# An ETF-based measure of stock price fragility

Renato Lazo-Paz *

July 15, 2023

**Abstract**

A growing literature employs equity mutual fund flows to measure stock's exposure to non-fundamental demand risk - stock price fragility. However, this approach may be biased by confounding fundamental information potentially leading to an underestimation of risk exposure. We propose an alternative estimation procedure that incorporates readily available primary market data from exchange-traded funds (ETFs). Our proposed procedure significantly enhances the predictive power of fragility in forecasting stock return volatility. Moreover, we find that our measure captures the influence of increase ETF activeness while partially capturing the effect of institutional investors' demand on price return volatility. Additionally, our analysis reveals a decrease in the explanatory power of mutual fund-based fragility. These results highlight the advantages of employing an ETF-based fragility measure that takes into account recent developments in the asset management industry, particularly the rise of *passive* investing.

**Keywords:** Non-fundamental demand risk, Fragility, Mutual funds, ETFs, volatility.
**JEL codes:** G12, G14, G23

*Renato Lazo-Paz is from Telfer School of Management, University of Ottawa, 55 Laurier Ave E, Ottawa, Ontario K1N 6N5, Canada. We thank Hamilton Galindo, Yinjie Shen, Wei Wang, Maksim Isakin, Fabio Moneta, Adelphe Ekponon, and seminar participants at Cleveland State University and The University of Ottawa for their valuable comments.

# 1 Introduction

Classic asset pricing theories state that stock prices fluctuate in response to fundamental shocks such as news. This argument is based on the assumption that trading unrelated to a firm's fundamentals (i.e., *non-fundamental demand shocks*) triggers a response by arbitrageurs who take the opposite side, canceling out any potential impact on security prices (e.g., Fama, 1965; Ross, 1976). However, a substantial body of empirical and theoretical literature demonstrates that trading driven by non-fundamental information can influence stock prices and that arbitrage activity faces various limitations that contribute to the potential persistence of mispricing[1] Thus, results of great interest to measure a stock's exposure to non-fundamental induced price variation.

But, how to measure the impact of non-fundamental shocks on stock prices? And how to measure a stock's exposure to such risk? The literature has addressed the first question by observing investor flows to mutual funds and their subsequent effect on stock prices (Coval and Stafford, 2007; Edmans et al., 2012; Dong et al., 2021). For the latter, an increasing number of studies[2] employ the stock price fragility of Greenwood and Thesmar (2011). This composite measure combines information on an asset's ownership structure and the owners non-fundamental driven trades to capture firm-level exposure to non-fundamental demand risk. To capture these liquidity-motivated trades, Greenwood and Thesmar (2011) also relies on mutual fund flows as proxy. Yet, despite the widely spread used of mutual fund flows as proxy of , there is a growing concern on their validity as instruments of non-fundamentally driven

[1]Seminal theoretical papers modeled the effect of noise traders (De Long et al., 1990), trading motivated by informational and noninformational motives (Wang, 1996) and the limits to arbitrage activity (Shleifer and Vishny, 1997) on stock prices and trading volume.

[2]In empirical corporate finance settings, studies have related stock price fragility to firm's financing costs (Francis et al., 2021), cash holdings and investment policies (Friberg et al., 2023). In the context of asset pricing factors, Huang et al. (2021) estimates the stock price fragility at the factor level to analyze the component of stock pricing factors returns driven by noise trading.

price pressure. Most importantly, the assumption of proportional trading and dumb money.

The focus of this article is to provide the literature with an alternative estimation of stock price fragility that potentially overcomes recent criticism while capturing a broader set of non-fundamental-driven sources of price variation. Motivated by the Exchange-traded Funds (ETF) trading model of Brown, Davies and Ringgenberg (2021), we propose an alternative estimation of Greenwood and Thesmar (2011) fragility that employs ETF primary market flows[3] and ownership composition. We argue that this approach does not suffer from several shortcomings of employing mutual fund data. Moreover, we relate our methodology to the theoretical framework of Ben-David et al. (2021) and show that using ETF data has the additional benefit of capturing the influence of both retail and institutional investor's demand effect on stock's volatility. We first construct the stock price fragility measure as in Greenwood and Thesmar (2011), $G^{MF}$, for the same sample period and extend it until the last quarter of 2018. We then proceed to reestimate the fragility measure employing ETFs primary market flows and ownership composition, $G^{ETF}$.

We highlight four main results. First, we find that the statistical and economic significance of $G^{MF}$ in explaining next quarter stock (excess) return volatility has significantly declined in the second part of our sample (2009-2018), which coincides with the *out-of-sample* period of Greenwood and Thesmar (2011) study. While we do not focus on studying the determinants of this decline, we observe that this behavior coincides with a period in which the equity mutual fund industry has experienced significant outflows, as shown in figure 1, and significant growth of the ETF industry

[3]An important distinction is between primary and secondary ETF trading markets. The primary market refers to the creation and redemption process between authorized participants (AP) and financial institutions. The secondary market refers to the intraday trading that occurs among investors that could be due to many different reasons. Madhavan (2014) and Ben-David et al. (2017) provide excellent reviews of the ETF industry.

in volume and trading participation (Dannhauser and Pontiff, 2019; Glosten et al., 2021; Easley et al., 2021). This finding is consistent with recent evidence showing that mutual fund flows might also include information about discretionary trades[4] (Huang et al., 2022; Berger, 2022) and time-varying specialized demand[5] (Rzeznik and Weber, 2022). Hence, raising significant concerns regarding the validity of the proportional trading assumption [6]

Second, we show that $G^{ETF}$ strongly predicts next quarter's stock (excess) return volatility. While the relationship is weak for the first part of our sample, 1989 - 2009, its economic and statistical significance increases in the second part, 2009 - 2018. This evidence is consistent with the rise of ETF trading volume (Ben-David et al., 2017) and participation in institutional and retail investors' portfolios (Dannhauser and Pontiff, 2019). Moreover, our findings are consistent with Brown, Davies and Ringgenberg (2021) model and empirical evidence of ETF primary flows as being indicatives of non-fundamental demand shocks.

Third, we show that the explanatory power of $G^{ETF}$ increased as the level of *ETF activeness* has risen in our sample. A potential concern is that our results are based on comparing two fundamentally distinct investment vehicles regarding their investment mandates. This is because equity mutual funds are actively managed while ETFs are, by construction, passive vehicles that aim at replicating a benchmark. We address this concern by estimating the activeness index of Easley

[4] *Discretionary* trades refer to those that contain fundamental information. This is, trades motivated by the fund managers' beliefs about stock mispricing that represent opportunities to generate alpha. Contrary to discretionary trades, *expected* trades assume that fund managers only expand (contract) their current portfolio in response to inflows(outflows).

[5] This refers to the demand from funds familiar with a specific set of assets that better allows them to price them adequately.

[6] For mutual fund flows to serve as a valid instrument for non-fundamental demand, it is essential that the information they convey remains independent and unrelated to any fundamental trading motive. This is possible if we assume that mutual funds trade (buy or sell) such that their initial allocation proportion does not change when faced with flows. This should be especially stronger when faced with extreme outflows or fire sales Coval and Stafford (2007); Edmans et al. (2012).

et al. (2021) for our sample of ETFs. We corroborate the authors' findings in a broader sample of ETFs and show that most ETFs are, in fact, *active invesments*[7]. Moreover, we decompose the $G^{ETF}$ into active and non-active components and find that our results are driven mainly by the active component. These results provides evidence supporting concerns that increased ETF industry activeness Easley et al. (2021) and that the spread of specialized ETFs may channel demand for overvalued stocks (Ben-David et al., 2023).

Fourth, we provide evidence that $G^{ETF}$ captures the effect of mid and small-sized institution ownership on stock price volatility. In a recent study, Ben-David et al. (2021) show that large institutional investors are less able to diversify away idiosyncratic shocks that affect their holdings due to the correlated behavior of their constitutient subunits (i.e., granularity). Thus, ownership by large institutional investors predicts higher stock price volatility. We follow the specification of Ben-David et al. (2021) and observe that $G^{ETF}$ remains significant when we include the effect of institutional ownership. Moreover, when including $G^{ETF}$ the explanatory power of mid and small institutional ownership disappear. We interpret this evidence as a result of the distinct ownership structure of the ETF industry. In contrast to the mutual fund industry, which is mostly owned by households (i.e., retail investors) ETFs are roughly equally owned by retail and institutional investors (Dannhauser and Pontiff, 2019). This fact can help explain why including $G^{ETF}$ subsumes the explanatory power of mid and small-sized institutions.

Overall, our results are consistent with the argument that ETF primary markets flows provide valid signals of non-fundamental demand shocks Brown, Davies and Ringgenberg (2021) and that not only retail-based ownership but institutional in-

[7] Easley et al. (2021) define ETF activeness as being in *form* or in *function*. A fund is active in *form* if it is designated to deliver out-performance or alpha. In *function* implies that a fund being passively or actively managed could serve as a building block of an active portfolio.

vestors ownership may contribute to stock return volatility (Koijen and Yogo, 2019; Bushee and Noe, 2000; Ben-David et al., 2021). Recent changes in the asset management industry (i.e., a rise of passive investing), greater accessibility to broader sets of data, and recent theoretical development and empirical evidence calls for a revision of the estimation of fragility, which we address in this paper.

This paper contributes to the empirical literature on non-fundamental demand and stock fragility in asset pricing. Furthermore, we contribute to the literature that studies the effect of ETF on the volatility (Ben-David et al., 2018) and activeness on the underlying assets (Easley et al., 2021). While the evidence that the ETS increases the volatility of the underlying stocks is vast, our analysis provides evidence that these effects are complemented by ownership structure.

The rest of this article is organized as follows. Section 2 describes the conceptual framework that supports our empirical approach. Section 3 describes the mutual fund (MF) and exchange-traded fund (ETF) data sources. Section 4 shows our main empirical results. Section 5 concludes and briefly discusses the implications of our results.

# 2 Conceptual framework

This section describes the theoretical framework that motivates our empirical methodology. We then review the theoretical model that illustrates the impact of non-fundamental demand on asset prices to describe how our proposed measure relates to previous studies.

## 2.1 Non-fundamental demand shocks

Non-fundamental demand shocks lead agents to trade an asset *independently* of fundamental information about deteriorating (improving) future growth perspectives or changes in risk aspects. While classic asset pricing theory regards those trades as *noise*, such flows can cause asset prices to deviate from their fundamental values (De Long et al., 1990). The financial economics literature that explores the reasons behind such trades is vast but can be broadly categorized into noise/liquidity-driven (De Long et al., 1990; Wang, 1994) and sentiment-driven[8] (Baker and Wurgler, 2006). While research acknowledges the influence of non-fundamental demand shocks on asset prices, it remains a challenge to identify those shocks since fundamental values are unobservable. To understand why ETF primary market flows provide clear signals of non-fundamental demand shocks, we begin by briefly describing the redemption/creation mechanism underlying ETF trading. We then show how this process helps us understand the intuition behind Brown, Davies and Ringgenberg (2021) model.

ETFs are commonly regarded as one of the most significant innovations in the asset management industry (Madhavan, 2014; Huang et al., 2020). Among the reasons for such success, are their low transaction costs and intraday liquidity. Moreover, the expansion of the ETF industry has led to the creation of a wide variety of products tracking a wide set of benchmarks which offers investors the possibility of gaining exposure to the broad market and specific sectors (Ben-David et al., 2023). Complementing these characteristics is the redemption/creating mechanism that sets ETFs apart from other investment vehicles.

ETFs are investment entities that issue securities (primary market) continuously

[8]The literature on investor sentiment include explanations based on over and under reaction, gambling-like behavior, disposition effect among among all the phenomena studied by behavioral finance approach

traded on public exchanges (secondary market). This specific setting leads to the possibility of ETFs shares diverging from the Net Asset Value (NAV) of the underlying securities that compose the benchmark. This arbitrage opportunity is corrected by Authorized Participants (AP). This is known as the creation-redemption mechanism. Brown, Davies and Ringgenberg (2021) argue that this temporary dislocation between ETF's NAV and the value of their underlying assets signals the appearance of a non-fundamental demand shock. Moreover, since these discrepancies are corrected through the redemption (creation) of ETF shares by APs, these changes in ETF shares (i.e., *flows*) allow researchers to observe those non-fundamentally driven trades.

Brown, Davies and Ringgenberg (2021) model the equilibrium ETF flow as follows

$$\lim_{N\to\infty} \Delta = \frac{\epsilon^{ETF} - \epsilon^{NAV}}{\lambda + \eta} \tag{1}$$

where $\Delta$ is the ETF flows, $\epsilon^{ETF}$ and $\epsilon^{NAV}$ represent aggregate non-fundamental demand shocks affecting the ETF shares price and the ETF underlying NAV, respectively. $\lambda$ proxies for investors' sensitivity to the measure of shares, and $\eta$ is a measure of price pressure. In other words, equation (1) states that the equilibrium ETF flows does not contain fundamental shocks[9] but that they are the product of net excess demand in either the ETF shares of the ETF underlying assets. The authors confirm their theoretical model by empirically showing that ETF flows predict future asset returns that later reverse and that these effect is strongests among leveraged ETFs and high activity ETFs (those with more active primary markets).

Brown, Davies and Ringgenberg (2021) relate their model to the well-known

[9]This is because while the demand for the ETF shares and the demand for the underlying assets both contain fundamental information, this component does not directly appear in the *relative mispricing*.

Berk and Green (2004) model that states that mutual fund flows reflect learning and adapting investor behavior regarding manager's skill. However, a notable difference is that ETFs are passively managed vehicles, thus flows cannot reflect learning about managerial skills by investors, but competition among APs that arbitrage away any misalingment between ETF shares value and its underlying NAV. We rely on this aspect as one of the main advantages of an ETF-based fragility measure over the one based on mutual fund data since ETF-flows are free of any discretionary skill-revealing information but rather signal arbitrage activity.

Overall, we argue that as shown by Brown, Davies and Ringgenberg (2021) the arbitrage mechanism that characterizes the ETF primary market provides two main benefits for fragility estimation (i) ETF shares creation and redemption (primary market) signals non-fundamental demand shocks, (ii) the mechanical correction of misalignments between ETF NAV and underlying assets avoids concerns regarding discretionary decisions that might confound fundamental information in fund flows.

## 2.2 Ownership structure and non-fundamental risk

Stock price fragility measures a security exposure to *shifts* in non-fundamental demand by capturing the joint influence of ownership composition and the variance-covariance matrix of liquidity-driven trades (*flows*) of the asset owners. Greenwood and Thesmar (2011) propose this measure based on a model that represents the changes in portfolio assets as the function of two components: i) that due to active rebalancing, and ii) flow-driven trading. Then, under the assumption of a stable relationship between aggregate flow-driven buys into security, its return can be modeled as follows:

$$r_{i,t+1} = \alpha + \lambda \frac{\sum_k w_{ikt} f_{kt}}{\theta_{it}} + \epsilon_{i,t+1} \tag{2}$$

Where $r_{i,t+1}$ is the return of security $i$ calculated from time $t$ to time $t+1$, $w_{ikt}$ is the dollar weight of security $i$ in investor's $k$ portfolio at time $t$, $f_{kt}$ net inflows to investor $k$, $\theta_{it}$ is a scaling factor usually proxied by the market capitalization of security $i$, $\lambda$ measures impact of liquidity trades, and $\epsilon_{i,t+1}$ and error term. Thus, a security expected return can be modeled as a function of aggregate flow-driven trades and an error term. In other words, the assumption of non-fundamental returns due to flows is captured by the first term on the right-hand side of this equation, while the error term $\epsilon$ can be seen as capturing information about security's $i$ fundamentals. If flow-driven demands cancel out across owners, prices should reflect fundamental information only. However, if that demand is not solved, it has the potential to exert temporary non-fundamental pressure on prices.

Another key assumption of this model is that flow-driven trading proxy for liquidity demand from investors rather than fundamentals. Therefore, the $f_{kt}$ and $\epsilon_{i,t+1}$ are uncorrelated. The stock fragility measure is then developed by first computing the conditional variance of $r_{i,t+1}$ as follows:

$$Var_t(r_{i,t+1}) = \underbrace{\lambda^2 \left(\frac{1}{\theta_{i,t}}\right)^2 W_{i,t}\Omega_t W_{i,t}}_{\text{Fragility (G)}} + \sigma^2_{i,t} \tag{3}$$

where $\sigma^2_{i,t}$ is the conditional variance of the $\epsilon_{i,t+1}$ term (i.e., fundamental news). Following the assumption of orthogonality between both components, the two components that determine the variance due to non-fundamental demand are: $W_{i,t}$ which is the vector of the weight of each investor in security $i$ and $\Omega_t$ the conditional variance-covariance matrix of flows from investors. Greenwood and Thesmar (2011) name their proposed variable $G$ stock price fragility,

Ben-David et al. (2021) focuses on studying the relationship between large institutions' ownership and asset prices. In principle, the authors argue that the demand by institutional demands influence stock return behaviors by means of their proportion of ownership and by the institutions' possibility of diversifying away any idiosyncratic shocks (i.e., granularity). Therefore, based on models of assymetric information and risk-averse market makers, Ben-David et al. (2021) derive the following expression of the determinants of the variance of stock returns:

$$Var_t(r_{i,t+1}) = \sigma_e^2 + \mu^2\epsilon_a^2 + \mu^2\epsilon_n^2 \sum \left(\frac{w_{ikt-1} f(A_{kt-1})}{m_{it-1}}\right)^2 \tag{4}$$

Where $\sigma_e^2$ represents the variance component explained by fundamental idiosyncratic shocks, just the $\sigma_{i,t}^2$ component in equation (3), $\mu^2\epsilon_a^2$ captures systematic aggregate shocks driving institutional trades, while the third component includes the effect of the structure of ownership $w_{ikt-1}$, stock's market capitalization $m_{it-1}$, an idiosyncratic component $\epsilon_n^2$ and a function of institutional investor $k$ demand for stock $i$, $f(A_{kt-1})$. $\mu$ is a measure of price impact[10]

While both Greenwood and Thesmar (2011) and Ben-David et al. (2021) models consider a stock's ownership structure as a determinant of stock volatility, the latter models this effect as a function of the degree of investor granularity and institutional investor size. Moreover, their model considers the effect that a common component may influence aggregate demand from institutional investors ($\mu^2\epsilon_a^2$).

We argue that an ETF-based fragility measure is able to capture both aggregate demand shocks to institutional investors while considering the effect of their ownership. This element is missing in Greenwood and Thesmar (2011) specification since mutual funds are mostly held by households while ETF are owned and traded by both institutional[11] and retail investors (Dannhauser and Pontiff, 2019). Also,

[10]While this component includes the consideration that demand shocks may affect stock prices differently, both Greenwood and Thesmar (2011) and Ben-David et al. (2021) models assume this parameter to be the same across stocks.

[11]On Appendix A shows the progressive inclusion of ETFs in 13F institutional investors' portfolios. We also include data on the adoption of leveraged and inverse-leveraged ETFs. We confirm

the literature on the widespread use of ETFs by institutional investors is fast growing showing their use as means to actively gain exposure to specific sectors (Easley et al., 2021), and by arbitrageurs circumventing short-sale constraints (Karmaziene and Sokolovski, 2022; Li and Zhu, 2022) and hedging industry risk (Huang et al., 2020).

### 2.3 An ETF-based stock price fragility ($G^{ETF}$)

Estimating $G_{it}$ faces two main empirical challenges: (i) observe a source of independent shocks to stock prices that is orthogonal to firm fundamentals; and (ii) access to *complete* data on assets' ownership structure. The first challenge, theoretically the most relevant one, has been widely studied in the financial economics literature. Starting with Coval and Stafford (2007), many studies have employed flow pressure from mutual fund sales as indicative of non-fundamental price shocks [12]. Among the reasons to use mutual fund data are evidence of mutual funds mechanically reducing portfolio holds when faced with significant outflows (i.e., fire-sales) and the well-known fact that the vast majority of mutual funds share owners are households[13] considered to be less financially sophisticated.

Greenwood and Thesmar (2011) cite this empirical evidence and rely on mutual fund data to estimate the stock fragility measure. It deserves to be emphasized that the fragility measure includes *all mutual fund flows* and does not rely on the most widely used MFFLOW measure proposed by Edmans et al. (2012). MFFLOW aims at capturing forced selling activity following large mutual fund outflows. While this approach does not directly suffer from the mechanical realized return mecha-

the findings in the literature by showing the widespread use of ETFs by institutional investors.

[12]A non-comprehensive list of related studies in empirical asset pricing area include Lou (2012); Edmans et al. (2012); Huang et al. (2021); Li (2022). See Wardlaw (2020) for a complete discussion of the related literature in empirical corporate finance.

[13]According to the 2020 Investment Company Institute (ICI) Fact Book, more than 89% of mutual fund assets in the US were held by households.

nism highlighted by Wardlaw (2020), the concerns that mutual funds flows convey fundamental information remain. Recent papers have raised concerns about using mutual fund outflow-induce price pressure measures. For instance, Berger (2022) show that the proportional trading assumption (i.e., managers sell off shares of their portfolio firms in proportion to their current portfolio weights) does not hold when empirically tested and leads to significantly biased inferences. Berger (2022) show that mutual fund managers discretionary sell specific firms. Huang et al. (2022) find similar results showing that fire sales also convey fundamental information . Overall, while empirical evidence shows that mutual fund flows, recent studies point out that employing mutual fund flows approach is, at most, a noisy measure of non-fundamental shocks.

Motivated by Brown, Davies and Ringgenberg (2021) evidence, we argue that an ETF-based stock price fragility ($G^{ETF}$) measure is not affected by the documented concerns related to mutual fund flows due to using ETF primary flows as signals of non-fundamental driven demand shocks. Moreover, as previously discussed, this measure has the potential to capture the effect of institutional demand on asset prices, an effect mostly overlooked by current methodology. Thus, we follow Greenwood and Thesmar (2011) and propose a fragility measure that employs only information (i.e., fund flows and ownership composition) from ETFs.

$$G_{it}^{ETF} = \left(\frac{1}{\theta_{i,t}}\right)^2 W_{i,t}^{ETF} \Omega_t^{ETF} W_{it}^{ETF}, \tag{5}$$

Where $W_{it}$ is the vector of weights of each ETF in security $i$ at time $t$, $\Omega_t$ is the conditional variance-covariance matrix of investors' dollar flows at time $t$, and $\theta_{it}$ is a scaling factor, usually proxied by security's market capitalization.

We also extend the expression in equation (5) to distinguish explicitly between Active and Passive ETFs as described by Easley et al. (2021). In this approach, we follow Greenwood and Thesmar (2011) decomposition and rewrite the fragility measure to include a term for each type of ETF, and a component that considers the holdings-weighted covariance between the two of them, as detailed in the following equation.

$$G_{it}^{ETF} = \left(\frac{1}{\theta_{it}}\right)^2 (W^{Act}\Omega^{Act}W^{Act} + W^{Pas}\Omega^{Pas}W^{Pas} + 2W^{Act}\Omega^{Act,Pas}W^{Pas}) \qquad (6)$$

Although this decomposition rules out any other possible investors by considering only ETFs, it offers the potential benefit of estimating the effect that fragility sources have on the full measure. This approach allows us to empirically test the concern brought up by Easley et al. (2021) regarding the increased activeness of the ETFs on price discovery. Moreover, the development of the ETF industry has been characterized by the creation of highly heterogenous products (Ben-David et al., 2023). We argue that our measure helps shed lights on these open questions regarding the impact of ETF trading on overall market efficiency. While $G_{it}^{ETF}$ represents a plausible improvement to the fragility measure, we are aware that it is still an imperfect proxy. Specifically, we rely on the same assumption of uncorrelated liquidity-driven trades from investors outside of our sample. We argue that they partially address some of the shortcomings of the fragility estimation and potentially help reconcile literature on mutual funds and ETF influence on underlying stocks return volatility. In the Online Appendix section 0A.1 we provide a theoretical model that provides a microfoundation of the measure of stock fragility and adds rigor to our discussion of the relationship between fragility and stock price return volatility.

# 3 Data and variable construction

We now describe our data sources and the data-retrieving process before explaining the methodology for estimating our fragility measure.

## 3.1 Mutual funds data

Our sample consists of US mutual funds from 1980 to 2018. To determine the sample period, we follow two criteria. First, since our motivation is to estimate the benefits of our proposed modified fragility measure, we closely follow Greenwood and Thesmar (2011) and start our sample period from the last quarter of 1989. This allows us to replicate their estimations. Second, although the first US-listed ETF, the SPDR, was launched in 1993, they became relevant investment vehicles in terms of the number of funds, assets under management (AUM), and participation in total volume traded in the period 2007-2009 (Madhavan, 2014). This period matches with the end of Greenwood and Thesmar (2011)'s sample period. Thus, to test the explanatory power of our proposed measure, we focus on the later part of our sample, starting in 2009, which allows us to capture the increase of ETF activity as well as perform an out-of-sample test of the original fragility measure in a context of raise of passive investing[14].

From the Center for Research in Security Prices (CRSP) Mutual Fund Database, we collect fund returns and total net assets (TNA). We then collect mutual funds' quarterly holdings data from the Thomson/Refinitiv Mutual Fund Database (*s12*). We merge both databases using the MFLinks database. As commonly done in the

[14]For instance, Madhavan (2014) highlights than the US ETF industry assets under management rose from \$70 billion in 2000 to \$1.7 trillion by mid-2014. Glosten et al. (2021) mentions that an increase in market participation has accompanied the rise in AUM since approximately 30% of US equity trading volume is attributable to ETFs. Regarding relocation from other investment vehicles, in 2017, the demand for equity ETFs resulted in \$186 billion net share issuance, whereas domestic equity mutual funds had net redemptions of \$236 billion.

literature, we proceed to clean our dataset only to include observations for which the FDATE matches RDATE. We follow Doshi et al. (2015) to identify and select US domestic equity mutual funds. Also, we exclude mutual funds with less than 5 million dollars in TNA (Friberg et al., 2022). Our fund sample includes 3,871 distinct US domestic equity mutual funds with 138,316 fund-quarter observations from the 1989-2018 period.

As commonly done in previous studies, we limit our holdings sample to include only stocks whose market capitalization is equal or above the NYSE market capitalization decile 5[15].

## 3.2 Exchange-traded funds (ETFs) data

To create our main ETF database, we use the list of ETFs identifiers from Brown, Davies and Ringgenberg (2021)[16] and combine it with data from Bloomberg and CRSP. From Bloomberg, we obtain data on outstanding shares and funds' net asset value (NAV). When the data is missing or incomplete, we supplement them with data from CRSP. We collect data on funds' prices and returns from CRSP. We obtain data on ETFs portfolio holdings using the Thomson/Refinitiv Mutual Fund Holdings (*s12*) and complement it with CRSP Mutual Fund Database data. Our sample of ETF data covers the period from 2000 to 2018.

We impose the same filters in stock size (market capitalization NYSE 5th decile or above) as in the sample of mutual funds to ensure comparability. In total, our sample includes 1,096 distinct ETFs for which we have both holdings data and

[15] Greenwood and Thesmar (2011) highlights two advantages of applying this filter: (1) Simplifies matrix computations (2) ensures that the estimation focuses on stocks of greater dollar importance more likely to be affected by liquidity-driven trades. Similarly, Francis et al. (2021) highlights that an empirical issue in fragility estimation is that it becomes highly noisy if a stock has low mutual fund ownership, which is precisely the case for stocks with smaller market capitalization. Thus, limiting the sample of stocks included in the holdings data reduces the possibility of distortions introduced by those noisy estimations.

[16] We thank David Brown for providing us with this data.

price/return data.

**[Table 1 Here]**

Table 1 (panels A and B) presents descriptive statistics of our Mutual Funds and ETFs sample. While in any given year, our sample includes more mutual funds (1,134) than ETFs (334), the ETFs are larger in terms of AUM and hold a larger number of stocks. Also, as detailed in previous studies, we observe a significant increase in ETF ownership over time (Da and Shive, 2018; Glosten et al., 2021). Specifically, it increased from 0.63% on average in the first part of the ETF sample period to 3.96% in the later part of our sample.

## 3.3 Mutual-fund-based Fragility

We estimate fragility as detailed in equation (5). The two main components of the fragility measure are a security ownership composition and the variance-covariance matrix of investors' non-fundamentally driven trades. The ownership structure is proxied by a vector of each mutual fund investor's portfolio allocation weight to stock $i$ relative to the fund's total net assets, as described in the following expression:

$$w_{i,j,t} = \frac{n_{i,j,t} P_{it}}{a_{j,t}}$$

where $n_{i,j,t}$ is the number of securities $i$ held by mutual fund $j$ at time $t$, $P_{it}$ is the price of security $i$, and $a_{j,t}$ is the total $j$ mutual fund portfolio value. We then proceed to sum over all the Mutual funds that hold shares of each security $i$.

Then, we proceed to calculate percentage flows for each mutual fund $i$ at the end of quarter $t$ as follows:

$$MFFlow_{j,t} = TNA_{j,t} - TNA_{j,t-1}(1 + R_{j,t})$$

where $TNA_{j,t}$ is the mutual fund $j$ Total Net Asset for quarter $t$ and $R_{j,t}$ is the fund's total return over that same quarter. Since we employ dollar positions of each fund in each security in matrix $W$, we require the covariance matrix $\Omega_t$ to be expressed in dollar terms. We follow Greenwood and Thesmar (2011) and rescale the $\Omega_t$ matrix by funds assets at time $t$ to obtain an estimate $\hat{\Omega}_t$:

$$\hat{\Omega}_t = diag(TNA_{j,t})\Omega_t diag(TNA_{j,t})$$

Each quarter $t$, we calculate $\hat{\Omega}_{j,t}$ on a five-year rolling window estimation starting from 1984:Q1. Another advantage of the fragility measure is that since it is estimated at the stock level, it is possible to aggregate it at the fund level (Friberg et al., 2023) or even at the factor level (Huang et al., 2021).

## 3.4 ETF-based Fragility

We estimate *ETF-based* fragility based on the same specification as in equation (5), however, we include ETF holdings and flows data only. The elements of matrix $W$ are estimated the same way as with the mutual fund data. Thus, this vector represents the ETFs portfolio allocation weights to each stock $i$ multiplied by the stock's $i$ price and divided by the total net assets of ETF $k$.

As with the *MF-based* fragility, we start by estimating ETF percentage flows as the change in shares outstanding for ETF fund $k$ at time $t$,

$$ETFFlow_{k,t} = \frac{SharesOutstanding_{j,t}}{SharesOutstanding_{j,t-1}} - 1$$

Greenwood and Thesmar (2011) warns about using dollar units to construct the variance-covariance matrix of flows since it would induce heteroskedasticity. Thus, as performed with the mutual fund data, we normalize ETF fund flows covariance matrix $\Omega_{k,t}$ as follows.

$$\hat{\Omega}_k = diag(TNA_{k,t})\Omega_{k,t}diag(TNA_{k,t})$$

For consistency with the *MF-based* fragility, we estimate $\hat{\Omega}_k$ based on a five-year rolling window. We select this starting year to ensure that our estimations include enough data and avoid noisy outcomes.

**[Table 2 Here]**

Table 2 shows descriptive statistics of the variables that compose the fragility measure (Panels A, B, and C) as well as for the square root of MF-based and ETF-based fragility (Panel D). Both the number of mutual funds and ETFs holding the same stocks have increased over time, especially in the ETF sample for the later part of our sample period. The average stock in the mutual fund sample is held by 50 funds, while in the ETF sample, it is approximately 25 funds.

Panel C of Table 2 reports the sample variation of flow's volatility estimated as the standard deviation of percentage mutual fund(ETF) flows. The volatility of Mutual fund flows increased in the first part of our sample period, 1989 to 2009, in line with the results of Greenwood and Thesmar (2011). Nonetheless, the volatility decreased in the out-of-sample period from 2010 to 2018. On the other hand, the volatility of ETF flows has increased significantly over the full sample period, especially in the later part, 2014-2018.

Concerning the correlation between flows, panel C shows a decrease in mean values for the mutual funds and ETF sample. It is worth mentioning that both the bottom and top quintile of flows correlation is considerably similar for both mutual fund flows and ETF flows. In relation to the squared root value of fragility, panel D, the increase in the mean value for the mutual fund sample is significant for the period between 1989 and 2009, from 0.039 to 0.143. Nonetheless, this value decreased in the later part of our sample to 0.102 on average. On the contrary, the mean $\sqrt{G}$ has continuously increased for the ETF sample.

A potential concern is that the estimated fragility values are driven by significant differences in the stocks' characteristics included in each sample. We check for this possibility in table A1 of the Online Appendix. In each quarter *t,* stocks are sorted into 5 quintile portfolios based on their MF-based $G^{MF}$ (Panel A), ETF-based $G^{ETF}$ (Panel B). Then, we estimate the time-series averages of the cross-sectional mean of the following variables for each quintile: price, market capitalization, ratio book-to-market, past 12-month stock returns, firm age (in years), turnover, and the average number of owners. These are variables that previous studies have shown to be correlated with future stock return volatility. Our results reveal that values and dispersion among quintiles are very similar for most variables for the three different panels and closely resemble the values reported by Greenwood and Thesmar (2011). This is consistent with the assumption that differences among the sample of underlying securities do not drive fragility values.

# 4 Empirical Results

In this section, we perform our main analysis and estimate the fragility measures influence on the future volatility of asset returns in a regression setting.

## 4.1 Fragility and stock return volatility

We test for this predictive power by estimating the following Fama and MacBeth (1973) regression[17].

[17]We perform Fama and MacBeth (1973) regressions to control for the effect of common trends like increasing ownership of Mutual Funds and ETF.

$$\sigma_{i,t+1} = \alpha + \beta\sqrt{G}_{i,t} + \delta Z_{i,t} + \mu_{i,t} \tag{7}$$

Where $\sigma_{i,t+1}$ is the one-quarter-ahead standard deviation of daily stock returns or excess returns. For the excess return volatility, we estimate risk-adjusted returns from three models: (1) market-adjusted returns, (2) Fama and French (1993) three factors model, (3) Fama and French (1993) model augmented with the Carhart (1997) momentum factor. $Z_{i,t}$ represents the vector of control variables, including the log of unadjusted stock price, the natural logarithm of market capitalization, the ratio of book equity to market equity, the past 12-month stock return, lagged skewness of stock returns, the log of firm's age (in months) and share turnover. The coefficient $\beta$ measures the relationship between the current quarter's fragility and the next quarter's stock return volatility. A positive $\beta$ would indicate that an increase in stock fragility in the current quarter would forecast an increase in the next quarter's stock return volatility.

Equation (7) follows the main specification employed by Greenwood and Thesmar (2011). To examine the performance of $G^{MF}$ *out-of-sample* we repeat the four main specifications for the sample period between 2009-2018. For comparability, we run the regressions on $G^{ETF}$ for that same period. In addition to evaluating each measure relation with asset's return volatility, we also estimate the same previous specifications including the simultaneous influence of $G^{MF}$ and $G^{ETF}$. These results are reported in Table 3 and 4.

Columns (1) and (2) report the estimation results of the Fama-Macbeth regression of one-quarter-ahead standard deviation of daily stock returns on $\sqrt{G^{MF}}$ alone. As previously documented, we confirm that fragility is a strong predictor of future volatility. However, while the coefficient on $\sqrt{G^{MF}}$ remains positive and statistically significant in the second part of the sample period (2010-2018) we observe a reduction in both the coefficient and the $t$-statistic compared to that reported by Greenwood and Thesmar (2011).

In Columns (2) and (6), we confirm previous evidence[18] that daily stock returns are correlated with mutual fund ownership ($IO$). Nonetheless, the correlation between the number of owners and future volatility lost explanatory power in the later part of our sample (Column 6).

In columns (3) and (7), we check for the explanatory power of $\sqrt{G^{MF}}$ beyond ownership concentration, proxied by both mutual fund share and ownership Herfindal index. Consistent with Greenwood and Thesmar (2011), we observe that the coefficients on fragility are slightly affected by including such control, thus confirming fragility's ability to explain stock return volatility beyond ownership concentration alone. However, it is worth mentioning that the reduction in the coefficient values is more pronounced in the second part of our sample.

In the specifications shown in columns (4) and (8), we include the complete set of control variables. For these cases, the coefficient on $\sqrt{G^{MF}}$ drops considerably for the full sample, 0.072 ($t$-stat=2.75), and especially in the later part of the sample, 0.018 ($t$-stat=1.70). While in both cases, the coefficient remains statistically significant, the reduction in $t$-statistic value is substantial.

**[Table 3 Here]**

The remaining columns in table 3 repeat the previous analysis employing $\sqrt{G^{ETF}}$. Column (9) repeats the specifications of Columns (1) and (5). As seen, $\sqrt{G^{ETF}}$ is a strongly positive predictor of next quarter standard deviation of daily stock returns. Moreover, both the coefficient and t-stats of the variable are significantly higher that of the $\sqrt{G^{MF}}$ for the same period. This result remains after we include the full

[18] For instance, Bushee and Noe (2000), Sias (1996)

set of controls, column (12). In column (10) we confirm the findings of Ben-David et al. (2017) that higher ETF ownership is related with higher volatility.

Our results show a decline in the $\sqrt{G^{MF}}$ measure while $\sqrt{G^{ETF}}$ remains a significant predictor of return's volatility. As discussed in section 2.3, in contrast to mutual funds, ETFs are owned almost equally by households and institutional investors. Thus, it can be expected that the influence of non-fundamentally driven demand that $G^{ETF}$ captures is different to that of $G^{MF}$. In other words, while the ETF-based measure captures a similar component as the MF-based fragility (i.e., retail investors' demand), it also includes institutional investors' demand. We begin by exploring these differences by repeating the analysis of table 3 but including both $\sqrt{G^{ETF}}$ and $\sqrt{G^{MF}}$ simultaneously.

**[Table 4 Here]**

Table 4 shows the results of such analysis. We limit the analysis for the second part of the sample (2008-2018) to ensure comparability between measures. In column (1) we observe that the coefficient on $\sqrt{G^{ETF}}$ is reduced when we include $\sqrt{G^{MF}}$, nonetheless, remains statistically significant[19]. In columns (3) and (4) we sequentially add control variables and observe that $\sqrt{G^{ETF}}$ coefficient is lower but statistically significant while $\sqrt{G^{MF}}$ is no longer significant.

**[Table 5 Here]**

We also explore the relationship between fragility and the volatility of returns in excess of several asset pricing factors. Greenwood and Thesmar (2011) state that under this specification we can expect to get weaker results since aggregated

[19] In table A4 of the Online Appendix we replicate this analysis in a different specification setting as performed by Friberg et al. (2022). In this panel regression setting, our results remain mostly unchanged.

versions of fragility may predict the volatility of risk factors themselves. Recent empirical evidence support this claim by showing that a significant component of asset pricing factor return movement is not due to changes in economic risk but rather the result of price pressure from flows (Huang et al., 2021; Li, 2022). Table 5 shows the results of this analysis. We confirm Greenwood and Thesmar (2011) initial findings. Moreover, in line with our previous results, we observe a significant decline in coefficient magnitude of $\sqrt{G^{MF}}$ variable for the second part of our sample of Mutual funds (Panel A). We then turn to estimating the relationship with $\sqrt{G^{ETF}}$ and for a specification that includes both measures simultanously. Panel B of table 5 shows this results. We see that the coefficients of $\sqrt{G^{ETF}}$ are significantly higher than those of $\sqrt{G^{MF}}$. Moreover, the inclusion of $\sqrt{G^{MF}}$ reduces $\sqrt{G^{ETF}}$ coefficient only by a slight amount. As seen, the estimated relationship between fragility and excess return volatilty is significant for both $\sqrt{G^{MF}}$ and $\sqrt{G^{ETF}}$. Nonetheless, this evidence points out to the ETF-based measure to showing a stronger forecast power.

In sum, our regression results suggest that the difference in the coefficients and t-stats between $G^{ETF}$ and $G^{MF}$ is non-trivial and point out to an increased ability of $G^{ETF}$ to forecast volatility. These results support the notion that ETF flows can capture non-fundamentally driven shocks that later lead to increase return volatility. Additionally, we explored the possibility that the difference may be due to $G^{ETF}$ capturing the effect of institutional ownership on volatility. We expand on this potential explanation in the next section.

## 4.2 Fragility and large institutional investors

In a recent paper, Ben-David et al. (2021) shows that trading by large institutional investors induces greater return volatility due to their inability to diversify idiosyncratic shocks among their subunits. In other words, the subunits that conform large institutional investors show correlated behavior during those shocks increasing its effect on asset price volatility. In this section, we separately include the ownership of large institutional investors, as well as mid and bottom institutions in terms of assets under management, and evaluate the impact that the predictive power of $G^{MF}$ and $G^{ETF}$.

Thus, we follow Ben-David et al. (2023) perform the following panel regression

$$\sigma_{i,t+1} = \beta_1 \text{TopIO}_{i,t} + \beta_2 \text{MidIO}_{i,t} + \beta_3 \text{BottomIO}_{i,t} + \delta Z_{i,t} + \beta_4 G_{i,t} + \alpha_i + \theta_t + \mu_{i,t} \quad (8)$$

where $\sigma_{i,t+1}$ is the next quarter $t$ stock $i$ volatility. $\text{TopIO}_{i,t}$ is the fraction of shares outstanding collectively held by the top institutions ranked based on the money value of portfolio holdings over the previous four quarters. $\text{BottomIO}_{i,t}$ represents the aggregate stock's $i$ ownership of the smallest institutional investors whose aggregate money holdings value equals that of the top institutions. $\text{MidIO}_{i,t}$ is collective ownership by those institutions not classified as top neither as bottom. $Z_{i,t}$ is the vector of control variables that include the log of market capitalization, book-to-mark ratio, past 6-month momentum returns, the inverse of price ratio (1/price), and the Amihud illiquidity measure (Amihud, 2002). $\alpha_i$ is the stock fixed effect, and $\theta_t$ is the time (calendar-quarter) fixed effect.

**[Table 6 Here]**

Table 6 shows the results for two specifications: considering top 3 and top 10 institutional investors. Our results show that $G^{MF}$ is able to predict next-quarter stock volatility while including the effect of institutional investors' ownership for the full sample period. This is in line with the results shown by Ben-David et al. (2021). However, we observe that this relationship is significantly weaker in the later part of our sample period (2008-2018).

We find that $G^{ETF}$ is positive and statistically significant for the full period, but the effect is stronger for the later part of the sample. Interestingly, we observe that including $G^{ETF}$ reduces the explanatory power of bottom institutions on next quarter stock return volatility. We replicate the results for alternative grouping of top institutional investors, specifically top 5 and top 10 in the online Appendix, table A3. Our results remain qualitative the same. Overall, our evidence confirms our assumption that $G^{ETF}$ measure partially captures the effect of institutional investors demand on volatility and that $G^{MF}$ does not consider. Our results are in line with Koijen and Yogo (2019) that shows that while top institutions hold a significant portion of the stock market capitalization, smaller size institutions and households account for most of stock's volatility .

### 4.3 ETF activeness

A valid concern of our empirical analysis is that we combine data from two *distinct* investment vehicles. This is in terms of their investment mandates. We follow Easley et al. (2021) and estimate their *Activenss Index* to estimate which fraction of our sample of ETFs can be considered to be active.

$$\mathrm{ActivenessIndex_{i,t}} = \sum_{s=1}^{N} w_{i,s,t} - w_{market,s,t} \tag{9}$$

In table 7, for our larger sample of ETFs, we corroborate Easley et al. (2021) findings and show that most ETFs can be considered as active investment vehicles.

**[Table 7 Here]**

Is it possible that more active ETFs drive our results? those with an activeness index above 50%? While Brown, Howard and Lundblad (2021) does not differentiate ETFs by activeness level Easley et al. (2021) shows concern that the growing activeness of ETFs may harm price discovery. While, we do not directly test for the effect on price discovery, it is possible to think that more active ETFs could be specially important to channel fragility. We follow Easley et al. (2021) and split our sample based on according to activeness index (50% threshold), since this cutoff most likely includes both *active-in-form* and *active-in-function* ETFs.

**[Table 8 Here]**

Table 8 repeates the main specifications of tables 4 and 5 but considering the decomposition of $G^{ETF}$ into active and passive ETFs. Column (1) shows that most of the observed relationship between $G^{ETF}$ and volatility steams from active ETFs component. Column(2) examines the same relationship if we include $G^{MF}$ while Column (3) shows the results when we include the full set of control variables. Our results confirm the concerns brought up by Easley et al. (2021) regarding the role that increased ETF activeness play on price informativenes. We show that the active ETF component of the ETF-based fragility measure is responsible for most of the observed relationship.

# 5 Conclusion

Numerous studies in financial economics have focused on measuring the influence of non-fundamentally driven trades on stock prices. These studies strive to empirically capture mispricing signals, gauge its impact on prices, and examine the reactions of market participants to such occurrences.A widely adopted empirical approach to capture a stock's exposure to non-fundamentally driven demand shocks is to estimate the Greenwood and Thesmar (2011) fragility. However, its estimation is based on strong assumptions that may raise concerns about data suitability and the validity of the inferences drawn from this approach.

Recently literature points out to ETF primary market data as clearly reflecting non-fundamentally driven shocks to stock prices (Brown, Davies and Ringgenberg, 2021). This paper proposes an ETF-based fragility measure that takes advantage of readily available data on ETFs flows and holdings. In doing so, we aimed at providing a measure free of most concerns that suround the use of mutual fund data. We also address recent concerns about the effect of increased ETF activeness on return's volatility.

Our empirical evidence supports our main prediction: an ETF-based fragility measure strongly predict price volatility. Not only are the findings both statistically and economically significant, but are stronger for the second part of sample period and remain significant for different specification and to the inclusion of a broad set of controls. Moreover, we document a decrease in the Greenwood and Thesmar (2011) explanatory power to forecast price volatility out-of-sample for a period that coincides with the raise of the ETF industry, both in terms of assets under management and volume traded.

We argue that avoiding accounting for changes in the current asset management industry (i.e., raise of *passive* investing) significantly bias stock's exposure to non-fundamental demand risk. Although our approach does not completely resolve the limitations associated with empirically estimating stock fragility, it represents a significant advancement by eliminating concerns of the inclusion of confounding fundamental information while offering researchers a more accurate proxy for assessing firm-level exposure to non-fundamental demand risk.

# References


Amihud, Yakov (2002), 'Illiquidity and stock returns: Cross-section and time-series effects', *Journal of Financial Markets* **5**(1), 31–56.

Amihud, Yakov (2019), 'Illiquidity and stock returns: A revisit', *Critical Finance Review* **8**(1-2), 203–221.

Baker, Malcolm and Jeffrey Wurgler (2006), 'Investor sentiment and the cross-section of stock returns', *Journal of Finance* **61**(4), 1645–1680.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim and Rabih Moussawi (2023), 'Competition for Attention in the ETF Space', *Review of Financial Studies* **36**(3), 987–1042.

Ben-David, Itzhak, Francesco Franzoni and Rabih Moussawi (2017), 'Exchange-Traded Funds', *Annual Review of Financial Economics* **9**(1), 169–189.

Ben-David, Itzhak, Francesco Franzoni and Rabih Moussawi (2018), 'Do ETFs Increase Volatility?: Do ETFs Increase Volatility?', *The Journal of Finance* **73**(6), 2471–2535.

Ben-David, Itzhak, Francesco Franzoni, Rabih Moussawi and John Sedunov (2021), 'The Granular Nature of Large Institutional Investors', *Management Science* **67**(11), 6629–6659.

Berger, Elizabeth A. (2022), 'Selection Bias in Mutual Fund Fire Sales', *Journal of Financial and Quantitative Analysis* **58**(3), 1039–1077.

Berk, Jonathan B. and Richard C. Green (2004), 'Mutual fund flows and performance in rational markets', *Journal of Political Economy* **112**(6), 1269–1295.

Brown, David C, Shaun William Davies and Matthew C Ringgenberg (2021), 'ETF Arbitrage, Non-Fundamental Demand, and Return Predictability', *Review of Finance* **25**(4), 937–972.

Brown, Gregory W, Philip Howard and Christian T Lundblad (2021), 'Crowded Trades and Tail Risk', *The Review of Financial Studies* **00**, 1–41.

Bushee, Brian J. (2001), 'Do Institutional Investors Prefer Near-Term Earnings over Long-Run Value?', *Contemporary Accounting Research* **18**(2), 207–246.

Bushee, Brian J and Christopher F Noe (2000), 'Corporate Disclosure Practices, Institutional Investors, and Stock Return Volatility', *Journal of Accounting Research* **38**(2000), 171–202.

Carhart, Mark M. (1997), 'On persistence in mutual fund performance', *Journal of Finance* **52**(1), 57–82.

Coval, Joshua and Erik Stafford (2007), 'Asset fire sales (and purchases) in equity markets', *Journal of Financial Economics* **86**(2), 479–512.

Da, Zhi and Sophie Shive (2018), 'Exchange traded funds and asset return correlations', *European Financial Management* **24**(1), 136–168.

Dannhauser, Caitlin D and Jeffrey Pontiff (2019), 'Flow', *SSRN Electronic Journal* pp. 12–26.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers and Robert J Waldmann (1990), 'Noise trader risk in financial markets', *Journal of Political Economy* **98**(4), 703–738.

Dong, Ming, David Hirshleifer and Siew Hong Teoh (2021), 'Misvaluation and Corporate Inventiveness', *Journal of Financial and Quantitative Analysis* **56**(8), 2605–2633.

Doshi, Hitesh, Redouane Elkamhi and Mikhail Simutin (2015), 'Managerial Activeness and Mutual Fund Performance', *Review of Asset Pricing Studies* **5**(2), 156–184.

Easley, David, David Michayluk, Maureen O'Hara and Tālis J Putniņš (2021), 'The Active World of Passive Investing', *Review of Finance* **25**(5), 1433–1471.

Edmans, Alex, Itay Goldstein and Wei Jiang (2012), 'The Real Effects of Financial Markets: The Impact of Prices on Takeovers', *The Journal of Finance* **67**(3), 933–971.

Fama, E. F . (1965), 'The Behavior of Stock-Market Prices Author ( s ): Eugene F . Fama Published by : The University of Chicago Press Stable', *The Journal of Business* **38**(1), 34–105.

Fama, Eugene F. and James D. MacBeth (1973), 'Risk, Return, and Equilibrium: Empirical Tests', *Journal of Political Economy* **81**(3), 607–636.

Fama, Eugene F. and Kenneth R. French (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3–56.

Fama, Eugene F. and Kenneth R. French (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1–22.

Francis, Bill, Iftekhar Hasan, Yinjie (Victor) Shen and Pengfei Ye (2021), 'Stock price fragility and the cost of bank loans', *Journal of Empirical Finance* **63**, 118–135.

Friberg, Richard, Itay Goldstein and Kristine W Hankins (2023), 'Corporate Responses to Stock Price Fragility *'.

Friberg, Richard, Itay Goldstein and Kristine Watson Hankins (2022), 'Corporate Responses to Stock Price Fragility', *SSRN Electronic Journal* .

Glosten, Lawrence, Suresh Nallareddy and Yuan Zou (2021), 'ETF Activity and Informational Efficiency of Underlying Securities', *Management Science* **67**(1), 22–47.

Greenwood, Robin and David Thesmar (2011), 'Stock price fragility', *Journal of Financial Economics* **102**(3), 471–490.

Huang, Sheng, Matthew C. Ringgenberg and Zhe Zhang (2022), 'The Information in Asset Fire Sales', *Management Science* (July 2023).

Huang, Shiyang, Maureen O'Hara and Zhuo Zhong (2020), 'Innovation and Informed Trading: Evidence from Industry ETFs', *The Review of Financial Studies* **00**(0), 37.

Huang, Shiyang, Yang Song and Hong Xiang (2021), 'Noise Trading and Asset Pricing Factors '.

Karmaziene, Egle and Valeri Sokolovski (2022), *Short-Selling Equity Exchange Traded Funds and Its Effect on Stock Market Liquidity*, Vol. 57.

Koijen, Ralph S.J. and Motohiro Yogo (2019), 'A demand system approach to asset pricing', *Journal of Political Economy* **127**(4), 1475–1515.

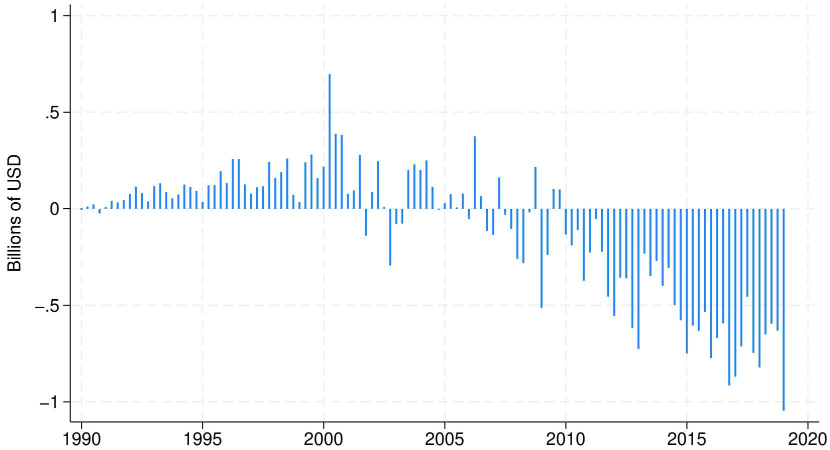
Li, Frank Weikai and Qifei Zhu (2022), 'Short Selling ETFs', *The Review of Asset*

*Pricing Studies* **00**(February), 1–39.
Li, Jiacui (2022), 'What Drives the Size and Value Factors?', *Review of Asset Pricing Studies* **12**(4), 845–885.
Lou, Dong (2012), 'A Flow-Based Explanation for Return Predictability', *Review of Financial Studies* **25**(12), 3457–3489.
Madhavan, Ananth (2014), 'Exchange-Traded Funds: An Overview of Institutions, Trading, and Impacts', *Annual Review of Financial Economics* **6**(1), 311–341.
Merton, Robert C (1971), 'Optimum consumption and portfolio rules in a continuous-time model', *Journal of Economic Theory* **3**(4), 373–413.
Ross, Stephen A. (1976), 'The arbitrage theory of capital asset pricing', *Journal of Economic Theory* **13**(3), 341–360.
Rzeznik, Aleksandra and Rudiger Weber (2022), 'Money in the right hands: the price effects of specialized demand', p. 62.
Shleifer, Andrei and Robert W. Vishny (1997), 'The Limits of Arbitrage', *The Journal of Finance* **52**(1), 35–55.
Sias, Richard W. (1996), 'Volatility and the institutional investor', *Financial Analysts Journal* **52**(2), 13–20.
Wang, Jiang (1994), 'A Model of Competitive Stock Trading Volume', *Journal of Political Economy* **102**(1), 127–168.
Wang, Jiang (1996), 'The term structure of interest rates in a pure exchange economy with heterogeneous investors', *Journal of Financial Economics* p. 36.
Wardlaw, Malcolm (2020), 'Measuring Mutual Fund Flow Pressure as Shock to Stock Returns', *The Journal of Finance* **75**(6), 3221–3243.
Yan, Xuemin and Zhe Zhang (2009), 'Institutional investors and equity returns: Are short-term institutions better informed', *Review of Financial Studies* **22**(2), 893–924.

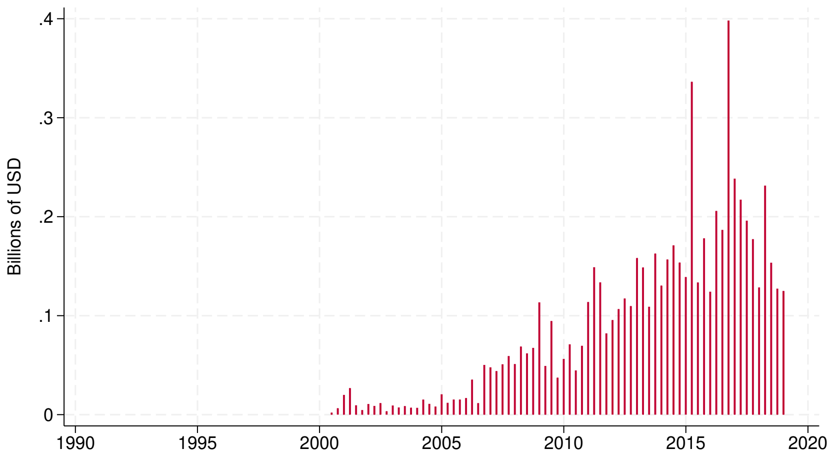
# 6 Tables and Figures

Figure 1: **Flows to Equity Mutual Funds and Exchange-Traded Funds (ETFs)**

This figure plots the total new cash flows to our sample of equity mutual funds in Panel A and to the exchange-traded funds (ETFs) in Panel B. The sample period for mutual fund data covers the period from 1989:Q4 to 2018:Q4. For ETFs data, the sample is from 2000:Q1 to 2018:Q4



(a) Mutual Funds



(b) Exchange−Traded funds (ETFs)

Table 1: **Descriptive Statistics**

This table reports the time-series average of the cross-sectional mean, median, standard deviation, and first and third quartiles of several variables for our sample of mutual funds and exchange-traded funds (ETFs). *Number of funds* is the average of the total number of funds per quarter. *Number of holdings* represents the average number of stocks in the fund's portfolio. *TNA* is the fund's total net assets at quarter's end, in millions of USD dollars. *Ownership* is the percentage of shares outstanding owned by all equity mutual funds (ETFs) in our sample. *NYSE Decile* is the average NYSE size decile of a mutual fund(ETF) portfolio stock. Panel A reports the descriptive statistics for our sample of mutual funds. Panel B shows the results for the sample of exchange-traded funds (ETFs). Only stocks with market capitalization equal to or higher than the NYSE size decile 5 are included. The *Full sample* covers the period from 1989:Q4 to 2018:Q4.

Panel A: Mutual Funds

| | Full Sample | | | | | Mean by period | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | p25 | Median | p75 | 1989-1999 | 2000-2009 | 2010-2018 |
| Number of funds | 1,138 | 501 | 690 | 1352 | 1537 | 524 | 1494 | 1441 |
| Number of holdings | 80 | 85.071 | 36 | 58 | 90 | 66 | 80 | 85 |
| TNA (in MM of USD) | 879.82 | 3764.83 | 30.80 | 132.78 | 532.74 | 467.22 | 733.43 | 1219.15 |
| Ownership (%) | 8.71 | 12.29 | 1.49 | 5.15 | 11.86 | 4.28 | 10.95 | 11.20 |
| NYSE decile | 8.05 | 0.11 | 7.99 | 8.03 | 8.11 | 8.08 | 8.10 | 7.97 |

Panel B: ETFs

| | Full Sample | | | | | Mean by period | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std | p25 | Median | p75 | 1989-1999 | 2000-2009 | 2010-2018 |
| Number of funds | 334 | 276 | 94 | 112 | 571 | | 89 | 606 |
| Number of holdings | 116 | 188 | 18 | 48 | 110 | | 93 | 120 |
| TNA (in MM of USD) | 1,760.5 | 9,280.1 | 34.3 | 157.8 | 689.1 | | 1,000.3 | 1,766.9 |
| Ownership (%) | 2.27 | 2.97 | 0.14 | 0.91 | 3.64 | | 0.63 | 3.96 |
| NYSE decile | 7.41 | 1.73 | 6.00 | 7.00 | 9.00 | | 7.46 | 7.37 |

Table 2: **Fragility and fragility components descriptive statistics**

This table reports the time-series statistics of cross-section averages mean, median, standard deviation, and first and third quartiles of the following variables: *Number of owners* is the total number of funds holding the same stock. *Flow volatility* represents the standard deviation of mutual (ETF) fund flows. *Flow correlation* is the Pearson correlation of fund flows at the fund pair level, each quarter. *Fragility* (sqrt) is the square root of the fragility measure estimated as in equation 3. Only stocks whose market capitalization is equal to or higher than the NYSE size decile 5 are included. The sample period for the equity mutual funds is from 1989:Q4 to 2018:Q4, while for the exchange-traded funds (ETFs) is from 2000:Q1 to 2018:Q4. Fragility is winsorized at the 1% and 99% levels.

| | Mutual funds | | | | | | ETFs | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Std | p25 | Median | p75 | | Mean | Std | p25 | Median | p75 |
| Panel A: Number of owners | | | | | | | | | | | |
| 1989-1999 | 22 | 26 | 7 | 15 | 27 | 2000-2008 | 5 | 4 | 2 | 4 | 7 |
| 2000-2009 | 76 | 71 | 28 | 59 | 100 | 2009-2013 | 31 | 24 | 7 | 31 | 50 |
| 2010-2018 | 82 | 65 | 40 | 72 | 108 | 2014-2018 | 51 | 31 | 29 | 48 | 73 |
| Full sample | 50 | 61 | 7 | 27 | 73 | Full sample | 25 | 29 | 4 | 9 | 44 |
| Panel B: Flow volatility | | | | | | | | | | | |
| 1989-1999 | 4.664 | 11.505 | 0.399 | 0.870 | 2.749 | 2000-2008 | 0.351 | 0.491 | 0.058 | 0.177 | 0.369 |
| 2000-2009 | 5.498 | 17.178 | 0.408 | 0.895 | 3.933 | 2009-2013 | 0.824 | 0.963 | 0.342 | 0.493 | 0.741 |
| 2010-2018 | 4.248 | 11.093 | 0.279 | 0.541 | 1.388 | 2014-2018 | 1.755 | 2.648 | 0.431 | 0.693 | 1.273 |
| Full sample | 4.821 | 13.500 | 0.331 | 0.650 | 2.472 | Full sample | 0.858 | 1.586 | 0.187 | 0.389 | 0.746 |
| Panel C: Flow correlation | | | | | | | | | | | |
| 1989-1999 | 0.097 | 0.646 | -0.384 | 0.133 | 0.653 | 2000-2008 | 0.066 | 0.633 | -0.441 | 0.058 | 0.615 |
| 2000-2009 | 0.069 | 0.485 | -0.215 | 0.069 | 0.386 | 2009-2013 | 0.027 | 0.460 | -0.238 | 0.004 | 0.306 |
| 2010-2018 | 0.035 | 0.417 | -0.179 | 0.033 | 0.260 | 2014-2018 | 0.025 | 0.433 | -0.225 | -0.006 | 0.273 |
| full sample | 0.072 | 0.432 | -0.149 | 0.063 | 0.319 | Full sample | 0.028 | 0.426 | -0.206 | -0.002 | 0.262 |
| Panel D: Fragility (sqrt) | | | | | | | | | | | |
| 1989-1999 | 0.039 | 0.207 | 0.000 | 0.001 | 0.005 | 2000-2008 | 0.001 | 0.006 | 0.000 | 0.000 | 0.000 |
| 2000-2009 | 0.143 | 0.434 | 0.001 | 0.006 | 0.051 | 2009-2013 | 0.010 | 0.041 | 0.000 | 0.000 | 0.000 |
| 2010-2018 | 0.102 | 0.217 | 0.001 | 0.022 | 0.114 | 2014-2018 | 0.064 | 0.130 | 0.000 | 0.001 | 0.047 |
| Full sample | 0.105 | 0.303 | 0.001 | 0.011 | 0.064 | Full sample | 0.028 | 0.089 | 0.000 | 0.001 | 0.001 |

Table 3: **Fragility and stock return volatility**

Standard deviation of daily stock returns over quarter $t$+1 ($\sigma_{t+1}$) are regressed on squared fragility $\sqrt{G}$ at quarter $t$ and a set of lagged control variables as detailed in Eq. (5) using the Fama and MacBeth (1973) methodology. This table reports the average slope coefficients and the Newey-West $t$-statistics in parentheses. Fragility is measured employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$), and ETF data only ($\sqrt{G}^{ETF}$). The control variables included are: the log of stock price, the log of market capitalization, the ratio of book equity to market equity, the past 12-month cumulative stock return, lagged skewness of monthly stock returns, the log of age, share turnover, and the lagged dependent variable ($\sigma$).

| | Mutual funds | | | | | | | | ETFs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Full sample | | | | 2009 - 2018 | | | | 2009 - 2018 | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| $\sqrt{G}^{MF}$ | 0.459*** | | 0.305*** | 0.072** | 0.325*** | | 0.189*** | 0.018* | | | | |
| | (11.82) | | (8.57) | (2.75) | (8.75) | | (6.26) | (1.70) | | | | |
| $\sqrt{G}^{ETF}$ | | | | | | | | | 0.825*** | | 0.722*** | 0.338*** |
| | | | | | | | | | (7.76) | | (7.10) | (5.93) |
| IO | | 0.015*** | | | | 0.014*** | | | | 0.003* | | |
| | | (15.64) | | | | (14.27) | | | | (2.35) | | |
| log(numb owners) | | 0.027 | | | | -0.033** | | | | -0.032*** | | |
| | | (1.26) | | | | (-2.82) | | | | (-3.37) | | |
| Own Herfindahl | | | -0.002*** | -0.001 | | | -0.004*** | -0.002*** | | | -0.001 | -0.011 |
| | | | (-4.27) | (-1.14) | | | (-6.51) | (-5.03) | | | (-1.00) | (-1.06) |
| Add Controls | No | No | No | Yes | No | No | No | Yes | No | No | No | Yes |
| $N$ | 148,342 | 148,342 | 148,342 | 137,283 | 58,377 | 58,377 | 58,377 | 54,633 | 45,078 | 45,078 | 44,808 | 42,776 |
| adj. $R^2$ | 0.010 | 0.049 | 0.045 | 0.486 | 0.007 | 0.045 | 0.043 | 0.376 | 0.013 | 0.025 | 0.024 | 0.373 |

Table 4: **MF and ETF Fragility and stock return volatility**

Standard deviation of daily stock returns over quarter *t+1* ($\sigma_{t+1}$) are regressed on squared fragility $\sqrt{G}$ at quarter $t$ and a set of lagged control variables as detailed in Eq. (5) using the Fama and MacBeth (1973) methodology. This table reports the average slope coefficients and the Newey-West $t$-statistics in parentheses. Fragility is measured employing only mutual fund flows and holdings data ($\sqrt{G}^{MF}$), and ETF data only ($\sqrt{G}^{ETF}$). The control variables included are: the log of stock price, the log of market capitalization, the ratio of book equity to market equity, the past 12-month cumulative stock return, lagged skewness of monthly stock returns, the log of age, share turnover, and the lagged dependent variable ($\sigma$).

| | 2009 - 2018 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $\sqrt{G}^{MF}$ | 0.067* | | 0.015 | 0.009 |
| | (1.99) | | (1.16) | (1.03) |
| $\sqrt{G}^{ETF}$ | 0.790*** | | 0.795*** | 0.426*** |
| | (7.77) | | (8.20) | (7.95) |
| $IO^{MF}$ | | 0.014*** | 0.012*** | 0.005*** |
| | | (11.11) | (12.37) | (7.47) |
| $IO^{ETF}$ | | -0.002* | -0.012*** | -0.007*** |
| | | (-2.03) | (-6.58) | (-4.96) |
| log (numb MF owners) | | -0.031* | | |
| | | (-2.25) | | |
| log (numb ETF owners) | | -0.032* | | |
| | | (-2.57) | | |
| Own MF Herfindahl | | | -0.004*** | -0.002*** |
| | | | (-10.74) | (-5.56) |
| Own ETF Herfindahl | | | 0.001 | -0.011 |
| | | | (0.77) | (-1.07) |
| Add Controls | No | No | No | Yes |
| Obs. | 44,956 | 44,956 | 44,956 | 44,956 |
| adj. $R^2$ | 0.015 | 0.025 | 0.034 | 0.376 |

Table 5: **Fragility and excess return volatility**

Standard deviation of *excess stock returns* over quarter *t+1* ($\sigma^{exc}_{t+1}$) are regressed on squared fragility $\sqrt{G}$ at quarter $t$. Excess returns are estimated based on the single-factor market model (1-Factor $\sigma$) the Fama and French (1993) three-factor model (3-Factor $\sigma$), and the Fama and French (1993) three-factor model augmented with the momentum factor of Carhart (1997) (4-Factor $\sigma$). This table reports the average slope coefficients and the Newey-West $t$-statistics in parentheses. In panel A, Fragility is measured based only on mutual fund flows and holding data ($\sqrt{G}^{MF}$). In panel B, Fragility is estimated as detailed in Eq. (5) based on ETF data only ($\sqrt{G}^{ETF}$).

Panel A: Mutual fund Fragility

| | Full sample | | | | 2009 - 2018 | | | |
|---|---|---|---|---|---|---|---|---|
| | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ | DGTW | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ | DGTW |
| $\sqrt{G}^{MF}$ | 0.530*** | 0.526*** | 0.527*** | 0.407*** | 0.400*** | 0.391*** | 0.397*** | 0.331*** |
| | (7.86) | (7.81) | (7.96) | (7.49) | (12.01) | (11.81) | (11.65) | (9.77) |
| Obs. | 148,337 | 148,337 | 148,337 | 111,704 | 58,373 | 58,373 | 58,373 | 41,459 |
| adj. $R^2$ | 0.010 | 0.010 | 0.010 | 0.010 | 0.011 | 0.010 | 0.010 | 0.012 |

Panel B: ETF and Mutual fund Fragility (2009-2018)

| | ETF | | | | MF and ETFs | | | |
|---|---|---|---|---|---|---|---|---|
| | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ | DGTW | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ | DGTW |
| $\sqrt{G}^{MF}$ | | | | | 0.245*** | 0.238*** | 0.245*** | 0.231*** |
| | | | | | (5.46) | (5.35) | (5.28) | (5.67) |
| $\sqrt{G}^{ETF}$ | 0.831*** | 0.804*** | 0.814*** | 0.774*** | 0.767*** | 0.744*** | 0.748*** | 0.619*** |
| | (9.18) | (9.08) | (9.27) | (7.48) | (8.62) | (8.73) | (8.86) | (6.74) |
| Obs. | 45,076 | 45,076 | 45,076 | 32,677 | 45,076 | 45,076 | 45,076 | 32,677 |
| adj. $R^2$ | 0.020 | 0.018 | 0.018 | 0.026 | 0.022 | 0.020 | 0.020 | 0.029 |

Table 6: **Stock return volatility, ownership by large 13F institutional investors, and stock price fragility**

This table presents the results of a panel regression of next quarter's stock volatility on a set of different aggregations of Institutional Ownership and stock price fragility estimated based on mutual fund data only ($G^{MF}$) or ETF data only ($G^{ETF}$). We estimate stock volatility as the standard deviation of daily stock returns within each quarter. *Top IO* represents the aggregate ownership of the largest institutional investors in a given stock. For specifications (1), (2), and (3), we sum the ownership of the top 3 institutions, while for specifications (4), (5), and (6), we take the top 10 institutions. *Bottom IO* represents the combined ownership of the smaller institutional investors whose equity holdings equal that of the top IO. *Middle IO* is the aggregated ownership of all institutional investors not considered neither in the top nor the bottom group of investors. The control variables include the Amihud (2002) illiquidity measure, the inverse of the stock price at quarter-end, book-to-market ratio, the log of the market capitalization of each stock estimated at quarter-end, and past 6-month momentum return over the previous two quarters. *t*-statistics are reported in parentheses and are based on standard errors clustered at the stock and quarter levels. $***$ , $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively. The *full sample* period is from 1989:Q4 to 2018:Q1.

| | Full Sample | | | | 2009-2018 | | | |
|---|---|---|---|---|---|---|---|---|
| | Top 3 Inst | | Top 10 Inst | | Top 3 Inst | | Top 10 Inst | |
| | (1) | (2) | (1) | (2) | (1) | (2) | (1) | (2) |
| Top IO | 0.471*** | 0.547*** | 0.263** | 0.424*** | 0.568*** | 0.617*** | 0.406*** | 0.424*** |
| | (2.71) | (4.37) | (2.37) | (4.44) | (5.00) | (4.37) | (4.29) | (4.44) |
| Mid IO | 0.163** | 0.115 | 0.184** | 0.048 | 0.164** | 0.115 | 0.158* | 0.048 |
| | (2.23) | (1.32) | (2.06) | (0.48) | (2.06) | (1.32) | (1.75) | (0.48) |
| Bottom IO | -0.466*** | 0.069 | 0.157 | 0.076 | 0.086 | 0.069 | 0.106 | 0.076 |
| | (-2.90) | (0.58) | (-1.45) | (0.72) | (0.72) | (0.58) | (1.08) | (0.72) |
| $G^{MF}$ | 0.071*** | | 0.070*** | | 0.039* | | 0.041* | |
| | (2.87) | | (2.85) | | (1.97) | | (1.91) | |
| $G^{ETF}$ | | 0.295** | | 0.208** | | 0.305** | | 0.288** |
| | | (2.09) | | (2.01) | | (2.25) | | (2.17) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Calendar-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Obs. | 131,040 | 68,960 | 131,678 | 69,217 | 77,054 | 68,960 | 77,421 | 69,217 |
| adj. $R^2$ | 0.668 | 0.675 | 0.667 | 0.674 | 0.675 | 0.675 | 0.674 | 0.674 |

Table 7: **Activeness of ETF sample**

This table reports the time-series averages of the cross-sectional mean, median, standard deviation, and 90th percentile of the activeness index (%) for the full sample period that covers the period from 2000:Q1 to 2018:Q4 as well as for three subperiods: before 2009, between 2009 and 2014, and from 2014 to 2018. For the same subperiods, the table shows the breakdown of the number of funds and assets under management (AUM) by the following four levels of activeness: Very Passive (VP) (activeness index < 25%), Moderately Passive (MP) (25% < activeness index < 50%), Moderately Active (MA), (50% < activeness index < 75%), and Very Active (VA) (activeness index > 75%).

| | Activeness index (%) | | | | Number of funds (%) | | | | AUM(%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Median | Std | P90 | VP | MP | MA | VA | VP | MP | MA | VA |
| Full sample | 89.41 | 97.38 | 17.48 | 99.95 | 0.92 | 4.69 | 10.27 | 84.53 | 19.91 | 11.98 | 6.99 | 62.09 |
| Before 2009 | 87.31 | 93.63 | 15.11 | 99.41 | 1.49 | 3.60 | 14.12 | 82.09 | 18.22 | 9.04 | 9.01 | 59.20 |
| 2009-2014 | 89.36 | 97.23 | 17.21 | 99.94 | 0.93 | 4.15 | 9.13 | 86.07 | 18.94 | 10.26 | 6.40 | 66.68 |
| 2014-2018 | 89.90 | 97.67 | 17.46 | 99.96 | 0.81 | 5.96 | 6.10 | 87.13 | 24.42 | 20.59 | 8.21 | 46.78 |

Table 8: **Stock return volatility, excess return volatility, and activeness of ETFs**

This table presents the results from Fama and MacBeth (1973) regressions of next quarter's total return volatility and excess return volatility on squared fragility of the current quarter. We estimate Fragility as detailed in Eq. (5). Following Easley et al. (2021), we classify ETFs according to their activeness index value into passive (Activeness index < 50%) and active (Activeness index > 50%) ETFs. The control variables included in specification (3) are: the log of stock price, the log of market capitalization, the ratio of book equity to market equity, the past 12-month cumulative stock return, lagged skewness of monthly stock returns, the log of age, share turnover, and the lagged dependent variable ($\sigma$). The sample period is from 2009:Q1 to 2018:Q4

| | Total return volatility | | | Excesss return volatility | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ | 1-Factor $\sigma$ | 3-Factor $\sigma$ | 4-Factor $\sigma$ |
| $\sqrt{G}^{ETF(Active)}$ | 0.801** | 0.727** | 0.381** | 0.887** | 0.817** | 0.745*** | 0.783** | 0.623** | 0.648*** |
| | (2.89) | (2.91) | (2.26) | (2.88) | (3.07) | (3.30) | (2.91) | (3.12) | (3.38) |
| $\sqrt{G}^{ETF(Passive)}$ | 0.128* | 0.130 | -0.170** | 0.164* | 0.162* | 0.116* | 0.127 | 0.0848 | 0.0873 |
| | (1.92) | (0.32) | (-1.97) | (2.10) | (1.85) | (2.06) | (0.32) | (0.11) | (0.22) |
| $\sqrt{G}^{MF}$ | | 0.387*** | 0.003 | | | | 0.236*** | 0.223*** | 0.230*** |
| | | (8.12) | (0.20) | | | | (5.32) | (5.11) | (4.96) |
| Add Controls | No | No | Yes | No | No | No | No | No | No |
| Obs. | 18,563 | 18,563 | 18,016 | 18,563 | 18,563 | 18,563 | 18,563 | 18,563 | 18,563 |
| adj. $R^2$ | 0.013 | 0.026 | 0.471 | 0.014 | 0.012 | 0.011 | 0.029 | 0.026 | 0.025 |

# Online Appendix

## OA1. A theoretical model of stock price fragililty

Our proposed model extends the Merton (1971)'s model to consider an idiosyncratic liquidity shock in an economy with two agents, which are heterogeneous in preferences.

### OA1.1. The Economic Setup

We first begin by defining that the agent's preferences are represented by the CRRA utility function as follows:

$$U_i\left(t, c_t\right) = e^{-\rho t}\left[\frac{c_{it}^{\gamma_i} - 1}{\gamma_i}\right], \quad i = 1, 2,$$

where $1 - \gamma_i$ is the relative risk aversion (RRA) of agent $i$, $\rho$ represents the impatience rate which is the same for both agents, and $c_{it}$ is the consumption rate per unit of time of agent $i$. Furthermore, the agents have access to two long-lived financial assets. The first asset is the risky one with a price $P_t$, and the second asset is the risk-free asset with a price $B_t$. The dynamic of asset prices is exogenous with the following dynamic:

$$\frac{dP_t}{P_t} = \alpha dt + \sigma dZ_t \quad (1)$$

$$dB_t = rB_t dt, \quad (2)$$

where $\alpha$ is the expected rate of return of the risky asset. We assume that this asset does not have dividends since it is common for mutual funds to reinvest all the profits in the portfolio. The volatility of risky asset returns is represented by $\sigma$, and

$r$ is the risk-free interest rate. The aggregate shock in this economy is represented by $dZ_t$, where $Z_t$ is a standard Brownian motion.

The wealth dynamic of the agent $i$ evolves according to Eq. (3).

$$dW_{it} = W_{it}\left[\theta_i(\alpha - r) + r - \frac{c_{it}}{W_{it}}\right]dt + W_{it}\theta_i\sigma dZ_t + W_{it}\sigma_{i,Liq}dZ_{i,Liq}, \tag{3}$$

where $\theta_i$ is the weight of the investment in the risky asset in the portfolio of agent $i$. We assume that an agent may experience surprise liquidity shocks such as a sudden drop in wealth. This shock is an idiosyncratic shock and is represented by $dZ_{i,Liq}$, where $Z_{i,Liq}$ is a standard Brownian motion. We also assume that these idiosyncratic shocks are not correlated between agents. Assuming that $\sigma_{i,Liq}$ is positive, a (negative) liquidity shock is when $dZ_{i,Liq}$ is negative, which means that the agent suddenly experiences a drop in his wealth. The Eq. (3) in compact form is

$$dW_{it} = W_{it}\mu_{it}dt + W_{it}q_i d\widetilde{Z}_i, \tag{4}$$

where

$$\mu_{it} = \theta_i(\alpha - r) + r - \frac{c_{it}}{W_{it}} \tag{5}$$

$$q_i = [\theta_i\sigma \quad \sigma_{i,Liq}] \tag{6}$$

$$d\widetilde{Z}_i = \left[dZ_t \quad dZ_{i,Liq}\right]' \tag{7}$$

We now define the consumption-portfolio choice problem for the agent $i$ as

$$\max_{\{c_{it},\theta_{it}\}} E_{0,W_{i0}}\left[\int_0^\infty U(t, c_{it})dt\right] \tag{8}$$

subject to

$$dW_{it} = W_{it}\mu_{it}dt + W_{it}q_i d\widetilde{Z}_i \tag{9}$$

with the following constraint

$$c_{it} \geq 0, \tag{10}$$

where $W_{i0}$ is the initial wealth of agent $i$.

The stochastic optimal control problem (Eq. 8, 9, and 10) can be transformed into a dynamic stochastic programming problem represented by the Hamilton-Jacobi-Bellman equation as follows.

$$\frac{\partial V_i\left(t, W_{it}\right)}{\partial t} + \sup_{c_{it}, \theta_{it}} \left\{ U\left(t, c_{it}\right) + \mathcal{A}(t) V_i\left(t, W_{it}\right) \right\} = 0, \tag{11}$$

where $V_i$ is the function value for the agent $i$ and $\mathcal{A}(t)$ is the second-order partial differential operator. We then use the first-order conditions to obtain the agent $i$'s optimal portfolio.

**Lemma A6.1.** *Given the optimal value function, $V_i$, that solves the Hamilton-Jacobi-Bellman equation, the optimal portfolio for agent $i = 1, 2$ is*

$$\theta_{it} = \left(\frac{\alpha - r}{\sigma^2}\right) \frac{1}{1 - \gamma_i} \tag{12}$$

### OA1.2.Non-fundamental Demand of the Risky Asset

We now calculate the total demand for the shares of the risky asset, which is $N^d$:

$$N^d = \sum_{i=1}^{2} N_i = N_1 + N_2, \tag{13}$$

where $N_i$ is the risky asset demand (in terms of the number of shares) of agent $i$. We know that the optimal portfolio, $\theta_{it}$, can also be written as

$$\theta_{it} = \frac{P_t N_{it}}{W_{it}} \tag{14}$$

Then, we can obtain the shares demand of agent $i$

$$N_{it} = \frac{W_{it}\theta_{it}}{P_t} \tag{15}$$

Introducing Eq. (15) into the aggregate risky asset demand (Eq. 13), we have

$$N^d = \sum_{i=1}^{2} N_{it} = \frac{W_{1t}\theta_{1t}}{P_t} + \frac{W_{2t}\theta_{2t}}{P_t}, \tag{16}$$

which is the share demand of the risky asset. Ordering the elements of Eq. (16), we have

$$N^d = \frac{1}{P_t}\big(W_{1t}\theta_{1t} + W_{2t}\theta_{2t}\big) \tag{17}$$

The Eq. (17) suggests that $N^d$ depends on three stochastic processes: $P_t$, $W_{1t}$, and $W_{2t}$.

$$N^d = f(P_t, W_{1t}, W_{2t})$$

Using the Itô's lemma, we find the dynamic of risky-shares demand, $dN^d$.

**Lemma A6.2.** *The dynamic of the risky asset demand is represented by the following stochastic differential equation*

$$dN^d = \frac{1}{P_t}g(W_{1t}, W_{2t})dt + \frac{1}{P_t}h(W_{1t}, W_{2t})dZ_t + \frac{1}{P_t}[\theta_{1t}W_{1t}\sigma_{1,Liq}]dZ_{1,Liq} + \frac{1}{P_t}[\theta_{2t}W_{2t}\sigma_{2,Liq}]dZ_{2,Liq} \tag{18}$$

*where*

$$g(W_{1t}, W_{2t}) = \tag{19}$$

$$h(W_{1t}, W_{2t}) = \tag{20}$$

We also can split the change in asset demand as the change in fundamental demand and non-fundamental demand as follows.

$$dN^d = \underbrace{\frac{1}{P_t}g(W_{1t}, W_{2t})dt + \frac{1}{P_t}h(W_{1t}, W_{2t})dZ_t}_{=dN_f\text{:change in fundamental Demand}} + \underbrace{\frac{1}{P_t}[\theta_{1t}W_{1t}\sigma_{1,Liq}]dZ_{1,Liq} + \frac{1}{P_t}[\theta_{2t}W_{2t}\sigma_{2,Liq}]dZ_{2,Liq}}_{=dN_f\text{:change in non-fundamental Demand}} \tag{21}$$

Then, Eq. (21) could be expressed as

$$dN^d = dN_f + dN_{nf}, \tag{22}$$

where the change in non-fundamental demand is driven by the agent's liquidity shocks.

$$dN_{nf} = \frac{1}{P_t}[\theta_{1t}W_{1t}\sigma_{1,Liq}]dZ_{1,Liq} + \frac{1}{P_t}[\theta_{2t}W_{2t}\sigma_{2,Liq}]dZ_{2,Liq} \tag{23}$$

We then use the definition of portfolio weights to obtain the number of shares of the risky asset per agent as follows.

$$\theta_{it} = \frac{P_t N_{it}}{W_{it}} \longrightarrow N_{it} = \frac{\theta_{it} W_{it}}{P_t} \tag{24}$$

We introduce the expression $\theta_{it} W_{it}/P_t$ into Eq. (23) resulting

$$dN_{nf} = N_{1t}\sigma_{1,Liq} dZ_{1,Liq} + N_{2t}\sigma_{2,Liq} dZ_{2,Liq} \tag{25}$$

This equation reflects the effects of liquidity shock of two agents in the total non-fundamental demand. For instance, if only agent 1 experiences a liquidity shock ($dZ_{1,Liq} < 0$), this will reduce the non-fundamental demand of the risky asset with intensity $\sigma_{1,Liq}$. We can also consider the ownership (or concentration) of the asset in the analysis. Dividing the Eq. (25) by the total shares outstanding, $N$, and considering that $\eta_{it}$ is the ownership of agent $i$ of the risky asset at time $t$: $\eta_{it} = N_{it}/N$, we have

$$dN_{nf} = N\eta_{1t}\sigma_{1,Liq} dZ_{1,Liq} + N\eta_{2t}\sigma_{2,Liq} dZ_{2,Liq} \tag{26}$$

Suppose that $\sigma_{1,Liq} = \sigma_{2,Liq}$, but agent 1 has more shares of the asset in his portfolio, i.e., $\eta_1 > \eta_2$. In this case, if agent 1 experiences a liquidity shock, the effect on non-fundamental demand would be higher than the case in which agent 2 experiences the same shock. The reason for that is agent 1 has more concentration of the asset in his portfolio. Therefore, ownership is relevant to understand the effects of liquidity shocks on asset demand and hence on asset prices.

### OA1.3. Stock Price Fragility

Greenwood and Thesmar (2011) define *fragility* as "the expected volatility of non-fundamental demand given an asset's ownership structure." In our theoretical model, shifts in non-fundamental demand are represented by Eq. (26). Although its expected value is equal to zero, $E(dN_{nf}) = 0$, its variance fits with the asset fragility definition of Greenwood and Thesmar (2011). Then, we define asset fragility as the variance of $dN_{nf}$ as follows

$$\text{Fragility} = Var(dN_{nf}) \tag{27}$$

In order to be explicit on the asset ownership and the Var-Cov matrix of liquidity shocks, we express $dN_{nf}$ in matrix form as follows.

$$dN_{nf} = \underbrace{\begin{bmatrix} N\eta_1 & N\eta_2 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} \sigma_{1,Liq} dZ_{1,Liq} \\ \sigma_{2,Liq} dZ_{2,Liq} \end{bmatrix}}_{Z} \equiv MZ \tag{28}$$

Then, $Var(dN_{nf})$ is defined as follows

$$\begin{aligned} Var(dN_{nf}) &= E\left[MZZ'M'\right], \quad \text{with } E[dN_{nf}] = 0 \\ &= ME\left[ZZ'\right]M' \\ &= N^2 \begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \Omega \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \end{aligned} \tag{29}$$

where $[\eta_1 \quad \eta_2]$ is a vector of asset ownership and $\Omega$ is the Var-Cov matrix of liquidity shocks defined as

$$E\left[ZZ'\right] = \Omega = \begin{bmatrix} \sigma^2_{1,Liq} Var(dZ_{1,Liq}) & \sigma_{1,Liq}\sigma_{2,Liq} Cov(dZ_{1,Liq}, dZ_{2,Liq}) \\ \sigma_{1,Liq}\sigma_{2,Liq} Cov(dZ_{1,Liq}, dZ_{2,Liq}) & \sigma^2_{2,Liq} Var(dZ_{2,Liq}) \end{bmatrix} \tag{30}$$

In our model, we assume that both idiosyncratic shocks are independent, then

$Cov(dZ_{1,Liq}, dZ_{2,Liq}) = 0$. However, the model can be easily extended to the case in which these shocks are correlated. With Eq. (30), our fragility definition would be

$$\text{Fragility} = Var(dN_{nf}) = N^2 \begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \Omega \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}, \tag{31}$$

which considers the effect of ownership and the Var-Cov matrix of liquidity shocks. This result provides a microfoundation of the measure of stock fragility of Greenwood and Thesmar (2011).

### OA1.4. Stock Price Fragility and Stock Return Volatility

We then analyze the connection between fragility and stock return volatility based on our model. First, we assume the supply side of the shares of the risky asset is represented by

$$N_t^s = AP_t, \quad A > 0 \tag{32}$$

In equilibrium, we have

$$dN_t^s = dN_t^d, \tag{33}$$

Using the Eq. (32) and the Eq. (22), the equilibrium condition (33) is equivalent to

$$d(AP_t) = dN_f + dN_{nf} \tag{34}$$

Dividing by $P_t$ and then applying the variance operator in Eq. (34), we have

$$A^2 \,\text{Var}\left[\frac{dP_t}{P_t}\right] = \frac{1}{P_t^2}\,\text{Var}[dN_f] + \frac{1}{P_t^2}\,\text{Var}[dN_{nf}], \tag{35}$$

which connects the volatility of the rate of return with the fragility measure.

**Lemma A6.3.** *Given the equilibrium condition in Eq. (33) and the assumption of the supply side of the risky asset, there exists a relationship between the volatility of the rate of return and the variance of the change of non-fundamental demand, which is the definition of stock fragility.*

$$\mathrm{Var}\left[\frac{dP_t}{P_t}\right] = \frac{1}{A^2 P_t^2}\,\mathrm{Var}[dN_f] + \frac{1}{A^2 P_t^2}\,\underbrace{\mathrm{Var}[dN_{nf}]}_{\textit{Fragility}}, \tag{36}$$

*where the rate of return of the risky asset is represented by* $dP_t/P_t$.

## OA2. Stock characteristics of quintile portfolios

Table A1: **Stock Characteristics**

For quarter $t$, stocks in our sample are sorted into 5 quintile portfolios based on their fragility. Fragility is defined as the conditional expected variance of flow-driven net buys into a stock. This table reports the time-series mean of the cross-sectional average of several stock-level characteristics for each fragility quintile portfolio. *PRC* is share price. *Market Cap* is the average stock's market capitalization (end-of-quarter share price times the total number of shares outstanding), expressed in millions of US dollars. *BM* is the book-to-market ratio. *Ret12* is the past 12-month stock return. *Age (years)* is the firm's Age is calculated as the number of years (months/12) since the first return appears in CRSP. *Turnover* is the average monthly share turnover (monthly volume traded over total shares outstanding) over the previous 3 months. *N Owners* is the average number of mutual funds (ETFs) that hold the same stock. Panel A shows the results for quintile portfolios sorted on fragility estimated as in Eq. (5) that consider flows and holdings data from mutual funds only ($G^{MF}$). Similarly, Panel B reports the average values for the characteristics sorted on fragility calculated using ETF data exclusively ($G^{ETF}$). The sample covers the period from 1989:Q4 to 2018:Q4.

Panel A: MF fragility ($G^{MF}$)

| Quintile | PRC | Market Cap | BM | RET12 | Age | Turnover | N Owners |
|---|---|---|---|---|---|---|---|
| 1(low) | 196.5 | 16,423.9 | 0.719 | 0.247 | 21.7 | 0.238 | 40.5 |
| 2 | 280.9 | 18,696.1 | 0.608 | 0.264 | 25.4 | 0.202 | 90.6 |
| 3 | 63.0 | 7,978.9 | 0.605 | 0.278 | 22.4 | 0.228 | 71.6 |
| 4 | 46.9 | 4,645.8 | 0.623 | 0.242 | 20.9 | 0.243 | 60.7 |
| 5 (high) | 41.0 | 3,095.3 | 0.632 | 0.207 | 19.7 | 0.266 | 55.5 |

Panel B: ETF fragility ($G^{ETF}$)

| Quintile | PRC | Market Cap | BM | RET12 | Age | Turnover | N Owners |
|---|---|---|---|---|---|---|---|
| 1(low) | 250.7 | 17,070.2 | 0.676 | 0.174 | 24.4 | 0.208 | 36.6 |
| 2 | 105.1 | 29,573.6 | 0.573 | 0.191 | 30.1 | 0.194 | 64.5 |
| 3 | 98.9 | 15,935.4 | 0.590 | 0.218 | 24.9 | 0.225 | 38.5 |
| 4 | 60.6 | 9,718.1 | 0.595 | 0.236 | 23.6 | 0.215 | 47.6 |
| 5 (high) | 48.2 | 12,771.7 | 0.670 | 0.176 | 24.9 | 0.224 | 28.9 |

## OA3. Sorting on Fragility

## Table A2: **Portfolio sorting on fragility**

At the end of each quarter, we form quintile stock portfolios based on one-quarter lag stock price fragility estimated using only mutual funds data ($G^{MF}$) or only ETF data ($G^{ETF}$) and track their monthly excess returns as the *value-weighted* average of excess returns on all stocks in each portfolio. Quintile 1 includes the stocks with the highest value of each stock price fragility measure, while Quintile Q5. Q1-Q5 is the spread portfolio that goes long Q1 and shorts Q5. Panel A presents the risk-adjusted one-quarter return of each portfolio. Panel B shows similar results when we track portfolio returns in the next two quarters. We adjust risk exposure using the three factors of Fama and French (1993) - FF3, the five factors of Fama and French (2015) - FF5, and the Fama-French five-factor model augmented with the *illiquid-minus-liquid* (IML) factor Amihud (2019) and the momentum factor of Carhart (1997) - FF5MA. Alphas are in percent per month. $***$ , $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels. The sample period is from 1989:Q1 to 2018:Q4.

Panel A: Equally-weighted

| | FF3 | | | FF5 | | | FF5MA | | |
|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 |
| $G^{MF}$ | 0.008 | -0.328 | 0.335 | -0.004 | -0.331 | 0.327 | -0.007 | -0.293 | 0.287 |
| | (0.08) | (-2.35) | (2.55) | (-0.04) | (-2.34) | (2.43) | (-0.07) | (-2.12) | (2.02) |
| $G^{ETF}$ | 0.072 | -0.333 | 0.404 | 0.052 | -0.320 | 0.372 | 0.019 | -0.278 | 0.297 |
| | (0.77) | (-2.45) | (3.20) | (0.56) | (-2.35) | (2.98) | (0.20) | (-2.15) | (2.34) |

Panel B: Value-Weighted

| | FF3 | | | FF5 | | | FF5MA | | |
|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 |
| $G^{MF}$ | 0.076 | -0.132 | 0.208 | 0.065 | -0.143 | 0.208 | 0.042 | -0.147 | 0.189 |
| | (0.81) | (-1.72) | (1.70) | (0.70) | (-1.82) | (1.68) | (0.45) | (-1.79) | (1.44) |
| $G^{ETF}$ | 0.029 | -0.251 | 0.279 | -0.023 | -0.259 | 0.277 | -0.013 | -0.251 | 0.238 |
| | (0.30) | (-2.32) | (2.27) | (-0.23) | (-2.37) | (2.23) | (-0.13) | (-2.20) | (1.91) |

Panel C: Two-quarters ahead Equally-weighted

| | FF3 | | | FF5 | | | FF5MA | | |
|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 |
| $G^{MF}$ | -0.096 | -0.337 | 0.241 | -0.103 | -0.337 | 0.234 | -0.102 | -0.284 | 0.182 |
| | (-0.99) | (-2.34) | (1.78) | (-1.05) | (-2.30) | (1.69) | (-1.02) | (-1.97) | (1.27) |
| $G^{ETF}$ | 0.112 | -0.391 | 0.502 | 0.108 | -0.388 | 0.495 | 0.041 | -0.335 | 0.376 |
| | (1.18) | (-2.75) | (3.96) | (1.12) | (-2.74) | (3.92) | (0.43) | (-2.52) | (3.09) |

Panel D: Two-quarters ahead Value-Weighted

| | FF3 | | | FF5 | | | FF5MA | | |
|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 | Q1 | Q5 | Q1-Q5 |
| $G^{MF}$ | -0.088 | -0.122 | 0.034 | -0.095 | -0.140 | 0.045 | -0.115 | -0.134 | 0.019 |
| | (-0.94) | (-1.41) | (0.28) | (-1.00) | (-1.60) | (0.36) | (-1.14) | (-1.45) | (0.14) |
| $G^{ETF}$ | 0.084 | -0.309 | 0.393 | 0.081 | -0.313 | 0.394 | -0.001 | -0.280 | 0.279 |
| | (0.85) | (-2.54) | (2.68) | (0.81) | (-2.54) | (2.64) | (-0.01) | (-2.18) | (1.82) |

## OA4. Alternative specifications - Large institutional investors, fragility and stock return volatility

Table A3: **Stock return volatility, ownership by large 13F institutional investors, and stock price fragility**

This table presents the results of a panel regression of next quarter's stock volatility on a set of different aggregations of Institutional Ownership and stock price fragility estimated based on mutual fund data only ($G^{MF}$) or ETF data only ($G^{ETF}$). We estimate stock volatility as the standard deviation of daily stock returns within each quarter. *Top IO* represents the aggregate ownership of the largest institutional investors in a given stock. For specifications (1), (2), and (3), we sum the ownership of the top 5 institutions, while for specifications (4), (5), and (6), we take the top 7 institutions. *Bottom IO* represents the combined ownership of the smaller institutional investors whose equity holdings equal that of the top IO. *Middle IO* is the aggregated ownership of all institutional investors not considered neither in the top nor the bottom group of investors. The control variables include the Amihud (2002) illiquidity measure, the inverse of the stock price at quarter-end, book-to-market ratio, the log of the market capitalization of each stock estimated at quarter-end, and past 6-month momentum return over the previous two quarters. $t$-statistics are reported in parentheses and are based on standard errors clustered at the stock and quarter levels. $***$ , $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively. The sample period is from 1989:Q4 to 2018:Q1

| | Full Sample | | | | 2009-2018 | | | |
|---|---|---|---|---|---|---|---|---|
| | Top 5 Inst | | Top 7 Inst | | Top 5 Inst | | Top 7 Inst | |
| | (1) | (2) | (1) | (2) | (1) | (2) | (1) | (2) |
| Top IO | 0.467*** | 0.578*** | 0.429*** | 0.534*** | 0.652*** | 0.678*** | 0.567*** | 0.594*** |
| | (3.35) | (5.84) | (3.45) | (5.83) | (5.62) | (5.84) | (5.65) | (5.83) |
| Mid IO | 0.131* | 0.054 | 0.125 | -0.004 | 0.116 | 0.053 | 0.075 | -0.003 |
| | (1.77) | (0.52) | (1.54) | (-0.04) | (1.23) | (0.52) | (0.79) | (-0.04) |
| Bottom IO | -0.284** | 0.106 | -0.227* | 0.124 | 0.139 | 0.106 | 0.155 | 0.124 |
| | (-2.18) | (1.00) | (-1.91) | (1.20) | (1.42) | (1.00) | (1.60) | (1.20) |
| $G^{MF}$ | 0.061*** | | 0.052*** | | 0.041* | | 0.036* | |
| | (2.89) | | (2.88) | | (1.94) | | (1.97) | |
| $G^{ETF}$ | | 0.304** | | 0.314** | | 0.384** | | 0.394** |
| | | (2.13) | | (2.06) | | (2.33) | | (2.36) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Calendar-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Obs. | 131,394 | 69,101 | 131,481 | 69,121 | 77,236 | 69,101 | 77,277 | 69,121 |
| adj $R^2$ | 0.668 | 0.675 | 0.668 | 0.675 | 0.675 | 0.675 | 0.675 | 0.675 |

Table A4: **Stock return volatility and fragility as in Friberg et al. (2023)**

This table presents the results of a panel regression of average daily return volatility over the next quarter on the square root of mutual fund fragility and ETF fragility following Friberg et al. (2023). We include as control variables the log of market capitalization, the inverse of stock price as well as Year-Quarter and firm fixed effects. $t$-statistics are reported in parentheses and are based on standard errors clustered at the stock levels. $***$ , $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively. To compare our results to those of Friberg et al. (2023) we define the sample period from 2001:Q1 to 2017:Q4

| | All firms | | | IO > 0.2 | | | Mkt cap > Median | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| $\sqrt{G^{MF}}$ | 0.065*** | | 0.032* | 0.060*** | | 0.046*** | 0.064*** | | 0.046*** |
| | (3.60) | | (1.86) | (3.37) | | (2.57) | (3.58) | | (2.59) |
| $\sqrt{G^{ETF}}$ | | 0.187** | 0.176** | | 0.191** | 0.179** | | 0.193** | 0.178** |
| | | (2.24) | (2.10) | | (2.26) | (2.20) | | (2.34) | (2.13) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Obs. | 98,304 | 69,776 | 69,776 | 95,923 | 68,744 | 68,744 | 98,283 | 69,772 | 69,772 |
| adj. $R^2$ | 0.662 | 0.683 | 0.689 | 0.661 | 0.683 | 0.688 | 0.662 | 0.683 | 0.691 |

## OA5. 13F Institutional investors ETF and leveraged/inverse holdings ETF

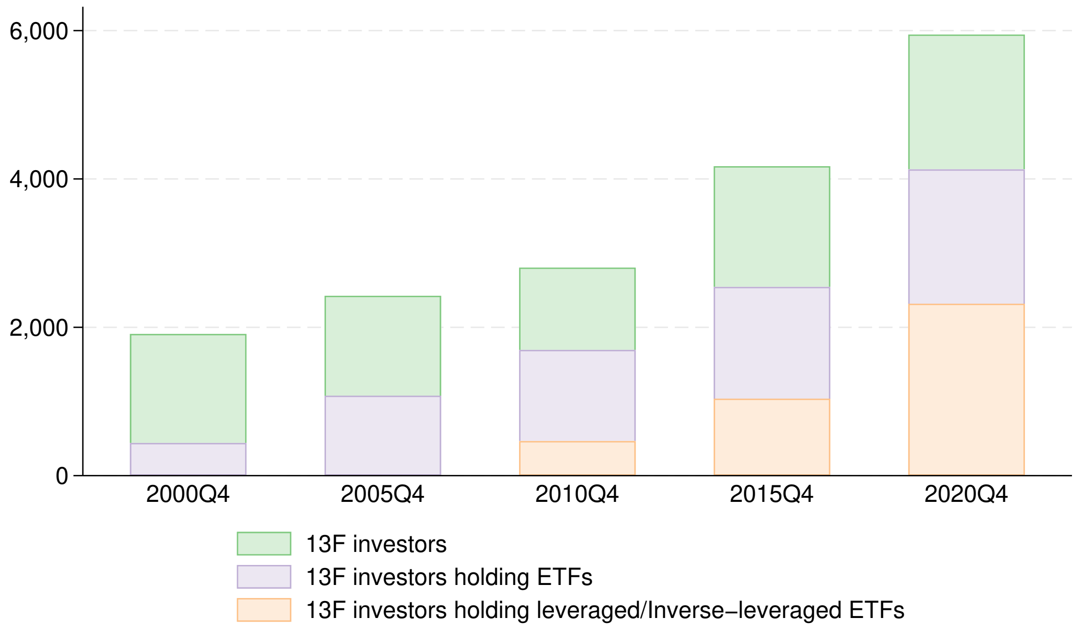
The first domestic ETF was introduced in 1993 while the first domestic leverage ETF was created in 2006. Since their inception, ETF attracted the attention of investors due to their hybrid design that combined characteristics of open and closed-end mutual funds while offering, initially, broad diversification at a lower cost.

Leverage ETFs were first launched to the market in 2006. Similarly to traditional ETFs, these funds offered exposure to a wide set of benchmarks, however, their replication method includes using derivatives. This mechanism allows ETFs fund managers to leverage the performance of the fund. While a positive exposure is possible (obtaining 1.5x or 2x the return of a specific benchmark) it is also possible to obtain a negative exposure. This is, investors can also buy ETFs that offered to provide the negative exposure by obtaining negative multiplier of the benchmark return, for instance -1.5x -2x of the return.

We identify leverage and inverse-leverage ETFs as those that include the following terms in their names: leverage, inverse, Double, Short, Ultra, UltraShort, 4x, 3x, 2.5x, 2x, 1.5x, 1.25x.

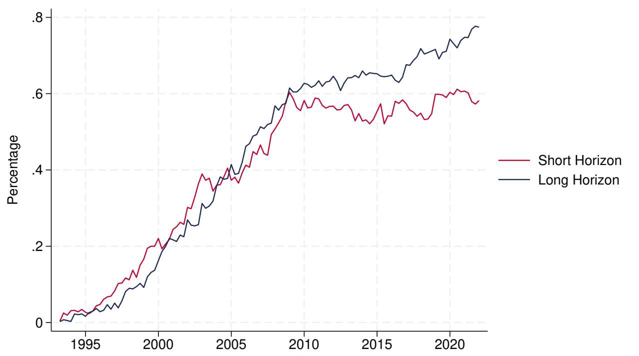
Figure A1: **13F Institutional Investors holding ETFs**

This figure plots the total number of 13F institutional investors, the number of 13F institutional investors that held ETFs and leveraged/inverse-leveraged ETFs in their portfolios at the last quarter of five different years.

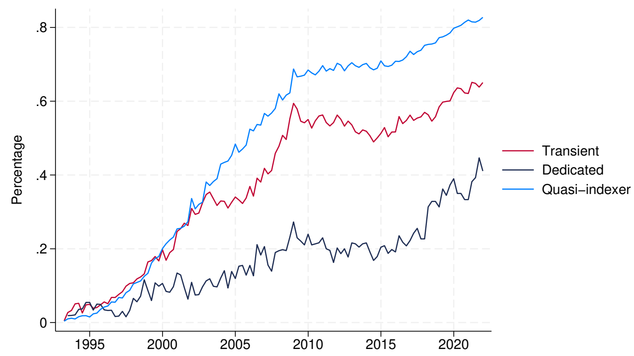


## Figure A2: **13F Institutional Investors holding ETFs**

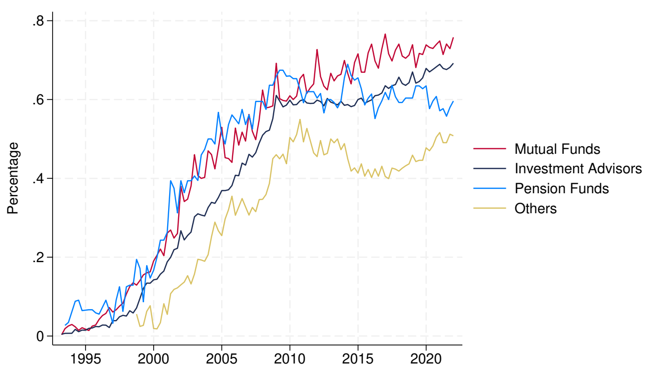
This figure shows the time series of the percentage of 13F institutional investors that held exchange-traded funds (ETFs) in their portfolios from 1993 to 2021. 13F institutional investors are classified based on three different criteria. In panel A, investors are classified into short-horizon and long-horizon based on the average churn ratio of Yan and Zhang (2009). In panel B, we group investors into transient (i.e., show high portfolio turnover and highly diversified portfolios), dedicated (i.e., characterized by large investments in portfolio firms and low portfolio turnover), and quasi-indexer (i.e., those with low portfolio turnover but more diversified portfolios) Bushee (2001). In panel C, we classify investors following Koijen and Yogo (2019). The 13F holdings data is obtained from Thomson/Refinitiv, while ETF data is collected from Bloomberg and CRSP.




(a) By Investment Horizon (Yan and Zhang, 2009)



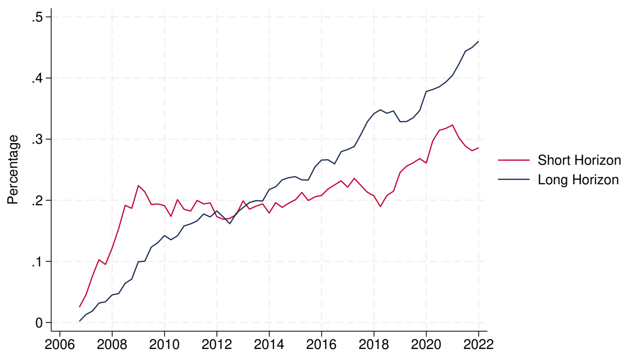

(b) By 13F Investor type (Bushee, 2001)



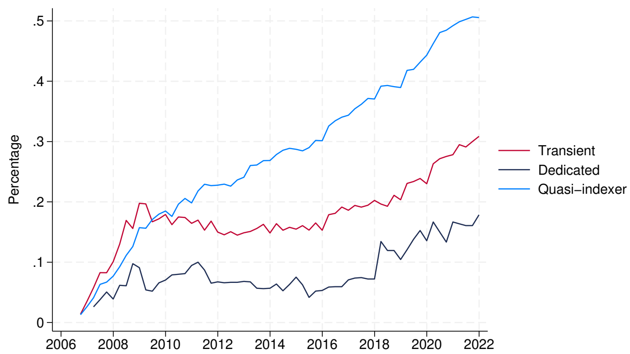

(c) By Investor type (Koijen and Yogo, 2019)

Figure A3: **13F Institutional Investors holding leveraged/inverse-leveraged ETFs**

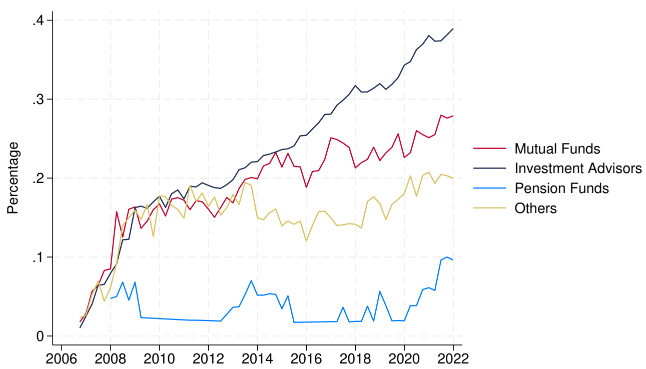
This figure shows the time series of the percentage of 13F institutional investors that held leverage or inverse-leveraged exchange-traded funds (ETFs) in their portfolios from 1993 to 2021. 13F institutional investors are classified based on three different criteria. In panel A, investors are classified into short-horizon and long-horizon based on the average churn ratio of Yan and Zhang (2009). In panel B, we group investors into transient (i.e., show high portfolio turnover and highly diversified portfolios), dedicated (i.e., characterized by large investments in portfolio firms and low portfolio turnover), and quasi-indexer (i.e., those with low portfolio turnover but more diversified portfolios) Bushee (2001). In panel C, we classify investors following Koijen and Yogo (2019). The 13F holdings data is obtained from Thomson/Refinitiv, while leveraged and inverse-leveraged ETF data is collected from Bloomberg and CRSP.




(a) By Investment Horizon (Yan and Zhang, 2009)




(b) By Investor type (Bushee, 2001)




(c) By Investor type (Koijen and Yogo, 2019)