losses, either because it shields some large funds from fire-sale externalities, or because it reduces the spread of the shock across asset classes. The left plot of the figure uses asset holdings data from the end of Q1 2020, when prices were already severely depressed by the Covid-19 shock, and it shows that the impact of first movers would be modest if the economy were subject to an additional exogenous shock. By contrast, for Q4 2020, spillover losses due to portfolio commonality are significantly higher in the presence of first movers. It is the first-mover advantage that, in this case, fuels the spread of shocks through the system via the contagion channel stemming from portfolio commonality. Portfolio commonality and first movers have compounding effects.

We next study how shocks spread across asset classes through the portfolio commonality channel. We consider a scenario in which a few assets are subject to a large initial shock, and we aggregate funds within each of the nine types. Figure 4.7 shows the total returns including fire-sale losses for all nine aggregate funds, after applying a shock to convertible, corporate and municipal bonds equal to a price change of 20% multiplied by the corresponding realized volatilities of these assets.

For low levels of flow-to-performance sensitivity, spillover losses are inconsequential and the shock does not spread across the system. In fact, the fund sectors that are most impacted are those holding the assets subject to the initial shock. This is not the case if investors react more strongly to fund performance. Large redemptions at funds that hold both fixed income assets — affected

---

[15]The fund types are equity domestic (ED), equity foreign (EF), fixed income municipal (IU), fixed income corporate (IC), fixed income government (IG), fixed income foreign (IF), other fixed income (I), mixed fixed income and equity (M), and other (O).
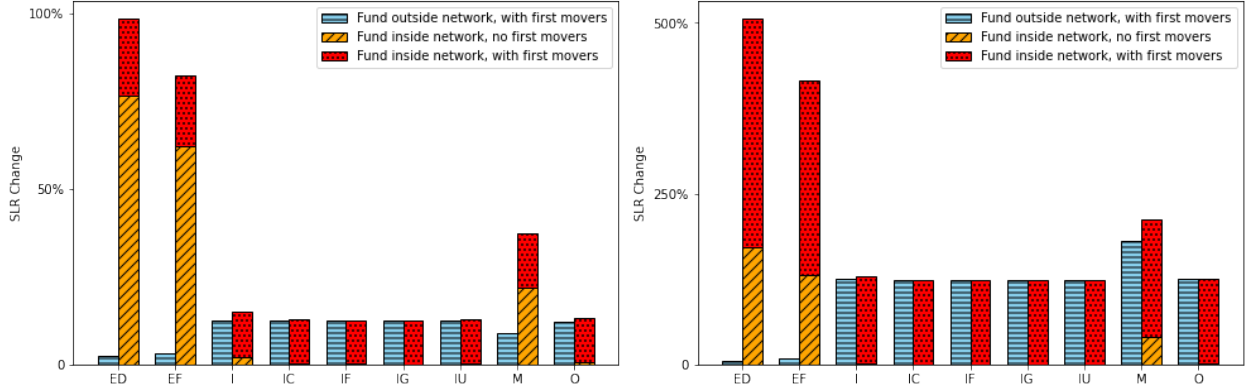
Figure 4.6: Change in spillover losses when each fund type is isolated from other fund types, with and without first movers. For each fund type, the bar with horizontal lines shows the increase in spillover losses due to first movers if funds of this type are isolated from others. The bar with diagonal lines and the dotted one show the increase in spillover losses if these funds are connected to the rest of the system — respectively with and without first movers — relative to the case in which they have no portfolio commonality with other funds and there are no first movers in the system. We set the flow-to-performance sensitivity to 45%. We apply initial shocks of -5% times the corresponding realized assets' relative volatilities. We consider portfolio holdings in the first quarter (left plot) and fourth quarter (right plot) of 2020.

by the exogenous shock — and equity assets may lead to sell-offs in asset classes not hit by the initial shock, and cause widespread spillover losses through the system. As the flow-to-performance sensitivity increases, equity funds become the most vulnerable to spillover losses, even though we applied the initial shock exclusively to fixed income assets. This is because the initial shock spills over to the equity asset class via the portfolio overlap of mixed funds.

5 Conclusion

We have developed a framework to quantify the vulnerability of the mutual fund sector to fire sales triggered by fund redemptions. The distinguishing feature of our framework is that it accounts for the liquidity mismatch that arises when mutual funds hold illiquid assets but provide same-day liquidity to their investors. We have constructed measures that quantify the mutual fund sector's vulnerability and its sensitivity to key parameters such as the distribution of first movers, shock size, and flow-to-performance sensitivity. We have evaluated these measures using mutual fund holdings data during stressed market conditions. Our framework can serve as a tool to test the impact of policies aimed at reducing spillover losses due to fund runs and common portfolio holdings.

We have shown that the first-mover incentive introduces a nonlinear dependence between spillover losses and the size of initial asset shocks. This nonlinearity can severely exacerbate the aggregate vulnerability of the system for large, yet plausible, sizes of initial shocks if first movers are concentrated in fewer funds or if the investor base of illiquid funds includes a high proportion of first movers.
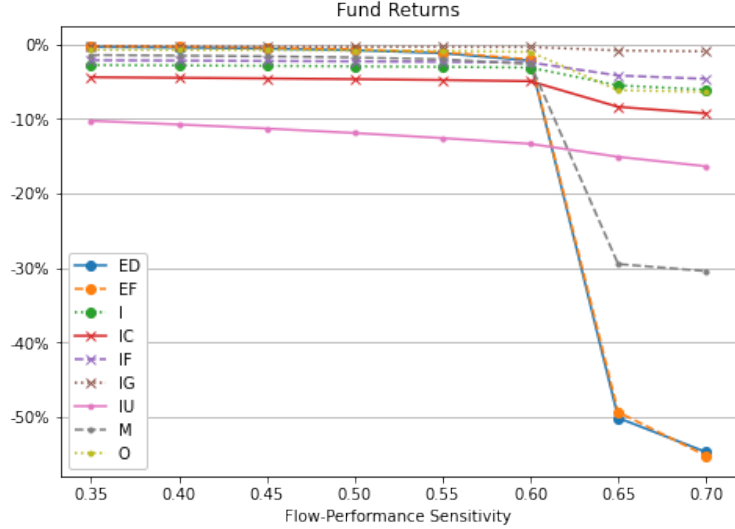
Figure 4.7: Fund returns for different values of fund-to-performance sensitivity for an exogenous shock to convertible, corporate and municipal bonds, equal to a price change of -20% multiplied by the realized volatility of each asset. We use asset holdings data from Q1 2020.

A   Cross-Sectional Regression of Municipal Bond Fund Flows

To check the robustness of the results in Section 2, we run the following cross-sectional regression using data in the CRSP database for Q1 2020:

$$Flow_i = \alpha + \beta_{inst} InstReturn_i + \beta_{retail} RetailReturn_i + \gamma Controls_i + \varepsilon_i,$$

where $Flow_i$ is the flow for fund i, defined as $\frac{TNA_i^{end} - TNA_i(1 + Return_i)}{TNA_i}$, where TNA_i^{end} is the total net asset value of fund share class i at the end of the quarter, TNA_i is the total net asset value at the beginning of the quarter, and $Return_i$ is the fund share class's return. For institutional fund share classes, $InstReturn$ is the fund share class's return, and $RetailReturn$ is set to 0. For retail fund share classes, $RetailReturn$ is the fund share class's return, and we set $InstReturn$ to 0. We control for lagged flow (the flow over the previous quarter), Log(TNA) (the logarithm of total net assets held by the fund at the beginning of the quarter), and Log(age) (the logarithm of the fund's age at the beginning of the quarter, expressed in years). We regress flows against contemporaneous returns (after fees), and not against returns over the previous quarter. Our specification is designed to capture the relation between the Covid-19 market shock and flows within the same quarter. In Table A.2, we report the summary statistics for the fund share classes in our sample. Table A.3 reports the results of the regression. The relation between flows and returns is statistically significant at the 1% level for both institutional and retail fund share classes. Returns are associated with outflows that are larger for institutional fund share classes (0.620) compared to retail fund share classes (0.381), consistent with the view that institutional investors react more strongly to negative returns than retail investors. The difference is on the borderline

of the conventional standard for significance: an F-test of the hypothesis that $\beta_{inst} = \beta_{retail}$ has a p-value of 0.060.

Table A.1: Summary statistics for characteristics of fund share classes in the sample for the panel regression in Section 2. We report the mean, median, standard deviation, 5th percentile (P5), 95th percentile (P95) and total number of observations (N).

	Institutional Fund Share Classes					
	Mean	Median	Std dev	P5	P95	N
Flow	-0.0105	0.0051	0.4716	-0.6846	0.5479	26,260
Return	-0.0219	0.0100	0.8984	-1.5800	0.6300	26,260
Log(TNA)	4.0762	4.4238	2.6682	-1.3934	7.6502	26,260
	Retail Fund Share Classes					
	Mean	Median	Std dev	P5	P95	N
Flow	-0.0176	-0.0009	0.4112	-0.5647	0.4318	49,259
Return	-0.0246	0.0100	0.9437	-1.6600	0.6600	49,259
Log(TNA)	3.8453	4.0842	2.5314	-0.4805	7.4571	49,259

Table A.2: Summary statistics of fund share classes' characteristics in our sample, used for the cross-sectional regression in Appendix A. We report the mean, median, standard deviation, 5th percentile (P5), 95th percentile (P95) and total number of observations (N).

	Mean	Median	Std dev	P5	P95	N
Flow	-0.0089	-0.0144	0.0747	-0.1114	0.1155	1436
InstReturn	-0.0238	-0.0182	0.0213	-0.0711	-0.0001	544
RetailReturn	-0.0238	-0.0191	0.0199	-0.0702	-0.0016	892
Lagged Flow	0.0340	0.0217	0.0858	-0.0810	0.1902	1436
Log(TNA)	4.7140	4.6250	1.7058	2.1604	7.6902	1436
Log(age)	2.5697	2.8396	0.8726	0.8823	3.5440	1436

B   Existence of a Fixed Point

In this section, we show that the procedure described in Section 3.2 has a fixed point. Before stating the main result, we state and prove a technical lemma which will be used in the proof of Proposition 1.

Lemma B.1. *Suppose that $f(x, y)$ is continuous in $(x, y) \in X \times Y$, and strictly monotone in x for each y, where $X \subset \mathbb{R}$ and $Y \subset \mathbb{R}^d$ are compact. Then for any sequence $(x_n, y_n) \in X \times Y$ with $\lim_{n\to\infty} y_n = y_0$ and $f(x_n, y_n) = 0$ for all n, there is an $x_0 \in X$ for which*

$$\lim_{n\to\infty} x_n = x_0, \quad f(x_0, y_0) = 0.$$

Table A.3: Relation between flows and returns in municipal bond funds. We source data from the CRSP database for Q1 2020. Flow is the proportional fund share class flow over Q1 2020. InstReturn is the return over Q1 2020 if the fund share class is institutional and 0 otherwise. RetailReturn is the return over Q1 2020 if the fund share class is retail and 0 otherwise. Lagged Flow is the flow over Q4 2019. Log(TNA) is the natural logarithm of total net assets at the beginning of Q1 2020. Log(age) is the natural logarithm of the fund share class age (expressed in years) at the beginning of Q1 2020. We removed index funds, ETFs, ETNs, fund share classes with TNA lower than 5 million dollars, and fund share classes less than one year old. Flows are winsorized at the 1st and 99th percentiles.

	Dependent Variable:
	Flow
Constant	0.041***
	(0.01)
InstReturn	0.620***
	(0.12)
RetailReturn	0.381***
	(0.10)
Lagged Flow	0.270***
	(0.02)
Log(TNA)	-0.001
	(0.00)
Log(age)	-0.016***
	(0.00)
N	1436
Adj. R^2	0.170

$^{***}p < 0.01, \, ^{**}p < 0.05, \, ^{*}p < 0.1$

Proof. Since X is compact, the sequence x_n has at least one limit point, and any limit point must be in X. Let $x_0 \in X$ be a limit point and let x_{n_k} be a subsequence through which $x_{n_k} \to x_0$. Then $(x_{n_k}, y_{n_k}) \to (x_0, y_0)$, and the continuity of f implies that

$$0 = \lim_{k \to \infty} f(x_{n_k}, y_{n_k}) = f(x_0, y_0).$$

Since $f(x, y)$ is strictly monotone in x for each y, x_0 is uniquely determined by y_0. Thus, x_n has just one limit point x_0, and we conclude that $x_n \to x_0$. $\qquad\square$

Proof of Proposition 1. Using first the expressions for n_i^{fm} and q_i^{fm} in (8) and then the expression for Δw_i^{fm} in (4), the ratios $U_i^{fm} = q_i^{fm}/n_i^{fm}$ in (9) become

$$U_i^{fm} = \frac{q_i^0 - \Delta q_i^{fm}}{n_i^0 - \Delta w_i^{fm}} = \frac{q_i^0 - \Delta q_i^{fm}}{n_i^0 + a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^*}. \tag{B.1}$$

The denominator is strictly positive because $n_i^0 = a_i$, $b_i < 1$ by hypothesis, $\pi_i \leq 1$, and $\Delta s_i^* \in [-1, 0]$. Substituting (B.1) into (10) and also substituting the expression for Δp_i^{fm} in (5) into (10), we find that the NAV change of fund i due to the exogenous shock and first movers' redemptions is given by

$$\Delta s_i^{fm} = \frac{n_i^0 - \Delta q_i^{fm}}{n_i^0 + a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^*} \cdot \left(p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \right) - s_i^0, \tag{B.2}$$

for $i = 1, \ldots, N$. We can similarly use (4) and (5) to write (6) as

$$\Delta q_i^{fm} \cdot \left(p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \right) = \Delta w_i^{fm} \cdot (s_i^0 + \Delta s_i^0)$$

$$= -a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^* \cdot (s_i^0 + \Delta s_i^0). \tag{B.3}$$

We will use (B.2) and (B.3) in showing the existence of a fixed point of the mapping $\Delta S^* \mapsto \Delta S^{\infty}(\Delta S^*)$ defined by (11). We will analyze the mapping from $(\Delta s_1^*, \ldots, \Delta s_N^*)$ to $(\Delta q_1^{fm}, \ldots, \Delta q_N^{fm})$ implicitly defined by (B.3), and then use that mapping in the mapping from $(\Delta s_1^*, \ldots, \Delta s_N^*)$ to $(\Delta s_1^{fm}, \ldots, \Delta s_N^{fm})$ defined by (B.2).

We will apply Brouwer's fixed point theorem to show the existence of a fixed point of the mapping $\Delta S^* \mapsto \Delta S^{\infty}(\Delta S^*)$ defined by (11). This boils down to proving the following two statements:

(i) The function in (11) is continuous w.r.t. the input $\Delta S^* \in [-1, 0]^N$.

(ii) For each input $\Delta S^* \in [-1, 0]^N$, the output of the function in (11) is also in $[-1, 0]^N$.

We next state and prove the following claim:

Claim: For sufficiently small $l_1, \ldots, l_K$, there exists a continuous mapping

$$\Phi : \ [-1,0]^N \ \to \ [0, n_1^0] \times \cdots \times [0, n_N^0]$$
$$\Phi(\Delta s_1^*, \Delta s_2^*, \ldots, \Delta s_N^*) = (\Delta q_1^{fm}, \Delta q_2^{fm}, \ldots, \Delta q_N^{fm})^\top,$$

such that $(\Delta q_1^{fm}, \Delta q_2^{fm}, \ldots, \Delta q_N^{fm})^\top$ solves (B.3).

Proof of the Claim. The system (B.3) can be regarded as a system of N quadratic equations, which can be solved sequentially.

Fix $\Delta s_1^* \in [-1, 0]$ and $(\Delta q_2, \ldots, \Delta q_N) \in [0, n_2^0] \times \cdots \times [0, n_N^0]$. Then, for $i = 1$, equation (B.3) is a quadratic equation in the variable Δq_1^{fm}. One solution of this equation is given by

$$\Delta q_1^{fm} = \Delta q_1^{fm}(\Delta s_1^*, \Delta q_2, \ldots, \Delta q_N) = \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\alpha_1 \gamma_1}}{2\alpha_1}, \tag{B.4}$$

where

$$\alpha_1 = -\sum_{k=1}^{K} m_{1k} l_k m_{1k}, \quad \beta_1 = \Delta p_1^0 + p_1^0 - \sum_{k=1}^{K} \sum_{j=2}^{N} m_{1k} l_k m_{jk} \Delta q_j,$$

and

$$\gamma_1 = a_1 \cdot \pi_1 \cdot b_1 \cdot \Delta s_1^* \cdot (s_1^0 + \Delta s_1^0).$$

Notice that, for sufficiently small $(l_k)_{k=1}^{K}$, the quantity $\beta_1^2 - 4\alpha_1 \gamma_1$ is strictly positive, and hence the function $\Delta q_1^{fm}(\cdot)$ as defined in (B.4) takes only real values. Moreover, it is positive because both the numerator and denominator of $\Delta q_1^{fm}(\cdot)$ are negative quantities. We claim that $\Delta q_1^{fm} \in [0, n_1^0]$. To see why, evaluate (B.3) at the endpoints of this interval. At $\Delta q_1^{fm} = 0$, the left side of (B.3) is zero and thus less than or equal to the right side of (B.3), which is nonnegative because $\Delta s_i^* \leq 0$ and $\Delta s_i^0 > -s_i^0$. At $\Delta q_1^{fm} = n_1^0$, the left side of (B.3) satisfies

$$n_1^0 \cdot \left(\Delta p_1^0 + p_1^0 - \sum_{k=1}^{K} m_{1k} l_k m_{1k} n_1^0 - \sum_{k=1}^{K} \sum_{j=2}^{N} m_{1k} l_k m_{jk} \Delta q_j \right)$$
$$\geq n_1^0 \cdot \left(\Delta p_1^0 + p_1^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{1k} l_k m_{jk} n_j^0 \right)$$
$$\geq a_1 \cdot \pi_1 \cdot b_1 \cdot (s_1^0 + \Delta s_1^0)$$
$$\geq -a_1 \cdot \pi_1 \cdot b_1 \cdot \Delta s_1^* \cdot (s_1^0 + \Delta s_1^0),$$

where the first inequality holds because $\Delta q_j \leq n_j^0$. The second inequality holds for sufficiently small $(l_k)_k$'s because $n_1^0 = a_1$, $\pi_1 \leq 1, b_1 < 1$. The last inequality follows from $\Delta s_1^* \in [-1, 0]$. Hence, by the intermediate value theorem, one of the two roots of (B.3) belongs to the interval $[0, n_1^0]$. Because the root $\frac{-\beta_1 - \sqrt{\beta_1^2 - 4\alpha_1 \gamma_1}}{2\alpha_1}$ can be arbitrarily large for sufficiently small $(l_k)_{k=1}^{K}$, we conclude

that $\Delta q_1^{fm}(\cdot)$ as defined in (B.4) takes values in $[0, n_1^0]$.

Next, we show that there exists a constant P_1, independent of l_1, ..., l_K, such that the following uniform bound holds:

$$P_1 \geq \sup_{\Delta s_1^* \in [-1,0], \Delta q_2 \in [0,n_2^0],...,\Delta q_N \in [0,n_N^0]} \max \left\{ \left| \frac{\partial \Delta q_1^{fm}}{\partial \Delta s_1^*} \right|, \left| \frac{\partial \Delta q_1^{fm}}{\partial \Delta q_2} \right|, ..., \left| \frac{\partial \Delta q_1^{fm}}{\partial \Delta q_N} \right| \right\}. \tag{B.5}$$

To see this, set $i = 1$ in (B.3), and rewrite the corresponding equation by treating Δq_1^{fm} as a function of $(\Delta s_1^*, \Delta q_2, \ldots, \Delta q_N)$. This yields

$$\Delta q_1^{fm}(\Delta s_1^*, \Delta q_2, \ldots, \Delta q_N) \cdot \left(\Delta p_1^0 + p_1^0 - \sum_{k=1}^K m_{1k} l_k m_{1k} \Delta q_1^{fm}(\Delta s_1^*, \Delta q_2, \ldots, \Delta q_N) \right.$$
$$\left. - \sum_{k=1}^K \sum_{j=2}^N m_{1k} l_k m_{jk} \Delta q_j \right) = -a_1 \cdot \pi_1 \cdot b_1 \cdot \Delta s_1^* \cdot (s_1^0 + \Delta s_1^0).$$

Differentiating the expression above with respect to Δq_2 on both sides leads to

$$\frac{\partial \Delta q_1^{fm}}{\partial \Delta q_2} \cdot \left(\Delta p_1^0 + p_1^0 - 2 \sum_{k=1}^K m_{1k} l_k m_{1k} \Delta q_1^{fm} - \sum_{k=1}^K \sum_{j=2}^N m_{1k} l_k m_{jk} \Delta q_j \right) - \Delta q_1^{fm} \sum_{k=1}^K m_{1k} l_k m_{2k} = 0.$$

Because we have previously shown that Δq_1^{fm} takes values in $[0, n_1^0]$, the equality above implies that $\frac{\partial \Delta q_1^{fm}}{\partial \Delta q_2}$ is uniformly bounded for sufficiently small l_1, ..., l_K as assumed in this proposition. The other derivatives appearing on the right-hand side of (B.5) can be estimated similarly, and we can thus conclude the existence of a uniform bound P_1 in (B.5).

Next, set $i = 2$ in (B.3). We want to show that, for any $(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N) \in [-1, 0]^2 \times [0, n_3^0] \times \cdots \times [0, n_N^0]$, there exists a function $\Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N)$ in the interval $[0, n_2^0]$. Rewriting equation (B.3) for $i = 2$, we get

$$\Delta q_2^{fm} \cdot \left(\Delta p_2^0 + p_2^0 - \sum_{k=1}^K m_{2k} l_k m_{1k} \Delta q_1^{fm}(\Delta s_1^*, \Delta q_2^{fm}, \Delta q_3, \ldots, \Delta q_N) - \sum_{k=1}^K m_{2k} l_k m_{2k} \Delta q_2^{fm} \right.$$
$$\left. - \sum_{k=1}^K \sum_{j=3}^N m_{2k} l_k m_{jk} \Delta q_j \right) = -a_2 \cdot \pi_2 \cdot b_2 \cdot \Delta s_2^* \cdot (s_2^0 + \Delta s_2^0). \tag{B.6}$$

We will show that $\Delta q_2^{fm} \in [0, n_2^0]$ by considering the values at the endpoints of this interval. If we set $\Delta q_2^{fm} = 0$ on the left side of (B.6), then the left side evaluates to zero, and it is less than or equal to the right side which is nonnegative. If we set $\Delta q_2^{fm} = n_2^0$ on the left side, then using

29

similar arguments as for the case $\Delta q_1^{fm} = n_1^0$, we obtain

$$n_2^0 \cdot \left(\Delta p_2^0 + p_2^0 - \sum_{k=1}^{K} m_{2k} l_k m_{1k} \Delta q_1^{fm}(\Delta s_1^*, n_2^0, \Delta q_3, \ldots, \Delta q_N) - \sum_{k=1}^{K} m_{2k} l_k m_{2k} n_2^0 \right.$$

$$\left. - \sum_{k=1}^{K} \sum_{j=3}^{N} m_{2k} l_k m_{jk} \Delta q_j \right)$$

$$\geq n_2^0 \cdot \left(\Delta p_2^0 + p_2^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{2k} l_k m_{jk} n_j^0 \right)$$

$$\geq a_2 \cdot \pi_2 \cdot b_2 \cdot (s_2^0 + \Delta s_2^0)$$

$$\geq -a_2 \cdot \pi_2 \cdot b_2 \cdot \Delta s_2^* \cdot (s_2^0 + \Delta s_2^0),$$

where in the first inequality we have used the previously established fact that $\Delta q_1^{fm} \in [0, n_1^0]$, and for the second inequality we have used that $a_2 = n_2^0$, $\pi_2 \leq 1$, and $b_2 < 1$.

Next, we show that the left side of equation (B.3) is an increasing function of Δq_2^{fm} by showing that its derivative with respect to Δq_2^{fm} is positive. Using the chain rule of differentiation, we find that the derivative of the left side of (B.3) with respect to Δq_2^{fm} is given by

$$\Delta p_2^0 + p_2^0 - \sum_{k=1}^{K} \sum_{j=1}^{2} m_{2k} l_k m_{jk} \Delta q_j^{fm} - \sum_{k=1}^{K} \sum_{j=3}^{N} m_{2k} l_k m_{jk} \Delta q_j$$

$$- \Delta q_2^{fm} \cdot \left(\sum_{k=1}^{K} m_{2k} l_k m_{1k} \frac{\partial \Delta q_1^{fm}}{\partial \Delta q_2^{fm}} + \sum_{k=1}^{K} m_{2k} l_k m_{2k} \right)$$

$$\geq \left(\Delta p_2^0 + p_2^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{2k} l_k m_{jk} n_j^0 \right) - n_2^0 \cdot \left(\sum_{k=1}^{K} m_{2k} l_k m_{1k} P_1 + \sum_{k=1}^{K} m_{2k} l_k m_{2k} \right) > 0,$$

where the first inequality holds because we have shown that the functions Δq_1^{fm} and Δq_2^{fm} satisfy $\Delta q_1^{fm} \in [0, n_1^0]$ and $\Delta q_2^{fm} \in [0, n_2^0]$, and because each input variable Δq_j is in the interval $[0, n_j^0]$, for $j \geq 3$. The last inequality holds for sufficiently small $(l_k)_k$. Since we have shown that the left side of (B.3) is strictly increasing in Δq_2^{fm} (with Δs_1^* fixed), it follows that (B.3) defines a unique implicit function $\Delta q_2^{fm} = \Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N) \in [0, n_2^0]$.

The continuity of the function $\Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N)$ follows by applying Lemma B.1 with $x = \Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N)$, $y = (\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N)$, and $f(x, y)$ as the difference between the left and right sides of (B.6). By replacing the input Δq_2 of Δq_1^{fm} with the function Δq_2^{fm}, we can write

$$\Delta q_1 = \Delta q_1^{fm}(\Delta s_1^*, \Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N), \Delta q_3, \ldots, \Delta q_N),$$

$$\Delta q_2 = \Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N).$$

Then equation (B.3) holds for $i = 1, 2$ simultaneously. Using the boundedness of Δq_1^{fm} and Δq_2^{fm}

30

together with the bound (B.5) (which is needed to bound the derivatives appearing in the chain rule of differentiation), we can show the existence of a uniform bound P_2 such that

$$P_2 \geq \sup_{\Delta s_1^*, \Delta s_2^* \in [-1,0], \Delta q_3 \in [0, n_2^0], \ldots, \Delta q_N \in [0, n_N^0]} \max \left\{ \left| \frac{\partial \Delta q_2^{fm}}{\partial \Delta s_1^*} \right|, \left| \frac{\partial \Delta q_2^{fm}}{\partial \Delta s_2^*} \right|, \left| \frac{\partial \Delta q_2^{fm}}{\partial \Delta q_3} \right|, \ldots, \left| \frac{\partial \Delta q_2^{fm}}{\partial \Delta q_N} \right| \right\},$$

(B.7)

using a similar method to that used to show the existence of P_1 in (B.5), and under the assumption that $l_1, \ldots, l_K$ are sufficiently small.

Repeating the steps above for $i = 3, \ldots, N$, and again assuming $l_1, \ldots, l_K$ are sufficiently small, we obtain N continuous functions

$$\Delta q_1^{fm}(\Delta s_1^*, \Delta q_2, \ldots, \Delta q_N) : [-1, 0] \times [0, n_2^0] \times \cdots \times [0, n_N^0] \to [0, n_1^0];$$
$$\Delta q_2^{fm}(\Delta s_1^*, \Delta s_2^*, \Delta q_3, \ldots, \Delta q_N) : [-1, 0]^2 \times [0, n_3^0] \times \cdots \times [0, n_N^0] \to [0, n_2^0];$$
$$\vdots$$
$$\Delta q_{N-1}^{fm}(\Delta s_1^*, \Delta s_2^*, \ldots, \Delta s_{N-1}^*, \Delta q_N) : [-1, 0]^{N-1} \times [0, n_N^0] \to [0, n_{N-1}^0];$$
$$\Delta q_N^{fm}(\Delta s_1^*, \Delta s_2^*, \ldots, \Delta s_N^*) : [-1, 0]^N \to [0, n_N^0].$$

If we replace each input variable Δq_j with the function Δq_j^{fm}, all of these functions become functions of $\Delta s_1^*, \ldots, \Delta s_N^*$. We have thus constructed a function $\Phi : [-1, 0]^N \to [0, n_1^0] \times \cdots \times [0, n_N^0]$ for which

$$\Phi(\Delta s_1^*, \Delta s_2^*, \ldots, \Delta s_N^*) = (\Delta q_1^{fm}, \Delta q_2^{fm}, \ldots, \Delta q_N^{fm})^\top$$

solves (B.3). This ends the proof of the claim. $\square$

Next, we show that for each input $\Delta S^* \in [-1, 0]^N$, the corresponding output given by (11) is still in $[-1, 0]^N$. Towards this goal, we will bound U^{fm}, and we begin by showing that

$$-a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^* \leq \Delta q_i^{fm}.$$

(B.8)

Since $\Delta q_j^{fm} \leq n_j^0$, we have

$$p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \geq p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta n_j^0 > 0,$$

where the last inequality holds for sufficiently small $(l_k)_k$. Hence, if $\Delta s_i^* = 0$, then $\Delta q_i^{fm} = 0$ and

(B.8) is satisfied. Assume now that $\Delta s_i^* < 0$. Then (B.3) yields

$$-a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^* \cdot (s_i^0 + \Delta s_i^0) = \Delta q_i^{fm} \cdot \left(p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \right)$$

$$\leq \Delta q_i^{fm} \cdot \left(p_i^0 + \Delta p_i^0 \right).$$

Hence, using $s_i^0 + \Delta s_i^0 = p_i^0 + \Delta p_i^0 > 0$ we find that (B.8) is again satisfied.

Applying (B.8) in (B.1) shows that $U_i^{fm} \leq 1$. In view of the definition in (9), we have that $U_i^{fm} \geq 0$. If $U_i^{fm} = 0$, it follows from (B.2) that $\Delta s_i^{fm} = -s_i^0 < 0$. If $0 < U_i^{fm} \leq 1$, then combining (B.1) and (B.2) yields

$$\Delta s_i^{fm} = \frac{n_i^0 - \Delta q_i^{fm}}{n_i^0 + a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^*} \cdot \left(p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \right) - s_i^0 \qquad \text{(B.9)}$$

$$\leq p_i^0 + \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} - s_i^0$$

$$= \Delta p_i^0 - \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} \Delta q_j^{fm} \leq 0,$$

where the last inequality follows from the fact that $\Delta p_i^0 < 0$ and $\Delta q_j^{fm} \in [0, n_j^0]$. Hence,

$$\|\Delta S^{fm}\|_\infty \leq \max_{i=1,\dots,K} \left\{ -\Delta p_i^0 + \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} n_j^0 \right\}.$$

Since $\|U^{fm}\|_\infty \leq 1$,

$$\|U^{fm} M L M^\top U^{fm} A(I - \Pi) B\|_\infty \leq \|M M^\top A(I - \Pi) B\|_\infty \cdot \max_k l_k.$$

Then (11) yields

$$\left\| \sum_{n=0}^{\infty} \Delta S^{n,sm} \right\|_\infty \leq \frac{\|\Delta S^{fm}\|_\infty}{1 - \|U^{fm} M L M^\top U^{fm} A(I - \Pi) B\|_\infty} \qquad \text{(B.10)}$$

$$\leq \max_{i=1,\dots,K} \left\{ -\Delta p_i^0 + \sum_{k=1}^{K} \sum_{j=1}^{N} m_{ik} l_k m_{jk} n_j^0 \right\} \Big/ \left(1 - \|M M^\top A(I - \Pi) B\|_\infty \cdot \max_k l_k \right)$$

$$\leq 1,$$

where the last inequality holds if $\max_k l_k$ is sufficiently small.

We know from (B.9) that $\Delta s_i^{fm} \le 0$. Moreover, for each $n \ge 1$,

$$\left(U^{fm} M L M^\top U^{fm} A(I - \Pi) B \right)^n \tag{B.11}$$

is the product of matrices with nonnegative entries and therefore has nonnegative entries. Hence,

$$\Delta S^{n,sm} = \left(U^{fm} M L M^\top U^{fm} A(I - \Pi) B \right)^n \Delta S^{fm} \in (-\infty, 0]^N,$$

and thus $\sum_{n=0}^\infty \Delta S^{n,sm} \in [-\infty, 0]^N$. Together with (B.10), we obtain that $\sum_{n=0}^\infty \Delta S^{n,sm} \in [-1, 0]^N$.

To establish the existence of a fixed point of (11) using Brouwer's fixed point theorem, it remains to show that the output in (11) is continuous in ΔS^*. In the Claim above, we established $(\Delta q_1^{fm}, \Delta q_2^{fm}, \dots, \Delta q_N^{fm})$ is a continuous function of ΔS^*. Let $\Delta S^*_{(k)} \in [-1, 0]^N$ be such that

$$\lim_{k \to +\infty} \Delta S^*_{(k)} = \Delta S^*.$$

For $\Delta S^*_{(k)}$, $k = 1, 2, \dots$, denote by $U_{(k)}^{fm}$, $\Delta S_{(k)}^{fm}$ the corresponding terms in (11). Because of the continuous dependence of $(\Delta q_1^{fm}, \Delta q_2^{fm}, \dots, \Delta q_N^{fm})$ on ΔS^*, it holds that

$$\lim_{k \to +\infty} U_{(k)}^{fm} = U^{fm}, \quad \lim_{k \to +\infty} \Delta S_{(k)}^{fm} = \Delta S^{fm}.$$

To see why $\lim_{k \to +\infty} U_{(k)}^{fm} = U^{fm}$, recall that $U_{(k),i}^{fm} = \frac{q_{(k),i}^{fm}}{n_{(k),i}^{fm}}$. Moreover, observe that

$$\lim_{k \to +\infty} n_{(k),i}^{fm} = n_i^0 + a_i \cdot \pi_i \cdot b_i \cdot \lim_{k \to +\infty} \Delta s^*_{(k),i} = n_i^0 + a_i \cdot \pi_i \cdot b_i \cdot \Delta s_i^* > 0.$$

Because the output $(\Delta q_1^{fm}, \Delta q_2^{fm}, \dots, \Delta q_N^{fm})$ depends continuously on the input ΔS^*, we have

$$\lim_{k \to +\infty} q_{(k),i}^{fm} = q_i^0 - \lim_{k \to +\infty} \Delta q_{(k),i}^{fm} = q_i^0 - \Delta q_i^{fm} = q_i^{fm}.$$

By combining the two limits above, we confirm that $\lim_{k \to +\infty} U_{(k)}^{fm} = U^{fm}$.

Notice also that

$$\left\| \left(U_{(k)}^{fm} M L M^\top U_{(k)}^{fm} A(I - \Pi) B \right)^n \right\|_\infty \le \left(\| M M^\top A(I - \Pi) B \|_\infty \cdot \max_k l_k \right)^n$$

Therefore the dominated convergence theorem implies that

$$\lim_{k \to +\infty} \sum_{n=0}^{+\infty} \left(U_{(k)}^{fm} M L M^\top U_{(k)}^{fm} A(I - \Pi) B \right)^n \Delta S_{(k)}^{fm} = \sum_{n=0}^{+\infty} \left(U^{fm} M L M^\top U^{fm} A(I - \Pi) B \right)^n \Delta S^{fm},$$

which shows the continuity of the output of (11) with respect to the input ΔS^*.

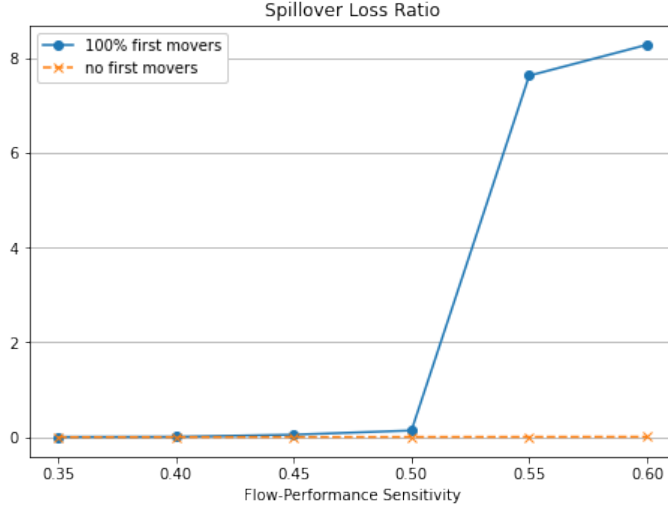In sum, the mapping (11) is continuous, and it maps $[-1, 0]^N$ into itself. According to Brouwer's

Figure C.1: The figure shows the Spillover Loss Ratio for different values of flow-to-performance ratio when funds follow a pecking order liquidation strategy both in the absence of first movers (dashed line) and if all investors are first movers (solid line). For each asset, we apply a shock equal to -5% price change times its relative volatility. We use asset holdings data from Q1 2020.

theorem, it follows that the mapping has a fixed point.   □

C   Spillover Losses and Pecking Order of Liquidation

In this section, we quantify the dependence of aggregate vulnerability to flow-to-performance sensitivity and first movers under a different asset liquidation strategy followed by funds. Specifically, we assume that funds follow a pecking order of liquidation, meaning that they sequentially liquidate assets in increasing order of price impact parameters. First, funds use cash, then they liquidate government bonds, and then sequentially the other assets. We assume that the assets labeled as "Other Securities," "Other Equities," and "Other Fixed-Income Securities" are the last ones to be liquidated because we do not have granular information on those assets.[16]

If all funds follow the pecking order liquidation strategy and the flow-to-performance sensitivity is low, aggregate spillover losses are significantly lower compared to the case of proportional liquidation. This is due to two compounding effects. First, the use of cash and the sale of liquid assets to repay redeeming investors reduces the downward impact on asset prices caused by redemptions and asset liquidation. Second, fire sales are concentrated in fewer (and more liquid) assets, which reduces asset price contagion. However, if the sensitivity of flow to performance is large, the first-mover advantage substantially increases the aggregate vulnerability of the mutual fund system. In Figure C.1, we compare the SLR in the two polar cases with no first movers, or all first-mover investors. Spillover losses are orders of magnitude larger in the system with only first

---

[16]For computational reasons, we aggregate funds within each of the nine types. Unlike the proportion liquidation strategy, we cannot use matrix algebra to compute the quantities needed to estimate the SLR.

movers compared to the system without first movers. The reason is that all funds first liquidate the same assets, severely impacting their prices and precipitating the spiral of redemptions and fire sales. This scenario is reminiscent of the disruption of Treasury markets during the Covid-19 crisis: as discussed in Ma et al. (2022), concentrated sales of their most liquid assets by fixed-income mutual funds led to a significant increase in Treasury yields.

D Analytical Results on First Mover Concentration

Consider a system with two funds holding identical portfolios. Let $\frac{\bar{\pi}}{2}$ be the proportion of first movers in the whole system, and let $\pi \in \left(\frac{\bar{\pi}}{2}, \bar{\pi}\right)$ be the proportion of first movers in the first fund. The proportion of first movers in the second fund is $\bar{\pi} - \pi$. Let $\ell := \sum_{k=1}^{K} m_{ik}^2 l_k$. Both funds are subject to an initial identical portfolio shock Δs^0. Let $\Delta Q^{fm} = x = (x_1, x_2)$ be the number of portfolio units each fund sells to repay first movers, and let $\Delta S^* = y = (y_1, y_2)$ be the aggregate shock to each fund's NAV. The amounts x_1 and x_2 are the solutions to the system

$$x_1(s^0 + \Delta s^0 - \ell(x_1 + x_2)) = -ab(s^0 + \Delta s^0)\pi y_1,$$
$$x_2(s^0 + \Delta s^0 - \ell(x_1 + x_2)) = -ab(s^0 + \Delta s^0)(\bar{\pi} - \pi)y_2,$$

where we assume without loss of generality that $p_0 = s_0$. More explicitly,

$$x_1 = f_{x,1}(y_1, y_2) := \pi y_1 \frac{s^0 + \Delta s^0 - \sqrt{(s^0 + \Delta s^0)(s^0 + \Delta s^0 + 4ab\ell(\pi y_1 + (\bar{\pi} - \pi)y_2))}}{2\ell(\pi y_1 + (\bar{\pi} - \pi)y_2)},$$
$$x_2 = f_{x,2}(y_1, y_2) := (\bar{\pi} - \pi)y_2 \frac{s^0 + \Delta s^0 - \sqrt{(s^0 + \Delta s^0)(s^0 + \Delta s^0 + 4ab\ell(\pi y_1 + (\bar{\pi} - \pi)y_2))}}{2\ell(\pi y_1 + (\bar{\pi} - \pi)y_2)},$$

where we have chosen the smallest roots, i.e., the ones corresponding to the least amount of assets funds would have to liquidate to meet first movers' redemptions. The NAV change of fund 1 observed by second movers is

$$\Delta s_1^{fm} = \frac{a - x_1}{a + ab\pi y_1}(s^0 + \Delta s^0 - \ell(x_1 + x_2)) - s^0.$$

We may rewrite the above expression, and obtain that each fund's NAV change observed by second movers is equal to

$$\Delta S^{fm} = (\Delta s^0 - \ell(x_1 + x_2))\begin{pmatrix} 1 \\ 1 \end{pmatrix} - (s^0 + \Delta s^0 - \ell(x_1 + x_2))\begin{pmatrix} \frac{x_1 + ab\pi y_1}{a + ab\pi y_1} \\ \frac{x_2 + ab(\bar{\pi} - \pi)y_2}{a + ab(\bar{\pi} - \pi)y_2} \end{pmatrix}.$$

Define $v_1 := \frac{a - x_1}{a + ab\pi y_1}$, $v_2 := \frac{a - x_2}{a + ab(\bar{\pi} - \pi)y_2}$. The matrix U^{fm} defined in Section 3.2 is then given by

$$U^{fm} = \begin{pmatrix} v_1 & 0 \\ 0 & v_2 \end{pmatrix},$$

and in each round of second movers' redemptions the NAV change is multiplied by the matrix $T := U^{fm} M L M^\top U^{fm} A(I - \Pi)B$. An explicit calculation yields

$$T = ab\ell \begin{pmatrix} (1-\pi)v_1^2 & (1-(\bar{\pi}-\pi))v_1 v_2 \\ (1-\pi)v_1 v_2 & (1-(\bar{\pi}-\pi))v_2^2 \end{pmatrix}.$$

For sufficiently small ℓ, the matrix $I - T$ is invertible and the aggregate impact on each fund's NAV is then given by

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_{y,1}(x,y) \\ f_{y,2}(x,y) \end{pmatrix} := \sum_{n=0}^{\infty} T^n \Delta S^{fm} = (I-T)^{-1} \Delta S^{fm}. \tag{D.12}$$

Therefore, to find the aggregate NAV change we need to solve the fixed point of the system $x = f_x(y)$ and $y = f_y(x,y)$. The component $y = (y_1, y_2)$ gives the aggregate NAV changes for funds 1 and 2, respectively.

Next, we restate Proposition 2 using the notation introduced in this section. In particular, the fixed point y^* plays the role of ΔS^* in Proposition 2. We make the dependence of the functions f_x and f_y on π explicit by writing f_x^π and f_y^π.

Proposition 2'. *Assume $\Delta s^0 \in (-1, 0)$ and $b \cdot \Delta s^0 \in (-1, 0)$. For sufficiently small ℓ, there exists a fixed point $y^*(\pi)$ for $f_y^\pi(f_x^\pi(y), y)$, where $\pi \in (\frac{\bar{\pi}}{2}, \bar{\pi})$ is the proportion of first movers in the first fund. Define $g(\pi) := \lim_{\ell \downarrow 0} \frac{1}{\ell}(y_1^*(\pi) + y_2^*(\pi) - 2\Delta s^0)$, i.e., $y_1^*(\pi) + y_2^*(\pi) = 2\Delta s^0 + \ell \cdot g(\pi) + o(\ell)$. The function $g(\pi)$ is decreasing in π.*

Proof. Fix $\pi \in (\frac{\bar{\pi}}{2}, \bar{\pi})$ and let $y^\pi := y^*(\pi)$ be the vector of aggregate NAV changes if the proportion of first movers for each fund is, respectively, π and $\bar{\pi} - \pi$. For $\ell = 0$, asset liquidation does not move prices, and therefore $y^\pi = (\Delta s^0, \Delta s^0)$. By continuity, a fixed point y^π exists for sufficiently small ℓ. Assume $y^\pi = (\Delta s^0, \Delta s^0) + \ell \cdot y^{\pi,1} + o(\ell^2)$, where $y^{\pi,1}$ is independent of ℓ. The first order expansion of $(x_1, x_2) = f_x^\pi(y^\pi)$ yields

$$x_1 = -ab\pi\Delta s^0 + \ell ab\pi \frac{ab\bar{\pi}(\Delta s^0)^2 - y_1^{\pi,1}(s^0 + \Delta s^0)}{s^0 + \Delta s^0} + o(\ell),$$

$$x_2 = -ab(\bar{\pi} - \pi)\Delta s^0 + \ell ab(\bar{\pi} - \pi) \frac{ab\bar{\pi}(\Delta s^0)^2 - y_2^{\pi,1}(s^0 + \Delta s^0)}{s^0 + \Delta s^0} + o(\ell).$$

After plugging the expansion for (x_1, x_2) into the right-hand side in equation (D.12), we obtain

$$(I-T)^{-1}\Delta S^{fm} = \Delta s^0 + \ell \cdot \begin{pmatrix} \frac{ab\Delta s^0(2 + b\pi\Delta s^0(2-\bar{\pi}))}{1 + b\pi\Delta s^0} \\ \frac{ab\Delta s^0(2 + b(\bar{\pi}-\pi)\Delta s^0(2-\bar{\pi}))}{1 + b(\bar{\pi}-\pi)\Delta s^0} \end{pmatrix} + o(\ell).$$

Hence, by comparing the terms of order ℓ in equation (D.12), we get

$$y_1^{\pi,1} = ab\Delta s^0 \frac{2 + b\pi\Delta s^0(2 - \bar{\pi})}{1 + b\pi\Delta s^0},$$

$$y_2^{\pi,1} = ab\Delta s^0 \frac{2 + b(\bar{\pi} - \pi)\Delta s^0(2 - \bar{\pi})}{1 + b(\bar{\pi} - \pi)\Delta s^0}.$$

In particular, $y_1^\pi + y_2^\pi = 2\Delta s^0 + \ell \cdot g(\pi) + o(\ell)$, where

$$g(\pi) = ab\Delta s^0 \left[\frac{2 + b\pi\Delta s^0(2 - \bar{\pi})}{1 + b\pi\Delta s^0} + \frac{2 + b(\bar{\pi} - \pi)\Delta s^0(2 - \bar{\pi})}{1 + b(\bar{\pi} - \pi)\Delta s^0} \right].$$

The first derivative of $g(\pi)$ is

$$ab^3(\Delta s^0)^3\bar{\pi} \frac{(2\pi - \bar{\pi})(2 + b\bar{\pi}\Delta s^0)}{(1 + b\pi\Delta s^0)^2(1 + b(\bar{\pi} - \pi)\Delta s^0)^2},$$

which is negative because $2\pi > \bar{\pi}$, $\Delta s^0 < 0$, and $b \cdot \Delta s^0 > -1$. This concludes the proof. $\qquad\square$

An increase in first mover concentration has a twofold effect on each fund's NAV. First, additional first-mover redemptions at the first fund negatively impact its NAV (and, conversely, fewer first mover redemptions at the second fund increase its NAV). This effect is symmetric across the two funds. Second, an increase in the proportion of first movers reduces the number of investors that bear the cost of first movers' redemptions, while this externality is spread over more investors for the fund with fewer first movers. This effect is asymmetric, because it exacerbates the impact of first movers on the first fund's NAV and reduces the benefit of having fewer first movers. Hence, the aggregate effect of first mover concentration on the system is negative.

References

Allen, F., Babus, A. and Carletti, E. (2009), 'Financial crises: theory and evidence', *Annual Review of Financial Economics* **1**, 97—-116.

Bouveret, A. and Yu, J. (2021), 'Risks and vulnerabilities in the US bond mutual fund industry', *IMF Working Papers* **2021**(109).

Capponi, A., Glasserman, P. and Weber, M. (2020), 'Swing pricing for mutual funds: Breaking the feedback loop between fire sales and fund redemptions', *Management Science* **66**(8), 3581–3602.

Capponi, A. and Larsson, M. (2015), 'Price contagion through balance sheet linkages', *Review of Asset Pricing Studies* **5**(2), 227–253.

Cetorelli, N., Duarte, F. and Eisenbach, T. (2016), 'Are asset managers vulnerable to fire sales?', *Liberty Street Economics* .

Chen, Q., Goldstein, I. and Jiang, W. (2010), 'Payoff complementarities and financial fragility: Evidence from mutual fund outflows', *Journal of Financial Economics* **97**(2), 239–262.

Chernenko, S. and Sunderam, A. (2020), Measuring the perceived liquidity of the corporate bond market, Technical report, National Bureau of Economic Research.

Choi, J., Hoseinzade, S., Shin, S. S. and Tehranian, H. (2020), 'Corporate bond mutual funds and asset fire sales', *Journal of Financial Economics* **138**(2), 432–457.

Duarte, F. and Eisenbach, T. M. (2021), 'Fire-sale spillovers and systemic risk', *Journal of Finance* **76**(3), 1251–1294.

Franzoni, F. and Schmalz, M. C. (2017), 'Fund flows and market states', *The Review of Financial Studies* **30**(8), 2621–2673.

Fricke, C. and Fricke, D. (2021), 'Vulnerable asset management? The case of mutual funds', *Journal of Financial Stability* **52**, 100800.

Gennaioli, N., Shleifer, A. and Vishny, R. (2013), 'A model of shadow banking', *Journal of Finance* **68**, 1331—-1363.

Goldstein, I., Jiang, H. and Ng, D. T. (2017), 'Investor flows and fragility in corporate bond funds', *Journal of Financial Economics* **126**(3), 592–613.

Greenwood, R., Landier, A. and Thesmar, D. (2015), 'Vulnerable banks', *Journal of Financial Economics* **115**(3), 471–485.

Jiang, H., Li, D. and Wang, A. (2021), 'Dynamic liquidity management by corporate bond mutual funds', *Journal of Financial and Quantitative Analysis* **56**(5), 1622–1652.

Jiang, H., Li, Y., Sun, Z. and Wang, A. (2022), 'Does mutual fund illiquidity introduce fragility into asset prices? Evidence from the corporate bond market', *Journal of Financial Economics* **143**(1), 277–302.

Jin, D., Kacperczyk, M., Kahraman, B. and Suntheim, F. (2022), 'Swing pricing and fragility in open-end mutual funds', *The Review of Financial Studies* **35**(1), 1–50.

Johnson, W. T. (2004), 'Predictable investment horizons and wealth transfers among mutual fund shareholders', *The Journal of Finance* **59**(5), 1979–2012.

Li, Y., O'Hara, M., Zhou, X. et al. (2020), 'Mutual fund fragility, dealer liquidity provisions, and the pricing of municipal bonds', *Working paper* .

Ma, Y., Xiao, K. and Zeng, Y. (2022), 'Mutual fund liquidity transformation and reverse flight to liquidity', *The Review of Financial Studies* **35**, 4674–4711.

Schmidt, L., Timmermann, A. and Wermers, R. (2016), 'Runs on money market mutual funds', *American Economic Review* **106**(9), 2625–57.